

THE STATISTICAL ANALYSIS OF TIME SERIES

T. W. ANDERSON

Professor of Statistics and Economics

Stanford University

JOHN WILEY & SONS, Inc.

New York London Sydney Toronto

1971

Т. Андерсон

СТАТИСТИЧЕСКИЙ
АНАЛИЗ
ВРЕМЕННЫХ
РЯДОВ

Перевод с английского

И. Г. Журбенко и В. П. Носко

Под редакцией

Ю. К. Беляева

ИЗДАТЕЛЬСТВО «МИР»

Москва 1976

Монография известного американского специалиста по математической статистике содержит обстоятельное изложение теории статистических выводов для различных вероятностных моделей. Излагаются методы представления временных рядов, оценивания параметров соответствующих вероятностных моделей, проверки гипотез относительно их структуры.

Собранный автором обширный материал, разбросанный ранее по различным источникам, делает книгу ценным руководством и справочником. Большое число задач удачно дополняет основной текст, позволяет ознакомиться с перспективами развития теории.

Эта книга весьма полезна студентам и аспирантам, специализирующимся в области теории вероятностей и математической статистики; она, несомненно, привлечет внимание инженеров, математиков и научных работников различных специальностей, интересующихся приложениями теории вероятностей.

Редакция литературы по математическим наукам

© Перевод на русский язык, «Мир», 1976

Автор книги профессор Станфордского университета Т. Андерсон знаком советскому читателю по изданной в русском переводе книге «Введение в многомерный статистический анализ». В новой его книге освещен широкий круг проблем, связанных со статистическим анализом последовательностей случайных величин. С задачами такого рода приходится обычно сталкиваться при анализе эмпирических данных.

В книге описаны многочисленные математические модели, в рамках которых отыскиваются рациональные методы получения оценок и проверки гипотез об адекватности выбранной математической модели обрабатываемым данным. Значительное внимание уделено моделям регрессии и авторегрессии с конечным числом неизвестных параметров. Подробно рассмотрены методы оценки спектральных плотностей стационарных случайных процессов, а также выявление трендов в последовательностях данных, «зашумленных» стационарными процессами.

Отличительной чертой книги Т. Андерсона является детальная проработка рассматриваемых проблем. Требования, предъявляемые к уровню математической подготовки читателя, вполне умеренные. Предполагается, что читатель знаком с основными понятиями теории вероятностей, математической статистики и теории матриц. Необходимые сведения по теории случайных процессов приведены в гл. 7.

Каждая глава книги завершается большим числом задач, что в равной степени полезно и читателям, индивидуально работающим с книгой, и преподавателям, которые могут использовать различные части книги в курсах лекций по теории случайных процессов и статистике. В приложении к книге приведены примеры анализа эмпирических временных рядов (ежегодных индексов цен на пшеницу, чисел солнечной активности) и рядов, полученных моделированием процессов авторегрессии.

Книга будет полезна не только математикам, работающим в области теории вероятностей и математической статистики, но также и широкому кругу специалистов, которые сталкиваются с необходимостью обработки измерений, рассматриваемых как случайные временные ряды.

Узнав о переводе книги на русский язык, Т. Андерсон любезно прислал нам список опечаток. Перевод гл. 7 и 8 сделан И. Г. Журбенко. Весь остальной материал переведен В. П. Носко.

Ю. К. Беляев

ИЗ ПРЕДИСЛОВИЯ АВТОРА

Авторы книг по статистическому анализу временных рядов по-разному могут подходить к их написанию. Я посвятил книгу математической теории статистических выводов относительно вероятностных моделей, лежащих (по предположению) в основе тех или иных временных рядов. Исходная вероятностная модель может включать детерминированный тренд и случайную составляющую, образующую стационарный случайный процесс. При этом естественно возникает целый ряд статистических задач, относящихся к характеристикам тренда и процесса. В книге всюду, где это возможно, отыскиваются оптимальные процедуры. Изучаются статистические свойства различных методов, причем во многих случаях свойства метода можно описать только для больших выборок, т. е. на основании совокупности наблюдений значений ряда в течение длительного периода времени. Как правило, дается математически строгий вывод соответствующих свойств.

В то время как теория развивается при строгих математических предположениях, приведенные в книге методы могут быть использованы и в тех случаях, когда указанные предположения выполнены не полностью. Часто бывает так, что свойства процедур меняются при этом несущественно. Как бы то ни было, точные результаты теорем служат руководством по использованию соответствующих процедур. В книге приведены примеры приложений развиваемых методов, обсуждаются вычислительные аспекты, вопросы использования и интерпретации этих методов. В то же время вопросы представления процедур в виде программ для вычислительных машин не затрагиваются.

Основой настоящей монографии явился курс лекций, который я читал в течение многих лет в Колумбийском университете обычно в течение одного, а иногда и двух семестров. Однако весь материал, в том виде как он представлен в книге, нельзя изложить в двухсеместровом курсе — лектор, использующий эту книгу как учебник, должен выбрать из нее разделы, представляющиеся ему наиболее интересными и важными. В книге приведено значительное число упражнений. Одни из них являются приложениями описанных методов, другие представляют собой разработку частных случаев общей теории. Часть упражнений связана с уточнением деталей сложных доказательств, остальные посвящены развитию теории.

Я надеюсь, что книга послужит не только учебником для начинающих, но будет полезна статистикам и специалистам, желающим познакомиться с анализом временных рядов, не обращаясь к фор-

мальным курсам. Чтение этой книги и выполнение упражнений поможет существенно пополнить запас статистических методов, полезных для анализа временных рядов. Книгу можно использовать также и как справочник. В ней достаточно систематизированно представлен обширный материал, ранее разбросанный по различным источникам. Приводятся также некоторые новые теоремы и методы. В ряде случаев за счет ослабления предположений известных ранее утверждений получены более сильные результаты.

Поскольку область анализа временных рядов весьма обширна, передо мной встала задача выбора разделов, которые следует включить в книгу. Материал, помещенный в этой книге, вместе с указанием соответствующих ограничений, при которых развивается теория, описан мною во введении (гл. 1). Определенную информацию дает также предметный указатель. Хочется надеяться, что основные и наиболее важные разделы анализа временных рядов оказались представленными, хотя полнота охвата тематики является делом вкуса. Постоянно создаются новые методы, меняются точки зрения, так что приведенные здесь результаты вряд ли можно считать окончательными. Фактически часть включенного материала уже в настоящее время представляет скорее исторический интерес.

Ввиду большого объема книги будет, по-видимому, полезно дать несколько советов читателям и преподавателям относительно выбора материала для самостоятельного изучения и для целей обучения. Глава 2 содержит краткое замкнутое изложение метода наименьших квадратов. Главы 3 и 4 имеют дело с моделями, в которых случайные составляющие независимы (о них иногда говорят как об «ошибках измерений»). Некоторые идеи и элементы анализа, приведенные в этих главах, используются в дальнейшем. Однако читатель, интересующийся именно последующими главами, может довольно многое пропустить (включая большую часть § 3.4, 4.3 и 4.4). Процессы авторегрессии, которые оказываются полезными в приложениях и служат примерами стационарных случайных процессов, исследуются в гл. 5. Посвященные теории больших выборок § 5.5 и 5.6 содержат результаты, доказательства которых включают значительное количество деталей и могут быть опущены. Статистические выводы для этих моделей служат основой анализа стационарных процессов «во временной области».

Глава 6 представляет собой обширное исследование сериальной корреляции и критериев для проверки независимости. Параграфы 6.3 и 6.4 имеют по преимуществу теоретический интерес. В § 6.5 изучается алгебра квадратичных форм и их отношений. Распределения, моменты и аппроксимирующие распределения получены в § 6.7 и 6.8. Там же даны таблицы процентных точек для соответствующих критериев. Первые пять параграфов гл. 7 служат введением в теорию стационарных случайных процессов и спектральные свойства этих процессов. В гл. 8 развивается теория статистик, относящихся

к стационарным случайным процессам. Глава 9 посвящена оцениванию спектральной плотности, являющемуся основой анализа стационарных процессов «в частотной области». В § 10.2 регрессионный анализ (гл. 2) распространяется на стационарные случайные составляющие. Результаты гл. 8 и 9 обобщаются на этот случай в § 10.3. В § 10.4 результаты гл. 6 обобщаются на случай, когда рассматриваются остатки от подобранного тренда. Можно выделить следующие части книги, которые образуют единое целое и которые можно читать независимо от других частей: (I) — гл. 2; (II) — гл. 3 и 4; (III) — гл. 5; (IV) — гл. 6; (V) — гл. 7; (VI) — гл. 8 и 9.

В практических приложениях статистический анализ временных рядов, кроме того, включает и менее формальную технику (которую теперь иногда называют «анализом данных»). Так, пониманию явления способствует простое графическое представление наблюдаемого временного ряда. Могут оказаться полезными те или иные преобразования измерений и связь последних с другими данными. Точно сформулированные процедуры, которые изучаются в этой книге, как правило, используются не в буквальном виде, а изменяются в зависимости от конкретных задач. Тем не менее для исследования статистических методов в рамках математической теории некоторые аспекты анализа необходимо формализовать. Например, задача об определении того, будет ли некоторый эффект достаточно большим, чтобы его можно было считать заметным, иногда формализуется как задача проверки гипотезы о том, что соответствующий параметр просто равен нулю.

По степени сложности эта книга примерно такая же, как и моя более ранняя книга «Введение в многомерный статистический анализ»; требуется лишь некоторое знание алгебры матриц. (Необходимый материал представлен в приложении В к упомянутой книге; дополнительные результаты приводятся в тексте и в упражнениях настоящей книги.) Полезно общее знакомство со статистическими методами. В частности, предполагается, что читатель знаком со стандартным материалом статистического анализа одной случайной переменной, включающим t -критерии и F -статистики, с многомерным нормальным распределением, с элементами оценивания и проверки гипотез. Более тонкие вопросы оценивания, проверки гипотез и теории решений, на которые имеются ссылки, излагаются в упражнениях. [Детальное и строгое изложение вопросов проверки гипотез читатель может найти у Лемана (1959).] Предполагается также знание основ курса высшей математики. Хотя в книге рассматриваются только действительные временные ряды, тем не менее иногда удобно записывать некоторые выражения с привлечением комплексных переменных. Однако фактически теория функций комплексного переменного в книге не применяется, поскольку (за исключением одной задачи) дело ограничивается формулой $e^{i\theta} = \cos \theta + i \sin \theta$. Теория вероятностей использована в объеме, включающем харак-

теристические функции и некоторые основные предельные теоремы. Теория случайных процессов развивается в книге только в необходимых пределах.

Как было отмечено выше, в конце каждой главы приводится большое число задач. Исключением служит лишь первая глава, являющаяся введением. Решения задач были подготовлены Полем Шаманом. Решения, на которые имеются ссылки в тексте или которые демонстрируют некоторые наиболее важные моменты, помещены в приложении В к этой книге.

Я весьма обязан и признателен П. Шаману за его большое содействие в подготовке этой книги. Он принимал участие в отборе материала и его редактировании, в подборе ссылок и задач, в усовершенствовании доказательств и изложения, в исправлении ошибок различного характера. Он прочел рукопись в нескольких вариантах. Мне всегда представлялось излишним говорить (как это обычно принято), что компетентный рецензент рукописи не несет ответственности за ошибки в публикации, так как очевидно, что всякий человек, любезно согласившийся просмотреть рукопись, не может брать на себя такой ответственности. Здесь же такой отказ от ответственности вполне уместен, поскольку П. Шаман исправил так много ошибок, что трудно поверить, чтобы какие-нибудь еще остались. Тем не менее я допускаю, что в подобном материале очень легко ошибиться, и читатель должен порицать за любую найденную ошибку автора (и информировать его об этом).

Я благодарен Д. Хинкли, Т. Сава и Дж. Стаяну, прочитавшим значительную часть рукописи и доказательств и принимавшим участие в подготовке библиографии и указателя. Я выражаю также признательность многим моим коллегам и ученикам за помощь различного рода. Дж. М. Крэдоку, С. У. Дж. Гренджеру, М. Дж. Кендаллу, А. Стюарту и Герману Вольду я весьма благодарен за то, что они разрешили мне использовать некоторые материалы.

В процессе работы над книгой весьма важным фактором была для меня длительная финансовая поддержка Военно-морского ведомства, оказывавшаяся в течение примерно десяти лет.

Отделения логики и математической статистики постоянно обеспечивали мне чрезвычайно благоприятные условия для работы.

Т. Андерсон

Станфордский университет,
Февраль 1970

ВВЕДЕНИЕ

Временным рядом называют последовательность наблюдений, обычно упорядоченную во времени, хотя возможно упорядочение и по какому-то другому параметру. Основной чертой, выделяющей анализ временных рядов среди других видов статистического анализа, является существенность порядка, в котором производятся наблюдения. Если во многих задачах наблюдения статистически независимы, то во временных рядах они, как правило, зависимы и характер этой зависимости может определяться положением наблюдений в последовательности. Природа ряда и структура порождающего ряд процесса могут предопределять порядок образования последовательности.

Почти в каждой области встречаются явления, которые интересно и важно изучать в их развитии и изменении во времени. В повседневной жизни могут представлять интерес, например, метеорологические условия, цены на тот или иной товар, те или иные характеристики состояния здоровья индивидуума и т. п. Все они изменяются во времени. Существуют также различные характеристики, относящиеся к целой нации и зависящие уже от совокупности характеристик отдельных индивидуумов, например экономические условия и народонаселение, которые эволюционируют и флуктуируют во времени. С течением времени изменяются деловая активность, режим протекания того или иного производственного процесса, глубина сна человека, восприятие телевизионной программы. Совокупность измерений какой-либо одной характеристики подобного рода в течение некоторого периода времени и представляет собой временной ряд. Это может быть, например, почасовая запись температуры в том или ином месте или ежегодное количество осадков, фиксируемое метеорологической станцией. Это могут быть также поквартальные данные о валовом национальном продукте.

К нескольким временным рядам приводит запись электрокардиограммы.

Цели изучения временных рядов могут быть различными. Можно, например, стремиться предсказать будущее на основании знания прошлого, управлять процессом, порождающим ряд, выяснить механизм, порождающий ряд, или просто сжато описать характерные особенности ряда. Как статистики мы будем интересоваться задачами статистических выводов. Именно, на основании ограниченного количества информации, временного ряда конечной длины, мы хотим делать выводы о вероятностном механизме, порождающем этот ряд, анализировать структуру, лежащую в его основе.

В принципе измерение многих величин, таких, как температура и напряжение, может производиться непрерывно. При этом наблюдения можно фиксировать в виде графика. Однако на практике измерения часто производятся все же в дискретные моменты времени. В других случаях, как, например, при подсчете урожайности зерновых культур, измерения вообще могут производиться только в определенные интервалы времени. Как бы то ни было, даже в том случае, когда изучаемые величины регистрируются непрерывно, при обработке их на цифровых вычислительных машинах реально используются только те значения, которые соответствуют дискретному множеству моментов времени. В настоящей книге мы ограничимся только временными рядами, представляющими собой дискретную последовательность наблюдений, производимых через регулярные промежутки времени, такими, как, например, почасовая запись атмосферного давления. Хотя часто имеет значение влияние одной характеристики на другую и взаимодействие во времени сразу нескольких величин, тем не менее во многих исследованиях существенных результатов можно достигнуть и при изучении только одного, отдельно взятого временного ряда. Настоящая книга почти целиком (за исключением того, что касается систем авторегрессии) посвящена статистическим методам анализа одномерных временных рядов. Иными словами, предполагается, что (один и тот же) объект или индивидуум подвергается повторным измерениям только одного типа. Мы будем предполагать, кроме того, что результатом измерения является действительное число, например температура, и что множество исходов не обязательно конечное (или счетное). О результате таких измерений часто говорят как о непрерывной переменной. Мы будем проводить математическое исследование результатов измерений подобного рода, обращаясь с ними так, как если бы сами измерения были непрерывными во времени. Например, ежегодный национальный доход можно измерить, в лучшем случае, с точностью до пенни. Однако его размеры могут быть столь велики, что не произойдет сколько-нибудь серьезной ошибки, если мы будем рассматривать эту величину как непрерывную переменную. Более того, мы будем рассматривать такие временные ряды, которые ве-

дут себя достаточно устойчивым образом, т. е. имеют тенденцию оставаться в определенных границах или по крайней мере меняются медленно, без резких, взрывных изменений. Так, мы рассматривали бы многие метеорологические переменные, но при этом исключили бы ударные волны.

Пусть наблюдаемым временным рядом является $y_1, y_2, \dots, y_T$. Мы будем понимать эту запись следующим образом. Имеется T чисел, представляющих собой наблюдение некоторой переменной в T равностоящих моментов времени. Эти моменты для удобства перенумерованы целыми числами $1, 2, \dots, T$. Достаточно общей математической (статистической, или вероятностной) моделью служит модель вида

$$(1) \quad y_t = f(t) + u_t, \quad t = 1, 2, \dots, T.$$

В этой модели наблюдаемый ряд рассматривается как сумма некоторой полностью детерминированной последовательности $\{f(t)\}$, которую можно назвать систематической составляющей, и случайной последовательности $\{u_t\}$, подчиняющейся некоторому вероятностному закону. (Иногда для этих двух составляющих используются соответственно термины сигнал и шум). Эти компоненты наблюдаемого ряда ненаблюдаемы; они являются теоретическими величинами. Например, если производится измерение количества ежедневно выпадаемых осадков, то $f(t)$ может представлять собой климатическую норму, получающуюся долговременным усреднением за большой период, а u_t те капризы и нерегулярности в погоде, которые характеризуют отклонения от климатической нормы. Точный смысл указанного разложения зависит не только от самих данных, но частично и от того, что понимается под повторением эксперимента, результатом которого являются эти данные. Мы используем здесь так называемую «частотную» интерпретацию. Мы полагаем, что по крайней мере принципиально можно повторять всю ситуацию целиком, получая новые совокупности наблюдений. При таком повторении эксперимента функция $f(t)$ должна была бы оставаться одной и той же, а случайные составляющие оказывались бы различными как различные реализации случайного процесса. Случайные составляющие, кроме всего прочего, могут включать в себя и ошибки наблюдений. (При этом ¹⁾ $f(t) = \mathbb{E}y_t$.)

Мы все имеем определенные интуитивные представления о том, что следовало бы понимать под временным параметром в подобных моделях или процессах. Одно из таких представлений состоит в том, что время течет в одном направлении. Другое — что события, близкие по времени, должны быть сравнительно сильно связаны, а события, разделенные большими промежутками времени, не должны

¹⁾ Символом $\mathbb{E}X$ в этой книге всюду обозначается математическое ожидание случайной величины X . — *Прим. перев.*

иметь сильной связи. Можно рассматривать различные варианты математической модели (1), в которых влияние времени может сказываться либо только на функции или последовательности $f(t)$, либо только на вероятностном процессе, определяющем случайную составляющую u_t , либо, наконец, на обеих этих компонентах. Первая часть книги посвящена анализу временных рядов, соответствующих так называемым моделям ошибок, в которых наблюдения рассматриваются как результат независимых случайных отклонений от некоторой функции, представляющей тренд. Во второй части книги мы будем иметь дело уже с последовательностями зависимых случайных величин, обычно со стационарными случайными процессами, выделяя при этом процессы авторегрессии. В конце книги будут рассмотрены модели, в которых на тренд накладывается случайная составляющая, образующая стационарный случайный процесс. Необходимые сведения о стационарных случайных процессах приводятся в гл. 7.

Во многих случаях модель можно определить с точностью до конечного числа параметров. Задачи статистических выводов будут связаны тогда именно с этими параметрами. В других ситуациях модель оказывается более неопределенной и приходится использовать непараметрические методы. Разумеется, модель должна достаточно хорошо представлять механизм образования соответствующего ряда. Однако, будучи математической абстракцией, она является лишь только приближением к реальному явлению. Сколь же точно можно определить модель, зависит от уровня знаний об исследуемом процессе и соответственно от той информации, которую мы можем получить с использованием статистического анализа, зависящего от характера этих знаний. В данной книге будет описано много методов и их свойств. Делается это для того, чтобы иметь возможность выбрать приемлемый метод, приводящий к полезным результатам. При этом внимание уделяется как самому статистическому выводу, так и его математическому обоснованию.

Первоначально анализ временных рядов базировался на моделях, в которых влияние временного параметра проявлялось только в систематической составляющей. Эту ситуацию можно было бы назвать классической, поскольку в известной степени она восходит к тем временам, когда Гаусс и другие развивали теорию и метод наименьших квадратов с целью применения их в астрономии и физике. В таких моделях мы предполагаем, что течение времени никак не отражается на случайной составляющей. Точнее говоря, мы предполагаем, что математическое ожидание (т. е. среднее значение) случайной составляющей тождественно равно нулю, дисперсия равна некоторой постоянной и что значения u_t в различные моменты времени некоррелированы. Такое определение приводит к тому, что всякую зависимость от времени приходится включать в систематическую составляющую $f(t)$. Последовательность $f(t)$ может

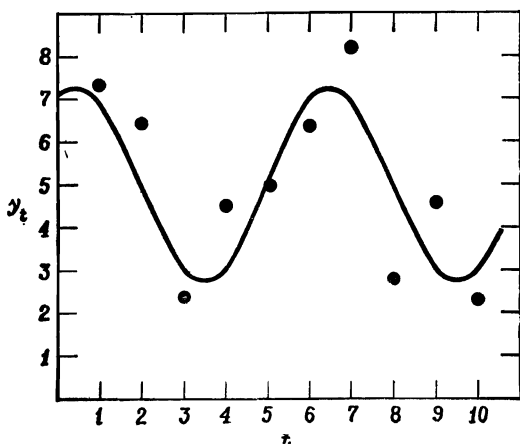


Рис. 1.1.

Ряд с тригонометрическим трендом.

зависеть от некоторых неизвестных коэффициентов и от известных величин, меняющихся со временем. В этом случае ее называют «функцией регрессии». Методы статистических выводов для коэффициентов функции регрессии оказываются полезными во многих областях статистики. Своеобразие же методов, относящихся именно к временным рядам, состоит в том, что здесь исследуются те модели, в которых упомянутые выше величины, меняющиеся со временем, являются известными функциями t .

В рамках сделанных ограничений можно выделить два типа временных последовательностей $f(t)$, часто называемых трендом. Один тип представляют медленно меняющиеся функции времени, примером которых могут служить полиномы достаточно низкой степени. К другому типу принадлежат циклические последовательности, например, конечные отрезки ряда Фурье, представляющие собой конечные суммы пар синусоидальных и косинусоидальных составляющих. Такой парой может являться $\alpha \cos \lambda t + \beta \sin \lambda t$ ($0 < \lambda < \pi$). Ее можно записать и с использованием одной только функции косинус, именно $\rho \cos(\lambda t - \theta)$. Период этой функции времени равен $2\pi/\lambda$, т. е. она повторяет свои значения всякий раз по прошествии времени $2\pi/\lambda$. Соответствующая частота, т. е. величина, обратная периоду, равна $\lambda/(2\pi)$. Коэффициент $\rho = \sqrt{\alpha^2 + \beta^2}$ является амплитудой, а θ — фазой указанной функции. Считается, что наблюдаемый ряд представляет собой сумму подобного отрезка ряда Фурье $f(t)$ и случайной составляющей. На рис. 1.1 представлены значения функции $y_t = 5 + 2 \cos 2\pi t/6 + \sin 2\pi t/6 + u_t$, где составляющая u_t нормально распределена с нулевым средним и единичной дисперсией. [Функция $f(t)$ представлена здесь в виде функ-

ции от непрерывного аргумента t .] Последовательные значения y_t разбросаны случайным образом по обе стороны от кривой $y = f(t)$. Если даже эта кривая известна и если известен закон распределения ошибки, то информация о значениях $y_1, \dots, y_{t-1}$ не оказывает в данной модели никакой помощи в предсказании значения y_t . Поведение графика функции $f(s)$ для $s > t - 1$ не зависит от значений $y_1, \dots, y_{t-1}$.

Подобная модель может оказаться приемлемой в ряде физических и экономических задач. В астрономии, например, $f(t)$ может описывать пространственное положение (по одной из координат) планеты в моменты времени t . Так как телескоп — прибор не идеальный, а состояние атмосферы постоянно изменяется, определение соответствующей координаты планеты производится с некоторой, хотя и достаточно малой, ошибкой. Эта ошибка наблюдения никак не влияет ни на последующие положения планеты, ни на реальные наблюдения этих положений. В случае свободно колеблющегося маятника его смещение (измеренное от нижнего положения) является тригонометрической функцией $\rho \cos(\lambda t - \theta)$.

Одной из общих моделей, в которой влияние временного параметра проявляется в случайной составляющей, является стационарный случайный процесс. Проиллюстрируем это примером процесса авторегрессии. Предположим, что y_1 имеет некоторое распределение с нулевым средним. Пусть y_1 и y_2 имеют совместное распределение, совпадающее с совместным распределением случайных величин y_1 и $\rho y_1 + u_2$, где u_2 не зависит от y_1 и имеет нулевое математическое ожидание. Совместное распределение $y_1, y_2, \dots, y_{t-1}, y_t$ для $t = 3, 4, \dots$ будем полагать в свою очередь таким же, как совместное распределение $y_1, y_2, \dots, y_{t-1}, \rho y_{t-1} + u_t$, причем предполагается, что случайная величина u_t не зависит от $y_1, \dots, y_{t-1}$ и имеет нулевое математическое ожидание. Если маргинальные распределения $u_2, u_3, \dots$ совпадают, а распределение y_1 выбрано надлежащим образом, то последовательность $\{y_t\}$ образует стационарный случайный процесс, именно процесс авторегрессии, и

$$(2) \quad y_t = \rho y_{t-1} + u_t$$

является стохастическим разностным уравнением первого порядка. Такое построение для $\rho = 1/2$ иллюстрирует рис. 1.2. В этой модели «возмущение» u_t оказывает влияние и на y_t , и на все последующие y_r . Из указанного построения вытекает, что условное математическое ожидание y_t при заданных значениях $y_1, \dots, y_{t-1}$ удовлетворяет равенству

$$(3) \quad \mathbb{E}(y_t | y_1, \dots, y_{t-1}) = \rho y_{t-1}.$$

(В действительности для процесса первого порядка значения y_t и $y_{t-2}, \dots, y_1$ условно независимы при заданном значении y_{t-1} .) Если

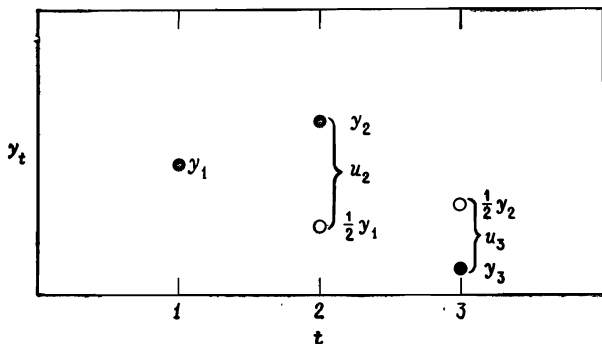


Рис. 1.2.

Построение ряда в модели авторегрессии.

мы хотим предсказать значение y_t по значениям $y_1, \dots, y_{t-1}$ и параметру ρ , то наилучшим прогнозом (в смысле минимума среднеквадратичной ошибки) будет ρy_{t-1} . Таким образом, в этой модели знание предшествующих наблюдений оказывает помощь в предсказывании y_t .

Процесс авторегрессии второго порядка получается, если взять совместное распределение $y_1, \dots, y_{t-1}, y_t$ таким же, как совместное распределение $y_1, y_2, \dots, y_{t-1}, \rho_1 y_{t-1} + \rho_2 y_{t-2} + u_t$, где u_t не зависит от $y_1, y_2, \dots, y_{t-1}$, $t = 3, 4, \dots$, а распределения y_1 и y_2 выбираются надлежащим образом. Графики подобных рядов представлены в приложении А.2. Графики других рядов, порожденных случайным моделированием, имеются у Кендалла и Стьюарта (1966, гл. 45) и у Вольда (1965, гл. 1). Переменная y_t может представлять собой смещение колеблющегося маятника, который подвержен некоторым случайным ударным воздействиям u_t . Тогда ряд y_t близок к тригонометрической функции $\rho \cos(\lambda t - \theta)$ с переменной амплитудой, переменной частотой и переменной фазой. Процесс авторегрессии четвертого порядка, порожденный моделью $y_t = \sum_{s=1}^4 \rho_s y_{t-s} + u_t$, будет походить уже на сумму двух тригонометрических функций с изменяющимися амплитудами, частотами и фазами.

Стационарный случайный процесс общего вида можно аппроксимировать процессом авторегрессии достаточно высокого порядка или процессом

$$(4) \quad \sum_{j=1}^q (A_j \cos \lambda_j t + B_j \sin \lambda_j t),$$

где $A_1, B_1, \dots, A_q, B_q$ — независимые случайные величины с $\mathcal{E}A_j = \mathcal{E}B_j = 0$ и $\mathcal{E}A_j^2 = \mathcal{E}B_j^2 = \varphi(\lambda_j)$. Последний является суммой

q тригонометрических функций со случайными амплитудами и фазами. Вклад тригонометрической функции с частотой $\lambda_j/(2\pi)$ в среднем пропорционален математическому ожиданию квадрата ее амплитуды, т. е. $2\phi(\lambda_j)$. Исходя из этого, стационарный случайный процесс (определенного класса) можно характеризовать *спектральной*

плотностью $f(\lambda)$, т. е. такой функцией, для которой $\int_a^b f(\lambda) d\lambda$ аппрок-

симируется суммой $\phi(\lambda_j)$ по всем λ_j , $a \leq \lambda_j < b$. Характерное свойство стационарных случайных процессов состоит в том, что ковариация $\mathcal{E}(y_t - \mathcal{E}y_t)(y_s - \mathcal{E}y_s)$ зависит только от разности $|t - s|$, так что ее можно обозначить через $\sigma(t - s)$. Ковариационная последовательность и спектральная плотность (если она существует) суть две альтернативные формы описания структуры моментов второго порядка стационарного случайного процесса. Ковариационная последовательность более удобна и информативна, когда большее значение имеет временной характер последовательности, как это, например, бывает во многих экономических рядах. Спектральная же плотность может оказаться более подходящей для других типов анализа. В частности, она весьма важна в физических науках, поскольку существо многих физических явлений может быть описано с помощью гармоник или тригонометрических функций времени. Так, поскольку изменение давления воздуха при наличии чистого тона выражается косинус-функцией, то для целей анализа звука естественно использовать анализ Фурье. В частности, удобным образом определяет высоту тона человеческое ухо.

Влияние времени может быть представлено в обеих составляющих модели так, что систематическая составляющая $f(t)$ является трендом во времени, а случайная составляющая u_t образует стационарный случайный процесс. К примеру экономический временной ряд может складываться из долговременного и сезонного изменений, которые вместе составляют $f(t)$, и из колебательной компоненты и других нерегулярностей, которые вместе образуют u_t и могут быть описаны процессом авторегрессии.

В тех случаях, когда тренд $f(t)$ имеет вполне определенную структуру и определяется конечным числом параметров, мы рассматриваем задачи статистических выводов о значениях этих параметров. Например, можно оценивать коэффициенты при синусах и косинусах в тригонометрических трендах. В первом случае может возникнуть вопрос о том, какую наивысшую степень t следует включить в рассмотрение, а во втором случае — вопрос о том, какие из нескольких слагаемых должны быть включены. Если же тренд не описывается столь точно, то для его оценивания можно использовать непараметрические методы, такие, как сглаживание.

Если случайный процесс описывается с помощью конечного

числа параметров, скажем, как процесс авторегрессии, то здесь также возникают задачи оценки коэффициентов, проверки гипотез относительно их значений или решения вопроса о том, какого порядка процесс следует использовать. Особый интерес представляет здесь задача проверки нулевой гипотезы о независимости случайных составляющих. Для этой цели может быть использован тот или иной сериальный коэффициент корреляции. Если процесс стационарный, но не описывается конечным числом параметров, то в этом случае можно оценивать ковариации $\{\sigma(h)\}$ или спектральную плотность. Соответствующие процедуры являются в основном непараметрическими.

Методы, представленные в настоящей книге, предназначены главным образом для получения выводов относительно структуры механизма, порождающего процесс. Указаны также методы предсказания последующих значений процесса для случая, когда структура известна. Если же структура случайного механизма оказывается неизвестной, ее можно оценить по имеющимся данным и затем уже для целей предсказания использовать найденные оценки.

ЛИТЕРАТУРА

Вольд (1965), Кендалл и Стьюарт (1966).

Глава 2

ИСПОЛЬЗОВАНИЕ РЕГРЕССИОННОГО АНАЛИЗА

2.1. ВВЕДЕНИЕ

Значительную часть статистических методов, используемых при анализе временных рядов, представляют методы регрессионного анализа (классической теории наименьших квадратов) или их видоизменения и аналоги. Независимые переменные могут быть заданными функциями времени, например степенями переменной t или тригонометрическими функциями от t . Сначала мы кратко опишем статистические процедуры в предположении некоррелированности случайных составляющих (§ 2.2 и 2.3). Соответствующие результаты будут использоваться при анализе тренда (гл. 3 и 4). Затем мы видоизменим эти процедуры применительно к произвольно коррелированным случайным составляющим, ковариационная матрица которых известна с точностью до постоянного множителя (§ 2.4).

В гл. 10 изучаются статистические процедуры, используемые при анализе тренда в предположении, что случайные составляющие образуют стационарный случайный процесс. При изложении регрессионного анализа определенное место уделяется асимптотической теории, которую можно использовать и при нарушении предположения о нормальном распределении случайных составляющих (§ 2.6).

Обобщения этих результатов полезны при оценивании коэффициентов стохастических разностных уравнений (гл. 5), поскольку в этом случае точные распределения использовать не удастся. Обобщения указанных результатов необходимы также, когда случайные составляющие образуют стационарный процесс более общего типа (§ 10.2).

2.2. ОБЩАЯ ТЕОРИЯ НАИМЕНЬШИХ КВАДРАТОВ

Рассмотрим некоррелированные случайные величины $y_1, y_2, \dots, y_T$, средние и дисперсии которых выражаются соотношениями

$$(1) \quad \mathbb{E} y_t = \sum_{i=1}^p \beta_i z_{it}, \quad t = 1, 2, \dots, T,$$

$$(2) \quad \mathbb{E} (y_t - \mathbb{E} y_t)^2 = \sigma^2, \quad t = 1, 2, \dots, T,$$

с заданными числовыми значениями z_{it} . При этом, z_{it} называются *независимыми* переменными, а y_t — *зависимыми* переменными. Используя векторные обозначения

$$(3) \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \mathbf{z}_t = \begin{pmatrix} z_{1t} \\ z_{2t} \\ \vdots \\ z_{pt} \end{pmatrix}, \quad t = 1, \dots, T,$$

выражение (1) можно записать в виде

$$(4) \quad \mathbb{E} y_t = \boldsymbol{\beta}' \mathbf{z}_t, \quad t = 1, \dots, T.$$

(Транспонирование вектора или матрицы $\mathbf{a}$ будет отмечаться штрихом: $\mathbf{a}'$.)

Будем обозначать через $\mathbf{b}$ оценку вектора $\boldsymbol{\beta}$, представляющую собой решение нормального уравнения

$$(5) \quad \mathbf{A} \mathbf{b} = \mathbf{c},$$

в котором

$$(6) \quad \mathbf{A} = \sum_{t=1}^T \mathbf{z}_t \mathbf{z}_t', \quad \mathbf{c} = \sum_{t=1}^T y_t \mathbf{z}_t,$$

и матрица $\mathbf{A}$ предполагается невырожденной (так что $T \geq p$). Вектор $\mathbf{b} = \mathbf{A}^{-1} \mathbf{c}$ минимизирует сумму $\sum_{t=1}^T (y_t - \tilde{\mathbf{b}}' \mathbf{z}_t)^2$ на множестве всех p -мерных векторов $\tilde{\mathbf{b}}$ и называется *оценкой наименьших квадратов* вектора $\boldsymbol{\beta}$. Несмещенная оценка s^2 для σ^2 может быть получена (при $T > p$) из соотношения

$$(7) \quad (T - p) s^2 = \sum_{t=1}^T (y_t - \mathbf{b}' \mathbf{z}_t)^2 = \sum_{t=1}^T y_t^2 - \mathbf{b}' \mathbf{A} \mathbf{b}.$$

Оценка наименьших квадратов $\mathbf{b}$ является несмещенной оценкой вектора $\boldsymbol{\beta}$,

$$(8) \quad \mathbb{E} \mathbf{b} = \boldsymbol{\beta},$$

и имеет ковариационную матрицу

$$(9) \quad \mathbb{E} (\mathbf{b} - \boldsymbol{\beta})(\mathbf{b} - \boldsymbol{\beta})' = \sigma^2 \mathbf{A}^{-1}.$$

Теорема Гаусса — Маркова утверждает, что компоненты вектора $\mathbf{b}$ являются наилучшими линейными несмещенными оценками соответствующих компонент вектора β в следующем смысле. Каждый элемент вектора $\mathbf{b}$ имеет дисперсию, наименьшую среди дисперсий всех несмещенных оценок соответствующего элемента вектора β , линейных по переменным $y_1, \dots, y_T$.

Если случайные величины $y_1, \dots, y_T$ независимы и нормально распределены, то $\mathbf{b}$ является оценкой максимального правдоподобия вектора β . Оценкой максимального правдоподобия для σ^2 служит в этом случае величина $(T - p) s^2 / T$. При этом оценка $\mathbf{b}$ распределена по многомерному нормальному закону $N(\beta, \sigma^2 \mathbf{A}^{-1})$ со средним β и ковариационной матрицей $\sigma^2 \mathbf{A}^{-1}$, а оценка $(T - p) s^2 / \sigma^2$ имеет χ^2 -распределение с числом степеней свободы, не зависящим от $\mathbf{b}$ и равным $T - p$. Вектор $\mathbf{b}$ и оценка s^2 образуют достаточное множество статистик для β и σ^2 .

В предположении нормальности случайных величин $y_1, \dots, y_T$ можно построить критерии для проверки гипотез относительно значений β_i и доверительные интервалы для β_i , $i = 1, \dots, p$. Пусть

$$(10) \quad \beta = \begin{pmatrix} \beta^{(1)} \\ \beta^{(2)} \end{pmatrix},$$

где

$$(11) \quad \beta^{(1)} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_r \end{pmatrix}, \quad \beta^{(2)} = \begin{pmatrix} \beta_{r+1} \\ \vdots \\ \beta_p \end{pmatrix}.$$

Аналогичным образом разобьем $\mathbf{z}_t$ и $\mathbf{b}$. Блочные векторы и матрицы рассмотрены у Т. Андерсона (1958, приложение 1, § 3). Тогда для проверки гипотезы H о том, что $\beta^{(2)} = \bar{\beta}^{(2)}$, где $\bar{\beta}^{(2)}$ — некоторый вектор, можно использовать F -статистику

$$(12) \quad \frac{(\mathbf{b}^{(2)} - \bar{\beta}^{(2)})' (\mathbf{A}^{22})^{-1} (\mathbf{b}^{(2)} - \bar{\beta}^{(2)})}{(p - r) s^2} = \frac{(\mathbf{b}^{(2)} - \bar{\beta}^{(2)})' (\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}) (\mathbf{b}^{(2)} - \bar{\beta}^{(2)})}{(p - r) s^2}.$$

Здесь $\mathbf{A}$ и $\mathbf{A}^{-1}$ разбиты на блоки с r и $p - r$ строками и столбцами:

$$(13) \quad \mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \quad \mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} \end{pmatrix}.$$

[По поводу равенства $\mathbf{A}^{22} = (\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12})^{-1}$ см. упр. 8.] Если выполнено предположение о нормальности и справедлива нулевая гипотеза, то статистика (12) имеет F -распределение с $p - r$ и $T - p$ степенями свободы. В общем случае в предположении нормальности

статистика (12) имеет нецентральное F -распределение с параметром нецентральности

$$(14) \quad \frac{(\beta^{(2)} - \bar{\beta}^{(2)})' (A_{22} - A_{21} A_{11}^{-1} A_{12}) (\beta^{(2)} - \bar{\beta}^{(2)})}{\sigma^2}.$$

Эти результаты являются следствием того, что $\mathbf{b}^{(2)}$ имеет нормальное распределение $N(\beta^{(2)}, \sigma^2 A^{22})$. [См., например, Т. Андерсон (1958, § 2.4).]

Если $\bar{\beta}^{(2)} = 0$, то нулевая гипотеза означает, что элементы $\mathbf{z}_i^{(2)}$ не входят в функцию регрессии. При этом говорят, что величины $y_1, \dots, y_T$ не зависят от векторов $\mathbf{z}_i^{(2)}$. В этом важном случае числитель в (12) есть просто $\mathbf{b}^{(2)'} (A_{22} - A_{21} A_{11}^{-1} A_{12}) \mathbf{b}^{(2)}$.

Доверительная область для $\beta^{(2)}$ с коэффициентом доверия $1 - \varepsilon$ имеет вид

$$(15) \quad \left\{ \beta^2 \left| \frac{(\mathbf{b}^{(2)} - \beta^{(2)})' (A_{22} - A_{21} A_{11}^{-1} A_{12}) (\mathbf{b}^{(2)} - \beta^{(2)})}{(p-r) s^2} \leq F_{p-r, T-p}(\varepsilon) \right. \right\},$$

где $F_{p-r, T-p}(\varepsilon)$ есть верхняя 100ε-процентная точка ¹⁾ F -распределения с $p-r$ и $T-p$ степенями свободы.

Если интерес представляет только один элемент вектора β , то вместо F -статистики можно использовать t -статистику. Пусть, например, нас интересует элемент β_p . Тогда $A^{22} = a^{pp}$ есть число, и отношение $(b_p - \beta_p)/(s \sqrt{a^{pp}})$ имеет t -распределение с $T-p$ степенями свободы.

Заметим, что остатки $\hat{y}_i = y_i - \mathbf{b}' \mathbf{z}_i$ не коррелированы с независимыми переменными $\mathbf{z}_i$ в выборке:

$$(16) \quad \sum_{i=1}^T \hat{y}_i \mathbf{z}_i = \sum_{i=1}^T y_i \mathbf{z}_i - \sum_{i=1}^T \mathbf{z}_i \mathbf{z}_i' \mathbf{b} = 0,$$

а множество этих остатков не коррелировано с вектором выборочной регрессии в генеральной совокупности. (См. упр. 7.)

Сказанное поясняет следующая геометрическая интерпретация (см. рис. 2.1). Пусть $\mathbf{y} = (y_1, \dots, y_T)'$ — вектор в T -мерном евклидовом пространстве, r столбцов матрицы $\mathbf{Z}_1 = (\mathbf{z}_1^{(1)}, \dots, \mathbf{z}_T^{(1)})'$ представляют собой r векторов в этом пространстве, а $p-r$ столбцов матрицы $\mathbf{Z}_2 = (\mathbf{z}_1^{(2)}, \dots, \mathbf{z}_T^{(2)})'$ суть $p-r$ его векторов. Пусть при этом $\mathbf{Z} = (\mathbf{Z}_1 \mathbf{Z}_2)$. Тогда математическое ожидание вектора $\mathbf{y}$ выражается в виде $\mathbf{Z}\beta$ и является вектором в p -мерном подпространстве, натянутом на столбцы матрицы $\mathbf{Z}$. Выборочная регрессия $\mathbf{Zb}$ является проекцией вектора $\mathbf{y}$ на это p -мерное подпространство.

¹⁾ То есть значение, выше которого лежит 100ε процентов распределения.—

Вектор остатков $y - Zb$ ортогонален каждому вектору этого p -мерного подпространства и является проекцией y на $(T - p)$ -мерное подпространство, ортогональное столбцам матрицы Z .

На рис. 2.2 представлено p -мерное подпространство, порожденное столбцами матрицы Z . Проекция вектора $Zb - Z\beta = Z(b - \beta)$ на r -мерное подпространство, порожденное столбцами матрицы Z_1 , равна $Z_1 b^{*(1)} - Z_1 \beta^{*(1)} = Z_1 (b^{*(1)} - \beta^{*(1)})$, где $b^{*(1)} = (Z_1' Z_1)^{-1} Z_1' y$. (См. § 2.3.) Проекция на $(p - r)$ -мерное подпространство Z , ортогональное Z_1 , равна $Z(b - \beta) - Z_1(b^{*(1)} - \beta^{*(1)}) = (Z_2 - Z_1 A_{11}^{-1} A_{12})(b^{(2)} - \beta^{(2)})$. Числитель F -статистики (12) равен квадрату длины последнего вектора, а знаменатель пропорционален квадрату длины вектора $y - Zb$.

2.3. ЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ НЕЗАВИСИМЫХ ПЕРЕМЕННЫХ; ОРТОГОНАЛЬНЫЕ НЕЗАВИСИМЫЕ ПЕРЕМЕННЫЕ

В представлении функции регрессии независимые переменные могут быть отнесены к различным p -координатным системам. При этом некоторые координатные системы могут оказаться предпочтительнее других.

Пусть $z_t^* = Gz_t$, $t = 1, \dots, T$, где G — произвольная невырожденная матрица, и пусть $\beta = G'\beta^*$. Тогда ξ_{y_t} можно записать в виде

$$(1) \quad \beta' z_t = \beta^{*'} G G^{-1} z_t^* = \beta^{*'} z_t^*.$$

Компоненты векторов z_t^* , $t = 1, \dots, T$, являются координатами векторов z_t в новой координатной системе. Оценки для β^* и σ^2 по наблюдениям $y_1, \dots, y_T, z_1, \dots, z_T$ выражаются соотношениями

$$(2) \quad b^* = A^{*-1} \sum_{t=1}^T z_t^* y_t = (GAG')^{-1} G \sum_{t=1}^T z_t y_t = (G')^{-1} b,$$

$$(3) \quad (T - p) s^{*2} = \sum_{t=1}^T (y_t - b^{*'} z_t^*)^2 = \sum_{t=1}^T (y_t - b' z_t)^2 = (T - p) s^2.$$

В последнем использовано равенство

$$(4) \quad b^{*'} z_t^* = [(G')^{-1} b]' G z_t = b' z_t.$$

Функции, которые взяты в качестве оценок для функции регрессии, в обеих координатных системах принимают одинаковые значения $b' z_t = b^{*'} z_t^*$.

Независимые переменные можно разбить на два множества, $z_t' = (z_t^{(1)'} z_t^{(2)'})$, и особо интересоваться множеством $z_t^{(2)'}$. Например, может представлять интерес проверка нулевой гипотезы $\beta^{(2)} = 0$.

В этом случае удобно преобразовать независимые переменные с помощью матрицы

$$(5) \quad \mathbf{G} = \begin{pmatrix} \mathbf{G}_{11} & 0 \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{pmatrix},$$

где матрица $\mathbf{G}_{11}$ квадратная. Тогда

$$(6) \quad \mathbf{G}^{-1} = \begin{pmatrix} \mathbf{G}_{11}^{-1} & 0 \\ -\mathbf{G}_{22}^{-1} \mathbf{G}_{21} \mathbf{G}_{11}^{-1} & \mathbf{G}_{22}^{-1} \end{pmatrix},$$

$$(7) \quad \begin{aligned} \beta^* &= \begin{pmatrix} \beta^{*(1)} \\ \beta^{*(2)} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_{11}^{-1} & 0 \\ 0 & (\mathbf{G}_{22}^{-1})' \end{pmatrix} \begin{pmatrix} \beta^{(1)} \\ \beta^{(2)} \end{pmatrix} = \\ &= \begin{pmatrix} (\mathbf{G}_{11}^{-1})' \beta^{(1)} - (\mathbf{G}_{22}^{-1} \mathbf{G}_{21} \mathbf{G}_{11}^{-1})' \beta^{(2)} \\ (\mathbf{G}_{22}^{-1})' \beta^{(2)} \end{pmatrix}. \end{aligned}$$

Гипотеза $H: \beta^{(2)} = 0$ эквивалентна, таким образом, гипотезе $H: \beta^{*(2)} = 0$. В результате преобразования вектора $\mathbf{z}_t$ получим

$$(8) \quad \begin{pmatrix} \mathbf{z}_t^{*(1)} \\ \mathbf{z}_t^{*(2)} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_{11} & 0 \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{z}_t^{(1)} \\ \mathbf{z}_t^{(2)} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_{11} \mathbf{z}_t^{(1)} \\ \mathbf{G}_{21} \mathbf{z}_t^{(1)} + \mathbf{G}_{22} \mathbf{z}_t^{(2)} \end{pmatrix}.$$

Поскольку $\mathbf{b}^{*(2)} = (\mathbf{G}_{22}^{-1})' \mathbf{b}^{(2)}$ и

$$(9) \quad \mathbf{A}_{22}^* - \mathbf{A}_{21}^* (\mathbf{A}_{11}^*)^{-1} \mathbf{A}_{12}^* = \mathbf{G}_{22} (\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}) \mathbf{G}_{22}',$$

то F -критерии для проверки обеих нулевых гипотез совпадают.

Если в качестве матрицы $\mathbf{G}$ взять матрицу

$$(10) \quad \begin{pmatrix} \mathbf{G}_{11} & 0 \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & 0 \\ -\mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{I} \end{pmatrix},$$

то векторы $\mathbf{z}_t^{*(1)} = \mathbf{z}_t^{(1)}$ и $\mathbf{z}_t^{*(2)} = \mathbf{z}_t^{(2)} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{z}_t^{(1)}$ становятся ортогональными, т. е.

$$(11) \quad \mathbf{A}_{21}^* = \mathbf{A}_{12}^* = \sum_{t=1}^T \mathbf{z}_t^{*(2)} \mathbf{z}_t^{*(1)'} = \mathbf{A}_{21} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{11} = 0.$$

Поскольку же

$$(12) \quad \mathbf{A}_{22}^* = \mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12},$$

то статистика F -критерия принимает в этом случае вид $\mathbf{b}^{*(2)'} \mathbf{A}_{22}^* \mathbf{b}^{*(2)} / [(p-r)s^2]$.

Линейное преобразование $\mathbf{G}$ можно выбрать таким образом, чтобы все компоненты вектора $\mathbf{z}_t^*$ были попарно ортогональны. Пусть

$$(13) \quad \mathbf{z}_t^* = \mathbf{\Gamma} \mathbf{z}_t,$$

где

$$(14) \quad \Gamma = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \gamma_{21} & 1 & 0 & \dots & 0 \\ \gamma_{31} & \gamma_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \gamma_{p1} & \gamma_{p2} & \gamma_{p3} & \dots & 1 \end{bmatrix}.$$

Тогда $z_{1l}^* = z_{1l}$ и

$$(15) \quad z_{kl}^* = z_{kl} + \sum_{j=1}^{k-1} \gamma_{kj} z_{jl}, \quad k = 2, \dots, p,$$

т. е. z_{kl}^* зависит только от тех z_{jl} , для которых $j \leq k$. Условия ортогональности вектора $(z_{k1}^*, \dots, z_{kT}^*)$ к $(z_{11}^*, \dots, z_{1T}^*), \dots, (z_{k-1,1}^*, \dots, z_{k-1,T}^*)$ совпадают с условиями ортогональности вектора $(z_{k1}, \dots, z_{kT})$ к $(z_{11}, \dots, z_{1T}), \dots, (z_{k-1,1}, \dots, z_{k-1,T})$ и имеют вид

$$(16) \quad (\gamma_{k1}, \dots, \gamma_{k,k-1}) \begin{pmatrix} a_{11} & \dots & a_{1,k-1} \\ \vdots & & \vdots \\ a_{k-1,1} & \dots & a_{k-1,k-1} \end{pmatrix} = (a_{k1}, \dots, a_{k,k-1}).$$

При этих условиях матрица

$$(17) \quad \sum_{i=1}^T z_i^* z_i^{**} = A^* = \Gamma A \Gamma' = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22}^* & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{pp}^* \end{bmatrix}$$

диагональна. Для компонент соответственно имеем соотношения

$$(18) \quad \sum_{i=1}^T z_{il}^* z_{il}^* = 0, \quad i \neq l.$$

(Если ортогональные переменные, кроме того, нормированы делением z_{il}^* на $\sqrt{a_{ii}^*}$, то соответствующая процедура ортогонализации известна под названием процесса ортогонализации Грама — Шмидта.)

Если независимые переменные ортогональны, то формулы и вычисления по ним значительно упрощаются. Поскольку

$$(19) \quad A^{*-1} = \begin{bmatrix} a_{11}^{*-1} & 0 & \dots & 0 \\ 0 & a_{22}^{*-1} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{pp}^{*-1} \end{bmatrix},$$

то элементы векторов $\mathbf{b}^*$ некоррелированы и имеют дисперсии, равные σ^2/a_{ii} , $i = 1, \dots, p$. При этом нормальные уравнения принимают простой вид

$$(20) \quad b_i^* = \frac{c_i^*}{a_{ii}^*} = \frac{1}{a_{ii}^*} \sum_{t=1}^T y_t z_{it}^*, \quad i = 1, \dots, p,$$

а формула для оценки дисперсии переходит в

$$(21) \quad (T-p) s^2 = \sum_{t=1}^T y_t^2 - \sum_{i=1}^p a_{ii}^* b_i^{*2}.$$

В этом случае F -статистика (12) из § 2.2 для проверки гипотезы $\beta^{*(2)} = \bar{\beta}^{*(2)}$ равна

$$(22) \quad \frac{\sum_{i=r+1}^p a_{ii}^* (b_i^* - \bar{b}_i^*)^2}{(p-r) s^2} = F_{p-r, T-p}.$$

Иногда независимые переменные с самого начала выбираются ортогональными. Как будет показано в гл. 4, тригонометрические последовательности $\{\cos 2\pi jt/T\}$, $j = 0, 1, \dots, \left[\frac{1}{2} T\right]$, и $\{\sin 2\pi kt/T\}$, $k = 1, \dots, \left[\frac{1}{2} (T-1)\right]$ ортогональны для $t = 1, \dots, T$ и могут быть использованы как компоненты вектора $\mathbf{z}_t$. Ортогонализировать можно любое множество независимых переменных, но проводить эту операцию не имеет особого смысла, если данное множество переменных используется только один раз. Напротив, если одна и та же совокупность независимых переменных используется многократно, то ортогонализация независимых переменных может дать большой выигрыш, сокращая объем вычислений.

Примером использования ортогонализации независимых переменных может служить полиномиальная регрессия. Предположим, что $z_{it} = t^{i-1}$ и

$$(23) \quad \mathcal{E} y_t = \beta_1 + \beta_2 t + \dots + \beta_p t^{p-1}, \quad t = 1, \dots, T.$$

Степени переменной t можно заменить ортогональными полиномами $\varphi_{0T}(t) = 1$, $\varphi_{1T}(t)$, $\dots$, $\varphi_{p-1,T}(t)$, имеющими вид

$$(24) \quad \varphi_{kT}(t) = t^k + C_{k-1}(k, T) t^{k-1} + \dots + C_1(k, T) t + C_0(k, T).$$

Здесь коэффициенты C зависят от длины ряда T и степени полинома k и определяются соотношениями

$$(25) \quad \sum_{t=1}^T \varphi_{iT}(t) \varphi_{jT}(t) = 0, \quad i \neq j.$$

При этом

$$(26) \quad \xi y_t = \gamma_0 \varphi_{0T}(t) + \gamma_1 \varphi_{1T}(t) + \dots + \gamma_{p-1} \varphi_{p-1,T}(t).$$

Более подробно ортогональные полиномы рассмотрены в § 3.2.

Большинство вычислительных методов решения нормальных уравнений $\mathbf{A}\mathbf{b} = \mathbf{c}$ включает в себя так называемые прямое и обратное решения¹⁾. Прямое решение состоит из последовательности операций над строками матрицы $(\mathbf{A}\mathbf{c})$, в результате которой $\mathbf{A}$ приводится к треугольному виду. При этом $(\mathbf{A}\mathbf{c})$ преобразуется в матрицу $(\bar{\mathbf{A}}\mathbf{c})$: $\mathbf{D}(\mathbf{A}\mathbf{c}) = (\bar{\mathbf{A}}\mathbf{c})$, или

$$(27) \quad \begin{bmatrix} 1 & 0 & \dots & 0 \\ d_{21} & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ d_{p1} & d_{p2} & \dots & 1 \end{bmatrix} (\mathbf{A}\mathbf{c}) = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} & \dots & \bar{a}_{1p} & \bar{c}_1 \\ 0 & \bar{a}_{22} & \dots & \bar{a}_{2p} & \bar{c}_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & \bar{a}_{pp} & \bar{c}_p \end{bmatrix}.$$

Матрица $\mathbf{D}$ имеет здесь форму матрицы Γ . Исследование выражения (27) показывает, что каждая строка из $\mathbf{D}$ совпадает с соответствующей строкой матрицы Γ , так что $\mathbf{D} = \Gamma$. [Различные методы приведения $\mathbf{A}$ к треугольному виду отличаются только последовательностью операций и являются в алгебраическом смысле эквивалентными при одинаково упорядоченных наборах переменных. Отметим, что прямое решение уравнения (27) для некоторого k является частью прямого решения для любого последующего k .] Таким образом,

$$(28) \quad \bar{a}_{kk} = \sum_{t=1}^T z_{kt} z_{kt}^* = \sum_{t=1}^T (z_{kt})^2 = a_{kk},$$

$$(29) \quad \bar{c}_k = \sum_{t=1}^T z_{kt} y_t^* = c_k.$$

Коэффициенты выборочной регрессии для ортогонализированных переменных $b_k^* = c_k^*/a_{kk}^*$ можно получить из прямого решения нормальных уравнений: $b_k^* = \bar{c}_k/\bar{a}_{kk}$. Фактически во многих вычислительных методах, таких, как метод Дулитла, каждая строка матрицы $(\bar{\mathbf{A}}\mathbf{c})$ делится на старший отличный от нуля элемент и запоминается. При этом последний элемент каждой строки является коэффициентом регрессии при соответствующей ортогональной переменной. Таким образом, прямое решение связано с теми же алгебраическими преобразованиями, которые используются при определении ортогональных переменных. Существенное отличие, конечно, состоит в том, что вычисление ортогональных переменных связано с полу-

¹⁾ В советской литературе их обычно называют *прямым* и *обратным ходом* решения. — Прим. перев.

чением pT чисел z_{kt}^* . Отметим в заключение, что значение выражения $b^{(2)'} (A_{22} - A_{21}A_{11}^{-1}A_{12}) b^{(2)}$ можно получить из прямого решения как $\sum_{k=r+1}^p b_k^* c_k^* = \sum_{k=r+1}^p \bar{c}_k^2 / \bar{a}_{kk}$. (См. упр. 12.)

2.4. КОРРЕЛИРОВАННЫЕ ПЕРЕМЕННЫЕ

Если зависимые переменные коррелированы и ковариационная матрица известна с точностью до постоянного множителя (или точно), то развитую выше теорию можно соответствующим образом видоизменить. Предположим, что

$$(1) \quad \xi y_t = \beta' z_t, \quad t = 1, \dots, T,$$

$$(2) \quad \xi (y_t - \beta' z_t) (y_s - \beta' z_s) = \sigma_{ts}, \quad t, s = 1, \dots, T.$$

Здесь z_t — известный вектор-столбец из p чисел, $t = 1, \dots, T$, а $\sigma_{ts} = \sigma^2 \psi_{ts}$, $t, s = 1, \dots, T$, где величины ψ_{ts} известны. Удобно записать эту модель в более компактной матричной форме. Пусть $y = (y_1, \dots, y_T)'$, $Z = (z_1, \dots, z_T)'$, $\Sigma = (\sigma_{ts})$ и $\Psi = (\psi_{ts})$.

Тогда (1) и (2) принимают вид

$$(3) \quad \xi y = Z\beta,$$

$$(4) \quad \xi (y - Z\beta) (y - Z\beta)' = \Sigma,$$

где $\Sigma = \sigma^2 \Psi$, а Ψ — известная матрица. Пусть матрица D удовлетворяет соотношению

$$(5) \quad D\Psi D' = I.$$

Положим $Dy = x = (x_1, \dots, x_T)'$ и $DZ = W = (w_1, \dots, w_T)'$. Умножая (3) на D слева, а (4) на D слева и на D' справа, приходим к модели

$$(6) \quad \xi x = W\beta,$$

$$(7) \quad \xi (x - W\beta) (x - W\beta)' = \sigma^2 I,$$

изучавшейся в § 2.2. Нормальное уравнение $Ab = c$, в котором $A = W'W$ и $c = W'x$, эквивалентно

$$(8) \quad Z'D'DZb = Z'D'Dy.$$

Из (5) видно, что $\Psi = D^{-1} (D')^{-1} = (D'D)^{-1}$ и $\Psi^{-1} = D'D$. Поэтому решением уравнения (8) является

$$(9) \quad b = (Z'\Psi^{-1}Z)^{-1} Z'\Psi^{-1}y.$$

Отсюда получаем

$$(10) \quad \xi b = (Z'\Psi^{-1}Z)^{-1} Z'\Psi^{-1}\xi y = \beta,$$

$$\begin{aligned}
 (11) \quad \mathcal{E}(\mathbf{b} - \beta)(\mathbf{b} - \beta)' &= (\mathbf{Z}'\Psi^{-1}\mathbf{Z})^{-1} \mathbf{Z}'\Psi^{-1} \times \\
 &\times \mathcal{E}(\mathbf{y} - \mathbf{Z}\beta)(\mathbf{y} - \mathbf{Z}\beta)' \Psi^{-1} \mathbf{Z}(\mathbf{Z}'\Psi^{-1}\mathbf{Z})^{-1} = \\
 &= \sigma^2 (\mathbf{Z}'\Psi^{-1}\mathbf{Z})^{-1}.
 \end{aligned}$$

Каждый элемент вектора $\mathbf{b}$ является наилучшей несмещенной линейной оценкой соответствующей компоненты вектора β (теорема Гаусса — Маркова). При этом вектор $\mathbf{b}$ называется марковской оценкой для β . Он минимизирует квадратичную форму $(\mathbf{y} - \mathbf{Z}\tilde{\mathbf{b}})' \times \times \Psi^{-1}(\mathbf{y} - \mathbf{Z}\tilde{\mathbf{b}})$. Если $y_1, \dots, y_T$ имеют совместное нормальное распределение, то $\mathbf{b}$ и $\hat{\sigma}^2 = (T - p) s^2/T$ являются оценками максимально правдоподобия для параметров β и σ^2 и образуют для этих параметров достаточное множество статистик.

Оценка наименьших квадратов

$$(12) \quad \mathbf{b}_L = (\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}'\mathbf{y}$$

является несмещенной, т. е.

$$(13) \quad \mathcal{E}\mathbf{b}_L = (\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}'\mathcal{E}\mathbf{y} = \beta,$$

и имеет ковариационную матрицу

$$\begin{aligned}
 (14) \quad \mathcal{E}(\mathbf{b}_L - \beta)(\mathbf{b}_L - \beta)' &= (\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}'\mathcal{E}(\mathbf{y} - \mathbf{Z}\beta)(\mathbf{y} - \mathbf{Z}\beta)' \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} = \\
 &= \sigma^2 (\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}'\Psi\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}.
 \end{aligned}$$

Если только столбцы матрицы $\mathbf{Z}$ не связаны специальным образом с Ψ , то любая линейная комбинация компонент вектора $\mathbf{b}_L$, например $\gamma'\mathbf{b}_L$, будет иметь дисперсию, большую чем дисперсия соответствующей линейной комбинации $\gamma'\mathbf{b}$ компонент вектора $\mathbf{b}$. В этом случае разность выражений (14) и (11) будет положительно полуопределенной матрицей.

ТЕОРЕМА 2.4.1. Если $\mathbf{Z} = \mathbf{V}^*\mathbf{C}$, причем p столбцов матрицы $\mathbf{V}^*$ являются линейно независимыми характеристическими векторами матрицы Ψ , а $\mathbf{C}$ — невырожденная матрица, то оценка наименьших квадратов (12) совпадает с марковской оценкой (9).

Доказательство. Условие на $\mathbf{V}^*$ означает, что

$$(15) \quad \Psi\mathbf{V}^* = \mathbf{V}^*\Lambda^*,$$

где Λ^* — диагональная матрица, состоящая из (положительных) характеристических корней матрицы Ψ , соответствующих столбцам матрицы $\mathbf{V}^*$. Тогда справедливо равенство

$$(16) \quad \mathbf{V}^*\Psi^{-1} = \Lambda^{*-1}\mathbf{V}^{**},$$

и $\mathbf{V}^{**}\Psi^{-1}\mathbf{V}^* = \Lambda^{*-1}\mathbf{V}^{**}\mathbf{V}^*$, а оценка наименьших квадратов равна

$$(17) \quad \mathbf{b}_L = \mathbf{C}^{-1}(\mathbf{V}^{**}\mathbf{V}^*)^{-1} \mathbf{V}^{**}\mathbf{y},$$

а марковская оценка

$$(18) \quad \begin{aligned} \mathbf{b} &= (\mathbf{C}'\mathbf{V}^*\mathbf{\Psi}^{-1}\mathbf{V}^*\mathbf{C})^{-1} \mathbf{C}'\mathbf{V}^*\mathbf{\Psi}^{-1}\mathbf{y} = \\ &= (\mathbf{C}'\mathbf{\Lambda}^{*-1}\mathbf{V}^*\mathbf{V}^*\mathbf{C})^{-1} \mathbf{C}'\mathbf{\Lambda}^{*-1}\mathbf{V}^*\mathbf{y}. \end{aligned}$$

Правая часть (18) идентична правой части (17). ■

В разд. 10.2.1 будет показано, что условие теоремы 2.4.1 является не только достаточным, но и необходимым. Его можно сформулировать и иначе: существует p линейно независимых линейных комбинаций столбцов матрицы $\mathbf{Z}$, являющихся характеристическими векторами матрицы $\mathbf{\Psi}$. Смысл теоремы заключается в том, что при выполнении указанных условий оценки наименьших квадратов (для случая, когда матрица $\mathbf{\Psi}$ неизвестна) являются несмещенными линейными оценками с наименьшей дисперсией. В разд. 10.2.1 будет рассмотрен случай, когда существует произвольное число линейно независимых комбинаций столбцов матрицы $\mathbf{Z}$, являющихся характеристическими векторами матрицы $\mathbf{\Psi}$. Оценки наименьших квадратов для коэффициентов этих линейных комбинаций совпадают с марковскими, если остальные независимые переменные ортогональны данным. Утверждение о том, что при выполнении условий теоремы 2.4.1 оценки наименьших квадратов являются и оценками максимального правдоподобия, было доказано для случая нормального распределения Т. Андерсоном (1948).

2.5. ПРОГНОЗИРОВАНИЕ

Займемся теперь прогнозированием значения y_τ в момент $t = \tau$. Если функция регрессии известна, то известно также ξy_τ , и оно будет наилучшим образом прогнозировать значение y_τ в том смысле, что при этом минимизируется среднеквадратичная ошибка прогноза.

Предположим теперь, что имеются наблюдения $y_1, \dots, y_T$, по которым мы хотим предсказать значение y_τ ($\tau > T$), причем $\xi y_\tau =$

$= \sum_{k=1}^p \beta_k z_{k\tau} = \beta' z_\tau$, где β неизвестный, а z_τ известный векторы. Представляется разумным оценивать вектор β с помощью оценки наименьших квадратов $\mathbf{b}$ и в качестве прогноза y_τ использовать $\mathbf{b}' z_\tau$. Займемся обоснованием такой процедуры. При этом будем рас-

сматривать только линейные прогнозы $\sum_{i=1}^T d_i y_i$. Коэффициенты d_i могут зависеть от $z_1, \dots, z_T$ и z_τ . Прежде всего потребуем, чтобы прогноз был несмещенным, т. е., чтобы

$$(1) \quad \xi \left(\sum_{i=1}^T d_i y_i - y_\tau \right) = 0.$$

Тем самым, прогноз должен быть несмещенной оценкой для $\mathbb{E}y_t = \beta'z_t$. Мы требуем также минимальности дисперсии (или, что равносильно, среднеквадратичной ошибки прогноза)

$$(2) \quad \mathbb{E} \left(\sum_{t=1}^T d_t y_t - y_t \right)^2 = \mathbb{E} \left[\left(\sum_{t=1}^T d_t y_t - \beta'z_t \right) - (y_t - \mathbb{E}y_t) \right]^2 = \\ = \mathbb{E} \left(\sum_{t=1}^T d_t y_t - \beta'z_t \right)^2 + \sigma^2.$$

Итак, задача состоит в том, чтобы отыскать несмещенную линейную оценку с наименьшей дисперсией для линейной комбинации коэффициентов регрессии $\beta'z_t$. Из теоремы Гаусса — Маркова следует, что такой оценкой является $b'z_t$. Это вытекает из предыдущих рассуждений, поскольку данную модель можно преобразовать таким образом, что $\beta'z_t$ будет компонентой β^* , если один из столбцов G^{-1} будет совпадать с z_t . Дисперсия оценки $b'z_t$ равна

$$(3) \quad \mathbb{E}(b'z_t - \beta'z_t)^2 = \mathbb{E}z_t' (b - \beta) (b - \beta)' z_t = \sigma^2 z_t' A^{-1} z_t.$$

Среднеквадратичная ошибка прогноза есть

$$(4) \quad \mathbb{E}(b'z_t - y_t)^2 = \sigma^2 (1 + z_t' A^{-1} z_t).$$

Указанное свойство прогноза $b'z_t$ можно сформулировать иначе. Этот прогноз минимизирует среднеквадратичную ошибку прогноза в классе всех линейных прогнозов, имеющих ограниченную среднеквадратичную ошибку. (См. упр. 15.)

Если предполагать, что наблюдения подчиняются нормальному закону, то можно построить доверительный интервал для y_t . При этом предположении случайная величина $b'z_t - y_t$ распределена нормально с нулевым средним и дисперсией (4) и не зависит от s^2 . Поэтому величина

$$(5) \quad \frac{b'z_t - y_t}{s \sqrt{1 + z_t' A^{-1} z_t}}$$

имеет t -распределение с $T - p$ степенями свободы. Доверительный интервал для y_t с коэффициентом доверия $1 - \varepsilon$ имеет вид

$$(6) \quad b'z_t - t_{T-p}(\varepsilon) s \sqrt{1 + z_t' A^{-1} z_t} \leq y_t \leq b'z_t + t_{T-p}(\varepsilon) \times \\ \times s \sqrt{1 + z_t' A^{-1} z_t}.$$

Здесь $t_{T-p}(\varepsilon)$ определяется из условия, что вероятность попадания случайной величины, имеющей t -распределение с $T - p$ степенями свободы, в интервал $(-t_{T-p}(\varepsilon), t_{T-p}(\varepsilon))$ равна $1 - \varepsilon$.

Изучим теперь ошибку прогноза для случая, когда некоторыми независимыми переменными пренебрегают. Предположим, что компоненты вектора z_t перенумерованы таким образом, что z_t' разби-

вается, как и прежде, на блоки $(z_t^{(1)'} z_t^{(2)'})$ и что вкладом от $z_t^{(2)}$ пренебрегают, т. е. предполагают ошибочно, что регрессия является линейной функцией только от $z_t^{(1)}$, а не от всего z_t . Запишем для удобства $\beta' z_t$ в виде $\beta^{*'} z_t^*$, где β^* определяется соотношением (7) § 2.3, причем $G_{11} = I$, $G_{21} = -A_{21}A_{11}^{-1}$, $G_{22} = I$ и

$$(7) \quad z_t^* = Gz_t = \begin{pmatrix} I & 0 \\ -A_{21} & A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} z_t^{(1)} \\ z_t^{(2)} \end{pmatrix} = \begin{pmatrix} z_t^{(1)} \\ z_t^{(2)} - A_{21}A_{11}^{-1}z_t^{(1)} \end{pmatrix} = \begin{pmatrix} z_t^{*(1)} \\ z_t^{*(2)} \end{pmatrix}.$$

При этом векторы $z_t^{*(1)}$ и $z_t^{*(2)}$ ортогональны, $t = 1, \dots, T$, и

$$(8) \quad A^* = \begin{pmatrix} A_{11}^* & 0 \\ 0 & A_{22}^* \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix}.$$

Оценкой для $\beta^{*(1)}$ является

$$(9) \quad b^{*(1)} = A_{11}^{-1}c^{(1)} = A_{11}^{-1} \sum_{t=1}^T z_t^{(1)} y_t.$$

Ее статистические свойства выражаются формулами

$$(10) \quad \mathbb{E} b^{*(1)} = A_{11}^{-1} \sum_{t=1}^T z_t^{(1)} (z_t^{(1)'} \beta^{*(1)} + z_t^{*(2)'} \beta^{*(2)}) =$$

$$= A_{11}^{-1} (A_{11} \beta^{*(1)} + 0 \beta^{*(2)}) = \beta^{*(1)},$$

$$(11) \quad \mathbb{E} (b^{*(1)} - \beta^{*(1)}) (b^{*(1)} - \beta^{*(1)})' = \sigma^2 A_{11}^{-1}.$$

Прогноз для момента τ ($\tau > T$) равен $b^{*(1)'} z_\tau^{*(1)} = b^{*(1)'} z_\tau^{(1)}$ и имеет смещение

$$(12) \quad \mathbb{E} (b^{*(1)'} z_\tau^{(1)} - y_\tau) = \beta^{*(1)'} z_\tau^{(1)} - \beta' z_\tau = -\beta^{*(2)'} z_\tau^{*(2)} = \\ = -\beta^{(2)'} (z_\tau^{(2)} - A_{21}A_{11}^{-1}z_\tau^{(1)}).$$

Его дисперсия равна

$$(13) \quad \mathbb{E} (b^{*(1)'} z_\tau^{(1)} - \beta^{*(1)'} z_\tau^{(1)})^2 = \sigma^2 z_\tau^{(1)'} A_{11}^{-1} z_\tau^{(1)},$$

а среднеквадратичная ошибка есть

$$(14) \quad \mathbb{E} (b^{*(1)'} z_\tau^{(1)} - y_\tau)^2 = \\ = \beta^{(2)'} (z_\tau^{(2)} - A_{21}A_{11}^{-1}z_\tau^{(1)}) (z_\tau^{(2)'} - z_\tau^{(1)'} A_{11}^{-1}A_{12}) \beta^{(2)} + \sigma^2 z_\tau^{(1)'} A_{11}^{-1} z_\tau^{(1)} + \sigma^2.$$

Пренебрежение вкладом от $z_t^{(2)}$ приводит, в общем случае, к смещению прогноза, но уменьшает его дисперсию. (См. упр. 18.)

В некоторых случаях можно ожидать, что прогнозирование проводится в ситуациях, когда векторы $\mathbf{z}_t$ для последующих значений t подобны векторам $\mathbf{z}_t$ для ранее наблюдававшихся значений. Сумма квадратов смещений

$$\begin{aligned}
 (15) \quad \sum_{t=1}^T [\beta^{(2)'} (z_t^{(2)} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} z_t^{(1)})]^2 &= \\
 &= \sum_{t=1}^T \beta^{(2)'} (z_t^{(2)} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} z_t^{(1)}) (z_t^{(2)'} - z_t^{(1)'} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}) \beta^{(2)} = \\
 &= \beta^{(2)'} (\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}) \beta^{(2)}
 \end{aligned}$$

пропорциональна параметру нецентральности распределения F -статистики для проверки гипотезы $\beta^{(2)} = 0$. Этот факт можно считать еще одним основанием для предпочтения F -критерия всем другим критериям для проверки этой гипотезы.

2.6. АСИМПТОТИЧЕСКАЯ ТЕОРИЯ

Приведенные процедуры проверки гипотез и построения доверительных областей были основаны на предположении о том, что наблюдения распределены нормально. Если предположение о нормальности не выполняется, то эти процедуры все же можно применять для больших выборок, используя асимптотическую теорию.

ТЕОРЕМА 2.6.1. Пусть $y_t = \beta' \mathbf{z}_t + u_t$, $t = 1, 2, \dots$, где все u_t независимы, имеют нулевые средние, дисперсии σ^2 и функции распределения $F_t(u)$, $t = 1, 2, \dots$. Положим $\mathbf{A}_T = \sum_{t=1}^T \mathbf{z}_t \mathbf{z}_t'$, $\mathbf{c}_T = \sum_{t=1}^T y_t \mathbf{z}_t$, $\mathbf{b}_T = \mathbf{A}_T^{-1} \mathbf{c}_T$,

$$(1) \quad \mathbf{D}_T = \begin{bmatrix} \sqrt{a_{11}^T} & 0 & \dots & 0 \\ 0 & \sqrt{a_{22}^T} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \sqrt{a_{pp}^T} \end{bmatrix}$$

и

$$(2) \quad \mathbf{R}_T = \mathbf{D}_T^{-1} \mathbf{A}_T \mathbf{D}_T^{-1}.$$

Предположим далее, что (i) $a_{ii}^T \rightarrow \infty$ при $T \rightarrow \infty$, $i = 1, \dots, p$, (ii) $z_{i,T+1}^2 / a_{ii}^T \rightarrow 0$ при $T \rightarrow \infty$, $i = 1, \dots, p$, (iii) $\mathbf{R}_T \rightarrow \mathbf{R}_\infty$ при

$T \rightarrow \infty$, (iv) $\mathbf{R}_\infty$ невырождена и (v)

$$(3) \quad \sup_{i=1,2,\dots} \int_{|u|>c} u^2 dF_i(u) \rightarrow 0$$

при $c \rightarrow \infty$. Тогда $\mathbf{D}_T(\mathbf{b}_T - \beta)$ имеет в пределе нормальное распределение с нулевым средним и ковариационной матрицей $\sigma^2 \mathbf{R}_\infty^{-1}$.

Доказательство. Прежде всего имеем

$$(4) \quad \begin{aligned} \mathbf{D}_T(\mathbf{b}_T - \beta) &= \mathbf{D}_T(\mathbf{A}_T^{-1} \mathbf{c}_T - \beta) = \\ &= \mathbf{D}_T \left[\mathbf{A}_T^{-1} \sum_{i=1}^T \mathbf{z}_i (u_i + \mathbf{z}_i' \beta) - \beta \right] = \\ &= (\mathbf{D}_T^{-1} \mathbf{A}_T \mathbf{D}_T^{-1})^{-1} \mathbf{D}_T^{-1} \sum_{i=1}^T \mathbf{z}_i u_i. \end{aligned}$$

Мы докажем, что $\mathbf{D}_T^{-1} \sum_{i=1}^T \mathbf{z}_i u_i$ имеет в пределе нормальное распределение с нулевым средним и ковариационной матрицей $\sigma^2 \mathbf{R}_\infty$, если удастся показать, что для любого вектора α , $\alpha \neq 0$, величина $\alpha' \mathbf{D}_T^{-1} \sum_{i=1}^T \mathbf{z}_i u_i$ имеет в пределе нормальное распределение с нулевым средним и дисперсией $\sigma^2 \alpha' \mathbf{R}_\infty \alpha$. (См. теорему 7.7.7.) Пусть $\gamma_i^T = \alpha' \mathbf{D}_T^{-1} \mathbf{z}_i = \sum_{j=1}^p (\alpha_j z_{ij} / \sqrt{a_{jj}^T})$, $i = 1, \dots, T$, $T = 1, 2, \dots$; тогда

$$(5) \quad \text{Var} \left(\sum_{i=1}^T \gamma_i^T u_i \right) = \sigma^2 \sum_{i=1}^T (\gamma_i^T)^2 = \sigma^2 \alpha' \mathbf{R}_T \alpha.$$

Положим $w_i^T = \gamma_i^T u_i / (\sigma \sqrt{\alpha' \mathbf{R}_T \alpha})$. Тогда $\mathbb{E} w_i^T = 0$, $\sum_{i=1}^T \text{Var}(w_i^T) = 1$ и

$$(6) \quad \begin{aligned} \sum_{i=1}^T \int_{|w|>\delta} w^2 dF_i^T(w) &= \sum_{i=1}^T \frac{(\gamma_i^T)^2}{\sigma^2 \alpha' \mathbf{R}_T \alpha} \int_{u^2 > \delta^2 [\alpha' \mathbf{R}_T \alpha / (\gamma_i^T)^2] \sigma^2} u^2 dF_i(u) \leq \\ &\leq \frac{1}{\sigma^2} \sup_{i=1,2,\dots} \int_{u^2 > \delta^2 [\alpha' \mathbf{R}_T \alpha / (\gamma_i^T)^2] \sigma^2} u^2 dF_i(u) \rightarrow 0, \end{aligned}$$

где $F_i^T(w)$ — функция распределения случайной величины w_i^T , $i = 1, \dots, T$. Сходимость к нулю правой части (6) вытекает из (3) в силу неравенства

$$(7) \quad \max_{i=1,\dots,T} |\gamma_i^T| = \max_{i=1,\dots,T} \left| \sum_{j=1}^p \alpha_j \frac{z_{ij}}{\sqrt{a_{jj}^T}} \right| \leq \sum_{j=1}^p |\alpha_j| \max_{i=1,\dots,T} \frac{|z_{ij}|}{\sqrt{\sum_{s=1}^T z_{js}^2}},$$

правая часть которого сходится к нулю в силу (i), (ii) и приводимой ниже леммы 2.6.1. Выполнение условия (6) Линдеберга — Феллера влечет за собой нормальность предельного распределения

$\alpha' \mathbf{D}_T^{-1} \sum_{i=1}^T \mathbf{z}_i u_i$. [См. Лозв (1963, § 21.2) или теорему 7.7.2.] Таким образом,

$\mathbf{D}_T^{-1} \sum_{i=1}^T \mathbf{z}_i u_i$ имеет в пределе нормальное распределение с нулевым средним и ковариационной матрицей $\sigma^2 \mathbf{R}_\infty$, а $\mathbf{D}_T (\mathbf{b}_T - \beta)$ имеет в пределе нормальное распределение с нулевым средним и ковариационной матрицей $\sigma^2 \mathbf{R}_\infty^{-1}$. ■

ЛЕММА 2.6.1. Если (i) $a_{ii}^T \rightarrow \infty$ и (ii) $z_{i,T+1}^2/a_{ii}^T \rightarrow 0$, то

$$(8) \quad \frac{\max_{i=1, \dots, T} z_{ii}^2}{a_{ii}^T} \rightarrow 0$$

при $T \rightarrow \infty$. Обратно, из условия (8) вытекают условия (i) и (ii).

ДОКАЗАТЕЛЬСТВО. Пусть $t(T)$ — наибольшее значение t , $t = 1, \dots, T$, для которого $z_{ii}^2(t) = \max_{i=1, \dots, T} z_{ii}^2$. При этом $\{t(T)\}$ — неубывающая последовательность целых чисел. Если она ограничена, то пусть ее максимум равен τ . Тогда

$$(9) \quad \frac{\max_{i=1, \dots, T} z_{ii}^2}{a_{ii}^T} \leq \frac{z_{i\tau}^{(2)}}{a_{ii}^T} \rightarrow 0.$$

в силу (i). Если же $t(T) \rightarrow \infty$ при $T \rightarrow \infty$, то

$$(10) \quad \frac{\max_{i=1, \dots, T} z_{ii}^2}{a_{ii}^T} \leq \frac{z_{ii}^2(t(T))}{a_{ii}^{t(T)-1}} \rightarrow 0$$

в силу (ii). Доказательство обратного утверждения предоставляется читателю. ■

СЛЕДСТВИЕ 2.6.1. Пусть $y_t = \beta' \mathbf{z}_t + u_t$, $t = 1, 2, \dots$, где все u_t независимы и одинаково распределены с нулевыми средними и дисперсиями σ^2 . Если выполнены условия (i) — (iv) теоремы 2.6.1, то $\mathbf{D}_T (\mathbf{b}_T - \beta)$ имеет в пределе нормальное распределение с нулевым средним и ковариационной матрицей $\sigma^2 \mathbf{R}_\infty^{-1}$.

СЛЕДСТВИЕ 2.6.2. Пусть $y_t = \beta' \mathbf{z}_t + u_t$, $t = 1, 2, \dots$, где все u_t независимы, имеют нулевые средние и дисперсии σ^2 , и пусть выполнены условия (i) — (iv) теоремы 2.6.1. Если существуют такие $\delta > 0$ и $M > 0$, что $\sum |u_t|^{2+\delta} \leq M$, $t = 1, 2, \dots$, то $\mathbf{D}_T (\mathbf{b}_T - \beta)$ имеет в пределе нормальное распределение с нулевым средним и ковариационной матрицей $\sigma^2 \mathbf{R}_\infty^{-1}$.

Теорема 2.6.1 в несколько отличной форме была получена Эйке-ром (1963). Оба следствия используют более ограничительные достаточные условия. Если в следствии 2.6.2 условие (ii) заменить условием равномерной ограниченности $\mathbf{z}_i, \mathbf{z}_i$, то соответствующий результат может быть достигнут прямым применением центральной предельной теоремы Ляпунова. (См. упр. 20.) Однако условие ограниченности $\mathbf{z}_i, \mathbf{z}_i$ слишком обременительно для наших целей, поскольку полиномы от t ему не удовлетворяют.

ТЕОРЕМА 2.6.2. В условиях теоремы 2.6.1, следствия 2.6.1 или следствия 2.6.2 статистика s^2 сходится по вероятности к σ^2 .

ДОКАЗАТЕЛЬСТВО Поскольку

$$(11) \quad \sum_{t=1}^T (y_t - \beta' \mathbf{z}_t)^2 = \sum_{t=1}^T [y_t - \mathbf{b}_T' \mathbf{z}_t + (\mathbf{b}_T - \beta)' \mathbf{z}_t]^2 = \\ = \sum_{t=1}^T (y_t - \mathbf{b}_T' \mathbf{z}_t)^2 + (\mathbf{b}_T - \beta)' \mathbf{A}_T (\mathbf{b}_T - \beta),$$

то

$$(12) \quad s^2 = \frac{T}{T-p} \frac{\sum_{t=1}^T (y_t - \beta' \mathbf{z}_t)^2}{T} - (\mathbf{b}_T - \beta)' \frac{1}{T-p} \mathbf{A}_T (\mathbf{b}_T - \beta).$$

Второй член в (12) неотрицателен. Его математическое ожидание равно

$$(13) \quad \frac{1}{T-p} \mathbb{E}(\mathbf{b}_T - \beta)' \mathbf{A}_T (\mathbf{b}_T - \beta) = \frac{1}{T-p} \text{tr } \mathbb{E} \mathbf{A}_T (\mathbf{b}_T - \beta) (\mathbf{b}_T - \beta)' = \\ = \frac{p}{T-p} \sigma^2$$

и стремится к нулю при $T \rightarrow \infty$. Отсюда в силу неравенства Чебышева следует, что второй член в (12) сходится по вероятности к нулю. Утверждение теоремы вытекает теперь из закона больших чисел:

$$(14) \quad \text{plim}_{T \rightarrow \infty} \frac{\sum_{t=1}^T (y_t - \beta' \mathbf{z}_t)^2}{T} = \text{plim}_{T \rightarrow \infty} \frac{\sum_{t=1}^T u_t^2}{T} = \text{plim}_{T \rightarrow \infty} \frac{\sum_{t=1}^T x_t}{T} = \sigma^2,$$

где $x_t = u_t^2$ и $\mathbb{E} x_t = \mathbb{E} u_t^2 = \sigma^2$. (Здесь plim означает сходимость по вероятности. — *Ред.*) Этот закон можно применить, поскольку в условиях леммы 2.6.1 x_t одинаково распределены, в условиях леммы 2.6.2 $\mathbb{E} |x_t|^{1+\delta/2} < M$ [см. Лоэв (1963, § 20.1)], а в условиях теоремы 2.6.1

$$(15) \quad \sup_{t=1,2,\dots} \int_{x>d} x dG_t(x) \rightarrow 0$$

при $d \rightarrow \infty$, где $G_i(x)$ — функция распределения случайной величины $x_i = u_i^2$. [См. Лозв (1963, § 20.2) и упр. 21.] ■

Значение приведенных теорем состоит в том, что, опираясь на них, обычную теорию для нормального случая при больших объемах выборок можно использовать с достаточной точностью и в тех ситуациях, когда наблюдения не являются нормально распределенными. Мы увидим в § 5.5, что в случае процесса авторегрессии, в котором $\beta'z_t$ заменяется линейной комбинацией наблюдений y_t при запаздывающих значениях переменной t , возможно дальнейшее развитие асимптотической теории, оправдывающей применение соответствующих процедур для больших выборок, когда предположение о нормальности не выполняется. В § 5.5 асимптотическая теория будет обобщена и представлена более подробно.

В разд. 10.2.4 подобные теоремы будут доказаны для последовательностей $\{u_t\}$, образующих стационарный случайный процесс типа скользящего среднего.

ЛИТЕРАТУРА

Теория регрессии рассматривается более полно в ряде монографий, в числе которых: Грейбилл (1961), Дрейпер и Смит (1966), Кемпторн (1952), Кендалл и Стьюарт (1946а), Плэккетт (1960), Уилкс (1962) и Уильямс (1959), Шеффе (1959).

§ 2.2. Т. Андерсон (1958).

§ 2.4. Т. Андерсон (1948).

§ 2.6. Лозв (1963), Эйкер (1963).

УПРАЖНЕНИЯ

1. (§ 2.2) Докажите, что вектор $\mathbf{b}$, определенный в (5), минимизирует $\sum_{t=1}^T (y_t - \mathbf{b}'\mathbf{z}_t)^2$ относительно $\mathbf{b}$. [Указание. Показать, что

$$\sum_{t=1}^T (y_t - \mathbf{b}'\mathbf{z}_t)^2 = \sum_{t=1}^T (y_t - \tilde{\mathbf{b}}'\mathbf{z}_t)^2 + (\mathbf{b} - \tilde{\mathbf{b}})' \mathbf{A} (\mathbf{b} - \tilde{\mathbf{b}}).]$$

2. (§ 2.2) Проверьте (8) и (9).

3. (§ 2.2) Докажите теорему Гаусса — Маркова. [Указание. Показать, что если компоненты $\sum_{t=1}^T \mathbf{a}_t y_t$ являются несмещенными оценками компонент вектора β , где $\mathbf{a}_t = \mathbf{A}^{-1} \mathbf{z}_t + \mathbf{d}_t$, $t = 1, \dots, T$, а $\mathbf{a}_1, \dots, \mathbf{a}_T$ и $\mathbf{d}_1, \dots, \mathbf{d}_T$ суть p -мерные векторы,

то $\sum_{t=1}^T d_t z_t' = 0$. Показать, что дисперсии этих оценок будут диагональными элементами матрицы $\sigma^2 A^{-1} + \sigma^2 \sum_{t=1}^T d_t d_t'$]

4. (§ 2.2) Покажите, что $\mathbf{b}$ и $\hat{\sigma}^2 = (T - p) s^2 / T$ являются оценками максимального правдоподобия для $\boldsymbol{\beta}$ и σ^2 , если $y_1, \dots, y_T$ независимы и нормально распределены.

5. (§ 2.2) Покажите, что $\mathbf{b}$ и s^2 образуют множество достаточных статистик для оценки значений $\boldsymbol{\beta}$ и σ^2 , если $y_1, \dots, y_T$ независимы и нормально распределены.

[Указание. Показать, что плотность имеет вид

$$(2\pi\sigma^2)^{-\frac{1}{2}T} \exp \left\{ -\frac{1}{2} [(\mathbf{b} - \boldsymbol{\beta})' \mathbf{A} (\mathbf{b} - \boldsymbol{\beta}) + (T - p) s^2] / \sigma^2 \right\}.$$

6. (§ 2.2) Докажите, что если $\xi \mathbf{y} = \mathbf{Z}\boldsymbol{\beta}$ и $\xi (\mathbf{y} - \mathbf{Z}\boldsymbol{\beta}) (\mathbf{y} - \mathbf{Z}\boldsymbol{\beta})' = \sigma^2 \mathbf{I}$, то ковариационная матрица остатков $\mathbf{y} - \mathbf{Z}\mathbf{b}$ равна $\sigma^2 (\mathbf{I} - \mathbf{Z}\mathbf{A}^{-1}\mathbf{Z}')$.

7. (§ 2.2) Докажите, что в условиях упр. 6 имеет место соотношение $\xi (\mathbf{y} - \mathbf{Z}\mathbf{b}) (\mathbf{b} - \boldsymbol{\beta})' = 0$.

8. (§ 2.2) Пусть невырожденная матрица $\mathbf{A}$ и обратная ей матрица $\mathbf{B} = \mathbf{A}^{-1}$ одинаковым образом разбиты на блоки

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix},$$

причем матрица $\mathbf{A}_{11}$ невырождена. Решая уравнение $\mathbf{A}\mathbf{B} = \mathbf{I}$ в блочном виде, убедитесь, что

$$(a) \quad \mathbf{B}_{12}\mathbf{B}_{22}^{-1} = -\mathbf{A}_{11}^{-1}\mathbf{A}_{12},$$

$$(b) \quad \mathbf{B}_{22} = (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}.$$

9. (§ 2.3) Покажите, что если матрица $\mathbf{G} = (g_{ij})$ невырождена и имеет треугольную форму, так что $g_{ij} = 0$, $i < j$, то таковой же является и матрица $\mathbf{G}^{-1}$.

10. (§ 2.3) Пусть $\tilde{y}_t$ — разность между y_t и выборочной регрессией на $z_t^{(1)}$ и $\tilde{z}_t^{(2)}$ — разность между $z_t^{(2)}$ и его формальной выборочной регрессией на $z_t^{(1)}$.

(a) Покажите, что $\xi \tilde{y}_t = \boldsymbol{\beta}^{(2)'} \tilde{z}_t^{(2)}$.

(b) Покажите, что оценка наименьших квадратов для $\boldsymbol{\beta}^{(2)}$, построенная по $\tilde{y}_t$ и $\tilde{z}_t^{(2)}$ в соответствии с (a), совпадает с оценкой, построенной по y_t и z_t .

11. (§ 2.3) Докажите, что $\mathbf{D} = \mathbf{\Gamma}$.

12. (§ 2.3) Пусть z_{kt}^* — переменные, полученные ортогонализацией переменных z_{kt} , $t = 1, \dots, T$, а $\boldsymbol{\beta}^{**} = (\boldsymbol{\beta}^{*(1)'} \boldsymbol{\beta}^{*(2)'})'$ — соответствующим образом преобразованный вектор $\boldsymbol{\beta}' = (\boldsymbol{\beta}^{(1)'} \boldsymbol{\beta}^{(2)'})'$. Докажите, что статистика (12) из § 2.2 для про-

верки гипотезы $H: \beta^{(2)} = 0$ имеет вид

$$\frac{\sum_{i=r+1}^p a_{ii}^* (b_i^*)^2}{(p-r)s^2} = \frac{\sum_{i=r+1}^p b_i^* c_i^*}{(p-r)s^2} = \frac{\sum_{i=r+1}^p (c_i^*)^2 / a_{ii}^*}{(p-r)s^2}.$$

13. (§ 2.4) Докажите алгебраически, что разность между (14) и (11) положительно полуопределена. [Указание. Эта разность отличается лишь множителем σ^2 от матрицы

$$(Z'Z)^{-1} [Z'\Psi Z - Z'Z (Z'\Psi^{-1}Z)^{-1} Z'Z] (Z'Z)^{-1},$$

которая положительно полуопределена, если таковой же является матрица

$$\begin{pmatrix} Z'\Psi^{-1}Z & Z'Z \\ Z'Z & Z'\Psi Z \end{pmatrix} = \begin{pmatrix} Z'\Psi^{-1} \\ Z' \end{pmatrix} \Psi (\Psi^{-1}Z \ Z).$$

14. (§ 2.5) Проверьте, что из теоремы Гаусса — Маркова вытекает, что $b'z_\tau$ является наилучшей линейной несмещенной оценкой для $\beta'z_\tau$.

15. (§ 2.5) Покажите, что если линейная оценка $\sum_{t=1}^T k_t y_t$ математического ожидания прогнозируемого значения является смещенной, то средняя квадратичная ошибка неограничена. [Указание. Если $\sum_{t=1}^T k_t y_t \neq \beta'z_\tau$ для $\beta = \gamma$, то рассмотреть $\beta = k\gamma$ при $k \rightarrow \infty$.]

16. (§ 2.5) Покажите, что односторонние доверительные интервалы для прогноза y_τ с коэффициентом доверия $(1 - \varepsilon)$ задаются соотношениями

$$y_\tau \leq b'z_\tau + t_{T-p}(2\varepsilon) s \sqrt{1 + z'_\tau A^{-1} z_\tau}$$

и

$$b'z_\tau - t_{T-p}(2\varepsilon) s \sqrt{1 + z'_\tau A^{-1} z_\tau} \leq y_\tau.$$

17. (§ 2.5) Покажите, что доверительная область для прогнозов y_τ и y_ρ ($\tau > T$, $\rho > T$, $\tau \neq \rho$) с коэффициентом доверия $(1 - \varepsilon)$ выражается соотношением

$$(b'z_\tau - y_\tau \quad b'z_\rho - y_\rho) \begin{pmatrix} 1 + z'_\tau A^{-1} z_\tau & z'_\tau A^{-1} z_\rho \\ z'_\rho A^{-1} z_\tau & 1 + z'_\rho A^{-1} z_\rho \end{pmatrix}^{-1} \begin{pmatrix} b'z_\tau - y_\tau \\ b'z_\rho - y_\rho \end{pmatrix} \leq 2s^2 F_{2, T-p}(\varepsilon)$$

18. (§ 2.5) Докажите, что

$$z_\tau^{(1)'} A_{11}^{-1} z_\tau^{(1)} \leq z'_\tau A^{-1} z_\tau.$$

[Указание. Использовать (7).]

19. (§ 2.6) Докажите следствие 2.6.2, используя теорему 2.6.1.

20. (§ 2.6) Докажите следствие 2.6.2 (без использования теоремы 2.6.1), в котором условие (ii) заменено условием (ii') существования константы L , такой, что $z'_t z_t \leq L$, $t = 1, 2, \dots$.

21. (§ 2.6) Докажите, что из (15) при $T \rightarrow \infty$ вытекает

$$(i) \quad \sum_{t=1}^T \int_{x>T} dG_t(x) \rightarrow 0,$$

$$(ii) \quad \frac{1}{T} \sum_{t=1}^T \int_{x<T} x dG_t(x) \rightarrow \sigma^2,$$

$$(iii) \quad \frac{1}{T^2} \sum_{t=1}^T \left\{ \int_{x<T} x^2 dG_t(x) - \left[\int_{x<T} x dG_t(x) \right]^2 \right\} \rightarrow 0$$

[условие Лозва (1963, § 20.2)]. [Указание. Для произвольной константы d' и $T > d'$ показать, что первый член последнего соотношения оценивается неравенством

$$\begin{aligned} \frac{1}{T^2} \sum_{t=1}^T \int_{0 \leq x \leq d'} x^2 dG_t(x) + \frac{1}{T^2} \sum_{t=1}^T \int_{d' < x < T} x^2 dG_t(x) \leq \\ \leq \frac{\sigma^2 d'}{T} + \frac{1}{T} \sum_{t=1}^T \int_{x>d'} x dG_t(x). \end{aligned}$$

ТРЕНДЫ И СГЛАЖИВАНИЕ

3.1. ВВЕДЕНИЕ

Мы будем изучать так называемые модели ошибок, в которых наблюдаемые временные ряды интерпретируются как сумма систематических составляющих, или тренда, и случайных составляющих, или ошибки. В этой главе мы рассмотрим модели, в которых подлежащий изучению тренд с течением времени гладко возрастает или убывает, но не повторяется регулярным образом. В следующей главе будут рассмотрены периодические тренды, поведение которых приблизительно одинаково на различных участках оси времени. Предполагается, что случайные составляющие имеют в каждый момент времени одинаковые дисперсии и некоррелированы. Они могут представлять собой ошибки наблюдения или нерегулярности иного рода. Предположения о равенстве дисперсий и отсутствии корреляции являются определенным приближением к действительному положению вещей. В гл. 10 мы рассмотрим трудности, возникающие в тех случаях, когда эти предположения не выполняются. Иногда наблюдения лучше соответствуют условиям равенства дисперсий и аддитивности ошибки, если преобразовать масштаб измерений изучаемой величины. Например, в ряде экономических исследований производится анализ не самих цен, а их логарифмов. Некоторые линейные комбинации значений переменной в последовательные моменты времени могут быть коррелированы в значительно меньшей степени, по сравнению со значениями самой переменной. Если коэффициенты такой линейной комбинации (например, конечной разности) известны, то ее можно использовать в качестве объекта изучения.

В ряде случаев исследуемый тренд является известной функцией времени или других наблюдаемых величин и некоторых (возможно неизвестных) параметров. Если эта функция зависит от параметров линейно, то имеем обычную регрессию, о которой упоминалось в

гл. 2. Однако некоторые функции времени, как, например, кривые роста, не линейны по параметрам. В таких случаях оценка параметров и проверка гипотез о значениях этих параметров являются далеко не простыми задачами. Мы кратко рассмотрим их в § 3.5.

В другой категории случаев тренд является функцией времени или каких-то иных величин, которая не известна, но может быть адекватно представлена в виде линейной комбинации известных функций времени. Периодический тренд может быть представлен в виде линейной комбинации тригонометрических составляющих (т. е. в виде конечного отрезка ряда Фурье). Статистические аспекты этого приближения рассматриваются в гл. 4. Если тренд флуктуирует, явно смещаясь в сторону возрастания или в сторону убывания, то такой тренд часто можно хорошо приблизить полиномами. Исследование подобного случая производится в § 3.2.

Иногда тренд является неизвестной функцией времени, вид которой настолько меняется за большой промежуток времени, что она не может быть приближена полиномом низкой степени или коротким отрезком ряда Фурье. Методы, использующие свойство «гладкости», изучаются в § 3.3 и 3.4. Эти методы являются непараметрическими.

3.2. ПОЛИНОМИАЛЬНЫЕ ТРЕНДЫ

3.2.1. Ортогональные полиномы

Встречающиеся во многих исследованиях временные ряды наряду с флуктуациями и нерегулярностями имеют некоторую общую тенденцию изменения. Например, многие временные ряды в экономике являются, в первом приближении, возрастающими. Такого рода общую тенденцию мы будем называть *трендом*. Во многих случаях желательно сделать выводы о тренде на основании самого наблюдаемого временного ряда, хотя при этом может быть известно, что тренд возникает в силу действия каких-то иных факторов. Так, рост временного ряда в экономике может быть обусловлен ростом численности населения. В ряде случаев, когда теория не может указать явный вид тренда как функции времени, тем не менее бывает возможно приблизить тренд полиномом от t достаточно низкой степени. В простейшем нетривиальном случае равномерного возрастания или убывания значений ряда адекватное представление тренда может дать полином первой степени, т. е. линейная функция.

Полиномиальный тренд есть в первую очередь средство описания. Он содержит в сжатой форме общие характеристики ряда. Для использования тренда в этом качестве полином должен иметь достаточно низкую степень. Во многих случаях коэффициентам полинома

нельзя придать никакого реального смысла. Такой полином служит заменой гораздо более сложной (но неизвестной) функции времени. Подобранный полином обычно может быть использован для интерполяции, однако использовать его для экстраполяции следует с осторожностью, поскольку вопрос о качестве приближения рассматриваемого тренда данным полиномом не может быть решен исходя лишь из наблюдаемых значений.

В некоторых случаях первостепенный интерес представляют короткие по времени периоды, флуктуации и нерегулярности. В таких случаях тренд желательно выделить, с тем чтобы получить базу, от которой можно было бы измерять быстрые изменения наблюдаемого ряда.

Наша основная модель ошибок состоит в том, что наблюдаемая величина y_i представлена в виде суммы временного тренда $f(t)$ и (ненаблюдаемой) ошибки u_i ; т. е. $y_i = f(t) + u_i$, где все u_i некоррелированы, имеют средние значения $\mathbb{E}u_i = 0$ и дисперсии $\mathbb{E}u_i^2 = \sigma^2$. В этом параграфе мы предполагаем, что тренд является полиномом степени q

$$(1) \quad f(t) = \alpha_0 + \alpha_1 t + \dots + \alpha_q t^q.$$

Полином первой степени отражает равномерное во времени возрастание или убывание значений ряда. Полином второй степени может выражать тенденцию возрастания и последующего убывания значений ряда или наоборот, и т. д. Обычно q будет мало по сравнению с T . В данной модели может быть использована обычная техника регрессионного анализа.

Как отмечалось в гл. 2, мы можем перейти от независимых переменных, в данном случае $1, t, \dots, t^q$, к ортогональным независимым переменным $\varphi_{0T}(t), \dots, \varphi_{qT}(t)$. Поскольку полиномы будут использоваться очень часто, имеет смысл ортогонализировать их заранее; тогда все последующие вычисления упрощаются. Степени переменной t — натуральные числа. Почти во всех случаях в регрессионной формуле вместе с некоторой степенью переменной t присутствуют и все более низкие ее степени.

Пусть ортогональный полином степени k записан в виде

$$(2) \quad \varphi_{kT}(t) = t^k + C_{k-1}(k, T)t^{k-1} + \dots + C_0(k, T), \\ k = 1, \dots, T-1,$$

и $\varphi_{0T}(t) = 1$. Поскольку $\varphi_{kT}(t)$ по предположению ортогонален $\varphi_{0T}(t), \dots, \varphi_{k-1,T}(t)$, или, что эквивалентно, ортогонален $1, t, \dots, t^{k-1}$, мы должны иметь

$$(3) \quad \sum_{i=1}^T \varphi_{kT}(t) t^i = 0, \quad i = 0, 1, \dots, k-1,$$

т. е.

$$(4) \quad C_0(k, T) \sum_{t=1}^T t^i + C_1(k, T) \sum_{t=1}^T t^{i+1} + \dots \\ \dots + C_{k-1}(k, T) \sum_{t=1}^T t^{i+k-1} = - \sum_{t=1}^T t^{i+k}, \quad i = 0, 1, \dots, k-1.$$

Эти уравнения представляют собой частный случай уравнений (16) § 2.3 (с очевидными изменениями в обозначениях). При $k \leq T-1$ коэффициенты $C_0(k, T), \dots, C_{k-1}(k, T)$ определяются однозначным образом из системы (4). Например,

$$\begin{aligned} \varphi_{0T} &= 1, \\ (5) \quad \varphi_{1T} &= t - \frac{1}{2}(T+1), \\ \varphi_{2T} &= t^2 - (T+1)t + (T+1)(T+2)/6. \end{aligned}$$

Мы определили ортогональные полиномы таким образом, что их главные члены есть просто высшие степени переменной t . Без потери свойства ортогональности

$$(6) \quad \sum_{t=1}^T \varphi_{iT}(t) \varphi_{kT}(t) = 0, \quad i \neq k, \quad i, k = 0, 1, \dots, T-1,$$

каждый $\varphi_{kT}(t)$ можно умножать на константу. Множество $\varphi_{0T}(t), \dots, \varphi_{T-1,T}(t)$ в целом задает невырожденное линейное преобразование переменных $1, t, \dots, t^{T-1}$ (максимального числа линейно независимых полиномов для $t = 1, \dots, T$).

С использованием ортогональных полиномов тренд (1) записывается следующим образом:

$$(7) \quad f(t) = \gamma_0 + \gamma_1 \varphi_{1T}(t) + \dots + \gamma_q \varphi_{qT}(t).$$

При этом

$$(8) \quad \alpha_k = \gamma_k + C_k(k+1, T) \gamma_{k+1} + \dots + C_k(q, T) \gamma_q, \\ k = 0, 1, \dots, q.$$

Оценками коэффициентов в представлении (7) по наблюдениям $y_1, \dots, y_T$ являются величины

$$(9) \quad c_k = \frac{\sum_{t=1}^T y_t \varphi_{kT}(t)}{\sum_{t=1}^T \varphi_{kT}^2(t)}, \quad k = 0, 1, \dots, q.$$

Несмещенной оценкой дисперсии служит

$$(10) \quad s^2 = \frac{\sum_{t=1}^T [y_t - c_0 - c_1 \varphi_{1T}(t) - \dots - c_q \varphi_{qT}(t)]^2}{T - q - 1} =$$

$$= \frac{\sum_{t=1}^T y_t^2 - c_0^2 T - c_1^2 \sum_{t=1}^T \varphi_{1T}^2(t) - \dots - c_q^2 \sum_{t=1}^T \varphi_{qT}^2(t)}{T - q - 1}.$$

Вычисления по формулам (9) и (10) обычно выполняются на клавишной вычислительной машине с помощью таблиц ортогональных полиномов. У Фишера и Изэйтса (1963) последние даны (табл. 23) для степеней до пяти включительно и для $T = 3, \dots, 52$, а у Андерсона и Хаузмана (1942) до значения $T = 104$. Для данного значения T (обозначаемого у Фишера и Изэйтса через n) табулированы величины

$$(11) \quad \xi'_{kt} = \lambda_k \varphi_{kT}(t),$$

где λ_k — наименьшее рациональное число, для которого величины $\lambda_k \varphi_{kT}(t)$ — целые числа. (Следует заметить, что коэффициенты $C_0(k, T), \dots, C_{k-1}(k, T)$ в выражении (4) — целые, а следовательно, и рациональные числа.) Для примера рассмотрим $\varphi_{1T}(t) = t - (T+1)/2$. Если T нечетное, то $\varphi_{1T}(t)$ целое и $\lambda_1 = 1$; если же T четное, то $\varphi_{1T}(t)$ равно целому плюс $1/2$ и $\lambda_1 = 2$. В конце каждой колонки таблиц даны значения

$$(12) \quad \sum_{t=1}^T (\xi'_{kt})^2 = \lambda_k^2 \sum_{t=1}^T \varphi_{kT}^2(t)$$

и λ_k . Цель использования полиномов с целочисленными коэффициентами состоит в том, чтобы не вносить ошибок округления при вычислениях в (9), вплоть до выполнения деления. Вычисления с табулированными величинами дают значения c_k/λ_k , т. е. оценки отношений γ_k/λ_k (являющихся коэффициентами при ξ'_{kt}). Числовые примеры приведены в конце этого параграфа.

В случае применения быстродействующих вычислительных машин пользоваться указанными таблицами нет необходимости. Как отмечалось в гл. 2, подбираемая регрессия будет одной и той же и для степеней переменной t и для ортогональных полиномов той же степени, причем одинаковыми будут и коэффициенты при высшей степени. Всякий метод накопления ведущих элементов дает и коэффициенты при ортогональных полиномах и коэффициенты при степенях переменной t . Коэффициенты ортогональных полиномов (нормированные делением на корни квадратные из сумм квадратов ортогональных полиномов) легко интерпретировать как показатели вклада, вносимого в регрессию соответствующей степени, без учета вклада, вносимого в регрессию членами более низкого порядка.

3.2.2. Определение степени полиномиального тренда

Из теории наименьших квадратов известно, что c_k является несмещенной оценкой для γ_k , а s^2 — несмещенной оценкой для σ^2 . Дисперсия оценки c_k равна

$$(13) \quad \mathcal{E}(c_k - \gamma_k)^2 = \frac{\sigma^2}{\sum_{t=1}^T \Phi_{kT}^2(t)},$$

а сами величины c_k некоррелированы. Будем предполагать, что y_t нормально распределены. При этом $c_0, c_1, \dots, c_q$ независимы и нормально распределены, а отношение $(T - q - 1) s^2 / \sigma^2$ имеет распределение χ^2 с $T - q - 1$ степенями свободы и не зависит от $c_0, c_1, \dots, c_q$, так что можно использовать обычные критерии и доверительные интервалы. Например, при проверке гипотезы $H: \gamma_q = 0$ против альтернативы $\gamma_q \neq 0$ с уровнем значимости ε нулевая гипотеза отвергается, если

$$(14) \quad \frac{|c_q| \sqrt{a_{qq}}}{s} > t_{T-q-1}(\varepsilon),$$

где

$$(15) \quad a_{qq} = \sum_{t=1}^T \Phi_{qT}^2(t),$$

а $t_{T-q-1}(\varepsilon)$ — двусторонняя 100 ε -процентная точка ¹⁾ t -распределения с $T - q - 1$ степенями свободы. Нулевая гипотеза здесь состоит в том, что полином, о котором известно, что его степень не больше q , на самом деле имеет степень меньшую, чем q . Указанный критерий для проверки гипотезы $H: \gamma_q = 0$ в предположении нормальности является наилучшим при уровне значимости ε в следующем смысле. Он является: (а) равномерно наиболее мощным несмещенным критерием; (б) равномерно наиболее мощным инвариантным критерием (инвариантным относительно преобразования изменения масштаба $c_i \rightarrow k c_i$, $i = 0, 1, \dots, q$, $s^2 \rightarrow k^2 s^2$, и сдвигов ²⁾ $c_i \rightarrow c_i + a_i$, $i = 0, 1, \dots, q - 1$); (в) равномерно наиболее мощным критерием, имеющим мощность, зависящую только от γ_q^2 / σ^2 ; (г) равномерно наиболее мощным симметричным (относительно c_q) подобным критерием; (д) равномерно наиболее мощным подобным критерием с мощностью, зависящей только от $|\gamma_q|$ (т. е. не зависящей от знака γ_q). [См. Леман (1959, гл. 5, разд. 2) и упр. 11 из гл. 6)].

¹⁾ То есть в пределах от $-t_{T-q-1}(\varepsilon)$ до $t_{T-q-1}(\varepsilon)$ заключено 100(1- ε) процентов распределения. — *Прим. перев.*

²⁾ Указанное изменение масштаба равносильно преобразованию $y_t \rightarrow k y_t$, $t = 1, \dots, T$, а совокупность преобразования сдвига эквивалентна аддитивному наложению на последовательность $\{y_t\}$ полинома степени $q-1$.

Часто исследователь не знает заранее, какой степени полином следует подбирать. Поэтому он не просто интересуется тем, является ли степень полинома величиной, не превосходящей некоторого натурального числа q , а стремится выбрать подходящую степень полинома среди некоторого множества возможных степеней. При этом определенные преимущества имеет выбор полинома более низкой степени. График его более гладкий, проще допускаемое толкование, более экономична запись функции. Однако если среднее значение наблюдаемой переменной плохо описывается полиномом низкой степени, то исследователь может оказаться вынужденным использовать полином более высокой степени. (В терминах статистики недостатком выбора слишком низкой степени является наличие смещения при оценивании тренда, а недостатком выбора слишком высокой степени — большая вариабельность при оценивании тренда.) Будем предполагать, что существует некоторое число m (возможно, что $m = 0$), которое является минимально возможной степенью полинома, и что степень полинома не может превосходить некоторого q , допуская при этом, что исследователь имеет какую-то априорную информацию подобного рода.

Таким образом, исследователь сталкивается с задачей со многими решениями. Именно, он должен решить, равна ли степень полинома m , $m + 1$, ..., $q - 1$ или q . Мы формализуем эту задачу, считая, что необходимо решить, к какому из следующих непересекающихся множеств принадлежит параметрическая точка $(\gamma_{m+1}, \dots, \gamma_q)$:

$$\begin{aligned}
 &H_q: \gamma_q \neq 0, \\
 &H_{q-1}: \gamma_q = 0, \quad \gamma_{q-1} \neq 0, \\
 &\vdots \\
 &H_{m+1}: \gamma_q = \dots = \gamma_{m+2} = 0, \quad \gamma_{m+1} \neq 0, \\
 &H_m: \gamma_q = \dots = \gamma_{m+1} = 0.
 \end{aligned}
 \tag{16}$$

Принадлежность параметрической точки множеству H_i означает, что полином имеет степень i . Альтернативная постановка задачи состоит в том, что необходимо решить, какие из следующих нулевых гипотез верны (и верна ли хотя бы одна из них):

$$\begin{aligned}
 &H_q^*: \gamma_q = 0, \\
 &H_{q,q-1}^*: \gamma_q = \gamma_{q-1} = 0, \\
 &\vdots \\
 &H_{q,q-1, \dots, m+1}^*: \gamma_q = \gamma_{q-1} = \dots = \gamma_{m+1} = 0.
 \end{aligned}
 \tag{17}$$

Если некоторая гипотеза из (17) верна, то предшествующие ей гипотезы также должны быть верны. Если же она неверна, то неверны

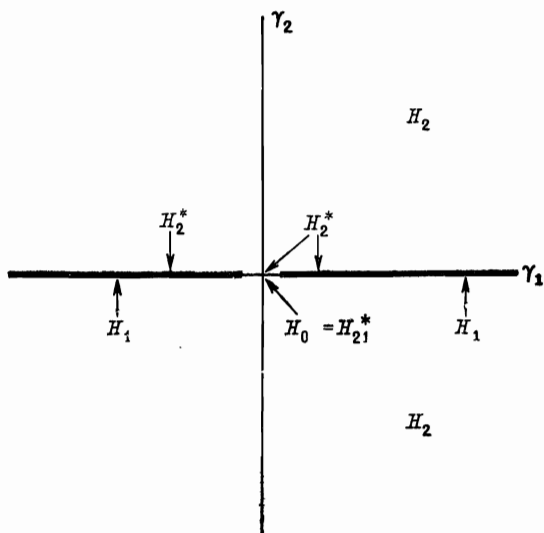


Рис. 3.1.

Семейства множеств в задаче с несколькими решениями.

и все последующие гипотезы. Таким образом, $H_{q,q-1}^*, \dots, m+1 \subset \dots \subset H_q^*$. Семейства множеств (16) и (17) связаны соотношениями

$$\begin{aligned}
 H_q^* &= H_{q-1} \cup \dots \cup H_m, \\
 H_{q,q-1}^* &= H_{q-2} \cup \dots \cup H_m, \\
 &\vdots \\
 H_{q,q-1, \dots, m+1}^* &= H_m.
 \end{aligned}
 \tag{18}$$

(См. рис. 3.1.)

Предположим, что исследователь намеревается ограничить вероятности ошибок, связанных с принятием решений о том, что коэффициенты отличны от нуля, в то время как на самом деле они равны нулю. Иными словами, ограничить вероятности ошибок, связанных с выбором степени полинома более высокой, чем это необходимо. При заданных значениях этих вероятностей он стремится минимизировать вероятности принятия решений о том, что коэффициенты равны нулю, в то время как на самом деле они отличны от нуля, т. е. минимизировать вероятности выбора степени полинома, меньшей чем требуемая. Мы предполагаем, что исследователь приписывает каждой из нулевых гипотез определенный уровень значимости:

$$\begin{aligned}
 p_q &= \text{Pr} \{ \text{отвергнуть } H_q^* \mid H_q^* \}, \\
 p_q + p_{q-1} &= \text{Pr} \{ \text{отвергнуть } H_{q,q-1}^* \mid H_{q,q-1}^* \}, \\
 &\vdots \\
 p_q + \dots + p_{m+1} &= \text{Pr} \{ \text{отвергнуть } H_{q,q-1}^*, \dots, H_{m+1}^* \mid H_{q,q-1}^*, \dots, H_{m+1}^* \} = \\
 &= \text{Pr} \{ \text{отвергнуть } H_m \mid H_m \},
 \end{aligned}
 \tag{19}$$

где $p_i \geq 0$ и $p_q + \dots + p_{m+1} \leq 1$. Поскольку каждая нулевая гипотеза включает в себя последующую (т. е. каждая последующая нулевая гипотеза является более сильной), то последовательность вероятностей отклонения в правых частях (19) берется монотонно неубывающей (т. е. вероятность отклонения более сильной нулевой гипотезы, когда она верна, не меньше, чем аналогичная вероятность для менее сильной гипотезы). Используя семейство попарно непересекающихся множеств H_i , запишем указанное разнесение гипотез по уровням значимости в виде

$$\begin{aligned}
 p_q &= \text{Pr} \{ \text{принять } H_q \mid H_{q-1} \cup \dots \cup H_m \}, \\
 p_{q-1} &= \text{Pr} \{ \text{принять } H_{q-1} \mid H_{q-2} \cup \dots \cup H_m \}, \\
 &\vdots \\
 p_{m+1} &= \text{Pr} \{ \text{принять } H_{m+1} \mid H_m \}.
 \end{aligned}
 \tag{20}$$

Пусть $p_m = 1 - p_q - \dots - p_{m+1} = \text{Pr} \{ \text{принять } H_m \mid H_m \}$. Посредством соотношений (20) исследователь приписывает определенные значения вероятностям принятия решений об использовании полинома степени i , когда на самом деле степень полинома меньше i (для каждого i).

Нерандомизированная статистическая процедура для этой задачи со многими решениями состоит в следующем. Имеется набор $q - m + 1$ областей, которые мы обозначим $R_m, R_{m+1}, \dots, R_q$, попарно непересекающихся и составляющих полную группу в пространстве значений $c_0, c_1, \dots, c_q$ и $S = (T - q - 1) s^2$ (или в исходном пространстве $y_1, \dots, y_T$). Если выборочная точка попадает в R_i , то принимается гипотеза H_i . Приписывание уровней значимости приводит к тому, что эти области становятся «подобными» в том смысле, что при $\gamma_i = \dots = \gamma_q = 0$ вероятности попадания выборочной точки в $R_i, \dots, R_q$ равны соответственно $p_i, \dots, p_q$ (независимо от $\gamma_0, \gamma_1, \dots, \gamma_{i-1}$ и σ^2). Другими словами, если степень полинома меньше i , то вероятность ошибки от приписывания ему степени i не зависит от того, какова его истинная (меньшая чем i) степень.

Во многих случаях исследователь стремится определить только, равен ли некоторый коэффициент нулю или отличен от нуля. При этом его не интересует знак коэффициента. Естественно в таком случае требовать, чтобы вероятности, связанные с указанной про-

цедурой, не меняли своих значений при изменении знака соответствующих коэффициентов, т. е. чтобы эти вероятности зависели от параметров только через $|\gamma_{m+1}|, \dots, |\gamma_q|$. (Из дальнейшего будет видно, что такое ограничение можно заменить присвоением ненулевым коэффициентам определенных знаков.)

Зафиксировав указанные ограничения, потребуем, чтобы области R_i были наилучшими, т. е. чтобы вероятности попадания в R_i при условии, что верна гипотеза H_i , $i = m+1, \dots, q$, были максимально возможными. Следует отметить, что мы хотим одновременно максимизировать вероятности попадания в $q - m$ различных областей (каждую для всех отличных от нуля значений соответствующего параметра). Далее будет показано, что при указанных выше условиях при подборе одной области с целью максимизации вероятности попадания в эту область не имеет значения то, как выбираются остальные области. Этот факт позволяет оптимизировать области $R_{m+1}, \dots, R_q$ одновременно. Следует только подчеркнуть, что возможность этого основана на предположении о нормальности $y_1, \dots, y_T$.

Как было отмечено выше, критерий с областью отклонения (14) является оптимальным для проверки гипотезы $H: \gamma_q = 0$, обозначаемой нами через H_q^* и являющейся дополнительной к H_q . Отсюда следует, что наилучшая область R_q определяется соотношением (14) с $\varepsilon = p_q$. Наилучшая процедура для проверки гипотезы $H: \gamma_i = 0$ с уровнем значимости ε_i , в предположении, что $\gamma_{i+1} = \dots = \gamma_q = 0$, состоит в ее отклонении при

$$(21) \quad \frac{c_{ii}^2 a_{ii}}{c_{i+1,i+1,i+1}^2 + \dots + c_{qq,qq}^2 + S} > \frac{t_{T-i-1}^2(\varepsilon_i)}{T-i-1},$$

где $S = (T - q - 1) s^2$ и

$$(22) \quad a_{ii} = \sum_{t=1}^T \varphi_{iT}^2(t).$$

Заметим далее, что если T_i — подобная область размера ε_i для проверки гипотезы $\gamma_i = 0$, т. е. если

$$(23) \quad \Pr\{T_i | \gamma_i = \gamma_{i+1} = \dots = \gamma_q = 0\} = \varepsilon_i$$

для всех $\gamma_0, \gamma_1, \dots, \gamma_{i-1}$ и $\sigma^2 (> 0)$, то T_i выделяет условную вероятность ε_i на почти всех комбинациях значений $c_0, c_1, \dots, c_{i-1}$ и $c_{ii}^2 a_{ii} + \dots + c_{qq,qq}^2 + S$ (достаточных статистик для $\gamma_0, \gamma_1, \dots, \gamma_{i-1}$ и σ^2 при $\gamma_i = \dots = \gamma_q = 0$). Иначе говоря,

$$(24) \quad \Pr\{T_i | c_0, c_1, \dots, c_{i-1}, c_{ii}^2 a_{ii} + \dots + c_{qq,qq}^2 + S; \gamma_i = \dots = \gamma_q = 0\} = \varepsilon_i$$

почти всюду (с вероятностью 1). Именно такой, как говорят, «неймановской» структурой подобных областей и определяются оптималь-

ные свойства t -критерия, отмеченные выше. [См. Леман (1959, гл. 4) или упр. 10.] Из требований (20) [или эквивалентных им требований (19)] вытекает, что такую структуру имеет R_i . Кроме того, такую структуру имеет $R_i \cup R_{i+1} \cup \dots \cup R_q$ — область отклонения гипотезы $H_{q,q-1,\dots,i}$, когда она верна. Мы используем этот факт для того, чтобы показать, что выбор $R_q, \dots, R_{i+1}$ (при допущении (20)) не влияет на выбор R_i в том смысле, что вероятность попадания в R_i (являющаяся функцией от $|\gamma_i|$) при $\gamma_{i+1} = \dots = \gamma_q = 0$ не зависит от того, какими были выбраны $R_q, \dots, R_{i+1}$. Отметим, что значение γ_i интересует нас только в том случае, когда $\gamma_{i+1} = \dots = \gamma_q = 0$. Если же какое-нибудь из $\gamma_{i+1}, \dots, \gamma_q$ отлично от нуля (т. е. если степень полинома больше i), то вопроса о том, отлично γ_i от нуля или нет, не возникает.

ЛЕММА 3.2.1. Пусть множество S_{i+1} в пространстве значений $c_0, \dots, c_q$ и S таково, что

$$(25) \quad \Pr \{S_{i+1} | \gamma_{i+1} = \dots = \gamma_q = 0\} = p_q + \dots + p_{i+1}$$

и пусть T_i — множество, определяемое значениями $c_0, \dots, c_i$ и $c_{i+1}^2 a_{i+1,i+1} + \dots + c_q^2 a_{qq} + S$, $S = (T - q - 1)s^2$. Тогда

$$(26) \quad \Pr \{\bar{S}_{i+1} \cap T_i | \gamma_{i+1} = \dots = \gamma_q = 0\} = \\ = (1 - p_q - \dots - p_{i+1}) \Pr \{T_i | \gamma_{i+1} = \dots = \gamma_q = 0\},$$

где $\bar{S}_{i+1}$ — множество, дополнительное к S_{i+1} .

Доказательство. Требование (25) означает, что S_{i+1} — подобная область (по отношению к $\gamma_0, \dots, \gamma_i$ и σ^2). В силу этого

$$(27) \quad \Pr \{S_{i+1} | c_0, \dots, c_i, c_{i+1}^2 a_{i+1,i+1} + \dots + c_q^2 a_{qq} + S; \\ \gamma_{i+1} = \dots = \gamma_q = 0\} = p_q + \dots + p_{i+1}$$

для почти всех $c_0, \dots, c_i$ и $c_{i+1}^2 a_{i+1,i+1} + \dots + c_q^2 a_{qq} + S$. Пусть $f_i(c_0, \dots, c_i, c_{i+1}^2 a_{i+1,i+1} + \dots + c_q^2 a_{qq} + S)$ — характеристическая функция множества T_i (т. е. $f_i = 1$, если аргумент принадлежит T_i , и $f_i = 0$ в противном случае). Тогда

$$(28) \quad \Pr \{\bar{S}_{i+1} \cap T_i\} = \\ = \mathbb{E}[\Pr \{\bar{S}_{i+1} | c_0, \dots, c_i, c_{i+1}^2 a_{i+1,i+1} + \dots + c_q^2 a_{qq} + S\} \times \\ \times f_i(c_0, \dots, c_i, c_{i+1}^2 a_{i+1,i+1} + \dots + c_q^2 a_{qq} + S)] = \\ = \mathbb{E}[(1 - p_q - \dots - p_{i+1}) f_i(c_0, \dots, c_i, c_{i+1}^2 a_{i+1,i+1} + \dots + c_q^2 a_{qq} + S)],$$

что и доказывает лемму, поскольку это соотношение совпадает с

(26). [Отметим, что это доказывает независимость t -критериев в (21).] ■

Смысл леммы состоит в том, что каким бы образом ни выбирались $R_q, \dots, R_{i+1}$ при условии (20), из которого следует (25) для $S_{i+1} = R_q \cup \dots \cup R_{i+1}$, вероятность попадания в область R_i , определенную как пересечение $\bar{S}_{i+1} \cap T_i$, будет зависеть только от T_i и не будет зависеть от S_{i+1} (когда верна гипотеза $H_{q,q-1, \dots, i+1}^*$).

Пусть теперь T_i^* — область, определяемая соотношением (21) при $\varepsilon_i = p_i/(1 - p_q - \dots - p_{i+1})$.

ЛЕММА 3.2.2. Пусть S_{i+1} удовлетворяет соотношению (25), а R_i — произвольное непересекающееся с S_{i+1} множество, для которого

$$(29) \quad \Pr \{R_i | \gamma_i = \gamma_{i+1} = \dots = \gamma_q = 0\} = p_i,$$

$$(30) \quad \Pr \{R_i | \gamma_i, \gamma_{i+1} = \dots = \gamma_q = 0\} = \\ = \Pr \{R_i | \gamma_i, \gamma_{i+1} = \dots = \gamma_q = 0\}.$$

Тогда

$$(31) \quad \Pr \{\bar{S}_{i+1} \cap T_i^* | \gamma_{i+1} = \dots = \gamma_q = 0\} \geq \\ \geq \Pr \{R_i | \gamma_{i+1} = \dots = \gamma_q = 0\}.$$

Доказательство. Предположим, что для некоторого значения параметра γ_i неравенство (31) нарушается. Покажем, что это противоречит приведенному выше утверждению о том, что $T_i^* = (S_{i+1} \cap T_i^*) \cup (\bar{S}_{i+1} \cap T_i^*)$ — равномерно наиболее мощный критерий с подобными областями для проверки гипотезы $\gamma_i = 0$ с мощностью, не зависящей от знака γ_i . Критические области T_i^* и $R_i \cup (S_{i+1} \cap T_i^*)$ имеют один и тот же размер. В то же время мощность второго критерия для указанного значения γ_i при сделанном предположении больше. Полученное противоречие доказывает лемму. ■

Из доказанных двух лемм вытекает, что какими бы ни были $R_q, \dots, R_{i+1}$, наилучший выбор множества R_i состоит в том, что это R_i должно являться частью множества T_i^* , не содержащей точек из $R_q, \dots, R_{i+1}$. При таком выборе R_i вероятность

$$(32) \quad \Pr \{R_i | \gamma_{i+1} = \dots = \gamma_q = 0\} = \\ = (1 - p_q - \dots - p_{i+1}) \Pr \{T_i^* | \gamma_{i+1} = \dots = \gamma_q = 0\}$$

не зависит от выбора $R_q, \dots, R_{i+1}$.

ТЕОРЕМА 3.2.1. Пусть $R_m, R_{m+1}, \dots, R_q$ — полная группа q — $m + 1$ непересекающихся областей в выборочном пространстве,

таких, что

$$(33) \Pr \{R_i | \gamma_i = \gamma_{i+1} = \dots = \gamma_q = 0\} = p_i, \quad i = m+1, \dots, q,$$

где $p_{m+1} + \dots + p_q \leq 1$ и

$$(34) \Pr \{R_i | \gamma_i, \gamma_{i+1} = \dots = \gamma_q = 0\} = \\ = \Pr \{R_i | -\gamma_i, \gamma_{i+1} = \dots = \gamma_q = 0\}, \quad i = m+1, \dots, q.$$

Тогда для каждого значения γ_i вероятность (34) принимает максимальное значение на множестве R_i , определяемом как пересечение множества (21) при $\varepsilon_i = p_i/(1 - p_q - \dots - p_{i+1})$ и дополнения к множеству $R_q \cup \dots \cup R_{i+1}$, $i = m+1, \dots, q$.

Оптимальной является, таким образом, следующая процедура:

$$(35) \quad \begin{aligned} R_q &= T_q^*, \\ R_{q-1} &= \bar{T}_q^* \cap T_{q-1}^*, \\ &\vdots \\ R_i &= \bar{T}_q^* \cap \bar{T}_{q-1}^* \cap \dots \cap \bar{T}_{i+1}^* \cap T_i^*, \\ &\vdots \\ R_{m+1} &= \bar{T}_q^* \cap \bar{T}_{q-1}^* \cap \dots \cap \bar{T}_{m+2}^* \cap T_{m+1}^*, \\ R_m &= \bar{T}_q^* \cap \bar{T}_{q-1}^* \cap \dots \cap \bar{T}_{m+2}^* \cap \bar{T}_{m+1}^*. \end{aligned}$$

Эта процедура по сути дела сводится к следующему. Поочередно проверяются гипотезы $\gamma_q = 0$, $\gamma_{q-1} = 0$, ... до тех пор, пока либо какая-то из гипотез не будет отвергнута, например гипотеза $\gamma_i = 0$, и будет решено, что верна H_i , либо будут приняты все гипотезы вплоть до H_m . Таким образом, указанная процедура является последовательной. Это обуславливается требованием независимости вероятности правильного решения о том, что степень полинома меньше заданного натурального числа, от истинной степени полинома.

У t -критерия имеется целый ряд оптимальных свойств, пять из которых мы отметили выше. Использование нескольких из них и приводит к утверждению сформулированной теоремы. Условие подобия (33) можно заменить условием несмещенности, а именно условием

$$(36) \Pr \{\text{отвергнуть } H_{q,q=1, \dots, i} | H_{q,q=1, \dots, i}^*\} \leq p_q + \dots + p_i, \\ i = m+1, \dots, q,$$

поскольку из несмещенности вытекает подобие соответствующих областей, и требованием независимости от знаков величин $\gamma_q, \dots, \gamma_{m+1}$ вероятностей попадания в эти области. [См. Леман (1959, гл. 4, разд. 1).]

Мы не останавливались еще на вопросе о том, как следует выбирать $p_{m+1}, \dots, p_q$. Если все ε_i фиксированы и равны, например, ε , то $p_q = \varepsilon$, $p_{q-1} = \varepsilon(1 - \varepsilon)$, $\dots$, $p_i = \varepsilon(1 - \varepsilon)^{q-i}$, $\dots$, $p_{m+1} = \varepsilon(1 - \varepsilon)^{q-m-1}$, $p_m = (1 - \varepsilon)^{q-m}$. Намерение не завышать степень полинома следует сбалансировать с чувствительностью процедуры к ненулевым коэффициентам. При разумном подходе следует устанавливать значение q довольно большим, но полагать ε_i для значений i , близких к q , весьма малыми. Если влияние i -й степени сравнительно велико, то при таком выборе будет иметься возможность это определить. Вероятность же принять решение о необходимости использования высоких степеней в том случае, когда этого не требуется, при указанном выборе мала. Например, пусть $\varepsilon_j = \varepsilon^*(q + 1 - j)/(q - m)$, $j = m + 1, \dots, q$. Тогда p_m приближенно равно $\exp\{-\varepsilon^*(q - m + 1)/2\} \sim 1 - \varepsilon^*(q - m + 1)/2$.

Описанная процедура не нова. Отличие здесь только в подходе. Другая процедура, предложенная для оценивания степени полиномиальной регрессии, также является последовательностью критериев значимости, но взятой в обратном порядке. Сначала с помощью t -критерия проверяется гипотеза $\gamma_{m+1} = 0$. Если она отклоняется, то проверяется гипотеза $\gamma_{m+2} = 0$ и т. д. В результате либо некоторая гипотеза принимается, либо все гипотезы вплоть до $\gamma_q = 0$ отклоняются. Недостатком этой процедуры является то, что если некоторое γ_i очень велико, то велика и вероятность слишком заниженного оценивания степени полинома. Например, если значение γ_2 велико, а значение γ_1 мало, то будет относительно велика вероятность того, что

$$(37) \quad \frac{c_{1a_{11}}^2}{c_{2a_{22}}^2 + \dots + c_{qa_{qq}}^2 + S} < k.$$

То есть с относительно большой вероятностью будет принята гипотеза $\gamma_1 = 0$ и решено, что степень полинома равна нулю. (Следует отметить, что эта процедура не удовлетворяет (19) или (20) при $\gamma_3 = \dots = \gamma_q = 0$.)

Практическое неудобство процедуры, которая в нашей формулировке является наилучшей, состоит в том, что она требует вычисления S и вследствие этого вычисления $c_0, c_1, \dots, c_q$ до значения q , выбранного заранее. В то же время во второй упомянутой последовательной процедуре вычисление $c_1, c_2, \dots$ производится последовательно ($c_{2a_{22}}^2 + \dots + c_{qa_{qq}}^2 + S = \sum_{t=1}^T y_t^2 - c_0^2 T - c_{1a_{11}}^2$ и т. д.) и

только до тех пор, пока гипотезы отвергаются. Однако отмеченное неудобство не очень существенно, поскольку коэффициенты регрессии при ортогональных полиномах относительно просто вычислять с помощью таблиц полиномов. К тому же обычно приходится использовать лишь небольшие значения q , поскольку в случаях, требую-

щих использования высоких степеней, применение ортогональных полиномов не приносит особой пользы. (На практике, если принимается H_q , следует дополнительно вычислять c_{q+1} для проверки правильности выбора q .) Как было указано ранее, при наличии быстродействующих вычислительных машин можно использовать и регрессию по степеням t . Заметим, что коэффициенты при ортогональных полиномах, используемые в процедуре, определяются в любом методе накопления ведущих элементов.

В § 2.3 было отмечено, что критерий проверки гипотезы о равенстве некоторого коэффициента регрессии нулю не изменится, если совершить линейное преобразование независимых переменных, оставляющее этот коэффициент без изменения в выражении регрессии с помощью новых переменных. Частным случаем такого преобразования является последовательная ортогонализация переменных, при которой интересующая нас переменная ортогонализируется в последнюю очередь. Таким образом, критерий для проверки гипотезы $\gamma_i = 0$ в предположении $\gamma_{i+1} = \dots = \gamma_q = 0$ в точности совпадает с критерием для проверки гипотезы $\alpha_i = 0$ в предположении, что $\alpha_{i+1} = \dots = \alpha_q = 0$ (здесь α_j — коэффициенты при степенях t в выражении (1) для полиномиального тренда). Приведенная выше процедура со многими решениями может поэтому выполняться на основе регрессионного анализа полиномиального тренда. (В действительности эта процедура применима к любому упорядоченному множеству независимых переменных.) Как было указано в § 2.3, критерий для проверки гипотезы о равенстве нулю некоторого регрессионного коэффициента требует только *прямого* решения нормальных уравнений. При выполнении процедуры со многими решениями прямое решение достаточно получить лишь для полной совокупности переменных, поскольку прямое решение для каждого подмножества, получаемого вычеркиванием последней переменной, является частью полного прямого решения.

Прогнозирование значения y_t для $t > T$ может быть произведено, как указано в § 2.5, т. е. в качестве прогноза можно брать $\sum_k c_k \Phi_{kt}(t)$. Однако обычно прогнозирование с использованием подобранного полиномиального тренда рискованно. Дело в том, что полином является лишь приближением к реальному тренду и поэтому нельзя быть уверенным, что это приближение останется удовлетворительным вне множества значений t , на котором производился подбор полинома.

Пример 3.1. В табл. 3.1 приведены данные¹⁾ о количестве мяса, потреблявшегося в США на душу населения, за период с 1919 по

¹⁾ Тинтнер (1952, стр. 195—198). Анализ этого ряда описан там же на стр. 177. Термин «мясо» объединяет здесь собственно мясо, птицу и рыбу.

Таблица 3.1

ЕЖЕГОДНОЕ ПОТРЕБЛЕНИЕ МЯСНЫХ ПРОДУКТОВ (ВКЛЮЧАЯ ПТИЦУ И РЫБУ)
В США С 1919 ПО 1941 Г.

t	y_t	ξ'_{1t}	ξ'_{2t}	ξ'_{3t}	ξ'_{4t}	ξ'_{5t}	Y_t
1919	171.5	-11	77	-77	1463	-209	165.805
1920	167.0	-10	56	-35	133	76	169.456
1921	164.5	-9	37	-3	-627	171	171.927
1922	169.3	-8	20	20	-950	152	173.350
1923	179.4	-7	5	35	-955	77	173.859
1924	179.2	-6	-8	43	-747	-12	173.585
1925	172.6	-5	-19	45	-417	-87	172.662
1926	170.5	-4	-28	42	-42	-132	171.223
1927	168.6	-3	-35	35	315	-141	169.399
1928	164.7	-2	-40	25	605	-116	167.325
1929	163.0	-1	-43	13	793	-65	165.132
1930	162.1	0	-44	0	858	0	162.954
1931	160.2	1	-43	-13	793	65	160.923
1932	161.2	2	-40	-25	605	116	159.172
1933	165.8	3	-35	-35	315	141	157.833
1934	163.5	4	-28	-42	-42	132	157.040
1935	146.7	5	-19	-45	-417	87	156.925
1936	160.2	6	-8	-43	-747	12	157.620
1937	156.8	7	5	-35	-955	-77	159.260
1938	156.8	8	20	-20	-950	-152	161.975
1939	165.4	9	37	3	-627	-171	165.900
1940	174.7	10	56	35	133	-76	171.167
1941	178.7	11	77	77	1463	209	177.908
λ_k		1	1	1/6	7/12	1/60	

Здесь t обозначает год, y_t обозначает потребление (в фунтах) на душу населения в t -й год, а Y_t является соответствующим значением подобранного полинома третьей степени.

Таблица 3.2

ВЫЧИСЛЕНИЯ С ИСПОЛЬЗОВАНИЕМ ОРТОГОНАЛЬНЫХ ПОЛИНОМОВ

Степень полинома k	Сумма квадратов $\sum_{t=1}^T \xi_{kt}^2$	Сумма произведений $\sum_{t=1}^T y_t \xi_{kt}$	Коэффициент $\frac{c_k}{\lambda_k}$	Квадрат * $c_k^2 a_{kk}$
0	23		166.1913	
1	1 012	-383.6	-0.379051	145.404
2	35 420	2 606.0	0.073574	191.735

Продолжение табл. 3.2

Степень полинома k	Сумма квадратов $\sum_{t=1}^T \xi_{kt}^2$	Сумма произведений $\sum_{t=1}^T y_t \xi_{kt}$	Коэффициент $\frac{c_k}{\lambda_k}$	Квадрат * $c_k^2 a_{kk}$
3	32 890	4 366.0	0.1327455	579.567
4	13 123 110	17 252.4	0.00131466	22.681
5	340 860	17.8	0.00005222	0.000929

k	Остаточная сумма квадратов $c_{k+1}^2 a_{k+1, k+1} + \dots + S$	Средняя остаточная сумма квадратов $\frac{c_{k+1}^2 a_{k+1, k+1} + \dots + S}{T - k - 1}$	Отношение t^2
0	1369.538	62.2517	
1	1224.413	58.3054	2.494
2	1032.400	51.6200	3.714
3	452.833	23.8333	24.318
4	430.152	23.8973	0.949
5	430.151	25.3030	0.000 037

$$\sum_{t=1}^T y_t = 3822.4, \quad \sum_{t=1}^T y_t^2 = 636 619.18, \quad S = 430.150796.$$

* Вычисляется как отношение $\left(\sum_{t=1}^T y_t \xi_{kt} \right)^2 / \sum_{t=1}^T \xi_{kt}^2$.

1941 г. ($T = 23$). В ней же даны значения ортогональных полиномов до пятой степени включительно [из Р. Л. Андерсона и Хаузмана (1942)]. Приведены также множители λ_k . Вычисления коэффициентов регрессии и t^2 -отношений сведены в табл. 3.2.

В этом случае подходящую степень полинома определить трудно. Возьмем $q = 5$. Если только p_5 берется не слишком близким к 1, то в соответствии с результатами вычислений мы принимаем гипотезу $\gamma_5 = 0$. Если, далее, $p_4/(1 - p_5)$ ненамного больше, чем $1/3$, то мы принимаем гипотезу $\gamma_4 = 0$. Если, наконец, $p_3/(1 - p_5 - p_4)$ не слишком мало, то мы принимаем гипотезу $\gamma_3 \neq 0$. Таким образом, мы выбираем степень, равную трем. В последней колонке табл. 3.1 приведены значения соответствующего кубического приближения.

Наблюдаемый ряд и подобранный тренд графически изображены на рис. 3.2. Следует заметить, что выравнивающая кривая дает хорошее согласие, причем большинство точек лежит близко от нее. Мы интерпретировали бы данную кривую как ожидаемое, или нормальное потребление мяса, если бы не возникающие год от года

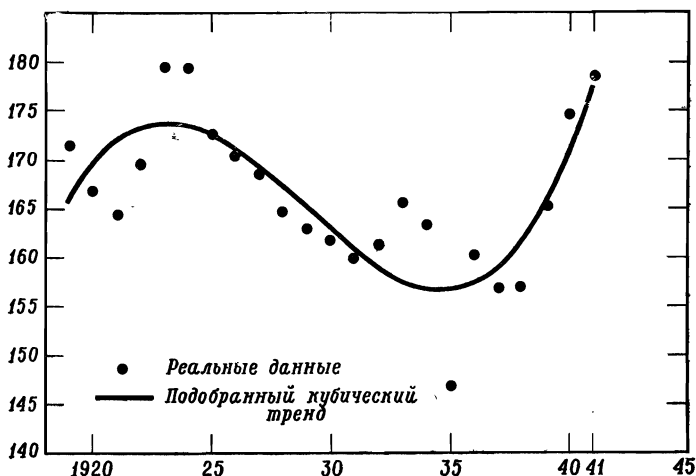


Рис. 3.2.

Ежегодное потребление мясных продуктов на душу населения в США с 1919 по 1941 г. и подобранный кубический тренд.

нерегулярности. Должно быть ясно, что изображенный полином третьей степени не дает хорошего прогноза. Во всяком случае, его можно использовать в этом качестве лишь для достаточно близкого будущего. Для значений же, расположенных значительно правее приведенных, этот полином растет, причем возрастает и его производная. Однако, даже без учета влияния войн, представляется не очень правдоподобным, чтобы потребление мяса на душу населения возрастало неограниченно и при все увеличивающемся темпе роста.

3.3. СГЛАЖИВАНИЕ

3.3.1. Сглаживающие процедуры

Иногда тренд является гладкой функцией времени, флуктуирует на любом коротком интервале времени незначительно и все же его невозможно представить простой функцией времени на всем рассматриваемом интервале. Тем не менее простая мысль о том, что тренд гладкий, а случайные ошибки, как правило, нерегулярны, приводит к естественной статистической технике сглаживания. Сглаживание временного ряда означает представление тренда в данной точке посредством взвешенного среднего значений, наблюдаемых в окрестности этой точки. При этом считается, что наблюдаемые значения являются суммой тренда и случайной ошибки. Грубо говоря, взвешенное среднее тренда совпадает со значением самого тренда в данной точке, а взвешенное среднее случайных составляю-

щих имеет тенденцию становиться весьма малой величиной (предполагается, что случайные составляющие независимы и имеют нулевые математические ожидания). Поэтому взвешенное среднее наблюдаемых значений будет оценивать тренд. Оно определяется для каждого момента времени, за исключением нескольких первых и нескольких последних точек. Тем самым довольно нерегулярный график наблюдений заменяется гладким графиком скользящего среднего.

Опишем более четко, что мы понимаем под сглаживанием. Под гладкой функцией мы подразумеваем функцию, которая может быть адекватно представлена полиномом достаточно низкой степени на некотором не слишком малом интервале времени. Это понятие до некоторой степени оправдывается теоремой Тейлора. (Нас фактически интересует здесь приближение только для нескольких целых значений аргумента.) Полиномы, аппроксимирующие функцию, не обязаны быть одинаковыми на различных интервалах. Полином, подобранный на одном интервале, может не иметь ничего общего с указанной гладкой функцией в любой другой части отрезка наблюдения. Приведем в качестве примера тренд, составленный из выпуклой и вогнутой парабол:

$$(1) \quad f(t) = \begin{cases} a[1 - (t - k)^2/k^2], & t = 0, 1, \dots, 2k, \\ -a[1 - (t - 3k)^2/k^2], & t = 2k, \dots, 4k, \end{cases}$$

для некоторого целого k . В точке $t = 2k$ значения и самих этих парабол и их первых производных совпадают. Так определенная функция будет гладкой. Она является полиномом первой или второй степени для любых трех последовательных точек и представима параболой на достаточно длинных интервалах времени. Однако параболы, представляющая функцию на первой половине всего интервала, весьма отличается от второй параболы. В действительности указанная функция довольно близка на интервале от 0 до $4k$ к полиному третьей степени. (См. упр. 20.) Смысл приведенных рассуждений состоит в том, что предположение о гладкости тренда является локальным свойством, в то время как предположение о полиномиальном характере тренда связано со всем интервалом $t = 1, \dots, T$. Соответственно предположение гладкости позволяет использовать для оценки тренда в данной точке только наблюдения вблизи этой точки, в то время как предположение о полиномиальном характере тренда приводит к тому, что для оценки полинома, представляющего тренд на всем интервале, используются все наблюдения.

Пусть имеются наблюдения $y_1, \dots, y_T$. Оценим тренд в точке t посредством величины

$$(2) \quad y_t^* = \sum_{s=-m}^m c_s y_{t+s}, \quad t = m+1, \dots, T-m,$$

являющейся взвешенным средним наблюдаемых значений y_t в ин-

тервале значений временного параметра t , отстоящих от t не более чем на m единиц. Веса c_s предполагаются нормированными, так что

$$(3) \quad \sum_{s=-m}^m c_s = 1.$$

(Пределы суммирования можно брать без потери общности симметричными относительно нуля, поскольку некоторые c_s могут быть равны нулю.) Полученная таким образом последовательность $y_{m+1}^*, \dots, y_{t-m}^*$ называется *скользящим средним* исходной последовательности $\{y_t\}$. Если $y_t = f(t) + u_t$, то

$$(4) \quad y_t^* = \sum_{s=-m}^m c_s f(t+s) + u_t^*,$$

где

$$(5) \quad u_t^* = \sum_{s=-m}^m c_s u_{t+s}.$$

Как и прежде, предполагается, что $\mathbb{E}u_t = 0$, $\mathbb{E}u_t^2 = \sigma^2$ и $\mathbb{E}u_t u_s = 0$, $t \neq s$. В силу этого u^* имеет дисперсию $\sigma^2 \sum_{s=-m}^m c_s^2$. Веса c_s выбираются так, чтобы дисперсия величины u_t^* была значительно меньше, чем дисперсия величины u_t . Если значения $f(t+s)$, $s = \pm 1, \dots, \pm m$, близки к $f(t)$, то и $\mathbb{E}y_t^*$ будет близко к $f(t)$. Поэтому скользящее среднее, или сглаженный ряд $\{y_t^*\}$, имеет примерно ту же последовательность математических ожиданий, что и $\{y_t\}$, но зато меньшую дисперсию (одинаковую для всех членов $\{y_t^*\}$). Однако последовательные члены сглаженного ряда в общем случае являются коррелированными. Именно

$$(6) \quad \mathbb{E}u_t^* u_{t+h}^* = \sum_{s=-m}^m \sum_{r=-m}^m c_s c_r \mathbb{E}u_{t+s} u_{t+h+r} = \\ = \begin{cases} \sigma^2 \sum_{s=-m+h}^m c_s c_{s-h}, & h = 0, 1, \dots, 2m, \\ 0, & h = 2m+1, \dots \end{cases}$$

Одним из частных случаев скользящего среднего является арифметическое среднее, для которого $c_s = 1/(2m+1)$. В этом случае сглаженное значение y_t^* выражается формулой

$$(7) \quad y_t^* = \frac{1}{2m+1} \sum_{s=-m}^m y_{t+s} = \\ = \frac{1}{2m+1} \sum_{s=-m}^m f(t+s) + \frac{1}{2m+1} \sum_{s=-m}^m u_{t+s}.$$

Слагаемое, представляющее усредненную ошибку, имеет дисперсию $\sigma^2/(2m+1)$ как среднее $2m+1$ некоррелированных переменных. Ковариации (последовательных) значений u_i^* суть

$$(8) \quad \mathcal{E} u_i^* u_{i+h}^* = \begin{cases} \frac{2m+1-h}{(2m+1)^2} \sigma^2, & h = 0, 1, \dots, 2m, \\ 0, & h = 2m+1, \dots \end{cases}$$

Общая основа для большинства формул сглаживания состоит фактически в подборе сглаживающего полинома по $(2m+1)$ последовательным наблюдениям и в использовании этого полинома для оценки тренда в средней точке. Поскольку оценки коэффициентов полинома зависят от наблюдаемых значений линейно, то линейной является и оценка тренда. Вследствие этого она имеет вид (2). Предположим теперь, что тренд $f(t+s)$ в точках $t+s = t-m, \dots, t+m$ можно приблизить полиномом

$$(9) \quad f_t(s) = \alpha_0 + \alpha_1 s + \dots + \alpha_q s^q, \quad s = -m, \dots, m.$$

(Коэффициенты α_s зависят от t , но мы не будем отмечать этого в обозначениях.) В частности, $f(t)$ приближенно равно $f_t(0) = \alpha_0$. Коэффициенты этого полинома можно оценить на основании наблюдений $y_{t-m}, \dots, y_{t+m}$, используя метод наименьших квадратов. Нормальные уравнения для оценок $a_0, a_1, \dots, a_q$ имеют вид

$$(10) \quad a_0 \sum_{s=-m}^m s^j + a_1 \sum_{s=-m}^m s^{j+1} + \dots + a_q \sum_{s=-m}^m s^{j+q} = \\ = \sum_{s=-m}^m s^j y_{t+s}, \quad j = 0, 1, \dots, q.$$

В силу симметрии для любых нечетных степеней k сумма величин s^k по всем значениям s от $-m$ до $+m$ равна нулю. Поэтому в соотношении (10) для четных j равны нулю коэффициенты при $a_1, a_3, \dots$, а для нечетных j равны нулю коэффициенты при $a_0, a_2, \dots$. Поскольку оценка для $f(t)$ должна быть и оценкой для $f_t(0) = \alpha_0$, то достаточно решить (10) относительно a_0 . С этой целью воспользуемся уравнениями с четными j , т. е.

$$(11) \quad a_0 \sum_{s=-m}^m s^{2i} + a_2 \sum_{s=-m}^m s^{2i+2} + \dots + a_{2[q/2]} \sum_{s=-m}^m s^{2i+2[q/2]} = \\ = \sum_{s=-m}^m s^{2i} y_{t+s}, \quad i = 0, 1, \dots, [q/2],$$

где

$$(12) \quad [q/2] = \begin{cases} q/2, & \text{когда } q \text{ четное,} \\ (q-1)/2, & \text{когда } q \text{ нечетное.} \end{cases}$$

Заметим, что уравнения, которые необходимо решить для определения a_0 при нечетном значении q , в точности совпадают с уравнениями, которые надо решить для определения a_0 при меньшем на единицу (четном) значении q . (Нам необходимо исследовать только степени 0, 2, 4 и т. д.) Пусть $[q/2] = k$. Тогда (11) можно записать в виде

$$(13) \quad (2m+1)a_0 + 2 \sum_{s=1}^m s^2 a_{2s} + \dots + 2 \sum_{s=1}^m s^{2k} a_{2k} = \sum_{s=-m}^m y_{t+s},$$

$$2 \sum_{s=1}^m s^{2i} a_0 + 2 \sum_{s=1}^m s^{2i+2} a_2 + \dots + 2 \sum_{s=1}^m s^{2i+2k} a_{2k} = \sum_{s=1}^m s^{2i} (y_{t-s} + y_{t+s}),$$

$$i = 1, \dots, k.$$

Коэффициенты в левых частях этих нормальных уравнений зависят только от m . Коэффициенты при y_{t-s} и y_{t+s} в правых частях совпадают, $s = 1, \dots, m$. Решением (13) относительно a_0 является

$$(14) \quad a_0 = \sum_{s=-m}^m c_s y_{t+s}$$

с $c_{-s} = c_s$. Заметим, что коэффициенты c_s зависят от m и k и являются полиномами от s . Из (14) вытекает, что $y_t^* = a_0$. Используя (14) и замечая, что если $y_{t+s} = \alpha$, $s = -m, \dots, m$, то $\hat{f}_t(s) \equiv \alpha$ — наилучший выравнивающий полином, можно получить соотношение

$$\sum_{s=-m}^m c_s = 1.$$

Положим для примера $m = 2$ и $q = 2$, т. е. используем для оценки $\mathcal{E}y_t$ значения $y_{t-2}, y_{t-1}, y_t, y_{t+1}, y_{t+2}$, предполагая, что некоторая парабола хорошо приближает значения $f(t-2), f(t-1), f(t), f(t+1), f(t+2)$. Нормальные уравнения для a_0 и a_2 имеют в этом случае вид

$$(15) \quad 5a_0 + 10a_2 = \sum_{s=-2}^2 y_{t+s},$$

$$10a_0 + 34a_2 = \sum_{s=1}^2 s^2 (y_{t-s} + y_{t+s}).$$

Отсюда

$$(16) \quad a_0 = \frac{34}{70} \sum_{s=-2}^2 y_{t+s} - \frac{10}{70} \sum_{s=1}^2 s^2 (y_{t-s} + y_{t+s}) =$$

$$= (-3y_{t-2} + 12y_{t-1} + 17y_t + 12y_{t+1} - 3y_{t+2})/35.$$

Если $k = 0$ ($q = 0$ или 1), то в (13) имеется всего одно уравнение относительно неизвестного a_0 . Его решение $a_0 = \sum_{s=-m}^m y_{t+s}/(2m+1)$,

т. е. $c_s = 1/(2m + 1)$, $s = -m, \dots, m$. Другими словами, скользящее среднее с равными весами есть частный случай полиномиального сглаживания. Он имеет место, когда степень полинома равна 0 или 1.

Таблица 3.3

КОЭФФИЦИЕНТЫ СГЛАЖИВАЮЩИХ ФОРМУЛ ДЛЯ $k = 1, 2$

m	$c_{-m} = c_s$	$c_{-4} = c_4$	$c_{-3} = c_3$	$c_{-2} = c_2$	$c_{-1} = c_1$	c_0
	$k = 1$					
2				$-\frac{3}{35}$	$\frac{12}{35}$	$\frac{17}{35}$
3			$-\frac{2}{21}$	$\frac{3}{21}$	$\frac{6}{21}$	$\frac{7}{21}$
4		$-\frac{21}{231}$	$\frac{14}{231}$	$\frac{39}{231}$	$\frac{54}{231}$	$\frac{59}{231}$
5	$-\frac{36}{429}$	$\frac{9}{429}$	$\frac{44}{429}$	$\frac{69}{429}$	$\frac{84}{429}$	$\frac{89}{429}$
	$k = 2$					
3			$\frac{5}{231}$	$-\frac{30}{231}$	$\frac{75}{231}$	$\frac{131}{231}$
4		$\frac{15}{429}$	$-\frac{55}{429}$	$\frac{30}{429}$	$\frac{135}{429}$	$\frac{179}{429}$
5	$\frac{18}{429}$	$-\frac{45}{429}$	$-\frac{10}{429}$	$\frac{60}{429}$	$\frac{120}{429}$	$\frac{143}{429}$

Если $k = 1$ ($q = 2$ или 3), то нормальные уравнения принимают вид (см. упр. 7)

$$\begin{aligned}
 (2m+1)a_0 + \frac{(2m+1)m(m+1)}{3}a_2 &= \sum_{s=-m}^m y_{t+s}, \\
 (17) \quad \frac{(2m+1)m(m+1)}{3}a_0 + \frac{(2m+1)m(m+1)(3m^2+3m-1)}{15}a_2 &= \\
 &= \sum_{s=1}^m s^2(y_{t-s} + y_{t+s}).
 \end{aligned}$$

Их решением является

$$(18) \quad a_0 = \frac{\sum_{s=-m}^m [3(3m^2+3m-1) - 15s^2] y_{t+s}}{(2m-1)(2m+1)(2m+3)}.$$

В табл. 3.3 приведены значения коэффициентов c_s для $k = 1$ и $k = 2$ и нескольких значений m . Случай $k = 1$ ($q = 2$ или 3) исследован выше, а случай $k = 2$ ($q = 4$ или 5) вынесен в упр. 22. Кендалл и Стьюарт (1966, разд. 46.5) приводят коэффициенты для других случаев и обсуждают связанные с ними формулы сглаживания.

Необходимо отметить, что при $m < k$ коэффициент a_0 остается неопределенным. При $m = k$ производится подбор полинома степени $2k + 1$ по $2m + 1 = 2k + 1$ точкам. Такой подбор выполняется точно и поэтому $a_0 = y_t$. Если $m > k$, то скользящее среднее нетривиально, т. е. в него входит несколько значений y_{t+s} . В простейшем из таких случаев ¹⁾, $m = k + 1$,

$$(19) \quad 1 - c_0 = C \binom{2k+2}{k+1},$$

$$(20) \quad c_{-s} = c_s = (-1)^{s+1} C \binom{2k+2}{k+1+s}, \quad s = 1, \dots, m,$$

где

$$(21) \quad C = \frac{\binom{2k+2}{k+1}}{\binom{4k+4}{2k+2}} = \frac{\frac{(2k+2)!}{(k+1)!^2}}{\frac{(4k+4)!}{(2k+2)!^2}} = \frac{(2k+2)!^3}{(k+1)!^2 (4k+4)!}.$$

Мы убедимся в этом в § 3.4, после того как получим ряд результатов, связанных с последовательностями разностей.

3.3.2. Свойства процедуры сглаживания

Перейдем теперь к рассмотрению некоторых свойств указанных методов сглаживания. Одна из основных задач сглаживания состоит в том, чтобы уменьшить случайную ошибку, т. е. сделать дисперсию сглаженной последовательности малой по сравнению с дисперсией исходной последовательности.

ТЕОРЕМА 3.3.1. Дисперсия величины $y_t^* = a_0$ равна

$$(22) \quad \sigma^2 b^{00} = \sigma^2 c_0,$$

где b^{00} — верхний левый угловой элемент матрицы $\mathbf{B}^{-1}$, обратной к матрице $\mathbf{B}$ коэффициентов уравнений (13), элементы которой суть

$$(23) \quad b_{ij} = 2 \sum_{s=1}^m s^{2(i+j)}, \quad i+j > 0, \quad i, j = 0, 1, \dots, k.$$

¹⁾ $\binom{h}{g} = \frac{h!}{g! (h-g)!}, \quad g = 0, 1, \dots, h; \quad 0! = 1.$

ДОКАЗАТЕЛЬСТВО. Если обозначить правые части уравнений (13) через Y_i , $i = 0, 1, \dots, k$, то эти уравнения можно записать в виде

$$(24) \quad \sum_{j=0}^k b_{ij} a_{2j} = Y_i, \quad i = 0, 1, \dots, k,$$

а их решения относительно a_0 в виде

$$(25) \quad a_0 = \sum_{j=0}^k b^{0j} Y_j = b^{00} \sum_{s=-m}^m y_{t+s} + \sum_{j=1}^k b^{0j} Y_j.$$

Здесь $(b^{ij}) = \mathbf{B}^{-1}$. Поскольку Y_j не содержат значения y_t для $j > 0$, то b^{00} является коэффициентом при y_t . Общая теория метода наименьших квадратов утверждает, однако, что дисперсия a_0 равна $\sigma^2 b^{00}$.

Таблица 3.4

ДИСПЕРСИИ СГЛАЖЕННЫХ ЗНАЧЕНИЙ ($\sigma^2 = 1$)

$2m+1$	$k=0$ $q=0,1$	$k=1$ $q=2,3$	$k=2$ $q=4,5$	$k=3$ $q=6,7$
3	$\frac{1}{3} = 0.333$			
5	$\frac{1}{5} = 0.200$	$\frac{17}{35} = 0.486$		
7	$\frac{1}{7} = 0.143$	$\frac{1}{3} = 0.333$	$\frac{131}{231} = 0.567$	
9	$\frac{1}{9} = 0.111$	$\frac{59}{231} = 0.255$	$\frac{179}{429} = 0.417$	$\frac{797}{1287} = 0.619$
11	$\frac{1}{11} = 0.091$	$\frac{89}{429} = 0.207$	$\frac{1}{3} = 0.333$	

Если $k = 0$, то дисперсия величины y_t^* есть $\sigma^2/(2m+1)$. Для $k = 1$ она равна

$$(26) \quad \frac{3(3m^2 + 3m - 1)}{(2m-1)(2m+1)(2m+3)} \sigma^2,$$

а для $m = k+1$

$$(27) \quad \sigma^2 \left[1 - C \left(\frac{2k+2}{k+1} \right) \right] = \sigma^2 \left[1 - \frac{(2k+2)!^4}{(4k+4)! (k+1)!^4} \right].$$

В табл. 3.4 приведены дисперсии некоторых сглаженных значений.

Для фиксированного k дисперсия уменьшается с ростом используемого числа точек. При фиксированном числе точек (т. е. при фиксированном m) дисперсия увеличивается с возрастанием k . Факти-

чески для фиксированного значения $m - k$, которое вдвое меньше разности числа точек и числа неявно подбираемых констант, дисперсия увеличивается с ростом k .

Отметим также, что разность наблюдаемого и сглаженного значений $(y_t - y_t^*)$ не коррелирована с y_t^* , поскольку оценки коэффициентов регрессии не коррелированы с остатками. (См упр. 7 гл. 2.) Поэтому

$$(28) \quad \text{Var}(y_t - y_t^*) = \sigma^2 - \text{Var} y_t^* = \sigma^2(1 - c_0).$$

Как было указано выше, последовательные сглаженные величины являются коррелированными. Например, корреляции y_t^* с y_{t-1}^* , y_{t-2}^* , y_{t-3}^* и y_{t-4}^* для случая $k = 1$ и $m = 2$ равны соответственно

$$(29) \quad \frac{336}{595} = 0.565; \quad \frac{42}{595} = 0.071; \quad -\frac{72}{595} = -0.121; \quad \frac{9}{595} = 0.015.$$

Мы изучим еще это явление в гл. 7, после того как разовьем более мощный математический аппарат.

Если $y_t = f(t) + u_t$ и используется сглаживающая формула с коэффициентами c_s , то систематическая ошибка сглаженной величины имеет вид

$$(30) \quad \mathcal{E}(y_t - y_t^*) = f(t) - \sum_{s=-m}^m c_s f(t+s).$$

Если сглаживающая формула основывается на полиноме степени q и тренд является полиномом той же (или меньшей) степени, то систематическая ошибка будет равна 0. В противном случае она отлична от нуля. Предположим, что $k = 0$ ($q = 0$ или 1) и коэффициенты те же. Тогда систематическая ошибка выражается соотношением

$$(31) \quad \mathcal{E}(y_t - y_t^*) = f(t) - \frac{1}{2m+1} \sum_{s=-m}^m f(t+s),$$

т. е. разностью между $f(t)$ и средним арифметическим соседних значений. Предположим, что $f(t+s)$, $s = -m, \dots, m$, записывается с помощью ортогональных полиномов степени, не превышающей $2m+1$ (ортогональных на множестве $-m, \dots, m$), в виде

$$(32) \quad f(t+s) = \gamma_0 + \gamma_1 \varphi_{1,2m+1}(s) + \dots + \gamma_{2m+1} \varphi_{2m+1,2m+1}(s),$$

$$s = -m, \dots, m.$$

Тогда использование сглаживающей формулы, основывающейся на полиноме степени $2k$ или $2k+1$, приводит к систематической

ошибке

$$(33) \quad \mathcal{E}(y_t - y_t^*) = \\ = \gamma_{2k+2} \Phi_{2k+2, 2m+1}^*(0) + \gamma_{2k+4} \Phi_{2k+4, 2m+1}^*(0) + \dots + \gamma_{2m} \Phi_{2m, 2m+1}^*(0),$$

поскольку выравнивающий полином состоит из элементов соотношения (32) степени до $2k$ или $2k + 1$ включительно и $\Phi_{i, 2m+1}^*(0) = 0$ для нечетных i . (См. упр. 30.)

В гл. 4 мы будем изучать случай, когда среднее значение является функцией $f(t) = \cos(\lambda t - \theta)$, т. е. косинусом с периодом $2\pi/\lambda$. Если при этом коэффициенты $c_s = 1/(2m + 1)$, то ожидаемое значение сглаженной переменной запишется в виде

$$(34) \quad \mathcal{E}y_t^* = \frac{1}{2m+1} \sum_{s=-m}^m \cos[\lambda(t+s) - \theta] = \\ = \frac{\sin \lambda(2m+1)/2}{(2m+1) \sin \lambda/2} \cos(\lambda t - \theta) = \\ = \frac{\sin \lambda(2m+1)/2}{(2m+1) \sin \lambda/2} f(t), \quad 0 < \lambda < 2\pi.$$

(См. упр. 31.) Таким образом, операция сглаживания здесь просто уменьшает амплитуду функции $f(t)$. Если λ мало (т. е. период велик), то и это уменьшение мало (упр. 32). При фиксированном λ большим значениям m (удовлетворяющим неравенству $(2m + 1)\lambda < 2\pi$) соответствует меньший коэффициент пропорциональности (упр. 33). Если $2m + 1 = 2\pi/\lambda$ (длина скользящего усреднения равна периоду), то сглаженное значение равно нулю.

Основная цель сглаживания состоит в оценивании тренда, или ожидаемого значения y_t с наименьшей ошибкой. Ошибка складывается здесь из смещения (30) и случайной составляющей $u_t^* =$

$= \sum_{s=-m}^m c_s u_{t+s}$. Первую составляющую можно измерить ее квадратом, а вторую — ее дисперсией $\sigma^2 b^{00} = \sigma^2 c_0$. При фиксированном k смещение с увеличением m в большинстве случаев возрастает, а дисперсия убывает. В то же время при фиксированном m смещение с увеличением k убывает, а дисперсия возрастает. Статистик, которому приходится использовать сглаживающую формулу, должен выбрать значения k и m . Он мог бы использовать в качестве меры ошибки среднеквадратичную ошибку, которая является суммой указанной дисперсии и среднего квадрата смещения. Если бы дисперсия σ^2 случайных ошибок u_t была известна и если бы были известны средние квадраты смещений для каждой комбинации k и m , то статистик смог бы выбрать комбинацию k и m , минимизирующую эту меру ошибки. Однако здесь трудно дать какую-либо рекомендацию, поскольку дисперсия и среднеквадратичное смещение ведут

себя в отношении k и m противоположным образом. Если σ^2 мало, то можно удовлетвориться относительно малым m . Чем более гладкой является $f(t)$, тем меньшим может быть выбрано k при фиксированном m (или тем большим выбрано m при фиксированном k). В действительности, конечно, эти характеристики не известны, а должны быть оценены по имеющимся данным. Поэтому выбор k и m является статистической задачей со многими решениями, которую трудно даже сформулировать, не говоря уже о ее строгом статистическом решении. Поэтому практик должен действовать здесь исходя из своей интуиции и накопленного опыта.

Другой подход состоит в том, чтобы выяснить, каково наименьшее k , такое, что средний квадрат смещения близок или равен нулю, когда m фиксировано или когда m — заданная функция переменной k , например $m = k + 1$. Мы рассмотрим этот подход в следующем параграфе.

Преимуществом сглаживания для оценки тренда является его гибкость в том смысле, что предположения, при которых его можно использовать, не очень обременительны. Однако, поскольку этот метод не основывается на явной вероятностной модели, свойства его не вполне определены и статистические выводы ограничены. Например, тренд здесь не определяется малым числом параметров, для которых можно было бы указать доверительные области. Невозможно проверять гипотезы относительно тренда. Нельзя непосредственно связать функцию, оценивающую тренд, с теорией или с моделью образования наблюдаемого ряда. При сглаживании оценивающая тренд функция годится скорее для целей описания, нежели для целей анализа ряда и его интерпретации. Из-за того, что этот метод не базируется на явной вероятностной модели, он не может быть изложен полностью и строго в терминах математической статистики (по крайней мере кратко).

Имеется и серьезная практическая трудность в применении сглаживания. Для того чтобы получить величину y_t^* , оценивающую тренд в точке t , необходимо использовать значения $y_{t-m}, \dots, y_{t+m}$. Поскольку эта процедура основывается на наблюдениях $y_1, \dots, y_T$, то первым сглаженным значением будет y_{m+1}^* , а последним y_{T-m}^* . Тем самым, мы не имеем оценок тренда в начале периода наблюдений и в его конце. Для оценки тренда в этих точках необходимо привлекать какие-то другие соображения.

Сглаживание само по себе, конечно, не дает средних прогнозирования. Экстраполяция оцененного тренда весьма ненадежна отчасти из-за того, что тренд не оценивается для последних m моментов времени.

Мы основывали сглаживание на нечетном числе членов с симметричными весами. Если используется четное число членов с симметричными весами, то сглаженное значение интерпретируется как

оценка тренда в точке, лежащей посередине между двумя средними точками. Это может оказаться неудобным.

Скольльзящее усреднение с равными весами ($k = 0$) можно легко осуществить на клавишной вычислительной машине, поскольку

сумма $\sum_{s=-m}^m y_{t+s}$ изменяется при каждом t путем вычитания одного члена и добавления другого. Эти суммы запоминаются и затем каждая делится на $2m + 1$ (или умножается на $1/(2m + 1)$). Представляет значительный интерес аппроксимация процедуры сглаживания с неравными весами последовательностью процедур сглаживания, использующих равные веса. Конечно, при наличии быстродействующей вычислительной машины нет никакой нужды упрощать коэффициенты.

Сглаживание с использованием скользящего среднего имеет длинную историю, причем к нему пришли первоначально с точки зрения, отличной от статистической. [См. Уиттекер и Робинсон (1926).] Иногда бывает необходимо интерполировать между точками, в которых наблюдения производились. В интерполяционных формулах используются последовательные разности. Для того чтобы эти разности вели себя гладким образом, перед интерполированием можно применить формулы сглаживания. С этой точки зрения две сглаживающие процедуры эквивалентны с точностью до некоторого порядка, если разности этого порядка согласуются для каждой пары сглаженных рядов, полученных в результате применения этих двух процедур. (См. § 3.4.) Говорят, что процедура является точной до разностей некоторого порядка, если она не нарушает разностей этого порядка для полиномов. Одной из часто используемых процедур, точных до разностей третьего порядка, является 15-точечная формула Спенсера. Эта процедура выполняется таким образом. Сначала вычисляются величины

$$(35) \quad y_t^* = (-3y_{t-2} + 3y_{t-1} + 4y_t + 3y_{t+1} - 3y_{t+2})/4,$$

затем усредняются (с равными весами) 5 последовательных y_t^* , далее — 4 последовательных члена полученного ряда и, наконец, усредняются 4 последовательных члена последнего ряда. Другой процедурой, сохраняющей разности третьего порядка, является 21-точечная формула Спенсера, соответствующая вычислению величин

$$(36) \quad y_t^* = (-y_{t-3} + y_{t-1} + 2y_t + y_{t+1} - y_{t+3})/2$$

и поочередному усреднению 7, затем 5 и 5 членов получающихся рядов. Обе эти процедуры сравнительно легко реализуются.

3.3.3. Сезонные изменения

Если на временной ряд накладываются регулярные периодические изменения, то полезным оказывается иной подход. Во многих ежемесячных временных рядах проявляется, например, сезонный фактор. Изучаемую функцию времени можно записать при этом в виде

$$(37) \quad \mathcal{E}y_t = f(t) = g(t) + h(t),$$

где функция $g(t)$ имеет период n (12 для ежемесячных данных, 4 для ежеквартальных и т. д.), т. е.

$$(38) \quad g(t+n) = g(t), \quad t = 1, \dots, T-n.$$

Мы можем нормировать $g(t)$ таким образом, чтобы

$$(39) \quad \sum_{t=1}^n g(t) = 0.$$

Из периодичности $g(t)$ следует при этом, что и

$$(40) \quad \sum_{t=1}^n g(t+s) = 0, \quad s = 0, \dots, T-n.$$

Обычно T выбирается таким, чтобы оно делилось нацело на n , $T = hn$. (Например, при ежемесячных данных за h лет $T = 12h$.) При произвольном выборе $f(t)$ данное выше описание не позволяет однозначно определить $g(t)$ до тех пор, пока не будут наложены определенные условия на функцию $h(t)$. Обычно предполагается, что она или является медленно меняющимся трендом, или циклическая.

Скользящее среднее с n членами и равными коэффициентами будет устранять сезонное колебание $g(t)$ в том смысле, что

$$(41) \quad \mathcal{E} \frac{1}{n} \sum_{s=1}^n y_{t+s} = \frac{1}{n} \sum_{s=1}^n h(t+s).$$

Если n четное, $n = 2m$, что обычно имеет место в экономических данных, то используем

$$(42) \quad y_t^* = \frac{1}{2m} \left[\sum_{s=-m}^{m-1} y_{t+s} + \frac{1}{2} y_{t-m} + \frac{1}{2} y_{t+m} \right],$$

$$t = m+1, \dots, T-m.$$

Тогда

$$(43) \quad \mathcal{E}y_t^* = \frac{1}{2m} \left[\sum_{s=-m}^{m-1} f(t+s) + \frac{1}{2} f(t-m) + \frac{1}{2} f(t+m) \right] =$$

$$= \frac{1}{2m} \left[\sum_{s=-m}^{m-1} h(t+s) + \frac{1}{2} h(t-m) + \frac{1}{2} h(t+m) \right].$$

Если долговременный тренд $h(t)$ меняется медленно, то (43) близко к $h(t)$. В частном случае, когда функция $h(t)$ линейна, (43) в точности равно $h(t)$.

При $T = hn$ функцию $g(t)$ можно определить однозначно соотношением

$$(44) \quad g(t) = \frac{1}{h} \sum_{j=0}^{h-1} f(t + nj) - \frac{1}{T} \sum_{s=1}^T f(s), \quad t = 1, \dots, n.$$

Сезонный эффект декабря равен, например, разности между средним по всем декаблям и средним по всем данным. Оценкой функции $g(t)$ служит в этом случае

$$(45) \quad \frac{1}{h} \sum_{j=0}^{h-1} y_{t+nj} - \frac{1}{T} \sum_{s=1}^T y_s, \quad t = 1, \dots, n.$$

Эта оценка является несмещенной и имеет дисперсию

$$(46) \quad \left[h \left(\frac{1}{h} - \frac{1}{T} \right)^2 + (T - h) \left(-\frac{1}{T} \right)^2 \right] \sigma^2 = \frac{n-1}{T} \sigma^2.$$

Альтернативный метод оценивания состоит в использовании отклонений от сглаженных значений. Если $n = 2m$, то соответствующие оценки для $g(t)$ следующие:

$$(47) \quad \begin{aligned} & \frac{1}{h-1} \sum_{j=1}^{h-1} (y_{t+nj} - y_{t+nj}^*), \quad t = 1, \dots, m, \\ & \frac{1}{h-1} \sum_{j=0}^{h-2} (y_{t+nj} - y_{t+nj}^*), \quad t = m+1, \dots, 2m. \end{aligned}$$

Ввиду того что в сглаживаемых рядах отсутствуют m первых и m последних членов, средние (47) должны основываться на $h-1$ отклонениях. В более подробной записи они имеют вид

$$(48) \quad \begin{aligned} & \frac{1}{h-1} \left[\sum_{j=1}^{h-1} y_{t+nj} - \frac{1}{2m} \left(\sum_{s=t+m+1}^{T-m-1+t} y_s + \frac{1}{2} y_{t+m} + \frac{1}{2} y_{T-m+t} \right) \right], \\ & \qquad \qquad \qquad t = 1, \dots, m, \\ & \frac{1}{h-1} \left[\sum_{j=0}^{h-2} y_{t+nj} - \frac{1}{2m} \left(\sum_{s=t-m+1}^{T-3m-1+t} y_s + \frac{1}{2} y_{t-m} + \frac{1}{2} y_{T-3m+t} \right) \right], \\ & \qquad \qquad \qquad t = m+1, \dots, 2m. \end{aligned}$$

Дисперсия оценки (48) равна

$$(49) \quad \left[\frac{n-1}{T-n} - \frac{1}{2(T-n)^2} \right] \sigma^2.$$

Может показаться, что оценка (48) несколько предпочтительнее, так как применение скользящего среднего является более гиб-

ким подходом, чем точное описание тренда. Однако разница между (48) и (45) заключается только в использовании крайних членов, именно в отбрасывании $(2m - 1)$ крайних членов и использовании половинных значений двух других крайних членов. Легко проверить, что (49) больше чем (46). [См. Дурбин (1963).]

При рассмотрении многих экономических временных рядов статистики полагают, что составляющие ряд воздействия перемножаются. Иными словами, они предполагают, что ряд имеет вид

$$(50) \quad Y(t) = G(t) H(t) U(t),$$

где $H(t)$ — тренд, $G(t)$ — сезонный фактор, $U(t)$ — случайный фактор, причем все сомножители положительны. Применяемое в этом случае так называемое «отношение к скользящему среднему» (ratio-to-moving average) заключается в последовательном образовании ряда скользящих средних (42) и использовании отношений $Y(t + nj)/Y^*(t + nj)$ вместо разностей из (47). Для получения оценки сезонного фактора, соответствующей суммам (47), производится усреднение образованных отношений при каждом значении t (или находится медиана этих отношений). Если произведение $2m$ средних (или медиан) указанных отношений не равно 1, как это обычно и бывает (в то время как сумма оценок (48) должна была бы всегда равняться нулю, если бы не влияние крайних членов), то каждое отношение умножается на одно и то же число, так чтобы результирующее произведение равнялось 1. Альтернатива к этой процедуре состоит в использовании логарифмов. При этом

$$(51) \quad \begin{aligned} y_t &= \log Y(t), & g(t) &= \log G(t), \\ h_t &= \log H(t), & u_t &= \log U(t). \end{aligned}$$

Если полученные ряды и их компоненты удовлетворяют сделанным ранее предположениям, то для оценивания функции $g(t) = \log G(t)$ можно использовать (45) или (48). Этот метод представляется более предпочтительным. Дело в том, что метод отношений к скользящему среднему не имеет строгого математического обоснования. Более того, чувствуется, что он не является безукоризненным в этом отношении, поскольку использует аддитивные методы при мультипликативных факторах (это проявляется в необходимости видоизменять оценки, принудительным образом домножая их произведение до 1).

Мы определили сезонное изменение $g(t)$ в (44) чисто формальным образом. Экономист обычно представляет себе сезонное изменение как проявление определенного поведения, обусловленного тем или иным временем года, а тренд ¹⁾ — как долговременную тенденцию,

¹⁾ Мы понимаем здесь под «трендом» всякое систематическое изменение ожидаемого значения, не связанное с сезонными изменениями. Экономисты часто понимают «тренд» как долговременное изменение, противопоставляя его «циклическим изменениям», коротким по времени.

обусловленную более устойчивыми воздействиями. При этом $f(t)$ является суммой сезонного воздействия $g(t)$ и долговременного тренда $h(t)$, причем (по крайней мере субъективно) $g(t)$ и $h(t)$ определяются независимо. При таком подходе функция $g(t)$ может и не быть строго периодической, медленно меняя со временем свой вид. Можно предположить, например, что

$$(52) \quad g(t + nj) = g^*(t) g^{**}(j),$$

где $g^{**}(j)$ — медленно меняющаяся функция. Мы, однако, не будем развивать далее это направление. Метод анализа сезонных изменений, основанный на подобной модели, предложен Вальдом (1936). Изменяющиеся сезонные колебания рассматривались затем Хеннаном (1964), Боксом и Дженкинсом (1970).

Статистики иногда рассматривают экономический временной ряд как ряд, складывающийся из долговременного тренда, циклического изменения, сезонного воздействия и нерегулярной составляющей. Тренд является долговременной тенденцией изменения, обусловленной ростом популяции, технологическими изменениями и другими достаточно долговременными воздействиями. Циклическое же изменение связано с колебаниями, известными под названием цикла деловой активности. С этой точки зрения циклическое изменение не обязательно периодически, как это имело место у сезонного изменения в (38). Тем не менее иногда предполагают, что скользящее среднее устраняет влияние циклического изменения при оценивании тренда. Вопрос об эффективности этой процедуры мы рассмотрим позднее.

Сглаживающие формулы могут быть основаны не только на полиномах, но и на других выравнивающих функциях. Кроме того, эти функции не обязательно должны подбираться с использованием равных весов. Значения в точках $s = 0, \pm 1, \dots, \pm m$ могут быть обработаны так, как если бы они имели возрастающую дисперсию.

Задачу получения сглаженных значений на концах рядов можно решать различными способами. Значения в начале ряда могут не играть особой роли. Значения же непосредственно в конце ряда обычно наиболее существенны. Если сглаживающая формула основывается на подборе полинома степени q по $2m + 1$ точкам, то полином, подобранный по значениям $y_{T-2m}, y_{T-2m+1}, \dots, y_T$, можно использовать для получения значений y_t^* сглаженного ряда и при $t = T - m + 1, T - m + 2, \dots, T$, и при $t = T - m$. Эти сглаженные значения y_t^* также являются линейными комбинациями значений $y_{T-2m}, \dots, y_T$. [Коуден (1962) приводит соответствующие коэффициенты для $q = 1, \dots, 5$ и $m = 1, \dots, 12$.] Другой подход состоит в применении сглаживающих процедур, основывающихся на использовании в конце ряда меньших значений m и k . При таком подходе $y_T^* = y_T$ (поскольку и m , и k должны быть равны в этом

случае нулю), что является не вполне удовлетворительным. Оба эти метода дают сглаженные значения, обладающие в конце ряда большей вариабельностью, чем в его середине.

3.4. МЕТОД ПЕРЕМЕННЫХ РАЗНОСТЕЙ

3.4.1. Введение

Метод переменных разностей был предложен для оценивания дисперсии случайной составляющей при гладком характере тренда. При дальнейшем развитии метод стал применяться и для решения вопроса о степени гладкости тренда. Обе эти статистические задачи рассматриваются применительно к модели $y_t = f(t) + u_t$, где u_t некоррелированы, имеют нулевые средние и дисперсии σ^2 , а функция $f(t)$ является гладкой в том смысле, что ее можно хорошо приблизить в последовательных интервалах времени полиномом небольшой степени.

Другое применение переменных разностей состоит в проверке отсутствия корреляции. Такие задачи рассматриваются в гл. 6 для иной вероятностной модели.

В настоящем параграфе будет развита некоторая общая теория, рассмотрены оценивание величины σ^2 и критерии для проверки гипотез о гладкости, а также будет выяснена связь указанного метода со сглаживанием. Некоторые из этих результатов используются в дальнейшем при изучении сериальной корреляции.

Впервые разности были использованы, по-видимому, Кейв-Браун-Кейвом (1904) и Хукером (1905) для изучения корреляции между двумя рядами. Метод был развит затем Стьюдентом (1914). О. Андерсон (1929) и Тинтнер (1940) интенсивно изучали его применительно к оцениванию дисперсии.

3.4.2. Вычисление последовательных разностей

Метод переменных разностей основан на вычислении последовательных разностей элементов временного ряда. Установим систему обозначений для таких вычислений и отметим некоторые их свойства. Пусть $\mathcal{P}$ — оператор, определяемый соотношением

$$(1) \quad \mathcal{P}u_t = u_{t+1}, \quad t = \dots, -1, 0, 1, \dots$$

По любой заданной последовательности $\dots, u_{-1}, u_0, u_1, \dots$ этот оператор строит новую последовательность, индексы в которой сдвинуты на единицу. Мы будем писать также $\mathcal{P}\{u_t\} = \{u_{t+1}\}$.

Оператор такого рода есть функция, аргументами и значениями которой являются последовательности. Оператор $\mathcal{O}$ называется

линейным, если для любой последовательности $\{u_i\}$ и для любого действительного числа c

$$(2) \quad \mathcal{O}\{cu_i\} = c\mathcal{O}\{u_i\}$$

и, кроме того, для любой пары последовательностей $\{u_i\}$ и $\{v_i\}$

$$(3) \quad \mathcal{O}\{u_i + v_i\} = \mathcal{O}\{u_i\} + \mathcal{O}\{v_i\}.$$

Очевидно, что $\mathcal{P}$ является линейным оператором. Мы будем использовать запись (1), понимая ее и как результат применения оператора $\mathcal{P}$ к целой последовательности, и как результат операции над отдельным элементом u_i последовательности.

Определим $\mathcal{P}^0 = 1$ как тождественный оператор ($\mathcal{P}^0 u_i = u_i$), а оператор $\mathcal{P}^n$ определим рекуррентно как $\mathcal{P}(\mathcal{P}^{n-1})$, т. е.

$$(4) \quad \mathcal{P}^n u_i = \mathcal{P}(\mathcal{P}^{n-1} u_i), \quad n = 2, 3, \dots$$

По индукции можно вывести, что

$$(5) \quad \mathcal{P}^n u_i = u_{i+n}.$$

Оператор $c\mathcal{P}$ задается соотношением

$$(6) \quad (c\mathcal{P}) u_i = c(\mathcal{P} u_i).$$

Определим далее сумму операторов соотношением

$$(7) \quad (c_1 \mathcal{P}^{n_1} + \dots + c_k \mathcal{P}^{n_k}) u_i = c_1 \mathcal{P}^{n_1} u_i + \dots + c_k \mathcal{P}^{n_k} u_i,$$

Как следует из (7),

$$(8) \quad (c_1 \mathcal{P}^{n_1} + \dots + c_k \mathcal{P}^{n_k}) u_i = c_1 u_{i+n_1} + \dots + c_k u_{i+n_k}.$$

Таким образом, определен полином от $\mathcal{P}$ (с действительными коэффициентами). Операции над этими полиномами (умножение и т. д.) согласуются с операциями над полиномами от абстрактной переменной. (Сумма бесконечного числа операторов есть предел левой части (8), если правая часть сходится в некотором смысле.)

В частности, представляет интерес полином $\Delta = \mathcal{P} - 1$ (разность первого порядка), действующий следующим образом:

$$(9) \quad \Delta u_i = (\mathcal{P} - 1) u_i = \mathcal{P} u_i - u_i = u_{i+1} - u_i.$$

Разность второго порядка есть

$$(10) \quad \Delta^2 u_i = (\mathcal{P} - 1)^2 u_i = (\mathcal{P}^2 - 2\mathcal{P} + 1) u_i,$$

или, что эквивалентно,

$$(11) \quad \Delta(\Delta u_i) = \Delta u_{i+1} - \Delta u_i = u_{i+2} - 2u_{i+1} + u_i.$$

Разности более высоких порядков определяются как

$$(12) \quad \Delta^r u_t = (\mathcal{P} - 1)^r u_t = \sum_{j=0}^r (-1)^{r-j} \binom{r}{j} \mathcal{P}^j u_t = \\ = \sum_{j=0}^r (-1)^{r-j} \binom{r}{j} u_{t+j}.$$

Для удобства вычислений мы можем предпочесть определение $\Delta^r u_t = \Delta (\Delta^{r-1} u_t)$. Для заданной последовательности $\{u_t\}$ в этом случае следовало бы поочередно вычислять: Δu_t , $\Delta^2 u_t = \Delta (\Delta u_t)$ и т. д.

Важным свойством операции вычисления разностей является результат ее воздействия на последовательность, образованную полиномами от t . Мы имеем

$$(13) \quad \Delta(\alpha_0 + \alpha_1 t + \dots + \alpha_k t^k) = \\ = \alpha_0 + \alpha_1(t+1) + \dots + \alpha_k(t+1)^k - \\ - [\alpha_0 + \alpha_1 t + \dots + \alpha_k t^k] = \\ = \alpha_1 + \alpha_2[(t+1)^2 - t^2] + \alpha_3[(t+1)^3 - t^3] + \dots \\ \dots + \alpha_k[(t+1)^k - t^k] = \\ = k\alpha_k t^{k-1} + \left[\binom{k}{2} \alpha_k + \binom{k-1}{1} \alpha_{k-1} \right] t^{k-2} + \dots \\ \dots + \left[\binom{k}{k-1} \alpha_k + \binom{k-1}{k-2} \alpha_{k-1} + \dots + \binom{2}{1} \alpha_2 \right] t + \\ + [\alpha_k + \alpha_{k-1} + \dots + \alpha_2 + \alpha_1].$$

Существенным моментом является здесь то, что применение разностного оператора к полиному понижает степень последнего на единицу. Из этого факта следует, что если $f(t)$ — полином степени k , то

$$(14) \quad \Delta^r f(t) = 0, \quad r = k+1, k+2, \dots$$

Ранее мы рассмотрели тренды, которые или являются полиномами, или хорошо аппроксимируются полиномами на интервалах. С помощью операции вычисления разностей такие тренды редуцируются к нулю или к функции, близкой к тождественному нулю. (Более полно об исчислении разностей см. Жордан (1939) или Миллер (1960)).

3.4.3. Вычисление последовательных разностей наблюдаемых рядов

Рассмотрим наблюдаемый временной ряд $\{y_t\}$, о котором предполагается, что он складывается из тренда $f(t)$ и случайной ошибки u_t . Поскольку оператор Δ линейный, то

$$(15) \quad \Delta y_t = \Delta [f(t) + u_t] = \Delta f(t) + \Delta u_t,$$

$$(16) \quad \Delta^r y_t = \Delta^r f(t) + \Delta^r u_t.$$

Если $f(t)$ — полином от t степени меньшей, чем r , то $\Delta^r f(t) = 0$ и $\mathbb{E} \Delta^r y_t = 0$. Во всяком случае,

$$(17) \quad \mathbb{E} \Delta^r y_t = \Delta^r f(t).$$

Обратимся теперь к дисперсии величины $\Delta^r y_t$. Она равна дисперсии случайной величины

$$(18) \quad \Delta^r u_t = (\mathcal{P} - 1)^r u_t = \left[\mathcal{P} - \binom{r}{1} \mathcal{P}^{-1} + \dots + (-1)^r \right] u_t = \\ = u_{t+r} - \binom{r}{1} u_{t+r-1} + \dots + (-1)^r u_t.$$

Для дальнейшего будет полезна

ЛЕММА 3.4.1. Пусть $P(x)$ и $Q(x)$ — полиномы степени p и q соответственно. Тогда $\mathbb{E} P(\mathcal{P}) u_t Q(\mathcal{P}) u_t$ равно умноженному на σ^2 коэффициенту при x^q в выражении $x^q P(x) Q(x^{-1})$ или, что равносильно, коэффициенту при x^p в $x^p P(x^{-1}) Q(x)$.

ДОКАЗАТЕЛЬСТВО. Пусть

$$(19) \quad P(x) = \sum_{i=0}^p a_i x^i, \quad Q(x) = \sum_{j=0}^q b_j x^j.$$

Тогда

$$(20) \quad \mathbb{E} P(\mathcal{P}) u_t Q(\mathcal{P}) u_t = \sum_{i=0}^p \sum_{j=0}^q a_i b_j \mathbb{E} u_{t+i} u_{t+j} = \sigma^2 \sum_{i=0}^{\min(p,q)} a_i b_i,$$

поскольку $\mathbb{E} u_t u_s = 0$, $t \neq s$. Однако это и есть умноженный на σ^2 коэффициент при x^q в выражении

$$(21) \quad P(x) x^q Q(x^{-1}) = \sum_{i=0}^p a_i x^i \sum_{j=0}^q b_j x^{q-j} = \sum_{i=0}^p \sum_{j=0}^q a_i b_j x^{i-j+q}.$$

Доказательство завершается выписыванием второго полинома. ■

Дисперсия величины $\Delta^r u_t$ выражается соотношением

$$(22) \quad \begin{aligned} \mathcal{E}(\Delta^r u_t)^2 &= \mathcal{E}(\mathcal{P} - 1)^r u_t (\mathcal{P} - 1)^r u_t = \\ &= \sigma^2 \binom{2r}{r} = \sigma^2 \frac{(2r)!}{(r!)^2}. \end{aligned}$$

Это связано с тем, что правая часть (22) есть умноженный на σ^2 коэффициент при x^r в выражении

$$(23) \quad \begin{aligned} (x - 1)^r (1 - x)^r &= (-1)^r (x - 1)^{2r} = \\ &= \sum_{j=0}^{2r} (-1)^{j+r} \binom{2r}{j} x^j. \end{aligned}$$

Ковариация величин $\Delta^r u_t$ и $\Delta^r u_{t+s}$ есть

$$(24) \quad \begin{aligned} \mathcal{E} \Delta^r u_t \Delta^r u_{t+s} &= \mathcal{E} (\mathcal{P} - 1)^r u_t (\mathcal{P} - 1)^r \mathcal{P}^s u_t = \\ &= \begin{cases} \sigma^2 (-1)^s \binom{2r}{r+s}, & s = 0, 1, \dots, r, \\ 0, & s = r+1, \dots \end{cases} \end{aligned}$$

Скользящее среднее является линейным оператором и может быть представлено в виде полинома от оператора $\mathcal{P}$:

$$(25) \quad \sum_{s=-m}^m c_s y_{t+s} = \left(\sum_{s=-m}^m c_s \mathcal{P}^{m+s} \right) y_{t-m}.$$

Аналогичным образом можно записать сглаживающие формулы из разд. 3.3.1, а также остаток от скользящего среднего

$$(26) \quad y_t - \sum_{s=-m}^m c_s y_{t+s} = \left(\mathcal{P}^m - \sum_{s=-m}^m c_s \mathcal{P}^{m+s} \right) y_{t-m}.$$

Если сглаживающая формула (25) использует полином степени $2k+1$, то оператор (26) аннулирует (т. е. обращает в нуль) всякий полином степени $2k+1$ или меньшей. Покажем, что отсюда следует, что этот оператор действует (исключая сдвиг во времени) как линейная комбинация разностей $\Delta^{2k+2}, \dots, \Delta^{2m}$.

ЛЕММА 3.4.2. Если $Q(x)$ — полином от x степени n и $Q(\mathcal{P})$ аннулирует всякий полином степени $p (< n)$, то $Q(\mathcal{P})$ можно представить в виде линейной комбинации операторов $\Delta^{p+1}, \dots, \Delta^n$.

ДОКАЗАТЕЛЬСТВО. Пусть $Q(y+1) = d_n y^n + d_{n-1} y^{n-1} + \dots + d_1 y + d_0$. Поскольку $\mathcal{P} = \Delta + 1$, то

$$(27) \quad Q(\mathcal{P})f(t) = d_n \Delta^n f(t) + d_{n-1} \Delta^{n-1} f(t) + \dots + d_1 \Delta f(t) + d_0 f(t).$$

Если $f(t) = t^q$, $0 \leq q \leq p$, то

$$(28) \quad Q(\mathcal{P})f(t) = d_q \Delta^q t^q + d_{q-1} \Delta^{q-1} t^q + \dots + d_1 \Delta t^q + d_0 t^q = 0$$

ввиду того, что $\Delta^r t^q = 0$ для $r > q$. Поочередное рассмотрение (28) для $q = 0, 1, \dots, p$ показывает, что $d_0 = 0, d_1 = 0, \dots, d_p = 0$. Это и доказывает лемму. ■

Следствие 3.4.1. Остаток (26) сглаживающей формулы с $2m + 1$ членами, использующей полином степени $2k$ или $2k + 1$, является результатом применения линейной комбинации операторов $\Delta^{2k+2}, \dots, \Delta^{2m}$ к y_{t-m} .

В случае $m = k + 1$ указанный оператор¹⁾ имеет степень $2m = 2k + 2$ и поэтому он должен быть пропорционален $(\mathcal{P} - 1)^{2k+2} = \Delta^{2k+2}$, т. е. равен $C' \Delta^{2k+2}$. Значение C' определяется путем сравнения двух выражений для дисперсии величины

$$(29) \quad u_t - \sum_{s=-(k+1)}^{k+1} c_s u_{t+s} = C' \Delta^{2k+2} u_{t-k-1}.$$

Мы имеем из (22)

$$(30) \quad \mathcal{E} (C' \Delta^{2k+2} u_{t-k-1})^2 = C'^2 \binom{4k+4}{2k+2} \sigma^2.$$

Далее,

$$(31) \quad \mathcal{E} \left(u_t - \sum_{s=-(k+1)}^{k+1} c_s u_{t+s} \right)^2 = \mathcal{E} u_t^2 - \mathcal{E} \left(\sum_{s=-(k+1)}^{k+1} c_s u_{t+s} \right)^2 = \sigma^2 - \sigma^2 c_0.$$

Это следует из теоремы 3.3.1 и из некоррелированности оценки регрессии и остатка от нее. Поскольку $1 - c_0$ является коэффициентом при x^{k+1} в разложении $C' (x - 1)^{2k+2}$, то

$$(32) \quad (-1)^{k+1} C' \binom{2k+2}{k+1} = C'^2 \binom{4k+4}{2k+2}.$$

Поэтому

$$(33) \quad C' = (-1)^{k+1} \frac{\binom{2k+2}{k+1}}{\binom{4k+4}{2k+2}},$$

$$(34) \quad 1 - c_0 = \frac{\binom{2k+2}{k+1}^2}{\binom{4k+4}{2k+2}}.$$

¹⁾ Оператор $\mathcal{P}^m - \sum_{s=-m}^m c_s \mathcal{P}^{m+s}$ в (26). — Прим. перев.

При этом

$$(35) \quad c_{-s} = c_s = (-1)^{s+1} \frac{\binom{2k+2}{k+1} \binom{2k+2}{k+1+s}}{\binom{4k+4}{2k+2}}.$$

Эти же выражения были приведены ранее в формулах (19), (20), (21) § 3.3.

3.4.4. Оценивание дисперсии ошибки

Мы показали, что если тренд $f(t)$ на коротких интервалах близок к полиному низкой степени, то полученная последовательность разностей имеет средние, близкие к нулю. Если эти средние значения равны нулю, то

$$(36) \quad V_r = \frac{\sum_{t=1}^{T-r} (\Delta^r y_t)^2}{(T-r) \binom{2r}{r}}$$

является несмещенной оценкой σ^2 . Для произвольного тренда

$$(37) \quad \mathcal{E} V_r = \sigma^2 + \frac{\sum_{t=1}^{T-r} [\Delta^r f(t)]^2}{(T-r) \binom{2r}{r}}.$$

Пусть $K_r = 1 / \left[(T-r) \binom{2r}{r} \right]$. Тогда дисперсия V_r при $\Delta^r f(t) = 0$, $t = 1, \dots, T-r$, равна произведению K_r^2 на величину

$$(38) \quad \mathcal{E} \left[\sum_{t=1}^{T-r} (\Delta^r u_t)^2 - (T-r) \binom{2r}{r} \sigma^2 \right]^2 = \\ = \sum_{t,s=1}^{T-r} \mathcal{E} (\Delta^r u_t)^2 (\Delta^r u_s)^2 - (T-r)^2 \binom{2r}{r}^2 \sigma^4.$$

Лемма 3.4.3. Пусть $\sum_{i,j=1}^n a_{ij} u_i u_j = \mathbf{u}' \mathbf{A} \mathbf{u}$, где $\mathbf{A}$ — симметричная матрица, $\mathcal{E} u_i = 0$, $\mathcal{E} u_i^2 = \sigma^2$, $\mathcal{E} u_i u_j = 0$, $i \neq j$, $\mathcal{E} u_i^4 = \kappa_4 + 3\sigma^2$, $\mathcal{E} u_i^2 u_j^2 = \sigma^4$, $i \neq j$, и $\mathcal{E} u_i u_j u_k u_l = 0$, если индексы не равны попарно. Тогда

$$(39) \quad \mathcal{E} \sum_{i,j=1}^n a_{ij} u_i u_j = \sigma^2 \sum_{i=1}^n a_{ii} = \sigma^2 \operatorname{tr} \mathbf{A},$$

$$\begin{aligned}
 (40) \quad \mathcal{E} \left(\sum_{i,j=1}^n a_{ij} u_i u_j \right)^2 &= \sum_{i,j,k,l=1}^n a_{ij} a_{kl} \mathcal{E} u_i u_j u_k u_l = \\
 &= \kappa_4 \sum_{i=1}^n a_{ii}^2 + \sigma^4 \left(\sum_{i=1}^n a_{ii} \right)^2 + 2\sigma^4 \sum_{i,j=1}^n a_{ij}^2
 \end{aligned}$$

ЛЕММА 3.4.4. В условиях леммы 3.4.3 дисперсия величины $\sum_{i,j=1}^n a_{ij} u_i u_j$ выражается соотношением

$$\begin{aligned}
 (41) \quad \text{Var} \left(\sum_{i,j=1}^n a_{ij} u_i u_j \right) &= \kappa_4 \sum_{i=1}^n a_{ii}^2 + 2\sigma^4 \sum_{i,j=1}^n a_{ij}^2 = \\
 &= \kappa_4 \sum_{i=1}^n a_{ii}^2 + 2\sigma^4 \text{tr} \mathbf{A}^2.
 \end{aligned}$$

Если u_i нормально распределены, то $\kappa_4 = 0$ и первый член в правой части (41) равен 0.

Положим

$$(42) \quad S_r = \sum_{t=1}^{T-r} (\Delta^r y_t)^2 = \sum_{t,s=1}^T a_{ts}^{(r)} y_t y_s,$$

и пусть $\mathbf{A}_r = (a_{ts}^{(r)})$. Тогда $\mathbf{A}_0 = \mathbf{I}$,

$$(43) \quad \mathbf{A}_1 = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix},$$

$$(44) \quad \mathbf{A}_2 = \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & 0 \\ -2 & 5 & -4 & 1 & \dots & 0 \\ 1 & -4 & 6 & -4 & \dots & 0 \\ 0 & 1 & -4 & 6 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix},$$

$$(45) \quad \mathbf{A}_3 = \begin{bmatrix} 1 & -3 & 3 & -1 & 0 & \dots & 0 \\ -3 & 10 & -12 & 6 & -1 & \dots & 0 \\ 3 & -12 & 19 & -15 & 6 & \dots & 0 \\ -1 & 6 & -15 & 20 & -15 & \dots & 0 \\ 0 & -1 & 6 & -15 & 20 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix},$$

$$(46) \quad A_4 = \begin{bmatrix} 1 & -4 & 6 & -4 & 1 & 0 & \dots & 0 \\ -4 & 17 & -28 & 22 & -8 & 1 & \dots & 0 \\ 6 & -28 & 53 & -52 & 28 & -8 & \dots & 0 \\ -4 & 22 & -52 & 69 & -56 & 28 & \dots & 0 \\ 1 & -8 & 28 & -56 & 70 & -56 & \dots & 0 \\ 0 & 1 & -8 & 28 & -56 & 70 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

Кроме того, $V_r = K_r S_r$. При $\Delta' f(t) = 0$ вычисления дают:

$$(47) \quad \text{Var } V_0 = \frac{1}{T} \{ \kappa_4 + 2\sigma^4 \},$$

$$(48) \quad \text{Var } V_1 = \frac{1}{T-1} \left\{ \left[1 - \frac{1}{2(T-1)} \right] \kappa_4 + 2 \left[\frac{3}{2} - \frac{1}{2(T-1)} \right] \sigma^4 \right\},$$

$$(49) \quad \text{Var } V_2 = \frac{1}{T-2} \left\{ \left[1 - \frac{5}{9(T-2)} \right] \kappa_4 + 2 \left[\frac{35}{18} - \frac{2}{2(T-2)} \right] \sigma^4 \right\},$$

$$(50) \quad \text{Var } V_3 = \frac{1}{T-3} \left\{ \left[1 - \frac{69}{100(T-3)} \right] \kappa_4 + \right. \\ \left. + 2 \left[\frac{231}{100} - \frac{3}{2(T-3)} \right] \sigma^4 \right\},$$

$$(51) \quad \text{Var } V_4 = \frac{1}{T-4} \left\{ \left[1 - \frac{194}{245(T-4)} \right] \kappa_4 + \right. \\ \left. + 2 \left[\frac{1287}{490} - \frac{4}{2(T-4)} \right] \sigma^4 \right\}.$$

Здесь κ_4 — четвертый семинвариант.

Общее выражение для дисперсий величин V_r найти трудно из-за того, что элементы в обоих концах каждой диагонали A_r отличаются от элементов, расположенных ближе к середине. Запишем S_r в виде

$$(52) \quad S_r = \sum_{t=1}^{T-r} (\Delta' y_t)^2 = \sum_{t=1}^{T-r} \sum_{\alpha, \beta=0}^r (-1)^{\alpha+\beta} \binom{r}{\alpha} \binom{r}{\beta} y_{t+\alpha} y_{t+\beta} = \\ = \sum_{s,u=1}^T (-1)^{s+u} \sum_{t=\max(s-r, u-r, 1)}^{\min(s, u, T-r)} \binom{r}{s-t} \binom{r}{u-t} y_s y_u.$$

Элемент $a_{ss}^{(r)}$ равен коэффициенту при y_s^2 в (52), так что

$$(53) \quad a_{ss}^{(r)} = \begin{cases} \sum_{\alpha=0}^r \binom{r}{\alpha}^2 = \binom{2r}{r}, & s = r+1, \dots, T-r, \\ \sum_{\alpha=0}^{s-1} \binom{r}{\alpha}^2 = a_{T-s+1, T-s+1}^{(r)}, & s = 1, \dots, r. \end{cases}$$

Первая сумма здесь вычислена с использованием тождества

$$(54) \quad \sum_{j=0}^{2r} (-1)^{j+r} \binom{2r}{j} x^j = (-1)^r (x-1)^{2r} = \\ = (x-1)^r (1-x)^r = \sum_{\alpha, \beta=0}^r (-1)^{\alpha+\beta} \binom{r}{\alpha} \binom{r}{\beta} x^{r-\alpha+\beta}$$

и равна коэффициенту при x^r в последнем выражении. Остальные отличные от нуля элементы суть

$$(55) \quad a_{s, s+k}^{(r)} = \begin{cases} (-1)^k \sum_{\alpha=0}^{r-k} \binom{r}{\alpha} \binom{r}{\alpha+k} = (-1)^k \binom{2r}{r+k}, & s = r-k+1, \dots, T-r, \\ (-1)^k \sum_{\alpha=0}^{s-1} \binom{r}{\alpha} \binom{r}{\alpha+k} = a_{T-s-k+1, T-s+1}^{(r)}, & s = 1, \dots, r-k, \end{cases}$$

для $k = 1, \dots, r$.

Теперь можно подсчитать и указанные дисперсии, однако вычисления утомительны. Для четных r , $r = 2n$,

$$(56) \quad \sum_{s=1}^T [a_{ss}^{(r)}]^2 = \sum_{s=1}^n \{[a_{ss}^{(r)}]^2 + [a_{T-s+1, T-s+1}^{(r)}]^2\} + \\ + \sum_{s=n+1}^{2n} \{[a_{ss}^{(r)}]^2 + [a_{T-s+1, T-s+1}^{(r)}]^2\} + (T-2r) \binom{2r}{r}^2 = \\ = (T-r) \binom{2r}{r}^2 + 4 \sum_{s=1}^n a_{ss}^{(r)} \left[a_{ss}^{(r)} - \binom{2r}{r} \right],$$

поскольку $a_{T-s+1, T-s+1}^{(r)} = a_{ss}^{(r)}$ и

$$(57) \quad a_{r+1-s, r+1-s}^{(r)} = \binom{2r}{r} - a_{ss}^{(r)}, \quad s = 1, \dots, r.$$

Подобным же образом, если $r = 2n + 1$, то

$$(58) \quad \sum_{s=1}^T [a_{ss}^{(r)}]^2 = (T-2r) \binom{2r}{r}^2 + 2 \sum_{s=1}^n [a_{ss}^{(r)}]^2 + 2 [a_{n+1, n+1}^{(r)}]^2 + \\ + 2 \sum_{s=n+2}^{2n+1} [a_{ss}^{(r)}]^2 = \\ = (T-r) \binom{2r}{r}^2 + 4 \sum_{s=1}^n a_{ss}^{(r)} \left[a_{ss}^{(r)} - \binom{2r}{r} \right] - \frac{1}{2} \binom{2r}{r}^2,$$

ввиду того что

$$(59) \quad a_{n+1,n+1}^{(r)} = \frac{1}{2} \binom{2r}{r}, \quad r = 2n + 1.$$

В обоих случаях главный член равен $(T - r) \binom{2r}{r}^2$.

Для $k = 1, \dots, r$ можно найти, что

$$(60) \quad \sum_{s=1}^{T-k} [a_{s,s+k}^{(r)}]^2 = (T - 2r + k) \binom{2r}{r+k}^2 + 2 \sum_{s=1}^{r-k} [a_{s,s+k}^{(r)}]^2 =$$

$$= \begin{cases} (T - r) \binom{2r}{r+k}^2 + 4 \sum_{s=1}^{(r-k)/2} a_{s,s+k}^{(r)} \left[a_{s,s+k}^{(r)} - \binom{2r}{r+k} \right], & \text{если } r - k \text{ четное,} \\ (T - r) \binom{2r}{r+k}^2 + 4 \sum_{s=1}^{(r-k-1)/2} a_{s,s+k}^{(r)} \left[a_{s,s+k}^{(r)} - \binom{2r}{r+k} \right] - \\ - \frac{1}{2} \binom{2r}{r+k}^2, & \text{если } r - k \text{ нечетное,} \end{cases}$$

(см. упр. 55), так что главный член в $\sum_{s,t=1}^T [a_{st}^{(r)}]^2$ запишется в виде

$$(61) \quad (T - r) \left[\binom{2r}{r}^2 + 2 \sum_{k=1}^r \binom{2r}{r+k}^2 \right] = (T - r) \sum_{k=-r}^r \binom{2r}{r+k}^2 =$$

$$= (T - r) \binom{4r}{2r}.$$

При $r = 2n$ второе слагаемое равно

$$(62) \quad 4 \sum_{s=1}^n a_{ss}^{(r)} \left[a_{ss}^{(r)} - \binom{2r}{r} \right] + 8 \sum_{s=1}^{n-1} a_{s,s+1}^{(r)} \left[a_{s,s+1}^{(r)} - \binom{2r}{r+1} \right] - \binom{2r}{r+1}^2 +$$

$$+ 8 \sum_{s=1}^{n-1} a_{s,s+2}^{(r)} \left[a_{s,s+2}^{(r)} - \binom{2r}{r+2} \right] + \dots$$

$$\dots + 8 a_{1,r-1}^{(r)} \left[a_{1,r-1}^{(r)} - \binom{2r}{2r-2} \right] - \binom{2r}{2r-1}^2.$$

Кендалл и Стьюарт [(1966, стр. 389)] приводят его к виду $-\beta_r =$
 $= -\frac{r}{2} \binom{2r}{r}^2$. Таким образом, при $\Delta^r f(t) = 0$

$$(63) \quad \text{Var } S_r = \left[(T - r) \binom{2r}{r}^2 - \alpha_r \right] \kappa_4 + 2 \left[(T - r) \binom{4r}{2r} - \beta_r \right] \sigma^4,$$

$$(64) \quad \text{Var } V_r =$$

$$= \frac{1}{T-r} \left\{ \left[-\frac{\alpha_r}{(T-r) \binom{2r}{r}^2} \right] \kappa_4 + 2 \left[\frac{\binom{4r}{2r}}{\binom{2r}{r}^2} - \frac{\beta_r}{(T-r) \binom{2r}{r}^2} \right] \sigma^4 \right\},$$

где κ_4 — четвертый семинвариант, а α_r определяется из (56) или (58). [Кендалл и Стьюарт (1966, стр. 389) приводят выражение и для α_r .]

Если T велико, то дисперсия V_r приблизительно равна

$$(65) \quad \frac{1}{T-r} \left\{ \kappa_4 + 2 \frac{\binom{4r}{2r}}{\binom{2r}{r}^2} \sigma^4 \right\}.$$

При нормально распределенных u_i семинвариант κ_4 равен 0.

Поскольку дисперсия при $T \rightarrow \infty$ стремится к нулю, то V_r является состоятельной оценкой для σ^2 . Можно показать, что, когда κ_4 существует, а u_i независимы и одинаково распределены, величина

$$(66) \quad \sqrt{T-r} \frac{V_r - \sigma^2}{\sqrt{\kappa_4 + 2\sigma^4 \frac{\binom{4r}{2r}}{\binom{2r}{r}^2}}}$$

имеет в пределе стандартное нормальное распределение со средним 0 и дисперсией 1. (См. разд. 3.4.5.) Если $\kappa_4 = 0$, то дисперсия

$\sqrt{T-r} V_r$ асимптотически эквивалентна $2\sigma^4 \frac{\binom{4r}{2r}}{\binom{2r}{r}^2}$ и последняя величина служит мерой эффективности V_r как оценки для σ^2 . Как и следовало ожидать, дисперсия V_r с ростом r возрастает.

ЛЕММА 3.4.5. Ковариация двух (симметричных) квадратичных форм $\sum_{i,j=1}^n a_{ij}u_iu_j = \mathbf{u}'\mathbf{A}\mathbf{u}$ и $\sum_{k,l=1}^n b_{kl}u_ku_l = \mathbf{u}'\mathbf{B}\mathbf{u}$ в условиях леммы 3.4.3 выражается соотношением

$$(67) \quad \text{Cov} \left(\sum_{i,j=1}^n a_{ij}u_iu_j, \sum_{k,l=1}^n b_{kl}u_ku_l \right) = \\ = \kappa_4 \sum_{i=1}^n a_{ii}b_{ii} + 2\sigma^4 \sum_{i,j=1}^n a_{ij}b_{ij} = \kappa_4 \sum_{i=1}^n a_{ii}b_{ii} + 2\sigma^4 \text{tr } \mathbf{AB}.$$

Можно показать, что ковариация V_r и V_q приближенно равна

$$(68) \quad \frac{1}{T-r} \left\{ \kappa_4 + 2 \frac{\binom{2r+2q}{r+q}}{\binom{2r}{r} \binom{2q}{q}} \sigma^4 \right\}.$$

Точная формула для ковариации отличается от (68) слагаемыми, возникающими из-за того, что коэффициенты членов, близких к концам ряда, отличаются от коэффициентов членов, близких к его середине. Кендалл и Стьюарт (1966, стр. 389) приводят точное выражение для этой ковариации в случае $q = r + 1$.

Кенуй (1953) указал, что линейная комбинация $c_p V_p + \dots + c_q V_q$, где $c_p + \dots + c_q = 1$, также является несмещенной оценкой для σ^2 (если

$\Delta^p f(t) = 0$). Если $p = 0$, то $V_0 = \sum_{i=1}^T y_i^2 / T$ — наилучшая несмещенная оценка для σ^2 , поскольку в этом случае $f(t) \equiv 0$. В предположении нормальности она является достаточной статистикой для σ^2 . (Если $p = 1$ и, следовательно, $f(t) = \text{const}$, то в предположении нормальности достаточное множество статистик образуют статисти-

ки $\bar{y}$ и $\sum_{i=1}^T (y_i - \bar{y})^2$.) Для $p \geq 1$ Кенуй поставил вопрос о том, какая линейная комбинация имеет минимальную дисперсию для $\kappa_4 = 0$ и различных значений p и q . Если $p = 1$, а $q = 4$, то «наилучшими» коэффициентами будут $c_1 = 7.5$, $c_2 = -20.25$, $c_3 = 22.50$, $c_4 = -8.75$. Соответствующая дисперсия равна $3/4$ дисперсии V_1 . Оказывается, что знаки коэффициентов этих линейных комбинаций чередуются. Таким образом, нет уверенности в том, что оценка получится положительной. Поскольку эти коэффициенты существенно больше единицы, то можно ожидать, что вероятность получения отрицательного значения оценки вовсе не мала.

Постановка этой задачи приводит к вопросу о причинах ограничений на числа V_r и, что более важно, о причинах, по которым статистик ограничивается только суммой квадратов переменных разностей. Что следует считать наилучшей оценкой σ^2 , когда тренд «гладкий»? Трудность состоит здесь в необходимости такого достаточно строгого определения «гладкости», при котором задача наилучшего оценивания была бы математически определена. Если гладким считать тренд, являющийся полиномом степени q , то наилучшей оценкой σ^2 будет оценка, приведенная в § 3.2. Любые другие определения гладкости являются либо недостаточно четкими, либо слишком сложными.

Кендалл (1946а, задача 30.8) предложил модифицировать метод переменных разностей путем введения фиктивных переменных $y_{-r+1} = y_{-r+2} = \dots = y_0 = 0$ и $y_{t+1} = \dots = y_{t+r} = 0$ и под-

счета сумм квадратов $\Delta^r y_t$ для значений t от $-r + 1$ до T . В таком случае существуют простые соотношения, связывающие эти суммы квадратов с суммами попарных произведений $\sum_{i=1}^{T-j} y_i y_{t+j}$. Кенуй (1953) предложил другую модификацию, связанную с изменением крайних членов. Так, например, у него m -я модифицированная сумма равна среднему двух $(m - 1)$ -х модифицированных сумм

$$(69) \quad \sum_{t=1}^T w_t = \frac{1}{2} \left[\sum_{t=1}^{T-1} w_t + \sum_{t=2}^T w_t \right],$$

где

$$(70) \quad \sum_{t=1}^T w_t = \sum_{t=1}^T w_t.$$

Далее, можно использовать модифицированные суммы $(\Delta^r y_t)^2$ и $y_t y_{t+j}$ и связать их. Иная модификация, приводящая к некоторым упрощениям, состоит в использовании A'_t вместо $A_{t,r}$. (См. гл. 6.)

3.4.5. Определение степени гладкости тренда

Можно интересоваться также вопросом о том, обладает ли тренд определенной степенью гладкости. От этого, например, зависит выбор сглаживающей формулы. Таким образом, может возникнуть необходимость выяснить, является ли тренд гладким в том смысле, что он в каждом интервале времени может быть адекватно представлен полиномом вполне определенной степени q . Это соответствует задаче проверки гипотезы о том, что данная степень является приемлемой для описания тренда, против альтернативы, состоящей в том, что данная степень недостаточна для его описания. Кроме того, можно рассмотреть задачу со многими решениями об определении приемлемой степени (в пределах между двумя заданными значениями m и q) полинома, аппроксимирующего тренд.

Если под точным соответствием понимать совпадение выравнивающей функции с единственным адекватным представлением тренда, то перечисленные задачи полностью равносильны изученным в § 3.2 для полиномиального тренда. Если допустить более широкое, но нечеткое толкование адекватности представления, то задачи теряют математическую определенность. Вполне возможно особое внимание уделить максимально допускаемому расхождению между действительным трендом и его полиномиальной аппроксимацией. Однако получаемые при этом математические задачи трудны для решения. Следует отметить, что общая теория задач со многими решениями, являющаяся обобщением теории § 3.2, здесь неприменима.

Рассмотрим указанные задачи, ограничиваясь использованием сумм квадратов переменных разностей. Из (37) видно, что математическое ожидание величины V_r зависит от суммы

$\sum_{t=r}^{T-r} [\Delta^r f(t)]^2$. Эта сумма равна нулю, если полином имеет степень, меньшую r , и близка к нулю, если функция $f(t)$ близка к полиному степени, меньшей r , для каждого набора $(r+1)$ последовательных значений t . Отсюда следует, что V_r можно использовать в упомянутых статистических задачах. Принимать решение о том, что $\Delta^r f(t)$ близка к нулю для всех значений t (т. е. что полиномы $(r-1)$ -й степени дают адекватное представление), против альтернативы близости к нулю только разности $\Delta^{r+1} f(t)$ для всех t , можно, например, убедившись в том, что V_r ненамного больше V_{r+1} .

Сглаживающие формулы, рассмотренные в § 3.3 и 3.4, были основаны на предположении о том, что полином степени $2k$ или $2k+1$ дает адекватное представление тренда в интервале $2m+1$ последовательных моментов времени. В частности, мы отмечали, что для $m=k+1$ смещение при оценивании тренда равно $C' \Delta^{2k+2} f(t)$. Таким образом, вопросы, которые мы сейчас изучаем, соответствуют задачам о выборе подходящих сглаживающих формул.

Рассмотрим проверку гипотезы $\Delta^r f(t) = 0$ в предположении, что $\Delta^q f(t) = 0$ ($q > r$). Мы отвергнем эту гипотезу, если V_r будет намного больше V_q . Такая процедура может быть основана на статистике

$$(71) \quad \sqrt{T-q} \frac{V_r - V_q}{V_q} = \sqrt{T-q} \left(\frac{V_r}{V_q} - 1 \right).$$

Из (68) видно, что дисперсия числителя $\sqrt{T-q} (V_r - V_q)$ приблизительно равна

$$(72) \quad \left\{ \kappa_4 + 2 \frac{\binom{4r}{2r}}{\binom{2r}{r}^2} \sigma^4 \right\} + \left\{ \kappa_4 + 2 \frac{\binom{4q}{2q}}{\binom{2q}{q}^2} \sigma^4 \right\} - \\ - 2 \left\{ \kappa_4 + 2 \frac{\binom{2q+2r}{q+r}}{\binom{2q}{q} \binom{2r}{r}} \sigma^4 \right\} = 2\sigma^4 \left[\frac{\binom{4r}{2r}}{\binom{2r}{r}^2} + \frac{\binom{4q}{2q}}{\binom{2q}{q}^2} - 2 \frac{\binom{2q+2r}{q+r}}{\binom{2q}{q} \binom{2r}{r}} \right].$$

Если указанная гипотеза верна, то математическое ожидание числителя (71) равно нулю; в противном случае оно положительно.

Разность $V_r - V_q$ можно записать в виде

$$(73) \quad V_r - V_q = \frac{1}{T-q} \sum_{t=1}^{T-q} \left[\frac{(\Delta' y_t)^2}{\binom{2r}{r}} - \frac{(\Delta^q y_t)^2}{\binom{2q}{q}} \right] + \\ + \left(\frac{1}{T-r} - \frac{1}{T-q} \right) \sum_{t=1}^{T-q} \frac{(\Delta' y_t)^2}{\binom{2r}{r}} + \frac{1}{T-r} \sum_{t=T-q+1}^{T-r} \frac{(\Delta' y_t)^2}{\binom{2r}{r}}.$$

Каждое из последних двух слагаемых, умноженное на $\sqrt{T-q}$, сходится по вероятности к нулю. (Отметим, что множитель, стоящий первым во втором слагаемом, имеет порядок $1/T^2$.) Первое слагаемое является средним для $T-q$ величин. Если u_t независимы и одинаково распределены, то рассматриваемые величины распределены также одинаково, но уже не являются независимыми. Тем не менее их последовательность образует стационарный случайный процесс (см. гл. 7), причем члены, отстоящие друг от друга более чем на q , являются независимыми. Это так называемый стационарный случайный процесс с конечной зависимостью. Теорема 7.7.5 утверждает, что $\sqrt{T-q} (V_r - V_q)$ имеет в пределе нормальное распределение. Поскольку V_q является состоятельной оценкой для σ^2 (вне зависимости от того, является ли нулевая гипотеза истинной или ложной), то и (71) имеет в пределе нормальное распределение, дисперсия которого получается из (72) опусканием σ^4 .

ТЕОРЕМА 3.4.1. Если $\Delta' f(t) = 0, t = 1, \dots, T-r$, а $u_1, u_2, \dots, u_T$ независимы и одинаково распределены с $\mathbb{E} u_t = 0$ и $\mathbb{E} u_t^4 < \infty$, то статистика

$$(74) \quad \sqrt{\frac{1}{2} (T-q) \left[\frac{\binom{4r}{2r}}{\binom{2r}{r}^2} + \frac{\binom{4q}{2q}}{\binom{2q}{q}^2} - 2 \frac{\binom{2q+2r}{q+r}}{\binom{2q}{q} \binom{2r}{r}} \right]} \cdot \frac{V_r - V_q}{V_q}$$

имеет в пределе нормальное распределение с нулевым средним и единичной дисперсией.

Поэтому в случае больших выборок мы отвергаем нулевую гипотезу с уровнем значимости ε , если вычисленное по выборке значение (74) превышает $t(2\varepsilon)$, где

$$(75) \quad \int_{t(2\varepsilon)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du = \varepsilon.$$

В теории больших выборок мы не использовали точную дисперсию, поскольку разность между ней и (72) стремится к нулю при

$T \rightarrow \infty$, а центральная предельная теорема, на основании которой строится указанная процедура, не чувствительна к подобным разностям. Возможно, что учет дополнительных слагаемых может привести к тому, что асимптотическое распределение будет лучше приближать точное распределение для данного T . Однако это неизвестно. Тинтнер (1940) брал более точные значения моментов (включая состоятельную оценку семинварианта κ_4) и привел таблицы, облегчающие вычисление V_r и использование предельного распределения.

Если u_t распределены нормально, то в случае истинности нулевой гипотезы статистика (74) имеет распределение, не зависящее от мешающих параметров. Распределения квадратичных форм от нормально распределенных переменных и отношений таких квадратичных форм будут детально изучены в гл. 6 в связи с рассмотрением сериальной корреляции. Там будет показано, что распределение величины $(V_r - V_q)/V_q$, или, эквивалентно, V_r/V_q , является весьма сложным и не может быть приведено к простой канонической форме. С целью упрощения отыскания этого распределения был предложен ряд модификаций. Тинтнер (1940, гл. 8) предложил заменить V_r и V_q соответственно суммами слагаемых $(\Delta^r y_t)^2$ и $(\Delta^q y_t)^2$, отбираемых таким образом, чтобы y_t не появлялось дважды. При этом и числитель, и знаменатель являются суммами квадратов независимых нормально распределенных величин и их нормированное отношение имеет F -распределение. Этот метод, однако, крайне неэффективен, поскольку число членов в каждой сумме составляет лишь $1/(q + r)$ от максимально возможного. Другая модификация, предложенная Тинтнером (1955), состоит в использовании циклического определения (см. гл. 6). Это упрощает задачу отыскания распределения, но может привести к значительному смещению, поскольку тренд в начале ряда часто бывает совершенно отличным от тренда в его конце. (Фактически смещение может возрастать с ростом T .) Камат (1955) и Гейссер (1956) предложили опускать одно или два средних слагаемых в выражении для V_r (или V_q) с целью упрощения распределения последних. Другая возможность состоит в замене A_r на A'_r . Мы обсудим эти задачи в дальнейшем в гл. 6 (где обозначение A_1 используется для других матриц).

Рассмотрим задачу со многими решениями о выборе одной из следующих гипотез:

$$\begin{aligned}
 & H_q: \Delta^q f(t) \neq 0, \\
 & H_{q-1}: \Delta^q f(t) = 0, \quad \Delta^{q-1} f(t) \neq 0, \\
 & \vdots \\
 & H_{m+1}: \Delta^{m+2} f(t) = 0, \quad \Delta^{m+1} f(t) \neq 0, \\
 & H_m: \Delta^{m+1} f(t) = 0.
 \end{aligned}
 \tag{76}$$

Здесь $\Delta^r f(t) = 0$ означает выполнение равенства для всех t , а $\Delta^r f(t) \neq 0$ означает нарушение равенства хотя бы для одного t . Эта совокупность предположений подобна (16) из § 3.2. Соответствующие решающие процедуры должны быть основаны на статистиках $V_{m+1}, \dots, V_{q+1}$. Предположим, что $\Delta^{q+1} f(t) = 0$. Как уже отмечалось, общая теория § 3.2 в данном случае неприменима. Дело в том, что V_r не являются здесь достаточными статистиками. Тем не менее в соответствии с методикой § 3.2 исследователь может действовать последовательно, начиная с проверки гипотезы $\Delta^q f(t) = 0$ с помощью отношения $(V_q - V_{q+1})/V_{q+1}$ так, как это было указано выше. Если эта гипотеза принимается, то проверяется гипотеза $\Delta^{q-1} f(t) = 0$ по значениям $(V_{q-1} - V_q)/V_q$. На каждом конкретном шаге используется $(V_r - V_{r+1})/V_{r+1}$. Если гипотеза $\Delta^{m+1} f(t) = 0$ принимается, то на этом процедура заканчивается. Для длинных рядов (т. е. при больших T) эти критерии могут основываться на асимптотической теории, приведенной выше. Однако числители не являются асимптотически независимыми. Поэтому предельная вероятность принятия гипотезы $\Delta^q f(t) = 0$, а затем гипотезы $\Delta^{q-1} f(t) = 0$ описывается двумерным нормальным распределением. Предельная вероятность для r решений требует привлечения r -мерного нормального распределения. На практике обычно обходят эту трудность, производя проверку по каждому критерию в отдельности.

Будем, как уже однажды предполагалось, совершать последовательные действия в обратном порядке. Сначала рассмотрим отношение $(V_{m+1} - V_{m+2})/V_{m+2}$. Если оно велико, то примем решение $\Delta^{m+1} f(t) \neq 0$. После этого рассмотрим $(V_{m+2} - V_{m+3})/V_{m+3}$ и т. д. Практически эта процедура может быть проведена следующим образом. Применяя оператор вычисления разностей к исходной последовательности, получают последовательность $\{\Delta y_t\}$ и вычисляют V_1 . Затем, применяя разностный оператор к $\{\Delta y_t\}$, получают последовательность $\{\Delta^2 y_t\}$ и вычисляют V_2 . Поскольку величины V_r определяются одна за другой, то каждую из них можно сравнивать с предыдущей. Критерии значимости могут быть применены, только когда задана последовательность уровней значимости. Однако, помимо трудности определения вероятностей ошибок из-за зависимости составляющих критериев значимости, осложняющим обстоятельством является еще и положительность вероятности того, что, например, величина $\sum_{t=1}^{T-r} [\Delta^r f(t)]^2$ мала, а $\sum_{t=1}^{T-r} [\Delta^{r+1} f(t)]^2$ ве-

лика, а это может привести к ошибочному заключению о том, что $\Delta^r f(t) = 0$.

Тинтнер (1952, стр. 320) привел для ряда из табл. 3.1 разности до 5-го порядка. (Последние числа в каждом столбце таблицы Тинтнера ошибочны.)

Таблица 3.5
ДИСПЕРСИИ ДЛЯ РАЗНОСТЕЙ ИЗ ТАБЛ. 3.1

Порядок	S_r	$(T-r) \binom{2r}{r}$	V_r
0			62.2517
1	905.64	44	20.5827
2	1 860.23	126	14.7637
3	5 411.77	400	13.5294
4	17 321.25	1330	13.0235
5	58 446.06	4536	12.8849

Значения статистик V_r , вычисленные по этим разностям, помещены в табл. 3.5. При $r = 0$ статистика V_r равна $\sum_{t=1}^{23} (y_t - \bar{y})^2 / 22$ и совпадает со средней остаточной суммой квадратов для $k = 0$ в нижней половине табл. 3.2. Заметим, что оценка V_r для σ^2 , основанная на аппроксимирующем полиноме степени $r - 1$, значительно меньше оценки, являющейся средней суммой квадратов остатков относительно выравнивающего полинома степени $r - 1$.

Тинтнер вместе с разностью $V_r - V_{r+1}$ в качестве состоятельной оценки для σ^2 использует вместо статистики V_{r+1} статистику V_r . Последняя представляется более предпочтительной ввиду того, что она состоятельна при $\Delta^r f(t) \neq 0$ и $\Delta^{r+1} f(t) = 0$.

3.5. НЕЛИНЕЙНЫЕ ТРЕНДЫ

Иногда во временных рядах проявляются тренды, которые лучше всего описываются функциями, нелинейными по параметрам, подлежащим оценке. Например, при изучении популяций часто обнаруживается характерный тренд, связанный с ростом популяций, который можно достаточно хорошо описать так называемой «логистической» кривой роста. [См. Дэвис (1941, стр. 247—271).] Логистическая кривая как функция времени выражается формулой

$$(1) \quad f(t; \alpha_1, \alpha_2, \alpha_3) = \frac{\alpha_1}{1 + \alpha_2 e^{-\alpha_3 t}}.$$

При построении модели ошибок предполагают, что наблюдаемые значения размера популяции $y_1, \dots, y_T$ отличаются от тренда некоррелированными случайными величинами, так что для оценивания параметров $\alpha_1, \alpha_2, \alpha_3$ можно привлечь критерий наименьших квадратов, так же как это было сделано в линейном случае. В связи с этим потребуем, чтобы выражение

$$(2) \quad \sum_{t=1}^T [y_t - f(t; a_1, a_2, a_3)]^2 = S(a_1, a_2, a_3)$$

принимало минимально возможное значение по параметрам a_1, a_2, a_3 . Минимизирующие значения a_1, a_2 и a_3 будут при этом оценками наименьших квадратов параметров α_1, α_2 и α_3 соответственно. Задачу вычисления этих оценок часто можно решить итерационным методом с использованием первых членов тейлоровского разложения $f(t; a_1, a_2, a_3)$ в окрестности некоторых подходящих значений a_1^0, a_2^0, a_3^0 :

$$(3) \quad f(t; a_1, a_2, a_3) = f(t; a_1^0, a_2^0, a_3^0) + \sum_{j=1}^3 (a_j - a_j^0) \frac{\partial f(t; a)}{\partial a_j} \Big|_{a_1^0, a_2^0, a_3^0} + R,$$

где R — остаточный член. При выбранных начальных оценках a_1^0, a_2^0, a_3^0 можно воспользоваться линейной техникой наименьших квадратов применительно к соотношению (3) и получить новые оценки a_1^1, a_2^1, a_3^1 . Линейные уравнения относительно $a_i^1 - a_i^0$, получаемые при отбрасывании остаточного члена R , имеют вид

$$(4) \quad \sum_{t=1}^T \frac{\partial f(t; a^0)}{\partial a^0} \left[\frac{\partial f(t; a^0)}{\partial a^0} \right]' (a - a^0) = \sum_{t=1}^T [y_t - f(t; a^0)] \frac{\partial f(t; a^0)}{\partial a^0},$$

где $f(t; a^0) = f(t; a_1^0, a_2^0, a_3^0)$, $(a - a^0)$ есть вектор-столбец с элементами $(a_i - a_i^0)$, а $\partial f(t; a^0)/\partial a^0$ — вектор, элементами которого являются частные производные по a_1, a_2, a_3 , вычисленные при выбранных начальных значениях этих параметров. Чтобы получить более точные решения, указанную процедуру можно повторить, используя вместо a_1^0, a_2^0 и a_3^0 соответственно значения a_1^1, a_2^1 и a_3^1 и т. д. Получаемая таким образом последовательность решений во многих случаях сходится к значениям, минимизирующим (2).

Можно поступать и иначе. Именно, формулу (2) использовать в точном виде, продифференцировать ее по a_1, a_2, a_3 и приравнять нулю соответствующие частные производные. При этом получаются соотношения

$$(5) \quad \sum_{t=1}^T f(t; a_1, a_2, a_3) \frac{\partial f(t, a)}{\partial a_i} = \sum_{t=1}^T y_t \frac{\partial f(t; a)}{\partial a_i}, \quad i = 1, 2, 3.$$

Если теперь разложить исходную функцию и ее частные производные в ряды Тейлора в окрестности точки (a_1^0, a_2^0, a_3^0) и опустить члены второго порядка малости относительно $(a_i - a_i^0)$, то в результате получатся уравнения

$$(6) \quad \left\{ \sum_{t=1}^T \frac{\partial f(t; \mathbf{a}^0)}{\partial \mathbf{a}^0} \left[\frac{\partial f(t; \mathbf{a}^0)}{\partial \mathbf{a}^0} \right]' - \sum_{t=1}^T [y_t - f(t; \mathbf{a}^0)] \frac{\partial^2 f(t; \mathbf{a}^0)}{\partial \mathbf{a}^0 \partial \mathbf{a}^{0'}} \right\} (\mathbf{a} - \mathbf{a}^0) = \\ = \sum_{t=1}^T [y_t - f(t; \mathbf{a}^0)] \frac{\partial f(t; \mathbf{a}^0)}{\partial \mathbf{a}^0},$$

в которых

$$(7) \quad \frac{\partial^2 f(t; \mathbf{a}^0)}{\partial \mathbf{a}^0 \partial \mathbf{a}^{0'}} = \left(\frac{\partial^2 f(t; \mathbf{a})}{\partial a_i \partial a_j} \Big|_{a_1^0, a_2^0, a_3^0} \right).$$

Если предполагается, что случайные составляющие — независимые нормально распределенные случайные величины с одинаковыми дисперсиями, то методы наименьших квадратов и максимального правдоподобия совпадают и данный итерационный метод эквивалентен методу Ньютона — Рафсона отыскания максимума. Если в (6) вместо (7) используется его математическое ожидание, то соответствующая процедура известна как метод меток. [См. С. Рао (1952, разд. 4с.2).]

Для минимизации выражения (2) можно воспользоваться и иными способами. В начальной точке $\mathbf{a}^0$ функция $S(a_1, a_2, a_3)$ убывает наиболее быстро в направлении $-\partial S/\partial a_1, -\partial S/\partial a_2, -\partial S/\partial a_3$. Изменение вектора $\mathbf{a}$ в этом направлении приводит к уменьшению $S(a_1, a_2, a_3)$. Для оценки перемещения, приводящего к наибольшему убыванию, применяются методы наискорейшего спуска. На практике частные производные могут быть приближенно заменены конечными разностями. Отправляясь от начального приближения (a_1^0, a_2^0, a_3^0) , вычисляют S для $a_i = -\lambda \partial S/\partial a_i$ при возрастающих значениях λ до тех пор, пока S не перестанет убывать. При этом получают вторую точку $\mathbf{a}^1$ и повторяют процедуру. Дрейпер и Смит (1966, разд. 10.3) обсуждают метод («компромисс Маркардта»), базирующийся одновременно на линеаризации функции $f(t; \mathbf{a})$ и на методе наискорейшего спуска. (Нелинейная регрессия рассматривалась также Уильямсом (1959).)

Мы указали на некоторые общие подходы к отысканию приближенных решений задачи наименьших квадратов итерационным путем. При этом формулы выписывались только для случая трех параметров лишь из соображений удобства. В этой области сделано

довольно многое, но в настоящей книге не представляется целесообразным давать обзор сделанного.

Много результатов получено также в области выравнивания наблюдаемых данных с помощью специальных функций, таких, как логистическая. [См., например, Дэвис (1941, стр. 250—254).] Мы не будем пытаться обрисовать эту деятельность даже в общих чертах. Некоторые из методов используют порядок наблюдений во времени. Например, можно минимизировать

$$(8) \quad \sum_{t=1}^{T-1} [\Delta y_t - \Delta f(t; a_1, a_2, a_3)]^2.$$

При этом допускается, что функция, подлежащая оценке, определяется другим способом. Такие методы могут основываться не только на модели с некоррелированными случайными ошибками, но и на других моделях, оказывающихся во многих случаях более подходящими.

3.6. ОБСУЖДЕНИЕ

В моделях, изучавшихся в настоящей главе, предполагалось, что наблюдения y_t получаются как результат наложения на некоторую функцию времени $f(t)$ не коррелированных с ней ошибок u_t . Кроме того, допускалось, что эта функция времени не возмущена нерегулярностями. Такое предположение может быть оправдано в тех ситуациях, когда случайный характер нерегулярностей связан с процессом измерения. Одной из классических областей, в которых исследуются временные ряды, является астрономия. Пусть производятся последовательные наблюдения расположения какой-либо планеты. Наблюдаемые нерегулярности обусловлены главным образом изменениями в земной атмосфере и в установке телескопа. Эти случайные факторы не влияют на курс планет. Более того, можно считать, что эти факторы изо дня в день или из недели в неделю являются некоррелированными.

В других прикладных задачах указанные предположения могут и не выполняться. Так, регистрируемое количество мяса, ежегодно потребляемого в Соединенных Штатах, подвержено ошибкам измерений. Часть продукта не учитывается, часть потребления ошибочно приписывается другому году и т. п. Эти ошибки в известной мере коррелированы. В гл. 10 мы рассмотрим влияние коррелированности ошибок. Будет развита теория больших выборок для случая, когда ошибки образуют стационарный случайный процесс.

Во временных рядах, описывающих потребление мяса, существуют, однако, и другие нерегулярности, которые могут привести к тому, что представление действительного потребления мяса простой функцией времени окажется неудовлетворительным. Суще-

вует, например, воздействие общих экономических условий, которые влияют и на доход потребителей, и на цены на мясо. Эти воздействия не относятся ни к упомянутой функции времени, ни к некоррелированным случайным членам. Если такие воздействия не принимаются в расчет явно (как переменные регрессии), то они попадают в рассмотрение неявным образом (в $f(t)$ или u_t) и представляют собой нерегулярности, которые оказываются включенными в изучаемую величину. Если потребление мяса в какой-то год возрастает в силу случайного стечения обстоятельств, то потребители могут изменить свои вкусы, склоняясь к длительному увеличению потребления мяса. Такое поведение не отражено в моделях, рассмотренных в настоящей главе, но будет учитываться в моделях, изучаемых ниже. В гл. 5 будут объединены регрессионная модель для систематической составляющей и процесс авторегрессии для случайной составляющей.

ЛИТЕРАТУРА

§ 3.2. Р. Андерсон и Хаузман (1942), Леман (1959), Тинтнер (1952), Фишер и Изйтс (1963).

§ 3.3. Бокс и Дженкинс (1970), Вальд (1936), Дурбин (1963), Кендалл и Стьюарт (1966), Коуден (1962), Уиттекер и Робинсон (1926), Хеннан (1964).

§ 3.4. О. Андерсон (1929), Гейссер (1956), Жордан (1939), Камат (1955), Кейв-Браун-Кейв (1904), Кендалл (1946а), Кендалл и Стьюарт (1966), Кенуй (1953), Миллер (1960), Тинтнер (1940, 1952, 1955), Стьюдент (1914), Хукер (1905).

§ 3.5. Дрейпер и Смит (1966), Дэвис (1941), С. Рао (1952), Уильямс (1959).

УПРАЖНЕНИЯ

1. (Разд. 3.2.1) В полиноме $f(t)$, заданном соотношением (1), замена переменного t на $t^* + \tau$ приводит к изменению начала отсчета. Найдите коэффициенты полученного полинома от t^* . (Заметьте, что если $\alpha_q \neq 0$, то в каждом полиноме, получающемся в результате такой подстановки, присутствуют все степени t^* , меньшие q .)

2. (Разд. 3.2.1) Найдите ортогональный полином третьей степени для $t = 1, \dots, T$.

3. (Разд. 3.2.1) Найдите ортогональный полином третьей степени для $t = 1, \dots, 17$.

4. (Разд. 3.2.1) Найдите ортогональный полином третьей степени для $t = 1, \dots, 23$.

5. (Разд. 3.2.1) Пусть $\psi_k(T) = \sum_{t=1}^T t^k$. Покажите, что

$$\sum_{k=0}^p \binom{p+1}{k} \psi_k(T) = (T+1)^{p+1} - 1.$$

(Указание. Показать, что обе части уравнения равны $\sum_{t=1}^T (t+1)^{p+1} - \sum_{t=1}^T t^{p+1}$.)

6. (Разд. 3.2.1) Покажите, что

$$2 \sum_{i=0}^q \binom{2q+1}{2i} \psi_{2i}(T) = (T+1)^{2q+1} + T^{2q+1} - 1,$$

где $\psi_k(T)$ определено в упр. 5. (Указание. Показать, что обе части уравнения равны

$$\sum_{t=1}^T (t+1)^{2q+1} - \sum_{t=1}^T (t-1)^{2q+1}.)$$

7. (Разд. 3.2.1) Проверьте приведенную ниже таблицу значений функций

$\psi_k(T) = \sum_{t=1}^T t^k$, $k = 0, \dots, 8$, используя упр. 5:

$$\sum_{t=1}^T 1 = T$$

$$\sum_{t=1}^T t = \frac{T(T+1)}{2},$$

$$\sum_{t=1}^T t^2 = \frac{(2T+1)T(T+1)}{6},$$

$$\sum_{t=1}^T t^3 = \frac{T^2(T+1)^2}{4},$$

$$\sum_{t=1}^T t^4 = \frac{(2T+1)T(T+1)(3T^2+3T-1)}{30},$$

$$\sum_{t=1}^T t^5 = \frac{T^2(T+1)^2(2T^2+2T-1)}{12},$$

$$\sum_{t=1}^T t^6 = \frac{(2T+1)T(T+1)(3T^4+6T^3-3T+1)}{42},$$

$$\sum_{t=1}^T t^7 = \frac{T^2(T+1)^2(3T^4+6T^3-T^2-4T+2)}{24},$$

$$\sum_{t=1}^T t^8 = \frac{T(T+1)(2T+1)(5T^6+15T^5+5T^4-15T^3-T^2+9T-3)}{90}.$$

8. (Разд. 3.2.1) Найдите значения $\psi_{2h}(T) = \sum_{t=1}^T t^{2h}$ для $h = 1, 2, 3, 4$, используя упр. 6.

9. (Разд. 3.2.2) Пусть $x_1, \dots, x_N$ — независимые, нормально распределенные случайные величины с дисперсией σ^2 и математическими ожиданиями $\mathbb{E}x_i = \mu_i$;

$i = 1, \dots, n, \mathbb{E}x_i = 0, i = n+1, \dots, N$. Докажите, что для любого измеримого множества R из N -мерного пространства $x_1, \dots, x_N$ соотношение

$$(i) \quad \Pr \left\{ R \mid x_1, \dots, x_n, \sum_{i=n+1}^N x_i^2 \right\} = \varepsilon$$

влечет за собой

$$(ii) \quad \Pr \{R\} = \varepsilon.$$

(Указание. Взять математическое ожидание (i) относительно распределения величин $x_1, \dots, x_n$ и $\sum_{i=n+1}^N x_i^2$.)

10. (Разд. 3.2.2) В формулировке упр. 9 докажите, что из (ii) (тождественного по $\mu_1, \dots, \mu_n$ и σ^2) следует, что (i) выполняется почти всюду (т. е. с вероятностью 1). (Указание. (ii) равно

$$(iii) \quad 1 = \int_0^\infty \int_{-\infty}^\infty \dots \int_{-\infty}^\infty \frac{\Pr \{R \mid x_1, \dots, x_n, v\}}{\varepsilon} \prod_{i=1}^n n(x_i \mid \mu_i, \sigma^2) h\left(\frac{v}{\sigma^2}\right) \times \\ \times \prod_{i=1}^n dx_i \frac{dv}{\sigma^2}$$

тождественно по $\mu_1, \dots, \mu_n$ и σ^2 , где h — плотность распределения χ_{N-n}^2 , а $n(x_i \mid \mu_i, \sigma^2)$ — плотность нормального распределения со средним μ_i и дисперсией σ^2 . Если в (iii) $1/\sigma^2$ заменить на $(1/\sigma^2) - 2\theta_0$ и μ_i на $[\mu_i/\sigma^2 + \theta_0]/[(1/\sigma^2) - 2\theta_0]$, то можно получить производящую функцию для моментов подынтегрального выражения; она оказывается производящей функцией для моментов плотности совместного распределения величин $x_1, \dots, x_n$ и v .)

11. (Разд. 3.2.2) Пусть $x_1, \dots, x_N$ независимы и нормально распределены с дисперсиями σ^2 и математическими ожиданиями $\mathbb{E}x_i = \mu_i, i = 1, \dots, n+1, \mathbb{E}x_i = 0, i = n+2, \dots, N$. Докажите, что $x_1, \dots, x_{n+1}$ и $\sum_{i=n+2}^N x_i^2$ образуют достаточное множество статистик для $\mu_1, \dots, \mu_{n+1}$ и σ^2 .

12. (Разд. 3.2.2) Докажите, что в формулировке упр. 11 наилучший критерий для проверки гипотезы $\mu_{n+1} = 0$ с уровнем значимости ε (т. е. с вероятностью ε отвержения гипотезы равномерно по $\mu_1, \dots, \mu_n$ и σ^2 при $\mu_{n+1} = 0$) против альтернативы $\mu_{n+1} > 0$ имеет критическую область вида

$$(iv) \quad \frac{x_{n+1}}{\sqrt{x_{n+1}^2 + \dots + x_N^2}} > k.$$

(Указание. Поскольку в упр. 10 показано, что эта критическая область удовлетворяет (i), рассмотрите плотность при данных $x_1, \dots, x_n$ и $\sum_{i=n+1}^N x_i^2$. Сравните пересечение множества $x_1 = c_1, \dots, x_n = c_n, x_{n+1}^2 + \dots + x_N^2 = c$ с областью R , определенной выше, и его пересечение с любой другой областью R^* , удовлетворяющей (i).)

13. (Разд. 3.2.2) Докажите, что в формулировке упр. 11 равномерно наиболее мощный критерий для проверки гипотезы $\mu_{n+1} = 0$ с уровнем значимости ε против альтернативы $\mu_{n+1} > 0$ имеет критическую область (iv) из упр. 12.

14. (Разд. 3.2.2) Пусть $x_1, \dots, x_n$ независимы и нормально распределены с нулевыми средними и единичными дисперсиями, а S не зависит от $x_1, \dots, x_n$ и имеет распределение χ^2 с m степенями свободы. Пусть

$$t_i = \frac{x_i}{\sqrt{\frac{x_{i+1}^2 + \dots + x_n^2 + S}{m + n - i}}}, \quad i = 1, \dots, n.$$

Докажите, что $t_1, \dots, t_n$ — независимые случайные величины, имеющие t -распределения с $m + n - i$ степенями свободы соответственно. (Указание. Преобразовать совместную плотность вероятностей значений $x_1, \dots, x_n$ и S к плотности вероятностей значений $t_1, \dots, t_n$ и $u = x_1^2 + \dots + x_n^2 + S$.)

15. (Разд. 3.2.2) Покажите, что соотношения (19) и (20) эквивалентны.

16. (Разд. 3.2.2) Покажите, что a_{ii} , определенное соотношением (22), равно

$$\frac{(i!)^4 T (T^2 - 1) (T^2 - 4) (T^2 - 9) \dots (T^2 - i^2)}{(2i)! (2i + 1)!}.$$

17. (Разд. 3.2.2) Представьте подобранный тренд из примера 3.1 в виде полинома от t (где t — год, предшествующий 1918-му).

18. (Разд. 3.2.2)

ИНДЕКС ДУУ-ДЖОНСА СРЕДНИХ ЦЕН НА
АКЦИИ РЯДА ПРОМЫШЛЕННЫХ
КОМПАНИЙ

Год	Цена	Год	Цена	Год	Цена
1897	45.5	1903	55.5	1909	92.8
1898	52.8	1904	55.1	1910	84.3
1899	71.6	1905	80.3	1911	82.4
1900	61.4	1906	93.9	1912	88.7
1901	69.9	1907	74.9	1913	79.2
1902	65.4	1908	75.6		

Используя эти данные, сделайте следующее:

(а) Изобразите данные графически.

(б) Принимая модель наименьших квадратов с $\xi_{it} = \gamma_0 + \gamma_1 \Phi_{1T}(t) + \gamma_2 \Phi_{2T}(t)$, где $\Phi_{iT}(t)$ — ортогональные полиномы, найдите оценки для γ_0, γ_1 и γ_2 .

(с) Найдите доверительный интервал для γ_2 с коэффициентом доверия $1 - \varepsilon = 0.95$.

(д) Проверьте гипотезу $H: \gamma_2 = 0$ с уровнем значимости 0.05.

(е) В предположении, что $\xi_{it} = \beta_0 + \beta_1 t + \beta_2 t^2$, найдите оценки наименьших квадратов для β_0, β_1 и β_2 . (Указание. Эти оценки можно найти из (б).)

(ф) Вычислите и изобразите графически остатки от оценки линейного тренда $\beta_0 + \beta_1 \Phi_{1T}(t)$.

19. (Разд. 3.2.2) Пусть регрессия полиномиальна

$$\xi_{it} = \beta_0 + \beta_1 t + \dots + \beta_q t^q$$

и требуется определить степень полинома по методу, изложенному в разд. 3.2.2. Покажите, как можно эффективно произвести вычисления с использованием прямого решения нормальных уравнений.

20. (Разд. 3.3.1) Пусть в (1) $a = 1$ и $k = 3$.

(a) Постройте таблицу значений и график $f(t)$.

(b) Проверьте, что производные df/dt обеих составляющих функций в точке $t = 2k = 6$ совпадают.

(c) Найдите параболу, которая в интервале $t = 0, 1, \dots, 12$ приближает $f(t)$ наилучшим образом в том смысле, что минимизируется сумма квадратов отклонений.

(d) Постройте таблицу значений и график этой параболы.

(e) Постройте таблицу отклонений.

(f) Какая парабола лучше выравнивает наблюдения в точках $t = 5, 6, 7$ и каковы ее значения в этих точках?

(g) Каково значение первой из составляющих парабол в точке $t = 12$?

(h) Какова парабола, наилучшим образом выравнивающая значения в точках $t = 4, 5, 6, 7, 8$, и каковы ее значения в этих точках?

(i) Каково наилучшее кубическое приближение в интервале $t = 0, 1, \dots, 12$?

(j) Постройте таблицу значений и график этого кубического приближения.

21. (Разд. 3.3.1) Покажите, что если $\sum_{s=-m}^m c_s = 1$, то

$$\sum_{s=-m}^m c_s^2 \geq \frac{1}{2m+1}.$$

Покажите, что если к тому же $c_s \geq 0$, то

$$\sum_{s=-m}^m c_s^2 \leq 1,$$

и что знак равенства имеет место в том и только в том случае, когда все c_s , кроме одного, равны нулю.

22. (Разд. 3.3.1) Найдите a_0 при $k = 2$. (Указание. Найти полиномы $\Phi_{i,2m+1}^*(s)$, $i = 0, 2, 4$, ортогональные на множестве $s = -m, \dots, m$, и представить a_0 в виде их соответствующей линейной комбинации.)

23. (Разд. 3.3.1) Произведите сглаживание ряда из примера 3.1 с помощью процедуры, основывающейся на $k = 1$ и $m = 2$. Сравните результат с соответствующим кубическим приближением.

24. (Разд. 3.3.1)

(a) Произведите сглаживание ряда в упр. 18 с помощью процедуры, основывающейся на $k = 0$ и $m = 3$.

(b) Произведите сглаживание с помощью процедуры, основывающейся на $k = 1$ и $m = 2$.

(c) Сравните сглаженный ряд с трендом, подобранным в упр. 18.

25. (Разд. 3.3.1) Пусть

$$C = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2^2 & 3^2 & \dots & m^2 \\ 1 & 2^4 & 3^4 & \dots & m^4 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 2^{2k} & 3^{2k} & \dots & m^{2k} \end{bmatrix}.$$

Докажите, что матрица коэффициентов в (13) равна

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} + 2CC'.$$

26. (Разд. 3.3.1) Пусть $D = 2CC'$, где C — матрица, определенная в упр. 25, и пусть D_{00} — алгебраическое дополнение элемента d_{00} матрицы D . Докажите, что ¹⁾

$$b^{00} = \frac{D_{00}}{D_{00} + |D|} = \frac{d^{00}}{1 + d^{00}}.$$

27. (Разд. 3.3.1) Проверьте правильность заполнения таблицы 3.3.

28. (Разд. 3.3.2) Найдите доверительный интервал для α_0 в момент t , используя независимую оценку s^2 для σ^2 , когда отношение s^2/σ^2 имеет χ^2 -распределение с n степенями свободы.

29. (Разд. 3.3.2) Пусть $\Phi_{i,2m+1}^*(s)$ — ортогональные многочлены на множестве $-m, \dots, m$. Найдите $\Phi_{1,5}^*(s)$ и $\Phi_{2,5}^*(s)$.

30. (Разд. 3.3.2) Пусть $\Phi_{i,2m+1}^*(s)$ — ортогональные полиномы на множестве $-m, \dots, m$. Докажите, что $\Phi_{i,2m+1}^*(0) = 0$ для нечетных i . (Указание. Пусть

$\varphi^*(s) = c_0 + c_1 s + \dots + c_{2r} s^{2r} + s^{2r+1}$. Докажите, что $\sum_{s=-m}^m \varphi^*(s) s^{2j} = 0$, $j = 0, 1, \dots, r$, — система $r+1$ однородных уравнений с $r+1$ неизвестными $c_0, c_2, \dots, c_{2r}$, имеющая невырожденную матрицу коэффициентов.)

31. (Разд. 3.3.2) Проверьте соотношение (34). (Указание. Положить $\cos[\lambda(t+s) - \theta] = \Re \exp[i(\lambda s + \lambda t - \theta)]$, где $\Re e^{iy}$ — действительная часть e^{iy} , и использовать формулу суммы геометрической прогрессии $\sum_{s=0}^{m-1} x^s = (1 - x^m)/(1 - x)$, $x \neq 1$ при $x = e^{iy}$.)

32. (Разд. 3.3.2) Докажите, что

$$\lim_{x \rightarrow 0} \frac{\sin ax}{\sin x} = a$$

33. (Разд. 3.3.2) Докажите, что $(\sin y)/y$ является убывающей функцией переменной y ($0 < y < \pi$).

34. (Разд. 3.3.2) Пусть

$$y_t^* = \frac{1}{2m+1} \sum_{s=-m}^m y_{t+s},$$

$$y_t^{**} = \frac{1}{2n+1} \sum_{r=-n}^n y_{t+r}^*.$$

¹⁾ Здесь b^{00} — левый верхний угловой элемент матрицы, обратной к матрице коэффициентов в (13), а d^{00} — левый верхний угловой элемент матрицы D^{-1} . — Прим. перев.

Найдите коэффициенты c_s в представлении

$$y_t^{**} = \sum_{s=-(m+n)}^{m+n} c_s y_{t+s}.$$

35. (Разд. 3.3.3) Покажите, что если $h(t) = \cos(\lambda t - \theta)$, то

$$\frac{\sin(2m+1)\lambda/2 + \sin(2m-1)\lambda/2}{4m \sin \lambda/2} \cos(\lambda t - \theta) = \frac{\sin \lambda m \cos \lambda/2}{2m \sin \lambda/2} \cos(\lambda t - \theta).$$

(Указание. См. упр. 31.)

36. (Разд. 3.3.3) Проверьте соотношение (49).

37. (Разд. 3.4.2) Проверьте соотношение (5).

38. (Разд. 3.4.2) Покажите, что если $\mathcal{O}_1$ и $\mathcal{O}_2$ — линейные операторы, то и $c_1 \mathcal{O}_1 + c_2 \mathcal{O}_2$ является линейным оператором.

39. (Разд. 3.4.2) Покажите, что если $\Delta \varphi_1(t) = \Delta \varphi_2(t)$ для $t = 1, 2, \dots$, где $\Delta \varphi_1(t) = \varphi_1(t+1) - \varphi_1(t)$, то разность $\varphi_1(t) - \varphi_2(t)$ постоянна.

40. (Разд. 3.4.2) Пусть $x^{(0)} = 1$, $x^{(r)} = x(x-1) \dots (x-r+1)$, $r = 1, 2, \dots$. Покажите, что

$$\Delta x^{(r)} = r x^{(r-1)},$$

где

$$\Delta x^{(r)} = (x+1)^{(r)} - x^{(r)}.$$

41. (Разд. 3.4.2) Покажите, что

$$\sum_{t=1}^T t^{(r)} = \frac{(T+1)^{(r+1)}}{r+1}, \quad r = 1, 2, \dots$$

42. (Разд. 3.4.2) Покажите, что для $x^{(r)}$, определенного в упр. 40,

$$x^n = \sum_{r=1}^n S_r^n x^{(r)}, \quad n = 1, 2, \dots,$$

где S_r^n — числа Стирлинга второго рода. Получите следующую таблицу:

Таблица чисел Стирлинга второго рода S_r^n

$n \backslash r$	1	2	3	4	5	6	7	8
1	1							
2	1	1						
3	1	3	1					
4	1	7	6	1				
5	1	15	25	10	1			
6	1	31	90	65	15	1		
7	1	63	301	350	140	21	1	
8	1	127	966	1701	1050	266	28	1

43. (Разд. 3.4.2) Вычислите $\psi_k(T) = \sum_{t=1}^T t^k$, используя результаты упр. 41 и 42, для $k = 1, \dots, 8$.

44. (Разд. 3.4.2) Покажите, что если $\Delta f(t)$ — полином степени $k-1$, то $f(t)$ — полином степени k . (Указание. С помощью (13) показать, что решение существует, а с помощью упр. 39 показать, что не существует никаких решений, отличных от указанного.)

45. (Разд. 3.4.2) Покажите, что $\psi_k(T) = \sum_{k=1}^T t^k$ — полином степени $k+1$.

46. (Разд. 3.4.2) Используя упр. 45, найдите выражения для $\psi_k(T) = \sum_{t=1}^T t^k$, $k = 1, \dots, 8$, приведенные в упр. 7.

47. (Разд. 3.4.3) Докажите, что

$$f(t+2) - \frac{1}{5} \sum_{s=-2}^2 f(t+2+s) = -\frac{1}{5} [\Delta^4 + 5\Delta^3 + 5\Delta^2] f(t).$$

48. (Разд. 3.4.3) Покажите, что разность $y_t - y_t^*$ между наблюдаемым рядом и скользящим средним, основывающимся на подборе полинома степени $q = 2k+1$ по $2m+1$ точкам, можно записать в виде

$$y_{t+k+1} - y_{t+k+1}^* = \Delta^{2(k+1)} \left[\sum_{s=-(m-k-1)}^{m-k-1} \beta_s y_{t+s} \right]$$

и что при этом $\beta_s = \beta_{-s}$, $s = 1, \dots, m-k-1$. Покажите, как можно вычислить коэффициенты β_r по коэффициентам c_r . (Указание. Записать

$$\mathcal{P}^{k+1} \left(1 - \sum_{s=-m}^m c_s \mathcal{P}^s \right) y_t = (\mathcal{P} - 1)^{2(k+1)} \left(\sum_{r=-(m-k-1)}^{m-k-1} \beta_r \mathcal{P}^r y_t \right).$$

49. (Разд. 3.4.3) В обозначениях упр. 48 покажите, что

$$\begin{aligned} \sum_{s=-1}^1 \beta_s y_{t+s} &= (2y_{t-1} + 5y_t + 2y_{t+1})/21 = \\ &= (2\Delta^2 + 9\mathcal{P}) y_{t-1}/21 \end{aligned}$$

для $q = 2k+1 = 3$ и $m = 3$.

50. (Разд. 3.4.3) В обозначениях упр. 48 покажите, что

$$\begin{aligned} \sum_{s=-2}^2 \beta_s y_{t+s} &= \frac{1}{231} (21y_{t-2} + 70y_{t-1} + 115y_t + 70y_{t+1} + 21y_{t+2}) = \\ &= \left(\frac{1}{11} \Delta^4 + \frac{2}{3} \Delta^2 \mathcal{P} + \frac{9}{7} \mathcal{P}^2 \right) y_{t-2} \end{aligned}$$

для $q = 2k+1 = 3$ и $m = 4$.

51. (Разд. 3.4.3) Убедитесь в том, что если оператор вычисления r -й разности аннулирует тренд, то тренд является полиномом степени, не большей $r-1$.

52. (Разд. 3.4.3) Покажите, что если тренд $f(t)$ является полиномом степени q для любых последовательных наборов из $q+2$ точек, то $f(t)$ является полиномом степени q на всем отрезке наблюдений.

53. (Разд. 3.4.4) Покажите, что если тренд является полиномиальным степени q , а остатки независимы и нормально распределены с дисперсией σ^2 , то величина s^2 , определяемая соотношением (10) § 3.2, является несмещенной оценкой для σ^2 с наименьшей дисперсией.

54. (Разд. 3.4.4) Докажите, что если тренд в точности аннулируется вычислением разностей порядка $q + 1$, а остатки независимы и нормально распределены с дисперсиями σ^2 , то величина s^2 , определенная соотношением (10) § 3.2, является несмещенной оценкой для σ^2 с наименьшей дисперсией (Указание. Свести к упр. 53.)

55. (Разд. 3.4.4) Докажите, что

$$a_{r+1-s-k, r+1-s}^{(r)} = (-1)^k \binom{2r}{r+k} - a_{s, s+k}^{(r)}, \quad s = 1, \dots, r-k, \\ k = 1, \dots, r-1, \\ a_{n+1, n+1+k}^{(r)} = \frac{1}{2} (-1)^k \binom{2r}{r+k}, \quad r-k = 2n+1.$$

56. (Разд. 3.4.4) Докажите, что величина (66) имеет в пределе нормальное распределение с нулевым средним и единичной дисперсией. (Указание. Последовательность $\{(\Delta^r u_t)^2\}$ образует стационарный случайный процесс с конечной зависимостью; применить теорему 7.7.5.)

57. (Разд. 3.4.4) Докажите лемму 3.4.5.

58. (Разд. 3.4.4) Покажите, что

$$\lim_{T \rightarrow \infty} \frac{1}{T-q} \sum_{s,t=1}^T a_{st}^{(r)} a_{st}^{(q)} = \binom{2r+2q}{r+q}.$$

(Указание. Показать, что

$$\sum_{k=-p}^p \binom{2r}{r+k} \binom{2q}{q+k} = \binom{2r+2q}{r+q},$$

где $p = \min(r, q)$, сравнивая коэффициенты при x^{r+q} в альтернативных разложениях в ряд $(x-1)^{2r} (1-x)^{2q}$.)

59. (Разд. 3.4.4) Докажите (68).

60. (Разд. 3.4.4) Используя (68) при $\kappa_4 = 0$, найдите c_1 и c_2 , удовлетворяющие условию $c_1 + c_2 = 1$, минимизирующие дисперсию (приближенную) линейной комбинации $c_1 V_1 + c_2 V_2$. Выпишите эту минимальную дисперсию.

61. (Разд. 3.4.4) Покажите, что $\sum_{t=1}^{T-r} (\Delta^r y_t)^2$ равна

$$\sum_{t=r+1}^{T-r} \sum_{j=-r}^r (-1)^j \binom{2r}{r+j} y_t y_{t+j}$$

с точностью до крайних членов.

62. (Разд. 3.4.5) Найдите предельное совместное распределение отношений $\sqrt{T-r-2} (V_{r+1} - V_{r+2})/V_{r+2}$ и $\sqrt{T-r-1} (V_r - V_{r+1})/V_{r+1}$ при $\Delta^r f(t) = 0$, $t = 1, 2, \dots$.

63. (Разд. 3.4.5) Пусть $u_1, \dots, u_T$ — независимые случайные величины с $\mathbb{E}u_t = 0$, $\mathbb{E}u_t^2 = \sigma^2$ и $\mathbb{E}u_t^4 = \kappa_4 + 3\sigma^4$. Пусть $Q_1 = \sum_{s,t=1}^T a_{st}u_su_t$ и $Q_2 = \sum_{s,t=1}^T b_{st}u_su_t$. Покажите, что

$$\text{Var} \left(Q_1 - \frac{\mathbb{E}Q_1}{\mathbb{E}Q_2} Q_2 \right) = \kappa_4 \sum_{t=1}^T \left(a_{tt} - \frac{\text{tr } \mathbf{A}}{\text{tr } \mathbf{B}} b_{tt} \right)^2 + 2\sigma^4 \text{tr} \left(\mathbf{A} - \frac{\text{tr } \mathbf{A}}{\text{tr } \mathbf{B}} \mathbf{B} \right)^2.$$

64. (Разд. 3.4.5) Пусть $\mathbb{E}u' \mathbf{A}_T u = \alpha \sigma^2$ и $\mathbb{E}u' \mathbf{B}_T u = \beta \sigma^2$, $\beta > 0$. Предположим, что $\sqrt{T} (u' \mathbf{A}_T u - \alpha \sigma^2)$ и $\sqrt{T} (u' \mathbf{B}_T u - \beta \sigma^2)$ имеют предельное двумерное нормальное распределение. Покажите, что $\sqrt{T} [u' \mathbf{A}_T u / u' \mathbf{B}_T u - \alpha / \beta]$ имеет предельное нормальное распределение с дисперсией

$$\lim T \left[\kappa_4 \sum_{t=1}^T \left(a_{tt}^{(T)} - \frac{\alpha}{\beta} b_{tt}^{(T)} \right)^2 + 2\sigma^4 \text{tr} \left(\mathbf{A}_T - \frac{\alpha}{\beta} \mathbf{B}_T \right)^2 \right] / (\beta^2 \sigma^4).$$

Покажите, что дисперсия этого предельного распределения не зависит от κ_4 , если $a_{11}^{(T)} / b_{11}^{(T)} = \dots = a_{TT}^{(T)} / b_{TT}^{(T)}$. (Указание. Показать, что

$$\text{Var} \left(u' \mathbf{A}_T u - \frac{\alpha}{\beta} u' \mathbf{B}_T u \right) = \kappa_4 \sum_{t=1}^T \left(a_{tt}^{(T)} - \frac{\alpha}{\beta} b_{tt}^{(T)} \right)^2 + 2\sigma^4 \text{tr} \left(\mathbf{A}_T - \frac{\alpha}{\beta} \mathbf{B}_T \right)^2.$$

65. (Разд. 3.4.5) Убедитесь, что последовательность $\Delta^r y_1, \Delta^q y_{r+2}, \Delta^r y_{r+q+3}, \dots$ является последовательностью независимых случайных величин. Используйте этот факт для построения F -критерия для проверки гипотезы $\Delta^r f(t) = 0$ в предположении нормальности распределения и при условии $\Delta^q f(t) = 0$, $r < q$.

66. (Разд. 3.4.5) По данным табл. 3.5 вычислите $(V_r - V_{r+1}) / V_{r+1}$, $r = 0, 1, \dots, 4$, перейдите к нормированным величинам и решите, будет ли $\Delta^r f(t) = 0$.

67. (Разд. 3.4.5) Найдите Δy_t , $\Delta^2 y_t$ и $\Delta^3 y_t$ по данным упр. 18. Вычислите V_1 , V_2 и V_3 .

68. (§ 3.5) Выпишите уравнения (4) для логистической функции (1) в явном виде.

69. (§ 3.5) С интервалами в 30 лет население Соединенных Штатов составляло (в млн.) соответственно

1820	9.6
1850	23.2
1880	50.2
1910	92.0
1940	131.4

Выводите эти данные с помощью логистической кривой. (Указание. Испытать несколько предположительных троек (a_1, a_2, a_3) на чертеже; затем использовать одну из таких троек в качестве исходной для одной итерации (4).)

ЦИКЛИЧЕСКИЕ ТРЕНДЫ

4.1. ВВЕДЕНИЕ

Наряду с основными долговременными изменениями, во временных рядах часто проявляются более или менее регулярные колебания. Эти изменения значений ряда в сторону увеличения и уменьшения могут быть строго периодическими или близкими к таковым, как, например, изменение времени захода солнца. С другой стороны, значения ряда могут колебаться и нерегулярным образом. Типичный случай последнего — экономические временные ряды, колебания которых отражают цикл деловой активности. В настоящей главе будут рассмотрены модели, в которых предполагается, что наблюдаемые временные ряды являются суммой периодического тренда $f(t)$ и случайной ошибки u_t . Периодичность тренда означает, что он в точности повторяет себя через определенный промежуток времени. Если период равен φ , то

$$(1) \quad f(t + \varphi) = f(t).$$

Такое повторение абсолютно регулярно и периодически. Если функция задана на каком-нибудь интервале длины φ , то она определена тем самым и на всем интервале наблюдений. Поскольку тренд предполагается равным заданной функции времени, то нерегулярные воздействия u_t не влияют на него. Эта модель противоположна другим моделям, которые мы будем изучать позднее и в которых колебания не являются регулярными, а случайные воздействия объединены в последовательность.

В настоящей главе анализ ряда производится с помощью линейных комбинаций функций времени — синусов и косинусов, причем коэффициенты линейных комбинаций рассматриваются как параметры. Позднее, в моделях стационарных случайных процессов, линейные комбинации синусов и косинусов будут рассмотрены

снова, но коэффициенты этих линейных комбинаций будут уже случайными величинами. Некоторые аспекты математической теории и многие вычислительные процедуры будут общими для обоих подходов.

4.2. ПРЕОБРАЗОВАНИЯ И ПРЕДСТАВЛЕНИЯ

4.2.1. Ортогональные периодические функции

Тригонометрические функции $\cos t$ и $\sin t$ являются периодическими с периодом 2π , т. е.

$$(1) \quad \cos(t + 2\pi) = \cos t, \quad \sin(t + 2\pi) = \sin t.$$

Отсюда следует, что

$$(2) \quad \cos(t + 2\pi k) = \cos t, \quad \sin(t + 2\pi k) = \sin t, \\ k = \pm 1, \pm 2, \dots$$

Мы можем линейно преобразовать аргументы, сохранив свойство периодичности. Функции $\cos(\lambda t - \theta)$ и $\sin(\lambda t - \theta)$ периодичны с периодом $2\pi/\lambda$, т. е.

$$(3) \quad \begin{aligned} \cos[\lambda(t + 2\pi/\lambda) - \theta] &= \cos[\lambda t + 2\pi - \theta] = \cos[\lambda t - \theta], \\ \sin[\lambda(t + 2\pi/\lambda) - \theta] &= \sin[\lambda t + 2\pi - \theta] = \sin[\lambda t - \theta]. \end{aligned}$$

Обратная величина $\lambda/(2\pi)$ называется *частотой*. Она равна числу периодов (не обязательно целому), содержащемуся в единичном интервале. Иными словами, именно такое число раз функция повторяет свои значения. Умножение на λ соответствует растяжению или сжатию масштаба времени, а вычитание θ — сдвигу графика косинуса или синуса. Функция $\cos(\lambda t - \theta)$ достигает максимума в точках $\lambda t = \theta + 2\pi k$, $k = 0, \pm 1, \dots$, т. е. при $t = (\theta + 2\pi k)/\lambda$. Угол θ называется *фазой*. Обычно θ выбирается так, чтобы первый максимум достигался в точке $t = \theta/\lambda$. В таком случае $0 \leq \theta < 2\pi$. При $t = 0$ указанные тригонометрические функции равны, соответственно $\cos \theta$ и $-\sin \theta$.

Сдвинутые косинусоида и синусоида являются линейными комбинациями обычной косинусоиды и обычной синусоиды и наоборот. Из тригонометрической формулы $\cos(a - b) = \cos a \cos b + \sin a \sin b$ имеем

$$(4) \quad \begin{aligned} \rho \cos(\lambda t - \theta) &= \rho(\cos \lambda t \cos \theta + \sin \lambda t \sin \theta) = \\ &= \alpha \cos \lambda t + \beta \sin \lambda t, \end{aligned}$$

где

$$(5) \quad \alpha = \rho \cos \theta, \quad \beta = \rho \sin \theta,$$

или, что эквивалентно,

$$(6) \quad \begin{aligned} \rho &= \sqrt{\alpha^2 + \beta^2}, \\ \operatorname{tg} \theta &= \frac{\beta}{\alpha}, \quad \theta = \operatorname{arctg} \frac{\beta}{\alpha}. \end{aligned}$$

Коэффициент ρ , являющийся максимумом функции $\rho \cos(\lambda t - \theta)$, называется *амплитудой* этой функции. Выражение (4) можно записать также в виде $\rho \sin(\lambda t + \varphi)$, где $\operatorname{tg} \varphi = \alpha/\beta$, но обычно предпочитают использовать функцию косинус.

С тригонометрическими функциями довольно удобно работать вследствие того, что они обладают определенными свойствами ортогональности. Мы рассмотрим здесь свойства ортогональности сумм на множестве $1, \dots, T$. Они соответствуют свойству ортогональности полиномов, разбиравшемуся ранее в разд. 3.2.1. Рассмотрим частоты $\lambda/(2\pi) = j/T$, $j = 0, 1, \dots, [T/2]$, где $[T/2] = T/2$ для четных T и $[T/2] = (T-1)/2$ для нечетных T . Период при этом равен $2\pi/\lambda = T/j$. На протяжении всего отрезка наблюдений T укладывается ровно j таких периодов. Функции косинус и синус с такими частотами являются ортогональными. Чтобы показать это, удобно воспользоваться соотношениями

$$(7) \quad \begin{aligned} e^{i\lambda} &= \cos \lambda + i \sin \lambda, \\ \cos \lambda &= (e^{i\lambda} + e^{-i\lambda})/2, \\ \sin \lambda &= (e^{i\lambda} - e^{-i\lambda})/2i. \end{aligned}$$

При этом

$$(8) \quad \begin{aligned} \sum_{t=1}^T \cos \frac{2\pi j}{T} t \left[\cos \frac{2\pi k}{T} t + i \sin \frac{2\pi k}{T} t \right] &= \\ &= \sum_{t=1}^T \frac{1}{2} [e^{i2\pi j t/T} + e^{-i2\pi j t/T}] e^{i2\pi k t/T} = \\ &= \frac{1}{2} e^{i2\pi (k+j)/T} \left[\sum_{t=0}^{T-1} e^{i2\pi (k+j) t/T} \right] + \\ &+ \frac{1}{2} e^{i2\pi (k-j)/T} \left[\sum_{t=0}^{T-1} e^{i2\pi (k-j) t/T} \right] = \end{aligned}$$

$$\begin{aligned}
& \left\{ \frac{1}{2} e^{i2\pi (k+j)/T} \left[\frac{1 - e^{i2\pi (k+j)}}{1 - e^{i2\pi (k+j)/T}} \right] + \frac{1}{2} e^{i2\pi (k-j)/T} \left[\frac{1 - e^{i2\pi (k-j)}}{1 - e^{i2\pi (k-j)/T}} \right], \right. \\
& \qquad \qquad \qquad 0 \leq k \neq j \leq [T/2], \\
& = \left\{ \frac{1}{2} e^{i2\pi (k+j)/T} \left[\frac{1 - e^{i2\pi (k+j)}}{1 - e^{i2\pi (k+j)/T}} \right] + \frac{1}{2} T, \right. \\
& \qquad \qquad \qquad 0 < k = j < T/2, \\
& \qquad \qquad \qquad T, \qquad \qquad \qquad k = j = 0, \quad T/2, \\
& = \left\{ \begin{array}{ll} 0, & 0 \leq k \neq j \leq [T/2], \\ T/2, & 0 < k = j < T/2, \\ T, & k = j = 0, \quad T/2. \end{array} \right.
\end{aligned}$$

Приравнивая действительные и мнимые части в соотношении (8), получаем

$$(9) \quad \sum_{t=1}^T \cos \frac{2\pi j}{T} t \cos \frac{2\pi k}{T} t = \begin{cases} 0, & 0 \leq k \neq j \leq [T/2], \\ T/2, & 0 < k = j < T/2, \\ T, & k = j = 0, \quad T/2, \end{cases}$$

$$(10) \quad \sum_{t=1}^T \cos \frac{2\pi j}{T} t \sin \frac{2\pi k}{T} t = 0, \quad k, j = 0, 1, \dots, [T/2].$$

Подобным же образом можно показать, что

$$(11) \quad \sum_{t=1}^T \sin \frac{2\pi j}{T} t \sin \frac{2\pi k}{T} t = \begin{cases} 0, & 0 \leq k \neq j \leq [T/2], \\ T/2, & 0 < k = j < T/2, \\ 0, & k = j = 0, \quad T/2. \end{cases}$$

Кроме того, полагая $j = 0$ в (9) и (10), получаем

$$(12) \quad \sum_{t=1}^T \cos \frac{2\pi k}{T} t = 0, \quad k = 1, \dots, [T/2],$$

$$(13) \quad \sum_{t=1}^T \sin \frac{2\pi k}{T} t = 0, \quad k = 0, 1, \dots, [T/2].$$

Если T — нечетное, то $\cos 2\pi 0t/T = 1$. При этом $\cos 2\pi kt/T$ и $\sin 2\pi kt/T$, $k = 1, \dots, (T-1)/2$, образуют множество из T последовательностей по T чисел, любые две из которых ортогональны. Если T — четное, то таким множеством является совокупность функций 1 , $\cos 2\pi kt/T$, $\sin 2\pi kt/T$, $k = 1, \dots, T/2 - 1$ и $\cos 2\pi (T/2) t/T = \cos \pi t = (-1)^t$. Сумма квадратов членов каждой последовательности равна $T/2$, за исключением последовательностей 1 и $(-1)^t$, у которых она равна T .

4.2.2. Представление конечной последовательности с помощью ее частот

Рассмотрим последовательность T чисел $y_1, \dots, y_T$ (которая не обязательно является наблюдаемым временным рядом). Эту последовательность можно рассматривать как совокупность координат некоторой точки в пространстве T измерений. Указанная точка может быть отнесена и к другим координатным системам. Иногда координаты относительно другой системы будут более осмысленными. Мы используем надлежащим образом нормированные ортогональные тригонометрические функции для того, чтобы определить ортогональную матрицу, посредством которой будут преобразованы координаты $y_1, \dots, y_T$. Если T четное, то матрицу M размера $T \times T$ определим равенством

$$(14) \quad M = \sqrt{\frac{2}{T}} \times \begin{bmatrix} \frac{1}{\sqrt{2}} \cos \frac{2\pi}{T} \sin \frac{2\pi}{T} \cos \frac{4\pi}{T} \dots \sin \frac{2\pi(T/2-1)}{T} - \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \cos \frac{4\pi}{T} \sin \frac{4\pi}{T} \cos \frac{8\pi}{T} \dots \sin \frac{4\pi(T/2-1)}{T} \frac{1}{\sqrt{2}} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \frac{1}{\sqrt{2}} \quad 1 \quad 0 \quad 1 \quad \dots \quad 0 \quad \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Из соотношений ортогональности (9) — (13) выводим, что

$$(15) \quad M'M = I.$$

Таким образом, матрица M ортогональна. Положим, далее,

$$(16) \quad y = (y_1, \dots, y_T)'$$

и

$$(17) \quad x = (x_1, \dots, x_T)' = M'y,$$

т. е.

$$(18) \quad \begin{aligned} x_1 &= \frac{1}{\sqrt{T}} \sum_{t=1}^T y_t, \\ x_{2k} &= \sqrt{\frac{2}{T}} \sum_{t=1}^T y_t \cos \frac{2\pi k}{T} t, \quad k = 1, \dots, T/2 - 1, \\ x_{2k+1} &= \sqrt{\frac{2}{T}} \sum_{t=1}^T y_t \sin \frac{2\pi k}{T} t, \quad k = 1, \dots, T/2 - 1, \\ x_T &= \frac{1}{\sqrt{T}} \sum_{t=1}^T y_t (-1)^t. \end{aligned}$$

Тогда, поскольку $\mathbf{M}\mathbf{M}' = \mathbf{I}$, имеем

$$(19) \quad \mathbf{y} = \mathbf{M}\mathbf{x},$$

$$(20) \quad y_t = \sqrt{\frac{2}{T}} \left(\frac{1}{\sqrt{2}} x_1 + x_2 \cos \frac{2\pi}{T} t + \dots + x_T \frac{(-1)^t}{\sqrt{2}} \right),$$

$$t = 1, \dots, T.$$

Разложение (20) называют иногда представлением Фурье последовательности $y_1, \dots, y_T$ с коэффициентами $x_1, \dots, x_T$.

Если T нечетное, то матрицу $\mathbf{M}$ определим равенством

$$(21) \quad \mathbf{M} = \sqrt{\frac{2}{T}} \times$$

$$\times \begin{bmatrix} \frac{1}{\sqrt{2}} \cos \frac{2\pi}{T} \sin \frac{2\pi}{T} \cos \frac{4\pi}{T} \dots \sin \frac{2\pi(T-1)/2}{T} \\ \frac{1}{\sqrt{2}} \cos \frac{4\pi}{T} \sin \frac{4\pi}{T} \cos \frac{8\pi}{T} \dots \sin \frac{4\pi(T-1)/2}{T} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \frac{1}{\sqrt{2}} \quad 1 \quad 0 \quad 1 \quad \dots \quad 0 \end{bmatrix}.$$

Соответствующее представление Фурье имеет вид

$$(22) \quad y_t = \sqrt{\frac{2}{T}} \left(\frac{1}{\sqrt{2}} x_1 + x_2 \cos \frac{2\pi}{T} t + \dots \right.$$

$$\left. \dots + x_T \sin \frac{2\pi(T-1)/2}{T} t \right), \quad t = 1, \dots, T.$$

(В некотором смысле проще оперировать с унитарной матрицей $\mathbf{N} = (n_{st}) = (e^{i2\pi st/T} / \sqrt{T})$, удовлетворяющей соотношению $\mathbf{N}\mathbf{N}' = \mathbf{I}$, в котором $\bar{N} = (\bar{n}_{st})$ и $\bar{n}_{st}$ — число, сопряженное с n_{st} . Отметим, что $\mathbf{N} = \mathbf{N}'$.)

4.2.3. Представление периодической последовательности в случае, когда период — целое число

Предположим, что числовая последовательность $y_1, \dots, y_T$ имеет период n , где n — целое число, т. е. что

$$(23) \quad y_{t+n} = y_t, \quad t = 1, \dots, T-n.$$

Покажем, как эту последовательность можно представить с помощью n тригонометрических функций. Положим

$$(24) \quad y^* = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Тогда

$$(25) \quad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{hn} \end{pmatrix} = \begin{pmatrix} y^* \\ y^* \\ \vdots \\ y^* \end{pmatrix},$$

где y^* в правом столбце повторяется h раз и $hn \leq T < (h+1)n$. (Допускается, что может быть $T - hn$ членов, составляющих лишь часть цикла.) Определим теперь матрицу M^* размера $n \times n$ таким же образом, как и в разд. 4.2.2, так что

$$(26) \quad y^* = M^* x^*, \quad x^* = M^{*'} y^*,$$

т. е.

$$(27) \quad \begin{aligned} x_1^* &= \frac{1}{\sqrt{n}} \sum_{t=1}^n y_t, \\ x_{2k}^* &= \sqrt{\frac{2}{n}} \sum_{t=1}^n y_t \cos \frac{2\pi k}{n} t, \quad k = 1, \dots, [(n-1)/2], \\ x_{2k+1}^* &= \sqrt{\frac{2}{n}} \sum_{t=1}^n y_t \sin \frac{2\pi k}{n} t, \quad k = 1, \dots, [(n-1)/2], \\ x_n^* &= \frac{1}{\sqrt{n}} \sum_{t=1}^n y_t (-1)^t, \quad n \text{ четное.} \end{aligned}$$

Тогда для $t = 1, \dots, n$ (четного) имеем

$$(28) \quad y_t = \sqrt{\frac{2}{n}} \left(\frac{1}{\sqrt{2}} x_1^* + x_2^* \cos \frac{2\pi}{n} t + \dots + x_n^* \frac{(-1)^t}{\sqrt{2}} \right).$$

Отметим, что для $t \leq n$

$$(29) \quad \begin{aligned} y_{t+n} &= y_t = \sqrt{\frac{2}{n}} \left(\frac{1}{\sqrt{2}} x_1^* + x_2^* \cos \frac{2\pi}{n} t + \dots + x_n^* \frac{(-1)^t}{\sqrt{2}} \right) = \\ &= \sqrt{\frac{2}{n}} \left(\frac{1}{\sqrt{2}} x_1^* + x_2^* \cos \frac{2\pi}{n} (t+n) + \dots \right. \\ &\quad \left. \dots + x_n^* \frac{(-1)^{t+n}}{\sqrt{2}} \right) \end{aligned}$$

в силу того, что $\cos 2\pi k(t+n)/n = \cos 2\pi kt/n$ и $\sin 2\pi k(t+n)/n = \sin 2\pi kt/n$. Подобным же образом можно показать, что (28) выполняется вообще для всех $t = 1, \dots, T$. Если n нечетное, то слагаемое, содержащее $(-1)^t$, опускается.

Мы показали, как произвольную периодическую функцию с целым периодом можно представить в виде линейной комбинации n тригонометрических функций. Следует отметить, что этот результат применим к любой периодической последовательности. Однако если данная периодическая последовательность является последовательностью значений функции, состоящей из синусоидальной и косинусоидальной компонент, и имеет период n (т. е. n является кратным минимальному периоду), то можно обойтись только одной или двумя тригонометрическими последовательностями (косинусов и синусов).

Тригонометрические функции с периодом n вовсе не обязаны быть ортогональными на множестве $1, \dots, T$, если только T не является кратным n , т. е. $T = hn$. В последнем случае

$$(30) \quad \begin{aligned} \cos \frac{2\pi j}{n} t &= \cos \frac{2\pi (jh)}{T} t, \\ \sin \frac{2\pi j}{n} t &= \sin \frac{2\pi (jh)}{T} t. \end{aligned}$$

Тригонометрические функции, входящие в (28), составляют лишь часть тригонометрических функций, входящих в (20) или (22). Поэтому некоторые (или многие) коэффициенты соответствующего представления Фурье равны нулю. Однако полное представление является все же удобным по той причине, что некоторые периодические последовательности являются его простыми частными случаями.

В § 3.3 мы рассмотрели модель, в которой тренд содержал периодическую компоненту $g(t)$. Последняя интерпретировалась как сезонное изменение ряда. Для полугодовых данных период равен 2, для ежеквартальных 4, а для ежемесячных 12. Такое сезонное изменение $g(t)$ можно выразить в виде линейной комбинации n тригонометрических членов.

Периодическая последовательность $y_1, \dots, y_T$ с периодом n и $T = hn$ может быть записана также в виде матрицы

$$(31) \quad \begin{pmatrix} y_1 & y_2 & \dots & y_n \\ y_{n+1} & y_{n+2} & \dots & y_{2n} \\ \vdots & \vdots & & \vdots \\ y_{(h-1)n+1} & y_{(h-1)n+2} & \dots & y_T \end{pmatrix},$$

строки которой совпадают.

4.2.4. Представление периодической функции

Рассмотрим теперь периодическую функцию $f(t)$, определенную для всех действительных t . Если ее период равен φ , то

$$(32) \quad f(t) = f(t + \varphi) = f(t + 2\varphi) = \dots$$

Можно ожидать, что такую функцию удастся выразить с помощью тригонометрических функций, имеющих тот же период φ , а именно с помощью функций $1, \cos 2\pi t/\varphi, \sin 2\pi t/\varphi, \cos 4\pi t/\varphi, \sin 4\pi t/\varphi, \dots$. Рассмотрим бесконечный ряд, состоящий из таких функций:

$$(33) \quad \alpha_0 + \left(\alpha_1 \cos \frac{2\pi}{\varphi} t + \beta_1 \sin \frac{2\pi}{\varphi} t \right) + \\ + \left(\alpha_2 \cos \frac{4\pi}{\varphi} t + \beta_2 \sin \frac{4\pi}{\varphi} t \right) + \dots$$

Если этот ряд сходится к $f(t)$ при некотором значении t , то он будет сходиться к $f(t)$ и при $t + \varphi$, поскольку

$$(34) \quad \cos \frac{2\pi k}{\varphi} (t + \varphi) = \cos \frac{2\pi k}{\varphi} t, \\ \sin \frac{2\pi k}{\varphi} (t + \varphi) = \sin \frac{2\pi k}{\varphi} t$$

и сумма ряда периодична с периодом φ . Тригонометрические функции удовлетворяют следующим соотношениям нормировки и ортогональности:

$$(35) \quad \int_0^{\varphi} \cos^2 \frac{2\pi j}{\varphi} t dt = \frac{\varphi}{2}, \quad j \neq 0,$$

$$(36) \quad \int_0^{\varphi} \sin^2 \frac{2\pi j}{\varphi} t dt = \frac{\varphi}{2}, \quad j \neq 0,$$

$$(37) \quad \int_0^{\varphi} \cos \frac{2\pi j}{\varphi} t \cos \frac{2\pi k}{\varphi} t dt = 0, \quad j \neq k,$$

$$(38) \quad \int_0^{\varphi} \sin \frac{2\pi j}{\varphi} t \cos \frac{2\pi k}{\varphi} t dt = 0, \quad j \neq k,$$

$$(39) \quad \int_0^{\varphi} \cos \frac{2\pi j}{\varphi} t \sin \frac{2\pi k}{\varphi} t dt = 0, \quad \text{для всех } j, k.$$

В этом можно убедиться, используя тригонометрические тождества. (См. упр. 1, 2, 8 и 9.) Например,

$$(40) \quad \int_0^{\varphi} \cos^2 \frac{2\pi j}{\varphi} t dt = \int_0^{\varphi} \left(\frac{1}{2} + \frac{1}{2} \cos \frac{4\pi j}{\varphi} t \right) dt = \\ = \frac{1}{2} \varphi + \frac{\varphi}{8\pi j} \int_0^{4\pi j} \cos s ds = \frac{1}{2} \varphi, \quad j \neq 0.$$

Если бесконечный ряд (33) сходится к $f(t)$ и если возможно его почленное интегрирование, то

$$(41) \quad \int_0^{\varphi} f(t) \cos \frac{2\pi k}{\varphi} t dt = \\ = \alpha_0 \int_0^{\varphi} \cos \frac{2\pi k}{\varphi} t dt + \int_0^{\varphi} \sum_{j=1}^{\infty} \left(\alpha_j \cos \frac{2\pi j}{\varphi} t + \beta_j \sin \frac{2\pi j}{\varphi} t \right) \cos \frac{2\pi k}{\varphi} t dt = \\ = \frac{1}{2} \varphi \alpha_k, \quad k \neq 0.$$

Отсюда определяется коэффициент α_k :

$$(42) \quad \alpha_k = \frac{2}{\varphi} \int_0^{\varphi} f(t) \cos \frac{2\pi k}{\varphi} t dt, \quad k \neq 0.$$

Подобным же образом

$$(43) \quad \beta_k = \frac{2}{\varphi} \int_0^{\varphi} f(t) \sin \frac{2\pi k}{\varphi} t dt, \quad k \neq 0,$$

$$(44) \quad \alpha_0 = \frac{1}{\varphi} \int_0^{\varphi} f(t) dt.$$

Если коэффициенты ряда (33) выбираются согласно (42), (43) и (44), то говорят, что этот ряд *представляет* $f(t)$. Существует несколько теорем о сходимости подобных рядов при различных условиях. Одна из них утверждает, что если $f(t)$ — функция ограниченной вариации на замкнутом интервале $[0, \varphi]$, то ряд (33) сходится к $f(t)$ в каждой точке непрерывности последней. [См., например, Уиттекер и Ватсон (1943, стр. 174—179).] Мы не ставим своей целью изложение анализа рядов Фурье, а просто хотим показать, как периодическую функцию можно представить с помощью тригонометрических выражений.

В том случае, когда функция $f(t)$ используется для описания тренда, существенны только ее значения при $t = 1, \dots, T$. Как было уже отмечено ранее, если функция $f(t)$ имеет период φ и если φ равно целому числу n , то $f(t)$ принимает только n значений, а именно $f(1), f(2), \dots, f(n)$. (Если отношение T/n не является целым, то каждое из этих n значений не может наблюдаться одно и то же число раз.) В таком случае эту функцию можно представить при $t = 1, \dots, T$ линейной комбинацией n тригонометрических функций. Однако если φ не является целым числом, то в представление $f(t)$ в T точках пришлось бы включить возможно до T членов. Важность приведенных соображений состоит в том, что даже если φ и не является целым числом, то $f(t)$ все-таки может быть приближена относительно небольшим числом тригонометрических членов.

4.2.5. Представление произвольной функции

Обратимся теперь к представлению на всей действительной оси функции с действительными значениями. Частным случаем, хорошо известным статистикам, является связь плотности вероятностей $p(x)$ с ее характеристической функцией $\varphi(t)$, определяемой соотношением

$$(45) \quad \varphi(t) = \int_{-\infty}^{\infty} e^{itx} p(x) dx = \int_{-\infty}^{\infty} [\cos tx + i \sin tx] p(x) dx = \\ = \varphi_1(t) + i\varphi_2(t),$$

в котором $\varphi_1(t)$ и $\varphi_2(t)$ — действительные функции. Из этого определения видно, что $\varphi_1(t) = \varphi_1(-t)$ четная, а $\varphi_2(t) = -\varphi_2(-t)$ нечетная функции. Обратное преобразование имеет вид

$$(46) \quad p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt = \\ = \frac{1}{2\pi} \int_{-\infty}^{\infty} [\cos(-tx) + i \sin(-tx)] [\varphi_1(t) + i\varphi_2(t)] dt = \\ = \frac{1}{2\pi} \int_{-\infty}^{\infty} [\cos tx \varphi_1(t) + \sin tx \varphi_2(t)] dt.$$

Кратко сформулируем результаты этого параграфа. Произвольную конечную последовательность с помощью ортогональной матрицы $\mathbf{M}$ можно преобразовать в конечное множество коэффициентов Фурье. Произвольную периодическую функцию можно предста-

вить конечным отрезком ряда Фурье, коэффициенты которого (при определенных условиях) являются тригонометрическими интегралами. Обратно: (при определенных условиях) ряд Фурье определяет некоторую периодическую функцию. Наконец, произвольную (непериодическую, интегрируемую) функцию можно выразить в виде интеграла Фурье. При этом подынтегральное выражение может быть получено как обратный интеграл Фурье.

4.3. СТАТИСТИЧЕСКИЕ ВЫВОДЫ ДЛЯ СЛУЧАЯ, КОГДА ПЕРИОДЫ ТРЕНДА ЯВЛЯЮТСЯ ДЕЛИТЕЛЯМИ ДЛИНЫ РЯДА

4.3.1. Оценки наименьших квадратов для коэффициентов и дисперсии

Рассмотрим теперь модель $y_t = f(t) + u_t$, $t = 1, \dots, T$, где $\mathbb{E}u_t = 0$, $\mathbb{E}u_t^2 = \sigma^2$, $\mathbb{E}u_t u_s = 0$, $t \neq s$, а $f(t)$ — периодическая функция с известными периодами, нацело делящими T . Таким образом, периоды $f(t)$ задаются числами T/k_j , $j = 1, \dots, q$, где $(k_1, \dots, k_q)$ — подмножество последовательности целых чисел $1, \dots, (T-1)/2$, если T нечетное, или $1, \dots, T/2-1$, если T четное. В последнем случае можно, кроме того рассмотреть и период 2. Отсюда следует, что для $f(t)$ возможно представление

$$(1) \quad f(t) = \alpha_0 + \sum_{j=1}^q \left[\alpha(k_j) \cos \frac{2\pi k_j}{T} t + \beta(k_j) \sin \frac{2\pi k_j}{T} t \right],$$

где слагаемое с периодом 2 не включено. Если же T четное, то с учетом слагаемого с периодом 2 $f(t)$ можно представить в виде

$$(2) \quad f(t) = \alpha_0 + \sum_{j=1}^q \left[\alpha(k_j) \cos \frac{2\pi k_j}{T} t + \beta(k_j) \sin \frac{2\pi k_j}{T} t \right] + \alpha_{T/2} (-1)^t.$$

Входящие сюда тригонометрические функции образуют некоторое подмножество функций, составляющих столбцы матрицы M . Вполне вероятно, что в (2) придется включить все T таких функций, но обычно их требуется меньше. Например, если ежемесячные данные накапливаются в течение h лет, то $T = 12h$ и для получения представления сезонного изменения следует взять константу и 11 членов с наименьшими периодами

$$(3) \quad \begin{aligned} \frac{T}{T/2} &= \frac{12}{6} = 2, & \frac{T}{5T/12} &= \frac{12}{5}, & \frac{T}{T/3} &= \frac{12}{4} = 3, \\ \frac{T}{T/4} &= \frac{12}{3} = 4, & \frac{T}{T/6} &= \frac{12}{2} = 6, & \frac{T}{T/12} &= 12. \end{aligned}$$

При этом соответствующие частоты равны $1/2 = 6/12$, $5/12$, $1/3 = 4/12$, $1/4 = 3/12$, $1/6 = 2/12$ и $1/12$. Если период равен 2, то в сумме (2) будет только одно слагаемое, отличное от константы. Это слагаемое вместе с константой полностью представляют цикл периода 2 (т. е. поочередное изменение между двумя фиксированными значениями). Каждому из следующих по величине периодов соответствуют два слагаемых, косинус и синус. Пара таких слагаемых представляет косинусоиду, как правило сдвинутую. В связи с этим можно переписать (1) и (2) соответственно в виде

$$(4) \quad f(t) = \alpha_0 + \sum_{j=1}^q \rho(k_j) \cos \left[\frac{2\pi k_j}{T} t - \theta(k_j) \right],$$

$$(5) \quad f(t) = \alpha_0 + \sum_{j=1}^q \rho(k_j) \cos \left[\frac{2\pi k_j}{T} t - \theta(k_j) \right] + \alpha_{T/2} (-1)^t,$$

где

$$(6) \quad \rho(k_j) = \sqrt{\alpha^2(k_j) + \beta^2(k_j)}, \quad \theta(k_j) = \arctg \frac{\beta(k_j)}{\alpha(k_j)},$$

$$(7) \quad \alpha(k_j) = \rho(k_j) \cos \theta(k_j), \quad \beta(k_j) = \rho(k_j) \sin \theta(k_j).$$

Рассмотрим теперь задачу оценивания параметров α_0 , $\alpha_{T/2}$, $\alpha(k_j)$ и $\beta(k_j)$, $j = 1, \dots, q$, методом наименьших квадратов. Пусть a_0 , $a_{T/2}$, $a(k_j)$ и $b(k_j)$ являются оценками наименьших квадратов для α_0 , $\alpha_{T/2}$, $\alpha(k_j)$ и $\beta(k_j)$ соответственно. Тогда нормальные уравнения для этих оценок имеют вид

$$(8) \quad \begin{bmatrix} T & 0 & 0 & \dots & 0 \\ 0 & T/2 & 0 & \dots & 0 \\ 0 & 0 & T/2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & T/2 \end{bmatrix} \begin{bmatrix} a_0 \\ a(k_1) \\ b(k_1) \\ \vdots \\ b(k_q) \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^T y_t \\ \sum_{t=1}^T y_t \cos \frac{2\pi k_1}{T} t \\ \sum_{t=1}^T y_t \sin \frac{2\pi k_1}{T} t \\ \vdots \\ \sum_{t=1}^T y_t \sin \frac{2\pi k_q}{T} t \end{bmatrix},$$

если $f(t)$ представляется в виде (1), или

$$(9) \quad \begin{bmatrix} T & 0 & 0 & \dots & 0 & 0 \\ 0 & T/2 & 0 & \dots & 0 & 0 \\ 0 & 0 & T/2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & T/2 & 0 \\ 0 & 0 & 0 & \dots & 0 & T \end{bmatrix} \begin{bmatrix} a_0 \\ a(k_1) \\ b(k_1) \\ \vdots \\ b(k_q) \\ a_{T/2} \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^T y_t \\ \sum_{t=1}^T y_t \cos \frac{2\pi k_1}{T} t \\ \sum_{t=1}^T y_t \sin \frac{2\pi k_1}{T} t \\ \vdots \\ \sum_{t=1}^T y_t \sin \frac{2\pi k_q}{T} t \\ \sum_{t=1}^T y_t (-1)^t \end{bmatrix},$$

если $f(t)$ представляется в виде (2). Вследствие ортогональности тригонометрических функций матрица коэффициентов при оценках диагональна. Поэтому решениями нормальных уравнений будут

$$(10) \quad a_0 = \frac{1}{T} \sum_{t=1}^T y_t = \bar{y},$$

$$(11) \quad a(k_j) = \frac{2}{T} \sum_{t=1}^T y_t \cos \frac{2\pi k_j}{T} t, \quad j = 1, \dots, q,$$

$$(12) \quad b(k_j) = \frac{2}{T} \sum_{t=1}^T y_t \sin \frac{2\pi k_j}{T} t, \quad j = 1, \dots, q,$$

и в случае (9)

$$(13) \quad a_{T/2} = \frac{1}{T} \sum_{t=1}^T y_t (-1)^t.$$

Интересно отметить, что выражения (10) — (13) можно получить также другим способом, подставляя в (1) или (2) вместо коэффициентов α и β их оценки a и b , а вместо $f(t)$ значения y_t . Применение обратной формулы Фурье дает выражения указанных коэффициентов через величины y_t .

Оценка наименьших квадратов для σ^2 дается формулой

$$(14) \quad s^2 = \frac{\sum_{t=1}^T y_t^2 - T\bar{y}^2 - \frac{T}{2} \sum_{j=1}^q [a^2(k_j) + b^2(k_j)]}{T - (2q + 1)},$$

если $f(t)$ представляется соотношением (1), или формулой

$$(15) \quad s^2 = \frac{\sum_{i=1}^T y_i^2 - T(\bar{y}^2 + a_{T/2}^2) - \frac{T}{2} \sum_{j=1}^q [a^2(k_j) + b^2(k_j)]}{T - (2q + 2)},$$

если $f(t)$ представляется соотношением (2).

Оценками параметров $\rho(k_j)$ и $\theta(k_j)$ являются соответственно

$$(16) \quad R(k_j) = \sqrt{a^2(k_j) + b^2(k_j)}, \quad \hat{\theta}(k_j) = \arctg \frac{b(k_j)}{a(k_j)},$$

$j = 1, \dots, q.$

Эти оценки можно получить непосредственно, минимизируя $\sum_{i=1}^T [y_i - f(t)]^2$ по параметрам $\rho(k_j)$ и $\theta(k_j)$, $j = 1, \dots, q$, и α_0 , когда $f(t)$ представляется в виде (4), или по тем же параметрам и по параметру $\alpha_{T/2}$, если $f(t)$ представляется соотношением (5).

Пример 4. 1. Рассмотрим данные [опубликованные департаментом сельского хозяйства США (1939, стр. 390)] табл. 4.1 о поступлении масла (в млн. фунтов) на пяти рынках (Бостон, Чикаго, Сан-Франциско, Милуоки и Сент-Луис).

Таблица 4.1
ПОСТУПЛЕНИЕ МАСЛА НА ПЯТИ РЫНКАХ

Месяц	Год			Общее количество m_t	Среднее
	1935	1936	1937		
Январь	48.9	48.3	42.4	139.6	46.5333
Февраль	43.4	47.1	41.4	131.9	43.9667
Март	43.8	52.4	49.0	145.2	48.4000
Апрель	50.8	55.3	50.8	156.9	52.3000
Май	67.6	64.7	65.8	198.1	66.0333
Июнь	83.7	79.5	85.9	249.1	83.0333
Июль	82.7	62.6	70.6	215.9	71.9667
Август	60.8	51.3	55.8	167.9	55.9667
Сентябрь	55.4	51.0	49.1	155.5	51.8333
Октябрь	48.4	54.0	45.7	148.1	49.3667
Ноябрь	37.7	45.2	43.8	126.7	42.2333
Декабрь	41.0	44.9	46.7	132.6	44.2000
Итого	664.2	656.3	647.0	1967.5	655.8333
Среднее	55.350000	54.691667	53.916667	163.958333	54.652778

Мы хотим оценить константу α_0 , коэффициент $\alpha_{T/2} = \alpha_{18}$, соответствующий периоду 2 (частоте $1/2 = 6/12$), и пары коэффициентов, соответствующих (наименьшим) периодам 12/5, 3, 4, 6 и 12 (частотам $5/12$, $1/3 = 4/12$, $1/4 = 3/12$, $1/6 = 2/12$ и $1/12$ соответственно). Пусть

$$(17) \quad m_t = \sum_{i=0}^2 y_{t+12i}$$

— сумма значений наблюдений для t -го месяца за три года. В силу периодичности тригонометрических функций имеем:

$$\begin{aligned} a_0 = \bar{y} &= \frac{1}{36} \sum_{t=1}^{36} y_t = \frac{1}{36} \sum_{t=1}^{12} m_t, \\ a_{3k} &= \frac{1}{18} \sum_{t=1}^{36} y_t \cos \frac{2\pi 3k}{36} t = \frac{1}{18} \sum_{t=1}^{12} m_t \cos \frac{2\pi k}{12} t, \\ (18) \quad b_{3k} &= \frac{1}{18} \sum_{t=1}^{36} y_t \sin \frac{2\pi 3k}{36} t = \frac{1}{18} \sum_{t=1}^{12} m_t \sin \frac{2\pi k}{12} t, \\ a_{18} &= \frac{1}{36} \sum_{t=1}^{36} y_t (-1)^t = \frac{1}{36} \sum_{t=1}^{12} m_t (-1)^t. \end{aligned}$$

$k = 1, 2, 3, 4, 5,$

$k = 1, 2, 3, 4, 5,$

Таблицы этих тригонометрических функций имеются у Кендалла (1946b). Таблица 4.1, расположенная по периодам и включающая суммарные величины, называется таблицей Бьюис-Баллота [Бьюис-Баллот (1847)].

Таблица 4.2
КОЭФФИЦИЕНТЫ ЦИКЛИЧЕСКИХ ТРЕНДОВ

k	0	3	6	9	12	15	18
a_k	54.65	-14.82	6.60	-3.98	2.21	-0.61	0.15
b_k		-2.02	1.23	0.30	1.73	0.60	
Минималь- ный период		12	6	4	3	12/5	2
Частота		1/12	1/6	1/4	1/3	5/12	1/2
R_k		14.96	6.71	3.99	2.81	0.86	0.15
R_k^2		223.7	45.1	15.9	7.88	0.73	0.02

Значения оценок параметров, вычисленные по данным табл. 4.1, приведены в табл. 4.2. Сумма квадратов отклонений равна 474.51, $s^2 = 19.771$. Наибольшая выборочная амплитуда $R_3 = 14.96$ в табл. 4.2 соответствует паре членов с минимальным периодом 12. Это отражает тенденцию ряда m_t к наличию у него единственного колебания, близкого к сумме синусоиды и косинусоиды.

4.3.2. Периодограмма и спектрограмма

График зависимости $R_k^2 = a_k^2 + b_k^2$ от периода T/k называется *периодограммой* (рис. 4.1). Применение ортогональных функций, рассмотренных выше, определяет периодограмму в точках $T/k = 2$ (если T четное), ..., $T/4$, $T/3$, $T/2$, T .

График зависимости R_k^2 от частоты k/T называется *спектрограммой* (рис. 4.2). Этот график удобнее, поскольку квадраты амплитуд, вычисленные для ортогональных слагаемых, располагаются на этом графике с одинаковыми промежутками, величина которых не зависит от T .

Аналогичные понятия можно ввести, используя квадрат теоретической амплитуды ρ_k^2 . Иногда этот квадрат амплитуды называют *интенсивностью*. С терминами периодограмма и спектрограмма в литературе обращаются весьма волно, вплоть до замены одного другим. В некоторых случаях при определении R_k^2 используют различные множители.

Следует заметить, что a_k и b_k суть значения параметров a и b , минимизирующие сумму

$$(19) \quad \sum_{t=1}^T \left[y_t - \left(a \cos \frac{2\pi k}{T} t + b \sin \frac{2\pi k}{T} t \right) \right]^2.$$

Поэтому $R_k = \sqrt{a_k^2 + b_k^2}$ и $\hat{\theta}_k = \text{arctg}(b_k/a_k)$ являются значениями величин R и θ , минимизирующими сумму

$$(20) \quad \sum_{t=1}^T \left[y_t - R \cos \left(\frac{2\pi k}{T} t - \theta \right) \right]^2.$$

Минимальное значение последней суммы есть

$$(21) \quad \sum_{t=1}^T y_t^2 - \frac{T}{2} (a_k^2 + b_k^2) = \sum_{t=1}^T y_t^2 - \frac{T}{2} R_k^2.$$

В этом смысле R_k^2 является мерой соответствия выравнивающей тригонометрической функции с частотой k/T наблюдаемым данным.

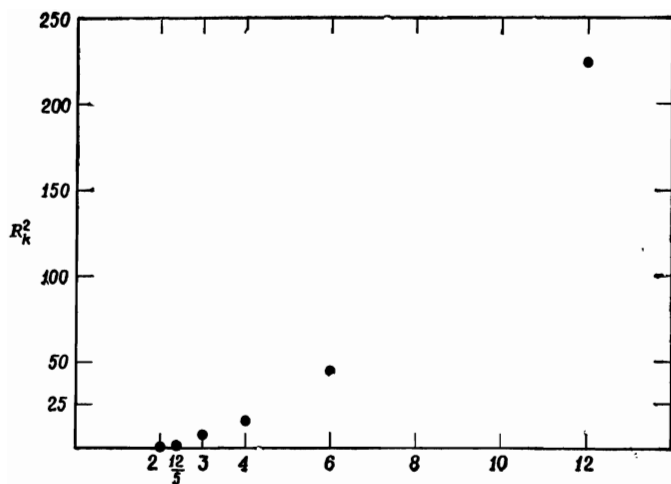


Рис. 4.1.

Периодограмма для примера 4.1 (период T/k).

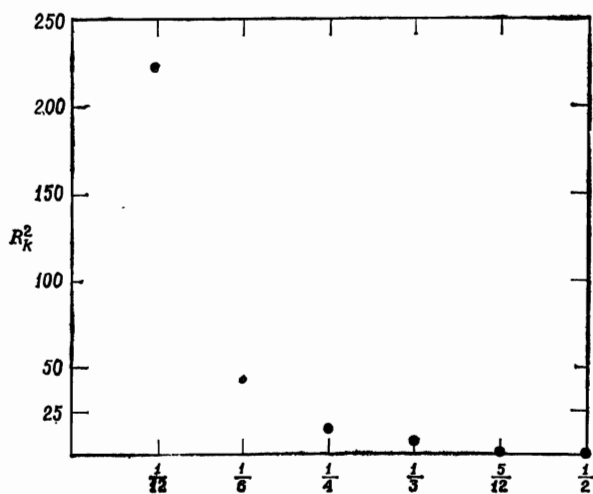


Рис. 4.2.

Спектрограмма для примера 4.1 (частота k/T).

Коэффициент корреляции между двумя последовательностями, $y_1, \dots, y_T$ и $z_1, \dots, z_T$, определяется отношением

$$(22) \quad \frac{\sum_{t=1}^T (y_t - \bar{y})(z_t - \bar{z})}{\left(\sum_{t=1}^T (y_t - \bar{y})^2\right)^{1/2} \left(\sum_{t=1}^T (z_t - \bar{z})^2\right)^{1/2}} = \frac{\sum_{t=1}^T y_t z_t - T \bar{y} \bar{z}}{\left(\sum_{t=1}^T y_t^2 - T \bar{y}^2\right)^{1/2} \left(\sum_{t=1}^T z_t^2 - T \bar{z}^2\right)^{1/2}},$$

в котором $\bar{y} = \sum_{t=1}^T y_t / T$, а $\bar{z} = \sum_{t=1}^T z_t / T$. Он является мерой подобия этих последовательностей, не меняющейся при сдвиге и изменении масштаба. Если $z_t = a \cos 2\pi kt / T + b \sin 2\pi kt / T = R \cos(2\pi kt / T - \theta)$, то соответствующий коэффициент корреляции равен

$$(23) \quad \frac{\sum_{t=1}^T y_t \left(a \cos \frac{2\pi k}{T} t + b \sin \frac{2\pi k}{T} t\right)}{\left(\sum_{t=1}^T (y_t - \bar{y})^2\right)^{1/2} \left(\frac{1}{2} T (a^2 + b^2)\right)^{1/2}} = \frac{\sum_{t=1}^T y_t \cos\left(\frac{2\pi k}{T} t - \theta\right)}{\left(\sum_{t=1}^T (y_t - \bar{y})^2\right)^{1/2} \left(\frac{1}{2} T\right)^{1/2}} = \frac{\left(\frac{1}{2} T\right)^{1/2}}{\left(\sum_{t=1}^T (y_t - \bar{y})^2\right)^{1/2}} [a_k \cos \theta + b_k \sin \theta].$$

Максимум его по всем θ достигается при значении $\hat{\theta}_k = \arctg \frac{b_k}{a_k}$

и равен $[T/2 \sum_{t=1}^T (y_t - \bar{y})^2]^{1/2} R_k$. Таким образом, максимальный (по отношению к параметру θ) коэффициент корреляции между последовательностями $\{y_t\}$ и $\{\cos(2\pi kt / T - \theta)\}$ пропорционален R_k .

4.3.3. Проверка гипотез

и построение доверительных областей для коэффициентов

Приведенные выше в соотношениях (10) — (13) оценки коэффициентов являются несмещенными, т. е.

$$(24) \quad \begin{aligned} \mathcal{E}a_0 &= \alpha_0, & \mathcal{E}a_{T/2} &= \alpha_{T/2}, \\ \mathcal{E}a(k_j) &= \alpha(k_j), & \mathcal{E}b(k_j) &= \beta(k_j), \end{aligned} \quad j = 1, \dots, q.$$

Дисперсии a_0 и $a_{T/2}$ равны σ^2/T , а дисперсии $a(k_j)$ и $b(k_j)$ равны $2\sigma^2/T$, $j = 1, \dots, q$. Эти оценки являются взаимно некоррелированными. Если величины y_t нормально распределены, то указанные оценки нормально распределены и независимы.

Статистика s^2 является несмещенной оценкой для σ^2 . Если величины y_t распределены нормально, то отношение s^2/σ^2 пропорционально величине, имеющей χ^2 -распределение, не зависящее от оценок коэффициентов. Если $f(t)$ представляется суммой (1), а s^2 задается соотношением (14), то отношение $(T - 2q - 1)s^2/\sigma^2$ имеет χ^2 -распределение с $(T - 2q - 1)$ степенями свободы. Если же $f(t)$ представляется в виде (2), а s^2 задается соотношением (15) (т. е. T четное и присутствует $a_{T/2}$), то отношение $(T - 2q - 2)s^2/\sigma^2$ имеет χ^2 -распределение с $(T - 2q - 2)$ степенями свободы.

В предположении нормальности можно использовать критерии и доверительные области, приведенные в гл. 2. При этом процедуры принимают особенно простой вид из-за ортогональности и нормировки используемых регрессионных переменных. Одной из нулевых гипотез, которые могут представлять для нас интерес, является гипотеза об отсутствии циклического слагаемого с заданным наименьшим периодом, скажем T/k_j . Эта нулевая гипотеза имеет вид

$$H: \alpha(k_j) = \beta(k_j) = 0.$$

Сформулированная гипотеза связана как с синусоидальной, так и с косинусоидальной составляющими потому, что в сумме они могут задавать сдвинутую функцию косинус. Мы здесь не затрагиваем вопроса о синхронизации, т. е. о фазе $\theta(k_j)$. Указанная выше гипотеза эквивалентна гипотезе $H: \rho(k_j) = 0$.

Если верна нулевая гипотеза, то величины $a(k_j)$ и $b(k_j)$ независимы и нормально распределены с нулевыми средними и дисперсиями $2\sigma^2/T$. Поэтому статистика

$$(26) \quad \frac{T}{2} \cdot \frac{a^2(k_j) + b^2(k_j)}{\sigma^2} = \frac{T}{2} \cdot \frac{R^2(k_j)}{\sigma^2}$$

имеет χ^2 -распределение с 2 степенями свободы. В противном случае она имеет нецентральное χ^2 -распределение с параметром нецентральности $Tr^2(k_j)/(2\sigma^2)$. (См. (49) ниже.) Пусть p — число оцениваемых коэффициентов ($2q + 1$ или $2q + 2$). Тогда при нулевой гипотезе статистика

$$(27) \quad \frac{\frac{T}{2} \cdot \frac{R^2(k_j)}{\sigma^2}}{2 \frac{s^2}{\sigma^2}} = \frac{TR^2(k_j)}{4s^2}$$

имеет F -распределение с 2 и $T - p$ степенями свободы. В общем случае (27) имеет нецентральное F -распределение с параметром

нецентральности $Tr^2(k_j)/(2\sigma^2)$. Нулевая гипотеза отвергается с уровнем значимости ε , если (27) больше, чем 100 ε -процентная точка F -распределения¹⁾. Соответствующая функция распределения равна

$$(28) \quad 1 - \left(1 + \frac{2}{T-p} F\right)^{-(T-p)/2}.$$

Задача проверки указанной гипотезы инвариантна относительно следующих преобразований оценок: $a_0^* = ka_0 + c_0$, $a^*(k_i) = ka(k_i) + c_i$, $b^*(k_i) = kb(k_i) + d_i$, $i \neq j$, $a_{T/2}^* = ka_{T/2} + c_{T/2}$ (если T четное), $a^*(k_i) = k[a(k_i)\cos\theta + b(k_i)\sin\theta]$, $b^*(k_j) = k[-a(k_j)\sin\theta + b(k_j)\cos\theta]$ и $s^{*2} = k^2s^2$ для произвольных $k \neq 0$, c_0 , c_i , d_i , $i \neq j$, $c_{T/2}$ и θ . Соответствующее преобразование переменных $y_1, \dots, y_T$ имеет вид

$$(29) \quad y_t^* = ky_t + c_0 + \sum_{\substack{i=1 \\ i \neq j}}^q \left(c_i \cos \frac{2\pi k_i}{T} t + d_i \sin \frac{2\pi k_i}{T} t \right) + \\ + \frac{2}{T} \sum_{s=1}^T \left\{ \cos \left[\frac{2\pi k_j}{T} (s-t) - \theta \right] - \cos \frac{2\pi k_j}{T} (s-t) \right\} ky_s + \\ + (-1)^t c_{T/2}, \quad t = 1, \dots, T.$$

Если T нечетное, то слагаемое $(-1)^t c_{T/2}$ отсутствует. (См. упр. 11 и 12.) Распределение переменных $y_1, \dots, y_T$ совпадает с распределением $y_1, \dots, y_T$ с соответственно преобразованными параметрами. Всякая функция параметров, инвариантная относительно введенных преобразований, является функцией от $\rho^2(k_j)/\sigma^2$. Всякая функция от достаточных статистик, инвариантная относительно введенных преобразований, является функцией от $R^2(k_j)/s^2$. Поскольку семейство распределений полно, единственными инвариантными критериями, основывающимися на достаточных статистиках и имеющими заданный уровень значимости, являются критерии, основанные на статистике $F = TR^2(k_j)/(4s^2)$. Поэтому равномерно наиболее мощным инвариантным критерием уровня ε будет критерий, который отвергает нулевую гипотезу, когда наблюдаемое значение F превосходит $F_{2,T-p}(\varepsilon)$ ²⁾. Этот критерий является равномерно наиболее мощным среди критериев, функции мощности которых зависят только от $\rho^2(k_j)/\sigma^2$.

¹⁾ Здесь и в дальнейшем под 100 ε -процентной точкой понимается *верхняя* 100 ε -процентная точка (см. примечание на стр. 23), если только не оговорено специально, что рассматривается *двусторонняя* 100 ε -процентная точка (см. примечание на стр. 48).— *Прим. перев.*

²⁾ Обозначение $F_{2,T-p}(\varepsilon)$ объясняется далее на стр. 130.— *Прим. перев.*

Гипотезы относительно значений α_0 и $\alpha_{T/2}$ можно проверять с помощью t -критериев. (См. упр. 14 и 15.) Если σ^2 известно, то F -критерии можно заменить χ^2 -критериями, а t -критерии — нормальными критериями.

Пусть $T = hn$, n — четное и априори предполагается, что $f(t) = f(t + n)$. Тогда для проверки гипотезы о том, что тренд ряда не содержит периодического изменения, мы должны проверить гипотезу о равенстве нулю коэффициентов при всех тригонометрических функциях, имеющих такой период (за исключением константы 1). Для этого можно использовать статистику

$$(30) \quad \frac{\frac{T}{2} \sum_{k=1}^{n/2-1} [a^2(kh) + b^2(kh)] + T\alpha_{T/2}^2}{(n-1)\sigma^2}.$$

Последняя при нулевой гипотезе, состоящей в том, что $f(t)$ постоянна, имеет F -распределение с $n-1$ и $T-n$ степенями свободы. В общей ситуации (когда $f(t)$ периодична с периодом n), статистика (30) имеет нецентральное F -распределение с параметром нецентральности

$$(31) \quad \frac{\frac{T}{2} \sum_{k=1}^{n/2-1} [\alpha^2(kh) + \beta^2(kh)] + T\alpha_{T/2}^2}{\sigma^2};$$

здесь $\alpha(kh)$ и $\beta(kh)$ — коэффициенты при $\cos 2\pi kht/T = \cos 2\pi kt/n$ и $\sin 2\pi kht/T = \sin 2\pi kt/n$, $k = 1, \dots, n/2 - 1$, в представлении $f(t)$, а $k_j = jh$.

Наблюдения можно расположить в виде таблицы Бьюис-Баллота

$$(32) \quad \begin{array}{cccc|c} y_1 & y_{n+1} & \dots & y_{(h-1)n+1} & \bar{y}_1 \\ y_2 & y_{n+2} & \dots & y_{(h-1)n+2} & \bar{y}_2 \\ \vdots & \vdots & & \vdots & \vdots \\ y_n & y_{2n} & \dots & y_{hn} & \bar{y}_n \end{array}$$

в которой

$$(33) \quad \bar{y}_t = \frac{1}{h} \sum_{j=0}^{h-1} y_{t+nj}, \quad t = 1, \dots, n.$$

Положим

$$(34) \quad \bar{y} = \frac{1}{n} \sum_{t=1}^n \bar{y}_t = \frac{1}{T} \sum_{t=1}^T y_t.$$

Имеем

$$(35) \quad \mathcal{E}y_{t+nj} = \mathcal{E}\bar{y}_t = f(t), \quad t = 1, \dots, n, \quad j = 0, \dots, h-1.$$

Нулевая гипотеза состоит в том, что $f(1) = f(2) = \dots = f(n)$. Это эквивалентно задаче о совпадении средних в дисперсионном анализе с равными числами наблюдений в каждом из классов однонарной классификации. Обычно для проверки этой нулевой гипотезы используется статистика

$$(36) \quad \frac{h \sum_{t=1}^n (\bar{y}_t - \bar{y})^2}{\sum_{t=1}^n \sum_{j=0}^{h-1} (y_{t+nj} - \bar{y}_t)^2} \cdot \frac{n(h-1)}{n-1},$$

имеющая F -распределение с $n-1$ и $n(h-1) = T-n$ степенями свободы при нулевой гипотезе и нецентральное F -распределение с параметром нецентральности

$$(37) \quad \frac{h \sum_{t=1}^n [f(t) - \bar{f}]^2}{\sigma^2},$$

где $\bar{f} = \sum_{t=1}^n f(t)/n$, в общем случае. Используя свойства тригонометрических функций, приведенные в разд. 4.2.1, можно показать, что F -статистики (30) и (36) совпадают. При этом совпадают и параметры нецентральности (31) и (37).

Доверительные области для параметров могут быть найдены обычным путем. Например, доверительная область для $\alpha(k_j)$ и $\beta(k_j)$ с коэффициентом доверия $1 - \varepsilon$ образуется парами чисел (α^*, β^*) , удовлетворяющими неравенству

$$(38) \quad [\alpha^* - a(k_j)]^2 + [\beta^* - b(k_j)]^2 \leq \frac{4s^2}{T} F_{2,T-p}(\varepsilon),$$

в котором $F_{2,T-p}(\varepsilon)$ есть 100\varepsilon-процентная точка F -распределения с 2 и $T-p$ степенями свободы. Эта доверительная область состоит из границы и внутренности круга с центром $[a(k_j), b(k_j)]$ и радиусом $\sqrt{4s^2 F_{2,T-p}(\varepsilon)/T}$. Точки граничной окружности и внутренности этого круга в полярной системе координат, $\rho^* = \sqrt{a^{*2} + \beta^{*2}}$, $\theta^* = \arctg(\beta^*/\alpha^*)$, образуют доверительную область для амплитуды $\rho(k_j)$ и фазы $\theta(k_j)$. Минимумом и максимумом ρ^* в круге являются граничные точки интервала

$$(39) \quad \left(R(k_j) - \sqrt{\frac{4s^2}{T} F_{2,T-p}(\varepsilon)}, \quad R(k_j) + \sqrt{\frac{4s^2}{T} F_{2,T-p}(\varepsilon)} \right).$$

Этот интервал является доверительным для $\rho(k_j)$ с коэффициентом доверия, большим $1 - \varepsilon$. (Если нижняя граница интервала (39) отрицательна, ее можно заменить нулем.) Используя нецентральное F -распределение статистики (27), можно построить доверитель-

ный интервал для параметра нецентральности $T\rho^2(k_j)/(2\sigma^2)$, но это не приносит особой пользы, поскольку σ^2 неизвестно.

Если $R(k_j) < \sqrt{4s^2 F_{2,T-p}(\varepsilon)/T}$, то начало координат попадает в доверительный круг и гипотеза $\rho^* = 0$ входит в число допустимых. Каждому значению угла θ^* соответствуют некоторые точки доверительного круга. При этом полезной является только верхняя граница (39). Нулевая гипотеза $\rho(k_j) = 0$ принимается с уровнем значимости ε . Если начало координат не принадлежит доверительному кругу, то можно определить доверительный интервал для $\theta(k_j)$, включая в него все углы θ^* , соответствующие этому кругу. Точнее говоря, в этот доверительный интервал включаются все углы θ^* , обладающие тем свойством, что луч, направленный из начала координат под углом θ^* к оси абсцисс, пересекает доверительный круг, и не включаются углы, не обладающие этим свойством. При этом длина доверительного интервала для θ меньше π , а коэффициент доверия больше $1 - \varepsilon$.

Другой подход использует тот факт, что величина $a(k_j) \operatorname{tg} \theta - b(k_j)$ распределена нормально со средним $\alpha(k_j) \operatorname{tg} \theta - \beta(k_j)$ и дисперсией $2\sigma^2(1 + \operatorname{tg}^2 \theta)/T = 2\sigma^2 \sec^2 \theta/T$. Если $\theta = \theta(k_j)$, то указанное среднее равно нулю. Поэтому

$$(40) \quad \Pr \left\{ \frac{[a(k_j) \operatorname{tg} \theta - b(k_j)]^2}{\frac{2}{T} s^2 (1 + \operatorname{tg}^2 \theta)} \leq F_{1,T-p}(\varepsilon) \right\} = 1 - \varepsilon.$$

Событие, заключенное в скобки, иначе можно записать в виде

$$(41) \quad \left[a^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon) \right] \left[\operatorname{tg} \theta - \frac{a(k_j) b(k_j)}{a^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon)} \right]^2 \leq \\ \leq \frac{\frac{2}{T} s^2 F_{1,T-p}(\varepsilon) \left[R^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon) \right]}{a^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon)}.$$

Из (41) получим следующее неравенство относительно $\operatorname{tg} \theta$:

$$(42) \quad \frac{a(k_j) b(k_j) - \sqrt{\frac{2}{T} s^2 F_{1,T-p}(\varepsilon) \left[R^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon) \right]}}{a^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon)} \leq \operatorname{tg} \theta \leq \\ \leq \frac{a(k_j) b(k_j) + \sqrt{\frac{2}{T} s^2 F_{1,T-p}(\varepsilon) \left[R^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon) \right]}}{a^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon)}.$$

если знаменатель в правой части (41) положителен (при этом подкоренное выражение неотрицательно), и

$$(43) \text{ или } \begin{aligned} \operatorname{tg} \theta &\geq - \frac{a(k_j) b(k_j) + \sqrt{\frac{2}{T} s^2 F_{1,T-p}(\varepsilon) \left[R^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon) \right]}}{a^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon)} \\ \operatorname{tg} \theta &\leq - \frac{a(k_j) b(k_j) - \sqrt{\frac{2}{T} s^2 F_{1,T-p}(\varepsilon) \left[R^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon) \right]}}{a^2(k_j) - \frac{2}{T} s^2 F_{1,T-p}(\varepsilon)}, \end{aligned}$$

если этот знаменатель отрицателен, а подкоренное выражение неотрицательно. Если же подкоренное выражение отрицательно, то оба отношения являются мнимыми. В таком случае неравенство (41) выполняется при всех значениях θ , поскольку при этом $a^2(k_j) < 2s^2 F_{1,T-p}(\varepsilon)/T$. Доверительное множество для $\operatorname{tg} \theta$ задается соотношением (42), если знаменатель правой части (41) положителен, или соотношением (43), если указанный знаменатель отрицателен, а подкоренное выражение неотрицательно. Это доверительное множество совпадает со всей действительной прямой, если указанное подкоренное выражение отрицательно. Вероятность того, что доверительный интервал совпадает со всей прямой, равна вероятности события

$$(44) \quad \frac{TR^2(k_j)}{4s^2} < \frac{1}{2} F_{1,T-p}(\varepsilon).$$

Левая часть (44) имеет нецентральное F -распределение с 2 и $T - p$ степенями свободы и параметром нецентральности $Tr^2(k_j)/(2\sigma^2)$. Вероятность события (44) мала, если $Tr^2(k_j)/(2\sigma^2)$ велико. Отметим, что доверительный интервал, основанный на (38), является тривиальным, если левая часть (44) меньше $F_{2,T-p}(\varepsilon)$, равного приблизительно $3/2$ правой части (44). [Шеффе (1970) предложил процедуру построения доверительного множества (38) для $\alpha(k_j)$ $\beta(k_j)$ в случае, когда это множество содержит начало координат, и построения интервалов, подобных (42) и (43), использующую вместо $F_{1,T-p}(\varepsilon)$ надлежащим образом выбранную монотонно убывающую функцию от $F_{2,T-p}(\varepsilon)$.]

Итак, имеется три типа доверительных множеств: (42), если $2s^2 F_{1,T-p}(\varepsilon)/T \leq a^2(k_j)$, (43), если $a^2(k_j) \leq 2s^2 F_{1,T-p}(\varepsilon)/T \leq a^2(k_j) + b^2(k_j)$, и вся прямая, если $a^2(k_j) + b^2(k_j) \leq 2s^2 F_{1,T-p}(\varepsilon)/T$. Они встречаются соответственно, когда $a^2(k_j)$ велико, когда $a^2(k_j)$ мало, а $b^2(k_j)$ велико, и когда обе эти величины малы.

Процедуры проверки гипотез и построения доверительных множеств, приведенные в этой главе, основываются на том, что

наблюдения независимы и нормально распределены. Тем не менее эти результаты сохраняются как асимптотические для фиксированных периодов T/k_j , если наблюдения независимы, их распределения имеют равномерно ограниченные (абсолютные) моменты порядка $2 + \delta$ для некоторого $\delta > 0$ или, в более общем случае, если выполняется условие Линдеберга — Феллера. (См. § 2.6.)

4.3.4. Решение вопроса о включении тригонометрических слагаемых

Трудной, но весьма существенной является задача о том, какие тригонометрические слагаемые следует включать в циклический тренд. Это задача с несколькими решениями, возникающая также во многих других областях применения регрессионных методов. Однако в данном случае она имеет некоторые особенности. Будем предполагать, что функция тренда содержит p членов. Точнее, это α_0 и, возможно, $\alpha_{T/2}(-1)^i$, если T четное, а также соответственно $(p-2)/2$ или $(p-1)/2$ пар тригонометрических членов с периодами T/k_j , $j = 1, \dots, (p-2)/2$ или $(p-1)/2$ соответственно. При этом мы располагаем оценкой σ^2 с $T - p$ степенями свободы. Пусть под вопросом находится включение в тренд q пар тригонометрических членов, тогда как $p - 2q$ членов уже определено должны быть включены в тренд. Мы хотим относительно каждой из q пар тригонометрических слагаемых решить, равна ли ее амплитуда нулю, т. е. для каждого $j = 1, \dots, q$ решить, будет ли $\rho(k_j) = 0$. Здесь возможны следующие решения:

$$\begin{aligned}
 H_0: & \rho(k_1) = \dots = \rho(k_q) = 0, \\
 H_i: & \rho(k_i) > 0, \rho(k_j) = 0, \quad j \neq i, \quad j = 1, \dots, q, \\
 (45) \quad H_{hi}: & \rho(k_h) > 0, \rho(k_i) > 0, \rho(k_j) = 0, \quad j \neq i, \quad j \neq h, \\
 & j = 1, \dots, q, \\
 & \vdots \\
 H_{12\dots q}: & \rho(k_h) > 0, \quad h = 1, \dots, q.
 \end{aligned}$$

Эта задача со многими решениями существенно отличается от рассмотренной в разд. 3.2.2 по той причине, что здесь уже не существует такого априорного естественного упорядочения периодов, каковым являлось упорядочение степеней полиномов.

Будем полагать, что о фазах априори ничего не известно и что их значения не представляют для нас интереса. Отсюда следует, что решающие процедуры должны основываться на статистиках $R^2(k_j)$. При этом они будут инвариантны относительно преобразований $\hat{a}_0 = a_0 + c_0$, $\hat{a}_{T/2} = a_{T/2} + c_{T/2}$ (если T четное), $\hat{a}^*(k_i) = a(k_i) + c_i$, $\hat{b}^*(k_i) = b(k_i) + d_i$, $i = q+1, \dots, (p-2)/2$ или $(p-1)/2$ и $\hat{a}^*(k_j) = a(k_j) \cos \theta_j + b(k_j) \sin \theta_j$, $\hat{b}^*(k_j) =$

$$= -a(k_j) \sin \theta_j + b(k_j) \cos \theta_j, \quad j = 1, \dots, q, \quad \text{т. е.}$$

$$(46) \quad y_t^* = y_t + c_0 + \sum_{i=q+1}^{[p/2-1]} \left(c_i \cos \frac{2\pi k_i}{T} t + d_i \sin \frac{2\pi k_i}{T} t \right) + \\ + \frac{2}{T} \sum_{s=1}^T \sum_{j=1}^q \left\{ \cos \left[\frac{2\pi k_j}{T} (s-t) - \theta_j \right] - \cos \frac{2\pi k_j}{T} (s-t) \right\} y_s + \\ + c_{T/2} (-1)^t.$$

Будем считать, кроме того, что заранее нельзя отдать предпочтение ни одной из рассматриваемых q периодических функций. В связи с этим естественно ограничиться только симметричными процедурами.

Рассмотрим сначала случай, когда σ^2 известно. Такая задача имеет определенный исторический интерес и является предельной для ряда других случаев. В предположении нормальности статистики $a_0, a(k_j), b(k_j), j = 1, \dots, (p-2)/2$ или $(p-1)/2$ и, возможно, $a_{T/2}$, образуют достаточное множество статистик для параметров $\alpha_0, \alpha(k_j), \beta(k_j), j = 1, \dots, (p-2)/2$ или $(p-1)/2$ и, соответственно, $\alpha_{T/2}$. Потребуем, чтобы решающие процедуры были инвариантны относительно преобразования (46). (Изменения фаз соответствуют вращениям пар $a(k_j), b(k_j)$.) Тогда они должны основываться на статистиках $R^2(k_1), \dots, R^2(k_q)$. Если $\rho(k_j) = 0$, то величина

$$(47) \quad z_j = \frac{T}{2\sigma^2} R^2(k_j)$$

имеет χ^2 -распределение с 2 степенями свободы. Плотность ее распределения и сама функция распределения равны соответственно

$$(48) \quad k(z) = \frac{1}{2} e^{-z/2}, \\ K(z) = \int_0^z k(v) dv = 1 - e^{-z/2}.$$

В противном случае z_j имеет нецентральное χ^2 -распределение с 2 степенями свободы и параметром нецентральности $\tau^2 = T\rho^2(k_j)/(\sigma^2)$. Соответствующая плотность¹⁾ имеет следующий

¹⁾ См., например, Т. Андерсон (1958, стр. 113). Параметр нецентральности получается заменой в квадратичной форме, имеющей χ^2 -распределение, всех входящих в нее нормальных случайных величин их математическими ожиданиями. Некоторые авторы называют параметром нецентральности вдвое меньшую величину.

вид:

$$\begin{aligned}
 (49) \quad k(z|\tau^2) &= \frac{1}{2\sqrt{\pi}} e^{-(\tau^2+z)/2} \sum_{\gamma=0}^{\infty} \frac{(\tau^2)^{\gamma} z^{\gamma} \Gamma\left(\gamma + \frac{1}{2}\right)}{(2\gamma)! \Gamma(\gamma+1)} = \\
 &= \frac{1}{2} e^{-(\tau^2+z)/2} \sum_{\gamma=0}^{\infty} \frac{(\tau^2)^{\gamma} z^{\gamma}}{2^{2\gamma} (\gamma!)^2} = \frac{1}{2} e^{-(\tau^2+z)/2} I_0(\tau\sqrt{z}),
 \end{aligned}$$

где

$$(50) \quad I_0(z) = \sum_{\gamma=0}^{\infty} \frac{z^{2\gamma}}{2^{2\gamma} (\gamma!)^2}$$

— бесселева функция первого рода с чисто мнимым аргументом порядка 0. В силу того что $I_0(z)$ выражается степенным рядом с положительными коэффициентами, она является монотонно возрастающей функцией аргумента z ($0 \leq z < \infty$). Поэтому $k(z|\tau^2)/k(z) = e^{-\tau^2/2} I_0(\tau\sqrt{z})$ — монотонно возрастающая функция от z ($0 \leq z < \infty$) при каждом $\tau^2 > 0$.

Рассмотрим процедуру проверки нулевой гипотезы о том, что величина некоторой амплитуды равна нулю, использующую оценку только этой амплитуды.

ТЕОРЕМА 4.3.1. *Равномерно наиболее мощный инвариантный критерий для проверки гипотезы $H_j^*: \rho(k_j) = 0$ против альтернативы $\rho(k_j) > 0$ с уровнем значимости ε_j имеет критическую область*

$$(51) \quad \frac{TR^2(k_j)}{2\sigma^2} > -2 \log \varepsilon_j.$$

Доказательство. Вероятность события (51) при нулевой гипотезе равна ε_j , поскольку функция распределения $z_j = TR^2(k_j)/(2\sigma^2)$ имеет вид $1 - e^{-z_j/2}$. Из фундаментальной леммы Неймана — Пирсона следует, что (51) является критической областью наиболее мощного критерия для проверки гипотезы H_j^* против любой простой альтернативы, например $\rho^2(k_j) = 2\sigma^2\tau^2/T$. Обозначим множество, определяемое соотношением (51), через A^* . Через A будем обозначать любую другую критическую область с $\text{Pr}\{A|H_j^*\} = \varepsilon_j$. Неравенство (51) эквивалентно неравенству

$$(52) \quad \frac{k(z_j|\tau^2)}{k(z_j)} = e^{-\tau^2/2} I_0(\tau\sqrt{z_j}) > e^{-\tau^2/2} I_0(\tau\sqrt{-2 \log \varepsilon_j}) = d_j,$$

правая часть которого обозначена для краткости символом d_j .

Поэтому

$$\begin{aligned}
 (53) \quad \Pr \{A^* | \tau^2\} &= \Pr \{A^* \cap \bar{A} | \tau^2\} + \Pr \{A^* \cap A | \tau^2\} = \\
 &= \int_{A^* \cap \bar{A}} k(z_j | \tau^2) dz_j + \Pr \{A^* \cap A | \tau^2\} > \\
 &> d_j \int_{A^* \cap \bar{A}} k(z_j) dz_j + \Pr \{A^* \cap A | \tau^2\} = \\
 &= d' \int_{\bar{A}^* \cap A} k(z_j) dz_j + \Pr \{A^* \cap A | \tau^2\} \geq \\
 &\geq \int_{\bar{A}^* \cap A} k(z_j | \tau^2) dz_j + \Pr \{A^* \cap A | \tau^2\} = \\
 &= \Pr \{A | \tau^2\}.
 \end{aligned}$$

Утверждение теоремы вытекает теперь из того, что неравенство (53) выполняется для всех $\tau^2 > 0$. ■

Приведенная процедура известна под названием *критерия Шустера* (Шустер (1898)).

Рассматриваемую задачу со многими решениями можно представить как комбинацию задач, включающих в себя гипотезы об отдельных амплитудах, а именно:

$$\begin{aligned}
 H_0 &= H_1^* \cap H_2^* \cap \dots \cap H_q^* = \bigcap_{i=1}^q H_i^*, \\
 H_i &= \bar{H}_i^* \cap \bigcap_{j=1}^q H_j^*, \quad i = 1, \dots, q, \\
 (54) \quad H_{hi} &= \bar{H}_h^* \cap \bar{H}_i^* \cap \bigcap_{j=1, j \neq h}^q H_j^*, \quad h < i, \quad h, i = 1, \dots, q, \\
 &\vdots \\
 H_{12 \dots q} &= \bigcap_{i=1}^q \bar{H}_i^*.
 \end{aligned}$$

Пусть R_0 , R_i и т. д. обозначают области из пространства значений $R^2(k_1), \dots, R^2(k_q)$, в которых соответственно принимаются гипотезы H_0 , H_i и т. д., а R_i^* , $i = 1, \dots, q$, — области, в которых принимаются гипотезы H_i^* . Эти два множества областей указанного выборочного пространства находятся точно в таком же соответствии, в каком находятся два введенных выше множества областей из пространства параметров, указанные в (54). Потребуем, чтобы

$$(55) \quad \Pr \{R_j^* | \rho(k_1), \dots, \rho(k_q); \rho(k_j) = 0\} = 1 - \varepsilon_j,$$

т. е. чтобы вероятность правильного решения $\rho(k_i) = 0$ не зависела от значений других амплитуд. Будем называть это свойство независимостью от мешающих параметров. Тогда рассуждения, подобные тем, которые были проведены в разд. 3.2.2 (мы их только наметим), показывают, что пересечения R_j^* и $R^2(k_i) = c_i$, $i \neq j$, имеют вероятности $1 - \varepsilon_j$ для почти всех наборов $q - 1$ неотрицательных значений $c_1, \dots, c_{j-1}, c_{j+1}, \dots, c_q$. Пусть $h[R^2(k_1), \dots, R^2(k_q)]$ равно 1 в области R_j^* и 0 вне этой области (т. е. h — критическая функция; ее можно продолжить до рандомизированной критической функции). Условное математическое ожидание h относительно $R^2(k_i)$ является ограниченной функцией остальных $q - 1$ аргументов. Из (55) следует, что математическое ожидание разности последнего и величины $(1 - \varepsilon_j)$ тождественно равно нулю, независимо от параметра нецентральности. Искомый результат вытекает теперь из ограниченной полноты семейства нецентральных χ^2 -распределений. (См. упр. 25 и 26.)

Из сказанного следует, что решение о $\rho(k_j)$ основывается в действительности только на $R^2(k_j)$. Если мы привлечем условие симметрии, то все ε_i , $i = 1, \dots, q$, будут равны. Обозначим это общее значение ε_i через ε . При данном ε наилучшая процедура для проверки H_j^* указана в теореме 4.3.1 (где $\varepsilon_j = \varepsilon$). Отметим, что

$$(56) \quad \Pr\{R_0 | H_0\} = \Pr\left\{\bigcap_{j=1}^q R_j^* \mid \rho(k_1) = \dots = \rho(k_q) = 0\right\} = \\ = \prod_{j=1}^q \Pr\{R_j^* \mid \rho(k_j) = 0\} = (1 - \varepsilon)^q.$$

Если задается значение именно этой вероятности в виде $1 - \delta$, то $1 - \varepsilon = (1 - \delta)^{1/q}$ или $\log(1 - \varepsilon) = [\log(1 - \delta)]/q$.

ТЕОРЕМА 4.3.2. Пусть задана вероятность $(1 - \varepsilon)^q$ принятия решения о том, что все рассматриваемые амплитуды равны 0, когда это в действительности имеет место. Тогда равномерно наиболее мощная симметричная инвариантная процедура выбора положительных амплитуд, не зависящая от мешающих параметров, состоит в том, что решение $\rho(k_i) > 0$ принимается при $TR^2(k_i)/(2\sigma^2) > -2 \log \varepsilon$.

Если все $\rho(k_j) = 0$, то ошибочным будет решение о том, что хотя бы одно $\rho(k_j) > 0$. Поскольку, согласно указанной процедуре, решение $\rho(k_i) > 0$ принимается при $z_i > C$, вероятность ошибки равна

$$(57) \quad \Pr\{\bar{R}_0 | H_0\} = 1 - (1 - e^{-C/2})^q.$$

Эта функция табулирована Дэвисом (1941, табл. 1), у которого $C/2$ и q соответствуют K и $N/2$. Указанная процедура в качестве

или меньше 1 отношение z_j/z_i . Обратимся к проблеме выбора между гипотезами H_j и H_i .

ЛЕММА 4.3.1. *Равномерно наиболее мощная симметричная инвариантная процедура выбора между гипотезами H_j и H_i состоит в принятии H_j при $z_j > z_i$ и принятии H_i при $z_j < z_i$.*

ДОКАЗАТЕЛЬСТВО. В данном случае под симметрией понимается симметрия по z_j и z_i , т. е. B_j и B_i симметричны, если в B_j замена z_j на z_i и обратно переводит B_j в B_i и обратно. Пусть $A_j^* = \{(z_1, \dots, z_q) | z_j > z_i\}$, $A_i^* = \{(z_1, \dots, z_q) | z_j < z_i\}$, и пусть A_j и A_i — два других взаимно дополнительных не пересекающихся симметричных множества. (Мы пренебрегаем множествами вероятности 0, такими, как $\{(z_1, \dots, z_q) | z_j = z_i\}$.) Тогда, из симметрии $A_j^* \cap A_i$ и $A_j \cap A_i^*$,

$$\begin{aligned}
 (63) \quad & \Pr \{A_j^* | \tau_j^2 = \tau^2, \tau_i^2 = 0\} = \\
 & = \Pr \{A_j^* \cap A_i | \tau_j^2 = \tau^2, \tau_i^2 = 0\} + \\
 & \quad + \Pr \{A_j^* \cap A_j | \tau_j^2 = \tau^2, \tau_i^2 = 0\} = \\
 & = \int \dots \int_{A_j^* \cap A_i} \prod_{h=1}^q k(z_h) e^{-\tau^2/2} I_0(\tau \sqrt{z_j}) \prod_{h=1}^q dz_h + \\
 & \quad + \Pr \{A_j^* \cap A_j | \tau_j^2 = \tau^2, \tau_i^2 = 0\} \geq \\
 & \geq \int \dots \int_{A_j^* \cap A_i} \prod_{h=1}^q k(z_h) e^{-\tau^2/2} I_0(\tau \sqrt{z_i}) \prod_{h=1}^q dz_h + \\
 & \quad + \Pr \{A_j^* \cap A_j | \tau_j^2 = \tau^2, \tau_i^2 = 0\} = \\
 & = \int \dots \int_{A_j \cap A_i^*} \prod_{h=1}^q k(z_h) e^{-\tau^2/2} I_0(\tau \sqrt{z_j}) \prod_{h=1}^q dz_h + \\
 & \quad + \Pr \{A_j^* \cap A_j | \tau_j^2 = \tau^2, \tau_i^2 = 0\} = \\
 & = \Pr \{A_j | \tau_j^2 = \tau^2, \tau_i^2 = 0\}.
 \end{aligned}$$

Это неравенство выполняется при всех значениях $\tau^2 > 0$. Отсюда следует, что указанная симметричная процедура является равномерно наиболее мощной по τ^2 . ■

Теорема 4.3.2 и лемма 4.3.1 утверждают, что в указанной выше задаче со многими решениями следует принимать гипотезу H_0 , если все выборочные амплитуды малы, и принимать гипотезу H_j , если $R^2(k_j)$ — максимальная и достаточно большая по величине

амплитуда. Соответствующая процедура симметрична в том смысле, что одновременная перестановка индексов у $H_1, \dots, H_q$ и у $R^2(k_1), \dots, R^2(k_q)$ не изменяет процедуры.

ТЕОРЕМА 4.3.3. *Равномерно наиболее мощная симметричная инвариантная решающая процедура выбора только одной из гипотез $H_0, H_1, \dots, H_q$ при заданной вероятности $\text{Pr}\{R_0|H_0\} = (1 - \varepsilon)^q$ и при условии независимости процедуры от мешающих параметров, состоит в принятии гипотезы H_0 , если $z_j = TR^2(k_j)/(2\sigma^2) \leq \leq -2 \log \varepsilon$, $j = 1, \dots, q$, и в противном случае — в принятии гипотезы H_i с индексом j , удовлетворяющим условию $z_j > z_i$, $i \neq j$, $i = 1, \dots, q$.*

Доказательство. Пусть $R_0, R_1, \dots, R_q$ — полная группа взаимно непересекающихся областей в пространстве значений $z_1, \dots, z_q$, в которых соответственно принимаются гипотезы $H_0, H_1, \dots, H_q$, отвечающая некоторой произвольной процедуре с $\text{Pr}\{R_0|H_0\} = (1 - \varepsilon)^q$. Положим $R_0 = R_0$ и

$$(64) \quad R'_j = \bar{R}_0 \cap \{(z_1, \dots, z_q) | z_j > z_i, i \neq j\}, \quad j = 1, \dots, q.$$

Тогда $\text{Pr}\{R'_0|H_0\} = (1 - \varepsilon)^q$ и

$$\begin{aligned} (65) \quad \text{Pr}\{R'_j | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} &= \\ &= \text{Pr}\{R'_j \cap \bar{R}_j | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} + \\ &\quad + \text{Pr}\{R'_j \cap R_j | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} = \\ &= \sum_{i \neq j} \text{Pr}\{R'_j \cap R_i | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} + \\ &\quad + \text{Pr}\{R'_j \cap R_j | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} \geq \\ &\geq \sum_{i \neq j} \text{Pr}\{R_j \cap R'_i | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} + \\ &\quad + \text{Pr}\{R'_j \cap R_j | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} = \\ &= \text{Pr}\{R_j | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\}, \quad j = 1, \dots, q, \end{aligned}$$

поскольку в силу симметрии и доказательства леммы 4.3.1

$$\begin{aligned} (66) \quad \text{Pr}\{R'_j \cap R_i | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} &\geq \\ &\geq \text{Pr}\{R_j \cap R'_i | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\}. \end{aligned}$$

Положим, далее,

$$(67) \quad R_0^* = \{(z_1, \dots, z_q) | z_i < -2 \log \varepsilon, i = 1, \dots, q\},$$

$$(68) \quad R_j^* = \bar{R}_0^* \cap \{(z_1, \dots, z_q) | z_j > z_i, i \neq j\}, \quad j = 1, \dots, q.$$

При этом $\Pr \{R_0^* | H_0\} = (1 - \varepsilon)^q$ и

$$\begin{aligned}
 (69) \quad \Pr \{R_j^* | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} &= \\
 &= \Pr \{R_j^* \cap R_0' | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} + \\
 &\quad + \Pr \{R_j^* \cap R_j' | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} \geq \\
 &\geq \Pr \{R_j' \cap R_0^* | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} + \\
 &\quad + \Pr \{R_j^* \cap R_j' | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\} = \\
 &= \Pr \{R_j' | \tau_j^2 = \tau^2, \tau_h^2 = 0, h \neq j\}
 \end{aligned}$$

аналогично доказательству теоремы 4.3.1. Поскольку (65) и (69) выполняются для всех τ^2 , отсюда следует утверждение теоремы. ■

ТЕОРЕМА 4.3.4. *Всякое симметричное байесовское решение состоит в принятии гипотезы H_0 при $TR^2(k_j)/(2\sigma^2) \leq C$, $j = 1, \dots, q$, и в противном случае — в принятии гипотезы H_j с индексом j , удовлетворяющим условию $z_j > z_i$, $i \neq j$, $i = 1, \dots, q$.*

ДОКАЗАТЕЛЬСТВО. Всякое байесовское решение определяется [см., например, Т. Андерсон (1958, § 6.6)] соотношениями

$$(70) \quad R_0: p_0 \prod_{j=1}^q k(z_j) > p_i k(z_i | \tau^2) \prod_{j \neq i} k(z_j), \quad i = 1, \dots, q,$$

$$(71) \quad R_h: \begin{cases} p_h k(z_h | \tau^2) \prod_{j \neq h} k(z_j) > p_i k(z_i | \tau^2) \prod_{j \neq i} k(z_j), \\ p_h k(z_h | \tau^2) \prod_{j \neq h} k(z_j) > p_0 \prod_{j=1}^q k(z_j), \end{cases} \quad i = 1, \dots, q, i \neq h,$$

где $p_i \geq 0$, $i = 0, 1, \dots, q$ и $\sum_{i=0}^q p_i = 1$. Эти области в силу их определения не пересекаются. Каждая точка на общей границе пары областей (попадание на которую есть событие нулевой вероятности) может быть приписана любой из них. Первая совокупность неравенств ($p_i > 0$) равносильна соотношению

$$(72) \quad R_0: \frac{k(z_i | \tau^2)}{k(z_i)} < \frac{p_0}{p_i}, \quad i = 1, \dots, q,$$

левая часть которого является монотонно возрастающей функцией от z_i . Определяемая этим соотношением область R_0 симметрична тогда и только тогда, когда $p_1 = \dots = p_q$. Поэтому R_0 имеет вид, указанный в теореме, и при надлежащем выборе значения

$p_0 = 1 - \sum_{i=1}^q p_i$ в точности таково же, как и в теореме. Аналогичные замечания справедливы и для остальных областей R_i . ■

Отметим, что эта процедура сводится к критерию значимости, приведенному выше, при следующем дополнительном условии. Если нулевая гипотеза H_0 отклоняется, то принимается решение о том, что наибольшая выборочная амплитуда соответствует единственной отличной от нуля теоретической амплитуде. Этот тип задачи исследовался общими методами Карлином и Труаксом (1960), а также Кудо (1960).

Поскольку в данной задаче класс байесовских решений совпадает с классом допустимых решений [см., например, Т. Андерсон (1958, теорема 6.6.4)], то теоремы 4.3.3 и 4.3.4 являются эквивалентными.

Перейдем теперь к более важному случаю неизвестного σ^2 . Надо различать ситуации, когда интересуются оценками некоторых q заданных частот и в распоряжении имеется оценка для σ^2 с $n = T - p$ степенями свободы, полученная независимым образом, и когда оценка для σ^2 зависит от оцениваемых частот. В первом случае существует оценка s^2 , такая, что $ns^2/\sigma^2 = S/\sigma^2$ имеет χ^2 -распределение с n степенями свободы независимо от q пар коэффициентов $a(k_1), b(k_1), \dots, a(k_q), b(k_q)$.

В первой из указанных ситуаций можно потребовать, чтобы процедура была симметричной и удовлетворяла соотношению (55), т. е. чтобы вероятность решения $\rho(k_j) = 0$ не зависела от величин других амплитуд, и чтобы вероятность правильного решения о равенстве всех амплитуд нулю не зависела бы от σ^2 . При этом наилучшей для проверки H_j^* будет процедура, отклоняющая эту гипотезу при $F_j = TR^2(k_j)/(4s^2) > c_j$. (Это вытекает из рассмотрения достаточных статистик, инвариантных относительно преобразований (29), и полноты соответствующего семейства распределений; см. замечание, следующее за (29).) Ввиду требования симметрии процедуры положим $c_1 = \dots = c_q = c$. Тогда

$$\begin{aligned}
 (73) \quad \Pr\{R_0 | H_0\} &= \Pr\left\{z_j < 2c \frac{s^2}{\sigma^2}, \quad j = 1, \dots, q | H_0\right\} = \\
 &= \int_0^\infty (1 - e^{-cu/n})^q \frac{u^{n/2-1} e^{-u/2}}{2^{n/2} \Gamma(n/2)} du = \\
 &= \int_0^\infty \sum_{j=0}^q (-1)^j \binom{q}{j} \frac{u^{n/2-1} e^{-[1+2cj/n]u/2}}{2^{n/2} \Gamma(n/2)} du = \\
 &= \sum_{j=0}^q (-1)^j \binom{q}{j} \left(1 + \frac{2cj}{n}\right)^{-n/2}.
 \end{aligned}$$

(Эта величина при $n \rightarrow \infty$ стремится к $(1 - e^{-c})^q$.) При этом c можно выбрать таким образом, чтобы вероятность (73) была равна заранее заданной величине.

ТЕОРЕМА 4.3.5. *Если σ^2 неизвестно, то равномерно наиболее мощная симметричная решающая процедура для выбора положительной амплитуды при заданной вероятности правильного решения о том, что все амплитуды равны нулю, и при условии независимости от мешающих параметров состоит в принятии решения $\rho(k_j) > 0$ при $TR^2(k_j)/(4s^2) > c$, где c определяется так, чтобы вероятность (73) была равна заданной величине.*

Если σ^2 неизвестно, приведенная процедура возможна только в случае существования такой оценки s^2 для σ^2 , распределение которой не зависит от тех $\alpha(k_1), \dots, \beta(k_q)$, которые включены в гипотезы. В противном случае невозможно выполнить требование независимости от мешающих параметров. Такая ситуация возникает, когда оказывается нежелательным предполагать равенство нулю даже хотя бы одной пары коэффициентов $\alpha(k_j), \beta(k_j)$, т. е. когда желательно допускать тригонометрические члены любой частоты (вида $j/T, j = 0, 1, \dots, [T/2]$). Тем не менее иногда можно предполагать, что в действительности отличны от нуля только некоторые из этих коэффициентов, в количестве, не большем некоторого малого числа. Исследуем этот простейший случай подробно.

Предположим, что отличной от нуля может быть только одна из q теоретических амплитуд. Задача состоит в том, чтобы решить, какая из гипотез $H_0, H_1, \dots, H_q$ верна. Потребуем, чтобы процедуры были симметричными и чтобы вероятность $\text{Pr}\{R_0|H_0\}$ определялась независимо от неизвестного σ^2 . Будем считать далее, что $r - q$ выборочных амплитуд $R^2(k_{q+1}), \dots, R^2(k_r)$ соответствуют нулевым теоретическим амплитудам. При этом $r - q$ может быть равным и нулю. Таким образом, в тренде остается еще $T - 2r$ коэффициентов, которые, возможно, и отличны от нуля, но нас не интересуют. (Такая постановка задачи является более общей, чем рассматривавшиеся другими авторами.) Если H_0 верна, то

$\sum_{j=1}^r R^2(k_j), a_0$ и все выборочные коэффициенты, соответствующие

другим теоретическим коэффициентам, возможно отличным от нуля, образуют достаточное множество статистик для соответствующих параметров. Отсюда следует, что для почти всех наборов значений

$\sum_{j=1}^r R^2(k_j), a_0$ и значений других указанных коэффициентов услов-

ная вероятность R_0 должна быть равна заданной вероятности $\text{Pr}\{R_0|H_0\}$ (см. упр. 10 гл. 3). Совместная плотность вероятностей

значений $z_j = TR^2(k_j)/(2\sigma^2), j = 1, \dots, r - 1$, при $\sum_{j=1}^r z_j = c$, т. е.

$z_r = c - \sum_{j=1}^{r-1} z_j$, равна

$$(74) \quad \prod_{j=1}^r k(z_j) = \begin{cases} \left(\frac{1}{2}\right)^r e^{-c/2}, & z_j \geq 0, \quad j = 1, \dots, r, \\ 0 & \text{в противном случае,} \end{cases}$$

когда верна гипотеза H_0 , и равна

$$(75) \quad k(z_j | \tau_j^2) \prod_{i \neq j} k(z_i) = \begin{cases} \left(\frac{1}{2}\right)^r e^{-c/2} e^{-\tau_j^2/2} I_0(\tau_j \sqrt{z_j}), & z_i \geq 0, \quad i = 1, \dots, r, \\ 0 & \text{в противном случае,} \end{cases}$$

когда верна гипотеза H_j . Мы используем здесь те же методы, что и при доказательстве теоремы 4.3.3. Пересечение областей R_0 и

$\sum_{j=1}^r z_j = c$ определяется неравенствами $z_i \leq g(c)$, $i = 1, \dots, q$, где $g(c)$ выбирается таким образом, чтобы условная вероятность совпадала с заданной $\Pr\{R_0 | H_0\}$. Пересечение областей R_h , $h = 1, \dots, q$, и $\sum_{j=1}^r z_j = c$ определяется неравенствами $z_h > g(c)$ и $z_h > z_i$, $i \neq h$, $i = 1, \dots, q$.

Условное совместное распределение значений $z_1, \dots, z_r$ при условии $\sum_{j=1}^r z_j = c$ является равномерным в области r -мерного пространства, определяемой соотношениями $z_j \geq 0$, $j = 1, \dots, r$, $\sum_{j=1}^r z_j = c$. Отсюда следует, что условное распределение значений $z_1/c, \dots, z_r/c$ при условии $\sum_{j=1}^r z_j/c = 1$ будет равномерным в области $z_j/c \geq 0$, $j = 1, \dots, r$, $\sum_{j=1}^r z_j/c = 1$. Если $g(c) = gc$ для каждого положительного g , то вероятность

$$(76) \quad \Pr\left\{z_j \leq gc, \quad j = 1, \dots, q \mid \sum_{j=1}^r z_j = c, H_0\right\} = \\ = \Pr\left\{\frac{z_j}{c} \leq g, \quad j = 1, \dots, q \mid \sum_{j=1}^r \frac{z_j}{c} = 1, H_0\right\}$$

не зависит от c и совпадает поэтому с безусловной вероятностью

$$(77) \quad \Pr\left\{\frac{z_j}{\sum_{i=1}^r z_i} \leq g, \quad j = 1, \dots, q \mid H_0\right\}.$$

ТЕОРЕМА 4.3.6. Если σ^2 неизвестно, то равномерно наиболее мощная симметричная инвариантная решающая процедура выбора среди $H_0, H_1, \dots, H_q$ при заданной вероятности $\Pr\{R_0 | H_0\}$ состоит в принятии H_0 , если

$$(78) \quad x_j = \frac{R^2(k_j)}{\sum_{i=1}^r R^2(k_i)} \leq g, \quad j = 1, \dots, q.$$

В противном случае принимается гипотеза H_j с индексом j , удовлетворяющим соотношению $R^2(k_j) > R^2(k_i)$, $i \neq j$, $i = 1, \dots, q$. При этом константа g выбирается таким образом, чтобы вероятность события (78) при гипотезе H_0 равнялась заданной $\Pr\{R_0 | H_0\}$.

Вычислим теперь вероятность события, дополнительного к (78), т. е.

$$(79) \quad \Pr\{\max_{j=1, \dots, q} x_j > g | H_0\} = \\ = \Pr\{x_j > g \text{ хотя бы для одного } j, \quad j = 1, \dots, q | H_0\}.$$

Обозначим через A_j событие $\{x_j > g\}$. Тогда при гипотезе H_0

$$(80) \quad \Pr\{\max_{j=1, \dots, q} x_j > g\} = \Pr\left\{\bigcup_{j=1}^q A_j\right\} = \\ = \sum_{j=1}^q \Pr\{A_j\} - \sum_{i < l} \Pr\{A_i \cap A_l\} + \dots + (-1)^{q-1} \Pr\left\{\bigcap_{j=1}^q A_j\right\}.$$

[См., например, Феллер (1968, гл. IV, § 1).] Далее, в силу одинаковой распределенности величин x_j имеем

$$(81) \quad \Pr\left\{\bigcup_{j=1}^q A_j\right\} = \\ = q \Pr\{A_1\} - \binom{q}{2} \Pr\{A_1 \cap A_2\} + \dots + (-1)^{q-1} \Pr\left\{\bigcap_{j=1}^q A_j\right\}.$$

Поскольку $\sum_{j=1}^r x_j = 1$ и $x_j \geq 0$, $j = 1, \dots, r$, то при $kg \geq 1$

$$(82) \quad \Pr\left\{\bigcap_{j=1}^k A_j\right\} = \Pr\{x_1 > g, \dots, x_k > g\} = 0.$$

Отсюда вытекает, что

$$(83) \quad \Pr\left\{\bigcup_{j=1}^q A_j\right\} = q \Pr\{A_1\} - \binom{q}{2} \Pr\{A_1 \cap A_2\} + \dots$$

$$\dots + (-1)^{n-1} \binom{q}{n} \Pr \left\{ \bigcap_{j=1}^n A_j \right\},$$

где n — меньшее из двух чисел q и $[1/g]$, а $[1/g]$ — наибольшее целое число, не превосходящее $1/g$.

Условное распределение значений $x_1, \dots, x_r$ при условии $\sum_{j=1}^r x_j = 1$ является равномерным в области значений $x_j \geq 0$, $j = 1, \dots, r$, $\sum_{j=1}^r x_j = 1$. Оно совпадает с безусловным распределением

величин $x_1, \dots, x_r$, поскольку условие $\sum_{j=1}^r x_j = 1$ выполняется автоматически в силу определения x_j . Координаты $x_1, \dots, x_r$ в $(r-1)$ -мерном пространстве $\sum_{j=1}^r x_j = 1$ являются *барицентрическими*. Вероятность попадания в произвольное подмножество этого пространства пропорциональна его $(r-1)$ -мерному объему. Имеем

$$\begin{aligned} (84) \quad \Pr \{A_1\} &= \Pr \{x_1 > g\} = \\ &= \Pr \left\{ g < x_1 \leq 1, 0 \leq x_j < 1-g, j \neq 1 \mid \sum_{j=1}^r x_j = 1 \right\} = \\ &= \frac{V \left\{ g < x_1 \leq 1, 0 \leq x_j < 1-g, j \neq 1, \sum_{j=1}^r x_j = 1 \right\}}{V \left\{ 0 \leq x_j \leq 1, j = 1, \dots, r, \sum_{j=1}^r x_j = 1 \right\}} = \\ &= \frac{V \left\{ 0 < x_1 - g \leq 1-g, 0 \leq x_j < 1-g, j \neq 1, x_1 - g + \sum_{j=2}^r x_j = 1-g \right\}}{V \left\{ 0 \leq x_j \leq 1, j = 1, \dots, r, \sum_{j=1}^r x_j = 1 \right\}}. \end{aligned}$$

(Здесь буквой V обозначен объем.— *Перев.*) Полагая в числителе $z_1 = x_1 - g$, $z_2 = x_2, \dots, z_r = x_r$, получаем

$$(85) \quad \Pr \{A_1\} = \frac{V \left\{ 0 \leq z_j \leq 1-g, j = 1, \dots, r, \sum_{j=1}^r z_j = 1-g \right\}}{V \left\{ 0 \leq x_j \leq 1, j = 1, \dots, r, \sum_{j=1}^r x_j = 1 \right\}}.$$

Правая часть (85) является отношением объемов двух подобных фигур $(r-1)$ -мерных подпространств, линейные размеры кото-

рых находятся в отношении $(1 - g) : 1$. Поэтому

$$(86) \quad \Pr \{A_1\} = (1 - g)^{r-1}.$$

Для определения вероятности

$$(87) \quad \Pr \{A_1 \cap A_2\} = \Pr \{x_1 > g, \quad x_2 > g\}$$

рассмотрим

$$(88) \quad V \left\{ g < x_j < 1 - g, \quad j = 1, 2, \quad 0 \leq x_j < 1 - 2g, \quad j = 3, \dots, r, \right. \\ \left. \sum_{j=1}^r x_j = 1 \right\} =$$

$$= V \left\{ 0 < x_j - g < 1 - 2g, \quad j = 1, 2, \quad 0 \leq x_j < 1 - 2g, \right. \\ \left. j = 3, \dots, r, \quad (x_1 - g) + (x_2 - g) + \sum_{j=3}^r x_j = 1 - 2g \right\}.$$

Обозначая $z_1 = x_1 - g$, $z_2 = x_2 - g$, $z_3 = x_3$, ..., $z_r = x_r$, получаем для (88) выражение

$$(89) \quad V \left\{ 0 \leq z_j \leq 1 - 2g, \quad j = 1, \dots, r, \quad \sum_{j=1}^r z_j = 1 - 2g \right\}.$$

Отсюда, из тех же соображений, имеем

$$(90) \quad \Pr \{A_1 \cap A_2\} = (1 - 2g)^{r-1}.$$

В общем случае,

$$(91) \quad \Pr \left\{ \bigcap_{j=1}^k A_j \right\} = (1 - kg)^{r-1}, \quad k \leq \frac{1}{g}.$$

Наконец, из этого соотношения и из (83) следует, что

$$(92) \quad \Pr \left\{ \max_{j=1, \dots, q} x_j > g \right\} = \Pr \left\{ \bigcup_{j=1}^q A_j \right\} = \\ = q(1 - g)^{r-1} - \binom{q}{2}(1 - 2g)^{r-1} + \dots \\ \dots + (-1)^{n-1} \binom{q}{n}(1 - ng)^{r-1},$$

где n — меньшее из чисел q и $[1/g]$.

Это распределение вместе с его таблицами приведено Фишером (1929). Указанная процедура называется *критерием Фишера*. Геометрическое представление для $q = r = 3$ предлагается в упр. 27. Уиттл (1951) использовал метод характеристических функций; Ирвин (1955) обсуждал различные методы.

Полученный результат относительно вероятности (79) можно обобщить. В этом направлении следующим шагом является вычисление вероятности

$$\begin{aligned}
 (93) \quad & \Pr \{ \text{второй по величине } x_j > g, \quad j = 1, \dots, q \mid H_0 \} = \\
 & = \Pr \{ \text{по крайней мере два } x_j > g, \quad j = 1, \dots, q \mid H_0 \} = \\
 & = \Pr \left\{ \bigcup_{i < j} (A_i \cap A_j) \right\} = \\
 & = \sum_{i < j} \Pr \{ A_i \cap A_j \} - 2 \sum_{i < j < k} \Pr \{ A_i \cap A_j \cap A_k \} + \\
 & \quad + 3 \sum_{h < i < j < k} \Pr \{ A_h \cap A_i \cap A_j \cap A_k \} - \dots + \\
 & \quad + (-1)^q (q-1) \Pr \left\{ \bigcap_{i=1}^q A_i \right\} = \\
 & = \binom{q}{2} \Pr \{ A_1 \cap A_2 \} - 2 \binom{q}{3} \Pr \{ A_1 \cap A_2 \cap A_3 \} + \\
 & \quad + 3 \binom{q}{4} \Pr \{ A_1 \cap A_2 \cap A_3 \cap A_4 \} - \dots \\
 & \quad \dots + (-1)^q (q-1) \Pr \left\{ \bigcap_{i=1}^q A_i \right\} = \\
 & = \binom{q}{2} (1-2g)^{r-1} - 2 \binom{q}{3} (1-3g)^{r-1} + 3 \binom{q}{4} (1-4g)^{r-1} - \dots \\
 & \quad \dots + (-1)^n (n-1) \binom{q}{n} (1-ng)^{r-1},
 \end{aligned}$$

где n — меньшее из чисел q и $[1/g]$. Разложение для $\Pr \left\{ \bigcup_{i < j} (A_i \cap A_j) \right\}$, использованное здесь, приведено Феллером (1968, гл. IV, § 5). Вообще имеет место следующий результат.

ТЕОРЕМА 4.3.7. Если $x_1, \dots, x_r$ равномерно распределены в $(r-1)$ -мерной области $x_j \geq 0, j = 1, \dots, r, \sum_{j=1}^r x_j = 1$, то

$$\begin{aligned}
 (94) \quad \Pr \{x_j \geq g \text{ не менее чем для } p \text{ индексов } j = 1, \dots, q\} = \\
 = \binom{q}{p} (1 - pg)^{r-1} - \binom{q}{p+1} \binom{p}{p-1} [1 - (p+1)g]^{r-1} + \\
 + \binom{q}{p+2} \binom{p+1}{p-1} [1 - (p+2)g]^{r-1} - \dots + \\
 + (-1)^{n-p} \binom{q}{n} \binom{n-1}{p-1} (1 - ng)^{r-1} = \\
 = q \binom{q-1}{p-1} \left\{ \frac{1}{p} (1 - pg)^{r-1} - \frac{1}{p+1} \binom{q-p}{1} \times \right. \\
 \times [1 - (p+1)g]^{r-1} + \frac{1}{p+2} \binom{q-p}{2} [1 - (p+2)g]^{r-1} - \\
 \left. - \dots + \frac{(-1)^{n-p}}{n} \binom{q-p}{n-p} (1 - ng)^{r-1} \right\},
 \end{aligned}$$

где $p \leq q$, а n равно меньшему из чисел q и $[1/g]$.

Стивенс (1939) вычислил эту вероятность при решении другой задачи и привел таблицы для $p = 2$. Фишер (1940) установил связь между задачей Стивенса и задачей, которую мы рассматриваем. Кроме того, он привел ряд таблиц.

Однако полученный результат имеет ограниченное применение. Например, для $p = 2$ указанное распределение дает вероятность того, что два отношения x_j будут больше заданного числа g при гипотезе H_0 . Это соответствует процедуре, в которой гипотеза H_0 принимается, если только не найдется по крайней мере двух больших, чем g , значений x_j . Фишер (1940) предположил, что такая процедура могла бы быть приемлемой тогда, когда предполагается, что реальный эффект проявляется скорее всего в двух амплитудах, а не в одной. В наших обозначениях предлагаемую процедуру можно перефразировать следующим образом. Гипотеза H_0 принимается за исключением случая, когда выполняются по крайней мере два неравенства $x_j \geq g$. В последнем случае принимается гипотеза H_{ij} с индексами i и j , такими, что $R^2(k_i)$ и $R^2(k_j)$ — наибольшие выборочные амплитуды, $i, j = 1, \dots, q$.

Сформулируем теперь байесовскую задачу решения о выборе среди гипотез $H_0, H_1, \dots, H_q, H_{12}, H_{13}, \dots, H_{q-1,q}$. При этом предположим, что σ^2 известно. Принимая во внимание симметрию задачи, будем брать априорные вероятности симметричными и

рассматривать гипотезы при симметричных значениях $\tau_1^2, \dots, \tau_q^2$. Пусть p_0 — априорная вероятность гипотезы H_0 , p_1 — априорная вероятность события $\tau_j^2 = \tau_1^2$, $\tau_i^2 = 0$, $i \neq j$ (назовем эту гипотезу H'_j) $j = 1, \dots, q$, а p_{11} — априорная вероятность гипотезы H'_{jh} , состоящей в том, что $\tau_j^2 = \tau_h^2 = \tau_{11}^2$, $\tau_i^2 = 0$, $i \neq j$, $i \neq h$, $j < h$, $j, h = 1, \dots, q$. Мы имеем $p_0 + qp_1 + q(q-1)p_{11}/2 = 1$. Поэтому апостериорные вероятности гипотез H_0 , H'_j и H'_{jh} соответственно пропорциональны величинам

$$(95) \quad p_0 \prod_{g=1}^q k(z_g),$$

$$(96) \quad p_1 e^{-\tau_1^2/2} I_0(\tau_1 \sqrt{z_j}) \prod_{g=1}^q k(z_g), \quad j = 1, \dots, q,$$

$$(97) \quad p_{11} e^{-\tau_{11}^2} I_0(\tau_{11} \sqrt{z_j}) I_0(\tau_{11} \sqrt{z_h}) \prod_{g=1}^q k(z_g), \quad j < h, \\ j, h = 1, \dots, q.$$

Следовательно, R_0 определяется неравенствами

$$(98) \quad I_0(\tau_1 \sqrt{z_j}) < \frac{p_0}{p_1} e^{\tau_1^2/2}, \quad j = 1, \dots, q,$$

$$(99) \quad I_0(\tau_{11} \sqrt{z_j}) I_0(\tau_{11} \sqrt{z_h}) < \frac{p_0}{p_{11}} e^{\tau_{11}^2}, \quad j < h, \quad j, h = 1, \dots, q.$$

Неравенства (98) эквивалентны неравенствам $z_j < c$, $j = 1, \dots, q$, если c определить соотношением

$$(100) \quad I_0(\tau_1 \sqrt{c}) = \frac{p_0}{p_1} e^{\tau_1^2/2}.$$

Поскольку $I_0(\tau_1 \sqrt{c}) \geq 1$, из (100) вытекает, что $p_0 e^{\tau_1^2/2} \geq p_1$. Если

$$(101) \quad I_0^2(\tau_{11} \sqrt{c}) \leq \frac{p_0}{p_{11}} e^{\tau_{11}^2},$$

то (99) следует из (98). (Если $\tau_1^2 = \tau_{11}^2$, то (101) принимает вид $p_0 p_{11} \leq p_1^2$). Так как $I_0(\tau_{11} \sqrt{z_h}) \geq 1$, из (99) вытекает

$$(102) \quad I_0(\tau_{11} \sqrt{z_j}) < \frac{p_0}{p_{11}} e^{\tau_{11}^2};$$

так что при выполнении

$$(103) \quad I_0(\tau_{11} \sqrt{c}) \geq \frac{p_0}{p_{11}} e^{\tau_{11}^2}$$

из (99) следует (98). (Если $\tau_I^2 = \tau_{II}^2$, то (103) есть просто $p_{II} \geq p_I e^{\tau_I^2/2}$.) Множества R_j определяются неравенствами $z_j > c$,

$$(104) \quad I_0(\tau_I \sqrt{z_j}) > I_0(\tau_I \sqrt{z_h}), \quad h \neq j, \quad h = 1, \dots, q,$$

$$(105) \quad p_I e^{-\tau_I^2/2} I_0(\tau_I \sqrt{z_j}) > p_{II} e^{-\tau_{II}^2} I_0(\tau_{II} \sqrt{z_i}) I_0(\tau_{II} \sqrt{z_h}),$$

$$i \neq h, \quad i, h = 1, \dots, q.$$

В силу определения они не пересекаются. Каждая точка на границе (попадание на границу имеет нулевую вероятность) может быть отнесена к любой области. Неравенства (104) эквивалентны неравенствам $z_j > z_h$, $h \neq j$, $h = 1, \dots, q$. Неравенство (105) при $i = j$ равносильно неравенству

$$(106) \quad \frac{I_0(\tau_I \sqrt{z_j})}{I_0(\tau_{II} \sqrt{z_j})} > \frac{p_{II}}{p_I} e^{\tau_I^2/2 - \tau_{II}^2} I_0(\tau_{II} \sqrt{z_h}).$$

Если $\tau_I^2 = \tau_{II}^2$, то (106) эквивалентно неравенству

$$(107) \quad I_0(\tau_I \sqrt{z_h}) < \frac{p_I}{p_{II}} e^{\tau_I^2/2},$$

которое в свою очередь равносильно неравенству $z_h < d$, где

$$(108) \quad I_0(\tau_I \sqrt{d}) = \frac{p_I}{p_{II}} e^{\tau_I^2/2}.$$

Если $d \leq c$ ($p_I^2 \leq p_0 p_{II}$), то из $z_j > c$ и $z_h < d$ вытекает неравенство $z_h < z_j$, которое совпадает с неравенством (104). Если $\tau_I^2 > \tau_{II}^2$, отношение $I_0(\tau_I \sqrt{z_j})/I_0(\tau_{II} \sqrt{z_j})$ монотонно возрастает, а ограничивающая его функция (106) имеет положительную производную dz_h/dz_j . Если же $\tau_I^2 < \tau_{II}^2$, отношение $I_0(\tau_I \sqrt{z_j})/I_0(\tau_{II} \sqrt{z_j})$ монотонно убывает и у ограничивающей его функции (106) производная dz_h/dz_j отрицательна.

Область R_{jh} , $j < h$, определяется неравенствами

$$(109) \quad I_0(\tau_{II} \sqrt{z_j}) I_0(\tau_{II} \sqrt{z_h}) > \frac{p_0}{p_{II}} e^{\tau_{II}^2},$$

$$(110) \quad p_{II} e^{-\tau_{II}^2} I_0(\tau_{II} \sqrt{z_j}) I_0(\tau_{II} \sqrt{z_h}) > p_I e^{-\tau_I^2/2} I_0(\tau_I \sqrt{z_i}),$$

$$i = 1, \dots, q,$$

и $z_j > z_i$, $z_h > z_i$, $i \neq j$, $i \neq h$, $i = 1, \dots, q$. Если $\tau_I^2 = \tau_{II}^2$, то (110) при $i = h$ равносильно неравенству $z_j > d$, а при $i = j$ — неравенству $z_h > d$.

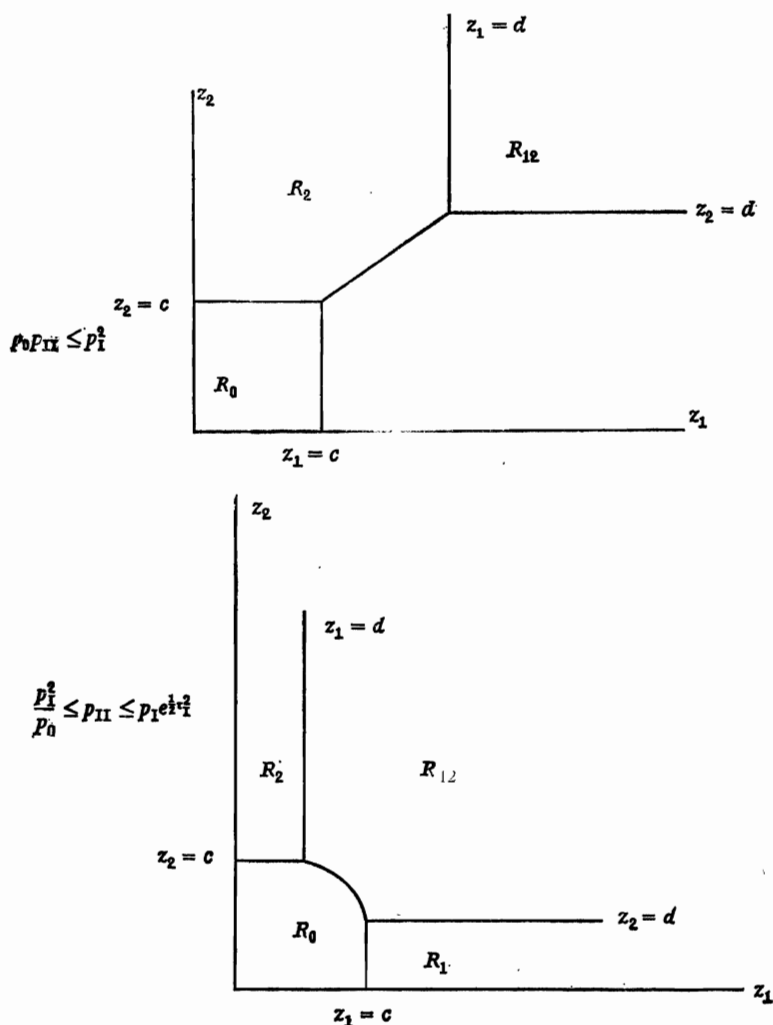


Рис. 4.3.

Области принятия гипотез H_0 , H'_1 , H'_2 , H'_{12} , когда $\tau_1^2 = \tau_{II}^2$.

На рис. 4.3 изображены R_0 , R_1 , R_2 и R_{12} при $\tau_1^2 = \tau_{II}^2$. Вид этих областей зависит от p_0 , p_1 , p_{II} и τ_1^2 . (При $\tau_1^2 = \tau_{II}^2$ знак неравенства в (103) невозможен.) При $\tau_1^2 > \tau_{II}^2$ прямые $z_1 = d$ и $z_2 = d$ заменяются кривыми с возрастающим наклоном, а при $\tau_1^2 < \tau_{II}^2$ кривыми с отрицательными тангенсами углов наклона,

Вид таких областей зависит от $p_0, p_1, p_{11}, \tau_1^2$ и τ_{11}^2 . Отметим, что не существует процедуры, оптимальной равномерно по τ_1^2 и τ_{11}^2 .

Указанные симметричные процедуры могут быть охарактеризованы вероятностями правильного решения $\Pr \{R_0 | H_0\}, \Pr \{R_1 | \tau_1^2 = \tau_1^2, \tau_i^2 = 0, i \geq 2\}$ и $\Pr \{R_{12} | \tau_1^2 = \tau_2^2 = \tau_{11}^2, \tau_i^2 = 0, i \geq 3\}$. Альтернативная постановка состоит в задании вероятностей $\Pr \{R_0 | H_0\}$ и $\Pr \{R_1 | \tau_1^2 = \tau_1^2, \tau_i^2 = 0, i \geq 2\}$ для некоторого заданного значения τ_1^2 и максимизации третьей вероятности для другого заданного значения τ_{11}^2 .

Если σ^2 неизвестно, можно потребовать, чтобы вероятность $\Pr \{R_0 | H_0\}$ не зависела от σ^2 . Это требование приводит, как и ранее, к процедурам, связанным с $\sum_{j=1}^r R^2(k_j)$. При каждом значении

$\sum_{j=1}^r z_j$ соответствующие области будут зависеть от $p_0, p_1, p_{11}, \tau_1^2$ и τ_{11}^2 . Бирнбаум (1959, 1961) использовал такой подход в задаче, частично совпадающей с нашей.

Другой подход [Уиттл (1952)] состоит в следующем. Гипотеза H_0 принимается, если $R^2(k_j) < g \sum_{i=1}^r R^2(k_i)$, $j = 1, \dots, q$, где g таково, что (92) равно ε . В противном случае принимается решение о том, что $\rho(k_j) > 0$, где $R^2(k_j) > g \sum_{i=1}^r R^2(k_i)$ и $R^2(k_j) > R^2(k_i)$, $i \neq j$. Затем остальные $r - 1$ значений $R^2(k_i)$ исследуются как бы в новой задаче того же рода, в которой только q заменяется на $q - 1$, а r на $r - 1$. Эта процедура может выполняться шаг за шагом, пока не будет решено, что остающиеся амплитуды $\rho(k_i)$ равны нулю. (Геометрическое представление о такой процедуре для $q = r = 3$ можно получить из упр. 28.)

С теоретической точки зрения указанная процедура не вполне удовлетворительна. Выбор определенного значения вероятности $\Pr \{R_0 | H_0\} = 1 - \varepsilon$ дает возможность контролировать вероятность одной из ошибок. Если только одна из амплитуд $\rho(k_j) > 0$, то вероятность правильного решения возрастает с ростом величины $\rho(k_j)$ (и это является удовлетворительным). Однако здесь допускается также решение о том, что положительными являются две или более амплитуд. При этом если положительными ровно две амплитуды и одна из них много больше другой, то вероятность решения о том, что эти две амплитуды положительны, в свою очередь может быть высокой. Если же две амплитуды положительны и близки друг другу, то вероятность того, что большая выборочная амплитуда будет столь велика по сравнению с соответствующей суммой, что

она обеспечит принятие решения об отклонении гипотезы H_0 , может оказаться и не очень большой.

Если гипотеза H_0 верна, то распределение значений x_j равномерно на множестве $x_j \geq 0$, $\sum_{j=1}^r x_j = 1$. Из геометрических соображений можно заметить, что распределение любых $r - 1$ величин x_j при условии, что их сумма равна единице минус величина не включенного в сумму x_j , превосходящего константу g , снова является равномерным (если g достаточно велико). Вероятность того, что отношение наибольшей из $(r - 1)$ величин x_j к их сумме будет больше определенного числа, может быть найдена подобным же образом. Однако если гипотеза H_0 неверна, то это распределение будет зависеть от значений $\rho^2(k_j)/\sigma^2$, отличных от нуля. Если какое-то из них отлично от нуля и столь велико, что велика и вероятность того, что соответствующее значение x_j является максимальным, то условное распределение других $r - 1$ величин x_j близко к равномерному. Возможны и другие случаи.

Отметим, что для использования процедуры теоремы 4.3.6 вовсе не обязательно отыскивать такое значение g , чтобы вероятность (92) равнялась требуемой, и затем определять, будет ли максимальное из наблюдаемых значений x_j больше этого числа. Вместо этого можно вычислить (92) для значения g , равного максимальному из x_j , и определить, будет ли вероятность (92) меньше требуемой.

4.3.5. Вычисление коэффициентов Фурье.

Быстрое преобразование Фурье

В большинстве ситуаций использование связей между тригонометрическими функциями может существенно упростить вычисление коэффициентов a_j и b_j . В качестве иллюстрации рассмотрим случай $T = T_1 T_2$. При этом можно записать $t = t_1 + (t_2 - 1) T_1$, $t_1 = 1, \dots, T_1$ и $t_2 = 1, \dots, T_2$, а также $j = j_1 + j_2 T_2$, $j_1 = 0, \dots, T_2 - 1$ и $j_2 = 0, \dots, T_1 - 1$. Отметим, что $a_{T-j} = a_j$ и $b_{T-j} = -b_j$. Искомые коэффициенты являются соответственно действительной и мнимой частями суммы

$$(111) \quad \sum_{t=1}^T y_t e^{i2\pi jt/T} = \sum_{t_1=1}^{T_1} \sum_{t_2=1}^{T_2} y_{t_1 + (t_2-1)T_1} \exp \left[i2\pi \left(\frac{j_1 t_1 + j_2 t_1 T_2}{T} + \frac{j_1 (t_2 - 1)}{T_2} \right) \right] =$$

$$= \sum_{t_1=1}^{T_1} \exp \left[i2\pi \left(\frac{j_1 + j_2 T_2}{T} \right) t_1 \right] \sum_{t_2=1}^{T_2} \exp \left[i2\pi \frac{j_1 (t_2 - 1)}{T_2} \right] y_{t_1 + (t_2 - 1) T_1}.$$

Для их отыскания вычислим сначала величины

$$(112) \quad c_{j_1, t_1} = \sum_{t_2=1}^{T_2} y_{t_1 + (t_2 - 1) T_1} \cos \frac{2\pi j_1}{T_2} (t_2 - 1),$$

$$(113) \quad s_{j_1, t_1} = \sum_{t_2=1}^{T_2} y_{t_1 + (t_2 - 1) T_1} \sin \frac{2\pi j_1}{T_2} (t_2 - 1),$$

для каждой пары (j_1, t_1) , $j_1 = 0, \dots, T_2 - 1$ и $t_1 = 1, \dots, T_1$. Поскольку $c_{T_2 - j_1, t_1} = c_{j_1, t_1}$ и $s_{T_2 - j_1, t_1} = -s_{j_1, t_1}$, то существует около $T_1 T_2 / 2 = T/2$ таких пар. После этого находятся действительная и мнимая части суммы (111):

$$(114) \quad \sum_{t_1=1}^{T_1} c_{j_1, t_1} \cos 2\pi \frac{j_1 + j_2 T_2}{T} t_1 - \sum_{t_1=1}^{T_1} s_{j_1, t_1} \sin 2\pi \frac{j_1 + j_2 T_2}{T} t_1,$$

$$(115) \quad \sum_{t_1=1}^{T_1} c_{j_1, t_1} \sin 2\pi \frac{j_1 + j_2 T_2}{T} t_1 + \sum_{t_1=1}^{T_1} s_{j_1, t_1} \cos 2\pi \frac{j_1 + j_2 T_2}{T} t_1$$

для $j = 0, 1, \dots, [T/2]$. Число операций умножения приблизительно равно $T_2 T$ в первых суммах и $2T_1 T$ во вторых суммах, так что число операций умножения в целом составляет примерно $T(T_2 + 2T_1)$. В то же время число произведений в определении коэффициентов a_j и b_j , $j = 0, 1, \dots, [T/2]$ примерно равно T^2 .

Для четных T существуют специальные формы вычислений. При $T = T_1 T_2 T_3$ можно использовать представление

$$(116) \quad t = t_1 + (t_2 - 1) T_1 + (t_3 - 1) T_1 T_2, \quad t_i = 1, \dots, T_i, \\ i = 1, 2, 3,$$

$$(117) \quad j = j_1 + j_2 T_3 + j_3 T_3 T_2, \quad j_1 = 0, 1, \dots, T_3 - 1, \\ j_2 = 0, \dots, T_2 - 1, \quad j_3 = 0, \dots, T_1 - 1.$$

При этом

$$(118) \quad \sum_{t=1}^T y_t e^{i2\pi j t / T} = \sum_{t_1=1}^{T_1} e^{i2\pi (j_1 + j_2 T_3 + j_3 T_3 T_2) t_1 / T} \times \\ \times \sum_{t_2=1}^{T_2} e^{i2\pi (j_1 + j_2 T_3) (t_2 - 1) / T_2} \sum_{t_3=1}^{T_3} e^{i2\pi j_1 (t_3 - 1) / T_3} y_{t_1 + (t_2 - 1) T_1 + (t_3 - 1) T_1 T_2}.$$

Мы не будем рассматривать эти методы более подробно. Детальное их описание приведено у Кули, Льюиса и Уэлша (1967). Сам метод был разработан Тьюки и его сотрудниками. [См. Кули и Тьюки (1965).] Идея метода восходит по-видимому к Рунге (1903).

4.4. СТАТИСТИЧЕСКИЕ ВЫВОДЫ ДЛЯ СЛУЧАЯ, КОГДА ПЕРИОДЫ ТРЕНДА НЕ ЯВЛЯЮТСЯ ДЕЛИТЕЛЯМИ ДЛИНЫ РЯДА

4.4.1. Определение тригонометрических коэффициентов и амплитуд для произвольных частот и их моменты

В предыдущем параграфе предполагалось, что изучаемый циклический тренд имеет периоды, нацело делящие длину ряда. Это предположение разумно в тех случаях, когда имеется предварительная информация о возможных периодах и длина ряда согласуется с этими периодами (так будет, например, если рассматривается сезонное изменение и имеются ежемесячные данные, накопленные за определенное число лет). Однако во многих случаях длина ряда определяется на основании каких-то факторов, никак не связанных с возможными периодами, а исследователь не имеет желания ограничиваться в своем рассмотрении только теми периодами, которые являются делителями длины ряда. Например, такова модель периодического тренда, предложенная для экономических рядов, обнаруживающих флуктуации из-за так называемого цикла деловой активности. Данные могут быть при этом ежегодными, а длина ряда определяется либо доступностью материала (например, тем, когда то или иное агентство начало собирать те или иные данные), либо внешними событиями (такими, как мировые войны). Цикл деловой активности может иметь периоды любой длины. Экономист может нуждаться в рассмотрении любых периодов, больших или равных 2. (В дальнейшем мы обсудим критику возможности применения указанной модели к экономическим временным рядам.)

В общей математической модели тренд представляется в виде суммы

$$(1) \quad f(t) = f_1(t) + \dots + f_m(t),$$

в которой $f_i(t)$ имеет (наименьший) период n_i . Однако такая модель является слишком общей для того, чтобы ее можно было изучать непосредственно на основе T наблюдений. Мы обсудим несколько более ограничительные формулировки. Рассмотрим простейший случай, когда $f(t)$ состоит из одной периодической и притом тригонометрической функции. Точнее, будем предполагать, что $y_t =$

$= f_t + u_t$, где $\mathcal{E}u_t = 0$, $\mathcal{E}u_t^2 = \sigma^2$, $\mathcal{E}u_t u_s = 0$, $t \neq s$ и

$$(2) \quad f(t) = \alpha \cos \lambda t + \beta \sin \lambda t = \rho \cos(\lambda t - \theta), \quad 0 \leq \lambda \leq \pi,$$

где

$$(3) \quad \rho^2 = \alpha^2 + \beta^2, \quad \operatorname{tg} \theta = \frac{\beta}{\alpha}.$$

Наблюдения $y_1, \dots, y_T$ можно взаимно однозначно заменить коэффициентами $a_0 = \bar{y}$,

$$(4) \quad a_j = \frac{2}{T} \sum_{t=1}^T y_t \cos \frac{2\pi j}{T} t, \quad j = 1, \dots, [(T-1)/2],$$

$$(5) \quad b_j = \frac{2}{T} \sum_{t=1}^T y_t \sin \frac{2\pi j}{T} t, \quad j = 1, \dots, [(T-1)/2],$$

$$(6) \quad a_{T/2} = \frac{1}{T} \sum_{t=1}^T y_t (-1)^t.$$

Последний коэффициент включается в рассмотрение, если T четное. Как уже было указано выше, это множество T коэффициентов есть непосредственный результат преобразования исходных данных. Дисперсии и ковариации этих коэффициентов равны, как и прежде,

$$\operatorname{Var}(a_0) = \frac{1}{T} \sigma^2 = \operatorname{Var}(a_{T/2}),$$

$$(7) \quad \operatorname{Var}(a_j) = \frac{2}{T} \sigma^2 = \operatorname{Var}(b_j), \quad j \neq 0, T/2,$$

$$\operatorname{Cov}(a_j, a_k) = 0 = \operatorname{Cov}(b_j, b_k), \quad j \neq k,$$

$$\operatorname{Cov}(a_j, b_k) = 0.$$

При желании можно рассмотреть величины

$$(8) \quad A(v) = \frac{2}{T} \sum_{t=1}^T y_t \cos vt, \quad 0 \leq v \leq \pi,$$

$$(9) \quad B(v) = \frac{2}{T} \sum_{t=1}^T y_t \sin vt, \quad 0 \leq v \leq \pi,$$

для всех v из отрезка $[0, \pi]$. При этом $a_j = A(2\pi j/T)$ и $b_j = B(2\pi j/T)$. Средние, дисперсии и ковариации этих коэффициентов являются суммами произведений тригонометрических функций.

ЛЕММА 4.4.1. Для $\nu + \lambda \neq 2\pi k$ и $\nu - \lambda \neq 2\pi k$, $k = 0, \pm 1, \dots$, выполняются соотношения

$$(10) \quad \sum_{t=1}^T \cos \lambda t \cos \nu t = \frac{1}{2} \frac{\sin \frac{1}{2} (\nu + \lambda) T}{\sin \frac{1}{2} (\nu + \lambda)} \cos \frac{1}{2} (\nu + \lambda) (T + 1) + \\ + \frac{1}{2} \frac{\sin \frac{1}{2} (\nu - \lambda) T}{\sin \frac{1}{2} (\nu - \lambda)} \cos \frac{1}{2} (\nu - \lambda) (T + 1),$$

$$(11) \quad \sum_{t=1}^T \cos \lambda t \sin \nu t = \frac{1}{2} \frac{\sin \frac{1}{2} (\nu + \lambda) T}{\sin \frac{1}{2} (\nu + \lambda)} \sin \frac{1}{2} (\nu + \lambda) (T + 1) + \\ + \frac{1}{2} \frac{\sin \frac{1}{2} (\nu - \lambda) T}{\sin \frac{1}{2} (\nu - \lambda)} \sin \frac{1}{2} (\nu - \lambda) (T + 1),$$

$$(12) \quad \sum_{t=1}^T \sin \lambda t \sin \nu t = -\frac{1}{2} \frac{\sin \frac{1}{2} (\nu + \lambda) T}{\sin \frac{1}{2} (\nu + \lambda)} \cos \frac{1}{2} (\nu + \lambda) (T + 1) + \\ + \frac{1}{2} \frac{\sin \frac{1}{2} (\nu - \lambda) T}{\sin \frac{1}{2} (\nu - \lambda)} \cos \frac{1}{2} (\nu - \lambda) (T + 1).$$

Для $\nu \neq \pi k$, $k = 0, \pm 1, \dots$, выполняются соотношения

$$(13) \quad \sum_{t=1}^T \cos^2 \nu t = \frac{1}{2} T + \frac{1}{2} \frac{\sin \nu T}{\sin \nu} \cos \nu (T + 1),$$

$$(14) \quad \sum_{t=1}^T \cos \nu t \sin \nu t = \frac{1}{2} \frac{\sin \nu T}{\sin \nu} \sin \nu (T + 1),$$

$$(15) \quad \sum_{t=1}^T \sin^2 \nu t = \frac{1}{2} T - \frac{1}{2} \frac{\sin \nu T}{\sin \nu} \cos \nu (T + 1).$$

Эти равенства получаются из рассмотрения сумм $\sum_{t=1}^T \cos \lambda t e^{i \nu t}$ для (10), (11), (13), (14) и $\sum_{t=1}^T \sin \nu t e^{i \lambda t}$ для (11), (12), (14), (15) с использованием метода (8) в § 4.2. (См. упр. 29.)

Средние, дисперсии и ковариации $A(v)$ и $B(v)$ приведены в теореме 4.4.1.

ТЕОРЕМА 4.4.1. Если $\mathcal{E}y_t = \alpha \cos \lambda t + \beta \sin \lambda t = \rho \cos(\lambda t - \theta)$, то

$$\begin{aligned}
 (16) \quad \mathcal{E}A(v) = \alpha(v) &= \frac{\alpha}{T} \left[\frac{\sin \frac{1}{2}(\lambda + v)T}{\sin \frac{1}{2}(\lambda + v)} \cos \frac{1}{2}(\lambda + v)(T+1) + \right. \\
 &\quad \left. + \frac{\sin \frac{1}{2}(\lambda - v)T}{\sin \frac{1}{2}(\lambda - v)} \cos \frac{1}{2}(\lambda - v)(T+1) \right] + \\
 &\quad + \frac{\beta}{T} \left[\frac{\sin \frac{1}{2}(\lambda + v)T}{\sin \frac{1}{2}(\lambda + v)} \sin \frac{1}{2}(\lambda + v)(T+1) + \right. \\
 &\quad \left. + \frac{\sin \frac{1}{2}(\lambda - v)T}{\sin \frac{1}{2}(\lambda - v)} \sin \frac{1}{2}(\lambda - v)(T+1) \right] = \\
 &= \frac{\rho}{T} \left\{ \frac{\sin \frac{1}{2}(\lambda + v)T}{\sin \frac{1}{2}(\lambda + v)} \cos \left[\frac{1}{2}(\lambda + v)(T+1) - \theta \right] + \right. \\
 &\quad \left. + \frac{\sin \frac{1}{2}(\lambda - v)T}{\sin \frac{1}{2}(\lambda - v)} \cos \left[\frac{1}{2}(\lambda - v)(T+1) - \theta \right] \right\}, \\
 &\quad v \neq \lambda, \quad 0 < v < \pi, \\
 &= \alpha + \frac{\rho}{T} \frac{\sin \lambda T}{\sin \lambda} \cos [\lambda(T+1) - \theta], \\
 &\quad 0 < \lambda = v < \pi,
 \end{aligned}$$

$$\begin{aligned}
 (17) \quad \mathcal{E}B(v) = \beta(v) &= \frac{\alpha}{T} \left[\frac{\sin \frac{1}{2}(\lambda + v)T}{\sin \frac{1}{2}(\lambda + v)} \sin \frac{1}{2}(\lambda + v)(T+1) - \right. \\
 &\quad \left. - \frac{\sin \frac{1}{2}(\lambda - v)T}{\sin \frac{1}{2}(\lambda - v)} \sin \frac{1}{2}(\lambda - v)(T+1) \right] +
 \end{aligned}$$

$$\begin{aligned}
& + \frac{\beta}{T} \left[- \frac{\sin \frac{1}{2} (\lambda + \nu) T}{\sin \frac{1}{2} (\lambda + \nu)} \cos \frac{1}{2} (\lambda + \nu) (T + 1) + \right. \\
& \left. + \frac{\sin \frac{1}{2} (\lambda - \nu) T}{\sin \frac{1}{2} (\lambda - \nu)} \cos \frac{1}{2} (\lambda - \nu) (T + 1) \right] = \\
& = \frac{\rho}{T} \left\{ \frac{\sin \frac{1}{2} (\lambda + \nu) T}{\sin \frac{1}{2} (\lambda + \nu)} \sin \left[\frac{1}{2} (\lambda + \nu) (T + 1) - \theta \right] - \right. \\
& \left. - \frac{\sin \frac{1}{2} (\lambda - \nu) T}{\sin \frac{1}{2} (\lambda - \nu)} \sin \left[\frac{1}{2} (\lambda - \nu) (T + 1) - \theta \right] \right\},
\end{aligned}$$

$$\nu \neq \lambda, \quad 0 < \nu < \pi,$$

$$= \beta + \frac{\rho}{T} \frac{\sin \lambda T}{\sin \lambda} \sin [\lambda (T + 1) - \theta],$$

$$0 < \lambda = \nu < \pi,$$

$$(18) \quad \text{Var} [A(\nu)] = \frac{2\sigma^2}{T} \left[1 + \frac{\sin \nu T}{T \sin \nu} \cos \nu (T + 1) \right],$$

$$0 < \nu < \pi,$$

$$(19) \quad \text{Var} [B(\nu)] = \frac{2\sigma^2}{T} \left[1 - \frac{\sin \nu T}{T \sin \nu} \cos \nu (T + 1) \right],$$

$$0 < \nu < \pi,$$

$$(20) \quad \text{Cov} [A(\nu), B(\nu)] = \frac{2\sigma^2}{T} \frac{\sin \nu T}{T \sin \nu} \sin \nu (T + 1),$$

$$0 < \nu < \pi.$$

Если $\nu = 2\pi j/T$, то соответствующие дисперсии и ковариации сводятся к (7). Математические ожидания также принимают более простой вид, поскольку $\sin(x \pm \pi j) = (-1)^j \sin x$ и $\cos(x \pm \pi j) = (-1)^j \cos x$.

Следствие 4.4.1.

$$(21) \quad \mathcal{E}A\left(\frac{2\pi j}{T}\right) = \mathcal{E}a_i = \alpha\left(\frac{2\pi j}{T}\right) =$$

$$\begin{aligned}
&= \frac{\rho}{T} \sin \frac{\lambda T}{2} \left\{ \frac{\cos \left[\frac{1}{2} \lambda (T+1) + \frac{\pi j}{T} - \theta \right]}{\sin \frac{1}{2} \left(\lambda + \frac{2\pi j}{T} \right)} + \right. \\
&\quad \left. + \frac{\cos \left[\frac{1}{2} \lambda (T+1) - \frac{\pi j}{T} - \theta \right]}{\sin \frac{1}{2} \left(\lambda - \frac{2\pi j}{T} \right)} \right\}, \\
&\quad \lambda \neq \frac{2\pi j}{T}, \quad j = 1, \dots, \lfloor (T-1)/2 \rfloor, \\
&= \alpha, \quad \lambda = \frac{2\pi j}{T}, \quad j = 1, \dots, \lfloor (T-1)/2 \rfloor.
\end{aligned}$$

$$\begin{aligned}
(22) \quad \mathfrak{E}B\left(\frac{2\pi j}{T}\right) &= \mathfrak{E}b_j = \beta\left(\frac{2\pi j}{T}\right) = \\
&= \frac{\rho}{T} \sin \frac{\lambda T}{2} \left\{ \frac{\sin \left[\frac{1}{2} \lambda (T+1) + \frac{\pi j}{T} - \theta \right]}{\sin \frac{1}{2} \left(\lambda + \frac{2\pi j}{T} \right)} - \right. \\
&\quad \left. - \frac{\sin \left[\frac{1}{2} \lambda (T+1) - \frac{\pi j}{T} - \theta \right]}{\sin \frac{1}{2} \left(\lambda - \frac{2\pi j}{T} \right)} \right\}, \\
&\quad \lambda \neq \frac{2\pi j}{T}, \quad j = 1, \dots, \lfloor (T-1)/2 \rfloor, \\
&= \beta, \quad \lambda = \frac{2\pi j}{T}, \quad j = 1, \dots, \lfloor (T-1)/2 \rfloor.
\end{aligned}$$

Отметим далее, что

$$(23) \quad \frac{1}{2} \mathfrak{E}A(0) = \mathfrak{E}a_0 = \alpha_0 = \frac{\rho}{T} \frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} \lambda} \cos \left[\frac{1}{2} \lambda (T+1) - \theta \right],$$

$\lambda \neq 0,$

и $\frac{1}{2} \mathfrak{E}A(0) = \mathfrak{E}a_0 = \alpha$ для $\lambda = 0$. Кроме того,

$$\begin{aligned}
(24) \quad \frac{1}{2} \mathfrak{E}A(\pi) &= \mathfrak{E}a_{T/2} = \alpha_{T/2} = \frac{\rho}{T} \frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} (\lambda + \pi)} \cos \left[\frac{1}{2} \lambda (T+1) + \right. \\
&\quad \left. + \frac{\pi}{2} - \theta \right], \quad \lambda \neq \pi,
\end{aligned}$$

для четного T и $\frac{1}{2}\mathcal{E}A(\pi) = \mathcal{E}a_{T/2} = \alpha$ для $\lambda = \pi$. Мы будем предполагать, что $0 \leq \lambda \leq \pi$. Дело в том, что если $\pi < \lambda < 2\pi$, то вместо λ можно рассмотреть $\lambda^* = 2\pi - \lambda$, поскольку

$$\begin{aligned} (25) \quad \alpha \cos \lambda t + \beta \sin \lambda t &= \alpha \cos [2\pi t - (2\pi - \lambda)t] + \beta \sin [2\pi t - \\ &\quad - (2\pi - \lambda)t] = \\ &= \alpha \cos [-\lambda^*t] + \beta \sin [-\lambda^*t] = \\ &= \alpha \cos \lambda^*t + (-\beta) \sin \lambda^*t. \end{aligned}$$

Выборочная функция интенсивности $R^2(v)$ определяется равенством

$$\begin{aligned} (26) \quad R^2(v) &= A^2(v) + B^2(v) = \\ &= \frac{4}{T^2} \sum_{s,t=1}^T y_s y_t (\cos vs \cos vt + \sin vs \sin vt) = \\ &= \frac{4}{T^2} \sum_{s,t=1}^T y_s y_t \cos v(s-t) = \\ &= \left| \frac{2}{T} \sum_{t=1}^T y_t e^{ivt} \right|^2, \quad 0 \leq v \leq \pi. \end{aligned}$$

Если $v = 2\pi j/T$, то $R^2(2\pi j/T) = R_j^2 = a_j^2 + b_j^2$, $j = 1, \dots, [(T-1)/2]$. Теоретической величиной, соответствующей $R^2(v)$, является

$$\begin{aligned} (27) \quad \rho^2(v) &= \alpha^2(v) + \beta^2(v) = \\ &= \frac{\rho^2}{T^2} \left\{ \frac{\sin^2 \frac{1}{2}(\lambda+v)T}{\sin^2 \frac{1}{2}(\lambda+v)} + \frac{\sin^2 \frac{1}{2}(\lambda-v)T}{\sin^2 \frac{1}{2}(\lambda-v)} + \right. \\ &\quad \left. + 2 \frac{\sin \frac{1}{2}(\lambda+v)T}{\sin \frac{1}{2}(\lambda+v)} \frac{\sin \frac{1}{2}(\lambda-v)T}{\sin \frac{1}{2}(\lambda-v)} \cos [\lambda(T+1) - 2\theta] \right\}, \\ &\quad v \neq \lambda, \quad 0 < v < \pi, \\ &= \rho^2 \left\{ 1 + \frac{\sin^2 \lambda T}{T^2 \sin^2 \lambda} + 2 \frac{\sin \lambda T}{T \sin \lambda} \cos [\lambda(T+1) - 2\theta] \right\}, \\ &\quad 0 < \lambda = v < \pi. \end{aligned}$$

При $v = 2\pi j/T$ (27) обращается в

$$(28) \quad \rho^2\left(\frac{2\pi j}{T}\right) = \rho_j^2 = \frac{\rho^2 \sin^2 \frac{1}{2} \lambda T}{T^2} \left\{ \frac{1}{\sin^2 \frac{1}{2} \left(\lambda + \frac{2\pi j}{T}\right)} + \right. \\ \left. + \frac{1}{\sin^2 \frac{1}{2} \left(\lambda - \frac{2\pi j}{T}\right)} + 2 \frac{\cos [\lambda (T+1) - 2\theta]}{\sin \frac{1}{2} \left(\lambda + \frac{2\pi j}{T}\right) \sin \frac{1}{2} \left(\lambda - \frac{2\pi j}{T}\right)} \right\}, \\ \lambda \neq \frac{2\pi j}{T}, \quad j = 1, \dots, [(T-1)/2], \\ = \rho^2, \quad \lambda = \frac{2\pi j}{T}, \quad j = 1, \dots, [(T-1)/2].$$

Заметим следующее. Пусть тригонометрический тренд имеет период, являющийся делителем длины ряда, т. е. $\lambda = 2\pi j/T$. Тогда математические ожидания выборочных тригонометрических коэффициентов A ($2\pi k/T$), B ($2\pi k/T$), $k = 1, \dots, [(T-1)/2]$, $k \neq j$, имеющих периоды, отличные от периода тренда, но являющиеся делителями длины ряда, равны нулю. Сумма квадратов этих математических ожиданий, т. е. квадрат теоретической амплитуды ρ_k^2 , также равен нулю в этом случае. Если $k = j$, то квадрат теоретической амплитуды равен одновременно квадрату амплитуды тренда ρ . Кроме того, $2\alpha_0^2 = 0$, и если T четное, то и $2\alpha_{T/2}^2 = 0$. (Если $\lambda = 0$, то $\rho_k^2 = 0$, $k = 1, \dots, [(T-1)/2]$, $2\alpha_0^2 = 2\alpha^2$ и, если T четное, $2\alpha_{T/2}^2 = 0$. Если $\lambda = \pi$ при четном T , то $\rho_k^2 = 0$, $k = 1, \dots, T/2-1$, $2\alpha_0^2 = 0$ и $2\alpha_{T/2}^2 = 2\alpha^2$.) Вообще отличными от нуля будут математические ожидания тех выборочных тригонометрических коэффициентов, периоды которых не являются делителями длины ряда ($v \neq 2\pi k/T$, $k = 1, \dots, [(T-1)/2]$). При этом будут отличны от нуля и соответствующие $\rho^2(v)$. Однако если период тренда не является делителем длины ряда, $\lambda \neq 2\pi j/T$, $j = 1, \dots, [(T-1)/2]$, то будут отличны от нуля математические ожидания и всех тех выборочных тригонометрических коэффициентов, периоды которых являются ее делителями, $v = 2\pi k/T$, $k = 1, \dots, [(T-1)/2]$, а также соответствующие теоретические амплитуды. Резюмируем: совокупность теоретических амплитуд $\rho_1^2, \dots, \rho_{[(T-1)/2]}^2$ обладает следующим специфическим свойством. Случай, когда ровно одна амплитуда равна ρ^2 , а остальные амплитуды равны 0, может наблюдаться, только если период тренда равен одному из значений: $T, T/2, \dots, T/[(T-1)/2]$.

В предположении, что величины y_t имеют совместное нормальное распределение, квадратичная форма ($0 < v < \pi$)

$$\begin{aligned}
 (29) \quad Q(v) &= \frac{T}{2\sigma^2} [A(v)B(v)] \times \\
 &\times \begin{pmatrix} 1 + \frac{\sin vT}{T \sin v} \cos v(T+1) & \frac{\sin vT}{T \sin v} \sin v(T+1) \\ \frac{\sin vT}{T \sin v} \sin v(T+1) & 1 - \frac{\sin vT}{T \sin v} \cos v(T+1) \end{pmatrix}^{-1} \begin{pmatrix} A(v) \\ B(v) \end{pmatrix} = \\
 &= \frac{T}{2\sigma^2 \left(1 - \frac{\sin^2 vT}{T^2 \sin^2 v}\right)} \{ A^2(v) + B^2(v) + \\
 &+ \frac{\sin vT}{T \sin v} [\cos v(T+1) [B^2(v) - A^2(v)] - \\
 &- 2 \sin v(T+1) A(v)B(v)] \}
 \end{aligned}$$

имеет нецентральное χ^2 -распределение с 2 степенями свободы и параметром нецентральности

$$\begin{aligned}
 (30) \quad &\frac{T}{2\sigma^2} [\alpha(v)\beta(v)] \times \\
 &\times \begin{pmatrix} 1 + \frac{\sin vT}{T \sin v} \cos v(T+1) & \frac{\sin vT}{T \sin v} \sin v(T+1) \\ \frac{\sin vT}{T \sin v} \sin v(T+1) & 1 - \frac{\sin vT}{T \sin v} \cos v(T+1) \end{pmatrix}^{-1} \begin{pmatrix} \alpha(v) \\ \beta(v) \end{pmatrix} = \\
 &= \frac{\rho^2}{2T\sigma^2 \left(1 - \frac{\sin^2 vT}{T^2 \sin^2 v}\right)} \left\{ \left[\frac{\sin^2 \frac{1}{2}(\lambda+v)T}{\sin^2 \frac{1}{2}(\lambda+v)} + \frac{\sin^2 \frac{1}{2}(\lambda-v)T}{\sin^2 \frac{1}{2}(\lambda-v)} \right] \times \right. \\
 &\times \left[1 - \frac{\sin vT}{T \sin v} \cos [\lambda(T+1) - 2\theta] \right] + \\
 &+ 2 \frac{\sin \frac{1}{2}(\lambda+v)T \sin \frac{1}{2}(\lambda-v)T}{\sin \frac{1}{2}(\lambda+v) \sin \frac{1}{2}(\lambda-v)} \left[\cos [\lambda(T+1) - 2\theta] - \right. \\
 &\left. \left. - \frac{\sin vT}{T \sin v} \right] \right\}, \quad v \neq \lambda, \quad 0 < v < \pi, \\
 &= \frac{T\rho^2}{2\sigma^2} \left\{ 1 + \frac{\sin \lambda T}{T \sin \lambda} \cos [\lambda(T+1) - 2\theta] \right\}, \quad 0 < \lambda = v < \pi.
 \end{aligned}$$

Если $\nu = 2\pi j/T$, то (29) будет равно $TR_j^2/(2\sigma^2)$, а (30) обратится в $Tr_j^2/(2\sigma^2)$, $j = 1, \dots, [(T-1)/2]$. Распределенные по (центральному или нецентральному) закону χ^2 величины, кратные $R^2(2\pi/T), \dots, R^2(2\pi[(T-1)/2]/T)$ независимы. Однако выборочные амплитуды для $\nu \neq 2\pi j/T$, $j = 1, \dots, [(T-1)/2]$, будут, вообще говоря, взаимно зависимыми.

4.4.2. Решающие процедуры, основанные на выборочных амплитудах, для периодов, являющихся делителями длины ряда

Рассмотрим теперь использование выборочных интенсивностей при $\nu = 2\pi/T, 4\pi/T, \dots, 2\pi[(T-1)/2]/T$. Если величины нормально распределены, то $R_1^2, \dots, R_{[(T-1)/2]}^2$ независимы и $TR_j^2/(2\sigma^2)$ имеет χ^2 -распределение с параметром нецентральности $Tr_j^2/(2\sigma^2)$, приведенное в (28).

Мы уже упоминали о том, что если $\lambda = 2\pi k/T$, где k — целое число, заключенное между 1 и $[(T-1)/2]$, то $\rho_k^2 = \rho^2$ и $\rho_j^2 = 0$ при $j \neq k$. Если же это не так (т. е. $\lambda T/(2\pi)$ нецелое), то $\rho_j^2 > 0$ для всех j . Иными словами, если период $2\pi/\lambda$ функции $f(t)$ не является периодом ни одной из ортогональных функций, табулированных столбцами в матрице $\mathbf{M}$, то все интенсивности, соответствующие периодам указанных функций, положительны. При этом можно ожидать, что интенсивность для $2\pi j$, близкого к λT , велика. Покажем, что если $2\pi k/T < \lambda < 2\pi(k+1)/T$, то тогда наибольшей интенсивностью является либо ρ_k^2 , либо ρ_{k+1}^2 .

Лемма 4.4.2. Если $|K| \leq 1$, то функция

$$(31) \quad h(x) = \frac{1}{\sin^2 \frac{1}{2}(\lambda + x)} + \frac{1}{\sin^2 \frac{1}{2}(\lambda - x)} + \\ + 2 \frac{K}{\sin \frac{1}{2}(\lambda + x) \sin \frac{1}{2}(\lambda - x)}$$

монотонно возрастает при $0 < x < \lambda$ и монотонно убывает при $\lambda < x < \pi$.

Доказательство. Запишем функцию $h(x)$ в виде

$$(32) \quad h(x) = \left[\frac{1}{\sin \frac{1}{2}(\lambda + x)} - \frac{1}{\sin \frac{1}{2}(\lambda - x)} \right]^2 +$$

$$\begin{aligned}
& + \frac{2(K+1)}{\sin \frac{1}{2}(\lambda+x) \sin \frac{1}{2}(\lambda-x)} = \\
& = \left[\frac{\sin \frac{1}{2}(\lambda-x) - \sin \frac{1}{2}(\lambda+x)}{\sin \frac{1}{2}(\lambda+x) \sin \frac{1}{2}(\lambda-x)} \right]^2 + \frac{2(K+1)}{\sin \frac{1}{2}(\lambda+x) \sin \frac{1}{2}(\lambda-x)} = \\
& = \left[\frac{-4 \cos \frac{1}{2}\lambda \sin \frac{1}{2}x}{\cos x - \cos \lambda} \right]^2 + \frac{4(K+1)}{\cos x - \cos \lambda} = \\
& = 4 \cos^2 \frac{1}{2}\lambda \left[\frac{\sin \frac{1}{2}x}{\sin^2 \frac{1}{2}\lambda - \sin^2 \frac{1}{2}x} \right]^2 + \frac{4(K+1)}{\cos x - \cos \lambda}.
\end{aligned}$$

Оба слагаемых в правой части возрастают с ростом x в интервале $(0, \lambda)$. Можно записать также, что

$$\begin{aligned}
(33) \quad h(x) &= \left[\frac{1}{\sin \frac{1}{2}(\lambda+x)} + \frac{1}{\sin \frac{1}{2}(\lambda-x)} \right]^2 - \\
&\quad - \frac{2(1-K)}{\sin \frac{1}{2}(\lambda+x) \sin \frac{1}{2}(\lambda-x)} = \\
&= \left[\frac{4 \sin \frac{1}{2}\lambda \cos \frac{1}{2}x}{\cos x - \cos \lambda} \right]^2 - \frac{4(1-K)}{\cos x - \cos \lambda} = \\
&= 4 \sin^2 \frac{1}{2}\lambda \left[\frac{\cos \frac{x}{2}}{\cos^2 \frac{x}{2} - \cos^2 \frac{\lambda}{2}} \right]^2 + \frac{4(1-K)}{\cos \lambda - \cos x}.
\end{aligned}$$

Здесь оба слагаемых в правой части убывают с ростом x в интервале (λ, π) . ■

Поскольку при $K = \cos[\lambda(T+1) - 2\theta]$

$$(34) \quad \rho_j^2 = \frac{\rho^2}{T^2} \sin^2 \frac{1}{2} \lambda T h \left(\frac{2\pi j}{T} \right),$$

то, согласно лемме, ρ_j^2 возрастает при $j = 1, \dots, k$ и убывает при $j = k+1, \dots, [(T-1)/2]$, если $2\pi k < \lambda T < 2\pi(k+1)$. Отсюда вытекает

ТЕОРЕМА 4.4.2. Если $2\pi k/T < \lambda < 2\pi(k+1)/T$ и ρ_l — (положительный) квадратный корень из величины (28), то

$$(35) \quad \begin{aligned} \rho_j < \rho_{j+1}, \quad j = 1, \dots, k-1, \\ \rho_j < \rho_{j-1}, \quad j = k+2, \dots, [(T-1)/2]. \end{aligned}$$

В соответствии с полученным результатом распределение с наибольшим параметром нецентральности имеет либо R_k^2 , либо R_{k+1}^2 . В некотором смысле худшим при данном k будет такое λ , $2\pi k/T < \lambda < 2\pi(k+1)/T$, при котором большая из величин ρ_k и ρ_{k+1} становится минимальной. Однако это значение λ и соответствующее ему наибольшее из чисел ρ_k и ρ_{k+1} являются весьма сложными функциями параметров k , T и θ .

Рассмотрим теперь аппроксимации ρ_k и ρ_{k+1} . Если $2\pi k/T < \lambda < 2\pi(k+1)/T$, то аргументы $\sin(\lambda/2 - k\pi/T)$ и $\sin[\lambda/2 - (k+1)\pi/T]$ по абсолютной величине меньше π/T . Поэтому при достаточно больших T первый член разложения

$$(36) \quad \sin x = x - \frac{x^3}{6} + \frac{x^5}{120} - \dots$$

дает достаточно хорошую аппроксимацию $\sin x$ для $x = \lambda/2 - k\pi/T$ и $x = \lambda/2 - (k+1)\pi/T$. Положим $\lambda = 2\pi(k+\varepsilon)/T$ ($0 < \varepsilon < 1$). Тогда

$$(37) \quad \begin{aligned} \rho_k^2 = \frac{\rho^2}{T^2} \sin^2 \pi \varepsilon \left\{ \frac{1}{\sin^2 \pi \frac{2k+\varepsilon}{T}} + \frac{1}{\sin^2 \pi \frac{\pi \varepsilon}{T}} + \right. \\ \left. + \frac{2 \cos [2\pi (T\varepsilon + \varepsilon + k)/T - 2\theta]}{\sin \pi \frac{2k+\varepsilon}{T} \sin \pi \frac{\varepsilon}{T}} \right\}, \end{aligned}$$

$$\begin{aligned} \rho_{k+1}^2 = \frac{\rho^2}{T^2} \sin^2 \pi \varepsilon \left\{ \frac{1}{\sin^2 \pi \frac{2k+1+\varepsilon}{T}} + \frac{1}{\sin^2 \pi \frac{\varepsilon-1}{T}} + \right. \\ \left. + \frac{2 \cos [2\pi (T\varepsilon + \varepsilon + k)/T - 2\theta]}{\sin \pi \frac{2k+1+\varepsilon}{T} \sin \pi \frac{\varepsilon-1}{T}} \right\}. \end{aligned}$$

Поскольку $0 < \sin \pi \varepsilon/T < \sin [\pi (2k+\varepsilon)/T]$ и $0 < \sin [\pi \times (1-\varepsilon)/T] < \sin [\pi (2k+1+\varepsilon)/T]$ ($k+1 < T/2$), то вторые члены в ρ_k^2 и ρ_{k+1}^2 соответственно больше первых и больше половин абсолютных значений третьих. То, в какой степени домини-

рует второй член, зависит в первом случае от отношения

$$(38) \quad \frac{\sin \pi \frac{2k + \varepsilon}{T}}{\sin \pi \frac{\varepsilon}{T}} = \frac{\pi \frac{2k + \varepsilon}{T} - \frac{1}{6} \left(\pi \frac{2k + \varepsilon}{T} \right)^3 + \dots}{\pi \frac{\varepsilon}{T} - \frac{1}{6} \left(\pi \frac{\varepsilon}{T} \right)^3 + \dots} =$$

$$= \frac{2k + \varepsilon}{\varepsilon} \cdot \frac{1 - \pi^2 \frac{(2k + \varepsilon)^2}{6T^2} + \dots}{1 - \frac{\pi^2 \varepsilon^2}{6T^2} + \dots},$$

равного примерно $2k/\varepsilon + 1$. За исключением случая, когда k очень мало (маленькая частота и большой период), это отношение велико и второе слагаемое в выражении для ρ_k^2 доминирует. Если k близко к $[(T - 1)/2]$, то первое слагаемое можно исследовать более точно. Аргумент $\pi(2k + \varepsilon)/T$ заменяется на $\pi - \pi \times (2k + \varepsilon)/T$ и (38) аппроксимируется посредством $(T - 2k - \varepsilon)/\varepsilon$.

Анализ вторых слагаемых в выражениях для ρ_k^2 и ρ_{k+1}^2 показывает, что большее из них достигает минимума при $\varepsilon = 1/2$. Соответствующей аппроксимацией для ρ_k^2 и ρ_{k+1}^2 является

$$(39) \quad \frac{\rho^2}{T^2} \cdot \frac{\sin^2 \frac{\pi}{2}}{\sin^2 \frac{\pi}{2T}} = \frac{\rho^2}{T^2 \left[\frac{\pi}{2T} - \frac{1}{6} \left(\frac{\pi}{2T} \right)^3 + \dots \right]^2} =$$

$$= \rho^2 \left(\frac{2}{\pi} \right)^2 \frac{1}{\left[1 - \frac{1}{6} \left(\frac{\pi}{2T} \right)^2 + \dots \right]^2}.$$

Если T достаточно велико, то (39) близко к $\rho^2 (2/\pi)^2 = \rho^2 (0.6366)^2 = \rho^2 \times 0.4053$. Иначе говоря, каждое из ρ_k^2 и ρ_{k+1}^2 составляет около 41% интенсивности гармонической составляющей, а сумма $\rho_k^2 + \rho_{k+1}^2$ составляет около 81%. Как будет видно из следующего абзаца, значения остальных ρ_j^2 необходимо должны быть малыми. (Если ε меньше $1/2$, то второе слагаемое в ρ_k^2 приближенно равно $\rho^2 (1 - \pi^2 \varepsilon^2 / 6)^2$.)

Полная сумма квадратов значений $\mathcal{E}y_t = f(t)$ равна

$$(40) \quad \sum_{t=1}^T f^2(t) = \alpha^2 \sum_{t=1}^T \cos^2 \lambda t + 2\alpha\beta \sum_{t=1}^T \cos \lambda t \sin \lambda t + \beta^2 \sum_{t=1}^T \sin^2 \lambda t =$$

$$= \begin{cases} \rho^2 \left\{ \frac{T}{2} + \frac{\sin \lambda T}{2 \sin \lambda} \cos [\lambda(T + 1) - 2\theta] \right\}, & 0 < \lambda < \pi, \\ T\alpha^2, & \lambda = 0, \pi. \end{cases}$$

Ее можно нормировать делением на $T/2$. Если $\lambda = 2\pi k/T$ или если $\lambda = 2\pi (k + 1/2)/T$, то второе слагаемое в правой части (40) равно 0 и

$$(41) \quad \frac{2}{T} \sum_{i=1}^T f^2(t) = \rho^2, \quad \lambda = \frac{2\pi k}{T}, \quad \lambda = \frac{2\pi (k + 1/2)}{T}.$$

Если $\lambda = 2\pi (k + \varepsilon)/T$, то последняя сумма равна

$$(42) \quad \frac{2}{T} \sum_{i=1}^T f^2(t) = \rho^2 \left\{ 1 + \frac{\sin 2\pi\varepsilon}{T \sin 2\pi \frac{(k + \varepsilon)}{T}} \cos \left[2\pi \left(\varepsilon + \frac{k + \varepsilon}{T} \right) - 2\theta \right] \right\}.$$

Второй член в (42) будет мал, в частности если k велико. Заметим, что в обоих случаях

$$(43) \quad \frac{2}{T} \sum_{i=1}^T f^2(t) = \sum_{j=1}^{[(T-1)/2]} \rho_j^2 + 2\alpha_0^2 + 2\alpha_{T/2}^2,$$

причем, если T нечетное, последнее слагаемое в правой части должно отсутствовать. Коэффициенты α_0 и $\alpha_{T/2}$ задаются соотношениями (23) и (24). Они обычно бывают малы (см. упр. 34 и 35) и вносят малый вклад в сумму квадратов (43). Отсюда следует, что нормированная сумма квадратов математических ожиданий наблюдаемых значений близка к сумме квадратов $[(T-1)/2]$ интенсивностей. Часть последней суммы, не входящая в наибольшую или в следующую за ней по величине амплитуду ρ_j^2 , относится к остальным ρ_j^2 . Предыдущие рассуждения показывают, что наибольшая из этих ρ_j^2 составляет около 41% указанной суммы квадратов.

Отметим, что сумма квадратов (40), деленная на σ^2 , является параметром нецентральности (30) при $v = \lambda$ ($\lambda \neq 0, \pi$). Статистика $Q(\lambda)$ является квадратичной формой относительно переменных $A(\lambda)$ и $B(\lambda)$. Ее можно рассматривать как сумму квадратов двух нормированных ортогональных линейных комбинаций значений $y_1, \dots, y_T$. Сумма квадратов математических ожиданий этих линейных комбинаций равна сумме квадратов математических ожиданий значений всех наблюдений. Это показывает, что параметр нецентральности распределения статистики $Q(v)$ принимает максимальное значение при $v = \lambda$.

Процедура, использующая статистики $R_1^2, \dots, R_q^2$ ($q = [(T-1)/2]$), состоит в следующем. Гипотеза $H_0: \rho^2 = 0$ принимается, если $R_j^2 \leq g \sum_{i=1}^q R_i^2$, $j = 1, \dots, q$, где g выбирается так,

чтобы вероятность (92) из § 4.3 была равна определенному уровню значимости ε . В противном случае принимается гипотеза H_j : $\rho^2 > 0$ и $\rho_j^2 > \rho_i^2$, $i \neq j$, $i = 1, \dots, q$, где $R_j^2 > g \sum_{i=1}^q R_i^2$ и $R_j^2 > R_i^2$, $i \neq j$, $i = 1, \dots, q$. Гипотеза H_j означает (теорема 4.4.2), что λ лежит в пределах между $2\pi/T$ и $2\pi j/T$ и находится в непосредственной близости к интервалу $(2\pi(j-1/2)/T, 2\pi(j+1/2)/T)$. Используя величины $z_j = TR_j^2/(2\sigma^2)$, области принятия указанных гипотез можно записать в виде

$$(44) \quad R_0: z_j \leq \frac{g}{1-g} \sum_{i \neq j} z_i, \quad j = 1, \dots, q,$$

$$(45) \quad R_j: z_j > \frac{g}{1-g} \sum_{i \neq j} z_i, \quad z_j > z_i, \quad i \neq j, \quad i = 1, \dots, q, \\ j = 1, \dots, q.$$

Покажем теперь, что если верна гипотеза H_j , то вероятность прийти к заключению о ее истинности больше вероятности решения о том, что верна какая-либо другая гипотеза H_i и что эта вероятность достигает максимума при $\lambda = 2\pi j/T$ (в последнем случае $\rho_i^2 = 0$ для $i \neq j$). Пусть $\tau_j^2 = TR_j^2/(2\sigma^2)$. Предположим, что наблюдения y_i нормально распределены.

ЛЕММА 4.4.3. Интеграл

$$(46) \quad \int_c^\infty k(z|\tau^2) dz = \int_c^\infty e^{-\tau^2/2} k(z) I_0(\tau\sqrt{z}) dz$$

является возрастающей функцией от τ^2 ($0 \leq \tau^2 < \infty$).

Доказательство. Если $\tau_1^2 < \tau_2^2$, то неравенство $z \geq c$ равносильно неравенству $k(z|\tau_2^2)/k(z|\tau_1^2) > d$ при надлежащим образом выбранном d . Тогда имеем

$$(47) \quad \int_c^\infty k(z|\tau_2^2) dz > d \int_c^\infty k(z|\tau_1^2) dz,$$

$$(48) \quad 1 - \int_c^\infty k(z|\tau_2^2) dz = \int_0^c k(z|\tau_2^2) dz < d \int_0^c k(z|\tau_1^2) dz = \\ = d \left[1 - \int_c^\infty k(z|\tau_1^2) dz \right].$$

Отсюда и следует утверждение леммы. ■

ЛЕММА 4.4.4. Вероятность $\Pr \{R_1 | \tau_1^2, \tau_2^2, \dots, \tau_q^2\}$ является возрастающей функцией от τ_1^2 при фиксированных значениях $\tau_2^2, \dots, \tau_q^2$ и убывающей функцией от τ_2^2 при фиксированных значениях $\tau_1^2, \tau_3^2, \dots, \tau_q^2$.

ДОКАЗАТЕЛЬСТВО. Пусть $h(z_2, \dots, z_q)$ — максимальное из чисел $z_2, \dots, z_q$ и $[g/(1-g)] \sum_{i=2}^q z_i$. Тогда

$$(49) \quad \Pr \{R_1 | \tau_1^2, \tau_2^2, \dots, \tau_q^2\} = \int_0^\infty \dots \int_0^\infty \prod_{i=2}^q k(z_i | \tau_i^2) \left\{ \int_{h(z_2, \dots, z_q)}^\infty k(z_1 | \tau_1^2) dz_1 \right\} \prod_{i=2}^q dz_i.$$

Интеграл в фигурных скобках является в силу леммы 4.4.3 возрастающей функцией от τ_1^2 . Аналогично, пусть $m(z_1, z_3, \dots, z_q)$ — минимальное из значений z_1 и $[(1-g)/g]z_1 - \sum_{i=3}^q z_i$. Тогда

$$(50) \quad \Pr \{R_1 | \tau_1^2, \tau_2^2, \dots, \tau_q^2\} = \int_0^\infty \dots \int_0^\infty \int_{\substack{\max z_i \\ i \geq 2}}^\infty \prod_{i \neq 2} k(z_i | \tau_i^2) \left\{ \int_0^{m(z_1, z_3, \dots, z_q)} k(z_2 | \tau_2^2) dz_2 \right\} \prod_{i \neq 2} dz_i.$$

Интеграл в фигурных скобках является в силу леммы 4.4.3 убывающей функцией от τ_2^2 . ■

ТЕОРЕМА 4.4.3. Если $\tau_j^2 > \tau_k^2$, то

$$(51) \quad \Pr \{R_j | \tau_1^2, \dots, \tau_q^2\} > \Pr \{R_k | \tau_1^2, \dots, \tau_q^2\}.$$

ДОКАЗАТЕЛЬСТВО. Левая часть (51) больше, чем вероятность события R_j с заменой в условии τ_j^2 на τ_k^2 . Последняя же равна вероятности R_k с заменой в условии τ_j^2 на τ_k^2 и больше, чем правая часть (51). ■

ТЕОРЕМА 4.4.4.

$$(52) \quad \Pr \{R_j | \tau_j^2 = \tau^2, \tau_i^2 = 0, i \neq j\} > \Pr \left\{ R_j | \tau_j^2 < \tau^2, \sum_{i=1}^q \tau_i^2 = \tau^2 \right\}.$$

ДОКАЗАТЕЛЬСТВО. Эта теорема вытекает из последовательного применения леммы 4.4.4 с заменой τ_1^2 на τ_j^2 и τ_2^2 на $\tau_i^2, i \neq j$. ■

Вероятности $\Pr \{R_1 | \tau_1^2, \dots, \tau_q^2\}, \dots, \Pr \{R_q | \tau_1^2, \dots, \tau_q^2\}$ зависят от $\tau_1^2, \dots, \tau_q^2$ и вычисление их весьма затруднительно.

4.4.3. Использование тригонометрических функций, периоды которых не являются делителями длины ряда

Ранее было замечено, что $y_1, \dots, y_T$ можно выразить с помощью $\bar{y}, a_1, b_1, \dots, a_{[(T-1)/2]}, b_{[(T-1)/2]}$ и $a_{T/2}$ (если T четное). Если $T = 2H + 1$, то

$$(53) \quad y_t = \bar{y} + \sum_{j=1}^H \left(a_j \cos \frac{2\pi j}{T} t + b_j \sin \frac{2\pi j}{T} t \right).$$

Если же $T = 2H + 2$, то

$$(54) \quad y_t = \bar{y} + \sum_{j=1}^H \left(a_j \cos \frac{2\pi j}{T} t + b_j \sin \frac{2\pi j}{T} t \right) + a_{T/2} (-1)^t.$$

Коэффициенты $A(v)$ и $B(v)$ также могут быть выражены с помощью этих T тригонометрических коэффициентов. Например, для $T = 2H + 1$

$$\begin{aligned} (55) \quad A(v) &= \frac{2}{T} \bar{y} \sum_{t=1}^T \cos vt + \\ &+ \frac{2}{T} \sum_{j=1}^H \left(a_j \sum_{t=1}^T \cos \frac{2\pi j}{T} t \cos vt + \right. \\ &\left. + b_j \sum_{t=1}^T \sin \frac{2\pi j}{T} t \cos vt \right) = \\ &= \frac{2}{T} \bar{y} \frac{\sin \frac{1}{2} v T}{\sin \frac{1}{2} v} \cos \frac{1}{2} v (T+1) + \\ &+ \frac{1}{T} \sum_{j=1}^H \left\{ a_j \left[\frac{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right) T}{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right)} \cos \left(\frac{1}{2} v + \frac{\pi j}{T} \right) (T+1) + \right. \right. \\ &\left. + \frac{\sin \left(\frac{1}{2} v - \frac{\pi j}{T} \right) T}{\sin \left(\frac{1}{2} v - \frac{\pi j}{T} \right)} \cos \left(\frac{1}{2} v - \frac{\pi j}{T} \right) (T+1) \right] + \\ &\left. + b_j \left[\frac{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right) T}{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right)} \sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right) (T+1) - \right. \right. \end{aligned}$$

$$\begin{aligned}
& - \frac{\sin\left(\frac{1}{2}v - \frac{\pi j}{T}\right) T}{\sin\left(\frac{1}{2}v - \frac{\pi j}{T}\right)} \sin\left(\frac{1}{2}v - \frac{\pi j}{T}\right) (T+1) \Bigg\} = \\
& = - \frac{\sin \frac{1}{2} v T}{T} \left[2\bar{y} \frac{\cos \frac{1}{2} v (T+1)}{\sin \frac{1}{2} v} + \right. \\
& + \sum_{j=1}^H \left\{ a_j \left[\frac{\cos \left[\frac{1}{2} v (T+1) + \frac{\pi j}{T} \right]}{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right)} + \right. \right. \\
& + \left. \left. \frac{\cos \left[\frac{1}{2} v (T+1) - \frac{\pi j}{T} \right]}{\sin \left(\frac{1}{2} v - \frac{\pi j}{T} \right)} \right] + \right. \\
& + b_j \left[\frac{\sin \left[\frac{1}{2} v (T+1) + \frac{\pi j}{T} \right]}{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right)} - \right. \\
& \left. \left. - \frac{\sin \left[\frac{1}{2} v (T+1) - \frac{\pi j}{T} \right]}{\sin \left(\frac{1}{2} v - \frac{\pi j}{T} \right)} \right] \right\} \Bigg],
\end{aligned}$$

$$\begin{aligned}
(56) \quad B(v) = & - \frac{\sin \frac{1}{2} v T}{T} \left[2\bar{y} \frac{\sin \frac{1}{2} v (T+1)}{\sin \frac{1}{2} v} + \right. \\
& + \sum_{j=1}^H \left\{ a_j \left[\frac{\sin \left[\frac{1}{2} v (T+1) + \frac{\pi j}{T} \right]}{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right)} + \right. \right. \\
& + \left. \left. \frac{\sin \left[\frac{1}{2} v (T+1) - \frac{\pi j}{T} \right]}{\sin \left(\frac{1}{2} v - \frac{\pi j}{T} \right)} \right] + \right.
\end{aligned}$$

$$+ b_j \left[- \frac{\cos \left[\frac{1}{2} v (T+1) + \frac{\pi j}{T} \right]}{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right)} + \right. \\ \left. + \frac{\cos \left[\frac{1}{2} v (T+1) - \frac{\pi j}{T} \right]}{\sin \left(\frac{1}{2} v - \frac{\pi j}{T} \right)} \right] \Bigg] \Bigg] .$$

Вычисление упрощается, если использовать следующие ортогональные преобразования величин $A(v)$, $B(v)$ и пар коэффициентов a_j и b_j , $j = 1, \dots, H$:

$$(57) \quad C(v) = A(v) \cos \frac{1}{2} v (T+1) + B(v) \sin \frac{1}{2} v (T+1) = \\ = - \frac{\sin \frac{1}{2} v T}{T} \left[\frac{2\bar{y}}{\sin \frac{1}{2} v} + \sum_{j=1}^H \left\{ \frac{a_j \cos \frac{\pi j}{T} + b_j \sin \frac{\pi j}{T}}{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right)} + \right. \right. \\ \left. \left. + \frac{a_j \cos \frac{\pi j}{T} + b_j \sin \frac{\pi j}{T}}{\sin \left(\frac{1}{2} v - \frac{\pi j}{T} \right)} \right\} \right] = \\ = - \frac{\sin \frac{1}{2} v T}{T} \left[\frac{2\bar{y}}{\sin \frac{1}{2} v} + \sum_{j=1}^H c_j \left(\frac{1}{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right)} + \right. \right. \\ \left. \left. + \frac{1}{\sin \left(\frac{1}{2} v - \frac{\pi j}{T} \right)} \right) \right] ,$$

$$(58) \quad D(v) = - A(v) \sin \frac{1}{2} v (T+1) + B(v) \cos \frac{1}{2} v (T+1) = \\ = - \frac{\sin \frac{1}{2} v T}{T} \sum_{j=1}^H \left\{ \frac{a_j \sin \frac{\pi j}{T} - b_j \cos \frac{\pi j}{T}}{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right)} + \right. \\ \left. - a_j \sin \frac{\pi j}{T} + b_j \cos \frac{\pi j}{T} \right. \\ \left. + \frac{\quad}{\sin \left(\frac{1}{2} v - \frac{\pi j}{T} \right)} \right\} =$$

$$= -\frac{\sin \frac{1}{2} \nu T}{T} \sum_{j=1}^H d_j \left\{ \frac{1}{\sin \left(\frac{1}{2} \nu - \frac{\pi j}{T} \right)} - \frac{1}{\sin \left(\frac{1}{2} \nu + \frac{\pi j}{T} \right)} \right\},$$

где

$$(59) \quad c_j = a_j \cos \frac{\pi j}{T} + b_j \sin \frac{\pi j}{T} = \frac{2}{T} \sum_{t=1}^T y_t \cos \frac{2\pi j}{T} \left(t - \frac{1}{2} \right),$$

$$(60) \quad d_j = -a_j \sin \frac{\pi j}{T} + b_j \cos \frac{\pi j}{T} = \frac{2}{T} \sum_{t=1}^T y_t \sin \frac{2\pi j}{T} \left(t - \frac{1}{2} \right).$$

Если $T = 2H + 2$, то $A(\nu)$ равно правой части (55) плюс $2a_{T/2} \sin \frac{1}{2} \nu T \cos \left[\frac{1}{2} \nu (T + 1) + \frac{1}{2} \pi \right] \left[T \sin \left(\frac{1}{2} \nu + \frac{1}{2} \pi \right) \right]^{-1}$, $B(\nu)$ равно правой части (56) плюс $2a_{T/2} \sin \frac{1}{2} \nu T \sin \left[\frac{1}{2} \nu (T + 1) + \frac{1}{2} \pi \right] \left[T \sin \left(\frac{1}{2} \nu + \frac{1}{2} \pi \right) \right]^{-1}$, $C(\nu)$ равно правой части (57), а $D(\nu)$ равно правой части (58) минус $2d_{T/2} \sin \frac{1}{2} \nu T \left[T \sin \left(\frac{1}{2} \nu + \frac{1}{2} \pi \right) \right]^{-1}$, где $d_{T/2} = -a_{T/2}$ (а $c_{T/2} = b_{T/2} = 0$).

В случае нормальности коэффициенты a_j и b_j некоррелированы и независимы. При этом некоррелированы и взаимно независимы среднее $\bar{y}$ и коэффициенты c_j и d_j . Поэтому некоррелированы также при любых ν и ν' величины $C(\nu)$ и $D(\nu')$ и совокупность $\{C(\nu)\}$ для всех ν ($0 \leq \nu \leq \pi$) не зависит от совокупности $\{D(\nu)\}$ для всех ν ($0 \leq \nu \leq \pi$). Поскольку преобразование, переводящее $A(\nu)$, $B(\nu)$ в $C(\nu)$, $D(\nu)$, ортогонально, то $R^2(\nu) = C^2(\nu) + D^2(\nu)$. К сожалению, представления $C(\nu)$ и $D(\nu)$ использовать нелегко.

Если $f(t) = \mu + \rho \cos(\lambda t - \theta)$, то можно изучить отклонения от выборочного среднего. Положим

$$(61) \quad A^*(\nu) = \frac{2}{T} \sum_{t=1}^T (y_t - \bar{y}) \cos \nu t,$$

$$(62) \quad B^*(\nu) = \frac{2}{T} \sum_{t=1}^T (y_t - \bar{y}) \sin \nu t.$$

ТЕОРЕМА 4.4.5. Если $\mathcal{E}y_t = \mu + \alpha \cos \lambda t + \beta \sin \lambda t = \mu + \rho \cos(\lambda t - \theta)$, то

$$(63) \quad \mathcal{E}A^*(\nu) = \frac{\rho}{T} \left\{ \frac{\sin \frac{1}{2} (\lambda + \nu) T}{\sin \frac{1}{2} (\lambda + \nu)} \cos \left[\frac{1}{2} (\lambda + \nu) (T + 1) - \theta \right] + \right.$$

$$\begin{aligned}
& + \frac{\sin \frac{1}{2} (\lambda - \nu) T}{\sin \frac{1}{2} (\lambda - \nu)} \cos \left[\frac{1}{2} (\lambda - \nu) (T + 1) - \theta \right] - \\
& - \frac{2}{T} \frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} \lambda} \cos \left[\frac{1}{2} \lambda (T + 1) - \theta \right] \times \\
& \times \frac{\sin \frac{1}{2} \nu T}{\sin \frac{1}{2} \nu} \cos \frac{1}{2} \nu (T + 1) \Bigg\}, \quad \nu \neq \lambda, \quad 0 < \nu < \pi, \\
& = \alpha + \frac{\rho}{T} \left\{ \frac{\sin \lambda T}{\sin \lambda} \cos [\lambda (T + 1) - \theta] - \right. \\
& - \frac{2}{T} \frac{\sin^2 \frac{1}{2} \lambda T}{\sin^2 \frac{1}{2} \lambda} \cos \left[\frac{1}{2} \lambda (T + 1) - \theta \right] \times \\
& \times \cos \frac{1}{2} \lambda (T + 1) \Bigg\}, \quad 0 < \lambda = \nu < \pi, \\
(64) \quad \mathcal{E}B^*(\nu) & = \frac{\rho}{T} \left\{ \frac{\sin \frac{1}{2} (\lambda + \nu) T}{\sin \frac{1}{2} (\lambda + \nu)} \sin \left[\frac{1}{2} (\lambda + \nu) (T + 1) - \theta \right] - \right. \\
& - \frac{\sin \frac{1}{2} (\lambda - \nu) T}{\sin \frac{1}{2} (\lambda - \nu)} \sin \left[\frac{1}{2} (\lambda - \nu) (T + 1) - \theta \right] - \\
& - \frac{2}{T} \frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} \lambda} \cos \left[\frac{1}{2} \lambda (T + 1) - \theta \right] \times \\
& \times \frac{\sin \frac{1}{2} \nu T}{\sin \frac{1}{2} \nu} \sin \frac{1}{2} \nu (T + 1) \Bigg\}, \quad \nu \neq \lambda, \quad 0 < \nu < \pi, \\
& = \beta + \frac{\rho}{T} \left\{ \frac{\sin \lambda T}{\sin \lambda} \sin [\lambda (T + 1) - \theta] - \right.
\end{aligned}$$

$$- \frac{2}{T} \frac{\sin^2 \frac{1}{2} \lambda T}{\sin^2 \frac{1}{2} \lambda} \cos \left[\frac{1}{2} \lambda (T+1) - \theta \right] \times \\ \times \sin \frac{1}{2} \lambda (T+1) \Big\}, \quad 0 < \lambda = \nu < \pi,$$

$$(65) \quad \text{Var} [A^*(\nu)] = \\ = \frac{2\sigma^2}{T} \left\{ 1 + \frac{\sin \nu T}{T \sin \nu} \cos \nu (T+1) - \right. \\ \left. - 2 \left[\frac{\sin \frac{1}{2} \nu T}{T \sin \frac{1}{2} \nu} \cos \frac{1}{2} \nu (T+1) \right]^2 \right\}, \quad 0 < \nu < \pi,$$

$$(66) \quad \text{Var} [B^*(\nu)] = \\ = \frac{2\sigma^2}{T} \left\{ 1 - \frac{\sin \nu T}{T \sin \nu} \cos \nu (T+1) - \right. \\ \left. - 2 \left[\frac{\sin \frac{1}{2} \nu T}{T \sin \frac{1}{2} \nu} \sin \frac{1}{2} \nu (T+1) \right]^2 \right\}, \quad 0 < \nu < \pi,$$

$$(67) \quad \text{Cov} [A^*(\nu), B^*(\nu)] = \frac{2\sigma^2}{T} \left\{ \frac{\sin \nu T}{T \sin \nu} \sin \nu (T+1) - \right. \\ \left. - 2 \left[\frac{\sin \frac{1}{2} \nu T}{T \sin \frac{1}{2} \nu} \right]^2 \cos \frac{1}{2} \nu (T+1) \sin \frac{1}{2} \nu (T+1) \right\}, \quad 0 < \nu < \pi.$$

Квадрат теоретической амплитуды равен

$$(68) \quad \rho^{*2}(\nu) = \frac{\rho^2}{T^2} \left\{ \frac{\sin^2 \frac{1}{2} (\lambda + \nu) T}{\sin^2 \frac{1}{2} (\lambda + \nu)} + \frac{\sin^2 \frac{1}{2} (\lambda - \nu) T}{\sin^2 \frac{1}{2} (\lambda - \nu)} + \right. \\ \left. + 2 \frac{\sin \frac{1}{2} (\lambda + \nu) T}{\sin \frac{1}{2} (\lambda + \nu)} \cdot \frac{\sin \frac{1}{2} (\lambda - \nu) T}{\sin \frac{1}{2} (\lambda - \nu)} \cos [\lambda (T+1) - 2\theta] - \right.$$

$$\begin{aligned}
& -\frac{4}{T} \frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} \lambda} \cos^2 \left[\frac{1}{2} \lambda (T+1) - \theta \right] \times \\
& \times \frac{\sin \frac{1}{2} \nu T}{\sin \frac{1}{2} \nu} \left[\frac{\sin \frac{1}{2} (\lambda + \nu) T}{\sin \frac{1}{2} (\lambda + \nu)} + \frac{\sin \frac{1}{2} (\lambda - \nu) T}{\sin \frac{1}{2} (\lambda - \nu)} \right] + \\
& + \frac{4}{T^2} \frac{\sin^2 \frac{1}{2} \lambda T}{\sin^2 \frac{1}{2} \lambda} \cos^2 \left[\frac{1}{2} \lambda (T+1) - \theta \right] \frac{\sin^2 \frac{1}{2} \nu T}{\sin^2 \frac{1}{2} \nu} \Bigg\}, \\
& \qquad \qquad \qquad \nu \neq \lambda, \quad 0 < \nu < \pi, \\
& = \rho^2 \left\{ 1 + \frac{\sin^2 \lambda T}{T^2 \sin^2 \lambda} + 2 \frac{\sin \lambda T}{T \sin \lambda} \cos [\lambda (T+1) - 2\theta] - \right. \\
& \quad - \frac{4}{T^2} \frac{\sin^2 \frac{1}{2} \lambda T}{\sin^2 \frac{1}{2} \lambda} \cos^2 \left[\frac{1}{2} \lambda (T+1) - \theta \right] \times \\
& \quad \times \left[1 + \frac{\sin \lambda T}{T \sin \lambda} - \frac{\sin^2 \frac{1}{2} \lambda T}{T^2 \sin^2 \frac{1}{2} \lambda} \right] \Bigg\}, \quad 0 < \lambda = \nu < \pi.
\end{aligned}$$

Отметим, что $A^*(2\pi j/T) = A(2\pi j/T)$, $B^*(2\pi j/T) = B(2\pi j/T)$, $R^{*2}(2\pi j/T) = R^2(2\pi j/T) = R_j^2$ и $\rho^{*2}(2\pi j/T) = \rho^2(2\pi j/T) = \rho_j^2$.

Критерий для проверки гипотезы о том, что $\rho = 0$ в предположении, что $\xi y_t = \rho \cos(\lambda t - \theta)$ для некоторого $0 < \lambda < \pi$, а значение σ^2 неизвестно, может состоять, например, в отклонении нулевой гипотезы, когда $\max R^2(\nu) / \sum_{t=1}^T y_t^2$ превышает некоторую постоянную величину. К сожалению, вероятность такого события при нулевой гипотезе не вычислена в нормальном случае. В силу непрерывности $R^2(\nu)$ максимум $R^2(\nu)$, $0 \leq \nu \leq \pi$, не будет сильно отличаться от максимума R_j^2 , $j = 1, \dots, [(T-1)/2]$. Если предполагается, что $\xi y_t = \mu + \rho \cos(\lambda t - \theta)$, то аналогичным образом можно использовать $\max R^{*2}(\nu) / \sum_{t=1}^T (y_t - \bar{y})^2$.

Обратимся теперь к задаче оценивания α , β и λ в предположении, что тренд является нетривиальной тригонометрической последовательностью. Для удобства будем считать, что $\mu = 0$. Если предполагать, что величины y_t распределены нормально, то логарифм соответствующей функции правдоподобия будет равен

$$(69) \quad -\frac{1}{2} T \log 2\pi - \frac{1}{2} T \log \sigma^{*2} - \frac{1}{2\sigma^{*2}} S(\alpha^*, \beta^*, \lambda^*),$$

где

$$(70) \quad S(\alpha^*, \beta^*, \lambda^*) = \sum_{t=1}^T (y_t - \alpha^* \cos \lambda^* t - \beta^* \sin \lambda^* t)^2.$$

Оценками максимального правдоподобия для α , β и λ являются значения α^* , β^* и λ^* , минимизирующие сумму $S(\alpha^*, \beta^*, \lambda^*)$. Для заданного значения λ^* экстремальные значения α^* и β^* удовлетворяют соотношениям

$$(71) \quad \begin{aligned} \left[1 + \frac{\sin \lambda^* T}{T \sin \lambda^*} \cos \lambda^* (T+1) \right] \alpha^* + \frac{\sin \lambda^* T}{T \sin \lambda^*} \sin \lambda^* (T+1) \beta^* &= A(\lambda^*), \\ \frac{\sin \lambda^* T}{T \sin \lambda^*} \sin \lambda^* (T+1) \alpha^* + \\ &+ \left[1 - \frac{\sin \lambda^* T}{T \sin \lambda^*} \cos \lambda^* (T+1) \right] \beta^* = B(\lambda^*). \end{aligned}$$

Разрешая их относительно α^* и β^* , получаем соответственно

$$(72) \quad \hat{\alpha} = \frac{A(\lambda^*) - \frac{\sin \lambda^* T}{T \sin \lambda^*} [\cos \lambda^* (T+1) A(\lambda^*) + \sin \lambda^* (T+1) B(\lambda^*)]}{1 - \frac{\sin^2 \lambda^* T}{T^2 \sin^2 \lambda^*}};$$

$$(73) \quad \hat{\beta} = \frac{B(\lambda^*) - \frac{\sin \lambda^* T}{T \sin \lambda^*} [\sin \lambda^* (T+1) A(\lambda^*) - \cos \lambda^* (T+1) B(\lambda^*)]}{1 - \frac{\sin^2 \lambda^* T}{T^2 \sin^2 \lambda^*}}.$$

Для этих значений α^* и β^* сумма квадратов (70) равна $\sum_{t=1}^T y_t^2 - \sigma^2 Q(\lambda^*)$, где $Q(\lambda)$ — квадратичная форма, определяемая соотношением (29). Значение λ^* , минимизирующее последнюю сумму квадратов, совпадает со значением λ^* , максимизирующим $Q(\lambda^*)$, и удовлетворяет уравнению $dQ(\lambda^*)/d\lambda^* = 0$. Однако при этом производная представима слишком сложным выражением и это уравнение нельзя решить в явном виде (см. упр. 40).

Рассмотрим теорию этих процедур для случая больших выборок.

ТЕОРЕМА 4.4.6. Если $\mathcal{E}y_t = \alpha \cos \lambda t + \beta \sin \lambda t$, $0 < \lambda < \pi$, $\text{Var}(y_t) = \sigma^2$, а $\text{Cov}(y_t, y_s) = 0$, $t \neq s$, то

$$(74) \quad \text{plim}_{T \rightarrow \infty} A(v) = \begin{cases} 0, & v \neq \lambda, \quad 0 < v < \pi, \\ \alpha, & v = \lambda, \end{cases}$$

$$(75) \quad \text{plim}_{T \rightarrow \infty} B(v) = \begin{cases} 0, & v \neq \lambda, \quad 0 < v < \pi, \\ \beta, & v = \lambda, \end{cases}$$

$$(76) \quad \text{plim}_{T \rightarrow \infty} R^2(v) = \begin{cases} 0, & v \neq \lambda, \quad 0 < v < \pi, \\ \rho^2, & v = \lambda. \end{cases}$$

ДОКАЗАТЕЛЬСТВО. В соответствии с теоремой 4.4.1 $\mathcal{E}A(v) \rightarrow 0$, $v \neq \lambda$, $\mathcal{E}A(v) \rightarrow \alpha$, $v = \lambda$, $\mathcal{E}B(v) \rightarrow 0$, $v \neq \lambda$, $\mathcal{E}B(v) \rightarrow \beta$, $v = \lambda$, $\text{Var}[A(v)] \rightarrow 0$ и $\text{Var}[B(v)] \rightarrow 0$. Поэтому в силу неравенства Чебышева справедливы соотношения (74) и (75). Из них в свою очередь вытекает соотношение (76), так как $R^2(v)$ непрерывно зависит от $A(v)$ и $B(v)$. ■

СЛЕДСТВИЕ 4.4.2. Если $\mathcal{E}y_t = \mu + \alpha \cos \lambda t + \beta \sin \lambda t$, $0 < \lambda < \pi$, $\text{Var}(y_t) = \sigma^2$ и $\text{Cov}(y_t, y_s) = 0$, $t \neq s$, то выводы теоремы 4.4.6 остаются справедливыми для $A^*(v)$, $B^*(v)$ и $R^*(v)$.

ТЕОРЕМА 4.4.7. Пусть $y_t = \alpha \cos \lambda t + \beta \sin \lambda t + u_t$, $0 < \lambda < \pi$, $\mathcal{E}u_t = 0$, $\mathcal{E}u_t^2 = \sigma^2$, случайные величины u_t независимы и их распределения удовлетворяют условию (3) § 2.6. Тогда величины $\sqrt{T/2} \times \times [A(\lambda) - \alpha]$ и $\sqrt{T/2} [B(\lambda) - \beta]$ имеют в пределе двумерное нормальное распределение с нулевыми средними, дисперсиями σ^2 и нулевой ковариацией. При этом $\sqrt{T/2} A(v)$ и $\sqrt{T/2} B(v)$, $v \neq \lambda$, $0 < v < \pi$, имеют то же самое предельное распределение, а $TR^2(v)/(2\sigma^2)$ имеет в пределе χ^2 -распределение с 2 степенями свободы.

СЛЕДСТВИЕ 4.4.3. Пусть $y_t = \mu + \alpha \cos \lambda t + \beta \sin \lambda t + u_t$, $0 < \lambda < \pi$, $\mathcal{E}u_t = 0$, $\mathcal{E}u_t^2 = \sigma^2$, и случайные величины u_t независимы, а их распределения удовлетворяют условию (3) § 2.6. Тогда утверждение теоремы 4.4.7 сохраняет силу для $A^*(v)$, $B^*(v)$ и $R^{*2}(v)$.

Теоремы 4.4.6 и 4.4.7 выполняются для $A(v)$, $B(v)$ и $R^2(v)$ при каждом значении v . Их можно обобщить таким образом, чтобы они выполнялись одновременно для $v = v_1, \dots, v_k$ при фиксированном k . Однако мы заинтересованы в аналогичных предельных результатах либо для всех $v = 2\pi j/T$, $j = 1, \dots, [(T-1)/2]$, либо для всех v в интервале $0 < v < \pi$. В любом из этих случаев число точек не является фиксированным конечным числом.

Уиттл (1952) высказал утверждение о том, что модифицированные оценки максимального правдоподобия $\hat{\alpha}$, $\hat{\beta}$ и $\hat{\lambda}$ являются со-

стоятельными [см. также М. Рао (1960)], а А. Уолкер (1968) дал строгое доказательство этого факта, показав, кроме того, что

$$(77) \quad \text{plim}_{T \rightarrow \infty} T(\hat{\lambda} - \lambda) = 0.$$

Уиттл утверждал также, что эти оценки распределены асимптотически нормально, с ковариационной матрицей, приводимой ниже (правда, Уиттл допустил незначительную ошибку). Уолкер дал строгое доказательство того, что $\sqrt{T}(\hat{\alpha} - \alpha)$, $\sqrt{T}(\hat{\beta} - \beta)$ и $T^{3/2}(\hat{\lambda} - \lambda)$ имеют в пределе совместное нормальное распределение с нулевыми средними и ковариационной матрицей

$$(78) \quad 2\sigma^2 \begin{bmatrix} 1 & 0 & \frac{1}{2}\beta \\ 0 & 1 & -\frac{1}{2}\alpha \\ \frac{1}{2}\beta & -\frac{1}{2}\alpha & \frac{1}{3}\rho^2 \end{bmatrix}^{-1} = \\ = \frac{2\sigma^2}{\rho^2} \begin{bmatrix} \alpha^2 + 4\beta^2 & -3\alpha\beta & -6\beta \\ -3\alpha\beta & 4\alpha^2 + \beta^2 & 6\alpha \\ -6\beta & 6\alpha & 12 \end{bmatrix}.$$

Доказательства этих результатов крайне длинны и сложны. Заслуживает внимания то обстоятельство, что дисперсия $\hat{\lambda}$ имеет порядок $1/T^3$ вместо обычного $1/T$ и иногда встречающегося $1/T^2$. Предположения Уолкера состоят в том, что $\lambda \neq 0, \pi$, а величины u_t независимы и одинаково распределены со средними 0 и дисперсиями σ^2 .

Существенным для получения соотношения (77) [А. Уолкер (1965)] является доказательство сходимости по вероятности к нулю величин

$$(79) \quad A_u(v) = \frac{2}{T} \sum_{t=1}^T u_t \cos vt, \quad B_u(v) = \frac{2}{T} \sum_{t=1}^T u_t \sin vt$$

одновременно для всех v . Пусть $R_u^2(v) = A_u^2(v) + B_u^2(v)$. Уолкер (1965) доказал следующий результат.

ТЕОРЕМА 4.4.8. Если величины u_t независимы, имеют нулевые средние и дисперсии σ^2 , то

$$(80) \quad \text{plim}_{T \rightarrow \infty} \max_{0 \leq v \leq \pi} R_u^2(v) = 0.$$

Доказательство. Имеем

$$(81) \quad \mathbb{E} \max_{0 \leq v \leq \pi} R_u^2(v) = \frac{4}{T^2} \mathbb{E} \max_{0 \leq v \leq \pi} \left| \sum_{t=1}^T u_t e^{ivt} \right|^2.$$

Однако

$$(82) \quad \left| \sum_{t=1}^T u_t e^{ivt} \right|^2 = \sum_{t, t'=1}^T e^{iv(t-t')} u_t u_{t'} = \\ = \sum_{s=-(T-1)}^{T-1} e^{ivs} \sum_{t=1}^{T-|s|} u_t u_{t+|s|} \leq \sum_{s=-(T-1)}^{T-1} \left| \sum_{t=1}^{T-|s|} u_t u_{t+|s|} \right|$$

для каждого v . Поэтому соответствующее математическое ожидание оценивается величиной

$$(83) \quad \mathbb{E} \sum_{t=1}^T u_t^2 + 2 \sum_{s=1}^{T-1} \mathbb{E} \left| \sum_{t=1}^{T-s} u_t u_{t+s} \right| \leq \\ \leq T\sigma^2 + 2 \sum_{s=1}^{T-1} \left[\mathbb{E} \left(\sum_{t=1}^{T-s} u_t u_{t+s} \right)^2 \right]^{1/2} = \\ = T\sigma^2 + \\ + 2 \sum_{s=1}^{T-1} \left[\mathbb{E} \sum_{t, t'=1}^{T-s} u_t u_{t+s} u_{t'} u_{t'+s} \right]^{1/2} = \\ = T\sigma^2 + 2 \sum_{s=1}^{T-1} [\sigma^4 (T-s)]^{1/2} = \\ = \sigma^2 \left\{ T + 2 \sum_{r=1}^{T-1} r^{1/2} \right\} \leq \\ \leq \sigma^2 \left\{ T + 2 \int_1^T r^{1/2} dr \right\} \leq \\ \leq \sigma^2 \left\{ T + \frac{4}{3} T^{3/2} \right\}.$$

Отсюда вытекает, что

$$(84) \quad \mathbb{E} \max_{0 \leq v \leq \pi} R_u^2(v) \leq 4\sigma^2 \left\{ \frac{1}{T} + \frac{4}{3} \frac{1}{\sqrt{T}} \right\}.$$

Утверждение теоремы следует из этого результата и из обобщенного неравенства Чебышева. ■

Уиттл (1959) при дополнительном предположении, что $\mathbb{E} |u_t|^{4+\delta} < \infty$ для некоторого $\delta > 0$, получил равенство

$$(85) \quad \text{plim}_{T \rightarrow \infty} \frac{T}{\log T} \max_{0 \leq v \leq \pi} R_u^2(v) = 4\sigma^2.$$

Из соотношения (57) § 4.3 вытекает, что в случае, когда величины u_t нормально распределены,

$$(86) \quad \lim_{T \rightarrow \infty} \frac{T}{\log T} \max_{j=1, \dots, [(T-1)/2]} R_u^2 \left(\frac{2\pi j}{T} \right) = 4\sigma^2.$$

Интересно отметить, что максимальное значение $R_u^2(v)$ по всем действительным значениям v ($0 \leq v \leq \pi$) стохастически подобно максимальному значению $R_u^2(v)$, взятому по множеству $v = 2\pi j/T$, $j = 1, \dots, [(T-1)/2]$.

4.5. ОБСУЖДЕНИЕ

{ Функции тренда, имеющие периодический характер, существенно используются во многих областях: экономике, метеорологии, связи, астрономии и др. Часто эти периодические функции являются тригонометрическими или могут быть выражены в виде линейной комбинации последних. В настоящей главе были рассмотрены некоторые проблемы статистических выводов относительно таких функций тренда.

Наблюдаемый временной ряд описывается и анализируется с помощью линейных комбинаций, коэффициентами в которых служат тригонометрические функции времени, т. е. выборочные тригонометрические коэффициенты $A(v)$ и $B(v)$. Как мы увидим в гл. 8, эти коэффициенты используются в спектральном анализе ковариационной структуры стационарных процессов.

Основным практическим недостатком процедур, рассмотренных в настоящей главе, является то, что они предполагают некоррелированность случайных членов и постоянство их дисперсий. В большинстве приложений, однако, от временного параметра зависит структура как систематической, так и случайной составляющей. Другим недостатком является то, что методы, которые можно математически строго обосновать, являются ограниченными и не применимы ко всему кругу практических задач.

Интересным, по крайней мере с исторической точки зрения, является ряд Бевериджа цен на пшеницу в Западной и Центральной Европе с 1500 по 1869 г. Индекс составлен на основе цен, имевшихся приблизительно в 50 пунктах различных стран. Беверидж (1921) приводит этот индекс (подобранный так, что среднее за период 1700—1745 гг. равно 100), но использует для анализа ряд процентных отношений индексов данного года к среднему за период в 31 год, для которого данный год является серединой. Этот последний ряд, который считается свободным от тренда, приведен в табл. А.1.1 приложения А.1. Позднее Беверидж (1922) дал периодограмму, вычисленную с использованием $NR^2(2\pi k/N)/300$, где

N — целое, не превосходящее 356, а k/N выбирается так, чтобы получить большое количество частот. Последовательность данных начиналась с 1545 г. Соответствующие результаты приведены в табл. А.1.3 приложения А.1. Беверидж обнаружил 19 периодов, заслуживающих дальнейшего анализа. Сам по себе этот анализ достаточно интересен. Однако исходная модель кажется здесь неуместной, потому что трудно оправдать лежащую в ее основе идею о том, что циклический тренд состоит из многих тригонометрических составляющих, которые остаются неизменными в течение более чем 300 лет. Как мы увидим позднее, более подходящей представляется модель стационарного процесса. При этом основой соответствующей статистической обработки является спектрограмма.

ЛИТЕРАТУРА

§ 4.2. Уиттекер и Ватсон (1943).

§ 4.3. Т. Андерсон (1958), Бирнбаум (1959), (1961), Бьюис-Баллот (1847), Департамент сельского хозяйства США (1939), Дэвис (1941), Ирвин (1955), Карлин и Труакс (1960), Кендалл (1946b), Кудо (1960), Кули, Льюис и Уэлш (1967), Кули и Тьюки (1965), Леман (1957), Рунге (1903), Шустер (1898), Шеффе (1970), Стивенс (1939), Уиттл (1951), (1952), Дж. Т. Уолкер (1914), Феллер (1968), Фишер (1929), (1940).

§ 4.4. М. Рао (1960, Уиттл (1952), (1959), А. Уолкер (1965), (1968).

§ 4.5. Беверидж (1921), (1922).

УПРАЖНЕНИЯ

1. (Разд. 4.2.1) Используя равенство $e^{i\lambda} = \cos \lambda + i \sin \lambda$, докажите, что
- $\cos(a - b) = \cos a \cos b + \sin a \sin b$,
 - $\cos(a + b) = \cos a \cos b - \sin a \sin b$,
 - $\sin(a + b) = \sin a \cos b + \cos a \sin b$,
 - $\sin(a - b) = \sin a \cos b - \cos a \sin b$.

2. (Разд. 4.2.1) Убедитесь в справедливости соотношений

- $\cos a \cos b = \frac{1}{2} \cos(a - b) + \frac{1}{2} \cos(a + b)$,
- $\cos^2 a = \frac{1}{2} + \frac{1}{2} \cos(2a)$,
- $\sin a \sin b = \frac{1}{2} \cos(a - b) - \frac{1}{2} \cos(a + b)$,
- $\sin^2 a = \frac{1}{2} - \frac{1}{2} \cos(2a)$,
- $\sin a \cos b = \frac{1}{2} \sin(a + b) + \frac{1}{2} \sin(a - b)$.

3. (Разд. 4.2.1) Докажите соотношение (11).

4. (Разд. 4.2.2) Выпишите матрицу M для $T = 4$ в числах (используя $\sqrt{2}$). Постройте графики соответствующих четырех функций.

5. (Разд. 4.2.2) Выпишите матрицу $\mathbf{M}$ для $T = 6$ в числах. Постройте графики соответствующих шести функций.

6. (Разд. 4.2.2) Пусть матрица $\mathbf{N} = (e^{i2\pi nst/T} / \sqrt{T})$, а $\mathbf{M}$ определяется соотношениями (14) и (21). Найдите матрицу $\mathbf{P}$, такую, что $\mathbf{N} = \mathbf{MP}$. Покажите, что $\mathbf{N}\mathbf{N}' = \mathbf{I}$.

7. (Разд. 4.2.3) Приведите все 12 функций, необходимых для представления периодической последовательности с периодом 12.

8. (Разд. 4.2.4) Докажите соотношения (36) — (39) методом, используемым в (40).

9. (Разд. 4.2.4) Докажите соотношения (35) — (39) по аналогии с доказательством (8).

10. (Разд. 4.3.3) Проверьте, что распределение величины (27) действительно имеет вид (28).

11. (Разд. 4.3.3) Покажите, что преобразование (29) индуцирует преобразование оценок, описанное в предыдущем предложении.

12. (Разд. 4.3.3) Покажите, что (29) эквивалентно

$$y_t^* = ky_t + c_0 + \sum_{\substack{i=1 \\ i \neq j}}^q \left(c_i \cos \frac{2\pi k_i}{T} t + d_i \sin \frac{2\pi k_i}{T} t \right) + \\ + \frac{4}{T} \sin \frac{1}{2} \theta \sum_{s=1}^T \sin \left[\frac{2\pi k_j}{T} (s-t) - \frac{1}{2} \theta \right] ky_s + c_{T/2} (-1)^t, \\ t = 1, \dots, T$$

13. (Разд. 4.3.3) Покажите, что критерий $F \geq F_{2, T-p}$ (е) является равномерным наиболее мощным для проверки гипотезы $H: \rho(k_j) = 0$ среди критериев, мощность которых зависит только от $\rho^2(k_j)/\sigma^2$.

14. (Разд. 4.3.3) При уровне значимости ϵ укажите t -критерии для проверки нулевой гипотезы $\alpha_0 = \alpha_0^*$, где α_0^* — заданное число, против альтернатив (а) $\alpha_0 > \alpha_0^*$, (б) $\alpha_0 < \alpha_0^*$ и (с) $\alpha_0 \neq \alpha_0^*$.

15. (Разд. 4.3.3) При уровне значимости ϵ укажите t -критерии для проверки нулевой гипотезы $\alpha_{T/2} = 0$ против альтернатив (а) $\alpha_{T/2} > 0$, (б) $\alpha_{T/2} < 0$ и (с) $\alpha_{T/2} \neq 0$.

16. (Разд. 4.3.3) Найдите группу преобразований, относительно которой статистики (30) (или (36)) и $\bar{y}$ являлись бы максимальными инвариантами, оставляющую инвариантной задачу проверки гипотез по значениям статистики (30).

17. (Разд. 4.3.3) Выразите $\alpha(kh)$ и $\beta(kh)$ в (31) в виде сумм, содержащих значения $f(t)$ при $t = 1, \dots, T = hn$, и в виде сумм, содержащих значения $f(t)$ при $t = 1, \dots, n$.

18. (Разд. 4.3.3) Докажите, что статистики (30) и (36) совпадают.

19. (Разд. 4.3.3) Докажите, что величины (31) и (37) совпадают.

20. (Разд. 4.3.3) Покажите, что граничные точки интервалов (42) и (43) с заменой $F_{1, T-p}$ (е) на $2F_{2, T-p}$ (е) являются тангенсами углов наклонов прямых, проходящих через начало координат и касающихся круга (38).

21. (Разд. 4.3.3) Докажите, что математическое ожидание величины $a(k_j) \cos \varphi + b(k_j) \sin \varphi$ равно $\rho(k_j) \cos[\varphi - \theta(k_j)]$.

22. (Разд. 4.3.3) Найдите двусторонний доверительный интервал для $\rho(k_j) \cos[\varphi - \theta(k_j)]$ с уровнем доверия $1 - \varepsilon$ при неизвестном σ^2 .

23. (Разд. 4.3.3) Покажите, что величины $\sqrt{T} [a_0 - \alpha_0]$, $\sqrt{T} [a_{T/2} - \alpha_{T/2}]$, $\sqrt{T} [a(k_j) - \alpha(k_j)]$, $\sqrt{T} [b(k_j) - \beta(k_j)]$, $j = 1, \dots, q$, имеют в пределе при $T \rightarrow \infty$ совместное нормальное распределение, если величины y_t независимы, имеют дисперсии σ^2 и равномерно ограниченные моменты порядка $2 + \delta$ для некоторого $\delta > 0$. (Указание. Использовать следствие 2.6.2.)

24. (Разд. 4.3.4) Докажите эквивалентность выражений для $k(z | \tau^2)$ в первых двух строчках соотношения (49), используя формулу удвоения для гамма-функции $\sqrt{\pi} \Gamma(2\beta + 1) = 2^{2\beta} \Gamma(\beta + \frac{1}{2}) \Gamma(\beta + 1)$.

25. (Разд. 4.3.4) Докажите, что нецентральное χ^2 -распределение с 2 степенями свободы (плотность которого приведена в (49)) является ограниченно полным, т. е. если

$$\int_0^\infty g(z) k(z | \tau^2) dz \equiv 0$$

тождественно относительно $\tau^2 > 0$, то $g(z) = 0$ почти всюду ($z \geq 0$) для каждой ограниченной функции $g(z)$. (Указание. Записанное выше тождество можно переписать в виде

$$0 \equiv \sum_{\nu=0}^{\infty} c_\nu (\tau^2)^\nu \int_0^\infty e^{-z/2} g(z) z^\nu dz,$$

где $c_\nu > 0$.)

26. (Разд. 4.3.4) Докажите, что произведение $q - 1$ нецентральных χ^2 -распределений, каждое из которых имеет 2 степени свободы, является ограниченно полным.

27. (Разд. 4.3.4) Пусть x_1, x_2, x_3 ($x_1 + x_2 + x_3 = 1$, $x_i \geq 0$) — барицентрические координаты (расстояния точки от сторон равностороннего треугольника с высотой 1). (a) Постройте на чертеже области, в которых $\max x_i \geq g$, $\frac{1}{3} \leq g \leq \frac{1}{2}$ и $\frac{1}{2} \leq g \leq 1$. (b) Постройте на чертеже области, в которых вторая по величине координата $x_j \geq g$, $0 \leq g \leq \frac{1}{3}$ и $\frac{1}{3} \leq g \leq \frac{1}{2}$.

28. (Разд. 4.3.4) Пусть x_1, x_2, x_3 ($x_1 + x_2 + x_3 = 1$, $x_i \geq 0$) — барицентрические координаты. Дайте графическое представление областей $R_0, R_1, R_2, R_3, R_{12}, R_{13}, R_{23}$, определяемых соотношениями

(i)	$R_0: x_i \leq g,$	$i = 1, 2, 3,$
(ii)	$R_i: x_i > g, \quad x_j \leq g^* x_k,$	$i \neq j \neq k \neq i,$
(iii)	$R_{ij}: x_i > g, \quad x_j > g^* x_k,$	$i \neq j \neq k \neq i,$

в которых $g^* \geq 1$ и $g \geq \frac{1}{2}$.

29. (Разд. 4.4.1) Докажите лемму 4.4.1.

30. (Разд. 4.4.1) Проверьте соотношение (23).

31. (Разд. 4.4.1) Проверьте соотношение (24).

32. (Разд. 4.4.1) Проверьте соотношение (30).

33. (Разд. 4.4.2) Покажите, что

$$2\alpha_0^2 + 2\alpha_{T/2}^2 = \frac{4\rho^2 \sin^2 \frac{1}{2} \lambda T}{T^2 \sin^2 \lambda} \{1 + \cos \lambda \cos [\lambda (T+1) - 2\theta]\}.$$

34. (Разд. 4.4.2) Покажите, что для $\lambda = 2\pi (k + \varepsilon)/T$

$$\begin{aligned} \alpha_0 &= \frac{\rho \sin \pi \varepsilon}{T \sin \pi \frac{k + \varepsilon}{T}} \cos \left(\pi \varepsilon + \pi \frac{k + \varepsilon}{T} - \theta \right) \sim \\ &\sim \rho \frac{\varepsilon}{k + \varepsilon} \frac{1 - \frac{\pi^2}{6} \varepsilon^2}{1 - \frac{\pi^2}{6} \cdot \frac{(k + \varepsilon)^2}{T^2}} \cos \left(\pi \varepsilon + \pi \frac{k + \varepsilon}{T} - \theta \right). \end{aligned}$$

35. (Разд. 4.4.2) Покажите, что для $\lambda = 2\pi (k + \varepsilon)/T$

$$\begin{aligned} \alpha_{T/2} &= - \frac{\rho \sin \pi \varepsilon}{T \cos \pi \frac{k + \varepsilon}{T}} \sin \left(\pi \varepsilon + \pi \frac{k + \varepsilon}{T} - \theta \right) \sim \\ &\sim - \frac{\rho}{T} \pi \varepsilon \left(1 - \frac{\pi^2 \varepsilon^2}{6} \right) \sin \left(\pi \varepsilon + \pi \frac{k + \varepsilon}{T} - \theta \right). \end{aligned}$$

36. (Разд. 4.4.3) Покажите, что

$$\begin{aligned} \frac{1}{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right)} + \frac{1}{\sin \left(\frac{1}{2} v - \frac{\pi j}{T} \right)} &= 4 \frac{\sin \frac{1}{2} v \cos \frac{\pi j}{T}}{\cos \frac{2\pi j}{T} - \cos v}, \\ \frac{1}{\sin \left(\frac{1}{2} v - \frac{\pi j}{T} \right)} - \frac{1}{\sin \left(\frac{1}{2} v + \frac{\pi j}{T} \right)} &= 4 \frac{\cos \frac{1}{2} v \sin \frac{\pi j}{T}}{\cos \frac{2\pi j}{T} - \cos v}. \end{aligned}$$

37. (Разд. 4.4.3) Покажите, что для $T = 2H + 1$

$$\begin{aligned} C(v) &= 4 \frac{\sin \frac{1}{2} v \sin \frac{1}{2} v T}{T} \left\{ \frac{\bar{y}}{1 - \cos v} + \sum_{j=1}^H \frac{c_j \cos \frac{\pi j}{T}}{\cos \frac{2\pi j}{T} - \cos v} \right\}, \\ D(v) &= 4 \frac{\cos \frac{1}{2} v \sin \frac{1}{2} v T}{T} \sum_{j=1}^H \frac{d_j \sin \frac{\pi j}{T}}{\cos \frac{2\pi j}{T} - \cos v}. \end{aligned}$$

38. (Разд. 4.4.3) Покажите, что для $K = \cos [\lambda (T + 1) - 2\theta]$

$$\begin{aligned} \frac{d\rho^2(v)}{dv} = & -\frac{\rho^2}{T^2} \left\{ -\frac{T}{2} \frac{\sin(\lambda + v)T}{\sin^2 \frac{1}{2}(\lambda + v)} + \right. \\ & + \frac{\sin^2 \frac{1}{2}(\lambda + v)T \cos \frac{1}{2}(\lambda + v)}{\sin^3 \frac{1}{2}(\lambda + v)} + \frac{T}{2} \frac{\sin(\lambda - v)T}{\sin^2 \frac{1}{2}(\lambda - v)} - \\ & - \frac{\sin^2 \frac{1}{2}(\lambda - v)T \cos \frac{1}{2}(\lambda - v)}{\sin^3 \frac{1}{2}(\lambda - v)} + \\ & + KT \frac{\sin vT}{\sin \frac{1}{2}(\lambda + v) \sin \frac{1}{2}(\lambda - v)} - \\ & \left. - K \frac{\sin \frac{1}{2}(\lambda + v)T \sin \frac{1}{2}(\lambda - v)T \sin v}{\sin^2 \frac{1}{2}(\lambda + v) \sin^2 \frac{1}{2}(\lambda - v)} \right\}, \quad v \neq \lambda, \quad 0 < v < \pi. \end{aligned}$$

39. (Разд. 4.4.3) Покажите, что для производной $d\rho^2(v)/dv$, определенной в упр. (38),

$$\lim_{v \rightarrow \lambda} \frac{d\rho^2(v)}{dv} = \rho^2 \frac{T \cos \lambda T \sin \lambda - \sin \lambda T \cos \lambda}{T \sin^2 \lambda} \left\{ \frac{\sin \lambda T}{T \sin \lambda} + K \right\}.$$

40. (Разд. 4.4.3) Запишите уравнение $dQ(\lambda^*)/d\lambda^* = 0$ в развернутом виде.

41. (Разд. 4.4.4) Докажите, что (86) вытекает из соотношения (57) § 4.3.

ЛИНЕЙНЫЕ ВЕРОЯТНОСТНЫЕ МОДЕЛИ С КОНЕЧНЫМ ЧИСЛОМ ПАРАМЕТРОВ

5.1. ВВЕДЕНИЕ

Обратимся теперь к моделям образования временных рядов, в которых характеристические и наиболее существенные свойства временной последовательности не сводятся к детерминированной функции среднего значения, а заключены в самой вероятностной структуре. В этом случае, например, уже не будет регулярных периодических циклов, а будут более или менее нерегулярные случайные изменения с теми или иными статистическими свойствами вариабельности. Такие модели обычно называют *случайными процессами*. Процессы, вероятностная структура которых не изменяется со временем, называются *стационарными*. (Более полно стационарные случайные процессы рассматриваются в гл. 7). В этой книге мы в основном изучаем процессы, которые или стационарны (близки к стационарным), или же таковы, что по крайней мере их случайная составляющая может приближенно считаться стационарной (в отличие от детерминированным образом меняющейся функции среднего значения).

Одной из наиболее простых и, по-видимому, наиболее часто используемой моделью такого рода является *процесс авторегрессии*, или *стохастическое разностное уравнение*. О последовательности случайных величин $y_1, y_2, \dots$ говорят, что она удовлетворяет стохастическому разностному уравнению, если существует такая линейная комбинация

$$(1) \quad y_t + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} = u_t, \quad t = p+1, \dots,$$

что последовательность $u_{p+1}, u_{p+2}, \dots$ является последовательностью независимых и одинаково распределенных случайных величин. (При этом мы часто будем предполагать $\mathbb{E} u_t = 0$.) Это определение удобно распространить на бесконечные в обе стороны

последовательности $\dots, y_{-1}, y_0, y_1, \dots$, связанные с бесконечной в обе стороны последовательностью $\dots, u_{-1}, u_0, u_1, \dots$. Такие процессы также называются процессами авторегрессии.

Используя оператор $\mathcal{P}$ (действующий по правилу $\mathcal{P}y_t = y_{t+1}$), соотношение (1) можно представить в виде

$$(2) \quad (\mathcal{P}^p + \beta_1 \mathcal{P}^{p-1} + \dots + \beta_p) y_{t-p} = u_t.$$

Поскольку $\mathcal{P} = \Delta + 1$, то оператор, применяемый в (2) к y_{t-p} , можно записать иначе как полином степени p от Δ , так что левая часть (1) будет линейной комбинацией величин $y_{t-p}, \Delta y_{t-p}, \dots, \Delta^p y_{t-p}$. Поэтому при $\beta_p \neq 0$ соотношение (1) называется стохастическим разностным уравнением *порядка* p . О соответствующем процессе говорят как о процессе авторегрессии *порядка* p . Разностный оператор Δ обсуждался ранее в § 3.4.

Стохастическое разностное уравнение (1) можно представлять и иначе. Если случайная величина u_t не зависит от $y_{t-1}, y_{t-2}, \dots$, то условное распределение случайной величины y_t при заданных значениях $y_{t-1}, y_{t-2}, \dots$ совпадает с распределением величины $u_t - (\beta_1 y_{t-1} + \dots + \beta_p y_{t-p})$. Поэтому совместное распределение случайных величин $y_t, y_{t-1}, \dots, y_1$ ($t > p$) можно получить, зная распределение u_t и совместное распределение величин $y_{t-1}, y_{t-2}, \dots, y_1$. Если задать распределение случайной величины u_t (одинаковое для всех t) и совместное распределение любых p последовательных величин y_t , то последовательным применением указанной процедуры можно получить совместное распределение для $y_s, s \leq t$, и любого числа последующих y_t . В следующем параграфе мы получим условия, при которых процесс, определяемый таким способом, будет стационарным (и при которых u_t не будет зависеть от $y_{t-1}, y_{t-2}, \dots$).

Приведенная модель полезна тем, что с ее помощью можно получить весьма обширный класс процессов. Влияние тренда легко учесть, добавляя в левую часть соотношения (1) слагаемое $\sum_i \gamma_i z_{it}$,

в котором z_{it} — известные величины (функции времени). Мы изучим некоторые свойства подобных моделей в § 5.2.

Многие задачи статистического вывода связаны с конечным числом параметров $\beta_1, \dots, \beta_p$ и дисперсией случайной величины u_t . Они могут состоять, например, в оценке указанных параметров и проверке гипотез относительно значений последних. При большой длине наблюдаемого ряда эти задачи могут быть рассмотрены в асимптотическом плане в соответствии с теорией наименьших квадратов, обзор которой был сделан в гл. 2. Теория сериальной корреляции для малых выборок рассматривается в гл. 6.

Другой простой моделью стационарного процесса является скользящее среднее

$$(3) \quad y_t = \alpha_0 v_t + \alpha_1 v_{t-1} + \dots + \alpha_q v_{t-q}.$$

Здесь v_t — независимые одинаково распределенные случайные величины. Эту модель можно объединить с моделью, рассматривавшейся ранее, приравнивая левую часть соотношения (1) правой части соотношения (3). Иными словами, возмущающий фактор в уравнении (1) будет при этом скользящим средним, определяемым правой частью (3). Каждую из описанных моделей можно модифицировать. Для этого предположим, например, что наблюдаемый процесс складывается из процесса, порождаемого одной из этих моделей, и независимой случайной «ошибки». Тогда члены наблюдаемой последовательности $\{x_t\}$ имеют вид $x_t = y_t + w_t$, где $\{w_t\}$ — последовательность независимых одинаково распределенных случайных величин. Статистические выводы во всех получаемых таким образом моделях сложнее (даже в асимптотическом плане), чем в случае процесса авторегрессии.

Если все совместные распределения нормальны, то совместное нормальное распределение случайных величин y_t полностью определяется заданием их средних, дисперсий и ковариаций. Если процесс к тому же стационарный, то эти распределения будут полностью определяться заданием среднего и дисперсии случайной величины u_t и коэффициентами β_i в случае (1), среднего и дисперсии случайной величины v_t и коэффициентами α_j в случае (3), среднего и дисперсии случайной величины v_t и коэффициентами α_j и β_i ($i = 0, \dots, p$, $j = 1, \dots, q$) в комбинированной модели. Если введенным выше определениям удовлетворяют лишь первые и вторые моменты, то иногда говорят, что эти определения выполняются в *широком смысле*. Например, говорят, что последовательность $\{y_t\}$ удовлетворяет стохастическому разностному уравнению в широком смысле, если величины u_t , определяемые соотношением (1), таковы, что $\mathbb{E} u_t = 0$, $\mathbb{E} u_t^2 = \sigma^2$, $\mathbb{E} u_t u_s = 0$, $t \neq s$. Подобным же образом можно говорить о моделях скользящего среднего и о комбинированных моделях, определенных в широком смысле с помощью величин v_t .

Юл (1927) предложил использовать процесс авторегрессии для анализа временных рядов и применил его к данным о числе солнечных пятен. Дж. Уолкер (1931) построил соответствующую теорию и применил ее к анализу атмосферных явлений.

5.2. ПРОЦЕССЫ АВТОРЕГРЕССИИ

5.2.1. Представление временного ряда с помощью бесконечного скользящего среднего

Во введении к этой главе было показано, каким образом по совместному распределению некоторых p последовательных величин y_t и по распределению (независимо и одинаково распределен-

ных) u_t в совокупности с соотношением

$$(1) \quad \sum_{r=0}^p \beta_r y_{t-r} = u_t, \quad \beta_0 = 1,$$

можно определить совместное распределение последующих случайных величин y_t . Покажем теперь, что каждую случайную величину y_t можно выразить в виде линейной комбинации p предшествующих ей случайных величин y_r и u_r . При выполнении формулируемых ниже условий каждую случайную величину y_t можно представить в виде бесконечной линейной комбинации случайной величины u_t и ей предшествующих случайных величин u_r .

Наиболее простым случаем является уравнение первого порядка

$$(2) \quad y_t = \rho y_{t-1} + u_t.$$

Если вместо y_{t-1} в это уравнение подставить его выражение $y_{t-1} = \rho y_{t-2} + u_{t-1}$ (получаемое из (2) заменой t на $t-1$), то (2) примет вид

$$(3) \quad y_t = u_t + \rho u_{t-1} + \rho^2 y_{t-2}.$$

Действуя далее подобным же образом, можно прийти к соотношению

$$(4) \quad y_t = u_t + \rho u_{t-1} + \dots + \rho^s u_{t-s} + \rho^{s+1} y_{t-(s+1)},$$

так что

$$(5) \quad y_t - (u_t + \rho u_{t-1} + \dots + \rho^s u_{t-s}) = \rho^{s+1} y_{t-(s+1)}.$$

Если $\{y_t\}$ — бесконечный в обе стороны по t стационарный процесс, то эта разность при $|\rho| < 1$ с ростом s становится малой. В частности (если существуют вторые моменты),

$$(6) \quad \mathbb{E}[y_t - (u_t + \rho u_{t-1} + \dots + \rho^s u_{t-s})]^2 = \rho^{2(s+1)} \mathbb{E} y_{t-(s+1)}^2$$

не зависит от t и с ростом s стремится к нулю. В связи с этим y_t здесь записывают в виде

$$(7) \quad y_t = \sum_{r=0}^{\infty} \rho^r u_{t-r}$$

и говорят, что ряд в правой части *сходится* к y_t *в среднем* (или в среднеквадратичном). (Сходимость в среднем обсуждается в разд. 7.6.1.)

Вернемся теперь к общему случаю. Запишем соотношение (1) для моментов t и $t-1$ в виде

$$(8) \quad y_t = u_t - \beta_1 y_{t-1} - \dots - \beta_p y_{t-p},$$

$$(9) \quad y_{t-1} = u_{t-1} - \beta_1 y_{t-2} - \dots - \beta_p y_{t-1-p};$$

Подстановка (9) в (8) дает

$$(10) \quad y_t = u_t - \beta_1 (u_{t-1} - \beta_1 y_{t-2} - \dots - \beta_p y_{t-1-p}) - \beta_2 y_{t-2} - \dots \\ \dots - \beta_p y_{t-p} = \\ = u_t - \beta_1 u_{t-1} - (\beta_2 - \beta_1^2) y_{t-2} - \dots + \beta_1 \beta_p y_{t-1-p}.$$

Повторяя подобную процедуру s раз, получим для y_t выражение

$$(11) \quad y_t = u_t + \delta_1^* u_{t-1} + \dots + \delta_s^* u_{t-s} + \alpha_{s1}^* y_{t-s-1} + \alpha_{s2}^* y_{t-s-2} + \dots \\ \dots + \alpha_{sp}^* y_{t-s-p}.$$

(При каждой подстановке справа остается ровно p последовательных y_r .) Подстановка в (11)

$$(12) \quad y_{t-s-1} = u_{t-s-1} - \beta_1 y_{t-s-2} - \dots - \beta_p y_{t-s-p-1}$$

приводит к равенству

$$(13) \quad y_t = u_t + \delta_1^* u_{t-1} + \dots + \delta_s^* u_{t-s} + \alpha_{s1}^* u_{t-s-1} + \\ + (\alpha_{s2}^* - \alpha_{s1}^* \beta_1) y_{t-s-2} + \dots + (\alpha_{sp}^* - \alpha_{s1}^* \beta_{p-1}) y_{t-s-p} - \\ - \alpha_{s1}^* \beta_p y_{t-s-p-1}.$$

Входящие сюда коэффициенты определяются рекуррентными соотношениями

$$(14) \quad \delta_{s+1}^* = \alpha_{s1}^*, \\ \alpha_{s+1,j}^* = \alpha_{s,j+1}^* - \alpha_{s1}^* \beta_j, \quad j = 1, \dots, p-1, \\ \alpha_{s+1,p}^* = -\alpha_{s1}^* \beta_p.$$

Дальнейшее применение указанной процедуры приводит к представлению

$$(15) \quad y_t = \sum_{i=0}^{\infty} \delta_i^* u_{t-i},$$

в котором $\delta_0^* = 1$. Условия, при которых ряд (15) сходится в среднем, мы выясним ниже. Во всяком случае, соотношения с (11) по (14) справедливы для любого s . А это и показывает, что совместное распределение некоторых p последовательных y_t и следующих за

ними величин u_i определяет совместное распределение последующих y_i .

Указанную процедуру можно изложить и формальным образом. Пусть $\mathcal{L}$ — оператор запаздывания, т. е.

$$(16) \quad \mathcal{L}y_t = y_{t-1}.$$

С использованием этого оператора разностное уравнение (1) может быть записано в виде

$$(17) \quad \sum_{r=0}^p \beta_r \mathcal{L}^r y_t = u_t.$$

Тогда формально

$$(18) \quad y_t = \left(\sum_{r=0}^p \beta_r \mathcal{L}^r \right)^{-1} u_t,$$

где

$$(19) \quad \left(\sum_{r=0}^p \beta_r \mathcal{L}^r \right)^{-1} = \sum_{r=0}^{\infty} \delta_r \mathcal{L}^r.$$

То есть δ_r являются коэффициентами при z^r в разложении

$$(20) \quad \left(\sum_{r=0}^p \beta_r z^r \right)^{-1} = \sum_{r=0}^{\infty} \delta_r z^r$$

и могут быть определены формальным делением.

Убедимся теперь в том, что $\delta_r^* = \delta_r$. Имеем

$$(21) \quad \frac{1}{1 + \beta_1 z + \dots + \beta_p z^p} = 1 - \frac{\beta_1 z + \dots + \beta_p z^p}{1 + \beta_1 z + \dots + \beta_p z^p} =$$

$$= 1 - \beta_1 z - \frac{(\beta_2 - \beta_1^2) z^2 + \dots + (\beta_p - \beta_1 \beta_{p-1}) z^p - \beta_1 \beta_p z^{p+1}}{1 + \beta_1 z + \dots + \beta_p z^p}.$$

Следует отметить, что коэффициент при z^j в числителе правой части (21) совпадает с коэффициентом при y_{t-j} в (10), $j = 2, \dots, p+1$. Продолжая этот процесс, получаем

$$(22) \quad \frac{1}{1 + \beta_1 z + \dots + \beta_p z^p} =$$

$$= 1 + \delta_1 z + \dots + \delta_s z^s + \frac{\alpha_{s1} z^{s+1} + \dots + \alpha_{sp} z^{s+p}}{1 + \beta_1 z + \dots + \beta_p z^p} =$$

$$= 1 + \delta_1 z + \dots + \delta_s z^s + \alpha_{s1} z^{s+1} +$$

$$+ \frac{(\alpha_{s2} - \alpha_{s1} \beta_1) z^{s+2} + \dots + (\alpha_{sp} - \alpha_{s1} \beta_{p-1}) z^{s+p} - \alpha_{s1} \beta_p z^{s+p+1}}{1 + \beta_1 z + \dots + \beta_p z^p}.$$

Таким образом, коэффициенты δ_r и α_{si} удовлетворяют тем же рекуррентным соотношениям, что и коэффициенты δ_r^* и α_{si}^* . Поскольку совпадают и начальные условия, отсюда следует, что $\delta_r^* = \delta_r$ и $\alpha_{si}^* = \alpha_{si}$. Уравнение

$$(23) \quad \sum_{r=0}^p \beta_r x^{p-r} = 0$$

называется *алгебраическим уравнением, присоединенным к (1)¹⁾*. Оно имеет p корней, которые мы обозначим $x_1, \dots, x_p$. Если $|x_i| < 1, i = 1, \dots, p$, то при $\beta_p \neq 0$ корни уравнения

$$(24) \quad \sum_{r=0}^p \beta_r z^r = 0$$

равны $z_i = 1/x_i$ и $|z_i| > 1$. Для всех z , таких, что $|z| < \min |z_i|$, ряд

$$(25) \quad \frac{1}{\sum_{r=0}^p \beta_r z^r} = \frac{1}{\prod_{i=1}^p \left(1 - \frac{z}{z_i}\right)} = \prod_{i=1}^p \sum_{v=0}^{\infty} \left(\frac{z}{z_i}\right)^v = \sum_{r=0}^{\infty} \delta_r z^r$$

сходится (абсолютно). Отсюда и из (22) следует, что

$$(26) \quad \frac{\alpha_{s1}z + \dots + \alpha_{sp}z^p}{1 + \beta_1z + \dots + \beta_pz^p}$$

сходится к нулю для $|z| < \min |z_i|$ (в частности, для $|z| = 1$). В силу этого $\alpha_{si} \rightarrow 0$ при $s \rightarrow \infty$ для каждого i . Таким образом, величина

$$(27) \quad \mathbb{E} \left(y_t - \sum_{r=0}^s \delta_r u_{t-r} \right)^2 = \mathbb{E} (\alpha_{s1}y_{t-s-1} + \dots + \alpha_{sp}y_{t-s-p})^2$$

сходится к нулю при $s \rightarrow \infty$. А это и означает, что

$$(28) \quad y_t = \sum_{r=0}^{\infty} \delta_r u_{t-r}$$

в смысле сходимости в среднем.

ТЕОРЕМА 5.2.1. Если все корни характеристического уравнения (23), соответствующего стохастическому разностному уравнению (1), по абсолютной величине меньше 1, то y_t представимо бесконечной линейной комбинацией случайных величин $u_t, u_{t-1}, \dots$.

¹⁾ Мы будем называть его для краткости *характеристическим уравнением* для (1). — Прим. перев.

Следствие 5.2.1. Если все корни характеристического уравнения по абсолютной величине меньше 1, то y_t не зависит от $u_{t+1}, u_{t+2}, \dots$.

Доказательство. Случайная величина y_t является линейной комбинацией случайных величин $u_t, u_{t-1}, \dots$, а они не зависят от $u_{t+1}, u_{t+2}, \dots$. ■

Рассмотрим кратко случай, когда некоторые корни характеристического уравнения по абсолютной величине превосходят 1. Предположим, что $|x_i| > 1, i = 1, \dots, q, |x_i| < 1, i = q+1, \dots, p$. Запишем исследуемое стохастическое разностное уравнение в виде

$$(29) \quad u_t = \sum_{r=0}^p \beta_r \mathcal{P}^{p-r} y_{t-p} = \prod_{i=1}^p (\mathcal{P} - x_i) y_{t-p}.$$

Обращая (29), получаем

$$\begin{aligned} (30) \quad y_{t-p} &= \prod_{i=1}^p (\mathcal{P} - x_i)^{-1} u_t = \\ &= \prod_{i=1}^q (\mathcal{P} - x_i)^{-1} \prod_{i=q+1}^p [\mathcal{P} (1 - x_i \mathcal{L})]^{-1} u_t = \\ &= \prod_{i=1}^q \left[\left(-\frac{1}{x_i} \right) \left(1 - \frac{1}{x_i} \mathcal{P} \right)^{-1} \right] \prod_{i=q+1}^p [\mathcal{L} (1 - x_i \mathcal{L})^{-1}] u_t, \end{aligned}$$

поскольку $\mathcal{L} = \mathcal{P}^{-1}$. Каждый член $(1 - x_i^{-1} \mathcal{P})^{-1}$ в последнем выражении можно разложить по степеням $\mathcal{P}$. Члены $(1 - x_i \mathcal{L})^{-1}$ разложим по степеням $\mathcal{L}$. Если $1 \leq q < p$, то весь оператор в правой части (30) будет степенным рядом по $\mathcal{L}$ ($= \mathcal{P}^{-1}$) и $\mathcal{P}$ ($= \mathcal{L}^{-1}$), так что при этом будут включены u_t , образующие бесконечный в обе стороны ряд $\dots, u_{t-1}, u_t, u_{t+1}, \dots$. Если же $q = p$ (все корни по абсолютной величине больше 1), то этот оператор будет степенным рядом только по $\mathcal{P}$ и соответственно будут включены только случайные величины $u_t, u_{t+1}, \dots$. Приведенное рассуждение является чисто формальным. Однако его можно обосновать таким же образом, как это было сделано в случае, когда все корни по абсолютной величине предполагались меньшими единицы. Мы увидим

в гл. 7 (упр. 22), что если линейную форму $\prod_{i=1}^p (\mathcal{P} - x_i) y_{t-p}$ заменить линейной формой (при любых x_i)

$$(31) \quad \prod_{i=1}^r (\mathcal{P} - x_i^{-1}) \prod_{i=r+1}^p (\mathcal{P} - x_i) y_{t-p} = u_t^*,$$

то остатки u_t^* также будут некоррелированными (но уже не обязательно независимыми).

Рассмотрим теперь случай, когда имеется только один корень, равный 1,

$$(32) \quad y_t = y_{t-1} + u_t.$$

Предположим, что $\mathbb{E}u_t = 0$, $\mathbb{E}u_t^2 = \sigma^2$. Тогда

$$(33) \quad y_t - y_{t-s} = u_t + u_{t-1} + \dots + u_{t-s+1}$$

и

$$(34) \quad \mathbb{E}(y_t - y_{t-s})^2 = \mathbb{E}y_t^2 + \mathbb{E}y_{t-s}^2 - 2\mathbb{E}y_t y_{t-s} = s\sigma^2.$$

Если процесс стационарный, то $\mathbb{E}y_t^2 = \mathbb{E}y_{t-s}^2$. Из (34) при этом получаем

$$(35) \quad \mathbb{E}y_t y_{t-s} = \mathbb{E}y_t^2 - \frac{1}{2}s\sigma^2, \quad s = 1, 2, \dots$$

Последнее может выполняться для всех $s > 0$ только при $\sigma^2 = 0$, и поэтому $y_t = y_{t-s}$ с вероятностью 1. Интуитивно ясно, что дисперсия y_t должна возрастать с ростом t , если только дисперсия u_t не равна нулю, а это противоречит стационарности.

Более общим является случай

$$(36) \quad (\mathcal{P} - 1) \prod_{i=2}^p (\mathcal{P} - x_i) y_{t-p} = u_t,$$

где $|x_i| \neq 1$, $i = 2, \dots, p$. Если $\prod_{i=2}^p (\mathcal{P} - x_i) y_{t-p} = z_{t-1}$, то

$(\mathcal{P} - 1) z_{t-1} = u_t$ и $z_t = z_{t-s} = z$ с вероятностью 1. Таким образом,

$$(37) \quad \prod_{i=2}^p (\mathcal{P} - x_i) y_{t-p} = z$$

и $y_t = \sum_{s=-\infty}^{\infty} \delta_s z$, т. е. $y_t = y_{t-s}$ с вероятностью 1 (переменная z здесь может быть случайной величиной).

ТЕОРЕМА 5.2.2. Если стационарный случайный процесс удовлетворяет стохастическому разностному уравнению, характеристическое уравнение которого имеет хотя бы один корень, равный единице, то с вероятностью 1 все значения этого процесса совпадают.

Начиная с настоящего момента, мы будем ограничиваться рассмотрением случая, когда все корни характеристического уравнения (23) по абсолютной величине меньше единицы. При этом случайная величина y_t не будет зависеть от u_{t+1} , u_{t+2} , ... и ее можно представить в виде $\sum_{r=0}^{\infty} \delta_r u_{t-r}$. Обратимся теперь к коэффи-

циентам δ_r . Из (25) вытекает, что

$$\begin{aligned}
 (38) \quad 1 &= \left(\sum_{r=0}^p \beta_r z^r \right)^{-1} \sum_{s=0}^p \beta_s z^s = \sum_{r=0}^{\infty} \delta_r z^r \sum_{s=0}^p \beta_s z^s = \\
 &= \sum_{r=0}^{\infty} \sum_{s=0}^p \beta_s \delta_r z^{s+r} = \\
 &= \sum_{t=0}^{p-1} \left(\sum_{s=0}^t \beta_s \delta_{t-s} \right) z^t + \sum_{t=p}^{\infty} \left(\sum_{s=0}^p \beta_s \delta_{t-s} \right) z^t.
 \end{aligned}$$

Здесь мы заменили r на $t - s$. Поскольку последнее соотношение является тождеством относительно z (для $|z| \leq 1$ сходимость рядов равномерна), то коэффициент при z^0 в его правой части равен 1, а коэффициенты при положительных степенях равны 0. Запишем это в развернутой форме:

$$\begin{aligned}
 1 &= \beta_0 \delta_0 = \delta_0, \\
 0 &= \beta_0 \delta_1 + \beta_1 \delta_0 = \delta_1 + \beta_1, \\
 (39) \quad &\vdots \\
 0 &= \beta_0 \delta_{p-1} + \dots + \beta_{p-1} \delta_0,
 \end{aligned}$$

$$(40) \quad 0 = \beta_0 \delta_t + \dots + \beta_p \delta_{t-p}, \quad t = p, p+1, \dots$$

Уравнение (40) является однородным разностным уравнением, соответствующим неоднородному разностному уравнению (1). Если все корни характеристического уравнения (23) различны, то общее решение однородного разностного уравнения (40) имеет вид

$$(41) \quad \delta_r = \sum_{i=1}^p k_i x_i^r, \quad r = 0, 1, \dots$$

При этом если x_i действительно, то действителен и коэффициент k_i . Если x_i и x_{i+1} комплексно сопряжены, то k_i и k_{i+1} также комплексно сопряжены, а сумма $k_i x_i^r + k_{i+1} x_{i+1}^r$ действительна, $r = 0, 1, \dots$. В случае наличия кратных корней общее решение можно построить, используя упр. 9 и 10.

Уравнения (39) задают p граничных условий. Решая их последовательно, определим $\delta_0, \delta_1, \dots, \delta_{p-1}$. Совокупность соотношений (41) для $r = 0, 1, \dots, p-1$ представляет собой систему p линейно независимых линейных уравнений с p неизвестными $k_1, \dots, k_p$. Эта система имеет единственное решение. (См. также упр. 4.)

Если $p = 1$, то $\delta_r = (-\beta_1)^r$ является показательной функцией от r . Если $p = 2$, а x_1 и x_2 различны, то $k_1 = x_1/(x_1 - x_2)$, $k_2 = -x_2/(x_1 - x_2)$ и

$$(42) \quad \delta_r = \frac{x_1^{r+1} - x_2^{r+1}}{x_1 - x_2}, \quad r = 0, 1, \dots$$

Если x_1 и x_2 действительны, то δ_r будет линейной комбинацией двух показательных функций переменного r . В противном случае запишем x_1 и x_2 в виде $x_1 = \alpha e^{i\theta}$, $x_2 = \alpha e^{-i\theta}$. При этом

$$(43) \quad k_1 = \frac{e^{i\theta}}{e^{i\theta} - e^{-i\theta}}, \quad k_2 = -\frac{e^{-i\theta}}{e^{i\theta} - e^{-i\theta}}$$

и коэффициенты

$$(44) \quad \begin{aligned} \delta_r &= k_1 x'_1 + k_2 x'_2 = \alpha' \frac{e^{i\theta(r+1)} - e^{-i\theta(r+1)}}{e^{i\theta} - e^{-i\theta}} = \\ &= \alpha' \frac{\sin \theta (r+1)}{\sin \theta} \end{aligned}$$

образуют затухающую функцию синусоидального типа, что аналогично (42) для комплексно сопряженных корней.

5.2.2. Моменты второго порядка. Ковариационная функция

Если совместные распределения нормальны, то они полностью характеризуются средними $\mathcal{E}y_t$, которые мы из соображений удобства временно будем полагать равными нулю, дисперсиями $\mathcal{E}y_t^2$ и ковариациями $\mathcal{E}y_t y_{t+s}$. В случае когда совместные распределения отличны от нормальных, эти моменты первого и второго порядков тем не менее несут существенную информацию о рассматриваемом процессе. Например, корреляция между y_t и y_{t+s} (т. е. отношение $\mathcal{E}y_t y_{t+s} / \sqrt{\mathcal{E}y_t^2 \mathcal{E}y_{t+s}^2}$) является мерой связи этих двух случайных величин.

Если процесс y_t стационарный (в частности, если y_t может быть представлен в виде (28)), то все дисперсии $\mathcal{E}y_t^2$ совпадают, а ковариации зависят только от разности индексов рассматриваемых случайных величин. Эти моменты

$$(45) \quad \mathcal{E}y_t y_{t+s} = \sigma(s), \quad s = \dots, -1, 0, 1, \dots,$$

образуют в совокупности так называемую *ковариационную функцию* (иногда называемую *корреляционной функцией*). В этой книге мы будем называть *корреляционной функцией* нормированную функцию $\sigma(s)/\sigma(0)$.

Покажем, что ковариационная функция $\sigma(h)$ удовлетворяет однородному разностному уравнению (40). Умножая (1) на (28) с заменой t на $t-s$, получаем

$$(46) \quad \sum_{r=0}^p \beta_r y_{t-r} y_{t-s} = \sum_{q=-\infty}^{\infty} \delta_q u_t u_{t-s-q}.$$

Поскольку $\mathcal{E}y_{t-r}y_{t-s} = \sigma(s-r)$, $\mathcal{E}u_t^2 = \sigma^2$, $\mathcal{E}u_t u_{t'} = 0$, $t \neq t'$, математическое ожидание обеих частей (46) удовлетворяет для $s = 0$ и $s > 0$ соответственно соотношениям

$$(47) \quad \sum_{r=0}^p \beta_r \sigma(-r) = \sigma^2,$$

$$(48) \quad \sum_{r=0}^p \beta_r \sigma(s-r) = 0, \quad s = 1, 2, \dots$$

Эти соотношения часто называют *уравнениями Юла — Уолкера*. Таким образом, последовательность $\sigma(1-p)$, $\sigma(2-p)$, ..., ..., $\sigma(0)$, ... удовлетворяет однородному разностному уравнению (48) и поэтому

$$(49) \quad \sigma(h) = \sum_{i=1}^p c_i x_i^h, \quad h = 1-p, 2-p, \dots,$$

если корни x_i различны и $\beta_p \neq 0$ (так что $x_i \neq 0$). Граничными условиями будут $p-1$ соотношений

$$(50) \quad \sigma(-h) = \sigma(h), \quad h = 1, \dots, p-1$$

и соотношение (47), в котором $\sigma(-p)$ заменяется на $\sigma(p)$ (причем $\sigma(-p)$ не обязательно имеет вид (49)). Последнее условие служит для определения величины коэффициента пропорциональности.

Если $p = 1$, то $\sigma(h) = (-\beta_1)^h \sigma^2 / (1 - \beta_1^2)$, $h = 0, 1, \dots$, является показательной функцией h . Если $p = 2$, а x_1 и x_2 различны, то

$$(51) \quad \sigma(h) = \frac{\sigma^2}{(x_1 - x_2)(1 - x_1 x_2)} \left(\frac{x_1^{h+1}}{1 - x_1^2} - \frac{x_2^{h+1}}{1 - x_2^2} \right).$$

Если x_1 и x_2 действительны, то $\sigma(h)$ будет линейной комбинацией двух показательных функций. Если же они комплексно сопряжены и равны $\alpha e^{\pm i\theta}$, то (51) принимает вид

$$(52) \quad \sigma(h) = \frac{\sigma^2 \alpha^h [\sin \theta (h+1) - \alpha^2 \sin \theta (h-1)]}{(1 - \alpha^2) \sin \theta [1 - 2\alpha^2 \cos 2\theta + \alpha^4]}$$

и представляет собой затухающую линейную комбинацию функций от h синусоидального типа.

Поскольку $\sigma(h)$ является линейной комбинацией возведенных в степень h корней характеристического уравнения и все эти корни лежат в единичном круге, то $|\sigma(h)|$ ограничена сверху показательной убывающей функцией. Именно, для надлежащим образом выбранного $K > 0$, $|\sigma(h)| < K (\max_i |x_i|)^h$. Положительный корень вносит вклад в виде убывающей показательной функции; отрицательный — в виде знакопеременной показательной функции, убывающей по абсолютной величине. Пара комплексно сопряженных кор-

ней дает осциллирующую тригонометрическую функцию, убывающую по абсолютной величине. При этом период ее колебаний зависит от аргумента этих комплексных корней.

5.2.3. Флуктуации временных рядов

Типичный временной ряд, образованный с помощью модели стохастического разностного уравнения, флуктуирует. Наблюдаемые при этом колебания значений ряда нерегулярны. Однако они обнаруживают тенденцию к определенной средней продолжительности, значение которой зависит от параметров порождающего процесс разностного уравнения. Если считать, что последовательные значения ряда образуются согласно (1), то каждое множество из p последовательных y_t непосредственно влияет на следующее значение y_t . Так, при $p = 2$ можно записать y_t в виде

$$(53) \quad \begin{aligned} y_t &= -\beta_1 y_{t-1} - \beta_2 y_{t-2} + u_t = \\ &= -(\beta_1 + \beta_2) y_{t-1} + \beta_2 (y_{t-1} - y_{t-2}) + u_t. \end{aligned}$$

Из этого выражения видно, что непосредственное воздействие на значение y_t оказывают здесь значение y_{t-1} и разность $y_{t-2} - y_{t-1}$. Обычно благодаря этому и возникает тенденция к флуктуациям.

На указанный процесс можно взглянуть с иной точки зрения, используя представление $y_t = \sum_{r=0}^{\infty} \delta_r u_{t-r}$. Каждое u_s вносит в последующее y_{s+q} вклад, определяемый коэффициентом δ_q . Поскольку эти коэффициенты осциллируют, то воздействие на последовательные y_t флуктуирует, вызывая тенденцию к флуктуации у наблюдаемого ряда y_t .

На подобный характер поведения ряда указывает, кроме того, и ковариационная функция. Поскольку (обычно) $\sigma(s) \neq 0$, то между y_t и y_{t+s} имеется определенная статистическая связь, хотя и имеющая тенденцию к уменьшению при увеличении $s > 0$. Эта связь (измеряемая коэффициентом корреляции) и ведет к возникновению флуктуаций. Если характеристическое уравнение имеет пару комплексных корней, то соответствующие тригонометрические функции осциллируют и указанная связь может на некотором интервале изменения s возрасть вместе с s .

5.2.4. Введение «независимых» переменных

Пусть известно, что на рассматриваемый временной ряд могут оказывать воздействие некоторые внешние переменные $z_{1t}, \dots, z_{qt}$. Их влияние можно учесть, заменяя (1) более сложной моделью

$$(54) \quad \sum_{r=0}^p \beta_r y_{t-r} + \sum_{i=1}^q \gamma_i z_{it} = u_t.$$

В частности, подобным образом в модель всегда можно ввести константу (беря $q = 1$, $z_{1t} = 1$) и положить при этом $\mathbb{E}u_t = 0$.

Выражение (54) можно записать иначе:

$$\begin{aligned}
 (55) \quad u_t &= \prod_{i=1}^p (\mathcal{P} - x_i) y_{t-p} + \sum_{i=1}^q \gamma_i z_{it} = \\
 &= \prod_{i=1}^p (\mathcal{P} - x_i) \left[y_{t-p} + \sum_{i=1}^q \gamma_i \prod_{j=1}^p (\mathcal{P} - x_j)^{-1} z_{it} \right] = \\
 &= \prod_{j=1}^p (\mathcal{P} - x_j) \left[y_{t-p} + \sum_{i=1}^q \gamma_i \sum_{s=0}^{\infty} \delta_s z_{i,t-p-s} \right] = \\
 &= \sum_{r=0}^p \beta_r \left[y_{t-r} + \sum_{i=1}^q \gamma_i \sum_{s=0}^{\infty} \delta_s z_{i,t-r-s} \right].
 \end{aligned}$$

При $q = 1$, $z_{1t} = 1$ оно принимает вид

$$(56) \quad \sum_{r=0}^p \beta_r [y_{t-r} - \mu] = u_t,$$

где $\mu = -\gamma_1 \sum_{s=0}^{\infty} \delta_s$.

5.2.5. Прогнозирование

Стохастическое разностное уравнение (1) можно переписать в виде

$$(57) \quad y_t = -\beta_1 y_{t-1} - \dots - \beta_p y_{t-p} + u_t,$$

где u_t не зависит от $y_{t-1}, y_{t-2}, \dots$. Условное математическое ожидание случайной величины y_t при заданных значениях $y_{t-1}, y_{t-2}, \dots$ есть

$$(58) \quad \mathbb{E}\{y_t | y_{t-1}, y_{t-2}, \dots\} = -\beta_1 y_{t-1} - \dots - \beta_p y_{t-p}.$$

Это условное математическое ожидание можно использовать в качестве прогноза случайной величины y_t . Дисперсия его выражается соотношением

$$(59) \quad \mathbb{E}[-\beta_1 y_{t-1} - \dots - \beta_p y_{t-p} - y_t]^2 = \mathbb{E}u_t^2 = \sigma^2;$$

среднеквадратичная ошибка любого другого прогноза $f(y_{t-1}, y_{t-2}, \dots)$ будет иметь вид

$$\begin{aligned}
 (60) \quad \mathbb{E}[f(y_{t-1}, y_{t-2}, \dots) - y_t]^2 &= \\
 &= \mathbb{E}[f(y_{t-1}, y_{t-2}, \dots) + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} - u_t]^2 = \\
 &= \mathbb{E}u_t^2 + \mathbb{E}[f(y_{t-1}, y_{t-2}, \dots) + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p}]^2.
 \end{aligned}$$

ТЕОРЕМА 5.2.3. Если стационарный процесс y_t удовлетворяет стохастическому разностному уравнению (1) и все характеристические корни этого уравнения лежат в единичном круге, то прогноз (58) имеет наименьшую среднеквадратичную ошибку среди всех прогнозов, использующих значения $y_{t-1}, y_{t-2}, \dots$.

Наилучшим прогнозом случайной величины y_t , использующим значения $y_{t-s-1}, y_{t-s-2}, \dots$ ($s \geq 0$), будет прогноз

$$(61) \quad \mathbb{E}\{y_t | y_{t-s-1}, y_{t-s-2}, \dots\} = \alpha_{s1}y_{t-s-1} + \dots + \alpha_{sp}y_{t-s-p}.$$

Коэффициенты $\alpha_{s1} = \alpha_{s1}^*, \dots, \alpha_{sp} = \alpha_{sp}^*$ задаются соотношениями (11) и (22).

5.3. РЕДУКЦИЯ ОБЩЕГО СКАЛЯРНОГО УРАВНЕНИЯ К ВЕКТОРНОМУ УРАВНЕНИЮ ПЕРВОГО ПОРЯДКА

Для изучения стохастического разностного уравнения (1) § 5.2 его удобно записать в форме векторного уравнения первого порядка. Положим

$$(1) \quad \tilde{y}_t = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix},$$

$$(2) \quad \tilde{u}_t = \begin{pmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (p \text{ строк}),$$

$$(3) \quad \tilde{B} = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \dots & \beta_{p-1} & \beta_p \\ -1 & 0 & 0 & \dots & 0 & 0 \\ 0 & -1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 0 \end{bmatrix}.$$

Тогда указанное стохастическое разностное уравнение можно переписать в виде

$$(4) \quad \tilde{y}_t + \tilde{B}\tilde{y}_{t-1} = \tilde{u}_t.$$

Первое скалярное уравнение из (4) есть в точности уравнение (1) из § 5.2, а все остальные являются тождествами. Таким образом, скалярное стохастическое разностное уравнение порядка p оказывается частным случаем векторного стохастического разностного уравнения первого порядка.

Рассмотрим теперь общее стохастическое разностное уравнение первого порядка относительно вектора с p компонентами:

$$(5) \quad \mathbf{y}_t + \mathbf{B}\mathbf{y}_{t-1} = \mathbf{u}_t.$$

Здесь $\mathbf{u}_t$ — случайный вектор с нулевым средним и ковариационной матрицей $\mathbf{E}\mathbf{u}_t\mathbf{u}_t' = \Sigma$, причем $\mathbf{E}\mathbf{u}_t\mathbf{u}_s' = \mathbf{0}$, $t \neq s$, а матрица $\mathbf{B} = (\beta_{ij})$. Если в (5) последовательно производить подстановку

$$(6) \quad \mathbf{y}_s = -\mathbf{B}\mathbf{y}_{s-1} + \mathbf{u}_s$$

для $s = t-1, t-2, \dots$, то в конце концов придем к представлению

$$(7) \quad \mathbf{y}_t = \sum_{\tau=0}^{\infty} (-\mathbf{B})^{\tau} \mathbf{u}_{t-\tau}.$$

Пусть $\lambda_1, \lambda_2, \dots, \lambda_p$ — характеристические корни матрицы $-\mathbf{B}$, т. е. корни уравнения

$$(8) \quad |\mathbf{B} + \lambda \mathbf{I}| = 0.$$

Если все эти корни различны, то найдется матрица $\mathbf{C}$, такая, что

$$(9) \quad -\mathbf{B} = \mathbf{C}\mathbf{A}\mathbf{C}^{-1},$$

где $\mathbf{A}$ — диагональная матрица ¹⁾

$$(10) \quad \mathbf{A} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_p \end{pmatrix}.$$

В этом параграфе, а также в § 5.5 и 5.6 мы получим ряд результатов, справедливых в случае, когда все характеристические корни

¹⁾ Если существует p линейно независимых характеристических векторов, то (10) является канонической формой матрицы $-\mathbf{B}$. Более общей является Жорданова каноническая форма, имеющая диагональные блоки вида

$$\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix},$$

в которых λ — корень характеристического уравнения (8). (См. Халмош (1958) или Тернбулл и Эйкен (1952). Они называют такую форму классической канонической формой.) Каждому из линейно независимых характеристических векторов будет при этом соответствовать ровно один такой блок. (См. упр. 17 и 18.)

матрицы $\mathbf{B}$ (или $\tilde{\mathbf{B}}$) лежат в единичном круге. Так как вычисления с использованием общей жордановой канонической формы трудоемки, будем доказывать соответствующие результаты исходя из диагонального характера матрицы $\mathbf{A}$. Что касается общего случая, то здесь мы предлагаем читателю обратиться к подстрочным примечаниям и к упражнениям.

Из соотношения (9) имеем

$$(11) \quad (-\mathbf{B})^\tau = \mathbf{C} \mathbf{A}^\tau \mathbf{C}^{-1},$$

где ¹⁾

$$(12) \quad \mathbf{A}^\tau = \begin{pmatrix} \lambda_1^\tau & 0 & \dots & 0 \\ 0 & \lambda_2^\tau & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_p^\tau \end{pmatrix}.$$

Поэтому (7) можно записать в виде

$$(13) \quad \mathbf{y}_t = \mathbf{C} \sum_{\tau=0}^{\infty} \mathbf{A}^\tau \mathbf{C}^{-1} \mathbf{u}_{t-\tau}.$$

Ряд в правой части сходится тогда и только тогда (в среднеквадратичном, см. разд. 7.6.1), когда $|\lambda_i| < 1$, $i = 1, \dots, p$. Действительно,

$$(14) \quad \mathbf{y}_t - \sum_{\tau=0}^s (-\mathbf{B})^\tau \mathbf{u}_{t-\tau} = \mathbf{y}_t - \mathbf{C} \sum_{\tau=0}^s \mathbf{A}^\tau \mathbf{C}^{-1} \mathbf{u}_{t-\tau} = \\ = \mathbf{C} \mathbf{A}^{s+1} \mathbf{C}^{-1} \mathbf{y}_{t-(s+1)}.$$

В предположении стационарности ковариационная матрица

$$(15) \quad \mathbb{E} \mathbf{y}_{t-(s+1)} \mathbf{y}_{t-(s+1)}' = \mathbf{F}$$

не зависит ни от t , ни от s . Поэтому ковариационная матрица разности (14) равна

$$(16) \quad \mathbf{C} \mathbf{A}^{s+1} \mathbf{C}^{-1} \mathbf{F} (\mathbf{C}^{-1})' \mathbf{A}^{s+1} \mathbf{C}.$$

Поскольку каждый элемент матрицы (16) является линейной комбинацией элементов $(\lambda_i \lambda_j)^{s+1}$, он будет сходиться к нулю при $s \rightarrow \infty$, только когда каждое $|\lambda_i| < 1$ ²⁾. (Для того чтобы было верно и

¹⁾ В случае когда матрица $\mathbf{A}$ имеет общую жорданову каноническую форму, диагональный блок размера $m \times m$ матрицы $\mathbf{A}^\tau$, соответствующий характеристическому корню λ_i , будет иметь следующий вид. Все его элементы, расположенные на j мест выше главной диагонали, равны $\binom{\tau}{j} \lambda_i^{\tau-j}$, $j = 0, 1, \dots, m-1$. Элементы же, расположенные под главной диагональю, равны нулю. (См. упр. 17.)

²⁾ В случае общей жордановой канонической формы $(\lambda_i \lambda_j)^{s+1}$ заменяются полиномами от s (степени не выше чем $2p-2$), умноженными на $\lambda_i^{s+1-k} \lambda_j^{s+1-l}$.

обратное, т. е. чтобы условие $|\lambda_i| < 1$, $i = 1, \dots, p$, было также и необходимым для сходимости ряда в (13), требуется невырожденность ковариационной матрицы вектора y_t .)

Обратимся снова к соотношению (4), представляющему собой векторную форму записи скалярного уравнения порядка p . Покажем, что корни характеристического уравнения (23) из § 5.2. совпадают с корнями уравнения

$$(17) \quad |\tilde{\mathbf{B}} + \lambda \mathbf{I}| = 0.$$

Определитель

$$(18) \quad |\tilde{\mathbf{B}} + \lambda \mathbf{I}| = \begin{vmatrix} \beta_1 + \lambda & \beta_2 & \beta_3 & \dots & \beta_{p-1} & \beta_p \\ -1 & \lambda & 0 & \dots & 0 & 0 \\ 0 & -1 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 0 \\ 0 & 0 & 0 & \dots & -1 & \lambda \end{vmatrix}$$

легко вычислить последовательно, прибавляя к каждому $(i+1)$ -му столбцу i -й столбец, предварительно умноженный на λ , $i = 1, \dots, p-1$. При этом получим

$$(19) \quad \begin{vmatrix} \beta_1 + \lambda & \beta_2 + \beta_1 \lambda + \lambda^2 & \dots & \beta_p + \lambda \beta_{p-1} + \dots + \lambda^{p-1} \beta_1 + \lambda^p \\ -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{vmatrix} =$$

$$= \beta_p + \lambda \beta_{p-1} + \dots + \lambda^{p-1} \beta_1 + \lambda^p,$$

а это и есть характеристический полином из (23) § 5.2. Поскольку предполагается, что все корни этого полинома лежат в единичном круге, то таковыми же будут и характеристические корни матрицы $-\tilde{\mathbf{B}}$. Отсюда вытекает, что сумма

$$(20) \quad \tilde{\mathbf{y}}_t = \sum_{\tau=0}^{\infty} (-\tilde{\mathbf{B}})^{\tau} \tilde{\mathbf{u}}_{t-\tau}$$

является определенной (т. е. ряд (20) сходится в среднем).

Если все корни $x_1, \dots, x_p$ различны¹⁾, то найдется матрица C , такая, что $-\tilde{B} = C \Lambda C^{-1}$, где

$$(21) \quad \Lambda = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x_p \end{pmatrix}.$$

Тогда, как следует из предыдущего, (20) можно записать аналогично (13). При этом первая компонента равна

$$(22) \quad y_t = \sum_{\tau=0}^{\infty} \sum_{j=1}^p c_{1j} x_j^{\tau} c^{j1} u_{t-\tau},$$

где²⁾ $(c^{ji}) = C^{-1}$. Это соотношение эквивалентно представлению (28) § 5.2. Таким образом, представление (28) можно получить еще одним способом. Именно, для этого следует найти матрицу C , столбцы которой являются характеристическими векторами матрицы $-\tilde{B}$, и использовать (22). (См. упр. 20.)

В случае общего векторного уравнения соотношения (13) можно использовать для отыскания ковариационной матрицы вектора y_t при заданных матрицах B и Σ . Действительно (при $\xi y_t = 0$), искомая матрица имеет вид

$$(23) \quad \begin{aligned} \xi y_t y_t' &= \sum_{\tau, s=0}^{\infty} (-B)^{\tau} \xi u_{t-\tau} u_{t-s}' (-B')^s = \\ &= \sum_{\tau=0}^{\infty} B^{\tau} \Sigma B'^{\tau} = \\ &= C \sum_{\tau=0}^{\infty} \Lambda^{\tau} C^{-1} \Sigma (C^{-1})' \Lambda^{\tau} C' = F. \end{aligned}$$

Здесь³⁾

$$(24) \quad \begin{aligned} f_{ij} &= \sum_{g, h, k, l=1}^p c_{ig} \sum_{\tau=0}^{\infty} \lambda_g^{\tau} c^{gh} \sigma_{hk} c^{lk} \lambda_l^{\tau} c_{jl} = \\ &= \sum_{g, l=1}^p c_{ig} \frac{\sum_{h, k=1}^p c^{gh} \sigma_{hk} c^{lk}}{1 - \lambda_g \lambda_l} c_{jl}, \end{aligned}$$

¹⁾ Из упр. 19 вытекает, что если имеются кратные корни, то при этом не существует p линейно независимых характеристических векторов и каноническая форма состоит из диагональных блоков, как это было указано в предыдущих примечаниях.

²⁾ В случае общей жордановой канонической формы $\sum_{j=1}^p c_{1j} x_j^{\tau} c^{j1}$ следует надлежащим образом изменить.

³⁾ В случае общей жордановой канонической формы вместо суммирования по индексу τ произведений $\lambda_g^{\tau} \lambda_l^{\tau}$ возникает суммирование степеней $\lambda_g^{\tau-u} \lambda_l^{\tau-v}$, умноженных на полиномы от τ (степени не выше чем $2p - 2$). Такие суммы сходятся при $|\lambda_l| < 1$.

Если $\mathbf{C}$ и $\mathbf{A}$ определяются исходя из матрицы $\tilde{\mathbf{B}}$, то выражение (24) можно упростить: $\sigma(0) = f_{i0}$, $\sigma(1) = f_{i,1+1} = f_{i+1,1}$, ..., $\sigma(p-1) = f_{1p} = f_{p1}$.

Если умножить (5) на транспонированное (7), а также на транспонированное (7) с заменой в последнем индекса t на $t-s$, то получим

$$(25) \quad \mathbf{C} \mathbf{y}_t \mathbf{y}'_t + \mathbf{B} \mathbf{C} \mathbf{y}_{t-1} \mathbf{y}'_t = \Sigma,$$

$$(26) \quad \mathbf{C} \mathbf{y}_t \mathbf{y}'_{t-s} + \mathbf{B} \mathbf{C} \mathbf{y}_{t-1} \mathbf{y}'_{t-s} = 0, \quad s = 1, 2, \dots$$

Поскольку $\mathbf{C} \mathbf{y}_{t-1} \mathbf{y}'_{t-s} = \mathbf{C} \mathbf{y}_t \mathbf{y}'_{t-(s-1)}$, мы можем, действуя последовательно, получить из (26)

$$(27) \quad \mathbf{C} \mathbf{y}_t \mathbf{y}'_{t-s} = (-\mathbf{B})^s \mathbf{C} \mathbf{y}_t \mathbf{y}'_t.$$

Беря последнее выражение при $s = 1$ и подставляя его в транспонированном виде в (25) вместо $\mathbf{C} \mathbf{y}_{t-1} \mathbf{y}'_t$, получим соотношение

$$(28) \quad \mathbf{C} \mathbf{y}_t \mathbf{y}'_t - \mathbf{B} \mathbf{C} \mathbf{y}_t \mathbf{y}'_t \mathbf{B}' = \Sigma$$

или

$$(29) \quad \mathbf{F} - \mathbf{B} \mathbf{F} \mathbf{B}' = \Sigma.$$

Решение системы линейных уравнений (29) является, таким образом, еще одним способом отыскания матрицы $\mathbf{F}$. При этом $\mathbf{F}_0 = \Sigma$ можно взять в качестве начального приближения и применить итерационный метод, вычисляя последовательные приближения

$\mathbf{F}_i = \Sigma + \mathbf{B} \mathbf{F}_{i-1} \mathbf{B}'$. Если $\mathbf{B} = \tilde{\mathbf{B}}$ и

$$(30) \quad \Sigma = \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$$

это даст нам возможность вычисления $\sigma(h)$.

В векторное уравнение независимые переменные можно ввести, полагая

$$(31) \quad \tilde{\Gamma} = \begin{pmatrix} \gamma_1 & \gamma_2 & \dots & \gamma_q \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \quad \mathbf{z}_t = \begin{pmatrix} z_{1t} \\ z_{2t} \\ \vdots \\ z_{qt} \end{pmatrix}.$$

При этом модель принимает вид

$$(32) \quad \tilde{\mathbf{y}}_t + \tilde{\mathbf{B}} \tilde{\mathbf{y}}_{t-1} + \tilde{\Gamma} \mathbf{z}_t = \tilde{\mathbf{u}}_t$$

и

$$(33) \quad \tilde{\mathbf{y}}_t = \sum_{\tau=0}^{\infty} (-\tilde{\mathbf{B}})^{\tau} \tilde{\mathbf{u}}_{t-\tau} - \sum_{\tau=0}^{\infty} (-\tilde{\mathbf{B}})^{\tau} \tilde{\Gamma} \mathbf{z}_{t-\tau}.$$

5.4. ОЦЕНКИ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ В СЛУЧАЕ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

Одной из первоочередных задач при статистическом исследовании модели стохастического разностного уравнения является оценка коэффициентов $\beta_1, \dots, \beta_p$ и дисперсии σ^2 по наблюдениям отрезка ряда $y_1, \dots, y_T$. Если u_t распределены нормально и модель описывается разностным уравнением (1) из § 5.2 для всех t , то набор этих коэффициентов вместе с дисперсией σ^2 (при $\mathbb{E}u_t = 0$) полностью определяет распределение величин y_t . Мы займемся более общей задачей оценки коэффициентов в модели

$$(1) \quad \sum_{r=0}^p \beta_r y_{t-r} + \sum_{i=1}^q \gamma_i z_{it} = u_t, \quad t = 1, \dots, T.$$

Найдем оценки максимального правдоподобия параметров $\beta_1, \dots, \beta_p, \gamma_1, \dots, \gamma_q$ и σ^2 . При этом будем предполагать, что величины u_t независимы и нормально распределены с нулевыми средними и дисперсиями $\mathbb{E}u_t^2 = \sigma^2$. Исходное предположение о стационарности процесса мы изменим, полагая (помимо того, что мы ввели в модель переменные z_{it}), что наблюдения y_t начинаются в точке $t = -(p-1)$ и что при этом $y_{-(p-1)}, y_{-(p-2)}, \dots, y_0$ — заданные известные числа. Тогда совместное распределение величин $u_1, \dots, u_T$ будет полностью определять совместное распределение наблюдаемых величин $y_1, \dots, y_T$ (заметим, что $z_{11}, \dots, z_{q1}, \dots, z_{1T}, \dots, z_{qT}$ — заданные числа). Указанное предположение сделано для удобства отыскания оценок максимального правдоподобия. (В гл. 6 рассмотрены трудности, возникающие при отсутствии такого предположения.) Заметим, что любая процедура оценивания для этой модели является оценкой максимального правдоподобия или ее незначительной модификацией. Во всяком случае при больших T влияние такого предположения невелико, а в асимптотической теории оно и вовсе не требуется. Оценки максимального правдоподобия и соответствующая асимптотическая теория для рассматриваемой модели были предложены Манном и Вальдом (1943b).

Обозначим

$$(2) \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_q \end{pmatrix}, \quad \tilde{y}_t = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix}, \quad z_t = \begin{pmatrix} z_{1t} \\ z_{2t} \\ \vdots \\ z_{qt} \end{pmatrix}.$$

Полагая $\beta_0 = 1$, уравнение (1) запишем в виде

$$(3) \quad y_t + \beta' \tilde{y}_{t-1} + \gamma' z_t = u_t, \quad t = 1, \dots, T.$$

Поскольку $u_1, \dots, u_T$ не зависят от $y_{-p+1}, \dots, y_0$, то условная плотность распределения случайных величин $u_1, \dots, u_T$ при заданных значениях $y_{-p+1}, \dots, y_0$ совпадает с безусловной и равна

$$(4) \quad \frac{1}{(2\pi\sigma^2)^{T/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^T u_t^2\right).$$

При фиксированных значениях $y_{-p+1}, \dots, y_0$ соотношение (3) определяет взаимно однозначное отображение переменных $y_1, \dots, y_T$ на переменные $u_1, \dots, u_T$. Якобиан этого преобразования есть

$$(5) \quad \left| \frac{\partial u_t}{\partial y_s} \right| = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ \beta_1 & 1 & 0 & \dots & 0 \\ \beta_2 & \beta_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix} = 1.$$

Отсюда следует, что при заданных значениях $y_{-p+1}, \dots, y_0$ совместная плотность вероятностей значений $y_1, \dots, y_T$ есть

$$(6) \quad \frac{1}{(2\pi\sigma^2)^{T/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{t=1}^T (y_t + \beta' \tilde{y}_{t-1} + \gamma' z_t)^2\right].$$

Если это выражение рассматривать как функцию правдоподобия по отношению к параметрам β и γ , то для получения оценок максимального правдоподобия для этих параметров необходимо максимизировать последнее выражение на множестве всех возможных значений β и γ . Для этого в свою очередь достаточно найти значения параметров β и γ , при которых достигает минимума сумма

$$(7) \quad \sum_{t=1}^T (y_t + \beta' \tilde{y}_{t-1} + \gamma' z_t)^2 = \sum_{t=1}^T \left[y_t - (-\beta' - \gamma') (\tilde{y}_{t-1}) \right]^2.$$

Таким образом, мы приходим к обычной задаче наименьших квадратов. Нормальными уравнениями относительно значений $\hat{\beta} = \hat{\beta}$ и $\hat{\gamma} = \hat{\gamma}$, минимизирующих сумму (7), будут здесь уравнения

$$(8) \quad \sum_{t=1}^T \begin{pmatrix} \tilde{y}_{t-1} \\ z_t \end{pmatrix} (\tilde{y}_{t-1}' z_t') \begin{pmatrix} -\hat{\beta} \\ -\hat{\gamma} \end{pmatrix} = \sum_{t=1}^T y_t \begin{pmatrix} \tilde{y}_{t-1} \\ z_t \end{pmatrix}.$$

Их можно записать в матричной форме следующим образом:

$$(9) \quad \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = - \begin{pmatrix} a_0 \\ d \end{pmatrix},$$

где

$$(10) \quad \begin{aligned} A_{11} &= \sum_{t=1}^T \tilde{y}_{t-1} \tilde{y}'_{t-1}, & A_{12} &= A'_{21} = \sum_{t=1}^T \tilde{y}_{t-1} z'_t, \\ A_{22} &= \sum_{t=1}^T z_t z'_t, & a_0 &= \sum_{t=1}^T y_t \tilde{y}_{t-1}, & d &= \sum_{t=1}^T y_t z_t. \end{aligned}$$

Оценкой максимального правдоподобия для σ^2 будет

$$(11) \quad \hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^T (y_t + \hat{\beta}' \tilde{y}_{t-1} + \hat{\gamma}' z_t)^2.$$

Непосредственно видно, что плотность (6) зависит от значений наблюдаемых переменных через выражение (7). Последнее же равно

$$(12) \quad \sum_{t=1}^T y_t^2 + 2\beta' a_0 + 2\gamma' d + \beta' A_{11} \beta + 2\beta' A_{12} \gamma + \gamma' A_{22} \gamma.$$

Отсюда следует, что $\sum_{t=1}^T y_t^2$, a_0 , d , A_{11} и A_{12} образуют достаточное множество статистик.

В частном случае $q = 1$, $z_{1t} = 1$ положим $\gamma_1 = \gamma$; уравнения для оценок примут вид

$$(13) \quad \sum_{j=1}^p a_{ij} \hat{\beta}_j + T \bar{y}_{(i)} \hat{\gamma} = -a_{i0}, \quad i = 1, \dots, p,$$

$$(14) \quad T \sum_{j=1}^p \bar{y}_{(j)} \hat{\beta}_j + T \hat{\gamma} = -T \bar{y},$$

где

$$(15) \quad a_{ij} = \sum_{t=1}^T y_{t-i} y_{t-j}, \quad i, j = 1, \dots, p,$$

$$(16) \quad a_{i0} = \sum_{t=1}^T y_{t-i} y_t, \quad i = 1, \dots, p,$$

$$(17) \quad T \bar{y}_{(i)} = \sum_{t=1}^T y_{t-i}, \quad i = 0, 1, \dots, p,$$

и $\bar{y} = \bar{y}_{(0)}$. Если из i -го уравнения (13) вычесть умноженное на $y_{(i)}$ уравнение (14), то получим

$$(18) \quad \sum_{j=1}^p \left[\sum_{t=1}^T y_{t-i} y_{t-j} - T \bar{y}_{(i)} \bar{y}_{(j)} \right] \hat{\beta}_j = - \sum_{t=1}^T y_{t-i} y_t + T \bar{y}_{(i)} \bar{y},$$

$i = 1, \dots, p,$

или, что равносильно,

$$(19) \quad \sum_{j=1}^p \sum_{i=1}^T (y_{t-i} - \bar{y}_{(i)}) (y_{t-j} - \bar{y}_{(j)}) \hat{\beta}_j = - \sum_{i=1}^T (y_{t-i} - \bar{y}_{(i)}) (y_t - \bar{y}),$$

$$i = 1, \dots, p.$$

Суммы по t в (19), соответствующие одному и тому же значению разности $(i - j)$, отличаются только добавлением или вычитанием крайних членов. Эти суммы можно различным образом изменять. Например, можно вместо (19) записать уравнения

$$(20) \quad \sum_{j=1}^p C_{i-j}^* b_j = -C_i^*, \quad i = 1, \dots, p,$$

$$(21) \quad \sum_{j=1}^p c_{i-j}^* b_j = -c_i^*, \quad i = 1, \dots, p,$$

$$(22) \quad \sum_{j=1}^p r_{i-j}^* b_j = -r_i^*, \quad i = 1, \dots, p.$$

Из соображений удобства и для согласования с последующим изложением определим входящие сюда величины следующим образом:

$$(23) \quad C_h^* = C_{-h}^* = \frac{1}{T-h} \sum_{i=1}^{T-h} (y_i - \bar{y}) (y_{i+h} - \bar{y}),$$

$$h = 0, 1, \dots, T-1,$$

$$(24) \quad c_h^* = c_{-h}^* = \frac{1}{T} \sum_{i=1}^{T-h} (y_i - \bar{y}) (y_{i+h} - \bar{y}),$$

$$h = 0, 1, \dots, T-1,$$

$r_h^* = c_h^*/c_0^*$, $h = 0, 1, \dots, T-1$, и $\bar{y} = \sum_{i=1}^T y_i/T$. Отметим, что (22) получается из (21) делением всех c_h^* на c_0^* .

Если T относительно велико, то различие между уравнениями максимального правдоподобия (19) (при наблюдениях $y_{-p+1}, \dots, y_T$) и уравнениями (20), (21), (22) (включающими только наблюдения $y_1, \dots, y_T$) становится незначительным. При этом последние уравнения обладают большей симметрией. Каждая диагональ матрицы коэффициентов состоит в этом случае из одинаковых элементов. Что касается правых частей, то они образованы из тех же элементов, к которым добавляется еще один.

Если обозначить

$$(25) \quad \mathbf{R}^* = \begin{bmatrix} 1 & r_1^* & \dots & r_{p-1}^* \\ r_1^* & 1 & \dots & r_{p-2}^* \\ \vdots & \vdots & \ddots & \vdots \\ r_{p-1}^* & r_{p-2}^* & \dots & 1 \end{bmatrix}, \quad \mathbf{r}^* = \begin{bmatrix} r_1^* \\ r_2^* \\ \vdots \\ r_p^* \end{bmatrix},$$

то (22) можно записать в виде

$$(26) \quad \mathbf{R}^* \mathbf{b} = -\mathbf{r}^*.$$

Решением уравнений (26) будет $\mathbf{b} = -(\mathbf{R}^*)^{-1} \mathbf{r}^*$. Эти уравнения обладают тем преимуществом, что они используют лишь p статистик, а матрица $\mathbf{R}^*$ является положительно определенной. (Матрицу $\mathbf{R}^*$ можно рассматривать как пропорциональную матрицу, у которой элемент, находящийся на пересечении i -й строки и j -го столбца, равен $\sum_{t=-p}^T (y_{t+i} - \bar{y})(y_{t+j} - \bar{y})$, и разность $y_{t+i} - \bar{y}$ полагается равной нулю при $t+i \leq 0$ и $t+i > T$.) Кроме того, матрица $\mathbf{R}^*$ симметрична и относительно главной диагонали и диагонали, идущей из левого нижнего в правый верхний угол матрицы.

Специфическая форма $\mathbf{R}^*$ и $\mathbf{r}^*$ приводит к простому рекуррентному способу вычисления коэффициентов уравнения $(p+1)$ -го порядка по решению соответствующего уравнения p -го порядка. Уравнения относительно вектора коэффициентов уравнения $(p+1)$ -го порядка, который мы обозначим $\mathbf{b}(p+1)$, имеют вид

$$(27) \quad \mathbf{R}_{p+1}^* \mathbf{b}(p+1) = -\mathbf{r}_{p+1}^*,$$

где

$$(28) \quad \mathbf{R}_{p+1}^* = \begin{pmatrix} \mathbf{R}_p^* & \tilde{\mathbf{r}}_p^* \\ \tilde{\mathbf{r}}_p^* & 1 \end{pmatrix}, \quad \mathbf{b}(p+1) = \begin{pmatrix} \mathbf{b}^{(1)}(p+1) \\ b_{p+1}(p+1) \end{pmatrix},$$

$$\mathbf{r}_{p+1}^* = \begin{pmatrix} \mathbf{r}_p^* \\ r_{p+1}^* \end{pmatrix},$$

$\mathbf{R}_p^*$ и $\mathbf{r}_p^*$ задаются соотношениями (25), а $\tilde{\mathbf{r}}_p^* = (r_p^*, r_{p-1}^*, \dots, r_1^*)'$ есть просто $\mathbf{r}_p^*$ с координатами, записанными в обратном порядке. При этом (27) распадается на два уравнения

$$(29) \quad \mathbf{R}_p^* \mathbf{b}^{(1)}(p+1) + \tilde{\mathbf{r}}_p^* b_{p+1}(p+1) = -\mathbf{r}_p^*,$$

$$(30) \quad \tilde{\mathbf{r}}_p^* \mathbf{b}^{(1)}(p+1) + b_{p+1}(p+1) = -r_{p+1}^*.$$

Исключив $\mathbf{b}^{(1)}(p+1)$ из (29) и (30), получим

$$(31) \quad b_{p+1}(p+1) = - \frac{r_{p+1}^* - \tilde{r}_p^{*'} (R_p^*)^{-1} r_p^*}{1 - \tilde{r}_p^{*'} (R_p^*)^{-1} \tilde{r}_p^*} = - \frac{r_{p+1}^* + \tilde{r}_p^{*'} \tilde{\mathbf{b}}(p)}{1 + \tilde{r}_p^{*'} \tilde{\mathbf{b}}(p)},$$

где вектор $\tilde{\mathbf{b}}(p) = (b_p(p), b_{p-1}(p), \dots, b_1(p))'$ получается из вектора $\mathbf{b}(p)$ изменением порядка его координат на противоположный. Подстановка (31) в (29) дает

$$(32) \quad \mathbf{b}^{(1)}(p+1) = - (R_p^*)^{-1} r_p^* - (R_p^*)^{-1} \tilde{r}_p^{*'} b_{p+1}(p+1) = \\ = \mathbf{b}(p) + \tilde{\mathbf{b}}(p) b_{p+1}(p+1).$$

В покомпонентной записи (31) и (32) имеют вид

$$(33) \quad b_{p+1}(p+1) = - \frac{r_{p+1}^* + \sum_{j=1}^p r_{p+1-j}^* b_j(p)}{1 + \sum_{j=1}^p r_j^* b_j(p)},$$

$$(34) \quad b_j(p+1) = b_j(p) + b_{p+1-j}(p) b_{p+1}(p+1), \\ j = 1, \dots, p.$$

Полученные уравнения позволяют легко вычислять оценки для процесса заданного порядка. Промежуточные вычисления дают коэффициенты для всех процессов более низкого порядка. В § 5.6 будет показано, что $b_{p+1}(p+1)$ есть частная корреляция между y_t и y_{t+p+1} , «при фиксированных величинах $y_{t+1}, \dots, y_{t+p}$ ». Этот метод соответствует стандартному методу добавления в уравнение регрессии независимой переменной. Упрощенные формулы для этого случая были предложены Дурбином (1960а).

Мы уже замечали, что уравнения (9) для оценок коэффициентов стохастического разностного уравнения формально совпадают с уравнениями, используемыми в методе наименьших квадратов. В случае обычной регрессии оценки наименьших квадратов являются наилучшими несмещенными линейными оценками. Получаемые здесь оценки не линейны по y_t . Обращаясь к уравнениям (9), Дурбин (1960b) назвал их линейными (относительно оценок) и несмещенными, в том смысле, что при замене в (9) оценок на истинные значения параметров математические ожидания правой и левой частей совпадают. В пределах такого класса уравнения (9) являются наилучшими в том смысле, что дисперсия разности правой и левой частей минимальна.

5.5. АСИМПТОТИЧЕСКОЕ РАСПРЕДЕЛЕНИЕ ОЦЕНОК МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

5.5.1. Обсуждение асимптотических свойств

В этом параграфе мы покажем, что оценки максимального правдоподобия, полученные в § 5.4, состоятельны и асимптотически нормальны при неограниченном возрастании периода наблюдений. Хотя доказательства оказываются здесь более сложными, соответствующие условия и результаты подобны сформулированным в теореме 2.6.1 и следствиях 2.6.1, 2.6.2 для оценок наименьших квадратов. В связи с этим при больших выборках и в настоящем случае можно целиком использовать обычную технику регрессионного анализа.

В § 5.3 было показано, что скалярное стохастическое разностное уравнение порядка p можно рассматривать как частный случай p -компонентного векторного стохастического разностного уравнения первого порядка

$$(1) \quad \mathbf{y}_t + \mathbf{B}\mathbf{y}_{t-1} = \mathbf{u}_t,$$

где

$$(2) \quad \mathbf{y}_t = \begin{pmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{pt} \end{pmatrix}, \quad \mathbf{u}_t = \begin{pmatrix} u_{1t} \\ u_{2t} \\ \vdots \\ u_{pt} \end{pmatrix}.$$

Здесь $\mathbf{B}$ — матрица коэффициентов размера $p \times p$, $\mathbb{E}\mathbf{u}_t = \mathbf{0}$, $\mathbb{E}\mathbf{u}_t\mathbf{u}_t' = \Sigma$, $\mathbf{u}_t$ независимы, а характеристические корни матрицы $-\mathbf{B}$ лежат в единичном круге. Уравнения, из которых определяются оценки максимального правдоподобия для скалярной модели произвольного порядка, являются частным случаем уравнений, используемых для получения таких оценок в векторной модели первого порядка:

$$(3) \quad \frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{y}_{t-1}' \hat{\mathbf{B}}' = -\frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{y}_t',$$

$$(4) \quad \hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T (\mathbf{y}_t + \hat{\mathbf{B}}\mathbf{y}_{t-1})(\mathbf{y}_t + \hat{\mathbf{B}}\mathbf{y}_{t-1})'.$$

Если $\mathbf{u}_t$ распределены нормально, а $\mathbf{y}_0$ считается заданным вектором, то $\hat{\mathbf{B}}$, определяемая из (3), есть оценка максимального правдоподобия. Нам будет удобно сначала получить асимптотические свойства этих оценок, а затем уже перенести полученные результаты на случай скалярного уравнения p -го порядка.

Прежде всего мы покажем, что с точки зрения асимптотической (при $T \rightarrow \infty$) теории, случай, когда (1) выполняется при $t = 1, 2, \dots$, а вектор y_0 фиксирован, ничем не отличается от случая, когда (1) выполняется при всех t ($t = \dots, -1, 0, 1, \dots$). Затем будет показано, что оценка $\hat{B}$ состоятельна и асимптотически нормальна. Далее находятся дисперсии и ковариации предельного распределения и показывается, что их можно состоятельно оценить. Эти результаты будут распространены на модель скользящего среднего, а также на модель стохастического разностного уравнения, в которой возмущение является скользящим средним.

При проведении доказательств предполагается, что все характеристические корни скалярного стохастического разностного уравнения различны и что в векторном случае матрицу $-B$ можно привести к диагональному виду. Однако соответствующие результаты будут иметь силу и без этих предположений. Важно только, чтобы характеристические корни лежали в единичном круге. Указания на необходимые изменения в доказательствах вынесены в подстрочные примечания.

5.5.2. Асимптотическая эквивалентность выборочных моментов второго порядка двух процессов

Рассмотрим сумму $\sum_{t=1}^T y_t y_t'$, в которой y_t , $t = 1, \dots, T$, порождается с помощью (1) при заданном y_0 , так что

$$(5) \quad y_t = \sum_{s=0}^{t-1} (-B)^s u_{t-s} + (-B)^t y_0.$$

Наряду с ней рассмотрим сумму $\sum_{t=1}^T y_t^* y_t^{*'}$, в которой y_t^* удовлетворяет соотношению (1) при всех $t = \dots, -1, 0, 1, \dots$, так что

$$(6) \quad y_t^* = \sum_{s=0}^{\infty} (-B)^s u_{t-s}.$$

Покажем, что матрица $\sum_{t=1}^T y_t^* y_t^{*'} - \sum_{t=1}^T y_t y_t'$, нормированная делением на $\sqrt{T}$ или на T , сходится по вероятности к нулевой матрице.

Из (5) и (6) следует, что $y_t^* - y_t = (-B)^t (y_0^* - y_0)$, где

$$(7) \quad y_0^* = \sum_{s=0}^{\infty} (-B)^s u_{-s}.$$

Поэтому

$$\begin{aligned}
 (8) \quad \sum_{t=1}^T y_t^* y_t^{*'} - \sum_{t=1}^T y_t y_t' &= \\
 &= \sum_{t=1}^T (-B)^t (y_0^* - y_0) y_t' + \sum_{t=1}^T y_t (y_0^* - y_0)' (-B')^t + \\
 &\quad + \sum_{t=1}^T (-B)^t (y_0^* - y_0) (y_0^* - y_0)' (-B')^t.
 \end{aligned}$$

Сумма математических ожиданий квадратов элементов матрицы $\sum_{t=1}^T (-B)^t (y_0^* - y_0) y_t'$ равна

$$\begin{aligned}
 (9) \quad \text{tr} \mathbb{E} \sum_{t=1}^T (-B)^t (y_0^* - y_0) y_t' \left[\sum_{s=1}^T (-B)^s (y_0^* - y_0) y_s' \right]' &= \\
 &= \text{tr} \mathbb{E} \sum_{t,s=1}^T (-B)^t (y_0^* - y_0) y_t' y_s (y_0^* - y_0)' (-B')^s = \\
 &= \sum_{t,s=1}^T \text{tr} [(-B)^t (F + y_0 y_0') (-B')^s] \mathbb{E} y_t' y_s,
 \end{aligned}$$

в силу того, что вектор y_0^* (зависящий от $u_0, u_{-1}, \dots$) не зависит от $u_1, \dots, u_T$ (поскольку последние зависят только от $u_1, \dots, u_T$) и $\mathbb{E} y_0 y_0' = F$.

ЛЕММА 5.5.1. Для любого λ , большего абсолютной величины максимального по абсолютной величине характеристического корня матрицы $-B$, найдется такая константа c , что все элементы матрицы $(-B)^t$ будут меньше $c\lambda^t$, $t = 0, 1, \dots$.

ДОКАЗАТЕЛЬСТВО. Матрицу $(-B)^t$ можно записать в виде

$$(10) \quad (-B)^t = C \Lambda^t C^{-1},$$

где матрица Λ диагональная и ее диагональ состоит из характеристических корней матрицы $-B$, а C — невырожденная матрица. В качестве указанной константы c можно взять ¹⁾ произведение следующих трех чисел: порядка p , максимального абсолютного значения элементов матрицы C и максимального абсолютного значения элементов матрицы C^{-1} . ■

¹⁾ В случае общей жордановой канонической формы при $|\lambda_i| < \lambda$ найдется константа k_i , такая, что $\binom{t}{l} \lambda_i^{t-l} < k_i \lambda^t$. При этом в c будет входить $\max_i k_i$. (См. упр. 17.)

Поскольку

$$(11) \quad \mathbb{E} y_t y_t' = \sum_{r=0}^{t-1} (-B)^r \Sigma (-B')^r + (-B)^t y_0 y_0' (-B')^t,$$

для $\lambda < 1$ и удовлетворяющих условию леммы 5.5.1,

$$(12) \quad \mathbb{E} y_t' y_t = \text{tr } \mathbb{E} y_t y_t' \leq \sum_{r=0}^{t-1} c^2 \lambda^{2r} p^3 \max |\sigma_{ij}| + c^2 \lambda^{2t} p^2 y_0' y_0 = \\ = c^2 p^3 \max |\sigma_{ij}| \frac{1 - \lambda^{2t}}{1 - \lambda^2} + c^2 p^2 y_0' y_0 \lambda^{2t} \leq \\ \leq \frac{c^2 p^3 \max |\sigma_{ij}|}{1 - \lambda^2} + c^2 p^2 y_0' y_0,$$

где c — константа из леммы 5.5.1. Оценивая $|\mathbb{E} y_t' y_s|$ произведением $\sqrt{\mathbb{E} y_t' y_t} \sqrt{\mathbb{E} y_s' y_s}$ и используя для оценки последнего неравенство (12), мы можем оценить абсолютную величину второго из сомножителей, стоящих под знаком суммы в правой части (9).

Что касается первого сомножителя, то его абсолютная величина ограничена значением $K \lambda^{t+s}$, где K равно произведению c^2 , p^3 и максимума абсолютных значений элементов матрицы $F + y_0 y_0'$. Поэтому

$$(13) \quad \sum_{s,t=1}^T \text{tr} [(-B)^t (F + y_0 y_0') (-B')^s] \leq K \lambda^2 \left(\frac{1 - \lambda^T}{1 - \lambda} \right)^2 \leq \frac{K \lambda^2}{(1 - \lambda)^2}.$$

Таким образом, правая часть (9) не превосходит некоторой не зависящей от T постоянной. Отсюда и из неравенства Чебышева следует, что первая и вторая суммы в правой части (8), нормированные делением на $\sqrt{T}$, сходятся по вероятности к нулю.

Третья сумма в правой части (8) является положительно полуопределенной. Ее математическое ожидание равно

$$(14) \quad \mathbb{E} \sum_{t=1}^T (-B)^t (y_0^* - y_0) (y_0^* - y_0)' (-B')^t = \\ = \sum_{t=1}^T (-B)^t (F + y_0 y_0') (-B')^t.$$

След этой матрицы ограничен произведением p^3 , максимального абсолютного значения элементов матрицы $F + y_0 y_0'$ и $\lambda^2 c^2 (1 - \lambda^{2T}) / (1 - \lambda^2)$, так что он равномерно ограничен по T . Применение неравенства Чебышева в форме, используемой для неотрицательных случайных величин, показывает, что сумма диагональных элементов матрицы, представляющей третью сумму в правой части (8), деленная на $\sqrt{T}$, сходится по вероятности к нулю. По-

сколькxу эта матрица положительно полуопределенная, отсюда следует, что вообще все ее элементы сходятся по вероятности к нулю. Таким образом, показано, что разность $\sum_{t=1}^T y_t^* y_t' - \sum_{t=1}^T y_t y_t'$, деленная на $\sqrt{T}$, сходится по вероятности к нулю.

Рассмотрим теперь разность

$$\begin{aligned}
 (15) \quad & \sum_{t=1}^T y_{t-1}^* y_t'' - \sum_{t=1}^T y_{t-1} y_t' = \\
 & = \sum_{t=1}^T y_{t-1}^* [y_{t-1}' (-B') + u_t'] - \\
 & - \sum_{t=1}^T y_{t-1} [y_{t-1}' (-B') + u_t'] = \\
 & = \left[\sum_{t=1}^T y_{t-1}^* y_{t-1}' - \sum_{t=1}^T y_{t-1} y_{t-1}' \right] (-B)' + \\
 & + \sum_{t=1}^T (y_{t-1}^* - y_{t-1}) u_t'.
 \end{aligned}$$

Сумма математических ожиданий квадратов элементов второго слагаемого правой части равна

$$\begin{aligned}
 (16) \quad & \text{tr } \mathbb{E} \sum_{t,s=1}^T (-B)^{t-1} (y_0^* - y_0) u_t' u_s (y_0^* - y_0)' (-B')^{s-1} = \\
 & = (\text{tr } \Sigma) \sum_{t=1}^T \text{tr} (-B)^{t-1} (F + y_0 y_0') (-B')^{t-1}
 \end{aligned}$$

и ограничена произведением чисел $\text{tr } \Sigma$, p^3 , максимального абсолютного значения элементов матрицы $F + y_0 y_0'$ и $c^2(1 - \lambda^{2T})/(1 - \lambda^2)$. Поэтому она равномерно ограничена по T . Отсюда следует, что второе слагаемое правой части (15), деленное на $\sqrt{T}$, сходится по вероятности к нулю.

Поскольку

$$\begin{aligned}
 (17) \quad & \frac{1}{\sqrt{T}} \left[\sum_{t=1}^T y_{t-1}^* y_{t-1}' - \sum_{t=1}^T y_{t-1} y_{t-1}' \right] = \\
 & = \frac{1}{\sqrt{T}} \left[\sum_{t=1}^T y_t^* y_t' - \sum_{t=1}^T y_t y_t' + y_0^* y_0' - y_T^* y_T' - y_0 y_0' + y_T y_T' \right]
 \end{aligned}$$

и последние четыре члена в правой части последнего соотношения сходятся по вероятности к нулю, то левая часть (17) сходится к нулю. Объединяя этот результат с предыдущим, получаем, что разность (15), деленная на $\sqrt{T}$, сходится по вероятности к нулю.

ТЕОРЕМА 5.5.1. Пусть процесс $\{y_t\}$ удовлетворяет соотношению (1) для $t = 1, 2, \dots$, y_0 — фиксированный вектор, а процесс $\{y_t^*\}$ удовлетворяет соотношению (1) для всех $t = \dots, -1, 0, 1, \dots$; кроме того, случайные величины u_t независимы, $\xi u_t = 0$, $\xi u_t u_t' = \Sigma$ и все характеристические корни матрицы $-B$ лежат в единичном круге. Тогда

$$(18) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T y_{t-1}^* y_{t-1}' - \frac{1}{\sqrt{T}} \sum_{t=1}^T y_{t-1} y_{t-1}' \right] = 0,$$

$$(19) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T y_{t-1}^* y_t' - \frac{1}{\sqrt{T}} \sum_{t=1}^T y_{t-1} y_t' \right] = 0,$$

$$(20) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T y_t^* y_t' - \frac{1}{\sqrt{T}} \sum_{t=1}^T y_t y_t' \right] = 0.$$

Смысл этой теоремы заключается в том, что асимптотические свойства матриц $\hat{B}$ и $\hat{\Sigma}$ (состоятельность и асимптотическая нормальность) будут одними и теми же и для процесса $\{y_t\}$, и для процесса $\{y_t^*\}$. Иметь дело с процессом $\{y_t^*\}$ более удобно, поскольку он является стационарным.

5.5.3. Состоятельность оценок максимального правдоподобия

Покажем теперь, что матрица коэффициентов при оценках максимального правдоподобия в (3) сходится по вероятности к некоторой невырожденной матрице, матрица в правой части (3) также сходится по вероятности к некоторой матрице и что результирующее уравнение имеет в качестве единственного решения матрицу параметров B . В этом разделе обозначение y_t будет указывать на то, что процесс удовлетворяет (1) при всех $t = \dots, -1, 0, 1, \dots$, так что

$$(21) \quad y_t = \sum_{s=0}^{\infty} (-B)^s u_{t-s}, \quad t = \dots, -1, 0, 1, \dots,$$

и

$$(22) \quad \xi y_t y_t' = \sum_{s=0}^{\infty} (-B)^s \Sigma (-B')^s = F.$$

ЛЕММА 5.5.2. Пусть случайные векторы $u_1, u_2, \dots$ независимы, $\xi u_t = 0$ и $\xi u_t u_t' = \Sigma$. Тогда если u_t одинаково распределены или $\xi |u_{it}|^{2+\varepsilon} < m$, $i = 1, \dots, p$, $t = 1, 2, \dots$, для некоторых $\varepsilon > 0$ и m , то

$$(23) \quad \text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T u_t u_t' = \Sigma.$$

Доказательство. В случае одинаково распределенных случайных векторов утверждение леммы следует непосредственно из закона больших чисел для одинаково распределенных случайных величин, применяемого к каждой компоненте уравнения (23). При выполнении второго условия элемент матрицы $u_i u'_j$, стоящий на пересечении i -й строки и j -го столбца, именно $u_{it} u_{jt}$, имеет математическое ожидание σ_{ij} и при этом

$$(24) \quad \mathbb{E} |u_{it} u_{jt}|^{1+\varepsilon/2} \leq \sqrt{\mathbb{E} |u_{it}|^{2+\varepsilon} \mathbb{E} |u_{jt}|^{2+\varepsilon}} < m.$$

Утверждение леммы вытекает в этом случае из закона больших чисел, известного под названием теоремы Маркова. [См. Лозв (1963, § 20.1).] ■

Лемма 5.5.3. Пусть y_t определяется соотношением (21), все характеристические корни матрицы $-B$ лежат в единичном круге, случайные величины u_t независимы и $\mathbb{E} u_t = 0$, $\mathbb{E} u_t u'_t = \Sigma$. Тогда

$$(25) \quad \text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T u_t y'_{t-1} = 0.$$

Доказательство. Сумма математических ожиданий квадратов элементов матрицы $(1/T) \sum_{t=1}^T u_t y'_{t-1}$ равна

$$\begin{aligned} (26) \quad \text{tr } \mathbb{E} \frac{1}{T} \sum_{s=1}^T y_{s-1} u'_s \frac{1}{T} \sum_{t=1}^T u_t y'_{t-1} &= \\ &= \frac{1}{T^2} \sum_{s,t=1}^T \mathbb{E} \text{tr } y_{s-1} u'_s u_t y'_{t-1} = \frac{1}{T^2} \sum_{t=1}^T \mathbb{E} u'_t u_t \mathbb{E} y'_{t-1} y_{t-1} = \\ &= \frac{1}{T^2} \sum_{t=1}^T \text{tr } \Sigma \text{tr } F = \frac{1}{T} \text{tr } \Sigma \text{tr } F \end{aligned}$$

в силу того, что $\mathbb{E} y_{s-1} u'_s u_t y'_{t-1} = 0$ при $t \neq s$ (поскольку u_s не зависит от y_{s-1} и $u_t y'_{t-1}$ при $s > t$ и u_t не зависит от остальных сомножителей при $s < t$). Утверждение леммы следует из неравенства Чебышева. ■

В соответствии с леммой 5.5.3

$$(27) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{1}{T} \sum_{t=1}^T y_t y'_{t-1} + B \frac{1}{T} \sum_{t=1}^T y_{t-1} y'_{t-1} \right] = 0,$$

а в соответствии с леммой 5.5.2

$$(28) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{1}{T} \sum_{t=1}^T y_t y'_t + \mathbf{B} \frac{1}{T} \sum_{t=1}^T y_{t-1} y'_t + \frac{1}{T} \sum_{t=1}^T y_t y'_{t-1} \mathbf{B}' + \right. \\ \left. + \mathbf{B} \frac{1}{T} \sum_{t=1}^T y_{t-1} y'_{t-1} \mathbf{B}' \right] = \Sigma.$$

Если вычесть из (28) произведение (27) на $\mathbf{B}'$, а также произведение $\mathbf{B}$ на транспонированное соотношение (27), то получим

$$(29) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{1}{T} \sum_{t=1}^T y_t y'_t - \mathbf{B} \frac{1}{T} \sum_{t=1}^T y_t y'_t \mathbf{B}' \right] = \Sigma.$$

Здесь использован тот факт, что разность

$$(30) \quad \frac{1}{T} \sum_{t=1}^T y_t y'_t - \frac{1}{T} \sum_{t=1}^T y_{t-1} y'_{t-1} = \frac{1}{T} (y_T y'_T - y_0 y'_0)$$

сходится по вероятности к нулю. Поскольку уравнение

$$(31) \quad \mathbf{A} - \mathbf{B} \mathbf{A} \mathbf{B}' = \Sigma$$

имеет единственное решение относительно $\mathbf{A}$ (см. упр. 27), то (29) определяет предел по вероятности матрицы $(1/T) \sum_{t=1}^T y_t y'_t$ и этот предел равен (22).

ТЕОРЕМА 5.5.2. Пусть y_t определяется соотношением (21) для $t = 1, 2, \dots$, причем все характеристические корни матрицы $\mathbf{B}$ лежат в единичном круге, а случайные величины u_t независимы и $\mathbb{E} u_t = 0$, $\mathbb{E} u_t u'_t = \Sigma$. Пусть, кроме того, u_t либо одинаково распределены, либо $\mathbb{E} |u_{it}|^{2+\varepsilon} < m$, $i = 1, \dots, p$, $t = 1, 2, \dots$, для некоторых $\varepsilon > 0$ и m . Тогда

$$(32) \quad \text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T y_t y'_t = \text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T y_{t-1} y'_{t-1} = \mathbf{F},$$

где $\mathbf{F}$ задается соотношением (22).

ЛЕММА 5.5.4. Если матрица Σ положительно определена, то такова же и матрица $\mathbf{F}$.

ДОКАЗАТЕЛЬСТВО. Матрица $\mathbf{F}$, во всяком случае, будет положительно полуопределенной как ковариационная матрица. Поскольку же $\mathbf{F} = \Sigma + \mathbf{B} \mathbf{F} \mathbf{B}'$, она является положительно определенной как сумма положительно определенной и положительно полуопределенной матриц. (См. упр. 28). ■

ТЕОРЕМА 5.5.3. Пусть выполнены условия теоремы 5.5.2 и матрица $\mathbf{F}$ положительно определена. Тогда

$$(33) \quad \text{plim}_{T \rightarrow \infty} \hat{\mathbf{B}} = \mathbf{B},$$

$$(34) \quad \text{plim}_{T \rightarrow \infty} \hat{\Sigma} = \Sigma.$$

ДОКАЗАТЕЛЬСТВО. Для доказательства (33) заметим, что

$$(35) \quad \begin{aligned} \text{plim}_{T \rightarrow \infty} \hat{\mathbf{B}}' - \mathbf{B}' &= \text{plim}_{T \rightarrow \infty} \left(\frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{y}_{t-1}' \right)^{-1} \times \\ &\times \left(-\frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{y}_t' - \frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{y}_{t-1}' \mathbf{B}' \right) = \\ &= -\text{plim}_{T \rightarrow \infty} \left(\frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{y}_{t-1}' \right)^{-1} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{u}_t' = 0. \end{aligned}$$

Тогда (34) вытекает из представления

$$(36) \quad \begin{aligned} \hat{\Sigma} &= \frac{1}{T} \sum_{t=1}^T (\mathbf{y}_t + \hat{\mathbf{B}} \mathbf{y}_{t-1}) (\mathbf{y}_t + \hat{\mathbf{B}} \mathbf{y}_{t-1})' = \\ &= \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t' + \hat{\mathbf{B}} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{y}_t' + \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_{t-1}' \hat{\mathbf{B}}' + \\ &\quad + \hat{\mathbf{B}} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{y}_{t-1}' \hat{\mathbf{B}}', \end{aligned}$$

и из соотношений (28) и (33). ■

Теорема 5.5.3 показывает, что указанные оценки для $\mathbf{B}$ и Σ состоятельны при довольно общих условиях, в частности более общих, чем те, при которых они являются оценками максимального правдоподобия. Мы выбрали эти условия потому, что они относительно просты и тоже приводят к асимптотической нормальности. Предположение о том, что все характеристические корни лежат в единичном круге, не является необходимым для состоятельности оценки матрицы $\mathbf{B}$. Однако оно необходимо для общей теоремы об асимптотической нормальности. (См. Т. Андерсон (1959).)

Чтобы доказать состоятельность оценок для скалярного уравнения порядка p , остается показать невырожденность матрицы $\tilde{\mathbf{F}}$.

ЛЕММА 5.5.5. Если $\sigma^2 > 0$, то матрица

$$(37) \quad \tilde{\mathbf{F}} = \sum_{s=0}^{\infty} \tilde{\mathbf{B}}^s \tilde{\Sigma} \tilde{\mathbf{B}}^{s'},$$

положительно определена, где

$$(38) \quad \tilde{\Sigma} = \begin{bmatrix} \sigma^2 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad \tilde{\mathbf{B}} = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \dots & \beta_p \\ -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

Доказательство. Если для некоторого $\mathbf{x}$ оказывается, что

$$(39) \quad \mathbf{x}'\tilde{\mathbf{F}}\mathbf{x} = \sum_{s=0}^{\infty} \mathbf{x}'\tilde{\mathbf{B}}^s\tilde{\Sigma}\tilde{\mathbf{B}}^{s'}\mathbf{x} = 0,$$

то при этом должны выполняться соотношения

$$(40) \quad \begin{aligned} 0 &= \mathbf{x}'\tilde{\Sigma}\mathbf{x} = x_1^2\sigma^2, \\ 0 &= \mathbf{x}'\tilde{\mathbf{B}}^s\tilde{\Sigma}(\mathbf{x}'\tilde{\mathbf{B}}^s)', \quad s = 1, 2, \dots \end{aligned}$$

Но из последних непосредственно следует, что $x_1 = 0$ и равны нулю первые компоненты векторов $\mathbf{x}'\tilde{\mathbf{B}}^s$, $s = 1, 2, \dots$. Первая же компонента вектора $\mathbf{x}'\tilde{\mathbf{B}}$ равна $-x_2$, так что $x_2 = 0$. Подобным же образом показывается, что равны нулю все компоненты вектора $\mathbf{x}$. Итак, если $\mathbf{x}'\tilde{\mathbf{F}}\mathbf{x} = 0$, то обязательно $\mathbf{x} = 0$, что и требовалось доказать. ■

ТЕОРЕМА 5.5.4. Пусть y_t определяется соотношением (1) из § 5.1 для $t = \dots, -1, 0, 1, \dots$, причем все характеристические корни этого уравнения лежат в единичном круге, а случайные величины u_t независимы и $\mathbb{E}u_t = 0$, $\mathbb{E}u_t^2 = \sigma^2 > 0$. Пусть, кроме того, u_t либо имеют одинаковые распределения, либо $\mathbb{E}|u_t|^{2+\varepsilon} < \infty$, $t = 1, 2, \dots$, для некоторых $\varepsilon > 0$ и m . Тогда

$$(41) \quad \text{plim}_{T \rightarrow \infty} \hat{\beta} = \beta,$$

$$(42) \quad \text{plim}_{T \rightarrow \infty} \hat{\sigma}^2 = \sigma^2.$$

5.5.4. Асимптотическая нормальность оценок

Предельное распределение матрицы

$$(43) \quad \sqrt{T}(\hat{\mathbf{B}}' - \mathbf{B}') = -\sqrt{T} \left(\frac{1}{T} \sum_{t=1}^T y_{t-1} y'_{t-1} \right)^{-1} \frac{1}{T} \sum_{t=1}^T y_{t-1} u'_t$$

совпадает с предельным распределением матрицы $-F^{-1} (1/\sqrt{T}) \times \times \sum_{t=1}^T y_{t-1} u'_t$. Покажем, что предельное распределение произвольной линейной комбинации элементов последней матрицы, например

$$(44) \quad S_T = \frac{1}{\sqrt{T}} \sum_{t=1}^T u'_t \Phi y_{t-1},$$

является нормальным.

Положим

$$(45) \quad y_t^{(k)} = \sum_{s=0}^k (-B)^s u_{t-s},$$

$$(46) \quad Z_{kT} = \frac{1}{\sqrt{T}} \sum_{t=1}^T u'_t \Phi y_{t-1}^{(k)},$$

$$(47) \quad X_{kT} = \frac{1}{\sqrt{T}} \sum_{t=1}^T u'_t \Phi \sum_{s=k+1}^{\infty} (-B)^s u_{t-1-s}.$$

Тогда

$$\begin{aligned} (48) \quad \mathbb{E} X_{kT}^2 &= \frac{1}{T} \sum_{t,\tau=1}^T \sum_{s,\sigma=k+1}^{\infty} \mathbb{E} u'_t \Phi (-B)^s u_{t-1-s} u'_\tau \Phi (-B)^\sigma u_{\tau-1-\sigma} = \\ &= \frac{1}{T} \sum_{t,\tau=1}^T \sum_{s,\sigma=k+1}^{\infty} \mathbb{E} \operatorname{tr} (-B)^s u_{t-1-s} u'_{\tau-1-\sigma} (-B)^\sigma \Phi' u_\tau u'_t \Phi = \\ &= \frac{1}{T} \sum_{t=1}^T \sum_{s=k+1}^{\infty} \operatorname{tr} (-B)^s \Sigma (-B')^s \Phi' \Sigma \Phi = \\ &= \operatorname{tr} \Phi' \Sigma \Phi \sum_{s=k+1}^{\infty} (-B)^s \Sigma (-B')^s = M_k. \end{aligned}$$

(Здесь использовано соотношение $x'x = \operatorname{tr}xx'$.) Поскольку ряд $\sum_{s=0}^{\infty} (-B)^s \Sigma (-B')^s$ сходится, то $M_k \rightarrow 0$ при $k \rightarrow \infty$. Для фиксированного k случайная величина

$$(49) \quad u'_t \Phi y_{t-1}^{(k)} = u'_t \Phi \sum_{s=0}^k (-B)^s u_{t-1-s}$$

имеет нулевое среднее и дисперсию

$$\begin{aligned} (50) \quad \mathbb{E} \sum_{s,r=0}^k u'_t \Phi (-B)^s u_{t-1-s} u'_t \Phi (-B)^r u_{t-1-r} &= \\ &= \sum_{s=0}^k \operatorname{tr} (-B)^s \Sigma (-B')^s \Phi' \Sigma \Phi = \sigma_k^2. \end{aligned}$$

Ковариация двух таких величин, взятых в точках t и τ ($t \neq \tau$), равна

$$(51) \quad \mathbb{E} \sum_{s,r=0}^k \mathbf{u}'_t \Phi(-\mathbf{B})^s \mathbf{u}_{t-1-s} \mathbf{u}'_{\tau} \Phi(-\mathbf{B})^r \mathbf{u}_{\tau-1-r} = 0, \quad \tau \neq t.$$

Кроме того, последовательность $\{\mathbf{u}'_t \Phi_{t-1}^{(k)}\}$ обладает тем свойством, что любое конечное подмножество ее элементов, соответствующих значениям $t = t_1, \dots, t_n$, $t_1 < \dots < t_n$, не зависит от элементов, соответствующих $t < t_1 - k - 1$ и $t > t_n + k + 1$. Использование теоремы 7.7.5 приводит к следующей лемме.

ЛЕММА 5.5.6. Пусть $\mathbf{y}_t^{(k)}$ определены соотношением (45), причем случайные величины $\mathbf{u}_t$ независимы и одинаково распределены с нулевыми средними и ковариационными матрицами $\mathbb{E} \mathbf{u}_t \mathbf{u}'_t = \Sigma$. Тогда распределение случайной величины Z_{kT} , определенной соотношением (46),

$$(52) \quad \Pr \{Z_{kT} \leq z\} = F_{kT}(z)$$

имеет при $T \rightarrow \infty$ предел

$$(53) \quad \lim_{T \rightarrow \infty} F_{kT}(z) = \Phi\left(\frac{z}{\sigma_k}\right),$$

где

$$(54) \quad \Phi(v) = \int_{-\infty}^v \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw,$$

а σ_k^2 указаны в (50).

Далее,

$$(55) \quad \lim_{k \rightarrow \infty} \Phi\left(\frac{z}{\sigma_k}\right) = \Phi\left(\frac{z}{\sigma}\right),$$

где

$$(56) \quad \sigma^2 = \lim_{k \rightarrow \infty} \sigma_k^2 = \text{tr } \Sigma \Phi \Phi'.$$

Здесь мы используем следствие 7.7.1.

ЛЕММА 5.5.7. Пусть выполнены условия леммы 5.5.6 и характеристические корни матрицы $-\mathbf{B}$ лежат в единичном круге. Тогда случайная величина S_T , определенная в (44), имеет в пределе при $T \rightarrow \infty$ нормальное распределение с нулевым средним и дисперсией σ^2 .

Применение теоремы 7.7.7 приводит к искомому результату. Матрицу с элементами $f_{ij} \sigma_{ki}$ обозначим для краткости $\mathbf{F} \otimes \Sigma$. Это есть кронекерово произведение матриц $\mathbf{F}$ и Σ .

ТЕОРЕМА 5.5.5. Если выполнены условия лемм 5.5.6 и 5.5.7, то $(1/\sqrt{T}) \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{u}'_t$ имеет в пределе нормальное распределение с нулевым средним и ковариационной матрицей $\mathbf{F} \otimes \Sigma$.

ТЕОРЕМА 5.5.6. Пусть $\{y_t\}$ — последовательность случайных векторов, удовлетворяющих уравнению (1), в котором u_t — независимые и одинаково распределенные случайные векторы с $E u_t = 0$ и $E u_t u_t' = \Sigma$, и пусть характеристические корни матрицы $-B$ лежат в единичном круге, а матрица F положительно определена. Тогда $\sqrt{T}(\hat{B}' - B')$ имеет в пределе нормальное распределение с нулевым средним и ковариационной матрицей $F^{-1} \otimes \Sigma$.

ДОКАЗАТЕЛЬСТВО. Этот результат вытекает из теоремы 5.5.5 и (43). (См. упр. 29.) ■

ТЕОРЕМА 5.5.7. Пусть $\{y_t\}$ — последовательность случайных величин, удовлетворяющая уравнению (1) из § 5.1 для $t = \dots, -1, 0, 1, \dots$, и пусть корни характеристического уравнения, соответствующего (1), лежат в единичном круге, а случайные величины u_t независимы и одинаково распределены с нулевыми средними и дисперсиями σ^2 . Тогда $\sqrt{T}(\hat{\beta} - \beta)$ имеет в пределе нормальное распределение с нулевым средним и ковариационной матрицей $\sigma^2 \tilde{F}^{-1}$.

В формулировках теорем 5.5.6 и 5.5.7 условие одинаковой распределенности случайных величин можно заменить условием равномерной ограниченности их моментов порядка $2 + \varepsilon$ или использовать вместо него условие типа Линдеберга.

5.5.5. Случай неизвестного среднего

Предположим теперь, что векторное разностное уравнение имеет вид

$$(57) \quad y_t + B y_{t-1} + v = u_t$$

или, что эквивалентно,

$$(58) \quad (y_t - \mu) + B(y_{t-1} - \mu) = u_t,$$

где

$$(59) \quad v = -\mu - B\mu.$$

Тогда уравнениями максимального правдоподобия для B , v и Σ будут

$$(60) \quad \begin{bmatrix} \sum_{t=1}^T y_{t-1} y_{t-1}' & \sum_{t=1}^T y_{t-1}' \\ \sum_{t=1}^T y_{t-1}' & T \end{bmatrix} \begin{bmatrix} \hat{B}' \\ \hat{v}' \end{bmatrix} = - \begin{bmatrix} \sum_{t=1}^T y_{t-1} y_t' \\ \sum_{t=1}^T y_t' \end{bmatrix},$$

$$(61) \quad \hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T (y_t + \hat{B} y_{t-1} + \hat{v})(y_t + \hat{B} y_{t-1} + \hat{v})'.$$

Уравнение (60) приводит к следующей паре уравнений:

$$(62) \quad \left(\frac{1}{T} \sum_{t=1}^T y_{t-1} y'_{t-1} - \bar{y}_{(1)} \bar{y}'_{(1)} \right) \hat{\mathbf{B}}' = - \left(\frac{1}{T} \sum_{t=1}^T y_{t-1} y'_t - \bar{y}_{(1)} \bar{y}' \right),$$

$$(63) \quad \hat{\mathbf{v}} = -\bar{y} - \hat{\mathbf{B}} \bar{y}_{(1)},$$

где $\bar{y} = (1/T) \sum_{t=1}^T y_t$ и $\bar{y}_{(1)} = (1/T) \sum_{t=1}^T y_{t-1}$. Мы хотим показать, что

асимптотическое поведение оценок $\hat{\mathbf{B}}$, являющихся решениями уравнений (62) и (3) соответственно, одинаково при $\mu = 0$. Для этого достаточно в свою очередь показать, что при $\mu = 0$ величины $\sqrt{T} \bar{y}_{(1)} \bar{y}'_{(1)}$ и $\sqrt{T} \bar{y}_{(1)} \bar{y}'$ сходятся по вероятности к нулю.

Рассмотрим сначала случай, когда (57) выполняется для $t = \dots, -1, 0, 1, \dots$. При этом

$$(64) \quad y_t^* = \mu + \sum_{s=0}^{\infty} (-\mathbf{B})^s u_{t-s},$$

как что $\mathcal{E} y_t^* = \mu$. Ковариационная матрица векторов y_t^* и $y_{t'}^*$ для $t \leq t'$ есть

$$(65) \quad \mathcal{E} (y_t^* - \mu) (y_{t'}^* - \mu)' = \sum_{s, s'=0}^{\infty} (-\mathbf{B})^s \mathcal{E} u_{t-s} u_{t'-s'}' (-\mathbf{B}')^{s'} = \\ = \sum_{s=0}^{\infty} (-\mathbf{B})^s \Sigma (-\mathbf{B}')^{s+t'-t} = \mathbf{F} (-\mathbf{B}')^{t'-t}.$$

Отсюда следует, что

$$(66) \quad T \mathcal{E} (\bar{y}^* - \mu) (\bar{y}^* - \mu)' = \frac{1}{T} \mathcal{E} \sum_{t, t'=1}^T (y_t^* - \mu) (y_{t'}^* - \mu)' = \\ = \mathbf{F} + \frac{1}{T} \sum_{s=1}^{T-1} (T-s) \mathbf{F} (-\mathbf{B}')^s + \frac{1}{T} \sum_{s=1}^{T-1} (T-s) (-\mathbf{B})^s \mathbf{F}$$

и сходится при $T \rightarrow \infty$ к

$$(67) \quad \lim_{T \rightarrow \infty} T \mathcal{E} (\bar{y}^* - \mu) (\bar{y}^* - \mu)' = \mathbf{F} + \mathbf{F} \sum_{s=1}^{\infty} (-\mathbf{B}')^s + \sum_{s=1}^{\infty} (-\mathbf{B})^s \mathbf{F} = \\ = \mathbf{F} + \mathbf{F} (-\mathbf{B}') (\mathbf{I} + \mathbf{B}')^{-1} + \\ + (\mathbf{I} + \mathbf{B})^{-1} (-\mathbf{B}) \mathbf{F} = \\ = \mathbf{F} (\mathbf{I} + \mathbf{B}')^{-1} + (\mathbf{I} + \mathbf{B})^{-1} \mathbf{F} - \mathbf{F}.$$

В силу неравенства Чебышева $\sqrt{T} (\bar{y}^* - \mu) (\bar{y}^* - \mu)'$ сходится по вероятности к нулевой матрице.

В случае когда (57) выполняется только для $t = 1, 2, \dots$ и y_0 — фиксированный вектор,

$$(68) \quad y_t = \mu + (-B)^t (y_0 - \mu) + \sum_{s=0}^{t-1} (-B)^s u_{t-s}.$$

Разность между y_t^* , определяемым (64), и y_t , определяемым (68), равна

$$(69) \quad (-B)^t [y_0^* - (y_0 - \mu)],$$

где y_0^* задается соотношением (7), так что соответствующая разность между $\sqrt{T} \bar{y}^*$ и $\sqrt{T} \bar{y}$ равна

$$(70) \quad \frac{1}{\sqrt{T}} \sum_{t=1}^T (-B)^t [y_0^* - (y_0 - \mu)].$$

Сумма же математических ожиданий квадратов элементов матрицы (70) сходится к нулю.

Из приведенных результатов следует, что $\sqrt{T} \bar{y}_{(1)} \bar{y}_{(1)}'$ и $\sqrt{T} \bar{y}_{(1)} \bar{y}'$ сходятся при $\mu = 0$ по вероятности к нулю и во втором случае. Оценки $\hat{B}$ для обеих моделей, определяемые (62) и (3), имеют одинаковые асимптотические свойства.

ТЕОРЕМА 5.5.8. Пусть y_t определяется соотношениями (64) или (68), в которых u_t независимы и одинаково распределены с $\xi u_t = 0$, $\xi u_t u_t' = \Sigma$, и все характеристические корни матрицы $-B$ лежат в единичном круге. Тогда предельное распределение случайной величины $\sqrt{T} (\bar{y} - \mu)$ нормальное, имеет нулевое среднее и ковариационную матрицу (67).

Доказательство. Пусть $S_T = \sqrt{T} \Phi' (\bar{y} - \mu)$, $y_t^{(k)}$ равно правой части (45) плюс μ , $Z_{kT} = (1/\sqrt{T}) \sum_{t=1}^T \Phi' (y_t^{(k)} - \mu)$, $X_{kT} = S_T - Z_{kT}$. Тогда, если y_t определяется соотношением (64), то

$$(71) \quad \begin{aligned} \xi X_{kT}^2 = \Phi' (-B)^{k+1} & \left[F + \frac{1}{T} \sum_{s=1}^{T-1} (T-s) F (-B')^s + \right. \\ & \left. + \frac{1}{T} \sum_{s=1}^{T-1} (T-s) (-B)^s F \right] (-B')^{k+1} \Phi. \end{aligned}$$

Правая часть (71) сходится к нулю, когда $k \rightarrow \infty$. Используя теорему 7.7.5, находим, что Z_{kT} имеет в пределе при $T \rightarrow \infty$ нормальное распределение с нулевым средним и дисперсией, сходящейся при $k \rightarrow \infty$ к некоторой постоянной. Применение следствия 7.7.1 приводит к заключению о том, что S_T имеет в пределе нормальное

распределение. Поскольку вектор Φ был произвольным, то последнее означает нормальность в пределе случайного вектора $\sqrt{T}(\bar{y} - \mu)$. Если y_t определяется соотношением (68), то надо использовать (70) и еще раз применить следствие 7.7.1. ■

Скалярное стохастическое разностное уравнение порядка p можно записать в векторной форме, беря в качестве v вектор $\tilde{v}$, у которого отлична от нуля и равна v только первая компонента.

Действительно, уравнения (60) для модели (57), в которой B заменяется на $\tilde{B}$, y_t на $\tilde{y}_t$ и u_t на $\tilde{u}_t$, показывают, что при этом необходимо должны равняться нулю последние $p-1$ компонент $\tilde{v}$. Тогда из (59) будет следовать, что все компоненты вектора $\tilde{\mu}$ равны $\mu = v/(1 + \sum_{i=1}^p \beta_i)$. Применяя полученные результаты для

этого случая, находим, что $\sqrt{T}(\tilde{y} - \tilde{\mu})$ имеет в пределе распределение, совпадающее с предельным распределением $\sqrt{T}(\bar{y} - \mu) \varepsilon$, где $\varepsilon = (1, \dots, 1)'$.

5.5.6. Случай фиксированных переменных

Оценки максимального правдоподобия были получены у нас для модели

$$(72) \quad \sum_{s=0}^p \beta_s y_{t-s} + \sum_{j=1}^q \gamma_j z_{jt} = u_t.$$

Она может рассматриваться как частный случай p -компонентной векторной модели

$$(73) \quad y_t + B y_{t-1} + \Gamma z_t = u_t$$

в которой Γ — матрица размера $p \times q$, состоящая из констант, и $z_t = (z_{1t}, \dots, z_{qt})'$. Оценки максимального правдоподобия для B , Γ и Σ определяются из соотношений

$$(74) \quad \begin{bmatrix} \frac{1}{T} \sum_{t=1}^T y_{t-1} y'_{t-1} & \frac{1}{T} \sum_{t=1}^T y_{t-1} z'_t \\ \frac{1}{T} \sum_{t=1}^T z_t y_{t-1} & \frac{1}{T} \sum_{t=1}^T z_t z'_t \end{bmatrix} \begin{bmatrix} \hat{B}' \\ \hat{\Gamma}' \end{bmatrix} = - \begin{bmatrix} \frac{1}{T} \sum_{t=1}^T y_{t-1} y'_t \\ \frac{1}{T} \sum_{t=1}^T z_t y'_t \end{bmatrix},$$

$$(75) \quad \hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T (y_t + \hat{B} y_{t-1} + \hat{\Gamma} z_t)(y_t + \hat{B} y_{t-1} + \hat{\Gamma} z_t)'$$

Асимптотическая теория, развитая для $\hat{\mathbf{B}}$ и $\hat{\Sigma}$, определяемых уравнениями (3) и (4), может быть распространена и на $\hat{\mathbf{B}}, \hat{\Gamma}, \hat{\Sigma}$, определяемые из (74) и (75). Однако подробно изложить такую теорию довольно сложно. Поэтому мы наметим ее лишь в общих чертах. (См. Т. Андерсон и Рубин (1950), а также Купменс, Рубин и Лейпник (1950).)

Если (73) выполняется для $t = 1, 2, \dots$ и $\mathbf{y}_0$ — фиксированный вектор, то

$$(76) \quad \mathbf{y}_t = \sum_{s=0}^{t-1} (-\mathbf{B})^s \mathbf{u}_{t-s} + \mathbf{w}_t + (-\mathbf{B})^t \mathbf{y}_0,$$

где

$$(77) \quad \mathbf{w}_t = - \sum_{s=0}^{t-1} (-\mathbf{B})^s \mathbf{\Gamma} \mathbf{z}_{t-s}.$$

Если (73) выполняется для всех $t = \dots, -1, 0, 1, \dots$ и $\mathbf{z}_t = \mathbf{0}$ для $t = 0, -1, \dots$, то решение (73) будет иметь вид

$$(78) \quad \mathbf{y}_t^* = \sum_{s=0}^{\infty} (-\mathbf{B})^s \mathbf{u}_{t-s} + \mathbf{w}_t.$$

Разность $\mathbf{y}_t^* - \mathbf{y}_t$ равна при этом $(-\mathbf{B})^t (\mathbf{y}_0^* - \mathbf{y}_0)$, где $\mathbf{y}_0^*$ определено в (7).

ТЕОРЕМА 5.5.9. Пусть $\mathbf{y}_t$ определяется соотношением (73) для $t = 1, 2, \dots$ при фиксированном векторе $\mathbf{y}_0$, а $\mathbf{y}_t^*$ определяется тем же соотношением (73) для всех $t = \dots, -1, 0, 1, \dots$. Пусть при этом случайные величины $\mathbf{u}_t$ независимы, $\mathbf{E} \mathbf{u}_t = \mathbf{0}$, $\mathbf{E} \mathbf{u}_t \mathbf{u}_t' = \Sigma$, а характеристические корни матрицы $-\mathbf{B}$ лежат в единичном круге. Тогда, если $\mathbf{z}_t = \mathbf{0}$ для $t = 0, -1, \dots$ и $\mathbf{z}_t' \mathbf{z}_t < N$, $t = 1, 2, \dots$ для некоторого N , то справедливы соотношения (18), (19), (20) и

$$(79) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{z}_t' - \frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{z}_t' \right] = \mathbf{0},$$

$$(80) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbf{z}_t \mathbf{y}_t^{*'} - \frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbf{z}_t \mathbf{y}_t' \right] = \mathbf{0}.$$

ДОКАЗАТЕЛЬСТВО. Мы имеем

$$(81) \quad \begin{aligned} \sum_{t=1}^T \mathbf{y}_t^* \mathbf{y}_t^{*'} - \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t' &= \\ &= \sum_{t=1}^T (-\mathbf{B})^t (\mathbf{y}_0^* - \mathbf{y}_0) \left[\sum_{s=0}^{t-1} (-\mathbf{B})^s \mathbf{u}_{t-s} + (-\mathbf{B})^t \mathbf{y}_0 \right]' + \end{aligned}$$

$$\begin{aligned}
& + \sum_{t=1}^T (-B)^t (y_0^* - y_0) w_t' + \\
& + \sum_{t=1}^T \left[\sum_{s=0}^{t-1} (-B)^s u_{t-s} + (-B)^t y_0 \right] (y_0^* - y_0)' (-B')^t + \\
& + \sum_{t=1}^T w_t (y_0^* - y_0)' (-B')^t + \\
& + \sum_{t=1}^T (-B)^t (y_0^* - y_0) (y_0^* - y_0)' (-B')^t.
\end{aligned}$$

Первый, третий и пятый члены в правой части, деленные на $\sqrt{T}$, сходятся по вероятности к нулю в силу доказательства теоремы 5.5.1. Сумма математических ожиданий квадратов компонент второго члена равна

$$\begin{aligned}
(82) \quad \text{tr} \mathbb{E} \sum_{t, t'=1}^T (-B)^t (y_0^* - y_0) w_t' w_{t'}' (y_0^* - y_0)' (-B')^{t'} &= \\
&= \sum_{t, t'=1}^T w_t' w_{t'}' \text{tr} (-B)^t (F + y_0 y_0') (-B')^{t'}.
\end{aligned}$$

Поскольку $|w_t' w_{t'}'| \leq \sqrt{w_t' w_t} \sqrt{w_{t'}' w_{t'}}$,

$$(83) \quad w_t' w_t = \sum_{s, s'=0}^{t-1} z_{t-s}' \Gamma' (-B')^s (-B)^{s'} \Gamma z_{t-s}$$

и каждый элемент вектора z_t меньше $\sqrt{N}$, то сумма (82) равномерно ограничена (по лемме 5.5.1). Из этого следует, что разность (81), деленная на $\sqrt{T}$, сходится по вероятности к нулю.

Рассмотрим теперь

$$(84) \quad \sum_{t=1}^T y_{t-1}^* z_t' - \sum_{t=1}^T y_{t-1} z_t' = \sum_{t=1}^T (-B)^{t-1} (y_0^* - y_0) z_t'.$$

Сумма математических ожиданий квадратов компонент (84) ограничена по тем же причинам, что и выше (надо только w_t заменить на z_t). Это доказывает (79).

Далее,

$$\begin{aligned}
(85) \quad \sum_{t=1}^T y_{t-1}^* y_t' - \sum_{t=1}^T y_{t-1} y_t' &= - \left[\sum_{t=1}^T y_{t-1}^* y_{t-1}' - \sum_{t=1}^T y_{t-1} y_{t-1}' \right] B' - \\
&- \left[\sum_{t=1}^T y_{t-1}^* z_t' - \sum_{t=1}^T y_{t-1} z_t' \right] \Gamma' + \sum_{t=1}^T (y_{t-1}^* - y_{t-1}) u_t'.
\end{aligned}$$

Сумма математических ожиданий квадратов компонент последней матрицы также равномерно ограничена. (См. доказательство теоремы 5.5.1.)

Наконец,

$$(86) \quad \sum_{t=1}^T y_t^* z_t' - \sum_{t=1}^T y_t z_t' = -B \left[\sum_{t=1}^T y_{t-1}^* z_t' - \sum_{t=1}^T y_{t-1} z_t' \right].$$

Таким образом, теорема доказана. ■

Как и в разд. 5.5.2, асимптотические свойства оценок $\hat{B}$, $\hat{\Gamma}$ и $\hat{\Sigma}$ (при соответствующих предположениях) одинаковы и для y_t , определяемого (73) при $t = 1, 2, \dots$ с фиксированным вектором y_0 , и для y_t^* , определяемого (73) при всех $t = \dots, -1, 0, 1, \dots$. В связи с этим мы в дальнейшем будем рассматривать только последний случай, используя обозначение y_t для (78).

Лемма 5.5.8. Пусть y_t определяется соотношением (78), в котором u_t независимы, $\xi u_t = 0$, $\xi u_t u_t' = \Sigma$, а матрица $-B$ такова, что все ее характеристические корни лежат в единичном круге. Тогда, если $z_t z_t' < N$, $t = 1, 2, \dots$, для некоторого N , то

$$(87) \quad \text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T u_t y_{t-1}' = 0,$$

$$(88) \quad \text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T u_t z_t' = 0.$$

Доказательство. Сумма математических ожиданий квадратов компонент допредельной матрицы в (87) равна

$$(89) \quad \frac{1}{T^2} \text{tr} \sum_{t,t'=1}^T \xi y_{t-1} u_t' u_{t'} y_{t'-1}' = \frac{1}{T^2} \text{tr} \Sigma \text{tr} \sum_{t=1}^T \xi y_{t-1} y_{t-1}' = \\ = \frac{1}{T} \text{tr} \Sigma \left(\text{tr} F + \frac{1}{T} \sum_{t=1}^T w_{t-1}' w_{t-1} \right)$$

и сходится к нулю при $T \rightarrow \infty$. Сумма математических ожиданий квадратов компонент допредельной матрицы в (88) равна

$$(90) \quad \frac{1}{T^2} \text{tr} \sum_{t,t'=1}^T \xi u_t z_t' z_{t'} u_{t'}' = \frac{1}{T^2} \text{tr} \Sigma \sum_{t=1}^T z_t' z_t < \frac{1}{T} \text{tr} \Sigma N$$

и также сходится к нулю. ■

Из леммы 5.58 вытекает, что

$$(91) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{1}{T} \sum_{t=1}^T y_t y'_{t-1} + B \frac{1}{T} \sum_{t=1}^T y_{t-1} y'_{t-1} + \Gamma \frac{1}{T} \sum_{t=1}^T z_t y'_{t-1} \right] = 0,$$

$$(92) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{1}{T} \sum_{t=1}^T y_t z'_t + B \frac{1}{T} \sum_{t=1}^T y_{t-1} z'_t + \Gamma \frac{1}{T} \sum_{t=1}^T z_t z'_t \right] = 0.$$

Последние соотношения в совокупности с леммой 5.5.2 дают

$$(93) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{1}{T} \sum_{t=1}^T y_t y'_t + B \frac{1}{T} \sum_{t=1}^T y_{t-1} y'_t + \Gamma \frac{1}{T} \sum_{t=1}^T z_t y'_t \right] = \Sigma.$$

Предположим теперь, что существуют пределы

$$(94) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T z_t z'_t = M,$$

$$(95) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T w_{t-1} z'_t = L.$$

Тогда

$$(96) \quad \mathcal{E} \frac{1}{T} \sum_{t=1}^T y_{t-1} z'_t = \frac{1}{T} \sum_{t=1}^T w_{t-1} z'_t \rightarrow L$$

и

$$(97) \quad \begin{aligned} \text{tr } \mathcal{E} \left[\frac{1}{T} \sum_{t=1}^T (y_{t-1} - w_{t-1}) z'_t \right] \left[\frac{1}{T} \sum_{t'=1}^T (y_{t'-1} - w_{t'-1}) z'_{t'} \right]' &= \\ &= \frac{1}{T^2} \sum_{t, t'=1}^T z'_t z'_{t'} \text{tr } \mathcal{E} (y_{t-1} - w_{t-1}) (y_{t'-1} - w_{t'-1})' < \\ &< N \frac{1}{T} \text{tr} [F(I + B')^{-1} + (I + B)^{-1} F - F]. \end{aligned}$$

Правая часть последнего неравенства сходится к нулю. Таким образом,

$$(98) \quad \text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T y_{t-1} z'_t = L.$$

Из (92) и (98) выводим, что

$$(99) \quad \text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T y_t z'_t = -BL - \Gamma M.$$

Из (91), (93) и того факта, что (30) сходится по вероятности к нулевой матрице, получаем

$$(100) \quad \Sigma = \text{plim}_{T \rightarrow \infty} \left(\frac{1}{T} \sum_{t=1}^T y_t y'_t - B \frac{1}{T} \sum_{t=1}^T y_{t-1} y'_{t-1} B' - \right.$$

$$\begin{aligned}
& - \mathbf{B} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{z}_t' \Gamma' + \Gamma \frac{1}{T} \sum_{t=1}^T \mathbf{z}_t \mathbf{y}_t' \Big) = \\
& = \text{plim}_{T \rightarrow \infty} \left(\frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t' - \mathbf{B} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t' \mathbf{B}' \right) - \mathbf{B} \mathbf{L} \Gamma' - \Gamma \mathbf{L}' \mathbf{B}' - \\
& - \Gamma \mathbf{M} \Gamma'.
\end{aligned}$$

Это показывает, что $(1/T) \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t'$ имеет предел по вероятности и что этот последний должен быть равен

$$(101) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t' = \mathbf{F} + \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbf{w}_t \mathbf{w}_t' = \mathbf{F} + \mathbf{H}.$$

Здесь мы предположили существование предела

$$(102) \quad \mathbf{H} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbf{w}_t \mathbf{w}_t'.$$

ТЕОРЕМА 5.5.10. Пусть характеристические корни матрицы $-\mathbf{B}$ лежат в единичном круге, пределы (94), (95) и (102) существуют и $\mathbf{z}_t \mathbf{z}_t' < N$ при $t = 1, 2, \dots$ для некоторого N . Пусть при этом либо $\mathbf{u}_t$ одинаково распределены, либо $\mathbb{E} |\mathbf{u}_{it}|^{2+\varepsilon} < m$, $i = 1, \dots, p$, $t = 1, 2, \dots$, для некоторого $\varepsilon > 0$. Тогда, если $\mathbf{u}_t$ независимы, $\mathbb{E} \mathbf{u}_t = 0$ и $\mathbb{E} \mathbf{u}_t \mathbf{u}_t' = \Sigma$, то

$$(103) \quad \text{plim}_{T \rightarrow \infty} \begin{bmatrix} \frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{y}_{t-1}' & \frac{1}{T} \sum_{t=1}^T \mathbf{y}_{t-1} \mathbf{z}_t' \\ \frac{1}{T} \sum_{t=1}^T \mathbf{z}_t \mathbf{y}_{t-1}' & \frac{1}{T} \sum_{t=1}^T \mathbf{z}_t \mathbf{z}_t' \end{bmatrix} = \begin{bmatrix} \mathbf{F} + \mathbf{H} & \mathbf{L} \\ \mathbf{L}' & \mathbf{M} \end{bmatrix}.$$

Для состоятельности оценок $\hat{\mathbf{B}}$ и $\hat{\Gamma}$ при этих условиях нужно еще, чтобы матрица (103) была невырожденной.

ЛЕММА 5.5.9. Если матрицы $\mathbf{F}$ и $\mathbf{M}$ положительно определены, то и матрица (103) положительно определена.

ДОКАЗАТЕЛЬСТВО. Поскольку

$$\begin{aligned}
(104) \quad (\mathbf{x}' \quad \mathbf{y}') \begin{pmatrix} \mathbf{H} & \mathbf{L} \\ \mathbf{L}' & \mathbf{M} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T (\mathbf{x}' \quad \mathbf{y}') \begin{pmatrix} \mathbf{w}_{t-1} \\ \mathbf{z}_t \end{pmatrix} (\mathbf{w}_{t-1}' \quad \mathbf{z}_t') \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \\
&= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T [\mathbf{x}' \mathbf{w}_{t-1} + \mathbf{y}' \mathbf{z}_t]^2 \geq 0,
\end{aligned}$$

матрица этой квадратичной формы положительно полуопределена. Квадратичная форма

$$(105) \quad (\mathbf{x}'\mathbf{y}') \begin{pmatrix} \mathbf{F} + \mathbf{H} & \mathbf{L} \\ \mathbf{L}' & \mathbf{M} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{x}'\mathbf{F}\mathbf{x} + (\mathbf{x}'\mathbf{y}') \begin{pmatrix} \mathbf{H} & \mathbf{L} \\ \mathbf{L}' & \mathbf{M} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}$$

является суммой двух неотрицательных квадратичных форм и поэтому может равняться нулю, только если обе они обращаются в нуль. Поскольку $\mathbf{F}$ положительно определена, то из $\mathbf{x}'\mathbf{F}\mathbf{x} = 0$ следует, что $\mathbf{x} = 0$. Поэтому, для того чтобы (105) равнялось нулю, необходимо, чтобы $\mathbf{y}'\mathbf{M}\mathbf{y} = 0$. Последнее выполняется лишь при $\mathbf{y} = 0$. ■

ТЕОРЕМА 5.5.11. Если в условиях теоремы 5.5.10 матрица $\mathbf{F}$ положительно определена, а матрица $\mathbf{M}$ из соотношения (94) не вырождена, то

$$(106) \quad \text{plim}_{T \rightarrow \infty} \hat{\mathbf{B}} = \mathbf{B},$$

$$(107) \quad \text{plim}_{T \rightarrow \infty} \hat{\mathbf{\Gamma}} = \mathbf{\Gamma},$$

$$(108) \quad \text{plim}_{T \rightarrow \infty} \hat{\mathbf{\Sigma}} = \mathbf{\Sigma}.$$

Обратимся теперь к асимптотической нормальности $(1/\sqrt{T}) \times \sum_{t=1}^T \mathbf{y}_{t-1}\mathbf{u}'_t$ и $(1/\sqrt{T}) \sum_{t=1}^T \mathbf{z}_t\mathbf{u}'_t$. Средние этих случайных величин равны нулю. Выкладки, подобные проведенным в (89) и (90), показывают, что

$$(109) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \otimes \sum_{t=1}^T \mathbf{y}_{t-1}\mathbf{u}'_t \otimes \sum_{t=1}^T \mathbf{u}_t\mathbf{y}'_{t-1} = (\mathbf{F} + \mathbf{H}) \otimes \mathbf{\Sigma},$$

$$(110) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \otimes \sum_{t=1}^T \mathbf{y}_{t-1}\mathbf{u}'_t \otimes \sum_{t=1}^T \mathbf{u}_t\mathbf{z}'_t = \mathbf{L} \otimes \mathbf{\Sigma},$$

$$(111) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \otimes \sum_{t=1}^T \mathbf{z}_t\mathbf{u}'_t \otimes \sum_{t=1}^T \mathbf{u}_t\mathbf{z}'_t = \mathbf{M} \otimes \mathbf{\Sigma}.$$

Положим

$$(112) \quad S_T = \frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbf{u}'_t (\Phi \mathbf{y}_{t-1} + \Theta \mathbf{z}_t),$$

$$(113) \quad Z_{kT} = \frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbf{u}'_t (\Phi \mathbf{y}_{t-1}^{(k)} + \Theta \mathbf{z}_t),$$

где $\mathbf{y}_t^{(k)}$ и X_{kT} те же, что и раньше. В данном случае слагаемые в Z_{kT} не будут одинаково распределенными, даже если $\mathbf{u}_t$ одинаково

распределены. Если моменты четвертого порядка u_t равномерно ограничены по t , легко показать, что будут равномерно ограниченными и четвертые моменты слагаемых в $\sqrt{T}Z_{kt}$. Сложнее показать (с использованием неравенства Гёльдера), что если $\mathbb{E}|u_{it}|^{2+\varepsilon}$ равномерно ограничены, то будут равномерно ограниченными и моменты порядка $2 + \varepsilon$ этих слагаемых.

ТЕОРЕМА 5.5.12. Пусть y_t определяется соотношениями (76) или (78), случайные векторы u_t независимы, $\mathbb{E}u_t = 0$, $\mathbb{E}u_t u_t' = \Sigma$ и все характеристические корни матрицы $-B$ лежат в единичном круге. Пусть, кроме того, $z_t z_t' < N$, $t = 1, 2, \dots$, для некоторого N и существуют пределы (94), (95) и (102). Тогда, если $\mathbb{E}u_{it}^4 < N^*$, $i = 1, \dots, p$, $t = 1, 2, \dots$, для некоторого N^* , то

$$(114) \quad \begin{bmatrix} \frac{1}{\sqrt{T}} \sum_{t=1}^T y_{t-1} u_t' \\ \frac{1}{\sqrt{T}} \sum_{t=1}^T z_t u_t \end{bmatrix}$$

имеет в пределе при $T \rightarrow \infty$ нормальное распределение с нулевым средним и ковариационной матрицей

$$(115) \quad \begin{pmatrix} F + H & L \\ L' & M \end{pmatrix} \otimes \Sigma.$$

ТЕОРЕМА 5.5.13. Если выполнены условия теоремы 5.5.12 и если матрицы $F = \sum_{s=0}^{\infty} B^s \Sigma B^{s'}$ и $M = \lim_{T \rightarrow \infty} (1/T) \sum_{t=1}^T z_t z_t'$ положительно определены, то

$$(116) \quad \sqrt{T}(\hat{B}' - B'), \quad \sqrt{T}(\hat{\Gamma}' - \Gamma')$$

имеет в пределе нормальное распределение с нулевым средним и ковариационной матрицей

$$(117) \quad \begin{pmatrix} F + H & L \\ L' & M \end{pmatrix}^{-1} \otimes \Sigma.$$

Невырожденность матриц F и M предполагается, так что обратная матрица в (117) существует. Если Σ вырождена, то вырожденным будет и предельное нормальное распределение. Если z_t есть скалярная постоянная, то утверждение теоремы будет вытекать из одинаковой распределенности u_t с ковариационной матрицей Σ , без привлечения условий на моменты более высокого порядка. Дело в том, что при этом $u_t' (\Phi y_{t-1}^{(k)} + \Theta z_t)$ образуют стационарный процесс с конечной зависимостью. В общем случае условия

на $\mathbf{z}_t$ и $\mathbf{u}_t$ можно ослабить в том смысле, что можно требовать только, чтобы $\mathbf{u}_t' (\Phi \mathbf{y}_{t-1}^{(k)} + \Theta \mathbf{z}_t)$ удовлетворяли условию типа Линдеберга. Условие $\mathbf{z}_t' \mathbf{z}_t < N$ можно также ослабить, заменив его условием $\sum_{t=1}^{\infty} \lambda^t |\mathbf{z}_{jt}| < \infty, j = 1, \dots, q$.

ТЕОРЕМА 5.5.14. Пусть y_t определяется соотношением (1) из § 5.4, в котором $\mathbf{u}_t$ независимо распределены, $\mathbb{E} \mathbf{u}_t = 0$, $\mathbb{E} \mathbf{u}_t^2 = \sigma^2 > 0$, причем все характеристические корни матрицы $-\mathbf{B}$ лежат в единичном круге. Пусть, кроме того, $\mathbf{z}_t' \mathbf{z}_t < N$, $t = 1, 2, \dots$, для некоторого N , а $\mathbb{E} \mathbf{u}_t^4 < m$, $t = 1, 2, \dots$, для некоторого m , матрица $\mathbf{M}$ из (94) положительно определена и существуют пределы (95) и (102), где

$$(118) \quad \tilde{\mathbf{w}}_t = - \sum_{s=0}^{t-1} (-\tilde{\mathbf{B}})^s \begin{pmatrix} \gamma' \mathbf{z}_{t-s} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Тогда $\sqrt{T}(\hat{\beta} - \beta)$, $\sqrt{T}(\hat{\gamma} - \gamma)$ имеет в пределе при $T \rightarrow \infty$ нормальное распределение с нулевым средним и ковариационной матрицей

$$(119) \quad \sigma^2 \begin{pmatrix} \tilde{\mathbf{F}} + \tilde{\mathbf{H}} & \tilde{\mathbf{L}} \\ \tilde{\mathbf{L}}' & \mathbf{M} \end{pmatrix}^{-1}.$$

5.6. СТАТИСТИЧЕСКИЕ ВЫВОДЫ О МОДЕЛЯХ АВТОРЕГРЕССИИ, ОСНОВАННЫЕ НА ТЕОРИИ БОЛЬШИХ ВЫБОРОК

5.6.1. Связь с регрессионными методами

В § 5.4 было показано, что оценки максимального правдоподобия векторов $\beta = (\beta_1, \dots, \beta_p)'$, $\gamma = (\gamma_1, \dots, \gamma_q)'$ и дисперсии σ^2 получаются при минимизации выражения

$$(1) \quad \sum_{t=1}^T \left(y_t + \sum_{i=1}^p b_i y_{t-i} + \sum_{j=1}^q c_j z_{jt} \right)^2$$

и потому формально являются оценками наименьших квадратов. Следствием этого является то, что в этих моделях можно использовать все вычислительные процедуры обычного регрессионного анализа.

В § 5.5 мы показали, что при надлежащим образом выбранных условиях соответствующие оценки $\hat{\beta}$ и $\hat{\gamma}$ имеют асимптотически нормальное распределение, как и в случае обычной регрессии. Таким образом, при больших выборках (т. е. при больших значениях T) эти оценки можно рассматривать как нормально распределенные и использовать при этом процедуры обычного регрессионного анализа. Аппроксимирующее многомерное нормальное распределение статистик $\sqrt{T}(\hat{\beta} - \beta)$ и $\sqrt{T}(\hat{\gamma} - \gamma)$ имеет ковариационную матрицу

$$(2) \quad \sigma^2 \begin{pmatrix} F + H & L \\ L' & M \end{pmatrix}^{-1}.$$

В свою очередь матрица (2) аппроксимируется матрицей

$$(3) \quad s^2 \begin{pmatrix} A_T & L_T \\ L'_T & M_T \end{pmatrix}^{-1},$$

где

$$(4) \quad s^2 = \frac{1}{T-p-q} \sum_{t=1}^T (y_t + \hat{\beta}'\tilde{y}_{t-1} + \hat{\gamma}'z_t)^2 = \frac{T\hat{\sigma}^2}{T-p-q},$$

$$(5) \quad A_T = \frac{1}{T} \sum_{t=1}^T \tilde{y}_{t-1}\tilde{y}_{t-1}' =$$

$$= \begin{bmatrix} \frac{1}{T} \sum_{t=1}^T y_{t-1}^2 & \frac{1}{T} \sum_{t=1}^T y_{t-1}y_{t-2} & \dots & \frac{1}{T} \sum_{t=1}^T y_{t-1}y_{t-p} \\ \frac{1}{T} \sum_{t=1}^T y_{t-2}y_{t-1} & \frac{1}{T} \sum_{t=1}^T y_{t-2}^2 & \dots & \frac{1}{T} \sum_{t=1}^T y_{t-2}y_{t-p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{T} \sum_{t=1}^T y_{t-p}y_{t-1} & \frac{1}{T} \sum_{t=1}^T y_{t-p}y_{t-2} & \dots & \frac{1}{T} \sum_{t=1}^T y_{t-p}^2 \end{bmatrix},$$

$$(6) \quad L_T = \frac{1}{T} \sum_{t=1}^T \tilde{y}_{t-1}z_t' =$$

$$= \begin{bmatrix} \frac{1}{T} \sum_{t=1}^T y_{t-1}z_{1t} & \frac{1}{T} \sum_{t=1}^T y_{t-1}z_{2t} & \dots & \frac{1}{T} \sum_{t=1}^T y_{t-1}z_{qt} \\ \frac{1}{T} \sum_{t=1}^T y_{t-2}z_{1t} & \frac{1}{T} \sum_{t=1}^T y_{t-2}z_{2t} & \dots & \frac{1}{T} \sum_{t=1}^T y_{t-2}z_{qt} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{T} \sum_{t=1}^T y_{t-p}z_{1t} & \frac{1}{T} \sum_{t=1}^T y_{t-p}z_{2t} & \dots & \frac{1}{T} \sum_{t=1}^T y_{t-p}z_{qt} \end{bmatrix},$$

$$(7) \quad \mathbf{M}_T = \frac{1}{T} \sum_{t=1}^T \mathbf{z}_t \mathbf{z}_t' =$$

$$= \begin{bmatrix} \frac{1}{T} \sum_{t=1}^T z_{1t}^2 & \frac{1}{T} \sum_{t=1}^T z_{1t} z_{2t} & \dots & \frac{1}{T} \sum_{t=1}^T z_{1t} z_{qt} \\ \frac{1}{T} \sum_{t=1}^T z_{2t} z_{1t} & \frac{1}{T} \sum_{t=1}^T z_{2t}^2 & \dots & \frac{1}{T} \sum_{t=1}^T z_{2t} z_{qt} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{T} \sum_{t=1}^T z_{qt} z_{1t} & \frac{1}{T} \sum_{t=1}^T z_{qt} z_{2t} & \dots & \frac{1}{T} \sum_{t=1}^T z_{qt}^2 \end{bmatrix}.$$

Обычно, полагая $z_{1t} = 1$, в модель вводят некоторую константу. При этом если $\hat{\gamma}_1$ исключить из уравнений для оценок, то все используемые суммы квадратов и попарных произведений будут включать отклонения от соответствующих средних. (Мы опустили знак $\sim$ у $\mathbf{F}$, $\mathbf{H}$, и $\mathbf{L}$, так как в § 5.6 общий векторный случай не рассматривается.)

Как было замечено в конце § 5.4, элементы каждой диагонали матрицы $(1/T) \sum_{t=1}^T \tilde{\mathbf{y}}_{t-1} \tilde{\mathbf{y}}_{t-1}'$ различаются только членами, относящимися к началу и концу ряда. Поскольку мы занимаемся теорией больших выборок, то этими «концевыми» эффектами можно пренебречь и использовать матрицу, на диагоналях которой стоят одинаковые элементы. Кроме того, можно изменить и нормировочный множитель $(1/T)$. Тогда каждая диагональ, отстоящая на s от главной диагонали, будет состоять из элементов

$$(8) \quad \frac{1}{T+p-s} \sum_{t=-p+s+1}^T y_t y_{t-s}, \quad s = 0, 1, \dots, p.$$

Если вычитается среднее $\bar{y}^* = \sum_{t=-(p-1)}^T y_t / (T+p)$, то сумма примет вид

$$(9) \quad \frac{1}{T+p-s} \sum_{t=-p+s+1}^T (y_t - \bar{y}^*) (y_{t-s} - \bar{y}^*).$$

Возможны и другие модификации, например $\sum_{t=-p+s+1}^T y_t y_{t-s} / (T+p-s) - \bar{y}^{*2}$. Отметим, что мы располагаем наблюдениями на отрезке от $-p+1$ до T .

5.6.2. Проверка гипотез и доверительные интервалы

Рассмотрим статистики

$$(10) \quad \sqrt{T} \frac{\hat{\beta}_i - \beta_i}{s \sqrt{\bar{a}^{ii}}}, \quad \sqrt{T} \frac{\hat{\gamma}_j - \gamma_j}{s \sqrt{\bar{m}^{jj}}},$$

где $s^2 \bar{a}^{ii}$ и $s^2 \bar{m}^{jj}$ соответственно i -й и $(p + j)$ -й диагональные элементы матрицы из (3). Из обычной теории регрессии можно вывести, что каждая из статистик в (10) имеет распределение, близкое к нормальному с нулевым средним и единичной дисперсией. Поэтому для проверки гипотез относительно β_i или γ_j и построения доверительных интервалов для этих параметров можно использовать обычные методы.

Пусть нас интересует некоторая совокупность коэффициентов, например $\beta_{i_1}, \dots, \beta_{i_g}, \gamma_{j_1}, \dots, \gamma_{j_h}$. Тогда можно воспользоваться тем, что статистика

$$(11) \quad \frac{1}{s^2} \left\{ \sum_{u,v=1}^g (\hat{\beta}_{i_u} - \beta_{i_u}) (\hat{\beta}_{i_v} - \beta_{i_v}) \bar{a}_{i_u i_v} + \right. \\ \left. + 2 \sum_{u=1}^g \sum_{v=1}^h (\hat{\beta}_{i_u} - \beta_{i_u}) (\hat{\gamma}_{j_v} - \gamma_{j_v}) \bar{l}_{i_u j_v} + \right. \\ \left. + \sum_{u,v=1}^h (\hat{\gamma}_{j_u} - \gamma_{j_u}) (\hat{\gamma}_{j_v} - \gamma_{j_v}) \bar{m}_{j_u j_v} \right\},$$

имеет распределение, близкое к распределению χ^2 с $g + h$ степенями свободы. Здесь $\bar{a}_{i_u i_v}$, $\bar{l}_{i_u j_v}$, $\bar{m}_{j_u j_v}$ образуют матрицу, обратную к матрице, составленной из строк и столбцов с номерами $i_1, \dots, i_g, j_1 + p, \dots, j_h + p$ матрицы $\begin{pmatrix} \mathbf{A}_T & \mathbf{L}_T \\ \mathbf{L}_T & \mathbf{M}_T \end{pmatrix}^{-1}$. Это

приводит к соответствующим критериям и доверительным интервалам. (Можно, кроме того, использовать и F -распределение с $g + h$ и $T - (p + q)$ степенями свободы.)

5.6.3. Проверка гипотез о порядке процесса авторегрессии

Порядок стохастического разностного уравнения определяется входящим в него (с ненулевым коэффициентом) переменным, имеющим наибольшее запаздывание. При той форме записи уравнения, которую мы используем, порядок равен просто p ($\beta_p \neq 0$). Если не считать влияния «независимых переменных» z_{it} , то непосредственно на значение y_t влияет самое большее p запаздывающих переменных, именно $y_{t-1}, \dots, y_{t-p}$. Поэтому для прогнозирования следует использовать только такие переменные.

В связи с этим представляет интерес определение порядка модели авторегрессии. Мы исследуем эту задачу для разностного уравнения

$$(12) \quad y_t + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \gamma = u_t.$$

Уравнения для оценок параметров $\beta_1, \dots, \beta_p$ имеют в этом случае вид

$$(13) \quad \sum_{j=1}^p a_{ij}^* \hat{\beta}_j = -a_{i0}^*, \quad i = 1, \dots, p,$$

где

$$(14) \quad a_{ij}^* = \sum_{t=1}^T y_{t-i} y_{t-j} - T \bar{y}_{(i)} \bar{y}_{(j)}, \quad i, j = 1, \dots, p,$$

$$(15) \quad a_{i0}^* = \sum_{t=1}^T y_{t-i} y_t - T \bar{y}_{(i)} \bar{y}, \quad i = 1, \dots, p,$$

$\bar{y} = \bar{y}_{(0)}$, а $\bar{y}_{(i)}$ задается соотношением (17) из § 5.4. [Уравнения (13) соответствуют (18) из § 5.4.]

Рассмотрим сначала задачу выбора между решениями о том, что порядок уравнения равен p или $p-1$ соответственно. Иными словами, требуется решить, будет ли $\beta_p = 0$. Мы поставим эту задачу как задачу проверки нулевой гипотезы $H: \beta_p = 0$ против альтернативы $\beta_p \neq 0$. Критерий отношения правдоподобия будет здесь аналогичен двустороннему t -критерию, использовавшемуся при исследовании регрессии. Мы отвергнем нулевую гипотезу, если

$$(16) \quad \frac{|\hat{\beta}_p|}{s \sqrt{a^{*pp}}} > t(\epsilon),$$

где $(a^{*ij}) = (a_{ij}^*)^{-1}$, а $t(\epsilon)$ — двусторонняя 100 ϵ -процентная точка стандартного нормального распределения. При достаточно больших значениях T уровень значимости этого критерия будет близок к ϵ .

Пусть нулевая гипотеза неверна. В общем случае мы имеем

$$(17) \quad \text{plim}_{T \rightarrow \infty} \left(\frac{1}{T} a_{ij}^* \right) = F,$$

где F выражается соотношением (37) § 5.5, $\text{plim}_{T \rightarrow \infty} s = \sigma$ и $\text{plim}_{T \rightarrow \infty} \hat{\beta}_p = \beta_p$. Статистика критерия, стоящая в левой части (16), ведет себя примерно как $\sqrt{T} |\beta_p| / (\sigma \sqrt{f^{pp}})$. Это утверждение можно сделать строгим и использовать его для доказательства состоятельности указанного критерия. Иными словами, можно доказать, что для заданной альтернативы $\beta_p \neq 0$ вероятность отвергнуть нулевую гипотезу с ростом T приближается к единице.

Бóльший интерес может представлять изучение мощности указанного критерия для последовательности конкурирующих гипотез β_{pT} , такой, что

$$(18) \quad \lim_{T \rightarrow \infty} \sqrt{T} \beta_{pT} = \varphi \neq 0$$

(при этом σ^2 , $\beta_1, \dots, \beta_{p-1}$ предполагаются фиксированными, а u_t , например, нормально распределенными). Тогда $\hat{\beta}_p / (s \sqrt{a^{*pp}})$ имеет в пределе нормальное распределение с единичной дисперсией и средним, равным $\varphi / (\sigma \sqrt{f^{pp}})$. Для того чтобы сделать это утверждение (а вместе с ним и ряд последующих) строгим, надо было бы доказать, что $\sqrt{T} (\hat{\mathbf{B}} - \mathbf{B}_T)$ имеет в пределе нормальное распределение с матрицей средних Φ и ковариационной матрицей, соответствующей $\mathbf{B}$, если $\lim_{T \rightarrow \infty} \sqrt{T} (\mathbf{B} - \mathbf{B}_T) = \Phi$ для надлежащим образом

выбранной матрицы $\mathbf{B}$. (Здесь $\hat{\mathbf{B}}$, $\mathbf{B}$ и $\mathbf{B}_T$ относятся к векторному марковскому случаю.)

ЛЕММА 5.6.1. Если $\beta_p = 0$, то

$$(19) \quad \sigma^2 f^{pp} = 1.$$

ДОКАЗАТЕЛЬСТВО. Прежде всего мы имеем (см. упр. 8 гл. 2)

$$(20) \quad \frac{1}{f^{pp}} = f_{pp} - (f_{p1} \dots f_{p,p-1}) \begin{pmatrix} f_{11} & \dots & f_{1,p-1} \\ \vdots & & \vdots \\ f_{p-1,1} & \dots & f_{p-1,p-1} \end{pmatrix}^{-1} \begin{pmatrix} f_{1p} \\ \vdots \\ f_{p-1,p} \end{pmatrix}.$$

В стационарном случае величины

$$(21) \quad \begin{aligned} f_{ij} &= \mathbb{E}(y_{t-i} - \mathbb{E}y_{t-i})(y_{t-j} - \mathbb{E}y_{t-j}) = \\ &= \mathbb{E}(y_s - \mathbb{E}y_s)(y_{s-j+i} - \mathbb{E}y_{s-j+i}) = \sigma(i-j), \\ & \qquad \qquad \qquad i, j = 1, \dots, p, \end{aligned}$$

при $\beta_p = 0$ соответствуют вполне определенному разностному уравнению порядка $p-1$, и они должны удовлетворять следующему соотношению [см. (48) из § 5.2]:

$$(22) \quad \begin{pmatrix} f_{11} & \dots & f_{1,p-1} \\ \vdots & & \vdots \\ f_{p-1,1} & \dots & f_{p-1,p-1} \end{pmatrix} \begin{pmatrix} \beta_{p-1} \\ \vdots \\ \beta_1 \end{pmatrix} = - \begin{pmatrix} f_{1p} \\ \vdots \\ f_{p-1,p} \end{pmatrix}.$$

Соотношение (20) принимает при этом вид [см. (47) из § 5.2]

$$(23) \quad \frac{1}{f^{pp}} = \sigma(0) + \sigma(p-1)\beta_{p-1} + \dots + \sigma(1)\beta_1 = \sigma^2.$$

Это и доказывает лемму. ■

В свою очередь доказанная лемма приводит к следующей теореме.

ТЕОРЕМА 5.6.1. Если $\beta_p = 0$, то в условиях теоремы 5.5.7 статистика $\sqrt{T}\hat{\beta}_p$ имеет в пределе при $T \rightarrow \infty$ нормальное распределение с нулевым средним и единичной дисперсией.

Рассмотрим теперь задачу проверки нулевой гипотезы о том, что порядок стохастического разностного уравнения равен m , против конкурирующей гипотезы, состоящей в том, что порядок этого уравнения равен p ($> m$). Пусть матрица $A^* = (a_{ij}^*)$ разбита на блоки с выделением m и $p - m$ строк и столбцов и пусть соответствующим образом разбивается вектор $\hat{\beta}$, так что

$$(24) \quad A^* = \begin{pmatrix} A_{11}^* & A_{12}^* \\ A_{21}^* & A_{22}^* \end{pmatrix}, \quad \hat{\beta} = \begin{pmatrix} \hat{\beta}^{(1)} \\ \hat{\beta}^{(2)} \end{pmatrix}.$$

Тогда статистика критерия отношения правдоподобия для проверки гипотезы о том, что $\beta_{m+1} = \dots = \beta_p = 0$, будет монотонной функцией статистики

$$(25) \quad \frac{\hat{\beta}^{(2)'} (A_{22}^* - A_{21}^* (A_{11}^*)^{-1} A_{12}^*) \hat{\beta}^{(2)}}{s^2},$$

имеющей в случае истинности нулевой гипотезы предельное χ^2 -распределение с $p - m$ степенями свободы. Имеющаяся здесь аналогия с обычными регрессионными методами позволяет предположить, что мощность этого критерия зависит главным образом от некоторой функции параметров, измеряющей увеличение среднеквадратичной ошибки прогнозирования из-за использования m составляющих вместо p . [См. § 2.5. Это было отмечено А. Уолкером (1952).]

Другой подход к проверке гипотезы о том, что $\beta_{m+1} = \dots = \beta_p = 0$, был предложен Кенуем (1947) и развит затем Барлеттом и Дианандой (1950), а также другими авторами. Для простоты изложения положим $\gamma = 0$ и используем в связи с этим матрицу A вместо A^* . Если $\{y_t\}$ является стационарным процессом авторегрессии произвольного порядка, например p , с коэффициентами $\beta_1, \dots, \beta_m, 0, \dots, 0$, то величины

$$(26) \quad g_i = \frac{1}{\sqrt{T}} \sum_{t=1}^T y_{t-i} u_t = \\ = \frac{1}{\sqrt{T}} \left[\sum_{t=1}^T y_t y_{t-i} + \right. \\ \left. + \beta_1 \sum_{t=1}^T y_{t-1} y_{t-i} + \dots + \beta_m \sum_{t=1}^T y_{t-m} y_{t-i} \right] = \\ = \frac{1}{\sqrt{T}} \sum_{i=0}^m \beta_i a_{ii}, \quad i = 1, \dots, p,$$

имеют нулевые средние и матрица их ковариаций (порядка p) равна $\sigma^2 F = \sigma^2 [\sigma(i - j)]$. Из теоремы 5.5.5 вытекает, что $g_1, \dots, g_p$ имеют в пределе совместное нормальное распределение. Если y_t нормально распределены, то предельное совместное распределение любого набора случайных величин g_i совпадает с распределением аналогичного набора случайных величин σu_t . При этом совместное распределение случайных величин

$$(27) \quad \begin{aligned} h_j &= \sum_{k=0}^m \beta_k g_{j-k} = \frac{1}{\sqrt{T}} \sum_{k=0}^m \beta_k \sum_{t=1}^T y_{t-j+k} u_t = \\ &= \frac{1}{\sqrt{T}} \sum_{k,l=0}^m \beta_k \beta_l \sum_{t=1}^T y_{t-j+k} y_{t-l} = \frac{1}{\sqrt{T}} \sum_{k,l=0}^m \beta_k \beta_l a_{l,j-k}, \\ &\quad j = m+1, \dots, p, \end{aligned}$$

является в пределе нормальным. При любом T оно имеет нулевое среднее. Ковариации его вычисляются с использованием соотношения

$$(28) \quad \begin{aligned} \mathcal{E} h_j g_i &= \sum_{k=0}^m \beta_k \mathcal{E} g_{j-k} g_i = \sigma^2 \sum_{k=0}^m \beta_k \sigma (j - k - i) = \\ &= \begin{cases} 0, & j > i, \\ \sigma^4, & j = i, \end{cases} \end{aligned}$$

полученного с помощью (47) и (48) § 5.2, и равны

$$(29) \quad \mathcal{E} h_j h_{j'} = \sum_{k=0}^m \beta_k \mathcal{E} h_j g_{j'-k} = \begin{cases} 0, & j > j', \\ \sigma^4, & j = j'. \end{cases}$$

Таким образом, при нулевой гипотезе совместное распределение случайных величин $h_{m+1}/\sigma^2, h_{m+2}/\sigma^2, \dots, h_p/\sigma^2$ имеет нулевое среднее и единичную матрицу ковариаций. В пределе при $T \rightarrow \infty$ это распределение является стандартным нормальным.

Исходя из указанных результатов, статистики h_j можно было бы использовать для проверки гипотезы $\beta_{m+1} = \dots = \beta_p = 0$ при известных значениях σ^2 и $\beta_1, \dots, \beta_m$, поскольку статистика

$$(30) \quad \frac{1}{\sigma^4} \sum_{j=m+1}^p h_j^2$$

имеет χ^2 -распределение с $p - m$ степенями свободы. Для большей эффективности этой процедуры А. Уолкер (1952) предложил добавлять еще одну статистику с тем, чтобы предельное распределение критерия было χ^2 -распределением с p степенями свободы.

Для проверки гипотезы $\beta_{m+1} = \dots = \beta_p = 0$ при неизвестных значениях остальных параметров было предложено использовать статистику (30) в несколько измененном виде с заменой неизвестных

параметров в (30) их оценками. Положим

$$(31) \quad \hat{h}_i = \sum_{k=0}^m \hat{\beta}_k \hat{g}_{i-k} = \frac{1}{\sqrt{T}} \sum_{k,l=0}^m \hat{\beta}_k \hat{\beta}_l a_{i,j-k}, \quad i = m+1, \dots, p,$$

где $\hat{\beta}_0 = 1, \hat{\beta}_1, \dots, \hat{\beta}_m$ — оценки максимального правдоподобия, указанные в § 5.4 (с заменой p на m). Рассмотрим критерий $\sum_{j=m+1}^p \hat{h}_j^2/s^4$ и убедимся в том, что при нулевой гипотезе он асимптотически эквивалентен критерию (30). Действительно, поскольку $\hat{\beta}_k$ является состоятельной оценкой для β_k и $\text{plim}_{T \rightarrow \infty} a_{ij}/T = \sigma(i-j)$, то

$$(32) \quad \text{plim}_{T \rightarrow \infty} \sum_{k=0}^m \hat{\beta}_k \frac{1}{T} a_{i,j-k} = \sum_{k=0}^m \beta_k \sigma(j-k-l) = 0, \quad j > l,$$

$$(33) \quad \text{plim}_{T \rightarrow \infty} \sum_{l=0}^m \hat{\beta}_l \frac{1}{T} a_{i,j-k} = \sum_{l=0}^m \beta_l \sigma(j-k-l) = 0, \quad i > k.$$

Поэтому

$$\begin{aligned} (34) \quad \text{plim}_{T \rightarrow \infty} (\hat{h}_i - h_i) &= \\ &= \text{plim}_{T \rightarrow \infty} \sqrt{T} \left[\sum_{k,l=0}^m \hat{\beta}_k \hat{\beta}_l \frac{1}{T} a_{i,j-k} - \right. \\ &\quad \left. - \sum_{k,l=0}^m \beta_k \beta_l \frac{1}{T} a_{i,j-k} \right] = \\ &= \text{plim}_{T \rightarrow \infty} \sum_{k,l=0}^m \sqrt{T} (\hat{\beta}_k - \beta_k) \hat{\beta}_l \frac{1}{T} a_{i,j-k} + \\ &\quad + \text{plim}_{T \rightarrow \infty} \sum_{k,l=0}^m \beta_k \sqrt{T} (\hat{\beta}_l - \beta_l) \frac{1}{T} a_{i,j-k} = \\ &= 0, \quad i = m+1, \dots, p. \end{aligned}$$

Поскольку к тому же s^2 (построенная в предположении, что порядок стохастического разностного уравнения равен m) является состоятельной оценкой для σ^2 при нулевой гипотезе, то из этих двух фактов и следует упомянутая асимптотическая эквивалентность.

ТЕОРЕМА 5.6.2. Если все корни характеристического уравнения, соответствующего стохастическому разностному уравнению, по

рядка m , лежат в единичном круге, то статистика

$$(35) \quad \frac{\sum_{j=m+1}^p \hat{h}_j^2}{s_m^4},$$

имеет в пределе χ^2 -распределение с $p - m$ степенями свободы. Здесь $\hat{h}_j$ определяется соотношением (31), а s_m^2 является оценкой для σ^2 , соответствующей указанному разностному уравнению.

Свяжем теперь оба описанных подхода к задаче проверки гипотезы о порядке стохастического разностного уравнения.

ЛЕММА 5.6.2. Если $p = m + 1$, то

$$(36) \quad \sqrt{T} \hat{\beta}_{m+1}(m+1) = - \frac{\hat{h}_{m+1}}{\tilde{s}_m^2},$$

где $\hat{\beta}_{m+1}(m+1)$ — оценка для β_{m+1} в предположении, что процесс имеет порядок $m+1$, а $\tilde{s}_m^2$ — оценка для σ^2 , указанная ниже, в (45).

ДОКАЗАТЕЛЬСТВО. Пусть матрица сумм квадратов и перекрестных произведений с элементами

$$(37) \quad a_{ij} = \sum_{t=1}^T y_{t-i} y_{t-j}, \quad i, j = 0, 1, \dots, m+1,$$

разбита на блоки следующим образом:

$$(38) \quad \mathbf{A} = \begin{pmatrix} a_{00} & \mathbf{A}_{01} & a_{0,m+1} \\ \mathbf{A}_{10} & \mathbf{A}_{11} & \mathbf{A}_{1,m+1} \\ a_{m+1,0} & \mathbf{A}_{m+1,1} & a_{m+1,m+1} \end{pmatrix}.$$

Тогда оценкой вектора $\beta(m) = [\beta_1(m), \dots, \beta_m(m)]'$, в предположении, что разностное уравнение имеет порядок m , будет

$$(39) \quad \hat{\beta}(m) = -\mathbf{A}_{11}^{-1} \mathbf{A}_{10}.$$

Если выразить $\hat{h}_{m+1}$ через указанные блоки матрицы $\mathbf{A}$, то получим

$$(40) \quad \begin{aligned} \sqrt{T} \hat{h}_{m+1} = & a_{0,m+1} + \sum_{k=1}^m \hat{\beta}_k(m) a_{0,m+1-k} + \sum_{l=1}^m \hat{\beta}_l(m) a_{l,m+1} + \\ & + \sum_{k,l=1}^m \hat{\beta}_k(m) \hat{\beta}_l(m) a_{l,m+1-k} = \end{aligned}$$

$$\begin{aligned}
&= a_{0,m+1} + \sum_{j=1}^m \hat{\beta}_{m+1-j}(m) a_{0j} + A_{m+1,1} \hat{\beta}(m) + \\
&\quad + \sum_{i,l=1}^m \hat{\beta}_l(m) \hat{\beta}_{m+1-j}(m) a_{il} = \\
&= a_{0,m+1} - A_{m+1,1} A_{11}^{-1} A_{10}.
\end{aligned}$$

Уравнения для оценок величин $\beta_1, \dots, \beta_{m+1}$, использующих разностное уравнение $(m+1)$ -го порядка, имеют вид

$$(41) \quad \begin{pmatrix} A_{11} & A_{1,m+1} \\ A_{m+1,1} & a_{m+1,m+1} \end{pmatrix} \begin{pmatrix} \hat{\beta}^{(1)}(m+1) \\ \hat{\beta}_{m+1}(m+1) \end{pmatrix} = - \begin{pmatrix} A_{10} \\ a_{m+1,0} \end{pmatrix},$$

т.е.

$$(42) \quad A_{11} \hat{\beta}^{(1)}(m+1) + A_{1,m+1} \hat{\beta}_{m+1}(m+1) = -A_{10},$$

$$(43) \quad A_{m+1,1} \hat{\beta}^{(1)}(m+1) + a_{m+1,m+1} \hat{\beta}_{m+1}(m+1) = -a_{m+1,0}.$$

[Отметим, что последние соответствуют уравнениям (29) и (30) из § 5.4.] Если умножить (42) на $A_{m+1,1} A_{11}^{-1}$ и вычесть результат из (43), то получим соотношение

$$(44) \quad (a_{m+1,m+1} - A_{m+1,1} A_{11}^{-1} A_{1,m+1}) \hat{\beta}_{m+1}(m+1) = \\ = -(a_{m+1,0} - A_{m+1,1} A_{11}^{-1} A_{10}),$$

правая часть которого равна $-\sqrt{T} \hat{h}_{m+1}$. Коэффициент при $\hat{\beta}_{m+1}(m+1)$ в левой части в T раз больше величины

$$(45) \quad \frac{a_{m+1,m+1} - A_{m+1,1} A_{11}^{-1} A_{1,m+1}}{T} = \tilde{s}_m^2,$$

асимптотически эквивалентной $\hat{\sigma}_m^2$ (основанной на уравнении порядка m). Соответствующие суммы в матрицах, определяющих $\tilde{s}_m^2$ и $\hat{\sigma}_m^2$, отличаются только крайними членами. Это и доказывает лемму. ■

ЛЕММА 5.6.3. Если $\beta_{m+1} = \dots = \beta_p = 0$ и $\gamma = 0$, то

$$(46) \quad \text{plim}_{T \rightarrow \infty} \left[\frac{\hat{\beta}^{(2)'}(p) (A_{22} - A_{21} A_{11}^{-1} A_{12}) \hat{\beta}^{(2)}(p)}{s^2} - T \sum_{j=m+1}^p \hat{\beta}_j^2(j) \right] = 0.$$

ДОКАЗАТЕЛЬСТВО. Из доказательства леммы 5.6.2 видно, что $\hat{\beta}_j(j)$ является отношением правой части нормального уравнения к

коэффициенту при $\hat{\beta}_j$, если предварительно из системы нормальных уравнений исключить первые $j - 1$ неизвестных $\hat{\beta}_i$. Таким образом, мы имеем ситуацию, подобную соотношению (20) § 2.3, где $b_j^* = c_j/a_{jj}$. В данном случае коэффициент при $\hat{\beta}_j$ равен $T\tilde{s}_{j-1}^2$, а $\tilde{s}_{j-1}^2$ определяется соотношением (45) при $m = j - 1$. Поэтому, как показано в § 2.3, числитель уменьшаемого в (46) равен

$$(47) \quad \sum_{j=m+1}^p T\tilde{s}_{j-1}^2 \hat{\beta}_j^2(j).$$

Для завершения доказательства остается только заметить, что

$$(48) \quad \text{plim}_{T \rightarrow \infty} \frac{\tilde{s}_{j-1}^2}{s^2} = 1. \blacksquare$$

ТЕОРЕМА 5.6.3. Если $\beta_{m+1} = \dots = \beta_p = 0$ и $\gamma = 0$, то разности между (25), (30), (35) и

$$(49) \quad T \sum_{j=m+1}^p \hat{\beta}_j^2(j)$$

стремятся по вероятности к нулю при $T \rightarrow \infty$.

ДОКАЗАТЕЛЬСТВО. Это вытекает из лемм 5.6.2, 5.6.3 и из соотношений (48) и $\text{plim}_{T \rightarrow \infty} s^2 = \sigma^2$. ■

Если справедлива нулевая гипотеза, то каждый из этих критериев при $\gamma = 0$ имеет в пределе χ^2 -распределение с $p - m$ степенями свободы. Если $\gamma \neq 0$, то матрица A заменяется на A^* .

В критерии $\sum_{j=m+1}^p h_j^2/\sigma^4$ можно брать также величины h_j , определенные следующим образом:

$$(50) \quad h_j = \sum_{k=0}^m \beta_k g_{j+k} = \frac{1}{\sqrt{T}} \sum_{k=0}^m \beta_k \sum_{t=1}^T y_{t-j-k} u_t = \\ = \frac{1}{\sqrt{T}} \sum_{t=1}^T u_{t-j} u_t, \quad j = m+1, \dots, p.$$

Эти статистики имеют нулевые средние и дисперсии σ^4 . Они не коррелированы, а их предельное совместное распределение является нормальным. Однако если в таком определении h_j параметры $\beta_1, \dots, \beta_m$ заменить их оценками, то соответствующее предельное распределение изменится.

А. Уолкер (1952) рассмотрел предельную мощность перечисленных критериев против альтернатив, сближающихся с нулевой гипотезой.

Указанная асимптотическая теория не изменяется при изменении крайних членов сумм. В частности, при $\gamma \neq 0$ матрицу Λ^* можно заменить матрицей

$$(51) \quad \mathbf{C}^* = \begin{bmatrix} c_0^* & c_1^* & \dots & c_{p-1}^* & c_p^* \\ c_1^* & c_0^* & \dots & c_{p-1}^* & c_{p-1}^* \\ \vdots & \vdots & & \vdots & \vdots \\ c_{p-1}^* & c_{p-2}^* & \dots & c_0^* & c_1^* \\ c_p^* & c_{p-1}^* & \dots & c_1^* & c_0^* \end{bmatrix},$$

где $c_0^*, \dots, c_p^*$ определяются соотношениями (24) § 5.4 при наблюдениях $y_1, \dots, y_T$. В свою очередь c_j^* могут быть заменены отношениями $r_j^* = c_j^*/c_0^*$, $j = 0, 1, \dots, p$. Отметим, что уравнение (44) настоящего параграфа соответствует уравнению (31) § 5.4 (с переставленными индексами p и m).

Пусть матрица $\mathbf{C}^*$ разбита на блоки,

$$(52) \quad \mathbf{C}^* = \begin{bmatrix} c_0^* & c_{p-1}^{**} & c_p^* \\ c_{p-1}^* & \mathbf{C}_{p-1}^* & \tilde{c}_{p-1}^* \\ c_p^* & \tilde{c}_{p-1}^{**} & c_0^* \end{bmatrix}.$$

Тогда частная корреляция между y_t и y_{t+p} при фиксированных $y_{t+1}, \dots, y_{t+p-1}$ равна

$$(53) \quad \frac{c_p^* - \tilde{c}_{p-1}^{**} (\mathbf{C}_{p-1}^*)^{-1} \tilde{c}_{p-1}^*}{c_0^* - \tilde{c}_{p-1}^{**} (\mathbf{C}_{p-1}^*)^{-1} \tilde{c}_{p-1}^*}.$$

Нормальные уравнения для оценки $\beta(p)$, использующие в качестве оценок для $\sigma(j)$, $j = 0, 1, \dots, p$, статистики c_j^* , имеют вид

$$(54) \quad \begin{pmatrix} \mathbf{C}_{p-1}^* & \tilde{c}_{p-1}^* \\ \tilde{c}_{p-1}^{**} & c_0^* \end{pmatrix} \begin{pmatrix} b^{(1)}(p) \\ b_p(p) \end{pmatrix} = - \begin{pmatrix} c_{p-1}^* \\ c_p^* \end{pmatrix}.$$

Решением (54) относительно $-b_p(p)$ является (53). Таким образом, с точностью до эффекта, вносимого крайними членами, оценка параметра β_p в предположении, что порядок равен p , совпадает с коэффициентом частной корреляции между y_t и y_{t+p} при фиксированных значениях $y_{t+1}, \dots, y_{t+p-1}$.

Теперь следует рассмотреть вопрос о выборе надлежащей степени стохастического разностного уравнения. Аналогия с обычной регрессией (§ 3.2) наводит на мысль о том, что если исследователь

в состоянии установить наибольший p и наименьший m возможные порядки уравнения, то он может последовательно проверять гипотезы $\beta_p = 0$, $\beta_{p-1} = 0$ и т. д. Мы изучим этот вопрос в гл. 6 с помощью видоизмененной модели, позволяющей использовать точную теорию, подобную теории § 3.2.

При больших T можно использовать ту же теорию, что и в предположении нормальности. Максимальный порядок p может быть выбран достаточно большим, скажем $T/10$ или $T/4$, если значения ε_i , $i = p, p-1, \dots, m+1$, берутся достаточно малыми, такими, как 10^{-3} . Ряды изучаются, как правило, по той причине, что наблюдения оказываются зависимыми. Поэтому минимальный порядок m следует выбирать положительным, но малым, например равным 2. Для проверки гипотезы $\beta_p = 0$ при больших выборках нужна только оценка $\hat{\beta}_p(p)$. Фактически для выполнения процедуры со многими решениями здесь нужны только оценки $\hat{\beta}_j(j)$ для $j = p$ и для тех меньших значений j , при которых гипотеза $\beta_j = 0$ еще принимается, вплоть до значения j , при котором она впервые отвергается. Вычисление этих оценок исходя из $r_1^*, \dots, r_p^*$ рассматривается в § 5.4. Необходимые статистики можно также получить при прямом решении уравнения $R_p^* b(p) = -r_p^*$.

5.7. МОДЕЛЬ СКОЛЬЗЯЩЕГО СРЕДНЕГО

5.7.1. Модель

Пусть

$$(1) \quad y_t = \sum_{j=0}^q \alpha_j v_{t-j},$$

где $\alpha_0 = 1$, а $\{v_t\}$ — последовательность независимых случайных величин со средними $\mathbb{E}v_t = 0$ и дисперсиями $\mathbb{E}v_t^2 = \sigma^2$. Тогда $\mathbb{E}y_t = 0$ и

$$(2) \quad \mathbb{E}y_t y_s = \sigma(t-s) = \begin{cases} \sigma^2 \sum_{j=0}^{q-|t-s|} \alpha_j \alpha_{j+|t-s|}, & |t-s| \leq q, \\ 0, & |t-s| > q. \end{cases}$$

Такой вид моментов первого и второго порядков определяется по существу только некоррелированностью величин v_t . Соотношение (1) можно записать также в виде

$$(3) \quad y_t = \sum_{j=0}^q \alpha_j \mathcal{L}^j v_t = \mathcal{L}^q M(\mathcal{L}^{-1}) v_t,$$

где $\mathcal{L}$ — оператор запаздывания ($\mathcal{L}v_t = v_{t-1}$),

$$(4) \quad M(z) = \sum_{j=0}^q \alpha_j z^{q-j}$$

и $\mathcal{L}^q M(\mathcal{L}^{-1})$ — линейный оператор, полученный заменой z на $\mathcal{L}$ в выражении $z^q M(z^{-1}) = \sum_{j=0}^q \alpha_j z^j$.

Для произвольного стационарного процесса выражение $\sum_{h=-\infty}^{\infty} \sigma(h) z^h$ назовем *производящей функцией ковариаций*. Ковариация $\sigma(h)$ является коэффициентом при z^h в этом разложении. Для процесса скользящего среднего (1), имеющего ковариации (2), производящая функция ковариаций принимает вид

$$(5) \quad \sum_{h=-q}^q \sigma(h) z^h = \sigma^2 M(z) M(z^{-1}).$$

[Отметим, что это уравнение, а вместе с ним и соотношение (2) могут быть получены из леммы 3.4.1.] Если коэффициенты $\alpha_0 = 1, \alpha_1, \dots, \alpha_q$ заданы, то для вычисления ковариаций можно образовать произведение, стоящее в правой части (5), и найти коэффициенты при соответствующих степенях переменной z . Поскольку α_j действительны, то корни $z_1, \dots, z_q$ уравнения

$$(6) \quad M(z) = 0$$

либо действительны, либо образуют пары комплексно сопряженных корней. Если $\alpha_q \neq 0$ [и отсюда $\sigma(q) \neq 0$], то корнями уравнения

$$(7) \quad \sum_{g=0}^{2q} \sigma(g-q) z^g = 0$$

будут величины $z_1, \dots, z_q, 1/z_1, \dots, 1/z_q$.

Пусть нам дан произвольный набор значений ковариаций $\sigma(0), \sigma(1), \dots, \sigma(q) \neq 0$ стационарного процесса, причем $\sigma(q+1) = \dots = 0$. Если x_j является корнем уравнения (7), то, ввиду того что $\sigma(h) = \sigma(-h)$, корнем (7) будет также и $1/x_j$. Если корень по абсолютной величине меньше 1, то кратности этого корня и корня, обратного ему, совпадают (по той причине, что в этом случае при обращении в нуль производной некоторого порядка в точке x_j та же производная будет обращаться в нуль и в точке $1/x_j$). Таким образом, корни, отличные по абсолютной величине от единицы, можно объединить в пары $(x_j, 1/x_j)$. Более того, любой корень, лежащий на единичной окружности, должен иметь четную кратность. Дело в том, что если $|z| = 1$, например $z = e^{i\lambda}$, то

$\sum_{g=-q}^q \sigma(g) z^g$ является при этом спектральной плотностью, и поэтому

она будет действительной и неотрицательной (§ 7.3). (См. упр. 35.) Корни уравнения (7) можно перенумеровать и разбить на два таких множества $(x_1, \dots, x_q)$ и $(x_{q+1}, \dots, x_{2q})$, что если $|x_i| < 1$, то $x_{q+i} = 1/x_i$, а если $|x_i| = 1$, то $x_{q+i} = x_i$, $i = 1, \dots, q$. Поскольку коэффициенты в (7) действительны, то эти корни действительны или попарно сопряжены. Поэтому

$$(8) \quad M^*(z) = \prod_{j=1}^q (z - x_j) = \sum_{j=0}^q \alpha_j^* z^{q-j}$$

имеет действительные коэффициенты и (5) остается в силе, если $M(z)$ заменить на $M^*(z)$. Таким образом, последовательность ковариаций $\{\sigma(h)\}$, у которой $\sigma(q+1) = \dots = 0$, может быть порождена с помощью процесса конечного скользящего среднего (1), имеющего коэффициенты $\alpha_0^* = 1, \alpha_1^*, \dots, \alpha_q^*$ и $\text{Var}(v_t) = \sigma^2$.

Если заданы ковариации $\sigma(0), \sigma(1), \dots, \sigma(q)$ или дисперсия $\sigma(0)$ и корреляции $\rho_h = \sigma(h)/\sigma(0)$, $h = 1, \dots, q$, то коэффициенты соответствующего процесса скользящего среднего можно получить, исходя из изложенного, следующим образом. Решая уравнение (7), получаем значения $x_1, \dots, x_{2q}$. Затем образуем сумму $M^*(z)$ и находим коэффициенты при степенях z .

Если все корни уравнения (6) лежат в единичном круге, то ряд

$$(9) \quad \left(\sum_{j=0}^q \alpha_j x^j \right)^{-1} = \prod_{j=1}^q (1 - z_j x)^{-1} = \prod_{j=1}^q \sum_{v=0}^{\infty} (z_j x)^v = \sum_{r=0}^{\infty} \gamma_r x^r$$

сходится при $|x| < 1/\max_i |z_i|$. Мы можем записать

$$(10) \quad v_t = \left(\sum_{j=0}^q \alpha_j \mathcal{L}^j \right)^{-1} y_t = \sum_{r=0}^{\infty} \gamma_r \mathcal{L}^r y_t = \sum_{r=0}^{\infty} \gamma_r y_{t-r}.$$

Если все корни (6) лежат в единичном круге, то (10) будет сходиться в среднеквадратичном. Здесь можно использовать доводы, аналогичные применявшимся в § 5.2 для доказательства сходимости в среднем ряда (28). [См. также (34) — (36) из § 7.5.] Фактически

если все корни различны, то $\gamma_r = \sum_{i=1}^q k_i z_i^r$, где константы $k_1, \dots, k_q$ выбраны надлежащим образом. [См. (39) — (41) § 5.2.] Отметим, что последнему условию не удовлетворяют такие простые модели, как среднее $(v_t + v_{t-1})/2$ и разность $\Delta v_t = v_{t+1} - v_t$.

5.7.2. Оценивание параметров

Если случайные величины v_t распределены нормально, то наблюдения будут нормально распределенными с нулевыми средними и ковариациями (2). Мы займемся сейчас оцениванием $q+1$

параметров рассматриваемой модели по T наблюдениям: $y_1, \dots, y_T$. К сожалению, хотя ковариационная матрица имеет простой вид, этого нельзя сказать об обратной к ней матрице. Действительно, минимальное достаточное множество статистик состоит здесь из T компонент, а уравнения максимального правдоподобия весьма сложны и не могут быть решены непосредственно. [См. упр. 4 и 5 гл. 6.]

Подход, которому следовал А. Уолкер (1961), состоит в применении метода максимального правдоподобия, когда распределение некоторых n выборочных корреляций близко к нормальному. Как будет показано в теореме 5.7.1, вектор $\sqrt{T}(r_1 - \rho_1), \dots, \sqrt{T}(r_n - \rho_n)$ имеет в пределе нормальное распределение с нулевым средним и ковариациями $\omega_{gh} = \omega_{gh}(\rho_1, \dots, \rho_q)$, являющимися функциями от корреляций $\rho_h = \sigma(h)/\sigma(0)$, $h = 1, \dots, q$. Соответствующие выборочные корреляции равны $r_h = c_h/c_0$, где

$$(11) \quad c_h = c_{-h} = \frac{1}{T} \sum_{t=1}^{T-h} y_t y_{t+h}, \quad h = 0, 1, \dots, T-1.$$

Здесь n ($n \geq q$, $n \leq T-1$) фиксировано. По сути дела, тот факт, что максимизируется распределение выборочных корреляций, близкое к нормальному, служит здесь лишь мотивировкой для получения соответствующих уравнений относительно оценок. Свойства же процедур оценивания не зависят от этого приближения.

Положим $\mathbf{p} = (\rho_1, \dots, \rho_q)'$. Логарифм функции, приближающей функцию правдоподобия относительно $\mathbf{r} = (r_1, \dots, r_n)'$, выражается формулой

$$(12) \quad \log L = -\frac{1}{2} n \log 2\pi - \frac{1}{2} \log |\mathbf{W}| - \\ - \frac{1}{2} T (\mathbf{r}^{(1)'} - \mathbf{p}' \mathbf{r}^{(2)'}) \begin{pmatrix} \mathbf{W}^{11} & \mathbf{W}^{12} \\ \mathbf{W}^{21} & \mathbf{W}^{22} \end{pmatrix} \begin{pmatrix} \mathbf{r}^{(1)} - \mathbf{p} \\ \mathbf{r}^{(2)} \end{pmatrix},$$

где

$$(13) \quad \mathbf{r} = \begin{pmatrix} \mathbf{r}^{(1)} \\ \mathbf{r}^{(2)} \end{pmatrix}, \quad \mathbf{W}^{-1} = \begin{pmatrix} \mathbf{W}^{11} & \mathbf{W}^{12} \\ \mathbf{W}^{21} & \mathbf{W}^{22} \end{pmatrix}$$

разбиты на блоки с q и $n-q$ строками и столбцами. Вектор частных производных равен

$$(14) \quad \frac{\partial \log L}{\partial \mathbf{p}} = -\frac{1}{2} \frac{1}{|\mathbf{W}|} \frac{\partial |\mathbf{W}|}{\partial \mathbf{p}} - \frac{1}{2} T (\mathbf{r}^{(1)'} - \mathbf{p}' \mathbf{r}^{(2)'}) \times \\ \times \frac{\partial}{\partial \mathbf{p}} \begin{pmatrix} \mathbf{W}^{11} & \mathbf{W}^{12} \\ \mathbf{W}^{21} & \mathbf{W}^{22} \end{pmatrix} \begin{pmatrix} \mathbf{r}^{(1)} - \mathbf{p} \\ \mathbf{r}^{(2)} \end{pmatrix} + T \mathbf{W}^{11} (\mathbf{r}^{(1)} - \mathbf{p}) + T \mathbf{W}^{12} \mathbf{r}^{(2)}.$$

Приравняем этот вектор нулевому. Поскольку $\sqrt{T}(\mathbf{r}^{(1)} - \boldsymbol{\rho})$ и $\sqrt{T}\mathbf{r}^{(2)}$ имеют в пределе нормальное распределение, нормируем полученное уравнение делением обеих частей на $\sqrt{T}$. Первые два члена в результирующем уравнении будут иметь порядок $1/\sqrt{T}$ и сходиться по вероятности к нулю. Учитывая это, приходим к следующему соотношению для оценки вектора $\boldsymbol{\rho}$:

$$(15) \quad \hat{\boldsymbol{\rho}} = \mathbf{r}^{(1)} + (\mathbf{W}_{11})^{-1} \mathbf{W}_{12} \mathbf{r}^{(2)} = \mathbf{r}^{(1)} - \mathbf{W}_{12} \mathbf{W}_{22}^{-1} \mathbf{r}^{(2)} = \\ = \mathbf{r}^{(1)} - \mathbf{W}_{12}(\boldsymbol{\rho}) \mathbf{W}_{22}^{-1}(\boldsymbol{\rho}) \mathbf{r}^{(2)}.$$

(См. упр. 8 гл. 2.) Процедура оценивания состоит в том, что $\mathbf{r}^{(1)}$ используется в качестве состоятельной оценки для $\boldsymbol{\rho}$, затем вычисляется $\mathbf{W}_{12}(\mathbf{r}^{(1)}) \mathbf{W}_{22}^{-1}(\mathbf{r}^{(1)})$ и $\boldsymbol{\rho}$ оценивается по формуле

$$(16) \quad \hat{\boldsymbol{\rho}} = \mathbf{r}^{(1)} - \mathbf{W}_{12}(\mathbf{r}^{(1)}) \mathbf{W}_{22}^{-1}(\mathbf{r}^{(1)}) \mathbf{r}^{(2)}.$$

Статистика $\sqrt{T}(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho})$, вычисленная согласно (16), будет иметь то же самое предельное распределение, что и статистика $\sqrt{T}(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho})$, вычисленная в соответствии с (15), а именно нормальное распределение с нулевым средним и ковариационной матрицей

$$(17) \quad \mathbf{W}_{11} - \mathbf{W}_{12} \mathbf{W}_{22}^{-1} \mathbf{W}_{21}.$$

Указанную процедуру можно выполнять с помощью итерационного процесса. Пусть $\hat{\boldsymbol{\rho}}^{(0)}$ — начальная состоятельная оценка. Тогда последовательные итерации имеют вид

$$(18) \quad \hat{\boldsymbol{\rho}}^{(j)} = \mathbf{r}^{(1)} - \mathbf{W}_{12}(\hat{\boldsymbol{\rho}}^{(j-1)}) \mathbf{W}_{22}^{-1}(\hat{\boldsymbol{\rho}}^{(j-1)}) \mathbf{r}^{(2)}.$$

Уолкер приводит соответствующий пример для ряда из 100 наблюдений, полученного Дурбином (1959) с помощью модели, в которой $q = 1$ и $\alpha_1 = 1/2$. Первые пять выборочных корреляций равны 0.35005, -0.06174 , -0.08007 , -0.14116 и -0.15629 . Матрица $\mathbf{W}$, полученная из матрицы ковариаций первых пяти корреляций, равна

$$(19) \quad \begin{bmatrix} 1 - 3\rho^2 + 4\rho^4 & 2\rho(1 - \rho^2) & \rho^2 & 0 & 0 \\ 2\rho(1 - \rho^2) & 1 + 2\rho^2 & 2\rho & \rho^2 & 0 \\ \rho^2 & 2\rho & 1 + 2\rho^2 & 2\rho & \rho^2 \\ 0 & \rho^2 & 2\rho & 1 + 2\rho^2 & 2\rho \\ 0 & 0 & \rho^2 & 2\rho & 1 + 2\rho^2 \end{bmatrix},$$

где $\rho = \rho_1$. Для $n = 2$ и $q = 1$ формула (15) принимает вид

$$(20) \quad \hat{\rho} = r_1 - \frac{2\rho(1-\rho^2)}{1+\rho^2} r_2.$$

В качестве нулевого приближения берем $\hat{\rho}(0) = r_1$. Тогда $\hat{\rho}^{(1)} = 0.38051$ и $\hat{\rho}^{(2)} = 0.38121$ ($\rho = 0.4$). Оценка стандартного отклонения асимптотического распределения равна

$$(21) \quad \frac{1}{\sqrt{T}} \sqrt{w_{11}(\hat{\rho}) - w_{12}^2(\hat{\rho})/w_{22}(\hat{\rho})} = 0.0565.$$

Следует заметить, что (17) — монотонная функция от n в том смысле, что для каждого x квадратичная форма $x'(\mathbf{W}_{11} - \mathbf{W}_{12}\mathbf{W}_{22}^{-1}\mathbf{W}_{21})x$ не убывает с ростом n . Точнее говоря, выражение $x'[\mathbf{W}_{11}^{(n)} - \mathbf{W}_{12}^{(n)}(\mathbf{W}_{22}^{(n)})^{-1}\mathbf{W}_{21}^{(n)}]x$ является дисперсией предельного нормального распределения статистики $\sqrt{T}x'(\hat{\rho} - \rho)$ при использовании $r_1, \dots, r_n$. Эта асимптотическая дисперсия не убывает по n для каждого x . При этом предполагается, что n выбирается достаточно большим (T берется достаточно большим для того, чтобы соответствующие распределения хорошо аппроксимировались нормальным).

Оценки $\hat{\rho}$ и c_0 величин ρ и $\sigma(0)$ можно использовать для вычисления оценок параметров $\alpha_1, \dots, \alpha_q$ и σ^2 . Тогда вектор $\sqrt{T}(\hat{\alpha}_1 - \alpha_1), \dots, \sqrt{T}(\hat{\alpha}_q - \alpha_q)$ будет иметь предельное нормальное распределение.

Иной подход был использован Дурбином (1959). Он предложил бесконечную сумму (10) аппроксимировать конечной суммой

$$(22) \quad v_t^* = \sum_{r=0}^n \gamma_r y_{t-r},$$

где n достаточно велико. (Заметим, что $\gamma_0 = 1$.) Последовательность $\{v_t^*\}$ образует процесс конечного скользящего среднего. (Представление (22) фактически соответствует (11) из § 5.2 с надлежащими изменениями в обозначениях. Так, u_r заменяются на y_r , затем y_r — на v_r , а получающаяся в результате линейная комбинация переменных $v_{t-n-1}, \dots, v_{t-n-q}$ переносится в левую часть.) Статистика v_t^* состоит из v_t и линейной комбинации статистик $v_{t-n-1}, \dots, v_{t-n-q}$. Коэффициенты последней будут малы, если n велико. Они являются линейными комбинациями величин $\gamma_{n+1}, \dots, \gamma_{n+q}$, определяемых соотношениями (9) или (10). Последние представляют собой линейные комбинации степеней ($(n+1)$ -й по $(n+q)$ -ю) корней уравнения $M(z) = 0$, о которых предполагается, что все они лежат в еди-

ничном круге. [См. (14) или (22) и (26) § 5.2.] Отсюда следует, что при больших n случайные величины v_i будут близки к некоррелированным, а (22), грубо говоря, будет стохастическим разностным уравнением. Это наводит на мысль о том, что вектор $\gamma = (\gamma_1, \dots, \gamma_n)'$ следует оценивать исходя из уравнения

$$(23) \quad \begin{bmatrix} c_0 & c_1 & \dots & c_{n-1} \\ c_1 & c_0 & \dots & c_{n-2} \\ \vdots & \vdots & & \vdots \\ c_{n-1} & c_{n-2} & \dots & c_0 \end{bmatrix} \begin{bmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \\ \vdots \\ \hat{\gamma}_n \end{bmatrix} = - \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Последнее можно записать в эквивалентной форме с помощью корреляций $r_0 = 1, r_1, \dots, r_n$. Теорема 5.7.1 утверждает, что вектор $\sqrt{T} (r_1 - \rho_1), \dots, \sqrt{T} (r_n - \rho_n)$ имеет в пределе нормальное распределение и $\text{plim}_{T \rightarrow \infty} r_i = \rho_i, i = 1, \dots, n$. Поэтому нормальное рас-

пределение имеет в пределе и вектор $\sqrt{T} (\hat{\gamma} - \gamma^*)$, где $\gamma^* = -\Sigma^{-1}\sigma$, $\sigma = [\sigma(1), \dots, \sigma(n)]'$ и

$$(24) \quad \Sigma = \begin{bmatrix} \sigma(0) & \sigma(1) & \dots & \sigma(n-1) \\ \sigma(1) & \sigma(0) & \dots & \sigma(n-2) \\ \vdots & \vdots & & \vdots \\ \sigma(n-1) & \sigma(n-2) & \dots & \sigma(0) \end{bmatrix}.$$

Величины $\sigma(h)$ определены в (2). Вектор $\gamma^* (= -\Sigma^{-1}\sigma)$ является аппроксимирующим для γ (компоненты $\gamma_1, \dots, \gamma_n$ последнего определены в (9)). Отметим, что $\text{plim}_{T \rightarrow \infty} \hat{\gamma} = \gamma^*$. Сходство соотношения (22)

со стохастическим разностным уравнением приводит к предположению о том, что указанное предельное распределение имеет ковариационную матрицу, близкую к матрице, получающейся умножением матрицы, обратной к Σ , на σ^2 . Но тогда аппроксимирующая плотность должна иметь в показателе экспоненты квадратичную форму, равную умноженной на $-T/(2\sigma^2)$ форме

$$(25) \quad (\hat{\gamma} - \gamma^*)' \Sigma (\hat{\gamma} - \gamma^*) = \hat{\gamma}' \Sigma \hat{\gamma} - 2\hat{\gamma}' \Sigma \gamma^* + \gamma^{*'} \Sigma \gamma^* = \\ = \hat{\gamma}' \Sigma \hat{\gamma} + 2\hat{\gamma}' \sigma - \gamma^{*'} \sigma = \\ = (1 \ \hat{\gamma}') \begin{pmatrix} \sigma^*(0) - \sigma^2 & \sigma' \\ \sigma & \Sigma \end{pmatrix} \begin{pmatrix} 1 \\ \hat{\gamma} \end{pmatrix},$$

где $\sigma^*(0) = \sigma^2 - \gamma^{*\prime} \sigma = \sigma^2 - \sum_{r=1}^n \gamma_r^* \sigma(r)$. Если Σ и σ соответствуют процессу авторегрессии порядка n , то $\sigma^*(0)$ будет дисперсией этого процесса. [В § 5.2 соотношение (47) можно записать в виде $\sigma_a^2 = \sum_{r=0}^n \beta_{ar} \sigma_a(r)$, а (48) — в виде $\beta_a = -\Sigma_a^{-1} \sigma_a$, где индекс a указывает на процесс авторегрессии, изучаемый в настоящем разделе, и замену p на n .] Заменим в (25) $\sigma^*(0)$ на $\sigma(0)$. Тогда квадратичная форма в показателе экспоненты аппроксимирующей плотности будет равна умноженной на $-(1/2) T$ форме

$$\begin{aligned}
 (26) \quad & (\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \dots, \hat{\gamma}_n) \times \\
 & \times \begin{bmatrix} \sum_{j=1}^q \alpha_j^2 & \alpha_1 + \sum_{j=1}^{q-1} \alpha_j \alpha_{j+1} & \alpha_2 + \sum_{j=1}^{q-2} \alpha_j \alpha_{j+2} \dots & 0 \\ \alpha_1 + \sum_{j=1}^{q-1} \alpha_j \alpha_{j+1} & 1 + \sum_{j=1}^q \alpha_j^2 & \alpha_1 + \sum_{j=1}^{q-1} \alpha_j \alpha_{j+1} \dots & 0 \\ \alpha_2 + \sum_{j=1}^{q-2} \alpha_j \alpha_{j+2} & \alpha_1 + \sum_{j=1}^{q-1} \alpha_j \alpha_{j+1} & 1 + \sum_{j=1}^q \alpha_j^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 + \sum_{j=1}^q \alpha_j^2 \end{bmatrix} \times \\
 & \times \begin{bmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_1 \\ \hat{\gamma}_2 \\ \vdots \\ \hat{\gamma}_n \end{bmatrix} = \\
 & = \sum_{j=1}^q \alpha_j^2 + \left(1 + \sum_{j=1}^q \alpha_j^2\right) \sum_{u=1}^n \hat{\gamma}_u^2 + \\
 & + 2 \left(\alpha_1 + \sum_{j=1}^{q-1} \alpha_j \alpha_{j+1} \right) \sum_{u=0}^{n-1} \hat{\gamma}_u \hat{\gamma}_{u+1} + \\
 & + 2 \left(\alpha_2 + \sum_{j=1}^{q-2} \alpha_j \alpha_{j+2} \right) \sum_{u=0}^{n-2} \hat{\gamma}_u \hat{\gamma}_{u+2} + \dots + 2 \alpha_q \sum_{u=0}^{n-q} \hat{\gamma}_u \hat{\gamma}_{u+q},
 \end{aligned}$$

если $\hat{\gamma}_0 = 1$. Приравнявая нулю производные последней по α_i , $i = 1, \dots, q$, получаем

(27)

$$\begin{bmatrix} \sum_{u=0}^n \hat{\gamma}_u^2 & \sum_{u=0}^{n-1} \hat{\gamma}_u \hat{\gamma}_{u+1} & \sum_{u=0}^{n-2} \hat{\gamma}_u \hat{\gamma}_{u+2} & \dots & \sum_{u=0}^{n-q+1} \hat{\gamma}_u \hat{\gamma}_{u+q-1} \\ \sum_{u=0}^{n-1} \hat{\gamma}_u \hat{\gamma}_{u+1} & \sum_{u=0}^n \hat{\gamma}_u^2 & \sum_{u=0}^{n-1} \hat{\gamma}_u \hat{\gamma}_{u+1} & \dots & \sum_{u=0}^{n-q+2} \hat{\gamma}_u \hat{\gamma}_{u+q-2} \\ \sum_{u=0}^{n-2} \hat{\gamma}_u \hat{\gamma}_{u+2} & \sum_{u=0}^{n-1} \hat{\gamma}_u \hat{\gamma}_{u+1} & \sum_{u=0}^n \hat{\gamma}_u^2 & \dots & \sum_{u=0}^{n-q+3} \hat{\gamma}_u \hat{\gamma}_{u+q-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{u=0}^{n-q+1} \hat{\gamma}_u \hat{\gamma}_{u+q-1} & \sum_{u=0}^{n-q+2} \hat{\gamma}_u \hat{\gamma}_{u+q-2} & \sum_{u=0}^{n-q+3} \hat{\gamma}_u \hat{\gamma}_{u+q-3} & \dots & \sum_{u=0}^n \hat{\gamma}_u^2 \end{bmatrix} \times$$

$$\times \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\alpha}_3 \\ \vdots \\ \hat{\alpha}_q \end{bmatrix} = - \begin{bmatrix} \sum_{u=0}^{n-1} \hat{\gamma}_u \hat{\gamma}_{u+1} \\ \sum_{u=0}^{n-2} \hat{\gamma}_u \hat{\gamma}_{u+2} \\ \sum_{u=0}^{n-3} \hat{\gamma}_u \hat{\gamma}_{u+3} \\ \vdots \\ \sum_{u=0}^{n-q} \hat{\gamma}_u \hat{\gamma}_{u+q} \end{bmatrix}.$$

Эти уравнения совпадают с теми, которые можно было бы получить для q коэффициентов стохастического разностного уравнения, если бы наблюдения были соответственно $1, \hat{\gamma}_1, \dots, \hat{\gamma}_n$. Решение этих уравнений $\hat{\alpha} = (\hat{\alpha}_1, \dots, \hat{\alpha}_q)'$ является состоятельной оценкой вектора α^* , который в свою очередь является решением уравнения (27), если в последнем заменить $\hat{\gamma}$ на γ^* . Вектор α^* приближенно равен вектору α .

При $T \rightarrow \infty$ и фиксированном n вектор $\sqrt{T}(\hat{\alpha}_1 - \alpha_1^*), \dots, \sqrt{T}(\hat{\alpha}_q - \alpha_q^*)$ имеет в пределе нормальное распределение. Значения $\hat{\alpha}_1^*, \dots, \hat{\alpha}_q^*$, к которым сходятся по вероятности $\hat{\alpha}_1, \dots, \hat{\alpha}_q$, могут отличаться от $\alpha_1, \dots, \alpha_q$, но, по-видимому, это отличие будет незначительным, если n велико.

5.7.3. Центральная предельная теорема для процессов скользящего среднего

Для обоснования асимптотической нормальности оценок нам необходимо доказать, что выборочные корреляции имеют в пределе нормальное распределение. Хотя вторые моменты выборочных ковариаций зависят от четвертых моментов самого процесса, тем не менее оказывается, что вторые моменты асимптотических распределений коэффициентов корреляции зависят только от вторых его моментов. При этом факт асимптотической нормальности определяется только тем, будет ли конечной дисперсия.

ТЕОРЕМА 5.7.1. Пусть y_t , $t = 1, 2, \dots$, определяется соотношением (1), где v_t — независимые одинаково распределенные случайные величины с $\mathbb{E}v_t = 0$ и $\mathbb{E}v_t^2 = \sigma^2$. Тогда вектор $\sqrt{T}(r_1 - \rho_1), \dots, \sqrt{T}(r_n - \rho_n)$ имеет в пределе при $T \rightarrow \infty$ нормальное распределение с нулевым средним и ковариациями

$$(28) \quad w_{hh'} = \sum_{g=-(q+n)}^{q+n} (\rho_{g-h}\rho_{g-h'} + \rho_{g+h}\rho_{g+h'} - 2\rho_h\rho_g\rho_{g+h'} - \\ - 2\rho_h\rho_g\rho_{g+h} + 2\rho_h\rho_h\rho_g^2).$$

ДОКАЗАТЕЛЬСТВО. Пусть $0 < h \leq q$. Рассмотрим следующее выражение:

$$(29) \quad \sqrt{T}(c_h - \rho_h c_0) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} y_t y_{t+h} - \rho_h \frac{1}{\sqrt{T}} \sum_{t=1}^T y_t^2 = \\ = \frac{1}{\sqrt{T}} \left[\sum_{t=1}^{T-h} \sum_{i,j=0}^q \alpha_i \alpha_j v_{t-i} v_{t+h-j} - \right. \\ \left. - \rho_h \sum_{t=1}^T \sum_{i,j=0}^q \alpha_i \alpha_j v_{t-i} v_{t-j} \right] = \\ = \frac{1}{\sqrt{T}} \left[\sum_{t=1}^{T-h} \left(\sum_{i=0}^{q-h} \alpha_i \alpha_{i+h} v_{t-i}^2 + \sum_{\substack{i,j=0 \\ i \neq j-h}}^q \alpha_i \alpha_j v_{t-i} v_{t+h-j} \right) - \right. \\ \left. - \rho_h \sum_{t=1}^T \left(\sum_{i=0}^q \alpha_i^2 v_{t-i}^2 + \sum_{\substack{i,j=0 \\ i \neq j}}^q \alpha_i \alpha_j v_{t-i} v_{t-j} \right) \right].$$

Первая сумма в правой части равна

$$(30) \quad \frac{1}{\sqrt{T}} \sum_{i=0}^{q-h} \alpha_i \alpha_{i+h} \sum_{t=1}^{T-q} v_t^2$$

с точностью до самое большое $2q^2$ слагаемых. Подобным же образом и третья сумма в правой части (29) равна (30) с точностью до самое

большее $2q^3$ слагаемых, так как $\rho_h = \sum_{i=0}^{q-h} \alpha_i \alpha_{i+h} / \sum_{i=0}^q \alpha_i^2$. Поэтому предельное распределение величины $\sqrt{T} (c_h - \rho_h c_0)$ совпадает с предельным распределением статистики

$$(31) \quad \frac{1}{\sqrt{T}} \left[\sum_{t=1}^{T-h} \sum_{\substack{i,j=0 \\ i \neq j-h}}^q \alpha_i \alpha_j v_{t-i} v_{t+h-j} - \rho_h \sum_{t=1}^T \sum_{\substack{i,j=0 \\ i \neq j}}^q \alpha_i \alpha_j v_{t-i} v_{t-j} \right] = \\ = \frac{1}{\sqrt{T}} \left[\sum_{t=1}^{T-h} \sum_{i=0}^q \sum_{\substack{k=-h \\ k \neq i}}^{q-h} \alpha_i \alpha_{k+h} v_{t-i} v_{t-k} - \rho_h \sum_{t=1}^T \sum_{i=0}^q \sum_{\substack{k=0 \\ k \neq i}}^q \alpha_i \alpha_k v_{t-i} v_{t-k} \right].$$

За исключением, быть может, $4q^3$ слагаемых, (31) совпадает с

$$(32) \quad \frac{1}{\sqrt{T}} \left[\sum_{s=1}^{T-q} \sum_{i=0}^q \sum_{\substack{k=-h \\ k \neq i}}^{q-h} \alpha_i \alpha_{k+h} v_s v_{s+i-k} - \rho_h \sum_{s=1}^{T-q} \sum_{i=0}^q \sum_{\substack{k=0 \\ k \neq i}}^q \alpha_i \alpha_k v_s v_{s+i-k} \right] = \\ = \frac{1}{\sqrt{T}} \sum_{s=1}^{T-q} \sum_{i=0}^q \sum_{\substack{k=-h \\ k \neq i}}^q (\alpha'_i \alpha'_{k+h} - \rho_h \alpha'_i \alpha'_k) v_s v_{s+i-k},$$

где $\alpha'_i = 0$ для $i < 0$ и $i > q$, и $\alpha'_i = \alpha_i$ для $i = 0, 1, \dots, q$. Далее, (32) есть

$$(33) \quad \frac{1}{\sqrt{T}} \sum_{s=1}^{T-q} \sum_{\substack{g=-q \\ g \neq 0}}^{q+h} \sum_{i=0}^q (\alpha'_i \alpha'_{i+h-g} - \rho_h \alpha'_i \alpha'_{i-g}) v_s v_{s+g} = \\ = \frac{1}{\sqrt{T}} \sum_{s=1}^{T-q} \left[\sum_{g=1}^{q+h} \sum_{i=0}^q (\alpha'_i \alpha'_{i+h-g} - \rho_h \alpha'_i \alpha'_{i-g}) v_s v_{s+g} + \right. \\ \left. + \sum_{f=1}^q \sum_{i=0}^q (\alpha'_i \alpha'_{i+h+f} - \rho_h \alpha'_i \alpha'_{i+f}) v_s v_{s-f} \right].$$

Последнее отличается не более чем $2(q+1)^3$ слагаемыми от

$$(34) \quad \frac{1}{\sqrt{T}} \sum_{s=1}^{T-q} \sum_{g=1}^{q+h} \sum_{i=0}^q (\alpha'_i \alpha'_{i+h-g} + \alpha'_i \alpha'_{i+h+g} - 2\rho_h \alpha'_i \alpha'_{i-g}) v_s v_{s+g},$$

имеющего вид $\sum_{s=1}^{T-q} \left(\sum_{g=1}^n k_g v_s v_{s+g} \right)$. Слагаемые $\sum_{g=1}^n k_g v_s v_{s+g}$ имеют нулевые средние, дисперсии

$$(35) \quad \mathbb{E} \sum_{g,g'=1}^n k_g k_{g'} v_s v_{s+g} v_s v_{s+g'} = \sigma^4 \sum_{g=1}^n k_g^2$$

и нулевые корреляции. Последовательность их образует $(q + h)$ -зависимый стационарный случайный процесс. Асимптотическая нормальность (29) вытекает теперь ¹⁾ из теоремы 7.7.5.

Для $q < h \leq n$ имеем $\rho_h = 0$ и

$$(36) \quad \sqrt{T}c_h = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} y_t y_{t+h} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-h} \sum_{i,j=0}^q \alpha_i \alpha_j v_{t-i} v_{t+h-j},$$

а последнее отличается не более чем $2q(q+1)^2$ слагаемыми от

$$(37) \quad \frac{1}{\sqrt{T}} \sum_{s=1}^{T-h} \sum_{i,j=0}^q \alpha_i \alpha_j v_s v_{s+h+i-j} = \\ = \frac{1}{\sqrt{T}} \sum_{s=1}^{T-h} \sum_{g=-q+h}^{q+h} \sum_{i=0}^q \alpha'_i \alpha'_{i+h-g} v_s v_{s+g}.$$

Использование теоремы 7.7.5 приводит к тому, что и (37) имеет в пределе нормальное распределение. Отметим, что коэффициент при $v_i v_{t+i} + v_i v_{t-i}$ в (37) для $j \neq 0$ совпадает с коэффициентом при $v_i v_{t+j}$ в (34) по той причине, что $\alpha'_j = 0$ для $j > q$ и $\rho_h = 0$.

Поскольку

$$(38) \quad \sum_{i=0}^q \alpha'_i \alpha'_{i+k} = \left(\sum_{j=0}^q \alpha_j^2 \right) \rho_k,$$

ковариация между (34) и величиной, получающейся из (34) при замене h на h' , $0 < h \leq h'$, равна

$$(39) \quad \sigma^4 \left(\sum_{i=0}^q \alpha_i^2 \right)^2 \sum_{g=1}^{q+h'} (\rho_{h-g} + \rho_{g+h} - 2\rho_h \rho_g) (\rho_{h'-g} + \rho_{g+h'} - 2\rho_{h'} \rho_g) = \\ = \frac{1}{2} \sigma^4 \left(\sum_{i=0}^q \alpha_i^2 \right)^2 \sum_{g=-(q+h')}^{q+h'} (\rho_{h-g} + \rho_{g+h} - 2\rho_h \rho_g) (\rho_{h'-g} + \rho_{g+h'} - \\ - 2\rho_{h'} \rho_g) = \sigma^2(0) w_{hh'},$$

где $w_{hh'}$ определено в (28). [Если $\mathcal{E} u_t^4 < \infty$, то (39) есть $\lim_{T \rightarrow \infty} T \mathcal{E} (c_h - \rho_h c_0) (c_{h'} - \rho_{h'} c_0)$.] Имеем, далее,

$$(40) \quad c_0 = \frac{1}{T} \sum_{t=1}^T y_t^2 = \sum_{i,j=0}^q \alpha_i \alpha_j \frac{1}{T} \sum_{t=1}^T v_{t-i} v_{t-j} = \\ = \sum_{i=0}^q \alpha_i^2 \frac{1}{T} \sum_{t=1}^T v_{t-i}^2 + \sum_{\substack{i,j=0 \\ i \neq j}}^q \alpha_i \alpha_j \frac{1}{T} \sum_{t=1}^T v_{t-i} v_{t-j}.$$

¹⁾ Используемые здесь и далее результаты из гл. 7 не опираются на результаты первых шести глав.— *Прим. перев.*

Каждая сумма $(1/T) \sum_{i=1}^T v_{t-i}^2$ сходится по вероятности к σ^2 в силу закона больших чисел, а каждая сумма $(1/T) \sum_{i=1}^T v_{t-i} v_{t-j}$, $i \neq j$, сходится по вероятности к нулю, так как сходится к нулю при $i \neq j$

$$(41) \quad \mathbb{E} \left(\frac{1}{T} \sum_{i=1}^T v_{t-i} v_{t-j} \right)^2 = \frac{1}{T^2} \sum_{i,s=1}^T \mathbb{E} v_{t-i} v_{t-j} v_{t-s-i} v_{t-s-j} = \frac{1}{T} \sigma^4.$$

Таким образом, c_0 сходится по вероятности к $\sigma^2 \sum_{i=0}^q \alpha_i^2 = \sigma(0)$. Поэтому предельное распределение величины

$$(42) \quad \sqrt{T} (r_h - \rho_h) = \frac{\sqrt{T} (c_h - \rho_h c_0)}{c_0}$$

есть $N(0, \omega_{hh})$. Доказательство теоремы завершается аналогичным рассмотрением произвольной линейной комбинации $\sum_{h=1}^n k_h \sqrt{T} \times \times (r_h - \rho_h)$. ■

5.8. ПРОЦЕСС АВТОРЕГРЕССИИ С ОСТАТКАМИ В ВИДЕ СКОЛЬЗЯЩЕГО СРЕДНЕГО

5.8.1. Модель

Пусть $\{y_t\}$ удовлетворяет соотношению

$$(1) \quad \sum_{s=0}^p \beta_s y_{t-s} = \sum_{j=0}^q \alpha_j v_{t-j},$$

в котором $\beta_0 = \alpha_0 = 1$, а $\{v_t\}$ — последовательность независимых случайных величин с $\mathbb{E} v_t = 0$ и $\mathbb{E} v_t^2 = \sigma^2$. Чтобы исключить тривиальные случаи, будем полагать $\beta_p \neq 0$ и $\alpha_q \neq 0$. Тогда имеем $\mathbb{E} y_t = 0$. Соотношение (1) можно записать в виде

$$(2) \quad \sum_{s=0}^p \beta_s \mathcal{L}^s y_t = \sum_{j=0}^q \alpha_j \mathcal{L}^j v_t.$$

Пусть корни уравнения

$$(3) \quad \sum_{s=0}^p \beta_s x^{p-s} = 0$$

равны $x_1, \dots, x_p$, а корни уравнения

$$(4) \quad \sum_{j=0}^q \alpha_j z^{q-j} = 0$$

равны $z_1, \dots, z_q$. Тогда (2) можно переписать следующим образом:

$$(5) \quad \prod_{i=1}^p (1 - x_i \mathcal{L}) y_t = \prod_{j=1}^q (1 - z_j \mathcal{L}) v_t.$$

Соотношение между $\{y_t\}$ и $\{v_t\}$ не изменится, если обе части (5) разделить на общие множители. Поэтому без ограничения общности будем считать, что (3) и (4) не имеют общих корней.

Для того чтобы y_t можно было выразить в виде линейной комбинации $v_t, v_{t-1}, \dots$ (когда $\sigma^2 > 0$), необходимо, чтобы все корни уравнения (3) лежали в единичном круге. При этом

$$(6) \quad \begin{aligned} y_t &= \prod_{i=1}^p (1 - x_i \mathcal{L})^{-1} \prod_{j=1}^q (1 - z_j \mathcal{L}) v_t = \\ &= \sum_{s=0}^{\infty} \delta_s \sum_{j=0}^q \alpha_j v_{t-s-j} = \sum_{r=0}^{\infty} \gamma_r v_{t-r}. \end{aligned}$$

Коэффициенты $\{\delta_s\}$ здесь те же, что и в разд. 5.2.1. Коэффициенты $\{\gamma_r\}$ получаются следующим образом:

$$(7) \quad \begin{aligned} \gamma_0 &= \delta_0 \alpha_0 = 1, \\ \gamma_1 &= \delta_0 \alpha_1 + \delta_1 \alpha_0 = \alpha_1 + \delta_1, \\ &\vdots \\ \gamma_{q-1} &= \delta_0 \alpha_{q-1} + \delta_1 \alpha_{q-2} + \dots + \delta_{q-1} \alpha_0, \\ \gamma_r &= \delta_{r-q} \alpha_q + \delta_{r-q+1} \alpha_{q-1} + \dots + \delta_r \alpha_0, \quad r = q, q+1, \dots \end{aligned}$$

Умножим обе части (1) на y_{t-h} , $h = 0, 1, \dots$, и возьмем от полученных выражений математические ожидания. В результате придем к системе соотношений

$$(8) \quad \sum_{s=0}^p \beta_s \sigma(h-s) = \sum_{j=0}^q \sum_{r=0}^{\infty} \alpha_j \gamma_r \mathcal{E} v_{t-j} v_{t-h-r} = \sigma^2 \sum_{j=0}^q \sum_{r=0}^{\infty} \alpha_j \gamma_r \delta_{j,r+h},$$

$$h = 0, 1, \dots,$$

где $\delta_{j,r+h}$ есть дельта Кронекера. В частности, отсюда следует, что

$$(9) \quad \sum_{s=0}^p \beta_s \sigma(h-s) = 0, \quad h = q+1, \dots$$

Таким образом, последовательность $\{\sigma(r)\}$ удовлетворяет для $r = q+1-p, \dots$ однородному разностному уравнению. [В разд. 5.8.2. будет показано, что эти уравнения в совокупности с условием $\beta_0 = 1$ определяют $\beta_1, \dots, \beta_p$, если последовательность $\{\sigma(r)\}$ известна.]

Для некоторых целей более удобным оказывается рассматривать в качестве параметров процесса не $\beta_1, \dots, \beta_p, \sigma^2, \alpha_1, \dots, \alpha_q$, а $\beta_1, \dots, \beta_p, \sigma(0), \rho_1, \dots, \rho_q$, где $\rho_h = \sigma(h)/\sigma(0)$. Если обозначить левую часть выражения (1) через u_t , то математическое ожидание произведения $u_t u_{t+h}$ можно записать в виде

$$(10) \quad \mathbb{E} u_t u_{t+h} = \sum_{s,r=0}^p \beta_s \beta_r \mathbb{E} y_{t-s} y_{t+h-r} = \sum_{s,r=0}^p \beta_s \beta_r \sigma(h+s-r).$$

С другой стороны, используя правую часть (1), то же математическое ожидание можно записать в виде

$$(11) \quad \mathbb{E} u_t u_{t+h} = \begin{cases} \sigma^2 \sum_{j=0}^{q-|h|} \alpha_j \alpha_{j+|h|}, & h = 0, \pm 1, \dots, \pm q, \\ 0, & h = \pm(q+1), \dots \end{cases}$$

Использование (10) и (11) приводит к следующей записи производящей функции ковариаций случайной последовательности $\{u_t\}$:

$$(12) \quad \sigma(0) \sum_{h=-q}^q \sum_{s,r=0}^p \beta_s \beta_r \rho_{h+s-r} z^h = \sigma^2 M(z) M(z^{-1}).$$

Здесь $M(z)$ определяется посредством (4) из § 5.7. Используя (9), соотношение (12) можно переписать в виде

$$(13) \quad \sigma(0) \sum_{s,r=0}^p \beta_s \beta_r \sum_{h=-(q-r)}^{q-s} \rho_{h+s-r} z^h = \sigma^2 M(z) M(z^{-1}).$$

В левой части последнего соотношения присутствуют лишь параметры $\beta_1, \dots, \beta_p, \sigma(0), \rho_1, \dots, \rho_q$, а в правой — лишь параметры $\sigma^2, \alpha_1, \dots, \alpha_q$. Поэтому вторая совокупность параметров вполне определяется первой. Левая часть (13), умноженная на z^q , будет полиномом степени $2q$. Корни ее будут образовывать пары, как было указано в разд. 5.7.1. Одно из образующихся при этом двух множеств корней определяет $M(z)$ и через него значения $\alpha_1, \dots, \alpha_q$. [См. Дуб (1944).]

5.8.2. Оценивание параметров

Если случайные величины v_t распределены нормально, то нормально распределенными будут и y_t . Как и в случае обычного скользящего среднего, изучавшегося в § 5.7, матрица, обратная ковариационной матрице конечного множества случайных величин $y_1, \dots, y_T$, является отнюдь не простой, минимальное достаточное множество статистик имеет мощность T , а уравнения максимального правдоподобия сложны и не поддаются непосредственному решению. Тем не менее (как указывается в теореме 8.4.6) любое конечное множество

выборочных корреляций имеет асимптотически нормальное распределение. А. Уолкер (1962) обобщил на рассматриваемый случай метод оценивания, используемый для модели обычного скользящего среднего, т. е. максимизацию функции, аппроксимирующей функцию правдоподобия первых n выборочных сериальных корреляций.

Если случайные величины v_t независимы и одинаково распределены с $\mathcal{E}v_t = 0$ и $\mathcal{E}v_t^2 = \sigma^2$, то совместное предельное распределение статистик $\sqrt{T}(r_1 - \rho_1), \dots, \sqrt{T}(r_n - \rho_n)$ будет многомерным нормальным с нулевыми средними и ковариациями

$$(14) \quad \omega_{gh} = \sum_{r=-\infty}^{\infty} (\rho_{r+g}\rho_{r+h} + \rho_{r-g}\rho_{r+h} - 2\rho_h\rho_r\rho_{r+g} - 2\rho_g\rho_r\rho_{r+h} + 2\rho_g\rho_h\rho_r^2).$$

По этому поводу см. теорему 8.4.6. [Если все корни (3) лежат в единичном круге, то условия теоремы 8.4.6 будут выполнены. См. (41) в § 5.2, а также (6) и (7).] Для усовершенствования процедуры оценивания, использующей данное предельное распределение, удобно использовать переменные

$$(15) \quad \begin{aligned} x_j &= r_j, & j &= 1, \dots, q, \\ x_j &= \sum_{s=0}^p \beta_s r_{j-s}, & j &= q+1, \dots, q+p, \\ x_j &= \sum_{s,t=0}^p \beta_s \beta_t r_{j-s-t}, & j &= q+p+1, \dots, n, \end{aligned}$$

где $r_0 = 1$ и $r_{-j} = r_j$, $j = 1, 2, \dots$. Поскольку из (9) следует

$$(16) \quad \sum_{s=0}^p \beta_s \rho_{h-s} = 0, \quad h = q+1, \dots,$$

то совместное распределение статистик $\sqrt{T}(x_1 - \rho_1), \dots, \sqrt{T}(x_q - \rho_q), \sqrt{T}x_{q+1}, \dots, \sqrt{T}x_n$ будет нормальным с нулевыми средними и ковариациями Φ_{ij} , $i, j = 1, \dots, n$, выражения для которых можно получить из (ω_{ij}) . Пусть $\rho = (\rho_1, \dots, \rho_q)'$, $\mathbf{x}^{(1)} = (x_1, \dots, x_q)'$, $\mathbf{x}^{(2)} = (x_{q+1}, \dots, x_{q+p})'$, $\mathbf{x}^{(3)} = (x_{q+p+1}, \dots, x_n)'$, $\mathbf{x} = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)})'$. Соответственно этому разобьем матрицу $\Phi = (\Phi_{ij})$ и её обратную на блоки, состоящие из q , p и $n - (q + p)$ строк и столбцов,

$$(17) \quad \Phi = \begin{pmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} \\ \Phi_{21} & \Phi_{22} & \Phi_{23} \\ \Phi_{31} & \Phi_{32} & \Phi_{33} \end{pmatrix}, \quad \Phi^{-1} = \begin{pmatrix} \Phi^{11} & \Phi^{12} & \Phi^{13} \\ \Phi^{21} & \Phi^{22} & \Phi^{23} \\ \Phi^{31} & \Phi^{32} & \Phi^{33} \end{pmatrix}.$$

Тогда логарифм функции, аппроксимирующей функцию правдоподобия от $\mathbf{x}$, будет равен

$$(18) \quad \log L = -\frac{1}{2} n \log 2\pi - \frac{1}{2} \log |\Phi| - \\ - \frac{1}{2} T(\mathbf{x}^{(1)'} - \rho' \mathbf{x}^{(2)'} \mathbf{x}^{(3)'}) \begin{pmatrix} \Phi^{11} & \Phi^{12} & \Phi^{13} \\ \Phi^{21} & \Phi^{22} & \Phi^{23} \\ \Phi^{31} & \Phi^{32} & \Phi^{33} \end{pmatrix} \begin{pmatrix} \mathbf{x}^{(1)} - \rho \\ \mathbf{x}^{(2)} \\ \mathbf{x}^{(3)} \end{pmatrix}.$$

В качестве оценок мы возьмем такие значения $\beta_1, \dots, \beta_p, \rho_1, \dots, \rho_q$, которые будут удовлетворять дифференциальным уравнениям $\partial \log L / \partial \beta_s = 0$, $s = 1, \dots, p$, и $\partial \log L / \partial \rho_h = 0$, $h = 1, \dots, q$. В этих дифференциальных уравнениях члены, включающие частные производные ϕ_{ij} (или ϕ^{ij}) по β_s и ρ_h , имеют меньший порядок по T , чем другие. Если пренебречь этими членами и умножить на соответствующие константы, то в результате получим уравнения

$$(19) \quad \Phi^{11}(\mathbf{x}^{(1)} - \rho) + \Phi^{12}\mathbf{x}^{(2)} + \Phi^{13}\mathbf{x}^{(3)} = 0,$$

$$(20) \quad \frac{\partial \mathbf{x}^{(2)'}}{\partial \beta} [\Phi^{21}(\mathbf{x}^{(1)} - \rho) + \Phi^{22}\mathbf{x}^{(2)} + \Phi^{23}\mathbf{x}^{(3)}] + \\ + \frac{\partial \mathbf{x}^{(3)'}}{\partial \beta} [\Phi^{31}(\mathbf{x}^{(1)} - \rho) + \Phi^{32}\mathbf{x}^{(2)} + \Phi^{33}\mathbf{x}^{(3)}] = 0,$$

где

$$(21) \quad \frac{\partial \mathbf{x}^{(2)'}}{\partial \beta} = \left(\frac{\partial x_j}{\partial \beta_s} \right), \quad j = q+1, \dots, q+p, \quad s = 1, \dots, p,$$

$$(22) \quad \frac{\partial \mathbf{x}^{(3)'}}{\partial \beta} = \left(\frac{\partial x_j}{\partial \beta_s} \right), \quad j = q+p+1, \dots, n, \quad s = 1, \dots, p.$$

Для того чтобы упростить (20), заметим, что

$$(23) \quad \frac{\partial x_j}{\partial \beta_s} = r_{j-s}, \quad j = q+1, \dots, q+p, \quad s = 1, \dots, p,$$

$$(24) \quad \frac{\partial x_j}{\partial \beta_s} = 2 \sum_{t=0}^p \beta_t r_{j-s-t}, \quad j = q+p+1, \dots, n, \quad s = 1, \dots, p.$$

Поскольку статистики $\sqrt{T}(r_{j-s} - \rho_{j-s})$, $j = q+1, \dots, q+p$, $s = 1, \dots, p$, имеют предельные распределения, то $\partial x_j / \partial \beta_s$ в (20) можно для этих значений j и s заменить на ρ_{j-s} . Ввиду того что в силу (16) предельными распределениями обладают статистики

$\sqrt{T} \sum_{t=0}^p \beta_t r_{j-s-t}$ при $j = q+p+1, \dots, n$, $s = 1, \dots, p$, то для таких значений j и s производные $\partial x_j / \partial \beta_s$ можно заменить в (20) нулями. В результате получаем, что (20) асимптотически эквива-

лентно

$$(25) \quad \mathbf{R} [\Phi^{21} (\mathbf{x}^{(1)} - \boldsymbol{\rho}) + \Phi^{22} \mathbf{x}^{(2)} + \Phi^{23} \mathbf{x}^{(3)}] = \mathbf{0},$$

где

$$(26) \quad \mathbf{R} = (\rho_{j-s}) = \begin{pmatrix} \rho_q & \rho_{q+1} & \dots & \rho_{q+p-1} \\ \rho_{q-1} & \rho_q & \dots & \rho_{q+p-2} \\ \vdots & \vdots & & \vdots \\ \rho_{q-p+1} & \rho_{q-p+2} & \dots & \rho_q \end{pmatrix}.$$

Докажем, что матрица $\mathbf{R}$ невырождена. Пусть имеет место обратное. Тогда существует некоторый постоянный вектор $\mathbf{c} = (c_1, \dots, c_p)'$, такой, что

$$(27) \quad \sum_{s=1}^p c_s \rho_{j-s} = 0, \quad j = q+1, \dots, q+p.$$

Из (16) вытекает, что $\rho_h = -\sum_{t=1}^p \beta_t \rho_{h-t}$, $h = q+1, \dots$. Применение этого соотношения к (27) показывает, что (27) выполняется тогда и для $j = q+p+1, \dots$. Но это означало бы, что процесс можно представить моделью (1), в которой параметры $\beta_0, \dots, \beta_p$ заменены параметрами $c_1, \dots, c_p$, а значения параметров $\alpha_1, \dots, \alpha_q$ определяются через $c_1, \dots, c_p$ и $\rho_1, \dots, \rho_q$ в соответствии с (13). Но это противоречит предположению о том, что $\beta_p \neq 0$, в соответствии с которым левая часть (1) определяется не p , а $p+1$ постоянными.

Поскольку $\mathbf{R}$ невырожденная матрица, то (19) и (25) можно объединить в уравнение

$$(28) \quad \begin{pmatrix} \Phi^{11} & \Phi^{12} & \Phi^{13} \\ \Phi^{21} & \Phi^{22} & \Phi^{23} \end{pmatrix} \begin{pmatrix} \mathbf{x}^{(1)} - \boldsymbol{\rho} \\ \mathbf{x}^{(2)} \\ \mathbf{x}^{(3)} \end{pmatrix} = \mathbf{0}.$$

При этом

$$(29) \quad \begin{pmatrix} \mathbf{x}^{(1)} - \boldsymbol{\rho} \\ \mathbf{x}^{(2)} \end{pmatrix} = - \begin{pmatrix} \Phi^{11} & \Phi^{12} \\ \Phi^{21} & \Phi^{22} \end{pmatrix}^{-1} \begin{pmatrix} \Phi^{13} \\ \Phi^{23} \end{pmatrix} \mathbf{x}^{(3)} = \begin{pmatrix} \Phi_{13} \\ \Phi_{23} \end{pmatrix} \Phi_{33}^{-1} \mathbf{x}^{(3)}.$$

(См., например, упр. 8 гл. 2.) Правая часть (29) представляет регрессию векторов $\mathbf{x}^{(1)} - \boldsymbol{\rho}$ и $\mathbf{x}^{(2)}$ на $\mathbf{x}^{(3)}$, использующую ковариационную матрицу предельного распределения. Тогда

$$(30) \quad \hat{\boldsymbol{\rho}} = \mathbf{x}^{(1)} - \Phi_{13} \Phi_{33}^{-1} \mathbf{x}^{(3)}$$

состоит из (асимптотических) остатков корреляций $r_1, \dots, r_q$ на

$\sum_{s,t=0}^p \beta_s \beta_t r_{j-s-t}$, $j = q+p+1, \dots, n$. Остальные уравнения в (29)

говорят о том, что (асимптотические) остатки сумм $\sum_{s=0}^p \beta_s r_{j-s}$, $j = q + 1, \dots, q + p$ на $\sum_{s,t=0}^p \beta_s \beta_t r_{j-s-t}$, $j = q + p + 1, \dots, n$, равны нулю. Поэтому их можно записать в виде

$$(31) \quad \sum_{s=0}^p \hat{\beta}_s r_{j-s}(\mathbf{x}^{(3)}) = 0, \quad j = q + 1, \dots, q + p,$$

где $r_{j-s}(\mathbf{x}^{(3)})$ обозначает остаток r_{j-s} от его (асимптотической) регрессии на $\mathbf{x}^{(3)}$.

Полученные уравнения являются нелинейными, поскольку Φ и $\mathbf{x}^{(3)}$ зависят от неизвестных параметров. Тем не менее решения уравнений (30) и (31) можно получить методом итераций, вычисляя Φ и $\mathbf{x}^{(3)}$ для некоторых начальных оценок ρ и β_i и уточняя затем эти оценки путем решения (30) и (31) с полученными значениями Φ и $\mathbf{x}^{(3)}$, и т. д. В качестве начального приближения можно взять состоятельные оценки искомых параметров, определяемые соотношениями $\hat{\rho}_i^0 = r_i$, $i = 1, \dots, q$, и $\sum_{s=0}^p \hat{\beta}_s^0 r_{j-s} = 0$, $j = q + 1, \dots, q + p$.

Оценки, определяемые уравнениями (30) и (31), состоятельны и асимптотически нормальны. Вектор

$$(32) \quad \sqrt{T}(\hat{\rho} - \rho) = \sqrt{T}(\mathbf{x}^{(1)} - \rho) - \Phi_{13} \Phi_{33}^{-1} \sqrt{T} \mathbf{x}^{(3)}$$

имеет предельное нормальное распределение с нулевым средним и ковариационной матрицей $\Phi_{11} - \Phi_{13} \Phi_{33}^{-1} \Phi_{31}$. Уравнения (31) в силу $\hat{\beta}_0 = \beta_0 = 1$ можно записать в виде

$$(33) \quad \sum_{s=1}^p \hat{\beta}_s r_{j-s}(\mathbf{x}^{(3)}) = -r_j(\mathbf{x}^{(3)}), \quad j = q + 1, \dots, q + p.$$

Совместное распределение статистик $\sqrt{T}[r_i(\mathbf{x}^{(3)}) - \rho_i]$, $i = 1, \dots, q + p$, асимптотически нормально. Поэтому (33) эквивалентно уравнению

$$(34) \quad \sum_{s=1}^p r_{j-s}(\mathbf{x}^{(3)}) \sqrt{T}(\hat{\beta}_s - \beta_s) = -\sqrt{T} \sum_{s=0}^p r_{j-s}(\mathbf{x}^{(3)}) \beta_s, \\ j = q + 1, \dots, q + p.$$

С учетом (16) правая часть (34) может быть записана в виде

$$(35) \quad -\sum_{s=0}^p \sqrt{T}[r_{j-s}(\mathbf{x}^{(3)}) - \rho_{j-s}] \beta_s, \quad j = q + 1, \dots, q + p,$$

и имеет поэтому в пределе нормальное распределение с нулевыми средними и ковариационной матрицей $\Phi_{22} - \Phi_{23}\Phi_{33}^{-1}\Phi_{32}$. Поскольку $[r_{j-s}(\mathbf{x}^{(3)})]$ сходится по вероятности к $\mathbf{R}'$, то (34) показывает, что предельное распределение вектора $\sqrt{T}(\hat{\beta} - \beta)$, где $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_p)'$ и $\beta = (\beta_1, \dots, \beta_p)'$, совпадает с предельным распределением вектора

$$(36) \quad (\mathbf{R}')^{-1} \sqrt{T}(\mathbf{x}^{(2)} - \Phi_{23}\Phi_{33}^{-1}\mathbf{x}^{(3)}),$$

которое является нормальным, с нулевым средним и ковариационной матрицей

$$(37) \quad (\mathbf{R}')^{-1}(\Phi_{22} - \Phi_{23}\Phi_{33}^{-1}\Phi_{32})\mathbf{R}^{-1}.$$

Совместное распределение векторов $\sqrt{T}(\hat{\rho} - \rho)$ и $\sqrt{T}(\hat{\beta} - \beta)$ также сходится при $T \rightarrow \infty$ к нормальному. Ковариация этих двух векторов в предельном распределении равна

$$(38) \quad (\Phi_{12} - \Phi_{13}\Phi_{33}^{-1}\Phi_{32})\mathbf{R}^{-1}.$$

На рассматриваемый случай можно также распространить процедуру, предложенную Дурбином для модели скользящего среднего, несколько видоизменив ее (Дурбин (1960b)). Предположим, что мы исходим из некоторого начального приближения (начальных оценок) для параметров $\beta_1, \dots, \beta_p$. Обозначим соответствующие оценки $\hat{\beta}_1^0, \dots, \hat{\beta}_p^0$. Положим $u_t^{(0)} = \sum_{s=0}^p \hat{\beta}_s^0 y_{t-s}$. Используем теперь метод Дурбина, описанный в разд. 5.7.2, для оценки коэффициентов в модели скользящего среднего

$$(39) \quad u_t^{(0)} = \sum_{j=0}^q \alpha_j^{(0)} v_{t-j},$$

где $u_t^{(0)}$ рассматриваются как наблюдаемые значения. Обозначим получающиеся при этом оценки $\hat{\alpha}_1^{(0)}, \dots, \hat{\alpha}_q^{(0)}$.

Пусть случайный процесс $\{x_t\}$ таков, что

$$(40) \quad \sum_{s=0}^p \beta_s x_{t-s} = v_t, \quad t = \dots, -1, 0, 1, \dots$$

Тогда случайный процесс $\{y_t^*\}$, определяемый соотношением

$$(41) \quad y_t^* = x_t + \alpha_1 x_{t-1} + \dots + \alpha_q x_{t-q}, \quad t = \dots, -1, 0, 1, \dots,$$

будет удовлетворять уравнению (1) для $t = \dots, -1, 0, 1, \dots$. Если последовательность $\{x_t\}$ задать для $t = \dots, -1, 0$, а последовательность $\{y_t^*\}$ задать для $t = 1, 2, \dots$, то последовательность

$\{x_t\}$ можно продолжить, решая уравнение (41) последовательно относительно $x_1, x_2, \dots$. Это приводит к мысли определять $x_t^{(0)}$ для $t = 1, \dots, T$, используя $\hat{\alpha}_1^{(0)}, \dots, \hat{\alpha}_q^{(0)}$ и произвольные $x_0^{(0)}, \dots, x_{-q+1}^{(0)}$, следующим образом:

$$(42) \quad x_t^{(0)} = y_t - (\hat{\alpha}_1^{(0)} x_{t-1}^{(0)} + \dots + \hat{\alpha}_q^{(0)} x_{t-q}^{(0)}), \quad t = 1, \dots, T.$$

Параметры β_r можно оценить теперь исходя из последовательности $\{x_t^{(0)}\}$, удовлетворяющей (41). В результате получим новое приближение $\hat{\beta}_1^{(1)}, \dots, \hat{\beta}_p^{(1)}$. Дурбин предложил переходить поочередно от одной процедуры к другой таким образом, что после $\hat{\beta}_1^{(1)}, \dots, \hat{\beta}_p^{(1)}$ вычисляется $\hat{\alpha}_1^{(1)}, \dots, \hat{\alpha}_q^{(1)}$, затем $\hat{\beta}_1^{(2)}, \dots, \hat{\beta}_p^{(2)}$ и т. д. Можно ожидать, что при больших значениях T (и n) итерационный процесс будет сходиться.

5.9. НЕКОТОРЫЕ ПРИМЕРЫ

Вольд (1965) образовывал искусственные временные ряды с помощью разностных уравнений второго порядка, в которых u_t считались независимыми случайными нормальными отклонениями. Три таких ряда затабулированы в приложении А.2, там же приводятся их графики. При этом $\beta_1 = -\gamma$, $\beta_2 = \gamma^2$, корни характеристического уравнения равны $\gamma \exp \{\pm i2\pi/6\}$ и $\gamma = 0.25, 0.7$ и 0.9 соответственно. В каждом из этих случаев дисперсия величины u_t выбиралась так, чтобы дисперсия y_t равнялась единице. На приведенных графиках можно проследить тенденцию к наличию колебаний с периодом, близким к 6. Этот период выражен заметнее для больших значений γ .

Значения первых трех выборочных корреляций ($r_h = C_h/C_0$) указанных рядов для $T = 200$ и соответствующих им теоретических корреляций [$\rho_h = \sigma(h)/\sigma(0)$] приведены в табл. 5.1.

Таблица 5.1

ВЫБОРОЧНЫЕ И ТЕОРЕТИЧЕСКИЕ КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ ДЛЯ ТРЕХ ПРОЦЕССОВ АВТОРЕГРЕССИИ ($T = 200$)

γ	r_1	r_2	r_3	ρ_1	ρ_2	ρ_3
0.25	0.2473	0.1120	0.0492	0.23529	-0.00368	-0.015625
0.70	0.5113	-0.0473	-0.3001	0.46980	-0.16114	-0.34300
0.90	0.5011	-0.3858	-0.7726	0.49724	-0.36249	-0.72900

Оценки для β_1 и β_2 в предположении $p = 2$ и оценки для $\beta_1, \beta_2, \beta_3$ в предположении $p = 3$ получаются соответственно из уравнений

$$(1) \quad \begin{pmatrix} 1 & r_1 \\ r_1 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = - \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}, \quad \begin{pmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = - \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}.$$

Соответствующие результаты приведены в табл. 5.2.

Таблица 5.2

ОЦЕНКИ КОЭФФИЦИЕНТОВ ПРОЦЕССОВ АВТОРЕГРЕССИИ

γ	β_1	β_2	b_1	b_2	b_1	b_2	b_3
0.25	-0.25	0.0625	-0.2339	-0.0542	-0.2334	-0.0518	-0.0103
0.70	-0.70	0.49	-0.7250	0.4180	-0.6893	0.3561	0.0854
0.90	-0.90	0.81	-0.9273	0.8505	-0.9736	0.9010	-0.0545

Критерий для проверки нулевой гипотезы $\beta_3 = 0$ можно строить, используя статистику $\sqrt{T} b_3$, значения которой в указанных трех случаях равны соответственно -0.1458 , 1.208 и -0.7706 . При нулевой гипотезе значения статистики $\sqrt{T} b_3$ должны быть распределены приблизительно нормально с нулевым средним и единичной дисперсией. Если ограничиться разумными уровнями значимости, например 1%, 5% или 10%, то ни в одном из трех случаев нулевая гипотеза не будет отвергнута соответствующими двусторонними критериями.

Юл (1927) предложил использовать процесс авторегрессии в качестве более предпочтительной статистической модели по сравнению с моделью, в которой случайная ошибка накладывается на тригонометрический тренд. (Модель последнего типа рассматривалась нами в гл. 4.) Он применил его к числам Вольфа солнечной активности, которые представляют собой данные ежегодных измерений солнечной активности за период с 1749 по 1924 г. [Эти данные более широко и более подробно представлены Вальдмейером (1961).] Числа, использованные Юлом, затабулированы в табл. А.3.1 приложения А.3. Подобные данные анализировались Крэддоком (1967). Соответствующие результаты приведены ниже. В первом приближении эти ряды выглядят похожими на искусственные ряды Вольда.

Первые пять корреляций, приведенные Юлом, указаны в табл. 5.3.

Для процесса второго порядка Юл получил оценки $b_1 = -1.34254$, $b_2 = 0.65504$, $\hat{v} = -13.854$, $\hat{\sigma} = 15.41$. Корни характеристического уравнения равны $0.67127 \pm 0.45215i =$

$= 0.80935 \exp \{ \pm i 33.963^\circ \}$. В соответствующей модели имеется тенденция к сохранению периода $360^\circ / 33.963^\circ = 10.600$ (лет). Это несколько меньше, чем обычно принимаемый период, чуть больший 11 лет. [Шустер (1906) принимает его равным 11.125 лет.]

Таблица 5.3

КОРРЕЛЯЦИИ ДЛЯ ЧИСЕЛ СОЛНЕЧНОЙ АКТИВНОСТИ

Запаздывание	Корреляции
1	0.811180
2	0.433998
3	0.031574
4	-0.264463
5	-0.404119

Таблица 5.4

ПОСЛЕДНИЕ КОЭФФИЦИЕНТЫ

p	$b_p(p)$	$\sqrt{T} b_p(p)$
1	-0.811180	-10.761
2	0.655040	8.690
3	0.101043	1.340
4	-0.013531	-0.180
5	0.050001	0.663

Используя данные табл. 5.3, Юл находит оценку параметра β_p в предположении, что процесс имеет порядок p , (частный коэффициент корреляции) для $p = 1, 2, 3, 4, 5$. Эти оценки приведены в табл. 5.4. Следует отметить, что значения $\sqrt{T} b_p(p)$, которые должны быть приблизительно нормально распределенными при $\beta_p = \beta_{p+1} = \dots = 0$, оказываются относительно малыми по абсолютной величине для $p = 3, 4, 5$.

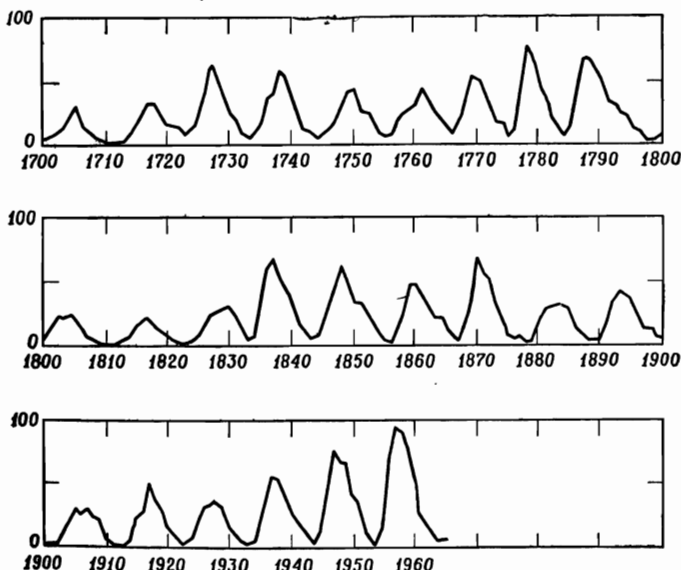


Рис. 5.1.

Солнечная активность за период с 1700 по 1965 г.

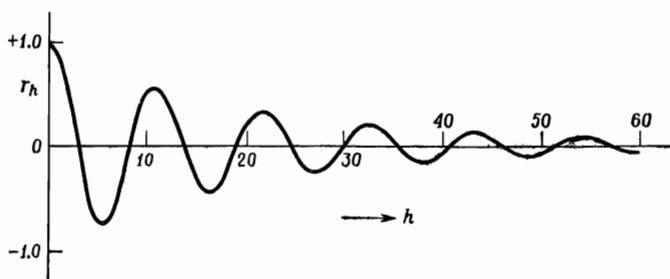


Рис. 5.2.

Коэффициенты корреляции для чисел солнечной активности по Крэддоку.

Изучение данных солнечной активности было предпринято также Крэддоком (1967), использовавшим ежегодные наблюдения с 1700 по 1965 г. График результатов наблюдений представлен на рис. 5.1. Соответствующие корреляции приведены на рис. 5.2. Крэддок производит подбор различных моделей авторегрессии для p от 1 до 30. На рис. 5.3 представлен график отношения $\hat{\sigma}_p^2/\hat{\sigma}_0^2$ (в процентах), где $\hat{\sigma}_p^2$ — оценка дисперсии на основании модели порядка p (с включенной в модель константой). Разность $T(\hat{\sigma}_{p-1}^2 - \hat{\sigma}_p^2)$ пропорциональна отношению $\hat{\beta}_p^2/\alpha^{*pp}$, которое используется при проверке гипотезы $\beta_p = 0$. [См. (16) § 5.6.] Фактически t^2 -статистикой является

$$(2) \quad (T-p) \frac{\hat{\sigma}_{p-1}^2 - \hat{\sigma}_p^2}{\hat{\sigma}_p^2} = (T-p) \left(\frac{\hat{\sigma}_{p-1}^2}{\hat{\sigma}_p^2} - 1 \right).$$

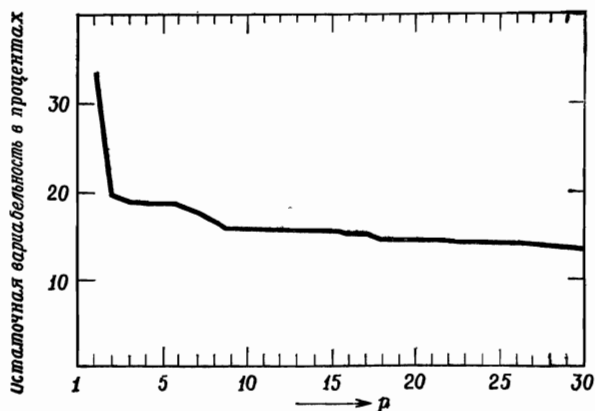


Рис. 5.3.

Остаточная вариабельность в моделях авторегрессии.

Приведенные данные весьма определенно указывают на то, что β_1 и β_2 должны быть взяты отличными от нуля. Далее, представляется сомнительным, чтобы значения $\beta_3, \dots, \beta_9$ были равны нулю. Значения $\beta_{10}, \dots, \beta_{18}$ более похожи на нулевые. Наконец, можно заключить, что $\beta_{19}, \dots, \beta_{30}$ будут нулями. В формальных терминах процедуры со многими решениями, описанной в конце разд. 5.6.3, возьмем $p = 30$ и $m = 0$. Коэффициенты $\sqrt{T}\hat{\beta}_j$, $j = 9, \dots, 30$,

Таблица 5.5

ОЦЕНКИ КОЭФФИЦИЕНТОВ ПРОЦЕССОВ АВТОРЕГРЕССИИ ДЛЯ ИНДЕКСА БЕВЕРИДЖА ЦЕН НА ПШЕНИЦУ С ВЫДЕЛЕННЫМ ТРЕНДОМ

p	$b_1(p)$	$b_2(p)$	$b_3(p)$	$b_4(p)$	$b_5(p)$	$b_6(p)$	$b_7(p)$	$b_8(p)$	$\sqrt{370} b_p(p)$	$\frac{\hat{\sigma}_p^2}{\hat{\sigma}_0^2}$
2	-0.7368	0.3110							5.984	0.6179
3	-0.7489	0.3397	-0.0388						-0.746	0.6170
4	-0.7503	0.3521	-0.0662	0.0367					0.706	0.6162
5	-0.7488	0.3494	-0.0516	0.0055	-0.0415				-0.798	0.6151
6	-0.7435	0.3501	-0.0582	0.0504	-0.0546	0.1284			2.470	0.6050
7	-0.7285	0.3437	-0.0524	0.0436	-0.0138	0.0417	0.1165		2.241	0.5968
8	-0.7123	0.3495	-0.0543	0.0496	-0.0211	0.0895	0.0153	0.1390	2.674	0.5853

равны приближенно 1.2, за исключением $\sqrt{T}\hat{\beta}_{18}$, равного примерно 2.2, и $\sqrt{T}\hat{\beta}_9$, примерно равного 3.8. Если последовательность ε_i , $i = 9, \dots, 30$, такова, что $0.0002 < \varepsilon_i < 0.02$, то порядок следует выбрать равным 9. Если же $0.03 < \varepsilon_i < 0.1$ для $i = 18, \dots, 30$, то порядок следует принять равным 18. ($\varepsilon_1 = \dots = \varepsilon_{30} = 0.0035$ соответствует $p_0 = 0.9$.)

Уиттл (1954) изучал солнечную активность на базе полугодовых данных за период с 1886 по 1945 г. и подыскал для них модель авторегрессии, использующую запаздывания на 1 и 22, т. е. на 6 месяцев и 11 лет. Было проведено и много других статистических исследований.

Рядами Вольфа и Уиттла занималась также Шерф (1954). Она подобрала модель авторегрессии с запаздываниями на 1, 2 и 9.

В качестве другого примера рассмотрим ряд Бевеиджа индексов цен на пшеницу с выделенным трендом, затабулированный в приложении А.1 вместе с его корреляциями. Оценки коэффициентов процесса авторегрессии порядка p даны в табл. 5.5 для $p = 2, 3, \dots, 8$. В предпоследнем столбце этой таблицы приведены значения статистики $\sqrt{T} b_p(p)$, которые могут быть использованы для проверки нулевой гипотезы $\beta_p = 0$ против альтернативы $\beta_p \neq 0$, в пред-

положении, что порядок процесса не выше p . Следует отметить, что при любом разумном уровне значимости нулевая гипотеза $\beta_2 = 0$ отвергается, в то время как нулевые гипотезы $\beta_p = 0$ для $p = 3, 4, 5$ принимаются. Нулевая гипотеза $\beta_p = 0$ отвергается для $p = 6, 7, 8$ с уровнем значимости 2,5% и принимается для $p = 6, 7$ с уровнем значимости 1%. Если полагать, что порядок не может быть больше p для $p = 6, 7$ или 8, и использовать предложенную ранее процедуру со многими решениями, то решение о том, что порядок равен p , принималось бы при уровне значимости, не меньшем 0.025. С другой стороны, если использовать процедуру со многими решениями, состоящую в поочередной проверке гипотез $\beta_j = 0$ для $j = 2, 3, \dots$ до тех пор, пока какая-то из этих гипотез не будет принята и будет решено, что порядок равен порядку, соответствующему последней отклоненной гипотезе, то можно было бы прийти к заключению, что процесс имеет порядок 2 [как это было у Саргана (1953)]. Однако, судя по табл. 5.5, для порядка рассматриваемого процесса может оказаться наиболее подходящим и значение, большее 8.

5.10. ОБСУЖДЕНИЕ

В настоящей главе мы обсудили математические модели стационарных процессов, определяемые конечным числом параметров. При этом модели были таковы, что о них можно говорить как о линейных, имея в виду линейность по коэффициентам, по наблюдаемым переменным и по гипотетическим случайным переменным. В ряде случаев модель соответствует теории образования наблюдаемых данных и ее коэффициенты имеют самостоятельный смысл. В других случаях модель является аппроксимацией, оказывающейся адекватной для многих целей.

Модель авторегрессии обладает рядом преимуществ по сравнению с моделью скользящего среднего и процессом авторегрессии с остатками в виде скользящего среднего, хотя последние в определенных случаях могут хорошо описывать образование наблюдаемых временных рядов. Оценки коэффициентов процесса авторегрессии легко вычисляются. Статистические процедуры для такого процесса, основывающиеся на теории больших выборок, легко выполнимы, поскольку они соответствуют обычной технике наименьших квадратов. Во многих случаях коэффициенты процесса авторегрессии допускают непосредственную интерпретацию, а линейные функции от запаздывающих переменных могут быть использованы для прогнозирования.

Как мы увидим позднее, каждую из этих моделей можно использовать для аппроксимации и оценивания спектральной плотности.

ЛИТЕРАТУРА

- § 5.1. Дж. Уолкер (1931), Юл (1927).
 § 5.3. Тернбулл и Эйкен (1952), Халмош (1958).
 § 5.4. Дурбин (1960a), (1960b), Манн и Вальд (1943b).
 § 5.5. Т. Андерсон (1959), Т. Андерсон и Рубин (1950), Купменс, Рубин и Лейпник (1950), Лозв (1963).
 § 5.6. Бартлетт и Диананда (1950), Кенуй (1947), А. Уолкер (1952).
 § 5.7. Дурбин (1959), А. Уолкер (1961).
 § 5.8. Дуб (1944), Дурбин (1960b), А. Уолкер (1962).
 § 5.9. Вальдмейер (1961), Вольд (1965), Крэдок (1967), Сарган (1953), Уиттл (1954), Шерф (1964), Шустер (1906), Юл (1927).
 Упражнения. Кузнец (1954), Хаавелмо (1947).

УПРАЖНЕНИЯ

1. (§ 5.1) Найдите коэффициенты $\alpha_1, \dots, \alpha_p$ в соотношении (2), если последняя записана в виде

$$(\Delta^p + \alpha_1 \Delta^{p-1} + \dots + \alpha_p) y_{t-p} = u_t.$$

2. (Разд. 5.2.1) Проведите подробно выкладки в (30) для случая $q = 1, p = 2$. Выразите u_t^* , определяемое выражением (31), с помощью u_t для $r = 1$. Убедитесь, что u_t^* не коррелированы.

3. (Разд. 5.2.1) Пусть $\{y_t\}$ — стационарный процесс, удовлетворяющий уравнению

$$y_t + y_{t-1} = u_t,$$

в котором $\xi u_t = 0$, $\xi u_t^2 = \sigma^2$ и $\xi u_t u_s = 0$, $t \neq s$. Докажите, что $y_t = (-1)^s y_{t-s}$ с вероятностью 1.

4. (Разд. 5.2.1) Докажите, что линейное разностное уравнение порядка p

$$\beta_0(t) z_t + \beta_1(t) z_{t-1} + \dots + \beta_p(t) z_{t-p} = \alpha(t),$$

где $\beta_0(t) \neq 0$, $\beta_p(t) \neq 0$ на множестве S последовательных значений t имеет одно и только одно решение z_t , для которого значения z_t при p последовательных значениях t равны заданным. (Указание. Доказать индукцией по t , что значения решения z_t при любых p последовательных целых значениях t определяют значение z_t для следующего целого t .)

5. (Разд. 5.2.1) Покажите, что t -я степень любого корня x_t характеристического уравнения (23), соответствующего стохастическому разностному уравнению

$$\beta_0 \omega_t + \beta_1 \omega_{t-1} + \dots + \beta_p \omega_{t-p} = 0,$$

является решением последнего ($\beta_0 \neq 0$, $\beta_p \neq 0$). (Указание. Сделать подстановку $\omega_t = x_t^t$.)

6. (Разд. 5.2.1) Покажите, что

$$(\mathcal{P} - a) P_q(t) a^t = P_{q-1}(t) a^{t+1},$$

где $P_q(t)$ и $P_{q-1}(t)$ — полиномы степени q и $q-1$ соответственно.

7. (Разд. 5.2.1) Покажите, что

$$(\mathcal{P} - a)^m P_{m-1}(t) a^t = 0.$$

8. (Разд. 5.2.1) Покажите, что выражение $c_0 a^t + c_1 t a^{t+1} + \dots + c_{m-1} t^{m-1} a^{t-m+1}$ можно записать в виде $P_{m-1}(t) a^t$ (для $a \neq 0$).

9. (Разд. 5.2.1) Докажите, что если a — действительный корень кратности m характеристического уравнения (23), соответствующего однородному разностному уравнению упр. 5, то $w_t = P_{m-1}(t) a^t$ будет решением этого разностного уравнения.

10. (Разд. 5.2.1) Докажите, что если $\alpha e^{i\theta}$ и $\alpha e^{-i\theta}$ составляют пару комплексно сопряженных корней кратности m алгебраического уравнения (23), соответствующего однородному разностному уравнению упр. 5, то

$$w_t = P_{m-1}(t) \alpha^t e^{it\theta} + \bar{P}_{m-1}(t) \alpha^t e^{-it\theta}$$

является решением последнего. Здесь $P_{m-1}(t)$ и $\bar{P}_{m-1}(t)$ — полиномы, соответствующие коэффициенты которых комплексно сопряжены.

11. (Разд. 5.2.1) Докажите, что если все корни уравнения (23) различны, то всякое решение разностного уравнения из упр. 5 имеет вид $w_t = \sum_{i=1}^p k_i x_i^t$, где коэффициент k_i действителен, если действителен x_i , и коэффициенты k_i и k_{i+1} комплексно сопряжены, если таковы же x_i и x_{i+1} . (Указание. Используя упр. 5, покажите, что w_t указанного вида является решением, и, используя упр. 4, покажите, что оно будет единственным для заданных p последовательных (по t) значений w_t .)

12. (Разд. 5.2.1) Выясните, является ли условие $\beta_1 = 0$ в модели разностного уравнения второго порядка необходимым и достаточным для того, чтобы y_t и y_{t-1} были независимыми.

13. (Разд. 5.2.2) Выведите из (51) следующие выражения:

$$\sigma(0) = \frac{\sigma^2}{(1 + \beta_2)^2 - \beta_1^2} \cdot \frac{1 + \beta_2}{1 - \beta_2},$$

$$\sigma(1) = \frac{\sigma^2}{(1 + \beta_2)^2 - \beta_1^2} \cdot \frac{-\beta_1}{1 - \beta_2}.$$

14. (Разд. 5.2.2) Получите выражения, указанные в упр. 13, используя (42) и соотношения

$$\sigma(0) = \mathbb{E} \left(\sum_{r=0}^{\infty} \delta_r u_{t-r} \right)^2 = \sigma^2 \sum_{r=0}^{\infty} \delta_r^2,$$

$$\sigma(1) = \mathbb{E} \sum_{r=0}^{\infty} \delta_r u_{t-r} \sum_{s=0}^{\infty} \delta_s u_{t-1-s} = \sigma^2 \sum_{r=0}^{\infty} \delta_r \delta_{r+1}.$$

15. (Разд. 5.2.2) Постройте график δ_r и $\sigma(r)$ для $\sigma^2 = 1$ и (а) $\beta_1 = 0.9$, (б) $\beta_1 = -0.9$, (с) $\beta_1 = 0.1$, (д) $\beta_1 = -0.1$.

16. (Разд. 5.2.2) Постройте график δ_r и $\sigma(r)$ для $\sigma^2 = 1$ и (а) $\beta_1 = 0.25$, $\beta_2 = 0.0625$, (б) $\beta_1 = 0.7$, $\beta_2 = 0.49$ и (с) $\beta_1 = 0.9$, $\beta_2 = 0.81$.

17. (§ 5.3) Найдите Λ^T , если

$$\Lambda = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix}.$$

18. (§ 5.3) Докажите, что кратность характеристического корня λ матрицы Λ в упр. 17 равна размеру этой матрицы и что у Λ имеется только один линейно независимый вектор.

19. (§ 5.3) Докажите, что вектор $(x_i^{p-1}, x_i^{p-2}, \dots, 1)'$ является и притом единственным характеристическим вектором матрицы $-\tilde{B}$ (с точностью до умножения на константу), соответствующим корню x_i произвольной кратности. Опираясь на этот факт, докажите, что если не все эти корни различны, то матрица $-\tilde{B}$ не может быть выражена в виде ΛC^{-1} , где матрица Λ диагональна.

20. (§ 5.3) Покажите, что соотношения (28) § 5.2 и (22) § 5.3 тождественны.

21. (§ 5.4) Найдите условия, при которых

$$\frac{1}{T} \sum_{t=1}^T \frac{y_t y_{t-1}}{y_{t-1}^2} = \frac{1}{T} \sum_{t=1}^T \frac{y_t}{y_{t-1}}$$

будет несмещенной оценкой для $-\beta_1$ при $p = 1$ (и $q = 0$).

22. (§ 5.4) Оценками в п. (b) упр. 18 гл. 3 служат $\hat{\gamma}_0 = 72.31$, $\hat{\gamma}_1 = 2.206$ и $\hat{\gamma}_2 = -0.1109$.

(a) Найдите оценки (максимального правдоподобия) коэффициентов β_1 , β_2 и дисперсии σ^2 в модели

$$\tilde{y}_t + \beta_1 \tilde{y}_{t-1} + \beta_2 \tilde{y}_{t-2} = u_t,$$

где u_t имеют нулевые средние и дисперсию σ^2 ,

$$\tilde{y}_t = y_t - \gamma_0 \varphi_{0T}(t) - \gamma_1 \varphi_{1T}(t)$$

и y_{-1} есть цена за 1897 г., y_0 — за 1898 г. и т. д.

(b) Найдите оценки корней x_1 и x_2 характеристического уравнения (23) § 5.2, оценки коэффициентов δ_r в (42) или (44) § 5.2 и оценки $\sigma(r)$ в (51) § 5.2.

(c) Покажите, что оценки для β_1 и β_2 в (a) близки к оценкам максимального правдоподобия в модели

$$y_t + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \alpha_0 + \alpha_1 t = u_t.$$

23. (§ 5.4) (a) Пусть оценки параметров β_1 и β_2 определены соотношением

$$\begin{pmatrix} 1 & r_1^* \\ r_1^* & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = - \begin{pmatrix} r_1^* \\ r_2^* \end{pmatrix}.$$

Покажите, что

$$b_2 = - \frac{r_2^* - r_1^{*2}}{1 - r_1^{*2}}.$$

(б) Пусть оценки параметров $\beta_1, \beta_2, \beta_3$ определены соотношениями

$$\begin{pmatrix} 1 & r_1^* & r_2^* \\ r_1^* & 1 & r_1^* \\ r_2^* & r_1^* & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = - \begin{pmatrix} r_1^* \\ r_2^* \\ r_3^* \end{pmatrix}.$$

Покажите, что

$$b_3 = - \frac{r_3^* - r_1^{*2}r_3^* + r_1^{*3} - 2r_1^*r_2^* + r_1^*r_2^{*2}}{1 - 2r_1^{*2} + 2r_1^{*2}r_2^* - r_2^{*2}}.$$

24. (Разд. 5.5.1) Распишите подробно уравнение (3) для $y_i = \tilde{y}_i$ и убедитесь, что это уравнение определяет те же самые оценки, что и уравнение (8) § 5.4.

25. (Разд. 5.5.1) Пусть в соотношении (1) u_i независимы и распределены согласно $N(0, \Sigma)$, а y_0 — заданный вектор. Покажите, что уравнения (3) и (4) определяют оценки максимального правдоподобия для $\mathbf{B}$ и Σ .

26. (Разд. 5.5.2) Покажите, что из того, что матрица $\mathbf{A}$ положительно полуопределена и $\text{tr} \mathbf{A} = 0$, вытекает $\mathbf{A} = \mathbf{0}$.

27. (Разд. 5.5.3) Докажите, что из уравнения (31) матрица $\mathbf{A}$ определяется однозначно. (Указание. Преобразовать (31) к виду

$$\mathbf{A}^* - \mathbf{A} \mathbf{A}^* \mathbf{A}' = \Sigma^*,$$

где $\mathbf{A}^* = \mathbf{C}^{-1} \mathbf{A} (\mathbf{C}')^{-1}$ и $\Sigma^* = \mathbf{C}^{-1} \Sigma (\mathbf{C}')^{-1}$.)

28. (Разд. 5.5.3) Покажите, что если $\mathbf{A} = \mathbf{B} + \mathbf{C}$ и при этом матрица $\mathbf{B}$ положительно определена, а матрица $\mathbf{C}$ положительно полуопределена, то матрица $\mathbf{A}$ будет положительно определенной. (Указание. $\mathbf{x}' \mathbf{A} \mathbf{x} = \mathbf{x}' \mathbf{B} \mathbf{x} + \mathbf{x}' \mathbf{C} \mathbf{x}$, а матрица $\mathbf{A}$ положительно определена тогда и только тогда, когда $\mathbf{x}' \mathbf{A} \mathbf{x} > 0$ и из $\mathbf{x}' \mathbf{A} \mathbf{x} = 0$ следует $\mathbf{x} = \mathbf{0}$.)

29. (Разд. 5.5.4) Убедитесь, что если случайная матрица $\mathbf{W}$ имеет ковариационную матрицу $\mathbf{A} \otimes \Sigma$, то матрица $\mathbf{A}^{-1} \mathbf{W}$ будет иметь ковариационную матрицу $\mathbf{A}^{-1} \otimes \Sigma$.

30. (Разд. 5.5.5) Пусть в соотношении (57) u_i независимы и распределены по закону $N(0, \Sigma)$, а y_0 — заданный вектор. Покажите, что уравнения (60) и (61) определяют оценки максимального правдоподобия для $\mathbf{B}$, $\mathbf{v}$ и Σ .

31. (Разд. 5.5.5) Вычислите (67) для $p = 2$ и $\mathbf{B} = \tilde{\mathbf{B}}$.

32. (Разд. 5.5.5) Покажите, что сумма математических ожиданий квадратов компонент (70) сходится к нулю.

33. (Разд. 5.6.3) Проверьте, что (50) имеет нулевое среднее, дисперсию σ^4 и что эти суммы при различных значениях индекса j не коррелированы.

34. (Разд. 5.7.1) Для $q = 1$ покажите, что $|\sigma(1)| \leq \sigma(0)/2$.

35. (Разд. 5.7.1) Докажите, что все корни уравнения (7), лежащие на единичной окружности, имеют четную кратность. (Указание. Для $z = e^{i\lambda}$ выражение

$$\sum_{g=-q}^q \sigma(g) z^g = \sum_{g=-q}^q \sigma(g) \cos \lambda g = f(\lambda),$$

является спектральной плотностью и поэтому неотрицательно. (См. § 7.3.) Покажите, что если $\lambda = \nu$ корень $f(\lambda)$, то он имеет четную кратность. Для этого надо показать, что наивысший порядок производной функции $f(\lambda)$, обращающейся в нуль при $\lambda = \nu$, должен быть четным.)

36. (Разд. 5.7.1) Докажите, что ряд (10) сходится в среднеквадратичном.

37. (Разд. 5.7.1) Для $q = 1$ покажите, что $|\Sigma_T| = \sigma^2 T (1 - \alpha_1^{2(T+1)}) \times \times (1 - \alpha_1^2)^{-1}$, где $\Sigma_T = [\sigma(t-s)]$ имеет размер $T \times T$. (Указание. См. разд. 6.7.7.)

38. (Разд. 5.7.1) Для $q = 1$ покажите, что

$$\lim_{T \rightarrow \infty} \sigma_T^{ts} = \frac{(-\alpha_1)^{|t-s|} (1 - \alpha_1^{2\min(t,s)})}{\sigma^2 (1 - \alpha_1^2)},$$

где (σ_T^{ts}) — матрица, обратная к матрице $[\sigma(t-s)] = \Sigma_T$ размера $T \times T$.

39. (Разд. 5.7.3) Для $q = 1$ и $\rho_1 = \rho$ покажите, что для положительных индексов

$$\begin{aligned} w_{11} &= 1 - 3\rho^2 + 4\rho^4, & w_{jj} &= 1 + 2\rho^2, & j &= 2, 3, \dots, \\ w_{12} &= 2\rho - 2\rho^3, & w_{j,j+1} &= 2\rho, & j &= 2, 3, \dots, \\ w_{j,j+2} &= \rho^2, & j &= 1, 2, \dots, & w_{ij} &= 0, & |i-j| > 2. \end{aligned}$$

40. (§ 5.9) В следующей таблице представлен график совокупного дохода, остающегося после уплаты налогов (в сопоставимых ценах), в США за период с 1919 по 1941 г., в млрд. долларов:

1919	44.32	1927	59.26	1935	57.47
1920	45.45	1928	61.58	1936	65.87
1921	44.08	1929	65.04	1937	67.39
1922	47.67	1930	59.18	1938	62.34
1923	54.10	1931	54.91	1939	68.09
1924	54.65	1932	46.72	1940	72.77
1925	56.28	1933	48.12	1941	84.29
1926	58.00	1934	53.25		

Эти данные взяты у Хаавелмо (1947) за 1922—1941 гг. и у Кузнеця (1954, стр. 151—153) за 1919—1921 гг.

(а) Начертите график ежегодных доходов.

(б) Найдите оценки (максимального правдоподобия) коэффициентов $\beta_1, \beta_2, \gamma_1, \gamma_2$ и стандартного отклонения σ в модели

$$y_t + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \gamma_1 + \gamma_2 t = u_t, \quad t = 1, \dots, 21,$$

где u_t имеют нулевые средние и дисперсии σ^2 , y_t представляет доход за $t+1920$ год, $t = -1, 0, 1, \dots, 21$, а y_{-1} и y_0 рассматриваются как заданные величины.

Матрица сумм квадратов и перекрестных произведений для $1, t, y_{t-1}, y_{t-2}, y_t$ равна

$$\begin{pmatrix} 21 & 231 & 1202 & 1174 & 1241 \\ 231 & 3311 & 13905 & 13551 & 14473 \\ 1202 & 13905 & 70112 & 68247 & 72365 \\ 1174 & 13551 & 68247 & 66781 & 70313 \\ 1241 & 14473 & 72365 & 70313 & 75151 \end{pmatrix}$$

(с) Найдите оценки (максимального правдоподобия) параметров μ и δ (коэффициентов тренда для модели, записанной в виде

$$y_t - (\mu + \delta t) + \beta_1 \{y_{t-1} - [\mu + \delta(t-1)]\} + \beta_2 \{y_{t-2} - [\mu + \delta(t-2)]\} = u_t.$$

(d) Найдите корни характеристического уравнения, соответствующего стохастическому разностному уравнению, имеющему в качестве коэффициентов полученные оценки, укажите аргументы и амплитуды этих (комплексных) корней.

СЕРИАЛЬНАЯ КОРРЕЛЯЦИЯ

6.1. ВВЕДЕНИЕ

Одна из первоочередных задач в анализе временных рядов состоит в том, чтобы решить, не является ли наблюдаемый ряд реализацией последовательности независимых случайных величин. В качестве простой альтернативы здесь можно использовать процесс, у которого последовательные наблюдения коррелированы. Примером такой альтернативы будет процесс, порожаемый стохастическим разностным уравнением первого порядка. В подобных случаях для проверки гипотезы о независимости может оказаться уместным использование различного вида сериальных корреляций первого порядка.

Пусть наблюдаемый ряд $y_1, \dots, y_T$ является реализацией процесса, имеющего нулевое математическое ожидание. Одно из возможных определений сериальной корреляции первого порядка таково:

$$(1) \quad r_1 = \frac{\sum_{t=2}^T y_t y_{t-1}}{\sum_{t=1}^T y_t^2}.$$

В этой главе рассматриваются различные определения сериальной корреляции. Если известно, что средние значения равны нулю, то эти определения являются модификациями (1). Если же средние значения неизвестны, то y_t в (1) заменяется отклонением от выборочного среднего (или от оценки тренда). Многие модификации связаны с изменением крайних членов, т. е. членов, содержащих y_1, y_2, y_{T-1}, y_T . Другие включают в себя введение дополнительного сомножителя, например $T/(T-1)$.

Сериальная корреляция служит мерой взаимной зависимости между элементами последовательности наблюдений, подобно тому как мерой зависимости между двумя множествами наблюдений $x_1, \dots, x_N$ и $z_1, \dots, z_N$ служит обычный коэффициент корреляции (средние предполагаются равными нулю)

$$(2) \quad r = \frac{\sum_{i=1}^N x_i z_i}{\left(\sum_{i=1}^N x_i^2 \right)^{1/2} \left(\sum_{i=1}^N z_i^2 \right)^{1/2}}.$$

В случае сериальной корреляции в числителе последнего выражения переменная x_i заменяется на y_t , а z_i на y_{t-1} . Поскольку $\sum_{t=2}^T y_t^2$ и $\sum_{t=2}^T y_{t-1}^2 = \sum_{t=1}^{T-1} y_t^2$ служат в случае стационарного процесса (с нулевым средним) оценками для $(T-1)/\sigma^2$, то эти статистики можно заменить на $\sum_{t=1}^T y_t^2$ (возможно, домножив еще на $(T-1)/T$). Таким образом, определение (1) аналогично определению (2). Сериальную корреляцию называют иногда *автокорреляцией*.

Сериальные корреляции более высокого порядка определяются аналогичным образом. Так, одним из определений сериальной корреляции ¹⁾ j -го порядка является

$$(3) \quad r_j = \frac{\sum_{t=j+1}^T y_t y_{t-j}}{\sum_{t=1}^T y_t^2}.$$

Эта величина измеряет взаимную зависимость наблюдений, сдвинутых на j единиц времени. Она может использоваться для проверки нулевой гипотезы о независимости против альтернативы, утверждающей существование зависимости между наблюдениями, отстоящими друг от друга на j единиц времени.

В большинстве ситуаций, однако, полагают, что наличие зависимости наблюдений, отстоящих друг от друга на i единиц времени, автоматически приводит к зависимости наблюдений, отстоящих

¹⁾ В дальнейшем автор обычно называет такие отношения *сериальными коэффициентами корреляции*.— Прим. перев.

друг от друга также на $1, 2, \dots, i - 1$ единиц времени. В связи с этим сериальную корреляцию i -го порядка обычно используют для проверки нулевой гипотезы, утверждающей существование зависимости $(i - 1)$ -го порядка, против альтернативы, утверждающей, что порядок зависимости равен i . Например, можно говорить, что процесс, порождаемый стохастическим разностным уравнением i -го порядка, соответствует зависимости порядка i . При этом сериальный коэффициент корреляции порядка i следовало бы использовать вместе с сериальными коэффициентами корреляции меньшего порядка.

Более общей является задача определения порядка зависимости для случая, когда предполагается, что этот порядок не меньше m (≥ 0) и не больше p ($> m$). Такую процедуру со многими решениями можно построить, используя сериальные коэффициенты корреляции вплоть до порядка p .

Вопросы проверки независимости, проверки порядка зависимости и установления наличия зависимости были по существу уже рассмотрены в гл. 5 при исследовании модели стохастического разностного уравнения (фактически стационарного гауссовского процесса конечного порядка). Свойства соответствующих процедур, распределения оценок и критерии, используемые для этих целей, были изложены в рамках теории больших выборок. В настоящей главе будут получены оптимальные процедуры и точные распределения для моделей, являющихся некоторыми модификациями модели стохастического разностного уравнения. (Например, оценка $-\hat{\beta}_1$ в модели стохастического разностного уравнения первого порядка имеет вид сериального коэффициента первого порядка.)

Общие типы моделей мы получим исходя из модели стохастического разностного уравнения (§ 6.2). Для этих экспоненциальных моделей предлагаются оптимальные (односторонние или несмещенные) критерии для проверки зависимости. Разрабатываются процедуры со многими решениями, использующие такие критерии (§ 6.4). Соответствующие построения аналогичны тем, которые использовались для определения степени полиномиального тренда (разд. 3.2.2); определяются некоторые точные модели как для случая известных средних (§ 6.5), так и для случая, когда средние неизвестны, включая модели с трендом (§ 6.6). Эти модели связываются с определениями сериальных коэффициентов корреляции. При их исследовании используются циклические сериальные коэффициенты корреляции, а также коэффициенты, основанные на сумме квадратов последовательных разностей. Мы получим точные распределения некоторых сериальных коэффициентов корреляции (§ 6.7) и их приближенные распределения (§ 6.8), а также найдем некоторые совместные распределения (§ 6.9). Использование сериальных корреляций в качестве оценок будет рассмотрено позднее (гл. 8).

6.2. ТИПЫ МОДЕЛЕЙ

В гл. 5 мы изучали временные ряды, порожденные стохастическим разностным уравнением порядка p :

$$(1) \quad y_t + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} = u_t,$$

где $\{u_t\}$ — последовательность независимых случайных величин с нулевыми средними и дисперсиями σ^2 . Если распределение величин $y_1, \dots, y_p$ задано, а u_t не зависит от $y_{t-1}, y_{t-2}, \dots, y_1$ для $t = p+1, \dots, T$, то распределение величин $u_{p+1}, \dots, u_T$ будет полностью определять совместное распределение $y_1, \dots, y_T$, являющихся наблюдаемыми величинами. Если, к тому же, случайные величины u_t являются одинаково распределенными, то совместное распределение величин $y_1, \dots, y_p$ можно подобрать таким образом, что y_t будут образовывать стационарный случайный процесс, если только все корни соответствующего характеристического уравнения лежат в единичном круге. [В разд. 5.2.1 было показано, что для того, чтобы случайный процесс $\dots, y_{-1}, y_0, y_1, \dots$, удовлетворяющий стохастическому разностному уравнению (1), был стационарным и величина y_t не зависела от $u_{t+1}, u_{t+2}, \dots$, необходимо и достаточно, чтобы все корни характеристического уравнения, соответствующего (1), лежали в единичном круге.] Если σ^2 является единственным неизвестным параметром распределения случайной величины u_t , то параметрами процесса будут $\beta_1, \dots, \beta_p$ и σ^2 .

Займемся рассмотрением временных рядов, имеющих совместные нормальные распределения с нулевыми средними. Случай, когда средние отличны от нуля, будет рассмотрен в этой главе позднее. Модели с нормальными распределениями называются *гауссовскими*. В таких моделях совместное распределение величин $y_1, \dots, y_p$ определяется полностью заданием ковариационной матрицы этих величин, которую будем обозначать Σ_{pp} . Пусть

$$(2) \quad \Sigma_{pp}^{-1} = (\sigma_{(p)}^{ts}).$$

Тогда плотность совместного распределения величин $y_1, \dots, y_p$ равна

$$(3) \quad \frac{1}{(2\pi)^{p/2} |\Sigma_{pp}|^{1/2}} \exp\left(-\frac{1}{2} \sum_{t,s=1}^p \sigma_{(p)}^{ts} y_t y_s\right),$$

а плотность совместного распределения $u_{p+1}, \dots, u_T$ равна

$$(4) \quad \frac{1}{(2\pi)^{(T-p)/2} (\sigma^2)^{(T-p)/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{t=p+1}^T u_t^2\right).$$

Поэтому плотность совместного распределения $y_1, \dots, y_T$ есть

$$(5) \quad \frac{1}{(2\pi)^{T/2} |\Sigma_{pp}|^{1/2} (\sigma^2)^{(T-p)/2}} \exp \left\{ -\frac{1}{2} \left[\sum_{t,s=1}^p \sigma_{(p)}^{ts} y_t y_s + \right. \right. \\ \left. \left. + \frac{1}{\sigma^2} \sum_{t=p+1}^T (y_t + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p})^2 \right] \right\}.$$

Выражение в показателе экспоненты, заключенное в квадратные скобки, равно

$$(6) \quad \sum_{t,s=1}^p \sigma_{(p)}^{ts} y_t y_s + \frac{1}{\sigma^2} \sum_{t=p+1}^T [y_t^2 + \beta_1^2 y_{t-1}^2 + \dots + \beta_p^2 y_{t-p}^2 + \\ + 2\beta_1 y_t y_{t-1} + 2\beta_1 \beta_2 y_{t-1} y_{t-2} + \dots + 2\beta_{p-1} \beta_p y_{t-p+1} y_{t-p} + \dots \\ \dots + 2\beta_p y_t y_{t-p}] = \\ = \frac{1}{\sigma^2} \left[(1 + \beta_1^2 + \dots + \beta_p^2) \sum_{t=p+1}^{T-p} y_t^2 + 2(\beta_1 + \beta_1 \beta_2 + \dots \right. \\ \left. \dots + \beta_{p-1} \beta_p) \sum_{t=p+1}^{T-p+1} y_t y_{t-1} + \dots + 2\beta_p \sum_{t=p+1}^T y_t y_{t-p} \right] + \\ + \sum_{t,s=1}^T a_{ts} y_t y_s + \sum_{t,s=1}^p b_{ts} y_{T+1-t} y_{T+1-s}.$$

(Мы предполагаем $T \geq 2p$.) Представим (6) в виде $y' \Sigma^{-1} y$, где $y = (y_1, \dots, y_T)'$.

Если Σ_{pp} выбирается так, чтобы процесс оказался стационарным, то из соотношения

$$(7) \quad \sigma(s-t) = \sigma[(T-s+1) - (T-t+1)]$$

следует, что элемент, стоящий на пересечении t -й строки и s -го столбца матрицы Σ^{-1} , совпадает с элементом $(T-t+1)$ -й строки и $(T-s+1)$ -го столбца той же матрицы. Ввиду этого указанная квадратичная форма по переменным $y_1, \dots, y_p$ будет иметь точно такой же вид, как и квадратичная форма по переменным $y_T, \dots, y_{T+1-p}$, т. е. $a_{ts} = b_{ts}$, $t, s = 1, \dots, p$. Коэффициенты последней можно получить из соотношения (6).

Заметим, что в квадратичной форме (6) отсутствуют слагаемые $y_t y_s$, для которых $|t-s| > p$. С тождеством до крайних членов эта квадратичная форма является линейной комбинацией сумм $\sum_t y_t y_{t-i}$,

$j = 0, 1, \dots, p$, именно

$$(8) \quad \frac{1 + \beta_1^2 + \dots + \beta_p^2}{\sigma^2} \sum_t y_t^2 + 2 \frac{\beta_1 + \beta_1 \beta_2 + \dots + \beta_{p-1} \beta_p}{\sigma^2} \sum_t y_t y_{t-1} + \dots + 2 \frac{\beta_p}{\sigma^2} \sum_t y_t y_{t-p}.$$

Коэффициенты этой квадратичной формы (в количестве $p + 1$) задают преобразование $p + 1$ параметров $\sigma^2, \beta_1, \dots, \beta_p$. Гипотезы и задачи со многими решениями можно формулировать, используя эти коэффициенты. Отметим, что $2\beta_p/\sigma^2$ является коэффициентом при $\sum_t y_t y_{t-p}$.

Последние члены в (6) [точнее две последние суммы в правой части (6)] также являются некоторыми функциями от $\sigma^2, \beta_1, \dots, \beta_p$. Отсюда следует, что достаточное множество статистик будут составлять здесь $p + 1$ сумм $\sum_t y_t y_{t-j}$, $j = 0, 1, \dots, p$, и, кроме того, не-

которые из произведений $y_t y_s$ и $y_{t+1-i} y_{t+1-s}$, $t, s = 1, \dots, p$. В свою очередь тот факт, что число статистик в минимальном достаточном множестве много больше, чем число оцениваемых параметров, не дает возможности поставить задачи статистических выводов таким образом, чтобы они приводили к относительно простым оптимальным процедурам.

Если значение T довольно велико по сравнению с p , то суммы $\sum_t y_t y_{t-j}$ будут доминировать над остальными членами. Это наводит на мысль заменить (5) такой плотностью, которая использовала бы те же самые суммы, но в которой крайние члены были бы изменены таким образом, чтобы можно было получить семейство только $p + 1$ достаточных статистик.

Предположим, что плотность совместного распределения величин $y_1, \dots, y_T$ имеет вид

$$(9) \quad f(y_1, \dots, y_T | \gamma_0, \dots, \gamma_p) = K \exp[-(\gamma_0 Q_0 + \dots + \gamma_p Q_p)/2],$$

где Q_j — квадратичная форма по $y = (y_1, \dots, y_T)'$,

$$(10) \quad Q_j = y' A_j y,$$

A_j — заданные симметрические матрицы, $j = 0, 1, \dots, p$, причем A_0 положительно определена, а $\gamma_0, \dots, \gamma_p$ — параметры, $\gamma_0 > 0$, такие, что квадратичная форма

$$(11) \quad Q = \gamma_0 Q_0 + \dots + \gamma_p Q_p = y' (\gamma_0 A_0 + \dots + \gamma_p A_p) y$$

положительно определена. При этих условиях (9) является многомерной нормальной плотностью. Константа $K = K(\gamma_0, \dots, \gamma_p)$

зависит от значений $\gamma_0, \dots, \gamma_p$ и имеет вид

$$(12) \quad K(\gamma_0, \dots, \gamma_p) = (2\pi)^{-T/2} |\gamma_0 A_0 + \dots + \gamma_p A_p|^{1/2}.$$

Обычно в качестве A_0 будет браться единичная матрица I .

Связь между параметрами в (9) и (5) следующая:

$$(13) \quad \begin{aligned} \gamma_0 &= \frac{1 + \beta_1^2 + \dots + \beta_p^2}{\sigma^2}, \\ \gamma_1 &= 2 \frac{\beta_1 + \beta_1 \beta_2 + \dots + \beta_{p-1} \beta_p}{\sigma^2}, \\ &\vdots \\ \gamma_{p-1} &= 2 \frac{\beta_{p-1} + \beta_1 \beta_p}{\sigma^2}, \\ \gamma_p &= 2 \frac{\beta_p}{\sigma^2}. \end{aligned}$$

Заметим, что $\gamma_{m+1} = \dots = \gamma_p = 0$ тогда и только тогда, когда $\beta_{m+1} = \dots = \beta_p = 0$. Обычно квадратичная форма Q_j равна

$\sum_{t=j+1}^T y_t y_{t-j}$ плюс квадратичная форма по переменным $y_1, \dots, y_j$ плюс подобная ей квадратичная форма по переменным $y_{T-j+1}, \dots, y_T$. Эти квадратичные формы выбираются соответствующим образом (например, с общими характеристическими векторами). Три примера систем квадратичных форм приводятся в § 6.5.

С помощью этой модели рассмотрим проверку гипотезы о том, что $\gamma_{m+1} = \dots = \gamma_p = 0$ для определенных заранее значений m и p . В частности, нас будет интересовать случай $m = p - 1$. Наиболее важным является случай, когда $p = 1$. Поэтому в литературе ему уделяется особое внимание. Мы найдем критерий, равномерно наиболее мощный против альтернатив $\gamma_p < 0$ (или, что эквивалентно, $\beta_p < 0$), а также равномерно наиболее мощный несмещенный критерий. Реальное использование этих критериев целиком зависит от степени разработанности теории и наличия необходимых таблиц.

Другой тип задач, которым мы будем заниматься, это процедура со многими решениями для выбора наибольшего индекса i , при котором еще $\gamma_i \neq 0$, среди $i = m, \dots, p$. Последний будет представлять собой порядок зависимости. Мы зададимся вероятностью превышения порядка для случая, когда порядок на самом деле равен i , $i = m, \dots, p - 1$, и потребуем, чтобы процедура была равномерно наиболее мощной против альтернатив о положительной зависимости ($\gamma_i < 0$) или равномерно наиболее мощной несмещенной процедурой. Такие процедуры строятся из ряда критериев проверки гипотез подобно тому, как это было сделано в разд. 3.2.2. [См. Т. Андерсон (1963).]

6.3. РАВНОМЕРНО НАИБОЛЕЕ МОЩНЫЕ КРИТЕРИИ ДЛЯ ПРОВЕРКИ ЗАДАННОГО ПОРЯДКА ЗАВИСИМОСТИ

6.3.1. Критерии, равномерно наиболее мощные против альтернатив о положительной зависимости

Исходной моделью для наблюдений $y_1, \dots, y_T$ будет модель с многомерной нормальной плотностью (9) из § 6.2. В настоящем параграфе мы будем интересоваться проверкой гипотезы о значении γ_i , $i = 1, \dots, p$, в предположении, что $\gamma_{i+1} = \dots = \gamma_p = 0$. В этом случае указанная плотность равна

$$(1) \quad f_i(y_1, \dots, y_T | \gamma_0, \dots, \gamma_i) = K_i(\gamma_0, \dots, \gamma_i) \times \\ \times \exp \left[-\frac{1}{2} (\gamma_0 Q_0 + \dots + \gamma_i Q_i) \right],$$

где

$$(2) \quad K_i(\gamma_0, \dots, \gamma_i) = K(\gamma_0, \dots, \gamma_i, 0, \dots, 0) = \\ = (2\pi)^{-T/2} |\gamma_0 A_0 + \dots + \gamma_i A_i|^{1/2},$$

а квадратичная форма $\gamma_0 Q_0 + \dots + \gamma_i Q_i$ положительно определена. Нулевая гипотеза состоит в том, что порядок зависимости меньше i (в предположении, что он не больше i). Если $i = 1$, то нулевая гипотеза $\gamma_i = 0$ есть просто гипотеза независимости (когда матрица A_0 диагональна, например когда $A_0 = I$). В этом разделе мы займемся проверкой гипотезы о том, что γ_i имеет заданное значение, против односторонней альтернативы. В случае когда $i = 1$ и выбрана нулевая гипотеза $\gamma_1 = 0$ ($\beta_1 = 0$), естественной является альтернатива $\gamma_1 < 0$ (т. е. $\beta_1 < 0$), которая соответствует положительной теоретической корреляции между последовательными наблюдениями. Если же $i > 1$, то альтернатива к нулевой гипотезе $\gamma_i = 0$ ($\beta_i = 0$) уже не обязательно будет иметь вид $\gamma_i < 0$ ($\beta_i < 0$), поскольку $-\beta_i$ представляет частную корреляцию между y_i и y_{i+1} при заданных корреляциях низшего порядка. Поскольку критерии, оптимальные против альтернатив $\gamma_i > 0$, строятся подобно критериям, оптимальным против альтернатив $\gamma_i < 0$, мы ограничимся рассмотрением только альтернатив, соответствующих положительной зависимости ($\gamma_i < 0$).

Для экспоненциальных семейств плотностей, таких, как (1), задачи статистических выводов относительно просты. Наиболее существенным является здесь тот факт, что минимальное достаточное множество статистик имеет в этом случае размерность, совпадающую с размерностью множества параметров.

ТЕОРЕМА 6.3.1. Если $y_1, \dots, y_T$ имеют плотность совместного распределения (1), то $Q_0, \dots, Q_i$ образуют достаточное множество статистик для параметров $\gamma_0, \dots, \gamma_i$.

Экспоненциальное семейство имеет классическую форму Купмена — Дармуа. Последняя, очевидно, удовлетворяет критерию факторизации, состоящему в том, что соответствующая плотность может быть представлена в виде произведения неотрицательной (измеримой) функции от параметров и достаточных статистик и неотрицательной (измеримой) функции от наблюдений, не зависящей от параметров. [См., например, Леман (1959, стр. 75) или Кендалл и Стьюарт (1961, стр. 41).] В рассматриваемом случае в качестве первой функции можно взять самую плотность, а в качестве второй использовать единицу. Из этой теоремы вытекает, что интересующие нас выводы могут быть сделаны на основании значений статистик $Q_0, \dots, Q_i$. Например, для любого критерия, использующего все наблюдения $y_1, \dots, y_t$, можно так подобрать критерий (быть может, рандомизированный), использующий только достаточные статистики $Q_0, \dots, Q_i$, что уровни значимости и функции мощности обоих критериев будут одинаковыми.

ТЕОРЕМА 6.3.2. Семейство распределений случайных величин $Q_0, \dots, Q_i$ в (1) при значениях параметров $\gamma_0, \dots, \gamma_i$, таких, что квадратичная форма $\gamma_0 Q_0 + \dots + \gamma_i Q_i$ положительно определена, является полным.

Свойство полноты также является общим свойством экспоненциальных семейств. [См., например, Леман (1959, стр. 183) или Кендалл и Стьюарт (1961, стр. 256).] Доказательство вынесено в упр. 7.

Из теоремы 6.3.2 следует, что критерий для проверки нулевой гипотезы о том, что γ_i равно заданному значению, например $\gamma_i^{(1)}$, имеющий заданный уровень значимости ε_i и определяемый подобной областью, обладает неймановской структурой. Иными словами, если S_i — множество в пространстве случайных величин $Q_0, \dots, Q_i$, такое, что

$$(3) \quad \text{Pr} \{S_i | \gamma_0, \gamma_1, \dots, \gamma_{i-1}, \gamma_i^{(1)}, 0, \dots, 0\} = \varepsilon_i,$$

то

$$(4) \quad \text{Pr} \{S_i | Q_0, \dots, Q_{i-1}; \gamma_0, \dots, \gamma_{i-1}, \gamma_i^{(1)}, 0, \dots, 0\} = \varepsilon_i$$

для почти всех возможных значений $Q_0, \dots, Q_{i-1}$. [См. Леман (1959, разд. 4.3) или Кендалл и Стьюарт (1961, п. 23.19).] Доказательство вынесено в упр. 9. Отметим, что, поскольку статистики $Q_0, \dots, Q_{i-1}$ достаточны для $\gamma_0, \dots, \gamma_{i-1}$, вероятность $\text{Pr} \{S_i | Q_0, \dots, Q_{i-1}; \gamma_0, \dots, \gamma_{i-1}, \gamma_i^{(1)}, 0, \dots, 0\}$ не зависит от значений $\gamma_0, \dots, \gamma_{i-1}$.

Пусть $h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_0, \dots, \gamma_i)$ — условная плотность распределения величин Q_i при заданных значениях $Q_0, \dots, Q_{i-1}$ и $\gamma_{i+1} = \dots = \gamma_p = 0$. Эта плотность будет интересовать нас на множестве V_{i-1} всех возможных значений $Q_0, \dots, Q_{i-1}$ (которое

характеризуется положительной маргинальной плотностью распределения величин $Q_0, \dots, Q_{i-1}$. Пусть $h_i(Q_0, \dots, Q_i | \gamma_0, \dots, \gamma_i)$ — плотность для $Q_0, \dots, Q_i$ при $\gamma_{i+1} = \dots = \gamma_p = 0$. Она выражается формулой

$$(5) \quad h_i(Q_0, \dots, Q_i | \gamma_0, \dots, \gamma_i) = K_i(\gamma_0, \dots, \gamma_i) \exp \left[-\frac{1}{2} (\gamma_0 Q_0 + \dots + \gamma_i Q_i) \right] k_i(Q_0, \dots, Q_i),$$

которая получается из соотношения для плотности величин $y_1, \dots, y_T$ преобразованием $y_1, \dots, y_T$ в величины $Q_0, \dots, Q_i$ и в $(T-i-1)$ величин y_t и последующим интегрированием этих $(T-i-1)$ величин y_t по области, определяемой значениями $Q_0, \dots, Q_i$. Выполнение этих операций одновременно определяет функцию $k_i(Q_0, \dots, Q_i)$ в (5). При этом V_i будет множеством значений $Q_0, \dots, Q_i$, для которых $k_i(Q_0, \dots, Q_i) > 0$. Маргинальная плотность величин $Q_0, \dots, Q_{i-1}$ будет равна

$$(6) \quad K_i(\gamma_0, \dots, \gamma_i) \exp \left[-\frac{1}{2} (\gamma_0 Q_0 + \dots + \gamma_{i-1} Q_{i-1}) \right] \times \\ \times m_i(Q_0, \dots, Q_{i-1}; \gamma_i),$$

где

$$(7) \quad m_i(Q_0, \dots, Q_{i-1}; \gamma_i) = \int_{-\infty}^{\infty} e^{-\gamma_i Q_i / 2} k_i(Q_0, \dots, Q_i) dQ_i.$$

(Заметим, что $m_i(Q_0, \dots, Q_{i-1}; \gamma_i) > 0$ для всех γ_i .) Таким образом, условная плотность для Q_i при заданных значениях статистик $Q_0, \dots, Q_{i-1}$ из V_{i-1} имеет вид

$$(8) \quad h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_0, \dots, \gamma_i) = e^{-\gamma_i Q_i / 2} \frac{k_i(Q_0, \dots, Q_i)}{m_i(Q_0, \dots, Q_{i-1}; \gamma_i)} = \\ = h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i).$$

Условная плотность распределения величины Q_i при заданных значениях $Q_0, \dots, Q_{i-1}$ представляет собой экспоненциальное семейство с параметром γ_i . Действительно, для фиксированных значений $Q_0, \dots, Q_{i-1}$ функция $k_i(Q_0, \dots, Q_i)$ будет зависеть только от Q_i , а функция $m_i(Q_0, \dots, Q_{i-1}; \gamma_i)$ будет функцией от γ_i , выполняя роль нормирующего множителя. Следует заметить, что условная плотность $h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i)$ может быть записана только для $Q_0, \dots, Q_{i-1}$ из V_{i-1} .

Поскольку критерий для проверки гипотезы $\gamma_i = \gamma_i^{(1)}$ имеет неймановскую структуру, мы можем применить фундаментальную лемму Неймана — Пирсона (упр. 10) к условному распределению величины Q_i . Наилучшим для проверки гипотезы $\gamma_i = \gamma_i^{(1)}$ против альтернативы $\gamma_i = \gamma_i^{(2)}$ оказывается критерий, имеющий критиче-

скую область

$$(9) \quad k < \frac{h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(2)})}{h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})} = \\ = \exp \left[\frac{1}{2} (\gamma_i^{(1)} - \gamma_i^{(2)}) Q_i \right] \frac{m_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})}{m_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(2)})}.$$

Константа k , зависящая от значений $Q_0, \dots, Q_{i-1}$, определяется так, чтобы условная вероятность события (8) при $\gamma_i = \gamma_i^{(1)}$ равнялась желаемому уровню значимости. Если $\gamma_i^{(1)} > \gamma_i^{(2)}$, то критическую область (9) можно записать в виде

$$(10) \quad Q_i > c_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}, \gamma_i^{(2)}).$$

В частности, такой вид будет иметь критерий для проверки гипотезы $\gamma_i = 0$ против альтернативы $\gamma_i = \gamma_i^{(2)} < 0$.

ТЕОРЕМА 6.3.3. *Наилучший подобный критерий для проверки нулевой гипотезы $\gamma_i = \gamma_i^{(1)}$ против альтернативы $\gamma_i = \gamma_i^{(2)} < \gamma_i^{(1)}$ с уровнем значимости ε_i имеет критическую область (10), где константа $c_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}, \gamma_i^{(2)})$ определяется так, чтобы при $\gamma_i = \gamma_i^{(1)}$ вероятность события (10), соответствующая плотности (8), была равна ε_i .*

Поскольку $c_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}, \gamma_i^{(2)})$ определяется из условной плотности (8) при $\gamma_i = \gamma_i^{(1)}$, то она не зависит от значения $\gamma_i^{(2)}$ и поэтому ее можно записать в виде $c_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})$. Таким образом, критическая область (10) не зависит от $\gamma_i^{(2)}$. Поэтому она определяет критерий, являющийся наилучшим против любой альтернативы $\gamma_i^{(2)} < \gamma_i^{(1)}$.

ТЕОРЕМА 6.3.4. *Равномерно наиболее мощный подобный критерий для проверки нулевой гипотезы $\gamma_i = \gamma_i^{(1)}$ против альтернатив $\gamma_i < \gamma_i^{(1)}$ с уровнем значимости ε_i имеет критическую область*

$$(11) \quad Q_i > c_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}).$$

Константа $c_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})$ определяется так, чтобы при $\gamma_i = \gamma_i^{(1)}$ вероятность события (11), вычисленная в соответствии с плотностью (8), равнялась ε_i .

Нас будет особенно интересовать случай $\gamma_i^{(1)} = 0$. В этом случае условная плотность Q_i при нулевой гипотезе равна

$$(12) \quad h_i(Q_i | Q_0, \dots, Q_{i-1}; 0) = \frac{k_i(Q_0, \dots, Q_i)}{k_{i-1}(Q_0, \dots, Q_{i-1})}.$$

Следствие 6.3.1. *Равномерно наиболее мощный подобный критерий для проверки нулевой гипотезы $\gamma_i = 0$ против альтернатив $\gamma_i < 0$ с уровнем значимости ε_i имеет критическую область*

$$(13) \quad Q_i > c_i(Q_0, \dots, Q_{i-1}).$$

Константа $c_i(Q_0, \dots, Q_{i-1})$ определяется так, чтобы интеграл от плотности (12) по множеству (13) был равен ε_i .

Когда мы перейдем к рассмотрению конкретных моделей, то увидим, что обычно критическая область (13) имеет вид

$$(14) \quad r_i > C_i(r_1, \dots, r_{i-1}),$$

где $r_1 = Q_1/Q_0, \dots, r_i = Q_i/Q_0$ — сериальные коэффициенты корреляции. В частности, при $i = 1$ критерий для проверки гипотезы о независимости против альтернатив о положительной зависимости имеет критическую область $r_1 > C_1$.

6.3.2. Равномерно наиболее мощные несмещенные критерии

При проверке гипотезы о том, что порядок зависимости меньше i (т. е. $\gamma_i = 0$), нас может интересовать альтернатива, состоящая в том, что порядок зависимости равен i (т. е. $\gamma_i \neq 0$). Эта альтернатива двусторонняя, а для двусторонних альтернатив не существует равномерно наиболее мощного критерия. Например, наилучший подобный критерий для проверки гипотезы $\gamma_i = \gamma_i^{(1)}$ против альтернативы $\gamma_i = \gamma_i^{(2)} < \gamma_i^{(1)}$ с уровнем значимости ε_i имеет критическую область (11). В то же время наилучший подобный критерий для проверки той же гипотезы, но при альтернативе $\gamma_i = \gamma_i^{(2)} > \gamma_i^{(1)}$ будет иметь критическую область

$$(15) \quad Q_i < c'_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}),$$

где $c'_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})$ определяется так, чтобы вероятность события (15), подсчитываемая согласно плотности (8), была равна ε_i при $\gamma_i = \gamma_i^{(1)}$. При рассмотрении двусторонних альтернатив потребуем, чтобы критерий был несмещенным, т. е. чтобы его функция мощности достигала минимума при $\gamma_i = \gamma_i^{(1)}$. В классе несмещенных критериев с данным уровнем значимости найдем равномерно наиболее мощный критерий. Такой критерий действительно будет существовать, поскольку семейство плотностей экспоненциально, и область принятия гипотезы будет интервалом. Опишем вывод равномерно наиболее мощного несмещенного критерия только в общих чертах. Полное и строгое доказательство приведено у Лемана (1959, гл. 4, разд. 4).

Для заданных $Q_0, \dots, Q_{i-1}$ пусть R будет критической областью с уровнем значимости ε_i , т. е.

$$(16) \quad \int_R h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) dQ_i = \varepsilon_i.$$

Пусть соответствующей функцией мощности будет $\beta(\gamma_i)$. При этом $\beta(\gamma_i^{(1)}) = \varepsilon_i$ и

$$(17) \quad \beta(\gamma_i) = \int_R \frac{k_i(Q_0, \dots, Q_i)}{m_i(Q_0, \dots, Q_{i-1}; \gamma_i)} e^{-\gamma_i Q_i / 2} dQ_i.$$

Для того чтобы критерий был локально несмещенным, надо приравнять нулю первую производную функции мощности в точке $\gamma_i = \gamma_i^{(1)}$. Дифференцируя под знаком интеграла, получаем

$$\begin{aligned} (18) \quad 0 &= \frac{d}{d\gamma_i} \beta(\gamma_i) \Big|_{\gamma_i = \gamma_i^{(1)}} = \\ &= -\frac{1}{2} \int_R Q_i \frac{k_i(Q_0, \dots, Q_i)}{m_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})} \exp\left(-\frac{\gamma_i^{(1)} Q_i}{2}\right) dQ_i + \\ &\quad + \frac{d}{d\gamma_i} \left[\frac{1}{m_i(Q_0, \dots, Q_{i-1}; \gamma_i)} \right] \Big|_{\gamma_i = \gamma_i^{(1)}} \times \\ &\quad \times \int_R k_i(Q_0, \dots, Q_i) \exp\left(-\frac{\gamma_i^{(1)} Q_i}{2}\right) dQ_i = \\ &= -\frac{1}{2} \int_R Q_i h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) dQ_i + \\ &\quad + \varepsilon_i m_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) \frac{d}{d\gamma_i} \left[\frac{1}{m_i(Q_0, \dots, Q_{i-1}; \gamma_i)} \right] \Big|_{\gamma_i = \gamma_i^{(1)}}. \end{aligned}$$

Дифференцируя подобным же образом

$$(19) \quad 1 = \int_{-\infty}^{\infty} \frac{k_i(Q_0, \dots, Q_i)}{m_i(Q_0, \dots, Q_{i-1}; \gamma_i)} e^{-\gamma_i Q_i / 2} dQ_i$$

в точке $\gamma_i = \gamma_i^{(1)}$ и сравнивая результат с (18), получаем соотношение

$$\begin{aligned} (20) \quad \int_R Q_i h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) dQ_i = \\ = \varepsilon_i \int_{-\infty}^{\infty} Q_i h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) dQ_i. \end{aligned}$$

Таким образом, (20) является условием локальной несмещенности критерия.

Мы хотим теперь найти такую критическую область R , чтобы она удовлетворяла соотношениям (16) и (20) и, кроме того, максимизировала функцию $\beta(\gamma_i)$ для некоторого частного значения γ_i , например для $\gamma_i = \gamma_i^{(2)} > \gamma_i^{(1)}$. В силу обобщенной фундаментальной леммы Неймана — Пирсона (упр. 12) такая область определяется неравенством

$$(21) \quad h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(2)}) > \\ > ah_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) + bQ_i h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}),$$

если существуют функции a и b от $Q_0, \dots, Q_{i-1}$, такие, что указанная область удовлетворяет (16) и (20). Неравенство (21) равносильно неравенству $g(Q_i) > 0$, где

$$(22) \quad g(Q_i) = \exp \left[-\frac{1}{2} (\gamma_i^{(2)} - \gamma_i^{(1)}) Q_i \right] - (A + BQ_i),$$

и

$$(23) \quad A = a \frac{m_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(2)})}{m_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})}, \quad B = b \frac{m_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(2)})}{m_i(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})}.$$

Поэтому

$$(24) \quad \frac{d}{dQ_i} g(Q_i) = -\frac{1}{2} (\gamma_i^{(2)} - \gamma_i^{(1)}) \exp \left[-\frac{1}{2} (\gamma_i^{(2)} - \gamma_i^{(1)}) Q_i \right] - B.$$

Если $B > 0$ (и $\gamma_i^{(2)} > \gamma_i^{(1)}$), то последняя производная отрицательна и неравенство (21) определяет некоторый полуинтервал $(-\infty, c)$. Однако такая критическая область не удовлетворяет соотношениям (16) и (20). (См. упр. 13.) Поэтому величину B следует брать отрицательной. Указанная производная будет непрерывно возрастающей функцией, принимающей нулевое значение только в одной точке. Эта точка будет точкой минимума функции $g(Q_i)$. Если этот минимум отрицателен (что имеет место, когда A достаточно велико), то (21) определяет два полуинтервала $(-\infty, c)$ и (d, ∞) . Область принятия в таком случае равна (c, d) . Отметим, что значения c и d определяются соотношениями (16) и (20), которые не зависят от $\gamma_i^{(2)}$. Подобным же образом при $\gamma_i^{(2)} < \gamma_i^{(1)}$ можно доказать, что константа B должна быть положительной и что область принятия будет интервалом, удовлетворяющим (16) и (20) независимо от значения $\gamma_i^{(2)}$. Таким образом, мы получили равномерно наиболее мощный несмещенный критерий.

ТЕОРЕМА 6.3.5. *Равномерно наиболее мощный критерий для проверки нулевой гипотезы $\gamma_i = \gamma_i^{(1)}$ против альтернативы $\gamma_i \neq \gamma_i^{(1)}$*

с уровнем значимости ε_i имеет критическую область

$$(25) \quad \begin{aligned} Q_i &> c_{U_i}(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) \\ Q_i &< c_{L_i}(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}), \end{aligned}$$

где $c_{U_i}(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})$ и $c_{L_i}(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})$ определяются так, чтобы

$$(26) \quad \int_{c_{L_i}(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})}^{c_{U_i}(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})} h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) dQ_i = 1 - \varepsilon_i$$

и

$$(27) \quad \begin{aligned} &\int_{c_{L_i}(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})}^{c_{U_i}(Q_0, \dots, Q_{i-1}; \gamma_i^{(1)})} Q_i h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) dQ_i = \\ &= (1 - \varepsilon_i) \int_{-\infty}^{\infty} Q_i h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) dQ_i. \end{aligned}$$

СЛЕДСТВИЕ 6.3.2. *Равномерно наиболее мощный критерий для проверки гипотезы $\gamma_i = 0$ против альтернатив $\gamma_i \neq 0$ с уровнем значимости ε_i имеет критическую область*

$$(28) \quad \begin{aligned} Q_i &> c_{U_i}(Q_0, \dots, Q_{i-1}), \\ Q_i &< c_{L_i}(Q_0, \dots, Q_{i-1}), \end{aligned}$$

где $c_{U_i}(Q_0, \dots, Q_{i-1})$ и $c_{L_i}(Q_0, \dots, Q_{i-1})$ определяются соотношениями

$$(29) \quad \int_{c_{L_i}(Q_0, \dots, Q_{i-1})}^{c_{U_i}(Q_0, \dots, Q_{i-1})} \frac{k_i(Q_0, \dots, Q_i)}{k_{i-1}(Q_0, \dots, Q_{i-1})} dQ_i = 1 - \varepsilon_i$$

и

$$(30) \quad \begin{aligned} &\int_{c_{L_i}(Q_0, \dots, Q_{i-1})}^{c_{U_i}(Q_0, \dots, Q_{i-1})} Q_i \frac{k_i(Q_0, \dots, Q_i)}{k_{i-1}(Q_0, \dots, Q_{i-1})} dQ_i = \\ &= (1 - \varepsilon_i) \int_{-\infty}^{\infty} Q_i \frac{k_i(Q_0, \dots, Q_i)}{k_{i-1}(Q_0, \dots, Q_{i-1})} dQ_i. \end{aligned}$$

Если условная плотность распределения величины Q_i при заданных $Q_0, \dots, Q_{i-1}$ для $\gamma_i = 0$ симметрична, то (30) автоматически

будет выполняться для

$$(31) \quad c_{L_i}(Q_0, \dots, Q_{i-1}) = -c_{U_i}(Q_0, \dots, Q_{i-1}),$$

поскольку тогда обе части (30) будут равны нулю. При этом (29) можно записать в виде

$$(32) \quad \int_{c_{U_i}(Q_0, \dots, Q_{i-1})}^{\infty} \frac{k_i(Q_0, \dots, Q_i)}{k_{i-1}(Q_0, \dots, Q_{i-1})} dQ_i = \frac{1}{2} \varepsilon_i.$$

Равномерно наиболее мощный односторонний подобный критерий можно называть также равномерно наиболее мощным односторонним несмещенным критерием, поскольку из несмещенности критерия вытекает подобие критической области.

Следует отметить, что условное распределение величины Q_i при заданных $Q_0, \dots, Q_{i-1}$ не зависит от значений параметров $\gamma_0, \dots, \gamma_{i-1}$. Ввиду этого последним можно приписать любые значения, которые окажутся более удобными. Например, можно взять $\gamma_0 = 1, \gamma_1 = \dots = \gamma_{i-1} = 0$.

Обычно Q_0 будет равно $\sum_{i=1}^T y_i^2$. Отношение

$$(33) \quad r_i = \frac{Q_i}{Q_0}$$

можно рассматривать как *сериальный коэффициент корреляции*. Если $\gamma_1 = \dots = \gamma_i = 0$, то совместное распределение $r_1, \dots, r_i$ не зависит от $\gamma_0 (> 0)$ и Q_0 . (См. теорему 6.7.2.) Оптимальные критерии (13) и (28) могут быть определены с использованием условного распределения $r_i = Q_i/Q_0$ при заданных $r_1, \dots, r_{i-1}$.

Частную сериальную корреляцию порядка i при неизменных сериальных корреляциях порядков $1, \dots, i-1$ можно определить обычным способом. Мы уже говорили о том, что этот частный сериальный коэффициент корреляции может быть использован для проверки гипотезы $\gamma_i = 0$. Поскольку он является линейной функцией аргумента r_i (с коэффициентами, зависящими от $r_1, \dots, r_{i-1}$), то использование его при заданных значениях $r_1, \dots, r_{i-1}$ эквивалентно использованию самого r_i при тех же $r_1, \dots, r_{i-1}$. (См. разд. 6.9.2.)

6.3.3. Случай, когда равномерно наиболее мощного критерия не существует

Имеется целый ряд задач сериальной корреляции, решения которых нельзя получить из теорем 6.3.4. и 6.3.5. Если бы мы хотели проверить нулевую гипотезу

$$(34) \quad \gamma_1 = \dots = \gamma_p = 0$$

против альтернативы $\gamma_1 = \gamma_1^0, \dots, \gamma_p = \gamma_p^0$, где $\gamma_1^0, \dots, \gamma_p^0$ — заданные числа, то равномерно наиболее мощный подобный критерий мог бы иметь критическую область

$$(35) \quad \sum_{j=1}^p \gamma_j^0 Q_j < c(Q_0),$$

где константа $c(Q_0)$ подбирается так, чтобы условная вероятность события (35) при нулевой гипотезе была равна ε . Если альтернативные значения параметров $\gamma_1, \dots, \gamma_p$ имеют вид

$$(36) \quad \gamma_j = k\gamma_j^*, \quad j = 1, \dots, p,$$

где γ_j^* фиксированы, то критическая область (35), с γ_j^0 замененными на γ_j^* , будет наилучшей для всех альтернатив $k > 0$. Однако, если альтернативные значения не лежат на прямой (36) в пространстве параметров, область (35) не будет одной и той же для всех допустимых альтернатив $(\gamma_1^0, \dots, \gamma_p^0)$. Но это и означает, что в данном случае не существует равномерно наиболее мощного критерия.

Рассмотрим стационарный случайный процесс, определяемый соотношением

$$(37) \quad y_t + \beta y_{t-1} = u_t, \quad t = \dots, -1, 0, 1, \dots,$$

в котором $|\beta| < 1$, а каждое u_t распределено нормально, с нулевым средним и дисперсией σ^2 , и не зависит от $y_{t-1}, y_{t-2}, \dots$ и $u_{t-1}, u_{t-2}, \dots$. При этом маргинальное распределение выборки $y_1, \dots, y_T$ задается выражением (1), в котором

$$A_0 = I, \\ A_1 = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix},$$

$$(38) \quad A_2 = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

Множество альтернатив для $\gamma_1 = 2\beta/\sigma^2$ и $\gamma_2 = \beta^2/\sigma^2$ не лежит на прямой. Поэтому здесь не существует равномерно наиболее мощного критерия даже для односторонних альтернатив.

Если $|\beta|$ мало (т. е. близко к нулю), этот случай подобен модели разд. 6.5.4. Если β близко к -1 , такой случай подобен модели разд. 6.5.3. Следовательно, можно предположить, что для проверки гипотезы $\beta = 0$ в этой модели (когда среднее значение известно) полезной статистикой должны оказаться одна или две соответствующие корреляции. Положительная зависимость соответствует $\beta < 0$.

Пусть (37) определяет распределение для $t = 1, \dots, T$ при условии $y_0 \equiv 0$. Тогда A_0 и A_1 будут такими же, как и в предыдущем случае, а

$$(39) \quad A_2 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & & \vdots & \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

Множество альтернатив для $\gamma_1 = 2\beta/\sigma^2$ и $\gamma_2 = \beta^2/\sigma^2$ не лежит на прямой, так что и здесь для проверки гипотезы $\beta = 0$ не существует равномерно наиболее мощного критерия.

Модель, определяемая разностным уравнением второго порядка

$$(40) \quad y_t + \beta_1 y_{t-1} + \beta_2 y_{t-2} = u_t, \quad t = 1, \dots, T,$$

с циклическим определением $y_0 \equiv y_T$, $y_{-1} \equiv y_{T-1}$, приводит к плотности (1) с

$$A_0 = I,$$

$$A_1 = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 1 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & & \vdots & \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 1 & 0 \end{bmatrix},$$

$$(41) \quad A_2 = \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 1 \\ 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & & \vdots & \\ 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

Если альтернативные гипотезы не связывают специальным образом β_1 и β_2 , то равномерно наиболее мощного критерия для проверки гипотезы $\beta_1 = \beta_2 = 0$ не существует.

Теория, изложенная в этом параграфе, была развита Т. Андерсоном (1948), (1963).

6.4. ВЫБОР ПОРЯДКА ЗАВИСИМОСТИ КАК ЗАДАЧА СО МНОГИМИ РЕШЕНИЯМИ

Займемся теперь вопросом об определении порядка зависимости на основании наблюдений $y_1, \dots, y_T$. Будем предполагать, что исследователь может определить некоторый наименьший m и некоторый наибольший p — возможные порядки зависимости. Во многих случаях m оказывается равным нулю (указывая тем самым на независимость). В других случаях исследователь может в силу каких-нибудь причин ожидать, что порядок зависимости не меньше некоторого положительного числа m . Ввиду того что имеющиеся наблюдения удалены друг от друга не более чем на $T - 1$ единиц времени, ясно, что максимальный порядок p не может быть больше чем $T - 1$. Однако ясно также, что для того, чтобы процедура обладала хорошими свойствами в отношении мощности, максимальный порядок p должен быть выбран по возможности много меньшим $T - 1$.

Исследователь имеет, таким образом, задачу со многими решениями. Именно, ему необходимо выбрать порядок $m, m + 1, \dots, p - 1$ или p . Эта задача подобна задаче выбора степени полинома, представляющего тренд, изучавшейся в разд. 3.2.2. Хотя распределения здесь уже другие, общая структура задачи и ее решение будут теми же, что и в разд. 3.2.2. Мы формализуем задачу, стоящую перед исследователем, считая, что ему необходимо решить, какому из перечисленных ниже взаимно не пересекающихся множеств принадлежит параметрическая точка $(\gamma_{m+1}, \dots, \gamma_p)$:

$$\begin{aligned}
 &H_p: \gamma_p \neq 0, \\
 &H_{p-1}: \gamma_p = 0, \quad \gamma_{p-1} \neq 0, \\
 &\vdots \\
 &H_{m+1}: \gamma_p = \dots = \gamma_{m+2} = 0, \quad \gamma_{m+1} \neq 0, \\
 &H_m: \gamma_p = \dots = \gamma_{m+1} = 0.
 \end{aligned}
 \tag{1}$$

Принадлежность параметрической точки множеству H_i означает, что порядок зависимости равен i . Альтернативная формулировка состоит в том, что исследователю необходимо решить, верна ли хотя бы одна и (если да) какая именно из следующих нулевых

гипотез:

$$\begin{aligned}
 &H_p^* : \gamma_p = 0, \\
 &H_{p,p-1}^* : \gamma_p = \gamma_{p-1} = 0, \\
 &\vdots \\
 &H_{p,p-1, \dots, m+1}^* : \gamma_p = \dots = \gamma_{m+1} = 0.
 \end{aligned}
 \tag{2}$$

Если верна какая-нибудь из гипотез (2), то будут верны и все ей предшествующие. Если же некоторая из гипотез (2) ложна, то будут ложными и все последующие. Иными словами,

$$H_{p,p-1, \dots, m+1}^* \subset \dots \subset H_p^*.$$

Эти два семейства множеств связаны соотношениями

$$\begin{aligned}
 H_p^* &= H_{p-1} \cup \dots \cup H_m, \\
 H_{p,p-1}^* &= H_{p-2} \cup \dots \cup H_m, \\
 &\vdots \\
 H_{p,p-1, \dots, m+1}^* &= H_m.
 \end{aligned}
 \tag{4}$$

Предположим, что исследователь намерен непосредственно контролировать вероятности ошибок, возникающих при неправильном решении о том, что коэффициенты отличны от нуля, когда они в действительности равны нулю, т. е. вероятности завышения порядка зависимости. В связи с этим будем считать, что исследователь приписывает некоторый уровень значимости каждой из нулевых гипотез:

$$\begin{aligned}
 p_p &= \text{Pr} \{ \text{отвергнуть } H_p^* | H_p^* \}, \\
 p_p + p_{p-1} &= \text{Pr} \{ \text{отвергнуть } H_{p,p-1}^* | H_{p,p-1}^* \}, \\
 &\vdots \\
 p_p + \dots + p_{m+1} &= \text{Pr} \{ \text{отвергнуть } H_{p,p-1, \dots, m+1}^* | H_{p,p-1, \dots, m+1}^* \} = \\
 &= \text{Pr} \{ \text{отвергнуть } H_m | H_m \},
 \end{aligned}
 \tag{5}$$

причем $p_i \geq 0$ и $p_p + \dots + p_{m+1} \leq 1$. Поскольку здесь каждая нулевая гипотеза включает в себя последующую (т. е. каждая последующая гипотеза является более сильной), вероятности отклонения берутся монотонно неубывающими (т. е. вероятность отклонения более сильной нулевой гипотезы, когда она верна, не меньше аналогичной вероятности для менее сильной гипотезы). С помощью взаимно непересекающихся множеств, указанное разнесение по уровням значимости записывается в виде

$$\begin{aligned}
 p_p &= \text{Pr} \{ \text{принять } H_p | H_{p-1} \cup \dots \cup H_m^* \}, \\
 p_{p-1} &= \text{Pr} \{ \text{принять } H_{p-1} | H_{p-2} \cup \dots \cup H_m \}, \\
 &\vdots \\
 p_{m+1} &= \text{Pr} \{ \text{принять } H_{m+1} | H_m \}.
 \end{aligned}
 \tag{6}$$

Пусть $p_m = 1 - p_q - \dots - p_{m+1} = \text{Pr} \{ \text{принять } H_m | H_m \}$. Посредством соотношений (6) исследователь приписывает определенные значения вероятностям решений о том, что порядок зависимости равен i , когда в действительности этот порядок меньше i , $i = m+1, \dots, p$.

Статистическая процедура для этой задачи со многими решениями состоит в следующем. Имеется набор $p - m + 1$ попарно не пересекающихся, составляющих полную группу областей в пространстве значений $Q_0, \dots, Q_p$ (или в исходном пространстве значений $y_1, \dots, y_T$). Мы обозначим эти области $R_m, R_{m+1}, \dots, R_p$. Если выборочная точка попадает в R_i , то принимается гипотеза H_i . Приписывание уровней значимости приводит к тому, что эти области становятся подобными в том смысле, что при $\gamma_i = \dots = \gamma_p = 0$ вероятности попадания выборочной точки в $R_i, \dots, R_p$ равны соответственно $p_i, \dots, p_p$ (независимо от $\gamma_0, \gamma_1, \dots, \gamma_{i-1}$). Иными словами, если порядок зависимости меньше i , то вероятность ошибок от приписывания зависимости порядка i не зависит от того, каков именно истинный (более низкий) порядок.

Рассмотрим сначала альтернативы, утверждающие факт положительной зависимости. Это соответствует значениям $\gamma_i < 0$. (Позднее мы рассмотрим несмещенные процедуры для $\gamma_i \neq 0$.)

Зафиксировав указанные ограничения, потребуем, чтобы области R_i были наилучшими в том смысле, чтобы были максимально возможными вероятности попадания в R_i при условии, что верна H_i , $i = m+1, \dots, p$. Мы одновременно пытаемся максимизировать вероятности попадания в $p - m$ различных областей (каждую для всех отрицательных значений соответствующего параметра). Далее будет показано, что в указанных выше условиях при подборе одной из областей с целью максимизации вероятности попадания в эту область не имеет значения то, как выбираются остальные области. Этот факт позволяет оптимизировать области $R_{m+1}, \dots, R_p$ одновременно.

Ограничения (5) или эквивалентные им ограничения (6) приводят к тому, что каждое множество R_i или, что равносильно, каждое множество $R_i \cup R_{i+1} \cup \dots \cup R_p$ — область отклонения гипотезы $H_{p,p-1,\dots,i}^*$ — имеет (ввиду свойств достаточности и полноты) неймановскую структуру. Иначе говоря, при $i = m+1, \dots, p$,

$$\text{Pr} \{ R_i | Q_0, \dots, Q_{i-1}; \gamma_i = \dots = \gamma_p = 0 \} = p_i
 \tag{7}$$

для почти всех возможных значений $Q_0, \dots, Q_{i-1}$. Исходя из этого можно показать, что выбор $R_p, \dots, R_{i+1}$ (при допущении (6)) не влияет на выбор R_i , т. е. что вероятность попадания в R_i (являющаяся функцией γ_i) при $\gamma_{i+1} = \dots = \gamma_p = 0$ не зависит от того, какими были выбраны $R_p, \dots, R_{i+1}$. Отметим, что значение γ_i интересует нас только в том случае, когда $\gamma_{i+1} = \dots = \gamma_p = 0$. Если же какое-нибудь из $\gamma_{i+1}, \dots, \gamma_p$ не равно нулю (т. е. если порядок зависимости больше i), то вопроса о значении γ_i не возникает и мы можем считать $\gamma_i \neq 0$.

ЛЕММА 6.4.1. Пусть множество S_{i+1} в пространстве значений $Q_0, \dots, Q_p$ таково, что

$$(8) \quad \Pr \{S_{i+1} | \gamma_{i+1} = \dots = \gamma_p = 0\} = p_p + \dots + p_{i+1},$$

и пусть T_i — множество, определяемое значениями $Q_0, \dots, Q_i$. Тогда

$$(9) \quad \Pr \{\bar{S}_{i+1} \cap T_i | \gamma_{i+1} = \dots = \gamma_p = 0\} = \\ = (1 - p_p - \dots - p_{i+1}) \Pr \{T_i | \gamma_{i+1} = \dots = \gamma_p = 0\},$$

где $\bar{S}_{i+1}$ — множество, дополнительное к S_{i+1} .

Смысл леммы состоит в следующем. Каким бы образом мы ни выбирали $R_p, \dots, R_{i+1}$ при условии (6), из которого следует (8) для $S_{i+1} = R_p \cup \dots \cup R_{i+1}$, вероятность попадания в область R_i , определенную как пересечение $\bar{S}_{i+1} \cap T_i$, будет зависеть только от T_i и не будет зависеть от S_{i+1} (когда верна гипотеза $H_{p,p-1, \dots, i+1}^*$). Доказательство леммы проведено в разд. 3.2.2 с использованием коэффициентов регрессии. Там же дано подобное доказательство сформулированной ниже леммы 6.4.2. Оба эти доказательства опираются только на свойства полноты и достаточности. При этом соответствующую терминологию и обозначения можно заменить терминологией и обозначениями настоящего параграфа.

Пусть T_i^* — область, определяемая неравенством

$$(10) \quad Q_i > c_i(Q_0, \dots, Q_{i-1}),$$

в котором $c_i(Q_0, \dots, Q_{i-1})$ выбирается так, чтобы условная вероятность (10) при заданных $Q_0, \dots, Q_{i-1}$ равнялась e_i для $\gamma_i = 0$, и пусть $e_i = p_i / (1 - p_p - \dots - p_{i+1})$.

ЛЕММА 6.4.2. Пусть S_{i+1} удовлетворяет соотношению (8), а R_i — произвольное не пересекающееся с S_{i+1} множество, для которого

$$(11) \quad \Pr \{R_i | \gamma_i = \gamma_{i+1} = \dots = \gamma_p = 0\} = p_i.$$

Тогда при $\gamma_i < 0$

$$(12) \quad \Pr \{ \bar{S}_{i+1} \cap T_i^* | \gamma_{i+1} = \dots = \gamma_p = 0 \} \geq \\ \geq \Pr \{ R_i | \gamma_{i+1} = \dots = \gamma_p = 0 \}.$$

Из приведенных двух лемм вытекает, что, какими бы ни были области $R_p, \dots, R_{i+1}$, наилучший выбор области R_i состоит в том, что R_i должно являться частью T_i^* , не содержащей точек $R_p, \dots, R_{i+1}$. При таком выборе R_i вероятность

$$(13) \quad \Pr \{ R_i | \gamma_{i+1} = \dots = \gamma_p = 0 \} = \\ = (1 - p_p - \dots - p_{i+1}) \Pr \{ T_i^* | \gamma_{i+1} = \dots = \gamma_p = 0 \}$$

не зависит от выбора $R_p, \dots, R_{i+1}$.

ТЕОРЕМА 6.4.1. Пусть непересекающиеся множества $R_m, R_{m+1}, \dots, R_p$ выборочного пространства таковы, что $R_m \cup R_{m+1} \cup \dots \cup R_p$ есть все это пространство и

$$(14) \quad \Pr \{ R_i | \gamma_i = \gamma_{i+1} = \dots = \gamma_p = 0 \} = p_i, \quad i = m+1, \dots, p$$

где $p_{m+1} + \dots + p_p \leq 1$. Тогда для каждого значения $\gamma_i < 0$ вероятность $\Pr \{ R_i | \gamma_i, \gamma_{i+1} = \dots = \gamma_p = 0 \}$ принимает максимальное значение на множестве R_i , определяемом как пересечение множества (10) при $e_i = p_i / (1 - p_p - \dots - p_{i+1})$ и дополнения к множеству $R_p \cup \dots \cup R_{i+1}$, $i = m+1, \dots, p$.

Оптимальной, таким образом, является следующая процедура

$$(15) \quad \begin{aligned} R_p &= T_p^* \\ R_{p-1} &= \bar{T}_p^* \cap T_{p-1}^*, \\ &\vdots \\ R_i &= \bar{T}_p^* \cap \bar{T}_{p-1}^* \cap \dots \cap \bar{T}_{i+1}^* \cap T_i^*, \\ &\vdots \\ R_{m+1} &= \bar{T}_p^* \cap \bar{T}_{p-1}^* \cap \dots \cap \bar{T}_{m+2}^* \cap T_{m+1}^*, \\ R_m &= \bar{T}_p^* \cap \bar{T}_{p-1}^* \cap \dots \cap \bar{T}_{m+2}^* \cap \bar{T}_{m+1}^*. \end{aligned}$$

Эта процедура сводится, по существу, к следующему. Поочередно проверяются гипотезы $\gamma_p = 0$, $\gamma_{p-1} = 0$ и т. д. до тех пор пока либо какая-то из гипотез, например $\gamma_i = 0$, не будет отклонена и будет решено, что верна H_i , либо будут приняты все гипотезы вплоть до H_m . Таким образом, указанная процедура является последовательностью критериев. Этот факт обусловлен требованием независимости от истинного порядка вероятности правильного решения о том, что порядок зависимости меньше данного натурального числа.

Рассмотрим теперь процедуры, использующие двусторонние критерии. При этом будем накладывать требование несмещенности.

ТЕОРЕМА 6.4.2. Пусть $R_m, R_{m+1}, \dots, R_p$ — непересекающиеся множества выборочного пространства, объединение которых $R_m \cup R_{m+1} \cup \dots \cup R_p$ совпадает со всем этим пространством и

$$(16) \quad \begin{aligned} \Pr \{R_i \cup \dots \cup R_p | \gamma_i = \gamma_{i+1} = \dots = \gamma_p = 0\} &\leq \\ &\leq p_i + \dots + p_p, \\ \Pr \{R_i \cup \dots \cup R_p | \gamma_i = \gamma_{i+1} = \dots = \gamma_p = 0\} &\leq \\ &\leq \Pr \{R_i \cup \dots \cup R_p | \gamma_i, \gamma_{i+1} = \dots = \gamma_p = 0\}, \\ &i = m+1, \dots, p, \end{aligned}$$

где $p_{m+1} + \dots + p_p \leq 1$. Тогда для любого значения γ_i вероятность $\Pr \{R_i | \gamma_i, \gamma_{i+1} = \dots = \gamma_p = 0\}$ будет максимальной на множестве R_i , представляющем собой пересечение множества, указанного в следствии 6.3.2, для $\varepsilon_i = p_i / (1 - p_p - \dots - p_{i+1})$, и дополнения к множеству $R_{i+1} \cup \dots \cup R_p$, $i = m+1, \dots, p$.

Мы не говорили еще о том, как следует выбирать вероятности $p_m, p_{m+1}, \dots, p_p$. Если все ε_i фиксированы и равны, например, ε , то $p_p = \varepsilon$, $p_{p-1} = \varepsilon(1 - \varepsilon)$, $\dots$, $p_i = \varepsilon(1 - \varepsilon)^{p-i}$, $\dots$, $p_{m+1} = \varepsilon(1 - \varepsilon)^{p-m-1}$, $p_m = (1 - \varepsilon)^{p-m}$. Вообще говоря, приходится идти на компромисс между нашими стремлениями не переоценить порядок зависимости и сохранить чувствительность процедуры к ненулевым коэффициентам. Альтернатива состоит здесь в выборе $p_p < p_{p-1} < \dots < p_{m+1}$. Причем, если для больших значений i значения p_i задавать весьма малыми, можно позволить себе в качестве максимального порядка p брать достаточно большое число.

Результаты этого параграфа приводят к критерию для проверки гипотезы

$$(17) \quad H: \gamma_{m+1} = \dots = \gamma_p = 0.$$

При этом процедуры, указанные в теоремах 6.4.1 и 6.4.2, можно интерпретировать как критерии с уровнем значимости $p_{m+1} + \dots + p_p$, согласно которым гипотеза (17) отклоняется, если выборочная точка попадает в любую область, отличную от R_m .

Процедура, указанная теоремой 6.4.1, состоит из односторонних критериев для проверки гипотез против односторонних альтернатив, утверждающих факт положительной зависимости. Если альтернативой к независимости служит отрицательная зависимость ($\gamma_i > 0$), в этом случае также употребляется процедура, аналогичная предыдущей, но состоящая уже из односторонних критериев вида (15) § 6.3. Кроме того, можно сконструировать составные

процедуры для случая, когда в качестве альтернативы к гипотезе $\gamma_i = 0$ выступает гипотеза $\gamma_i < 0$, $\gamma_i > 0$ или $\gamma_i \neq 0$, $i = m + 1, \dots, p$.

Замечания, сделанные в разд. 3.2.2 относительно такого типа процедур со многими решениями, используемых для определения степени полиномиального тренда, применимы и в настоящем случае. С теоретической точки зрения преимущество этой процедуры состоит в том, что исследователь контролирует здесь вероятность ошибочного завышения порядка зависимости. Неудобством ее является то, что максимальный порядок зависимости p должен быть установлен заранее. Частично это неудобство можно обойти, полагая значение p весьма большим, но выбирая зато вероятности p_i очень малыми для i , близких к p . Иными словами, устанавливаются весьма малыми значения вероятностей ошибочного решения о том, что порядок зависимости равен i , если это i велико.

Альтернативой к указанной последовательной процедуре является процедура, использующая критерии значимости в обратном порядке. Если считается, что порядок зависимости не меньше m , то исследователь может проверить сначала гипотезу $\gamma_{m+1} = 0$ с уровнем значимости ε_{m+1} . Если эта гипотеза будет принята, то выносится решение о том, что порядок зависимости равен m . В противном случае проверяется гипотеза $\gamma_{m+2} = 0$. Эту процедуру можно продолжать таким образом и дальше, используя указанную последовательность критериев значимости. Если впервые будет принята гипотеза $\gamma_i = 0$, то порядок зависимости принимается равным $i - 1$. Если гипотеза $\gamma_i = 0$ отвергается, то переходят к проверке гипотезы $\gamma_{i+1} = 0$. Если при этом отвергаются все гипотезы вплоть до гипотезы $\gamma_p = 0$ с номером p , заданным заранее, то порядок зависимости считается равным p . Преимущество этой процедуры состоит в том, что не всегда требуется вычислять все $p + 1$ квадратичных форм, поскольку последовательность критериев может оборваться прежде, чем придется проверять гипотезу $\gamma_p = 0$. Ввиду этого вопрос о том, как следует выбирать здесь p , является несущественным.

В разд. 3.2.2 было указано, что соответствующая процедура для определения степени полиномиального тренда имеет определенное практическое неудобство. Это неудобство состоит в том, что если коэффициент при степени i имеет достаточно большое значение, то возникает тенденция к завышению оценок дисперсии ошибки, используемых для проверки $\gamma_{m+1}, \dots, \gamma_{i-1}$. Это в свою очередь ведет к возрастанию вероятности принятия одной из гипотез $\gamma_{m+1} = 0, \dots, \gamma_{i-1} = 0$. В результате, может увеличиться вероятность ошибочного занижения степени полинома. Фактически чем больше γ_i , тем больше и вероятность такой ошибки (если считать остальные γ_j фиксированными). Этот случай проверки порядка зависимости представляется более трудным для анализа. Если зави-

симось соседних наблюдений, отстоящих на две единицы времени, велика, то вполне может оказаться, что вероятность отклонения гипотезы $\gamma_1 = 0$ не будет велика. Так что при таком подходе может быть ошибочно решено, что порядок зависимости равен единице, в то время как на самом деле он не меньше двух.

По-видимому, такую процедуру можно использовать, если для некоторого частного значения i значения параметров $\gamma_{m+1}, \dots, \gamma_{i-1}$ будут достаточны для того, чтобы нельзя было решить, что хотя бы один из них будет равен нулю без того, чтобы $\gamma_i \neq 0$. Например, такая процедура может оказаться подходящей для случая, когда значения параметров γ_i монотонно убывают (или просто быстро убывают). Однако никакого теоретического исследования оптимальных свойств здесь не имеется.

До сих пор в этой главе мы предполагали, что $\xi y_1 = \dots = \xi y_T = 0$. Если средние отличны от нуля, но известны, то при тех же предположениях в качестве объекта изучения можно рассматривать разности переменных и их математических ожиданий. Знание средних является, однако, предположением, далеким от реальности. Мы рассмотрим случай неизвестных средних в § 6.6, после того как обсудим более конкретные модели в § 6.5. (См. Т. Андерсон (1963).)

6.5. МОДЕЛИ: СИСТЕМЫ КВАДРАТИЧНЫХ ФОРМ

6.5.1. Общее обсуждение

Плотности распределений случайных величин $y_1, \dots, y_T$, составляющих наблюдаемые временные ряды, которые мы здесь рассматриваем, предполагаются нормальными. В показателе экспоненты этих плотностей будут стоять суммы — $\left(\sum_{j=0}^p \gamma_j Q_j\right)/2$, где $Q_j =$

$= y' A_j y$ — квадратичная форма по $y = (y_1, \dots, y_T)'$. Серийные коэффициенты корреляции, получаемые из таких сумм, равны $r_j = Q_j/Q_0$. В этом параграфе будет предложено несколько различных систем матриц $A_0, A_1, \dots, A_p$. В каждой из них Q_j будет близка к $\sum_i y_i y_{i-j}$, как это предполагалось в § 6.2. Коэффициенты $r_1, \dots, r_p$ в каждой из этих систем можно использовать для проверки гипотез о порядке зависимости, в соответствии с теорией § 6.3, и для выбора порядка зависимости согласно § 6.4.

Для того чтобы быть подготовленными к анализу указанных систем матриц, рассмотрим сначала множества матриц в общем плане. Характеристические корни $\lambda_1, \dots, \lambda_T$ и соответствующие им характеристические векторы $v_1, \dots, v_T$ матрицы A размера $T \times T$

удовлетворяют соотношениям

$$(1) \quad \mathbf{A}\mathbf{v}_s = \lambda_s \mathbf{v}_s, \quad s = 1, \dots, T.$$

Поскольку эти соотношения можно записать в виде $(\mathbf{A} - \lambda_s \mathbf{I}) \mathbf{v}_s = 0$ и при этом $\mathbf{v}_s \neq 0$, то матрица $\mathbf{A} - \lambda_s \mathbf{I}$ должна быть вырожденной. Иными словами, λ_s является корнем уравнения

$$(2) \quad |\mathbf{A} - \lambda \mathbf{I}| = 0.$$

Если матрица $\mathbf{A}$ симметрична (а это мы предполагаем для квадратичных форм $\mathbf{y}'\mathbf{A}\mathbf{y}$), то все корни (2) и компоненты характеристических векторов будут вещественными. Если все эти корни различны, то характеристические векторы ортогональны. Если m корней совпадают и равны, скажем, λ , то при этом существует m линейно независимых решений уравнения $(\mathbf{A} - \lambda \mathbf{I}) \mathbf{v} = 0$. Каждое такое множество может служить совокупностью характеристических векторов, соответствующих m -кратному корню λ . Удобно выбирать это множество таким образом, чтобы входящие в него векторы были взаимно ортогональны. При этом они необходимо будут ортогональны и остальным характеристическим векторам. Если каждый характеристический вектор нормировать так, чтобы сумма квадратов его компонент равнялась 1, то получим соотношения

$$(3) \quad \mathbf{v}_s' \mathbf{v}_t = \delta_{st}, \quad s, t = 1, \dots, T,$$

где $\delta_{ss} = 1$ и $\delta_{st} = 0$ при $s \neq t$.

Положим

$$(4) \quad \mathbf{A} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_T \end{bmatrix}, \quad \mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_T).$$

Тогда (1) и (3) можно записать соответственно в виде

$$(5) \quad \mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Lambda},$$

$$(6) \quad \mathbf{V}'\mathbf{V} = \mathbf{I}.$$

Если умножить обе части (5) слева на матрицу $\mathbf{V}'$, то придем к соотношению

$$(7) \quad \mathbf{V}'\mathbf{A}\mathbf{V} = \mathbf{\Lambda}.$$

Иными словами, ортогональная матрица $\mathbf{V}$ диагонализует симметрическую матрицу $\mathbf{A}$. При этом квадратичная форма $\mathbf{y}'\mathbf{A}\mathbf{y}$ приводится с помощью преобразования $\mathbf{y} = \mathbf{V}\mathbf{x}$ к диагональному виду $\mathbf{x}'\mathbf{\Lambda}\mathbf{x}$. Мы используем эти преобразования для того, чтобы упростить задачу отыскания распределений квадратичных форм и сериальных коэффициентов корреляции.

Нас будут интересовать системы матриц $A_0, A_1, \dots, A_p$. Обычно $A_0 = I$. При этом любой ненулевой вектор является характеристическим для I и любая ортогональная матрица V диагонализует I , так как $V'IV = I$. Оказывается весьма удобным, чтобы все A_j приводились к диагональному виду с помощью одной и той же матрицы V , т. е. чтобы

$$(8) \quad V'A_jV = \Lambda_j, \quad j = 1, \dots, p.$$

Это будет иметь место, если $v_1, \dots, v_T$ будут характеристическими корнями для каждой из матриц A_j . В связи с этим полезна следующая теорема.

ТЕОРЕМА 6.5.1. Если $\lambda_1, \dots, \lambda_T$ — характеристические корни, а $v_1, \dots, v_T$ — характеристические векторы матрицы A и если $P(A)$ полином от A , то характеристическими корнями и векторами матрицы $P(A)$ будут соответственно $P(\lambda_1), \dots, P(\lambda_T)$ и $v_1, \dots, v_T$.

Для доказательства этой теоремы мы используем следующие две леммы.

ЛЕММА 6.5.1. Пусть $\lambda_1, \dots, \lambda_T$ и $v_1, \dots, v_T$ — соответственно характеристические корни и характеристические векторы матрицы A . Тогда $\lambda_1^g, \dots, \lambda_T^g$ и $v_1, \dots, v_T$ будут соответственно характеристическими корнями и характеристическими векторами матрицы A^g , $g = 0, 1, \dots$.

ДОКАЗАТЕЛЬСТВО ЛЕММЫ Мы будем доказывать эту лемму по индукции. Прежде всего очевидно, что она справедлива для $g = 1$. Предположим, что она справедлива для $g = G - 1$. Тогда

$$(9) \quad \begin{aligned} A^G v_s &= A(A^{G-1} v_s) = A \lambda_s^{G-1} v_s = \\ &= \lambda_s^{G-1} (A v_s) = \lambda_s^{G-1} \lambda_s v_s = \\ &= \lambda_s^G v_s, \end{aligned} \quad s = 1, \dots, T. \blacksquare$$

ЛЕММА 6.5.2. Пусть $\lambda_{h1}, \dots, \lambda_{hT}$ и $v_1, \dots, v_T$ — соответственно характеристические корни и характеристические векторы матрицы A_h , $h = 0, \dots, H$. Тогда характеристическими корнями и характеристическими векторами матрицы $\sum_{h=0}^H a_h A_h$ будут соответственно $\sum_{h=0}^H a_h \lambda_{h1}, \dots, \sum_{h=0}^H a_h \lambda_{hT}$ и $v_1, \dots, v_T$.

ДОКАЗАТЕЛЬСТВО ЛЕММЫ. Имеем

$$(10) \quad \sum_{h=0}^H a_h A_h v_s = \sum_{h=0}^H a_h \lambda_{hs} v_s, \quad s = 1, \dots, T. \blacksquare$$

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ. Пусть

$$(11) \quad P(\mathbf{A}) = a_0 \mathbf{I} + a_1 \mathbf{A} + \dots + a_H \mathbf{A}^H \dots$$

Тогда утверждение теоремы вытекает из доказанных двух лемм. ■

6.5.2. Циклическая модель

Одной из возможных разновидностей сериальных корреляций, имеющей простые математические свойства и достаточно изученной, является так называемый циклический сериальный коэффициент корреляции. Он связан с некоторой циклической вероятностной моделью.

Пусть $\mathbf{B}$ есть циркулянт:

$$(12) \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Эта матрица ортогональна, $\mathbf{B}^{-1} = \mathbf{B}'$. Положим

$$(13) \quad \mathbf{A}_1 = \frac{1}{2} (\mathbf{B} + \mathbf{B}') = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

При этом имеем

$$(14) \quad Q_1 = \mathbf{y}' \mathbf{A}_1 \mathbf{y} = \frac{1}{2} \left(\sum_{t=1}^{T-1} y_t y_{t+1} + \sum_{t=2}^T y_t y_{t-1} + y_1 y_T + y_T y_1 \right) = \\ = \sum_{t=1}^T y_t y_{t-1},$$

где $y_0 \equiv y_T$. Последняя сумма равна $\sum_{t=2}^T y_t y_{t-1}$ плюс $y_1 y_T$. Последний член включается в Q_1 из соображений удобства вычислений.

Определим, далее,

$$(15) \quad \mathbf{A}_j = \frac{1}{2} [\mathbf{B}^j + (\mathbf{B}')^j] = \frac{1}{2} (\mathbf{B}^j + \mathbf{B}^{-j}), \quad j < \frac{1}{2} T.$$

Эта матрица имеет элементы, равные $1/2$, на диагоналях, расположенных соответственно на j единиц выше и ниже главной диагонали. Последние соответствуют квадратичной форме $\sum_t y_t y_{t-j}$, стоящей в показателе экспоненты плотности стационарного гауссовского процесса, имеющего порядок не ниже, чем j . Кроме того, матрица $\mathbf{A}_j$ имеет элементы, равные $1/2$, на диагоналях, расположенных соответственно на $T-j$ единиц выше и ниже главной диагонали. Вообще

$$(16) \quad \mathbf{A}_1^{2r} = \left(\frac{1}{2}\right)^{2r} [\mathbf{B} + \mathbf{B}^{-1}]^{2r} =$$

$$= \left(\frac{1}{2}\right)^{2r} \left[\mathbf{B}^{2r} + 2r\mathbf{B}^{2r-2} + \binom{2r}{2}\mathbf{B}^{2r-4} + \dots \right.$$

$$\left. \dots + 2r\mathbf{B}^{-(2r-2)} + \mathbf{B}^{-2r} \right] =$$

$$= \left(\frac{1}{2}\right)^{2r-1} \left[\mathbf{A}_{2r} + 2r\mathbf{A}_{2r-2} + \dots + \binom{2r}{r-1}\mathbf{A}_2 + \frac{1}{2}\binom{2r}{r}\mathbf{I} \right],$$

$$(17) \quad \mathbf{A}_1^{2r+1} = \left(\frac{1}{2}\right)^{2r} \left[\mathbf{A}_{2r+1} + (2r+1)\mathbf{A}_{2r-1} + \dots + \binom{2r+1}{r}\mathbf{A}_1 \right],$$

так что $\mathbf{A}_j$ можно записать в виде полинома от $\mathbf{A}_1$:

$$(18) \quad \begin{aligned} \mathbf{A}_2 &= 2\mathbf{A}_1^2 - \mathbf{I}, \\ \mathbf{A}_3 &= 4\mathbf{A}_1^3 - 3\mathbf{A}_1, \\ \mathbf{A}_4 &= 8\mathbf{A}_1^4 - 8\mathbf{A}_1^2 + \mathbf{I}, \\ \mathbf{A}_5 &= 16\mathbf{A}_1^5 - 20\mathbf{A}_1^3 + 5\mathbf{A}_1, \\ \mathbf{A}_6 &= 32\mathbf{A}_1^6 - 48\mathbf{A}_1^4 + 18\mathbf{A}_1^2 - \mathbf{I}. \end{aligned}$$

ТЕОРЕМА 6.5.2. Характеристические корни матрицы $\mathbf{A}_1$ равны

$$(19) \quad \cos 0 = 1, \cos \frac{2\pi}{T}, \cos \frac{2\pi}{T}, \cos \frac{4\pi}{T}, \cos \frac{4\pi}{T}, \dots,$$

$$\cos \frac{\pi(T-1)}{T}, \cos \frac{\pi(T-1)}{T}, \text{ если } T \text{ нечетное,}$$

$$(20) \quad \cos 0 = 1, \cos \frac{2\pi}{T}, \cos \frac{2\pi}{T}, \cos \frac{4\pi}{T}, \cos \frac{4\pi}{T}, \dots,$$

$$\cos \frac{\pi(T-2)}{T}, \cos \frac{\pi(T-2)}{T}, \cos \pi = -1, \text{ если } T \text{ четное.}$$

Характеристическим вектором, соответствующим характеристическому корню 1, является $(1, 1, \dots, 1)'$. Характеристические векторы, соответствующие характеристическим корням $\cos 2\pi s/T$ ($s \neq 0, T/2$), равны

$$(21) \quad \begin{bmatrix} \cos \frac{2\pi s}{T} \\ \cos \frac{4\pi s}{T} \\ \vdots \\ 1 \end{bmatrix}, \quad \begin{bmatrix} \sin \frac{2\pi s}{T} \\ \sin \frac{4\pi s}{T} \\ \vdots \\ 0 \end{bmatrix}.$$

Если T четное, то характеристическим вектором, соответствующим характеристическому корню (-1) , будет $(-1, 1, -1, \dots, 1)'$.

Доказательство. Уравнение, определяющее характеристические корни и характеристические векторы матрицы $\mathbf{B}$,

$$(22) \quad \mathbf{B}\mathbf{x} = \nu\mathbf{x},$$

в покомпонентной записи принимает вид

$$(23) \quad x_{t+1} = \nu x_t, \quad t = 1, \dots, T-1,$$

$$(24) \quad x_1 = \nu x_T.$$

Таким образом, из (23) следует, что

$$(25) \quad x_t = \nu^{t-1} x_1, \quad t = 2, \dots, T$$

Используя (24), получаем тогда

$$(26) \quad x_1 = \nu^T x_1,$$

$$(27) \quad \nu^T = 1.$$

Отсюда следует, что характеристические корни матрицы $\mathbf{B}$ являются корнями степени T из 1, т. е. они равны $1, e^{i2\pi/T}, e^{i4\pi/T}, \dots, e^{i2\pi(T-1)/T}$. Если характеристический вектор, соответствующий корню $e^{i2\pi s/T}$, $s = 0, 1, \dots, T-1$, имеет первую компоненту, равную $e^{i2\pi s/T}$, то его t -й компонентой будет $e^{i2\pi ts/T}$. Обозначим этот вектор через $\mathbf{u}_s$. Тогда

$$\begin{aligned} (28) \quad \mathbf{A}_1 \mathbf{u}_s &= \frac{1}{2} (\mathbf{B} + \mathbf{B}^{-1}) \mathbf{u}_s = \frac{1}{2} \mathbf{B} \mathbf{u}_s + \frac{1}{2} \mathbf{B}^{-1} \mathbf{u}_s = \\ &= \frac{1}{2} e^{i2\pi s/T} \mathbf{u}_s + \frac{1}{2} e^{-i2\pi s/T} \mathbf{u}_s = \\ &= \cos \frac{2\pi s}{T} \mathbf{u}_s. \end{aligned}$$

Если $s \neq 0, T/2$, то $\cos(2\pi s/T) = \cos[2\pi(T-s)/T]$ является двойным корнем. В качестве соответствующих ему характеристических векторов возьмем векторы $(\mathbf{u}_s + \mathbf{u}_{T-s})/2$ и $(\mathbf{u}_s - \mathbf{u}_{T-s})/(2i)$, t -е компоненты которых равны соответственно

$$(29) \quad \frac{1}{2} (e^{i2\pi ts/T} + e^{i2\pi t(T-s)/T}) = \cos \frac{2\pi ts}{T},$$

$$(30) \quad \frac{1}{2i} (e^{i2\pi ts/T} - e^{i2\pi t(T-s)/T}) = \sin \frac{2\pi ts}{T}.$$

Это и доказывает теорему. ■

ТЕОРЕМА 6.5.3. *Характеристические корни матрицы $\mathbf{A}_j$ равны*

$$(31) \quad \cos 0 = 1, \cos \frac{2\pi j}{T}, \cos \frac{2\pi j}{T}, \cos \frac{4\pi j}{T}, \cos \frac{4\pi j}{T}, \dots$$

$$\dots, \cos \frac{\pi j(T-1)}{T}, \cos \frac{\pi j(T-1)}{T}, \text{ если } T \text{ нечетное,}$$

$$(32) \quad \cos 0 = 1, \cos \frac{2\pi j}{T}, \cos \frac{2\pi j}{T}, \cos \frac{4\pi j}{T}, \cos \frac{4\pi j}{T}, \dots$$

$$\dots, \cos \frac{\pi j(T-2)}{T}, \cos \frac{\pi j(T-2)}{T}, \cos \pi j = (-1)^j, \text{ если } T \text{ четное.}$$

Характеристическим вектором, соответствующим характеристическому корню 1, является $(1, 1, \dots, 1)'$. Характеристические векторы, соответствующие характеристическим корням $\cos 2\pi js/T$ ($s \neq 0, T/2$), равны (21). Если T четное, то характеристическим вектором, соответствующим характеристическому корню $(-1)^j$, будет $(-1, 1, -1, \dots, 1)'$.

ДОКАЗАТЕЛЬСТВО. Поскольку $\mathbf{A}_j$ является полиномом от $\mathbf{A}_1$, то она имеет те же самые характеристические векторы, что и матрица $\mathbf{A}_1$, а ее характеристические корни являются значениями соответствующих полиномов от характеристических корней матрицы $\mathbf{A}_1$ (в силу теоремы 6.5.1). Аналогично, характеристические корни и векторы матрицы, являющейся степенью матрицы $\mathbf{B}$, будут характеристическими векторами и степенями характеристических корней матрицы $\mathbf{B}$. Поэтому

$$(33) \quad \mathbf{A}_j \mathbf{u}_s = \frac{1}{2} (\mathbf{B}^j + \mathbf{B}^{-j}) \mathbf{u}_s = \frac{1}{2} (e^{i2\pi js/T} + e^{-i2\pi js/T}) \mathbf{u}_s =$$

$$= \cos \frac{2\pi j}{T} s \mathbf{u}_s.$$

Следует отметить, что характеристические векторы образуют здесь последовательности тригонометрических функций, уже использовавшиеся нами при изучении циклического тренда в гл. 4.

Если эти векторы нормировать, они будут образовывать те же ортогональные матрицы, что и в разд. 4.2.2, а именно:

$$(34) \quad \mathbf{M} = \sqrt{\frac{2}{T}} \times$$

$$\times \begin{bmatrix} \frac{1}{\sqrt{2}} \cos \frac{2\pi}{T} & \sin \frac{2\pi}{T} & \cos \frac{4\pi}{T} & \dots & \sin \frac{\pi(T-2)}{T} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \cos \frac{4\pi}{T} & \sin \frac{4\pi}{T} & \cos \frac{8\pi}{T} & \dots & \sin \frac{2\pi(T-2)}{T} & \frac{1}{\sqrt{2}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{\sqrt{2}} & 1 & 0 & 1 & \dots & 0 & \frac{1}{\sqrt{2}} \end{bmatrix},$$

если T четное, и

$$(35) \quad \mathbf{M} = \sqrt{\frac{2}{T}} \begin{bmatrix} \frac{1}{\sqrt{2}} \cos \frac{2\pi}{T} & \sin \frac{2\pi}{T} & \cos \frac{4\pi}{T} & \dots & \sin \frac{\pi(T-1)}{T} \\ \frac{1}{\sqrt{2}} \cos \frac{4\pi}{T} & \sin \frac{4\pi}{T} & \cos \frac{8\pi}{T} & \dots & \sin \frac{2\pi(T-1)}{T} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{\sqrt{2}} & 1 & 0 & 1 & \dots & 0 \end{bmatrix},$$

если T нечетное.

Итак, все квадратичные формы Q_j диагонализуются одной и той же ортогональной матрицей. Преобразование вектора $(y_1, \dots, y_T)'$ с помощью этой матрицы к вектору $(z_1, \dots, z_T)'$ приводит j -ю квадратичную форму к виду

$$(36) \quad Q_j = \sum_{s=1}^T \cos \frac{2\pi j}{T} s z_s^2, \quad j = 0, 1, \dots, p.$$

(Заметим, что значения $\cos 2\pi js/T$ при $s=0$ и $s=T$ совпадают.) Если наблюдения y_t независимы и нормально распределены с нулевыми средними и дисперсиями σ^2 ($\gamma_0 = 1/\sigma^2$, $\gamma_1 = \dots = \gamma_p = 0$), то подобным же образом будут распределены и переменные z_t .

Циклический сериальный коэффициент корреляции j -го порядка определяется как

$$(37) \quad r_j = \frac{Q_j}{Q_0} = \frac{\sum_{t=1}^T y_t y_{t-j}}{\sum_{t=1}^T y_t^2} = \frac{\sum_{s=1}^T \cos \frac{2\pi j}{T} s z_s^2}{\sum_{s=1}^T z_s^2},$$

где $y_{t-j} = y_{t-j+T}$, $t = 1, \dots, j$, $j = 1, \dots, p$. Он называется так потому, что квадратичная форма Q_j строится здесь с участием всех попарных произведений, отстоящих на j единиц значений наблюдений y_t , и при этом сами наблюдения $y_1, \dots, y_T$ как бы располагаются вдоль некоторой окружности на равных расстояниях друг от друга.

Для большей наглядности можно расположить на единичной окружности T равноудаленных друг от друга точек, так чтобы одной из них была точка $(1, 0)$. Веса $\cos 2\pi s/T$ будут абсциссами последних.

Преимуществом циклической модели является то, что при рассмотрении в такой модели коэффициента сериальной корреляции первого порядка соответствующие характеристические корни образуют пары¹⁾, за исключением, быть может, одного или двух, и все используемые матрицы имеют одни и те же характеристические векторы. В § 6.7 будет показано, что если все корни двойные, за исключением, быть может, одного или двух, то распределение r_1 или r_1^* можно записать в явном виде. Практическое неудобство добавления слагаемых вида $y_1 y_T$ состоит в том, что такие слагаемые никак не связаны с взаимной зависимостью элементов последовательности. Если при этом имеется тренд (который мы игнорировали), то подобные слагаемые могут оказаться большими по абсолютной величине и это поведет к ошибкам в оценке меры зависимости.

6.5.3. Модель, использующая последовательные разности

Займемся теперь другой системой матриц $A_0, A_1, \dots, A_p$. В § 3.4 мы рассматривали использование сумм квадратов разностей для оценивания дисперсии временных рядов. Если $y_1, \dots, y_T$ имеют нулевые средние, дисперсии σ^2 и некоррелированы, то статистика

$$(38) \quad V_r = \frac{\sum_{t=1}^{T-r} (\Delta^r y_t)^2}{(T-r) \binom{2r}{r}}$$

¹⁾ Имеется в виду, что корни имеют кратность два (двойные корни).— Прим. перев.

является несмещенной оценкой для σ^2 . Если средние значения могут отличаться от нуля (но y_t по-прежнему некоррелированы), то отношение V_r/V_q ($q > r$) можно использовать для проверки нулевой гипотезы о том, что тренд $f(t) = \mathbb{E}y_t$ удовлетворяет условию $\Delta^r f(t) = 0$, в предположении, что $\Delta^q f(t) = 0$. Если средние равны нулю, а величины y_t могут быть коррелированы, то изменение разностей будет отражать эту корреляцию. В частности, для $r = 1$ математическое ожидание $2(T-1)V_1$ равно

$$(39) \quad \mathbb{E} \sum_{i=1}^{T-1} (\Delta y_i)^2 = \mathbb{E} \sum_{i=1}^{T-1} (y_{i+1}^2 - 2y_{i+1}y_i + y_i^2) = \\ = \mathbb{E}y_1^2 + 2 \sum_{i=2}^{T-1} \mathbb{E}y_i^2 + \mathbb{E}y_T^2 - 2 \sum_{i=1}^{T-1} \mathbb{E}y_{i+1}y_i.$$

Если последовательные пары значений y_t будут положительно зависимыми, то использование статистики V_1 приведет к занижению величины дисперсии.

Для проверки гипотезы о сериальной корреляции можно использовать статистику, являющуюся отношением *суммы квадратов последовательных разностей* к сумме квадратов наблюдений. Если средние известны и равны нулю, то такая статистика имеет вид

$$(40) \quad \frac{\sum_{i=2}^T (y_i - y_{i-1})^2}{\sum_{i=1}^T y_i^2} = \frac{y_1^2 + 2 \sum_{i=2}^{T-1} y_i^2 + y_T^2 - 2 \sum_{i=2}^T y_i y_{i-1}}{\sum_{i=1}^T y_i^2} = \\ = 2 \left[1 - \frac{\sum_{i=2}^T y_i y_{i-1} + \frac{1}{2} y_1^2 + \frac{1}{2} y_T^2}{\sum_{i=1}^T y_i^2} \right].$$

Если эту статистику рассматривать как альтернативу к циклическим квадратичным формам, то (40) приводит к $Q_0 = \sum_{i=1}^T y_i^2$ и Q_1 , равной числителю дроби в правой части (40). Матрицей квадратичной формы Q_1 является

$$(41) \quad A_1 = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & 1 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

Она отличается от матрицы квадратичной формы $\sum_{t=2}^T y_t y_{t-1}$ наличием $1/2$ в верхнем левом и в правом нижнем углах. Последующие матрицы в этой модели можно образовать с помощью A_1 , используя соотношения (18) (вытекающие из (16) и (17)). Первые две полученные при этом матрицы имеют вид

$$(42) \quad A_2 = 2A_1^2 - I = \frac{1}{2} \begin{bmatrix} 0 & 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & 1 & \dots & 0 \\ 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

$$(43) \quad A_3 = 4A_1^3 - 3A_1 = \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 1 & \dots & 0 \\ 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & 0 & & 0 \end{bmatrix}.$$

Следует отметить, что A_j отличается от желательной для нас матрицы квадратичной формы $\sum_t y_t y_{t-j}$ только добавлением по j элементов, равных $1/2$, в левый верхний и правый нижний углы.

Определение A_j , $j = 2, \dots, p$, как полиномов от A_1 , гарантирует совпадение характеристических векторов A_j с характеристическими векторами матрицы A_1 . Использование тех же полиномов от A_1 , что и в системе циклических сериальных коэффициентов корреляции, приводит к тому, что при малых значениях j матрицы A_j не будут сколько-нибудь значительно отличаться от матриц, у которых ненулевые элементы, равные $1/2$, стоят на диагоналях, удаленных на j единиц от главной. Дело в том, что используемая нами в настоящем случае матрица A_1 отличается от матрицы $(B + B')/2$ только угловыми элементами. Поэтому матрица A_j при малых j будет отличаться от матрицы $[B' + (B')']/2$ небольшим количеством элементов.

Для характеристического вектора x с компонентами $x_1, \dots, x_T$, соответствующего характеристическому корню λ , удовлетворяющего соотношению $A_1 x = \lambda x$, справедливы соотношения

$$(44) \quad \frac{1}{2} (x_1 + x_2) = \lambda x_1,$$

$$(45) \quad \frac{1}{2} (x_{t-1} + x_{t+1}) = \lambda x_t, \quad t = 2, \dots, T-1,$$

$$(46) \quad \frac{1}{2} (x_{T-1} + x_T) = \lambda x_T.$$

Уравнения (45) можно записать в виде

$$(47) \quad x_{t+1} - 2\lambda x_t + x_{t-1} = 0, \quad t = 2, \dots, T-1,$$

т. е. в виде разностного уравнения второго порядка. Решением последнего является

$$(48) \quad x_t = c_1 \xi_1^t + c_2 \xi_2^t,$$

где ξ_1 и ξ_2 — корни характеристического уравнения

$$(49) \quad x^2 - 2\lambda x + 1 = 0,$$

соответствующего этому разностному уравнению. Корни $\lambda \pm \pm \sqrt{\lambda^2 - 1}$ различны, если только λ не равно 1 или -1 . Поскольку $\xi_1 \xi_2 = 1$ и $\xi_1 + \xi_2 = 2\lambda$, то выражение (48) иначе можно записать как

$$(50) \quad x_t = c_1 \xi^t + c_2 \xi^{-t},$$

а $2\lambda = \xi + \xi^{-1}$. Тогда (44) примет вид

$$(51) \quad 0 = x_2 + (1 - 2\lambda) x_1 = c_1 \xi^2 + c_2 \xi^{-2} + \\ + (1 - \xi - \xi^{-1}) (c_1 \xi + c_2 \xi^{-1}) = \\ = c_1 (1 - \xi^{-1}) \xi + c_2 (1 - \xi) \xi^{-1} = (\xi - 1) (c_1 - c_2 \xi^{-1}).$$

Отсюда $c_2 = c_1 \xi$ и можно положить $c_1 = \xi^{-1/2}$, $c_2 = \xi^{1/2}$. При этом

$$(52) \quad x_t = \xi^{t-1/2} + \xi^{-(t-1/2)}.$$

Уравнение (46) из тех же соображений может быть записано в виде

$$(53) \quad 0 = (2\lambda - 1) x_T - x_{T-1} = \\ = (\xi + \xi^{-1} - 1) (\xi^{T-1/2} + \xi^{-(T-1/2)}) - (\xi^{T-3/2} + \xi^{-(T-3/2)}) = \\ = \xi^{T+1/2} + \xi^{-(T+1/2)} - \xi^{T-1/2} - \xi^{-(T-1/2)} = \\ = (\xi^T - \xi^{-T}) (\xi^{1/2} - \xi^{-1/2}),$$

так что либо $\xi = 1$, либо $\xi^{2T} = 1$. Корни уравнения $\xi^{2T} = 1$ в свою очередь равны $\xi = e^{i2\pi s/(2T)} = e^{i\pi s/T}$, $s = 0, 1, \dots, 2T-1$. Поскольку $e^{i2\pi s/(2T)} = e^{-i2\pi(2T-s)/(2T)}$, существует $T+1$ различных значений $\lambda = (\xi + \xi^{-1})/2$ для $\xi = e^{i2\pi s/(2T)}$, $s = 0, 1, \dots, T$. Однако значение $\lambda = -1$, соответствующее значению $\xi = -1$ для $s = T$, не является допустимым, так как при этом

$x_i \equiv 0$. Остальные значения λ являются характеристическими корнями матрицы $\mathbf{A}_1$, и, поскольку у нее существует всего T характеристических корней, эти значения λ образуют полную совокупность характеристических корней матрицы $\mathbf{A}_1$. Для получения t -й компоненты s -го характеристического вектора удобно домножить выражение $e^{i(2t-1)\pi s/(2T)} + e^{-i(2t-1)\pi s/(2T)}$ на $1/2$.

ТЕОРЕМА 6.5.4. *Характеристические корни матрицы $\mathbf{A}_1$, указанной в (41), равны $\cos \pi s/T$, $s = 0, 1, \dots, T-1$, а соответствующие им характеристические векторы равны*

$$(54) \quad \begin{bmatrix} \cos \frac{\pi s}{2T} \\ \cos \frac{3\pi s}{2T} \\ \vdots \\ \cos \frac{(2T-1)\pi s}{2T} \end{bmatrix}.$$

СЛЕДСТВИЕ 6.5.1. *Характеристические корни матрицы $\mathbf{A}_j$, получаемой по формулам (16) и (17) из матрицы $\mathbf{A}_1$, указанной в (41), равны $\cos \pi j s/T$, $s = 0, 1, \dots, T-1$. Соответствующие им характеристические векторы имеют вид (54).*

ДОКАЗАТЕЛЬСТВО. Матрица $\mathbf{A}_j$ определена как полином от $\mathbf{A}_1$, т. е. $\mathbf{A}_j = P_j(\mathbf{A}_1)$. Поэтому ее характеристические корни равны $P_j(\lambda_s)$, где λ_s — характеристические корни матрицы $\mathbf{A}_1$, $s = 0, 1, \dots, T-1$ (в силу теоремы 6.5.1). Из доказательства теоремы 6.5.3, где $\mathbf{A}_1 = (\mathbf{B} + \mathbf{B}')/2$, вытекает, что $P_j(\cos v) = \cos jv$. Используя теорему 6.5.4, получаем утверждение следствия. ■

СЛЕДСТВИЕ 6.5.2. *Матрица $\mathbf{A}_j$ из следствия 6.5.1 приводится к диагональному виду с помощью матрицы*

$$(55) \quad \sqrt{\frac{2}{T}} \begin{bmatrix} \frac{1}{\sqrt{2}} \cos \frac{\pi}{2T} & \cos \frac{2\pi}{2T} & \dots & \cos \frac{\pi(T-1)}{2T} \\ \frac{1}{\sqrt{2}} \cos \frac{3\pi}{2T} & \cos \frac{6\pi}{2T} & \dots & \cos \frac{3\pi(T-1)}{2T} \\ \vdots & \vdots & & \vdots \\ \frac{1}{\sqrt{2}} \cos \frac{(2T-1)\pi}{2T} & \cos \frac{(2T-1)2\pi}{2T} & \dots & \cos \frac{(2T-1)\pi(T-1)}{2T} \end{bmatrix}.$$

Таким образом, и здесь все матрицы приводятся к диагональному виду с помощью одной и той же ортогональной матрицы. Переходя к преобразованным переменным, найдем, что j -я квадратичная фор-

ма принимает в этих переменных вид

$$(56) \quad Q_j = \sum_{s=1}^T \cos \frac{\pi j}{T} (s-1) z_s^2, \quad j = 0, 1, \dots, p.$$

Матрицы, используемые в этом параграфе, тесно связаны с матрицами, использовавшимися в методе переменных разностей (§ 3.4). Положим

$$(57) \quad S_r = \sum_{t=1}^{T-r} (\Delta^r y_t)^2 = \sum_{s,t=1}^T c_{st}^{(r)} y_s y_t$$

и $C_r = (c_{st}^{(r)})$. Тогда матрица C_r приближенно равна C'_r . Действительно, эти матрицы отличаются только элементами с индексами $s, t \leq r$, и $(T-s+1), (T-t+1) \leq r$. Матрицы C_1, C_2, C_3 и C_4 были приведены в (43), (44), (45) и (46) § 3.4 (и обозначались там как A_1, A_2, A_3 и A_4 соответственно). Мы имеем:

$$(58) \quad C_1^2 = \begin{bmatrix} 2 & -3 & 1 & 0 & \dots & 0 \\ -3 & 6 & -4 & 1 & \dots & 0 \\ 1 & -4 & 6 & -4 & \dots & 0 \\ 0 & 1 & -4 & 6 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 2 \end{bmatrix},$$

$$(59) \quad C_1^3 = \begin{bmatrix} 5 & -9 & 5 & -1 & 0 & \dots & 0 \\ -9 & 19 & -15 & 6 & -1 & \dots & 0 \\ 5 & -15 & 20 & -15 & 6 & \dots & 0 \\ -1 & 6 & -15 & 20 & -15 & \dots & 0 \\ 0 & -1 & 6 & -15 & 20 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 5 \end{bmatrix},$$

$$(60) \quad C_1^4 = \begin{bmatrix} 14 & -28 & 20 & -7 & 1 & 0 & \dots & 0 \\ -28 & 62 & -55 & 28 & -8 & 1 & \dots & 0 \\ 20 & -55 & 70 & -56 & 28 & -8 & \dots & 0 \\ -7 & 28 & -56 & 70 & -56 & 28 & \dots & 0 \\ 1 & -8 & 28 & -56 & 70 & -56 & \dots & 0 \\ 0 & 1 & -8 & 28 & -56 & 70 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 14 \end{bmatrix}.$$

Поскольку $C_1 = 2(I - A_1)$, вектор (54) будет s -м характеристическим вектором матрицы $C_1^r = 2^r(I - A_1)^r$, соответствующим характеристическому корню $2^r(1 - \cos \pi s/T)^r$. А так как A_j является полиномом от A_1 степени j , то C_1^r представляет собой линейную комбинацию матриц $A_0 = I, A_1, \dots, A_r$:

$$(61) \quad C_1^r = \binom{2r}{r} I + 2 \sum_{j=1}^r (-1)^j \binom{2r}{r+j} A_j.$$

6.5.4. Другая модель

Представляет определенный интерес брать в качестве Q_1 квадратичную форму $\sum_{t=2}^T y_t y_{t-1}$, поскольку она и есть желательная сумма запаздывающих произведений. Соответствующая матрица

$$(62) \quad A_1 = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & 0 \end{bmatrix}.$$

Взяв эту матрицу в качестве исходной и воспользовавшись уравнениями (18) (вытекающими из (16) и (17)), получим новую систему матриц. Так,

$$(63) \quad A_2 = 2A_1^2 - I = \frac{1}{2} \begin{bmatrix} -1 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \dots & -1 \end{bmatrix},$$

$$(64) \quad A_3 = 4A_1^3 - 3A_1 = \frac{1}{2} \begin{bmatrix} 0 & -1 & 0 & 1 & 0 & \dots & 0 \\ -1 & 0 & 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

Следует отметить, что здесь матрицы A_j отличаются от нужного нам вида, соответствующего квадратичной форме $\sum_i y_i y_{t-j}$, только $j-1$ элементами, равными $1/2$, в левом верхнем углу матрицы и такими же $j-1$ элементами в ее правом нижнем углу.

Характеристический вектор x с компонентами $x_1, \dots, x_T$, соответствующий характеристическому корню λ , удовлетворяет разностному уравнению второго порядка (47) и поэтому имеет вид (50): $x_t = c_1 \xi^t + c_2 \xi^{-t}$, причем $2\lambda = \xi + \xi^{-1}$. Уравнение $x_2 = 2\lambda x_1$ принимает при этом вид

$$(65) \quad c_1 \xi^2 + c_2 \xi^{-2} = (\xi + \xi^{-1})(c_1 \xi + c_2 \xi^{-1}),$$

откуда следует, что $c_1 + c_2 = 0$. Мы можем взять, таким образом, $c_1 = 1$ и $c_2 = -1$. Тогда $x_t = \xi^t - \xi^{-t}$. Получаемое аналогично предыдущему соотношение $x_{T-1} = 2\lambda x_T$ дает

$$(66) \quad \xi^{T-1} - \xi^{-(T-1)} = (\xi + \xi^{-1})(\xi^T - \xi^{-T}),$$

откуда в свою очередь следует, что $\xi^{(T+1)} - \xi^{-(T+1)} = 0$, т. е. $\xi^{2(T+1)} = 1$. Корни этого уравнения являются корнями $2(T+1)$ -й степени из 1, именно $e^{i2\pi s/[2(T+1)]}$, $s = 0, 1, \dots, 2T+1$. (Значения $\xi = 1$ и $\xi = -1$ не являются допустимыми, поскольку при этом $x_t = 0$, $t = 1, \dots, T$.) Поэтому s -й характеристический корень равен

$$(67) \quad \lambda = \frac{\xi + \xi^{-1}}{2} = \frac{1}{2} (e^{i2\pi s/[2(T+1)]} + e^{-i2\pi s/[2(T+1)]}) = \\ = \cos \frac{\pi s}{T+1}, \quad s = 1, \dots, T,$$

а в качестве t -й компоненты соответствующего ему характеристического вектора можно взять

$$(68) \quad \frac{1}{2i} x_t = \frac{1}{2i} (e^{i2\pi st/[2(T+1)]} - e^{-i2\pi st/[2(T+1)]}) = \\ = \sin \frac{\pi st}{T+1}.$$

ТЕОРЕМА 6.5.5. *Характеристические корни матрицы A_1 , указанной в (62), равны $\cos \pi s/(T+1)$, $s = 1, \dots, T$, а соответствующие им характеристические векторы есть*

$$(69) \quad \begin{bmatrix} \sin \frac{\pi s}{T+1} \\ \sin \frac{2\pi s}{T+1} \\ \vdots \\ \sin \frac{T\pi s}{T+1} \end{bmatrix}.$$

Следствие 6.5.3. *Характеристические корни матрицы A_j , получаемой из матрицы A_1 , указанной в (62), по формулам (16) и (17), равны $\cos \pi j s / (T + 1)$, $s = 1, \dots, T$. Соответствующие им характеристические векторы имеют вид (69).*

Следствие 6.5.4. *Матрица A_j из следствия 6.5.3 приводится к диагональному виду с помощью матрицы*

$$(70) \quad V = \frac{1}{\sqrt{T+1}} \begin{bmatrix} \sin \frac{\pi}{T+1} & \sin \frac{2\pi}{T+1} & \cdots & \sin \frac{T\pi}{T+1} \\ \sin \frac{2\pi}{T+1} & \sin \frac{4\pi}{T+1} & \cdots & \sin \frac{2T\pi}{T+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sin \frac{T\pi}{T+1} & \sin \frac{2T\pi}{T+1} & \cdots & \sin \frac{T^2\pi}{T+1} \end{bmatrix}.$$

Применение преобразования с этой ортогональной матрицей приводит квадратичную форму Q_j к виду

$$(71) \quad Q_j = \sum_{s=1}^T \cos \frac{\pi j s}{T+1} s z_s^2, \quad j = 0, 1, \dots, p.$$

Следует отметить, что здесь корни для $T = T^*$ соответствуют корням из предыдущего случая для $T = T^* + 1$, за исключением того, что опускается корень 1 (соответствующий характеристическому вектору с равными компонентами). Однако характеристические векторы будут уже другими. Ортогональная матрица (70) симметрична.

6.5.5. Модели с двойными корнями

В 6.7 мы будем изучать распределения серийных корреляций. В случае независимости ($\gamma_1 = \dots = \gamma_p = 0$) мы найдем, что, если характеристические корни числителя квадратичной формы образуют пары, за исключением, быть может, одного простого корня, распределение серийной корреляции может быть выражено явно в относительно простом виде. Например, таким случаем будет циклическая модель для четного T и $j = 1$. В моделях двух других типов корни для $j = 1$ не образуют пар и распределение величины r_1 не выражается в явном виде для произвольных значений T .

Ватсон и Дурбин (1951) предложили выбирать матрицу A_1 таким образом, чтобы корни образовывали пары. Этого можно добиться, взяв матрицу A_1 в виде

$$(72) \quad A_1 = \begin{pmatrix} A_1^* & 0 \\ 0 & A_1^* \end{pmatrix},$$

если T четное, и в виде

$$(73) \quad A_1 = \begin{pmatrix} A_1^* & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & A_1^* \end{pmatrix},$$

если T нечетное. Матрица A_1^* такова, что все ее характеристические корни различны. В центре матрицы (73) нуль является скаляром. Все характеристические корни матрицы (72) разбиваются на пары, а характеристические корни матрицы (73) разбиваются на пары, за исключением корня 0. Этот корень будет простым, если 0 не является корнем матрицы A_1^* , и трехкратным, если верно обратное. Например, если в качестве A_1^* взять матрицу из разд. 6.5.4, то

$$(74) \quad Q_1 = \sum_{t=2}^{T/2} y_t y_{t-1} + \sum_{t=(T/2)+2}^T y_t y_{t-1}, \quad \text{если } T \text{ четное,}$$

$$(75) \quad Q_1 = \sum_{t=2}^{(T-1)/2} y_t y_{t-1} + \sum_{t=(T+1)/2+2}^T y_t y_{t-1}, \quad \text{если } T \text{ нечетное.}$$

Когда A_1 является полиномом $P_j(A_1)$ тогда соответственно

$$(76) \quad A_1 = P_j(A_1) = \begin{pmatrix} P_j(A_1^*) & 0 \\ 0 & P_j(A_1^*) \end{pmatrix},$$

или

$$(77) \quad A_1 = P_j(A_1) = \begin{pmatrix} P_j(A_1^*) & 0 & 0 \\ 0 & P_j(0) & 0 \\ 0 & 0 & P_j(A_1^*) \end{pmatrix}.$$

6.6. СЛУЧАИ, КОГДА СРЕДНИЕ ЗНАЧЕНИЯ НЕИЗВЕСТНЫ

6.6.1. Постоянное среднее

В большинстве случаев, представляющих статистический интерес, среднее значение ξy_t хотя и является постоянным, например $\xi y_t \equiv \mu$, тем не менее оно оказывается неизвестным. Для того чтобы учесть этот факт, плотность

$$(1) \quad K e^{-\frac{1}{2} (\gamma_0 Q_0 + \dots + \gamma_p Q_p)},$$

где $Q_j = y' A_j y$, видоизменяют, заменяя вектор y вектором $y - \mu \varepsilon$, в котором $\varepsilon = (1, 1, \dots, 1)'$. При этом показателем экспоненты

нормальной плотности будет умноженное на $-1/2$ выражение

$$(2) \quad \gamma_0(y - \mu\epsilon)' A_0(y - \mu\epsilon) + \dots + \gamma_p(y - \mu\epsilon)' A_p(y - \mu\epsilon) = \\ = \gamma_0 Q_0 + \dots + \gamma_p Q_p - 2\mu(\gamma_0 \epsilon' A_0 y + \dots + \gamma_p \epsilon' A_p y) + \\ + \mu^2 \epsilon' (\gamma_0 A_0 + \dots + \gamma_p A_p) \epsilon.$$

Отсюда следует, что достаточным множеством статистик для параметров $\gamma_0, \dots, \gamma_p$ и μ служит совокупность статистик $Q_0, \dots, Q_p$ и $\epsilon' A_0 y, \dots, \epsilon' A_p y$. Если при этом вектор ϵ является характеристическим вектором матриц A_j , соответствующим характеристическому корню λ_j , $j = 0, 1, \dots, p$, то (2) записывается в виде

$$(3) \quad \gamma_0 Q_0 + \dots + \gamma_p Q_p - 2\mu(\gamma_0 \epsilon' A_0 + \dots + \gamma_p \epsilon' A_p) y + \\ + \mu^2(\gamma_0 \epsilon' A_0 \epsilon + \dots + \gamma_p \epsilon' A_p \epsilon) = \\ = \gamma_0 Q_0 + \dots + \gamma_p Q_p - 2\mu(\gamma_0 \lambda_0 + \dots + \gamma_p \lambda_p) \epsilon' y + \\ + \mu^2(\gamma_0 \lambda_0 + \dots + \gamma_p \lambda_p) T,$$

так что достаточное множество статистик для параметров $\gamma_0, \dots, \gamma_p$ и μ в этом случае будут образовывать статистики $Q_0, \dots, Q_p$ и $y = \epsilon' y / T$. Эквивалентным ему достаточным множеством статистик является совокупность статистик $Q_0^*, \dots, Q_p^*$ и $\bar{y}$, где

$$(4) \quad Q_j^* = (y - \bar{y}\epsilon)' A_j (y - \bar{y}\epsilon) = y' A_j y - T \lambda_j \bar{y}^2.$$

Если же ϵ не является характеристическим вектором матриц A_j , $j = 0, 1, \dots, p$, то минимальное достаточное множество статистик для указанных параметров будет состоять из $Q_0, \dots, Q_p$ и линейно независимого подмножества (не обязательно собственного) статистик $\epsilon' A_0 y, \dots, \epsilon' A_p y$.

Например, в циклическом случае

$$(5) \quad Q_1^* = \sum_{t=2}^T y_t y_{t-1} + y_T y_1 - T \bar{y}^2.$$

В модели, основанной на последовательных разностях,

$$(6) \quad Q_1^* = \sum_{t=2}^T y_t y_{t-1} + \frac{1}{2} y_1^2 + \frac{1}{2} y_T^2 - T \bar{y}^2.$$

В третьей из рассматривавшихся моделей вектор ϵ не является характеристическим вектором матрицы A_1 и поэтому для Q_1^* нельзя получить выражение (4). В случаях моделей, основывающихся на использовании матрицы A_1^* в блочной матрице, положение будет следующим. Для четного T вектор ϵ (размерности T) будет характеристическим вектором матрицы A_1 , если вектор ϵ (размерности $T/2$) является характеристическим вектором матрицы A_1^* . Для нечетного T вектор ϵ (размерности T) будет характеристическим век-

тором матрицы A , если вектор ϵ (размерности $(T - 1)/2$) является характеристическим вектором матрицы A_1^* , соответствующим характеристическому корню, равному нулю. (Заметим, что $(T + 1)/2$ -я компонента вектора $A_1 \epsilon$ должна быть равна нулю.)

В оставшейся части разд. 6.6.1 будем предполагать, что ϵ является характеристическим вектором матрицы A_j , соответствующим характеристическому корню λ_j , $j = 0, 1, \dots, p$. При этом ϵ будет также характеристическим вектором матрицы $\sum_{i=0}^p \gamma_i A_i$, соответствующим характеристическому корню $\sum_{i=0}^p \gamma_i \lambda_i$. Оценка наименьших квадратов $\bar{y}$ для μ (минимизирующая $\sum_{i=1}^T (y_i - m)^2$) является марковской оценкой (см. § 2.4), т. е. $\bar{y}$ — наилучшая линейная несмещенная оценка μ .

ТЕОРЕМА 6.6.1. *Если ϵ является характеристическим вектором матрицы A_j , $j = 0, 1, \dots, p$, то распределения среднего $\bar{y}$ и вектора остатков $y - \bar{y}\epsilon$ независимы.*

ДОКАЗАТЕЛЬСТВО. Поскольку $\bar{y}$ и компоненты вектора остатков $y - \bar{y}\epsilon$ состоят из линейных комбинаций компонент вектора y , имеющего нормальное распределение, то $\bar{y}$ и $y - \bar{y}\epsilon$ будут иметь совместное (сингулярное) нормальное распределение. Ковариационная матрица Σ вектора y находится из соотношения $\Sigma^{-1} = \sum_{i=0}^p \gamma_i A_i$. Ввиду того что

$$\begin{aligned} (7) \quad & \mathbb{E} \bar{y} (y - \bar{y}\epsilon) = \mathbb{E} [(y - \mu\epsilon) - (\bar{y} - \mu)\epsilon] (\bar{y} - \mu) = \\ &= \mathbb{E} \left[\frac{1}{T} (y - \mu\epsilon) (y - \mu\epsilon)' \epsilon - \epsilon \frac{1}{T^2} \epsilon' (y - \mu\epsilon) (y - \mu\epsilon)' \epsilon \right] = \\ &= \frac{1}{T} \Sigma \epsilon - \epsilon \frac{1}{T^2} \epsilon' \Sigma \epsilon = \\ &= \frac{1}{T} \left(\sum_{i=0}^p \gamma_i \lambda_i \right)^{-1} \epsilon - \epsilon \frac{1}{T^2} \left(\sum_{i=0}^p \gamma_i \lambda_i \right)^{-1} T = 0, \end{aligned}$$

среднее $\bar{y}$ не коррелировано с компонентами вектора $y - \bar{y}\epsilon$ и поэтому не зависит от $y - \bar{y}\epsilon$. ■

Следствие 6.6.1. *Совместное распределение квадратичных форм $Q_0^*, \dots, Q_p^*$ от остатков и распределение среднего $\bar{y}$ независимы.*

Выводы относительно равномерно наиболее мощных критериев в § 6.3 и относительно процедур со многими решениями в § 6.4 могут

быть распространены и на модели, в которых $\mathcal{E}y_i = \mu$, а ε — характеристический вектор матриц A_j , $j = 0, 1, \dots, p$. В этом случае достаточное множество статистик для параметров $\gamma_0, \dots, \gamma_i$ и μ при $\gamma_{i+1} = \dots = \gamma_p = 0$ образуют статистики $Q_0^*, \dots, Q_i^*$ и $\bar{y}$, $i = 0, 1, \dots, p$. Критическая область S_i^* (в пространстве значений $Q_0^*, \dots, Q_i^*$ и $\bar{y}$) подобного критерия для проверки гипотезы $\gamma_i = \gamma_i^{(1)}$ с уровнем значимости ε_i (при $\gamma_{i+1} = \dots = \gamma_p = 0$) имеет неймановскую структуру. Иными словами, она удовлетворяет соотношению

$$(8) \Pr \{S_i^* | Q_0^*, \dots, Q_{i-1}^*, \bar{y}; \gamma_0, \dots, \gamma_{i-1}, \gamma_i^{(1)}, 0, \dots, 0, \mu\} = \varepsilon_i$$

для почти всех возможных значений $Q_0^*, \dots, Q_{i-1}^*$ и y . Поскольку $Q_0^*, \dots, Q_i^*$ статистически не зависят от $\bar{y}$, то условная вероятность того, что значение Q_i^* будет принадлежать S_i^* при заданных значениях $Q_0^*, \dots, Q_{i-1}^*$ и $\bar{y}$, не зависит от значения $\bar{y}$. Поскольку к тому же $Q_0^*, \dots, Q_i^*$ являются функциями остатков, распределение которых не зависит от параметра μ , то вероятность (8) не будет зависеть и от μ . Поэтому можно (8) записать в виде

$$(9) \Pr \{S_i^* | Q_0^*, \dots, Q_{i-1}^*; \gamma_0, \dots, \gamma_{i-1}, \gamma_i^{(1)}, 0, \dots, 0\} = \varepsilon_i.$$

Совместную плотность распределения величин $Q_0^*, \dots, Q_i^*$ (при $\gamma_{i+1} = \dots = \gamma_p = 0$) можно записать как

$$(10) h_i^*(Q_0^*, \dots, Q_i^* | \gamma_0, \dots, \gamma_i) = \\ = K_i^*(\gamma_0, \dots, \gamma_i) \exp \left[-\frac{1}{2} (\gamma_0 Q_0^* + \dots + \gamma_i Q_i^*) \right] k_i^*(Q_0^*, \dots, Q_i^*).$$

Маргинальная плотность для $Q_0^*, \dots, Q_{i-1}^*$ равна

$$(11) K_i^*(\gamma_0, \dots, \gamma_i) \exp \left[-\frac{1}{2} (\gamma_0 Q_0^* + \dots + \gamma_{i-1} Q_{i-1}^*) \right] \times \\ \times m_i^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i),$$

где

$$(12) m_i^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i) = \int_{-\infty}^{\infty} e^{-\frac{1}{2} \gamma_i Q_i^*} k_i^*(Q_0^*, \dots, Q_i^*) dQ_i^*.$$

Условная плотность для Q_i^* при заданных $Q_0^*, \dots, Q_{i-1}^*$ будет, таким образом, равна

$$(13) h_i^*(Q_i^* | Q_0^*, \dots, Q_{i-1}^*; \gamma_i) = \exp \left[-\frac{1}{2} \gamma_i Q_i^* \right] \times \\ \times \frac{k_i^*(Q_0^*, \dots, Q_i^*)}{m_i^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i)}.$$

Она может быть выписана только для тех наборов $Q_0^*, \dots, Q_{i-1}^*$, при которых $k_{i-1}^*(Q_0^*, \dots, Q_{i-1}^*) > 0$ и, значит, при которых $m_i^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i) > 0$. Мы можем воспользоваться теперь фундаментальной леммой Неймана — Пирсона.

ТЕОРЕМА 6.6.2. *Наилучший подобный критерий для проверки нулевой гипотезы $\gamma_i = \gamma_i^{(1)}$ против альтернатив $\gamma_i = \gamma_i^{(2)} < \gamma_i^{(1)}$ с уровнем значимости ε_i имеет критическую область*

$$(14) \quad Q_i^* > c_i^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)}),$$

где $c_i^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)})$ определяется таким образом, чтобы вероятность события (14), вычисляемая согласно плотности (13), при $\gamma_i = \gamma_i^{(1)}$ была равна ε_i .

ТЕОРЕМА 6.6.3. *Равномерно наиболее мощный подобный критерий для проверки нулевой гипотезы $\gamma_i = \gamma_i^{(1)}$ против альтернатив $\gamma_i < \gamma_i^{(1)}$ с уровнем значимости ε_i имеет критическую область (14), где $c_i^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)})$ определяется таким образом, чтобы вероятность события (14), вычисляемая согласно плотности (13), при $\gamma_i = \gamma_i^{(1)}$ была равна ε_i .*

СЛЕДСТВИЕ 6.6.2. *Равномерно наиболее мощный подобный критерий для проверки нулевой гипотезы $\gamma_i = 0$ против альтернатив $\gamma_i < 0$ с уровнем значимости ε_i имеет критическую область*

$$(15) \quad Q_i^* > c_i^*(Q_0^*, \dots, Q_{i-1}^*),$$

где $c_i^*(Q_0^*, \dots, Q_{i-1}^*)$ определяется таким образом, чтобы интеграл от плотности

$$(16) \quad h_i^*(Q_i^* | Q_0^*, \dots, Q_{i-1}^*; 0) = \frac{k_i^*(Q_0^*, \dots, Q_i^*)}{k_{i-1}^*(Q_0^*, \dots, Q_{i-1}^*)}$$

по множеству (15) был равен ε_i .

Наилучший критерий против альтернативы $\gamma_i = \gamma_i^{(2)} > \gamma_i^{(1)}$ будет иметь критическую область

$$(17) \quad Q_i^* < c_i^{*'}(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)}).$$

ТЕОРЕМА 6.6.4. *Равномерно наиболее мощный несмещенный критерий для проверки нулевой гипотезы $\gamma_i = \gamma_i^{(1)}$ против альтернатив $\gamma_i \neq \gamma_i^{(1)}$ с уровнем значимости ε_i имеет критическую область*

$$(18) \quad \begin{aligned} Q_i^* &> c_{U_i}^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)}), \\ Q_i^* &< c_{L_i}^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)}), \end{aligned}$$

где $c_{U_i}^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)})$ и $c_{L_i}^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)})$ определяются так, что

$$(19) \quad \int_{c_{L_i}^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)})}^{c_{U_i}^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)})} h_i^*(Q_i^* | Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)}) dQ_i^* = 1 - \varepsilon_i$$

$$(20) \quad \int_{c_{L_i}^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)})}^{c_{U_i}^*(Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)})} Q_i^* h_i^*(Q_i^* | Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)}) dQ_i^* =$$

$$= (1 - \varepsilon_i) \int_{-\infty}^{\infty} Q_i^* h_i^*(Q_i^* | Q_0^*, \dots, Q_{i-1}^*; \gamma_i^{(1)}) dQ_i^*.$$

СЛЕДСТВИЕ 6.6.3. Равномерно наиболее мощный несмещенный критерий для проверки нулевой гипотезы $\gamma_i = 0$ против альтернатив $\gamma_i \neq 0$ с уровнем значимости ε_i имеет критическую область

$$(21) \quad \begin{aligned} Q_i^* &> c_{U_i}^*(Q_0^*, \dots, Q_{i-1}^*), \\ Q_i^* &< c_{L_i}^*(Q_0^*, \dots, Q_{i-1}^*), \end{aligned}$$

где $c_{U_i}^*(Q_0^*, \dots, Q_{i-1}^*)$ и $c_{L_i}^*(Q_0^*, \dots, Q_{i-1}^*)$ определяются соотношениями

$$(22) \quad \int_{c_{L_i}^*(Q_0^*, \dots, Q_{i-1}^*)}^{c_{U_i}^*(Q_0^*, \dots, Q_{i-1}^*)} \frac{k_i^*(Q_0^*, \dots, Q_i^*)}{k_{i-1}^*(Q_0^*, \dots, Q_{i-1}^*)} dQ_i^* = 1 - \varepsilon_i$$

$$(23) \quad \int_{c_{L_i}^*(Q_0^*, \dots, Q_{i-1}^*)}^{c_{U_i}^*(Q_0^*, \dots, Q_{i-1}^*)} Q_i^* \frac{k_i^*(Q_0^*, \dots, Q_i^*)}{k_{i-1}^*(Q_0^*, \dots, Q_{i-1}^*)} dQ_i^* =$$

$$= (1 - \varepsilon_i) \int_{-\infty}^{\infty} Q_i^* \frac{k_i^*(Q_0^*, \dots, Q_i^*)}{k_{i-1}^*(Q_0^*, \dots, Q_{i-1}^*)} dQ_i^*.$$

Если условная плотность для Q_i^* при заданных значениях $Q_0^*, \dots, Q_{i-1}^*$ симметрична, когда $\gamma_i = 0$, соотношение (23) будет выполнено для

$$(24) \quad c_{L_i}^*(Q_0^*, \dots, Q_{i-1}^*) = -c_{U_i}^*(Q_0^*, \dots, Q_{i-1}^*),$$

поскольку обе части (23) при этом равны нулю. Соотношение (22) примет в этом случае вид

$$(25) \quad \int_{c_{U_i}^*(Q_0^*, \dots, Q_{i-1}^*)}^{\infty} \frac{k_i^*(Q_0^*, \dots, Q_i^*)}{k_{i-1}^*(Q_0^*, \dots, Q_{i-1}^*)} dQ_i^* = \frac{1}{2} \varepsilon_i.$$

Следует отметить, что условное распределение для Q_i^* при заданных значениях $Q_0^*, \dots, Q_{i-1}^*$ не зависит от значений параметров $\gamma_0, \dots, \gamma_{i-1}$. Поэтому можно приписать им любые удобные значения, например положить $\gamma_0 = 1, \gamma_1 = \dots = \gamma_{i-1} = 0$.

Как правило, Q_0^* будет равно $\sum_{t=1}^T (y_t - \bar{y})^2 = \sum_{t=1}^T y_t^2 - T\bar{y}^2$. Отношение (для $\lambda_j = 1$)

$$(26) \quad r_j^* = \frac{Q_j^*}{Q_0^*} = \frac{\sum_{s,t=1}^T a_{st}^{(j)} (y_s - \bar{y}) (y_t - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} = \frac{\sum_{s,t=1}^T a_{st}^{(j)} y_s y_t - T\bar{y}^2}{\sum_{t=1}^T y_t^2 - T\bar{y}^2}$$

можно рассматривать как *сериальный коэффициент корреляции*. Если $\gamma_0 > 0$, а $\gamma_1 = \dots = \gamma_i = 0$, то совместное распределение величин $r_1^*, \dots, r_i^*$ не зависит от γ_0 и Q_0^* . (См. теорему 6.7.2.) Оптимальные критерии (15) и (21) можно тогда определить, используя условное распределение r_i^* при заданных значениях $r_1^*, \dots, r_{i-1}^*$.

Теоремы § 6.4 относительно оптимальных процедур со многими решениями могут быть подобным же образом перефразированы в терминах $Q_0^*, \dots, Q_p^*$, на случай когда $\mathcal{E}y_t = \mu$, а ε является характеристическим вектором матриц $A_0, \dots, A_p$.

6.6.2. Некоторые функции регрессии

При изучении случая ненулевых средних можно рассматривать средние и более общего вида. Предположим, что

$$(27) \quad \mathcal{E}y = \sum_{v=1}^m \alpha_v \mathbf{v}_{sv},$$

где $\mathbf{v}_1, \dots, \mathbf{v}_T$ — характеристические векторы,

$$(28) \quad A_j \mathbf{v}_t = \lambda_{jt} \mathbf{v}_t, \quad t = 1, \dots, T, \quad j = 0, 1, \dots, p,$$

а $s_1, \dots, s_m$ — некоторое подмножество значений $1, \dots, T$. Векторы $\mathbf{v}_1, \dots, \mathbf{v}_T$ ортогональны: $\mathbf{v}_t \mathbf{v}_s = \delta_{ts}$. Показатель экспоненты соответствующей нормальной плотности равен тогда $-1/2$, умноженной

на

$$(29) \sum_{j=0}^p \gamma_j \left(y - \sum_{v=1}^m \alpha_v v_{sv} \right)' A_j \left(y - \sum_{v=1}^m \alpha_v v_{sv} \right) = \\ = \sum_{j=0}^p \gamma_j y' A_j y - 2 \sum_{j=0}^p \sum_{v=1}^m \gamma_j \lambda_{jsv} \alpha_v y' v_{sv} + \sum_{j=0}^p \sum_{v=1}^m \gamma_j \lambda_{jsv} \alpha_v^2 v_{sv}' v_{sv}.$$

Таким образом, $Q_0, \dots, Q_p$ и $y'v_{s_1}, \dots, y'v_{s_m}$ образуют достаточное множество статистик для параметров $\gamma_0, \dots, \gamma_p$ и $\alpha_1, \dots, \alpha_m$. Эквивалентным ему достаточным множеством статистик будет

$$(30) \left(y - \sum_{v=1}^m \hat{\alpha}_v v_{sv} \right)' A_j \left(y - \sum_{v=1}^m \hat{\alpha}_v v_{sv} \right) = y' A_j y - \sum_{v=1}^m \lambda_{jsv} \hat{\alpha}_v^2 v_{sv}' v_{sv}, \\ j = 0, 1, \dots, p,$$

и

$$(31) \hat{\alpha}_v = \frac{v_{sv}' y}{v_{sv}' v_{sv}}, \quad v = 1, \dots, m.$$

Заметим, что оценки наименьших квадратов (31) параметров α_v , $v = 1, \dots, m$, получаемые минимизацией $\left(y - \sum_{v=1}^m \alpha_v v_{sv} \right)' \times$
 $\times \left(y - \sum_{v=1}^m \alpha_v v_{sv} \right)$, также являются марковскими, т. е. наилучшими линейными несмещенными оценками, поскольку v_{sv} — характеристические векторы $\sum_{j=0}^p \gamma_j A_j$. (См. § 2.4.)

В качестве примера рассмотрим матрицы $A_0 = I$, $A_1, \dots, A_p$ из циклической модели разд. 6.5.2. Характеристические корни матрицы A_j равны $\cos 2\pi sj/T$, $s = 1, \dots, T$. Соответствующие характеристические векторы имеют в качестве t -х компонент величины $\cos 2\pi st/T$ и $\sin 2\pi st/T$ для $s = 1, \dots, (T-2)/2$, если T четное, или $s = 1, \dots, (T-1)/2$, если T нечетное. Характеристический вектор $(1, 1, \dots, 1)'$ соответствует корню $1 = \cos 2\pi Tj/T$. Если T четное, то имеется еще характеристический вектор $(-1, 1, \dots, 1)'$, соответствующий корню $(-1)^j = \cos 2\pi \frac{1}{2} Tj/T$. Эти характеристические векторы образуют последовательности, которые были использованы при рассмотрении периодических трендов в гл. 4. Предположим, что

$$(32) \mathcal{Y}_t = \mu_t = \alpha_0 + \sum_{h=1}^q \left(\alpha(v_h) \cos \frac{2\pi v_h}{T} t + \beta(v_h) \sin \frac{2\pi v_h}{T} t \right),$$

где $v_1, \dots, v_q$ составляют подмножество множества целых чисел $1, \dots, (T-2)/2$, если T четное, или $1, \dots, (T-1)/2$, если T нечетное.

Если T четное, можно предполагать также, что

$$(33) \quad \mathbb{E} y_t = \mu_t = \alpha_0 + \sum_{h=1}^q \left(\alpha(v_h) \cos \frac{2\pi v_h}{T} t + \beta(v_h) \sin \frac{2\pi v_h}{T} t \right) + \alpha_{T/2} (-1)^t.$$

Индекс m в (27) соответствует $2q+1$ в (32) и $2q+2$ в (33). Оценки наименьших квадратов коэффициентов даются соотношениями

$$(34) \quad \begin{aligned} a(v_h) &= \frac{2}{T} \sum_{t=1}^T y_t \cos \frac{2\pi v_h}{T} t, \\ b(v_h) &= \frac{2}{T} \sum_{t=1}^T y_t \sin \frac{2\pi v_h}{T} t, \quad v_h \neq 0, T/2, \end{aligned}$$

$$(35) \quad a_0 = \bar{y}$$

и, если T четное,

$$(36) \quad a_{T/2} = \frac{1}{T} \sum_{t=1}^T y_t (-1)^t.$$

Положим

$$(37) \quad m_t = a_0 + \sum_{h=1}^q \left(a(v_h) \cos \frac{2\pi v_h}{T} t + b(v_h) \sin \frac{2\pi v_h}{T} t \right)$$

или, если T четное,

$$(38) \quad m_t = a_0 + \sum_{h=1}^q \left(a(v_h) \cos \frac{2\pi v_h}{T} t + b(v_h) \sin \frac{2\pi v_h}{T} t \right) + a_{T/2} (-1)^t.$$

Циклический сериальный коэффициент корреляции j -го порядка, использующий остатки от подобранного периодического тренда, есть

$$(39) \quad \begin{aligned} r_j^* &= \frac{\sum_{t=1}^T (y_t - m_t) (y_{t-j} - m_{t-j})}{\sum_{t=1}^T (y_t - m_t)^2} = \\ &= \frac{\sum_{t=1}^T y_t y_{t-j} - T \bar{y}^2 - \frac{1}{2} T \sum_{h=1}^q \cos \frac{2\pi v_h j}{T} [a^2(v_h) + b^2(v_h)]}{\sum_{t=1}^T y_t^2 - T \bar{y}^2 - \frac{1}{2} T \sum_{h=1}^q [a^2(v_h) + b^2(v_h)]} \end{aligned}$$

или, если T четное,

$$(40) \quad r_j^* = \frac{\sum_{t=1}^T y_t y_{t-j} - T [\bar{y}^2 + (-1)^j a_{T/2}^2] - \frac{1}{2} T \sum_{h=1}^q \cos \frac{2\pi \nu_h j}{T} [a^2(\nu_h) + b^2(\nu_h)]}{\sum_{t=1}^T y_t^2 - T [\bar{y}^2 + a_{T/2}^2] - \frac{1}{2} T \sum_{h=1}^q [a^2(\nu_h) + b^2(\nu_h)]}.$$

[См. Р. Андерсон, Т. Андерсон (1950).]

Для рассмотренных здесь случаев также можно сформулировать наилучшие критерии и процедуры со многими решениями.

В примере 4.1 гл. 4 приведены ежемесячные данные о поступлении масла на пяти рынках в течении трех лет. Оценка тренда m_t , использующая двенадцать тригонометрических функций (включая константу), указана в табл. 4.1. Циклический сериальный коэффициент корреляции первого порядка, использующий остатки, равен

$$(41) \quad r_1^* = \frac{(2.4)(-0.6) + (-0.6)(-4.6) + \dots + (1.6)(2.5) + (2.5)(2.4)}{(2.4)^2 + (-0.6)^2 + \dots + (1.6)^2 + (2.5)^2} = \\ = \frac{232.18}{474.51} = 0.489.$$

Вычисление в (41) проведено с округленными остатками для облегчения чтения.

6.7. РАСПРЕДЕЛЕНИЯ СЕРИАЛЬНЫХ КОЭФФИЦИЕНТОВ КОРРЕЛЯЦИИ

6.7.1. Общая задача

Различные коэффициенты сериальной корреляции, рассматривавшиеся в предыдущих параграфах, представляют собой отношения квадратичных форм вида $s_j = P_j/P_0$, $j = 1, \dots, p$, где $P_j = \mathbf{y}'\mathbf{B}_j\mathbf{y}$, $j = 0, 1, \dots, p$. Нас будут интересовать их распределения в случае, когда $\mathbf{y}$ имеет многомерное нормальное распределение с ковариационной матрицей Σ . Имея в виду приложения к другим задачам, обсуждение вопроса будем вести в достаточно общих условиях. В настоящей главе при рассмотрении различных приложений соответствующая плотность будет иметь в показателе экспоненты квадратичную форму с матрицей $\Sigma^{-1} = \sum_{i=0}^p \gamma_i \mathbf{A}_i$. В простейших случаях $\mathbf{B}_j = \mathbf{A}_j$; в ряде других

$$(1) \quad \mathbf{B}_j = \left(\mathbf{I} - \frac{1}{T} \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}' \right) \mathbf{A}_j \left(\mathbf{I} - \frac{1}{T} \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}' \right),$$

где $\varepsilon = (1, 1, \dots, 1)'$, так что при этом

$$(2) \quad P_i = (y - \bar{y}\varepsilon)' A_i (y - \bar{y}\varepsilon).$$

В некоторых случаях матрицы B_i таковы, что P_i является квадратичной формой от разностей между y и некоторой более общей функцией регрессии. При этом

$$(3) \quad B_i = (I - V^* (V^{*'} V^*)^{-1} V^{*'}) A_i (I - V^* (V^{*'} V^*)^{-1} V^{*'}),$$

$$(4) \quad P_i = (y - V^* (V^{*'} V^*)^{-1} V^{*'} y)' A_i (y - V^* (V^{*'} V^*)^{-1} V^{*'} y),$$

где V^* — матрица с T строками и рангом, равным числу ее столбцов. Если $B_i = A_i$, мы будем писать Q_i вместо P_i и r_i вместо s_i . В тех случаях, когда B_i задается соотношением (1), вместо P_i и s_i будем писать соответственно Q_i^* и r_i^* .

ТЕОРЕМА 6.7.1. Если вектор y имеет распределение $N(V^* \beta, \Sigma)$, то совместное распределение статистик $P_0, \dots, P_p$, определяемых соотношением (4), не зависит от вектора β .

ДОКАЗАТЕЛЬСТВО. Пусть $y = u + V^* \beta$. Тогда вектор u распределен согласно $N(0, \Sigma)$ и

$$(5) \quad y - V^* (V^{*'} V^*)^{-1} V^{*'} y = (I - V^* (V^{*'} V^*)^{-1} V^{*'}) (u + V^* \beta) = \\ = (I - V^* (V^{*'} V^*)^{-1} V^{*'}) u,$$

$$(6) \quad P_i = (u - V^* (V^{*'} V^*)^{-1} V^{*'} u)' A_i (u - V^* (V^{*'} V^*)^{-1} V^{*'} u). \blacksquare$$

Если $\Sigma^{-1} = \sum_{j=0}^i \gamma_j A_j$ и $B_i = A_i$ ($P_i = Q_i$ и $s_i = r_i$), то условная плотность для Q_i при заданных значениях $Q_0, \dots, Q_{i-1}$ равна

$$(7) \quad h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i) = e^{-\frac{1}{2} \gamma_i Q_i} \frac{k_i(Q_0, \dots, Q_i)}{m_i(Q_0, \dots, Q_{i-1}; \gamma_i)},$$

где

$$(8) \quad m_i(Q_0, \dots, Q_{i-1}; \gamma_i) = \int_{-\infty}^{\infty} e^{-\frac{1}{2} \gamma_i Q_i} k_i(Q_0, \dots, Q_i) dQ_i$$

(последнее выражение предполагается положительным), так что она не зависит от параметров $\gamma_0, \dots, \gamma_{i-1}$. Это равносильно тому, что условное распределение величин r_i при заданных значениях Q_0 и $r_1, \dots, r_{i-1}$ не зависит от $\gamma_0, \dots, \gamma_{i-1}$. Именно это условное распределение будет использовано для проверки нулевой гипотезы о том, что порядок зависимости меньше i (т. е. $\gamma_i = 0$). Получать указанное распределение оказывается удобным, приписывая мешающим параметрам значения $\gamma_0 = 1, \gamma_1 = \dots = \gamma_{i-1} = 0$. Это условное

распределение при $\gamma_i = 0$ не зависит от Q_0 (поскольку $r_1, \dots, r_i$ инвариантны относительно преобразования масштаба $y_i \rightarrow cy_i$).

Следующие лемма и теорема будут полезны в дальнейшем.

ЛЕММА 6.7.1. *Условная плотность распределения величины $\tilde{Q}_i$ при заданных значениях $\tilde{Q}_0 = a_0, \dots, \tilde{Q}_{i-1} = a_{i-1}$, где $\tilde{Q}_0 = c^2 Q_0, \dots, \tilde{Q}_i = c^2 Q_i$, совпадает с условной плотностью $h_i(\tilde{Q}_i | \tilde{Q}_0, \dots, \tilde{Q}_{i-1}; \gamma_i/c^2)$, т. е. с условной плотностью распределения Q_i при заданных значениях $Q_0 = a_0, \dots, Q_{i-1} = a_{i-1}$, в которой параметр γ_i заменяется на γ_i/c^2 .*

Доказательство. Квадратичную форму $\tilde{Q}_i = c^2 y' A_i y = (cy)' A_i (cy)$ можно записать в виде $\tilde{Q}_i = x' A_i x$, $j = 0, \dots, i$, где вектор $x = cy$ имеет плотность, показатель экспоненты которой равен $-1/2$, умноженной на $\sum_{j=0}^i \gamma_j (c^{-1}x)' A_j (c^{-1}x) = \sum_{j=0}^i (\gamma_j/c^2) \tilde{Q}_j$. Таким образом, $\tilde{Q}_0, \dots, \tilde{Q}_i$ имеет совместную плотность, совпадающую с совместной плотностью распределения случайных величин $Q_0, \dots, Q_i$, в которой параметр γ_j заменен на γ_j/c^2 , $j = 0, \dots, i$. Отсюда и следует утверждение леммы. ■

ТЕОРЕМА 6.7.2. *Условная плотность распределения величины $r_i = Q_i/Q_0$ при заданных значениях Q_0 и $r_j = Q_j/Q_0$, $j=1, \dots, i-1$, равна*
(9)
$$h_i(r_i | 1, r_1, \dots, r_{i-1}; \gamma_i Q_0).$$

Если $\gamma_i = 0$, эта условная плотность не зависит от Q_0 .

Доказательство. В силу леммы 6.7.1, условная плотность распределения r_i при заданных значениях $Q_0 = b_0$ и $r_j = b_j$, $j = 1, \dots, i-1$, совпадает с условной плотностью распределения величины $c^2 Q_i$ для $c^2 = 1/Q_0$ при заданных значениях $c^2 Q_0 = 1$ и $r_j = c^2 Q_j = b_j$ (с заменой параметра γ_i на $\gamma_i/c^2 = \gamma_i Q_0$). Этим доказана первая часть теоремы. Если $\gamma_i = 0$, легко проверить, что условная плотность (9) не зависит от Q_0 . Для этого достаточно подставить в (7) и (8) аргументы выражения (9).

Если $\gamma_1 = \dots = \gamma_i = 0$, из теоремы 6.7.2 вытекает, что совместная условная плотность распределения $r_1, \dots, r_i$ при заданном значении Q_0 (являющаяся произведением i условных плотностей вида (9)) не зависит от Q_0 . Таким образом, условная плотность величины r_i при заданном значении Q_0 не зависит от Q_0 . ■

Следует отметить, что лемма 6.7.1 и теорема 6.7.2 также справедливы для Q_i^* и r_i^* .

6.7.2. Характеристические функции

Один из возможных подходов к отысканию распределений квадратичных форм $P_0, \dots, P_p$ или статистик $s_1, \dots, s_p$ состоит в использовании характеристической функции случайных величин $P_0, \dots, P_p$.

ТЕОРЕМА 6.7.3. Если вектор y распределен согласно $N(0, \Sigma)$, то многомерная характеристическая функция случайных величин $P_0, \dots, P_p$ равна

$$(10) \quad \mathbb{E} e^{i(t_0 P_0 + \dots + t_p P_p)} = \sqrt{\frac{|\Sigma^{-1}|}{|\Sigma^{-1} - 2i(t_0 B_0 + \dots + t_p B_p)|}} = \\ = \frac{1}{\sqrt{|I - 2i(t_0 B_0 + \dots + t_p B_p) \Sigma|}}.$$

ДОКАЗАТЕЛЬСТВО. Для чисто мнимых значений $t_0, \dots, t_p$, достаточно малых по абсолютной величине, равенство (10) можно получить, используя тот факт, что интеграл от $\exp \left\{ -\frac{1}{2} y' A y \right\}$ равен $A^{-1/2} (2\pi)^{T/2}$. Для доказательства справедливости соотношения (10) при действительных значениях $t_0, \dots, t_p$ редуцируем многомерный интеграл к произведению интегралов от одной переменной.

Для фиксированных вещественных значений $t_0, \dots, t_p$ пусть C будет такой матрицей, что

$$(11) \quad C' \Sigma^{-1} C = I,$$

$$(12) \quad C' (t_0 B_0 + \dots + t_p B_p) C = D,$$

где D — диагональная матрица. (Если матрица A положительно определена, а матрица B симметрична, то всегда найдется такая матрица C , что $C' A C = I$ и матрица $C' B C$ диагональна. См. упр. 30.) Тогда (при $y = Cz$)

$$(13) \quad \mathbb{E} e^{i(t_0 P_0 + \dots + t_p P_p)} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{T/2} |\Sigma|^{1/2}} \times \\ \times \exp \left[-\frac{1}{2} y' \Sigma^{-1} y + i(t_0 y' B_0 y + \dots + t_p y' B_p y) \right] dy_1 \dots dy_T = \\ = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{T/2} |\Sigma|^{1/2}} \times \\ \times \exp \left\{ -\frac{1}{2} y' [\Sigma^{-1} - 2i(t_0 B_0 + \dots + t_p B_p)] y \right\} dy_1 \dots dy_T =$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{T/2}} \exp \left[-\frac{1}{2} \mathbf{z}' (\mathbf{I} - 2i\mathbf{D}) \mathbf{z} \right] dz_1 \cdots dz_T = \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{T/2}} \exp \left[-\frac{1}{2} \sum_{s=1}^T z_s^2 + i \sum_{s=1}^T d_{ss} z_s^2 \right] dz_1 \cdots dz_T = \\
&= \prod_{s=1}^T \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} \exp \left[-\frac{1}{2} z_s^2 + i d_{ss} z_s^2 \right] dz_s = \\
&= \prod_{s=1}^T \frac{1}{\sqrt{1 - 2i d_{ss}}}
\end{aligned}$$

ввиду того, что если случайная величина z_s распределена по закону $N(0, 1)$, то случайная величина z_s^2 (распределенная по закону χ^2 с одной степенью свободы) имеет характеристическую функцию $1/\sqrt{1 - 2it}$. Используя (11) и (12), получаем из (13) соотношение

$$\begin{aligned}
(14) \quad \mathcal{E} e^{i(t_0 P_0 + \cdots + t_p P_p)} &= \left(\prod_{s=1}^T (1 - 2i d_{ss}) \right)^{-1/2} = (\mathbf{I} - 2i\mathbf{D})^{-1/2} = \\
&= (\|\mathbf{C}'\| \|\Sigma^{-1} - 2i(t_0 \mathbf{B}_0 + \cdots + t_p \mathbf{B}_p)\| \|\mathbf{C}\|)^{-1/2},
\end{aligned}$$

совпадающее с (10), что и требовалось. ■

Следствие 6.7.1. Если вектор $\mathbf{y}$ распределен с плотностью $K \exp \left\{ -\frac{1}{2} \sum_{j=0}^p \gamma_j Q_j \right\}$, то многомерная характеристическая функция квадратичных форм $Q_0 = \mathbf{y}' \mathbf{A}_0 \mathbf{y}$, ..., $Q_p = \mathbf{y}' \mathbf{A}_p \mathbf{y}$ равна

$$(15) \quad \mathcal{E} e^{i(t_0 Q_0 + \cdots + t_p Q_p)} = \sqrt{\frac{|\gamma_0 \mathbf{A}_0 + \cdots + \gamma_p \mathbf{A}_p|}{|(\gamma_0 - 2it_0) \mathbf{A}_0 + \cdots + (\gamma_p - 2it_p) \mathbf{A}_p|}},$$

В принципе для получения совместного распределения квадратичных форм $P_0, P_1, \dots, P_p$ можно взять обратное преобразование Фурье от найденной характеристической функции. Однако на практике это может быть выполнено далеко не всегда. Для нахождения совместного распределения статистик $s_1, \dots, s_p$ можно воспользоваться соотношением

$$\begin{aligned}
(16) \quad \Pr \{s_1 \leq w_1, \dots, s_p \leq w_p\} &= \Pr \left\{ \frac{P_1}{P_0} \leq w_1, \dots, \frac{P_p}{P_0} \leq w_p \right\} = \\
&= \Pr \{P_1 - w_1 P_0 \leq 0, \dots, P_p - \\
&\quad - w_p P_0 \leq 0\}.
\end{aligned}$$

Многомерную характеристическую функцию статистик $P_1 - w_1 P_0, \dots, P_p - w_p P_0$ получим, используя теорему 6.7.3, если в последней положим $t_0 = -w_1 t_1 - \dots - w_p t_p$.

6.7.3. Канонические формы

Как мы уже видели в § 6.5, если $\mathbf{v}_t$ является t -м нормированным характеристическим вектором матрицы $\mathbf{B}$, а φ_t — соответствующий ему характеристический корень, то

$$(17) \quad \mathbf{V}'\mathbf{B}\mathbf{V} = \Phi, \quad \mathbf{V}'\mathbf{V} = \mathbf{I},$$

где Φ — диагональная матрица, имеющая на диагонали элементы $\varphi_1, \dots, \varphi_T$, а $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_T)$. (Если условия ортогональности $\mathbf{v}_t' \mathbf{v}_s = 0$, $t \neq s$, не обязательно выполняются, то можно в качестве новых характеристических векторов взять линейные комбинации векторов $\mathbf{v}_t$, соответствующих общим характеристическим корням, и сделать это таким образом, что условия ортогональности будут выполнены.) Если столбцы матрицы $\mathbf{V}$ являются характеристическими векторами матриц $\mathbf{B}_0, \dots, \mathbf{B}_p$, отвечающими характеристическим корням $\varphi_{01}, \dots, \varphi_{0T}; \varphi_{11}, \dots, \varphi_{1T}; \dots; \varphi_{p1}, \dots, \varphi_{pT}$, соответственно, то

$$(18) \quad \mathbf{V}'\mathbf{B}_j\mathbf{V} = \Phi_j, \quad \mathbf{V}'\mathbf{V} = \mathbf{I}, \quad j = 0, 1, \dots, p,$$

где матрица Φ_j диагональна и имеет на диагонали элементы $\varphi_{j1}, \dots, \varphi_{jT}$. Если $\mathbf{B}_0 = \mathbf{I}$, то квадратичные формы P_j могут быть записаны в виде

$$(19) \quad P_0 = \sum_{t=1}^T z_t^2, \quad P_j = \sum_{t=1}^T \varphi_{jt} z_t^2, \quad j = 1, \dots, p,$$

где $\mathbf{y} = \mathbf{V}\mathbf{z}$. Необходимое и достаточное условие для того, чтобы матрицы $\mathbf{B}_1, \dots, \mathbf{B}_p$ имели одни и те же характеристические векторы, состоит в выполнении соотношения

$$(20) \quad \mathbf{B}_j \mathbf{B}_k = \mathbf{B}_k \mathbf{B}_j, \quad j, k = 1, \dots, p.$$

(См. упр. 31.)

Пусть Λ_j — диагональная матрица, определяемая соотношением $\mathbf{V}'\mathbf{A}_j\mathbf{V} = \Lambda_j$, $j = 0, 1, \dots, p$; матрица $\mathbf{V}^*$ размера $T \times m$ состоит из столбцов, являющихся какими-то m столбцами матрицы $\mathbf{V}$. При этом столбцы матрицы $\mathbf{V}^*$ будут характеристическими векторами матриц $\mathbf{A}_j$. Они будут соответствовать m характеристическим корням матрицы $\mathbf{A}_j$, которые представляют собой m диагональных элементов матрицы Λ_j , $j = 0, 1, \dots, p$. Мы будем предполагать, что характеристические корни и характеристические векторы матриц $\mathbf{A}_j$ перенумерованы таким образом, что указанные m столбцов матрицы $\mathbf{V}$ являются m ее последними столбцами, а соответствующие им характеристические корни матрицы $\mathbf{A}_j$ образуют диагональную

матрицу A_j^* , являющуюся правой нижней угловой подматрицей размера $m \times m$ матрицы A_j , $j = 0, 1, \dots, p$.

Нас будет особенно интересовать случай, когда $\Sigma^{-1} = \sum_{j=0}^p \gamma_j A_j$, $A_0 = I$, и

$$(21) \quad B_j = [I - V^* V^{**}] A_j [I - V^* V^{**}], \quad j = 0, 1, \dots, p.$$

В этом случае

$$(22) \quad V' V^* = \begin{pmatrix} 0 \\ I \end{pmatrix},$$

$$(23) \quad V' (I - V^* V^{**}) = V' - \begin{pmatrix} 0 \\ V^{**} \end{pmatrix},$$

$$(24) \quad V' B_j V = V' A_j V - \begin{pmatrix} 0 \\ V^{**} \end{pmatrix} A_j V - V' A_j \begin{pmatrix} 0 & V^* \end{pmatrix} + \begin{pmatrix} 0 \\ V^{**} \end{pmatrix} A_j \begin{pmatrix} 0 & V^* \end{pmatrix} = \\ = A_j - \begin{pmatrix} 0 & 0 \\ 0 & A_j^* \end{pmatrix}.$$

Таким образом (для $y = Vz$),

$$(25) \quad P_j = y' B_j y = z' V' B_j V z = \\ = z' A_j z - z' \begin{pmatrix} 0 & 0 \\ 0 & A_j^* \end{pmatrix} z = \\ = \sum_{t=1}^T \lambda_{jt} z_t^2 - \sum_{t=T-m+1}^T \lambda_{jt} z_t^2 = \sum_{t=1}^{T-m} \lambda_{jt} z_t^2, \\ (26) \quad P_0 = \sum_{t=1}^{T-m} z_t^2.$$

Если распределение вектора y есть $N(V^* \beta, \Sigma)$, где $\Sigma^{-1} = \sum_{j=0}^p \gamma_j A_j$, то плотностью распределения вектора $z = V^{-1}y$ будет

$$(27) \quad (2\pi)^{-T/2} \prod_{t=1}^T \left(\sum_{j=0}^p \gamma_j \lambda_{jt} \right)^{1/2} \times \\ \times \exp \left\{ -\frac{1}{2} \left[\sum_{t=1}^{T-m} \left(\sum_{j=0}^p \gamma_j \lambda_{jt} \right) z_t^2 + \sum_{t=T-m+1}^T \left(\sum_{j=0}^p \gamma_j \lambda_{jt} \right) (z_t - \beta_t)^2 \right] \right\},$$

где компоненты вектора β перенумерованы числами от $T - m + 1$ до T .

В случае циклической модели вектор $\varepsilon = (1, \dots, 1)'$ является характеристическим вектором матрицы A_j , соответствующим характеристическому корню 1 (T -й корень), $j = 0, 1, \dots, p$. Каноническим

видом квадратичной формы Q_j^* остатков от среднего значения будет поэтому

$$(28) \quad Q_j^* = \sum_{t=1}^{T-1} \cos \frac{2\pi j}{T} t z_t^2, \quad T = 2, 3, \dots$$

В случае модели, в которой используется сумма квадратов последовательных разностей, вектор $\mathbf{z}$ также будет характеристическим вектором матрицы $\mathbf{A}_j$, соответствующим характеристическому корню 1, и

$$(29) \quad Q_j^* = \sum_{t=1}^{T-1} \cos \frac{\pi j}{T} t z_t^2, \quad T = 2, 3, \dots$$

В третьей из рассматривавшихся моделей $\mathbf{z}$ уже не является характеристическим вектором матриц $\mathbf{A}_j$, $j = 1, \dots, p$, и поэтому редукция квадратичной формы Q_j^* к взвешенной сумме $T - 1$ квадратов не представляется возможной.

В общем случае мы можем иметь $p + 1$ квадратичных форм $P_0, \dots, P_p$, определяющих p сериальных коэффициентов корреляции $s_1 = P_1/P_0, \dots, s_p = P_p/P_0$, и располагать ковариационной матрицей Σ . Всегда можно найти такую матрицу $\mathbf{C}$, что $\mathbf{C}'\Sigma^{-1}\mathbf{C} = \mathbf{I}$ (эквивалентно $\mathbf{C}'\Sigma(\mathbf{C}')^{-1} = \mathbf{I}$), а $\mathbf{C}'\mathbf{B}_j\mathbf{C}$ диагональна для некоторого j . Однако, если только не имеется определенной связи между всеми этими матрицами, с помощью единственной матрицы $\mathbf{C}$ можно диагонализировать лишь одну из матриц $\mathbf{B}_j$. Таким образом, если матрица Σ преобразуется к единичной и $s_j = \mathbf{x}'(\mathbf{C}'\mathbf{B}_j\mathbf{C})\mathbf{x}/[\mathbf{x}'(\mathbf{C}'\mathbf{B}_0\mathbf{C})\mathbf{x}]$, то при этом или числитель или знаменатель обязательно можно привести к взвешенной сумме квадратов. Однако одновременное приведение и числителя и знаменателя к такому виду не всегда возможно.

Следующее утверждение является следствием теоремы 6.7.3.

Следствие 6.7.2. Пусть вектор $\mathbf{y}$ имеет многомерное нормальное распределение $N(\mathbf{V}^*\boldsymbol{\beta}, \Sigma)$, где $\Sigma^{-1} = \sum_{j=0}^p \gamma_j \mathbf{A}_j$, а столбцы матрицы $\mathbf{V}^*$ являются характеристическими векторами матрицы $\mathbf{A}_j$, соответствующими характеристическим корням λ_{jt} , $t = T - m + 1, \dots, T$, $j = 0, 1, \dots, p$. Тогда характеристическая функция квадратичной формы $P_j^* = (\mathbf{y} - \mathbf{V}^*\mathbf{b})'\mathbf{A}_j(\mathbf{y} - \mathbf{V}^*\mathbf{b})$, $j = 0, 1, \dots, p$, в которой $\mathbf{b}$ является оценкой наименьших квадратов вектора $\boldsymbol{\beta}$, имеет вид

$$(30) \quad \mathcal{E} e^{i(t_0 P_0^* + \dots + t_p P_p^*)} = \prod_{s=1}^{T-m} \left[\left(\sum_{j=0}^p \gamma_j \lambda_{js} \right) / \left(\sum_{j=0}^p (\gamma_j - 2it_j) \lambda_{js} \right) \right]^{1/2}.$$

6.7.4. Распределение сериальных коэффициентов корреляции при двойных корнях для случая независимых наблюдений

Для реализации процедур, полученных в § 6.3 и 6.4, нам необходимы условные распределения статистик r_i при заданных значениях статистик Q_0 и $r_1, \dots, r_{i-1}$. В частности, при построении критерия для проверки нулевой гипотезы $\gamma_1 = 0$ (если задано, что $\gamma_j = 0$, $j > 1$) с определенным уровнем значимости необходимо знать распределение случайной величины r_1 при $\gamma_1 = 0$. Это распределение не зависит от γ_0 . Рассмотрим в этой связи (маргинальное) распределение отдельного сериального коэффициента корреляции, когда $\gamma_1 = \dots = \gamma_p = 0$. Поскольку $A_0 = I$, то наблюдения независимы. Если характеристические корни образуют пары, распределение сериального коэффициента корреляции получить довольно просто. Ввиду того что это распределение может быть использовано и для других сериальных корреляций, проведем исследование в общем виде и поэтому опустим индекс. Позже мы обсудим условное распределение случайной величины r_i при заданных значениях $r_1, \dots, r_{i-1}$.

Предположим, что $T - m = 2H$ и что имеется ровно H различных корней

$$(31) \quad \lambda_{jh} = \lambda_{j, T-m+1-h} = v_h, \quad h = 1, \dots, H.$$

Это будет, например, при нечетном T в случае циклического сериального коэффициента корреляции первого порядка, использующего остатки от среднего значения, причем $m = 1$ и $v_h = \cos 2\pi h/T$. Положим

$$(32) \quad x_h = z_h^2 + z_{T-m+1-h}^2.$$

Тогда

$$(33) \quad r = \frac{\sum_{h=1}^H v_h x_h}{\sum_{h=1}^H x_h}.$$

Если $\gamma_j = 0$, $j > 0$, то случайные величины z_i независимы и имеют единичную дисперсию ($\gamma_0 = 1$). При этом условии случайные величины $x_1, \dots, x_H$ также независимы и каждая из них имеет χ^2 -распределение с 2 степенями свободы. Плотность распределения случайной величины x_h равна $\frac{1}{2} e^{-x_h/2}$, а совместная плотность распределения случайных величин $x_1, \dots, x_H$ имеет вид

$$(34) \quad \left(\frac{1}{2}\right)^H \exp\left(-\frac{1}{2} \sum_{h=1}^H x_h\right), \quad x_h \geq 0, \quad h = 1, \dots, H,$$

и равна 0 в противном случае. Поверхности постоянной плотности представляют собой части гиперплоскостей, заключенные в положительном ортанте:

$$(35) \quad \sum_{h=1}^H x_h = c, \quad x_h \geq 0, \quad h = 1, \dots, H,$$

для $c > 0$. Таким образом, условное распределение случайных величин $x_1, \dots, x_H$ при условии (35) будет равномерным распределением на соответствующем $(H - 1)$ -мерном правильном тетраэдре. (Если $H = 3$, то это распределение равномерно на равностороннем треугольнике.)

Соответствующая условная вероятность для коэффициента r , определенного в (33), равна

$$(36) \quad \Pr \left\{ r \geq R \left| \sum_{h=1}^H x_h = c \right. \right\} = \Pr \left\{ \frac{\sum_{h=1}^H v_h x_h}{\sum_{h=1}^H x_h} \geq R \left| \sum_{h=1}^H x_h = c \right. \right\} = \\ = \Pr \left\{ \sum_{h=1}^H v_h \frac{x_h}{c} \geq R \left| \sum_{h=1}^H \frac{x_h}{c} = 1 \right. \right\}.$$

Здесь величины $x_1/c, \dots, x_H/c$ имеют условное равномерное распределение на $(H - 1)$ -мерном тетраэдре $\sum_{h=1}^H (x_h/c) = 1, x_h/c \geq 0, h = 1, \dots, H$. (Это частный случай теоремы 6.7.2.) Таким образом, условная вероятность (36) не зависит от значения c , и из соображений удобства можно положить $c = 1$. Это показывает, что коэффициент r распределен независимо от суммы $\sum_{h=1}^H x_h$.

Исходя из сказанного, займемся отысканием вероятности

$$(37) \quad \Pr \{ r \geq R \} = \Pr \left\{ \sum_{h=1}^H v_h x_h \geq R \left| \sum_{h=1}^H x_h = 1 \right. \right\},$$

где случайные величины $x_1, \dots, x_H$ предполагаются равномерно распределенными на $(H - 1)$ -мерном тетраэдре

$$(38) \quad \sum_{h=1}^H x_h = 1, \quad x_h \geq 0, \quad h = 1, \dots, H.$$

Вероятность (37) равна отношению $(H - 1)$ -мерного объема тела, образованного пересечением тетраэдра (38) с множеством

$$(39) \quad \sum_{h=1}^H v_h x_h \geq R,$$

к $(H - 1)$ -мерному объему правильного тетраэдра (38).

Вершины правильного тетраэдра (38) находятся в точках $X_1, \dots, X_H$, причем все координаты точки X_j равны нулю, за исключением j -й координаты, которая равна 1, $j = 1, \dots, H$. Гиперплоскость $\sum_{h=1}^H v_h x_h = R$ пересекает ребро $X_i X_j$ (определяемое соотношениями $x_i + x_j = 1, x_k = 0, k \neq i, k \neq j, i \neq j$) или его продолжение за границу положительного ортанта в некоторой точке W_{ij} с координатами

$$(40) \quad x_i = \frac{R - v_j}{v_i - v_j}, \quad x_j = \frac{v_i - R}{v_i - v_j}, \quad x_k = 0, \quad k \neq i, k \neq j, i \neq j.$$

(Заметим, что $W_{ij} = W_{ji}$.) Характеристические числа удобно здесь перенумеровать так, чтобы

$$(41) \quad v_H < v_{H-1} < \dots < v_2 < v_1.$$

Отметим, что при этом $v_H \leq \sum_{h=1}^H v_h x_h \leq v_1$. Если $v_2 \leq R \leq v_1$, то указанная гиперплоскость пересекает каждое из ребер $X_1 X_j$, $j = 2, \dots, H$, в промежутке между вершинами X_1 и X_j . Иными словами, в этом случае координаты x_1 и x_j точек W_{1j} неотрицательны и отсекаемое от тетраэдра (38) условием $r \geq R$ множество является тетраэдром с вершинами $X_1, W_{12}, \dots, W_{1H}$. Покажем, что для произвольных значений R отсекаемое условием $r \geq R$ множество также может быть представлено с помощью тетраэдров. Тогда вероятность (37) можно выразить через объемы тетраэдров. См. рис. 6.1 и 6.2.

Лемма 6.7.2. Пусть T представляет собой $(H - 1)$ -мерный тетраэдр с вершинами в точках $V_1, V_2, V_3, \dots, V_H$, а $(H - 1)$ -мерный тетраэдр T^* имеет вершинами точки $V_1, V_2^*, V_3, \dots, V_H$, где точка

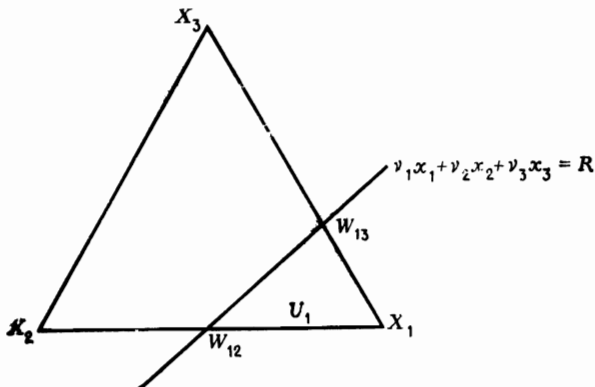


Рис. 6.1.

Области для $v_2 \leq R \leq v_1$.

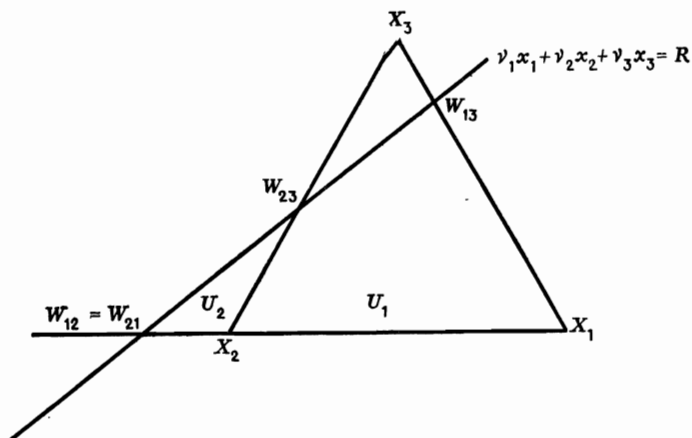


Рис. 6.2.

Области для $v_3 \leq R \leq v_2$.

V_2^* лежит на прямой V_1V_2 . Пусть длина отрезка $V_1V_2^*$ равна k_2^* длине отрезка V_1V_2 . Тогда $(H-1)$ -мерный объем тетраэдра T^* равен k_2^* , умноженному на $(H-1)$ -мерный объем тетраэдра T . ■

Доказательство. Пусть $y_1, \dots, y_{H-2}$ — произвольные (прямоугольные) координаты на $(H-2)$ -мерной гиперплоскости, ортогональной к прямой V_1V_2 . Пусть $h(y_1, \dots, y_{H-2})$ представляет собой длину части прямой, принадлежащей тетраэдру T , проходящей через точку $(y_1, \dots, y_{H-2})$ параллельно ребру V_1V_2 , а $h^*(y_1, \dots, y_{H-2})$ — длину части той же прямой, принадлежащей тетраэдру T^* .

(Если пересечение прямой с тетраэдром отсутствует, то соответствующая длина равна нулю.) Тогда $h^*(y_1, \dots, y_{H-2}) = k_2^* h(y_1, \dots, y_{H-2})$ и объем тетраэдра T^* равен

$$\begin{aligned}
 (42) \quad V(T^*) &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h^*(y_1, \dots, y_{H-2}) dy_1 \dots dy_{H-2} = \\
 &= k_2^* \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h(y_1, \dots, y_{H-2}) dy_1 \dots dy_{H-2} = \\
 &= k_2^* V(T). \blacksquare
 \end{aligned}$$

Лемма 6.7.3. Пусть T представляет собой $(H-1)$ -мерный тетраэдр с вершинами в точках $V_1, V_2, \dots, V_H$, а $(H-1)$ -мерный тетраэдр T^* имеет вершинами точки $V_1, V_2^*, \dots, V_H$, где точка V_j^*

лежит на прямой V_1V_j , $j = 2, \dots, H$. Пусть длина отрезка $V_1V_j^*$ равна k_j^* длинам отрезка V_1V_j , $j = 2, \dots, H$. Тогда $(H - 1)$ -мерный объем тетраэдра T^* равен $k_2^* \dots k_H^*$, умноженному на $(H - 1)$ -мерный объем тетраэдра T .

Доказательство. Эта лемма вытекает из леммы 6.7.2 по индукции. ■

Пусть множество S представляет собой пересечение (38) и (39).

Потребуем теперь, чтобы $\sum_{h=1}^H x_h = 1$ и предположим, что $2 \leq m \leq H - 1$. Если $v_{m+1} \leq R \leq v_m^-$, то вершины $X_1, \dots, X_m$ лежат в S , а вершины $X_{m+1}, \dots, X_H$ вне S . (Строго говоря, в S лежит и X_{m+1} , если $R = v_{m+1}$.) Определим множества $U_1, \dots, U_m$ следующим образом:

$$\begin{aligned}
 U_1: \quad & \sum_{h=1}^H v_h x_h \geq R, \quad x_2 \geq 0, \dots, x_H \geq 0, \\
 U_2: \quad & \sum_{h=1}^H v_h x_h \geq R, \quad x_1 < 0, x_3 \geq 0, \dots, x_H \geq 0, \\
 (43) \quad U_3: \quad & \sum_{h=1}^H v_h x_h \geq R, \quad x_1 < 0, x_2 < 0, x_4 \geq 0, \dots \\
 & \vdots \\
 & \dots, x_H \geq 0 \\
 U_{m-1}: \quad & \sum_{h=1}^H v_h x_h \geq R, \quad x_1 < 0, \dots, x_{m-2} < 0, x_m \geq \\
 & \geq 0, \dots, x_H \geq 0, \\
 U_m: \quad & \sum_{h=1}^H v_h x_h \geq R, \quad x_1 < 0, \dots, x_{m-1} < 0, x_{m+1} \geq \\
 & \geq 0, \dots, x_H \geq 0.
 \end{aligned}$$

Каждое из них является тетраэдром с вершинами

$$\begin{aligned}
 U_1: & X_1, W_{12}, W_{13}, \dots, W_{1H}, \\
 U_2: & W_{21}, X_2, W_{23}, \dots, W_{2H}, \\
 (44) \quad U_3: & W_{31}, W_{32}, X_3, \dots, W_{3H}, \\
 & \vdots \\
 U_{m-1}: & W_{m-1,1}, W_{m-1,2}, W_{m-1,3}, \dots, W_{m-1,m-2}, X_{m-1}, \\
 & W_{m-1,m}, \dots, W_{m-1,H}, \\
 U_m: & W_{m1}, W_{m2}, W_{m3}, \dots, W_{m,m-1}, X_m, W_{m,m+1}, \dots, W_{mH}.
 \end{aligned}$$

Поскольку $x_j \geq 0$ в U_{j-1} и $x_j < 0$ в U_{j+i} , $i \geq 1$, то

$$(45) \quad U_{j-1} \cap U_{j+i} = 0, \quad i = 1, \dots, m-j, \quad j = 2, \dots, m-1 (\geq 2)$$

пустое множество. Кроме того,

$$(46) \quad U_j \subset U_{j-1} \cup U_{j+1}, \quad j = 2, \dots, m-1 (\geq 2).$$

Поскольку $x_1 < 0, \dots, x_{m-1} < 0$ и $x_{m+1} \geq 0, \dots, x_H \geq 0$ в U_m , соотношения $\sum_{h=1}^H v_h x_h \geq R \geq v_{m+1}$ и $\sum_{h=1}^H x_h = 1$ влекут за собой неравенство

$$(47) \quad \sum_{h=m}^H (v_h - v_{m+1}) x_h \geq \sum_{h=1}^{m-1} (v_{m+1} - v_h) x_h > 0,$$

а из него вытекает, что $x_m > 0$ в U_m , так что

$$(48) \quad U_m \subset U_{m-1}.$$

ЛЕММА 6.7.4.

$$(49) \quad S = U_1 \cup U_3 \cup \dots \cup U_m - (U_2 \cup U_4 \cup \dots \cup U_{m-1}),$$

если m нечетное,

$$(50) \quad S = U_1 \cup U_3 \cup \dots \cup U_{m-1} - (U_2 \cup U_4 \cup \dots \cup U_m),$$

если m четное.

ДОКАЗАТЕЛЬСТВО. Из соотношений (46) и (48) вытекает, что обе операции вычитания осмысленны. Поскольку $S \subset U_1$ и $S \cap U_j = 0$, $j > 1$, то S содержится в правой части (49) или (50). С другой стороны, всякая точка, принадлежащая правой части одного из этих соотношений, должна содержаться в S . Действительно, всякая точка из правой части должна содержаться в одном из множеств U_{2j-1} (поскольку U_1, U_3 и т. д. — непересекающиеся множества) и не должна содержаться ни в каком из U_{2j} . Ввиду соотношений (46) и (48) такая точка обязана тогда принадлежать части множества U_1 , не содержащей точек множества U_2 , а потому лежит в S . Это и доказывает лемму. ■

ЛЕММА 6.7.5. Отношение $(H-1)$ -мерного объема тетраэдра U_i к $(H-1)$ -мерному объему тетраэдра (38) равно

$$(51) \quad (-1)^{i-1} \prod_{\substack{j=1 \\ j \neq i}}^H \frac{v_i - R}{v_i - v_j}.$$

Доказательство. Координаты вершины W_{ij} ($j \neq i$) тетраэдра U_i указаны в (40). Отношение длины ребра $X_i W_{ij}$ к длине ребра $X_i X_j$ равно $-(v_i - R)/(v_i - v_j)$ для $j < i$ и $(v_i - R)/(v_i - v_j)$ для $j > i$. Утверждение леммы вытекает теперь из леммы 6.7.3. ■

ТЕОРЕМА 6.7.4. *Распределение коэффициента r дается соотношениями*

$$\Pr \{r \geq R\} = 1, \quad R \leq v_H,$$

$$(52) \quad \Pr \{r \geq R\} = \sum_{i=1}^m \prod_{\substack{j=1 \\ j \neq i}}^H \frac{v_i - R}{v_i - v_j}, \quad v_{m+1} \leq R \leq v_m, \quad m = 1, \dots, H-1,$$

$$\Pr \{r \geq R\} = 0, \quad v_1 \leq R.$$

Доказательство. Утверждение теоремы вытекает из лемм 6.7.4, 6.7.5 и соотношения (45). ■

Следствие 6.7.3. *Плотность распределения коэффициента r равна*

$$(53) \quad (H-1) \sum_{i=1}^m \frac{(v_i - r)^{H-2}}{\prod_{\substack{j=1 \\ j \neq i}}^H (v_i - v_j)}, \quad v_{m+1} \leq r \leq v_m, \quad m = 1, \dots, H-1,$$

и равна нулю вне интервала (v_H, v_1) .

Распределение сериального коэффициента корреляции для случая, когда кроме двойных корней имеется еще и один простой, можно получить из (52). Представляется удобным доказать некоторый более общий результат, а из него уже в качестве следствия получить интересующее нас распределение. Пусть $\mu_1, \dots, \mu_M, \mu_{M+1}$ обозначают здесь произвольные веса.

ТЕОРЕМА 6.7.5. *Положим*

$$(54) \quad s = \frac{\sum_{j=1}^M \mu_j z_j^2}{\sum_{j=1}^M z_j^2}, \quad t = \frac{\sum_{j=1}^M \mu_j z_j^2 + \mu_{M+1} \sum_{j=M+1}^{M+L} z_j^2}{\sum_{j=1}^{M+L} z_j^2},$$

где $z_1, \dots, z_{M+L}$ — независимые случайные величины, каждая из которых распределена по стандартному нормальному закону $N(0, 1)$.

Тогда

$$(55) \Pr \{t \geq T\} =$$

$$= \int_0^{\infty} \Pr \{s \geq T + (T - \mu_{M+1}) u\} \frac{\Gamma \{(M+L)/2\}}{\Gamma(M/2) \Gamma(L/2)} \frac{u^{L/2-1}}{(1+u)^{(M+L)/2}} du.$$

Доказательство. Имеем

$$\begin{aligned} (56) \Pr \{t \geq T\} &= \Pr \left\{ \sum_{j=1}^M \mu_j z_j^2 + \mu_{M+1} \sum_{j=M+1}^{M+L} z_j^2 \geq T \sum_{j=1}^{M+L} z_j^2 \right\} = \\ &= \Pr \left\{ \sum_{j=1}^M \mu_j z_j^2 \geq T \sum_{j=1}^M z_j^2 + (T - \mu_{M+1}) \sum_{j=M+1}^{M+L} z_j^2 \right\} = \\ &= \Pr \left\{ \frac{\sum_{j=1}^M \mu_j z_j^2}{\sum_{j=1}^M z_j^2} \geq T + (T - \mu_{M+1}) \frac{\sum_{j=M+1}^{M+L} z_j^2}{\sum_{j=1}^M z_j^2} \right\} = \\ &= \Pr \{s \geq T + (T - \mu_{M+1}) u\}, \end{aligned}$$

где u есть отношение двух независимых случайных величин, имеющих распределения χ^2 с L и M степенями свободы соответственно. Таким образом, случайная величина $(M/L)u$ имеет F -распределение с L и M степенями свободы. В (56) использован тот факт, что отношение s распределено независимо от $\sum_{j=1}^M z_j^2$.

Следствие 6.7.4. Если $L = 1$, а $M = 2H$, то

$$(57) \Pr \{t \geq T\} =$$

$$= \int_0^{\infty} \Pr \{s \geq T + (T - \mu_{2H+1}) u\} \frac{\Gamma \left(H + \frac{1}{2} \right) u^{-1/2}}{\Gamma(H) \Gamma \left(\frac{1}{2} \right) (1+u)^{H+1/2}} du.$$

Для того чтобы вычислить (57) при s , равном r , используется частный случай следующей леммы.

Лемма 6.7.6.

$$\begin{aligned} (58) \int_0^{c/d} (c - du)^{M/2-1} \frac{\Gamma \{(M+L)/2\}}{\Gamma(M/2) \Gamma(L/2)} u^{L/2-1} (1+u)^{-(M+L)/2} du = \\ = c^{(M+L)/2-1} (c+d)^{-L/2}. \end{aligned}$$

Доказательство. Произведя в левой части (58) замену

$$(59) \quad w = \frac{c+d}{c} \cdot \frac{u}{1+u},$$

приведем ее к виду

$$(60) \quad \frac{\Gamma\{(M+L)/2\}}{\Gamma(M/2)\Gamma(L/2)} c^{M/2-1} \left(\frac{c}{c+d}\right)^{L/2} \int_0^1 (1-w)^{M/2-1} w^{L/2-1} dw.$$

Но последнее выражение есть в точности правая часть (58), поскольку входящий в него интеграл представляет собой бета-функцию $B(M/2, L/2)$.

Следствие 6.7.5.

$$(61) \quad \int_0^{c/d} (c-du)^{H-1} \frac{\Gamma\left(H + \frac{1}{2}\right)}{\Gamma(H) \Gamma\left(\frac{1}{2}\right)} u^{-1/2} (1+u)^{-(H+1/2)} du = \\ = c^{H-1/2} (c+d)^{-1/2}.$$

Теорема 6.7.6. Пусть

$$(62) \quad r = \frac{\sum_{h=1}^H v_h (z_{2h-1}^2 + z_{2h}^2) + v_{H+1} z_{2H+1}^2}{\sum_{h=1}^{2H+1} z_h^2},$$

где $z_1, \dots, z_{2H+1}$ — независимые случайные величины, каждая из которых распределена по стандартному нормальному закону $N(0, 1)$, и $v_{H+1} < v_H < \dots < v_1$. Тогда

$$(63) \quad \Pr\{r \geq R\} = \sum_{i=1}^n \frac{(v_i - R)^{H-1/2}}{\sqrt{v_i - v_{H+1}} \prod_{\substack{j=1 \\ j \neq i}}^H (v_i - v_j)}, \quad v_{n+1} \leq R \leq v_n, \\ n = 1, \dots, H.$$

Доказательство. Заменим в правой части (52) R на $R + (R - v_{H+1})u$. Если $v_{n+1} \leq R \leq v_n$, то $v_{n+1} \leq R + (R - v_{H+1})u$, ввиду того что $v_{H+1} \leq R$ и $u \geq 0$. Для $v_{m+1} \leq R + (R - v_{H+1})u \leq v_m$, $m = 1, \dots, n-1$, имеем

$$(64) \quad \frac{v_{m+1} - R}{R - v_{H+1}} \leq u \leq \frac{v_m - R}{R - v_{H+1}},$$

а для $v_{n+1} \leq R + (R - v_{H+1})u \leq v_n$ справедливо $0 \leq u \leq (v_n - R)/(R - v_{H+1})$. Отсюда, используя (57), получаем перестанов-

кой порядка суммирования по m и i для $n = 1, \dots, H-1$:

$$\begin{aligned}
 (65) \quad \Pr \{r \geq R\} &= \\
 &= \sum_{m=1}^{n-1} \int_{(v_{m+1}-R)/(R-v_{H+1})}^{(v_m-R)/(R-v_{H+1})} \sum_{i=1}^m \frac{[v_i - R - (R - v_{H+1})u]^{H-1}}{\prod_{\substack{j=1 \\ j \neq i}}^H (v_i - v_j)} \times \\
 &\quad \times \frac{\Gamma\left(H + \frac{1}{2}\right)}{\Gamma(H) \Gamma\left(\frac{1}{2}\right)} \frac{u^{-1/2}}{(1+u)^{H+1/2}} du + \\
 &+ \int_0^{(v_n-R)/(R-v_{H+1})} \sum_{i=1}^n \frac{[v_i - R - (R - v_{H+1})u]^{H-1}}{\prod_{\substack{j=1 \\ j \neq i}}^H (v_i - v_j)} \times \\
 &\quad \times \frac{\Gamma\left(H + \frac{1}{2}\right)}{\Gamma(H) \Gamma\left(\frac{1}{2}\right)} \frac{u^{-1/2}}{(1+u)^{H+1/2}} du = \\
 &= \sum_{i=1}^n \int_0^{(v_i-R)/(R-v_{H+1})} \frac{[v_i - R - (R - v_{H+1})u]^{H-1}}{\prod_{\substack{j=1 \\ j \neq i}}^H (v_i - v_j)} \times \\
 &\quad \times \frac{\Gamma\left(H + \frac{1}{2}\right)}{\Gamma(H) \Gamma\left(\frac{1}{2}\right)} \frac{u^{-1/2}}{(1+u)^{H+1/2}} du.
 \end{aligned}$$

Это доказывает (63) для $n = 1, \dots, H-1$. Чтобы доказать справедливость (63) и для $n = H$, необходимо произвести интегрирование по тем значениям u , для которых $R + (R - v_{H+1})u \leq v_H$. Чтобы обосновать справедливость (65) для этого случая, убедимся в том, что (52) при $m = H$ выполняется для $R \leq v_H$ и соответствует соотношению $\Pr \{r \geq R\} = 1$ для $R \leq v_H$. Последнее же есть частный случай следующей леммы. ■

ЛЕММА 6.7.7.

$$(66) \quad \sum_{i=1}^H \frac{(v_i - x)^{H-1}}{\prod_{\substack{j=1 \\ j \neq i}}^H (v_i - v_j)} \equiv 1.$$

Доказательство. Если в (52) положить $R = v_H$, получим $\Pr \{r \geq v_H\} = 1$. Если в этом соотношении заменить H на $H + 1$, а $v_1, \dots, v_H$ на $v_1, \dots, v_H, x$, оно примет вид

$$(67) \quad 1 = \sum_{i=1}^H \frac{(v_i - x)^H}{(v_i - x) \prod_{\substack{j=1 \\ j \neq i}}^H (v_i - v_j)},$$

что совпадает с (66). Поскольку это имеет место для всех $x \leq v_H$ и левая часть (66) представляет собой полином от x , то равенство (66) является тождеством и сохраняется для всех x . ■

Следует отметить, что вероятность $\Pr \{r < R\} = 1 - \Pr \{r \geq R\}$ может быть получена из леммы 6.7.7 и теоремы 6.7.4 и выражается формулой, подобной (52). (См. упр. 35.) Следует также отметить, что теорему 6.7.4 можно доказать по индукции, используя теорему 6.7.5. Впервые соотношение (52) было найдено именно таким методом Р. Андерсоном (1942) (при выводе совместного распределения величин r и Q_0). (См. упр. 36.)

Другой метод доказательства теоремы 6.7.4 состоит в обращении соответствующей характеристической функции. (См. Купменс (1942).) Именно

$$(68) \quad \Pr \{r \geq R\} = \Pr \left\{ \sum_{i=1}^H (v_i - R) x_i \geq 0 \right\}$$

и

$$(69) \quad \mathbb{E} \exp \left[it \sum_{i=1}^H (v_i - R) x_i \right] = \left(\prod_{i=1}^H |1 - 2it(v_i - R)| \right)^{-1}.$$

Ввиду того что корни λ_i образуют пары и имеется H различных корней v_i , то характеристическая функция (69) имеет простые полюсы и для ее обращения может быть использован обычный метод вычетов. (См. упр. 37.)

6.7.5. Распределения циклических серийных коэффициентов корреляции

В циклической модели характеристические корни матрицы $\mathbf{A}_1$ равны $\cos 2\pi t/T$, $t = 1, \dots, T$. Все они образуют пары, за исключением простого корня $\cos 2\pi T/T = 1$ и (в случае четного T) простого корня $\cos 2\pi (T/2)/T = -1$. Поскольку единичный вектор $\mathbf{e}$ является здесь характеристическим вектором матрицы $\mathbf{A}_1$, соответствующим характеристическому корню 1, то веса для остатков

от среднего значения равны

$$(70) \quad v_j = \cos \frac{2\pi j}{T}, \quad j = 1, \dots, (T-1)/2 = H, \text{ если } T \text{ нечетное,}$$

$$(71) \quad v_j = \cos \frac{2\pi j}{T}, \quad j = 1, \dots, T/2 - 1 = H, \quad v_{H+1} = -1, \\ \text{если } T \text{ четное.}$$

ТЕОРЕМА 6.7.7. Если случайные величины $y_1, \dots, y_T$ независимы и каждая из них распределена по закону $N(\mu, \sigma^2)$, а

$$(72) \quad r_1^* = \frac{\sum_{t=1}^T (y_t - \bar{y})(y_{t-1} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} = \frac{\sum_{t=1}^T y_t y_{t-1} - T\bar{y}^2}{\sum_{t=1}^T y_t^2 - T\bar{y}^2}, \quad y_0 \equiv y_T,$$

то

$$(73) \quad \Pr \{r_1^* \geq R\} = \sum_{i=1}^m \frac{(v_i - R)^{(T-3)/2}}{\alpha_i}, \quad v_{m+1} \leq R \leq v_m,$$

где $v_i = \cos 2\pi i/T$,

$$(74) \quad \alpha_i = \prod_{\substack{j=1 \\ j \neq i}}^{(T-1)/2} (v_i - v_j), \quad T = 3, 5, \dots,$$

$$(75) \quad \alpha_i = \sqrt{v_i + 1} \prod_{\substack{j=1 \\ j \neq i}}^{T/2-1} (v_i - v_j), \quad T = 4, 6, \dots,$$

а $m = 1, \dots, (T-3)/2$, если T нечетное, или $m = 1, \dots, T/2 - 1$ если T четное.

Р. Андерсон (1942) указал это распределение и вычислил таблицу значений R , при которых функция распределения принимает соответственно значения 0.01, 0.05, 0.95 и 0.99. Значения R приведены в табл. 6.1.

Распределение отношения $\sum_{t=1}^T y_t y_{t-1} / \sum_{t=1}^T y_t^2$ можно вывести из теоремы 6.7.6, если только T нечетное, поскольку при этом дважды встречаются все корни, за исключением лишь одного простого корня $\cos 2\pi T/T = 1$.

ТЕОРЕМА 6.7.8. Если случайные величины $y_1, \dots, y_T$ независимы и каждая из них распределена по закону $N(0, \sigma^2)$, причем T нечетное,

Таблица 6.1

ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ ЦИКЛИЧЕСКОГО СЕРИАЛЬНОГО
КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ, СТРОЯЩЕГОСЯ ПО ОСТАТКАМ ОТ
ВЫБОРОЧНОГО СРЕДНЕГО *

T	0.01	0.05	0.95	0.99
5	-0.798	-0.753	0.253	0.297
6	-0.863	-0.708	0.345	0.447
7	-0.799	-0.674	0.370	0.510
8	-0.764	-0.625	0.371	0.531
9	-0.737	-0.593	0.366	0.533
10	-0.705	-0.564	0.360	0.525
11	-0.679	-0.539	0.353	0.515
12	-0.655	-0.516	0.348	0.505
13	-0.634	-0.497	0.341	0.495
14	-0.615	-0.479	0.335	0.485
15	-0.597	-0.462	0.328	0.475
20	-0.524	-0.399	0.299	0.432
25	-0.473	-0.356	0.276	0.398
30	-0.433	-0.324	0.257	0.370
(35)	-0.401	-0.300	0.242	0.347
(40)	-0.376	-0.279	0.229	0.329
45	-0.356	-0.262	0.218	0.313
(50)	-0.339	-0.248	0.208	0.301
(55)	-0.324	-0.236	0.199	0.289
(60)	-0.310	-0.225	0.191	0.278
(65)	-0.298	-0.216	0.184	0.268
(70)	-0.287	-0.207	0.178	0.259
75	-0.276	-0.201	0.174	0.250

* Круглые скобки указывают на то, что соответствующие значения получены графической интерполяцией. Некоторые значения являются исправлениями Р. Андерсона соответствующих значений, имеющихся у Диксона (1944).

то для

$$(76) \quad r_1 = \frac{\sum_{t=1}^T y_t y_{t-1}}{\sum_{t=1}^T y_t^2}, \quad y_0 \equiv y_T,$$

имеем

$$(77) \quad \Pr \{r_1 \leq R^*\} = \sum_{i=m}^{(T-1)/2} \frac{(R^* - v_i^*)^{(T-2)/2}}{\alpha_i^*}, \quad v_m^* \leq R^* \leq v_{m-1}^*, \quad m = 1, \dots, (T-1)/2,$$

где $v^* = \cos 2\pi i/T$, $i = 0, \dots, (T-1)/2$, и

$$(78) \quad \alpha_i^* = \sqrt{1 - v_i^*} \prod_{\substack{j=1 \\ j \neq i}}^{(T-1)/2} (v_j^* - v_i^*).$$

ДОКАЗАТЕЛЬСТВО. Утверждение теоремы вытекает из теоремы 6.7.6, если в последней заменить v_i , v_{H+1} , R и H на $-v_{H+1-i}^*$, $-v_0^* = -1$, $-R^*$ и $(T-1)/2$ соответственно. [Р. Андерсон (1941).]

6.7.6. Другие распределения сериальных коэффициентов корреляции

В случае сериального коэффициента корреляции, использующего сумму квадратов последовательных разностей остатков от среднего значения, именно

$$(79) \quad r_1^* = \frac{\sum_{t=2}^T (y_t - \bar{y})(y_{t-1} - \bar{y}) + \frac{1}{2}(y_1 - \bar{y})^2 + \frac{1}{2}(y_T - \bar{y})^2}{\sum_{t=1}^T (y_t - \bar{y})^2},$$

характеристические корни равны $\cos \pi t/T$, $t = 1, \dots, T-1$, $T = 2, 3, \dots$, причем все различны. Положим

$$(80) \quad s = \frac{\sum_{t=1}^{T-1} \lambda_t z_t^2}{\sum_{t=1}^{T-1} z_t^2},$$

где λ_t — произвольные веса, удовлетворяющие условию $\lambda_{T-1} \leq \lambda_{T-2} \leq \dots \leq \lambda_1$. Нейман (1941) показал, что если $T-1 = 2H$ и $f(s)$ — плотность распределения s , то

$$(81) \quad \frac{d^{H-1}}{ds^{H-1}} f(s) = \begin{cases} 0, & \lambda_{2h+1} < s < \lambda_{2h}, \quad h = 1, \dots, H-1, \\ \frac{(-1)^{H-h} (H-1)!}{\pi \sqrt{-\prod_{t=1}^{2H} (s - \lambda_t)}}, & \lambda_{2h} < s < \lambda_{2h-1}, \quad h = 1, \dots, H, \end{cases}$$

для таких h , для которых указанные неравенства допустимы. (Например, если $\lambda_1 = \lambda_2 > \lambda_3 = \lambda_4 > \dots > \lambda_{2H-1} = \lambda_{2H}$, то неравенство $\lambda_{2h} < s < \lambda_{2h-1}$ невозможно.) В частности, в интервале $\lambda_{2h+1} < s < \lambda_{2h}$ плотность $f(s)$ будет полиномом степени не более $H-2$.

(В случае если все корни двойные и притом различные, этот результат согласуется с результатами разд. 6.7.4.)

Нейманом было показано, что если все веса λ_i различны, то производные $d^k f(s)/ds^k$ непрерывны на отрезке $\lambda_{2H} \leq s < \lambda_1$, $k = 0, 1, \dots, H-2$ (даже в точках $s = \lambda_i$, $i = 1, \dots, 2H$). В принципе соотношение (81) можно проинтегрировать и получить плотность $f(s)$. Постоянные интегрирования определяются тем фактом, что производные порядка, меньшего чем $H-1$, в смежных интервалах имеют совпадающие значения в общих концевых точках этих интервалов (и равны 0 при $s = \lambda_{2H}$ и $s = \lambda_1$). В интервалах $\lambda_{2h} < s < \lambda_{2h-1}$ знаменатель выражения для $(H-1)$ -й производной от плотности $f(s)$ представляет собой квадратный корень из полинома от s степени $2H$. Если $H = 1$ ($T = 3$), то эта плотность равна

$$(82) \quad f(s) = \begin{cases} \frac{1}{\pi \sqrt{\left(\frac{\lambda_1 - \lambda_2}{2}\right)^2 - \left(s - \frac{\lambda_1 + \lambda_2}{2}\right)^2}}, & \lambda_2 < s < \lambda_1, \\ 0 & \text{в противном случае,} \end{cases}$$

а соответствующая ей функция распределения имеет вид

$$(83) \quad F(s) = \int_{\lambda_2}^s f(t) dt = \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{2s - (\lambda_1 + \lambda_2)}{\lambda_1 - \lambda_2}, \quad \lambda_2 < s < \lambda_1,$$

и $F(s) = 1$ для $s \geq \lambda_1$. Если $H > 1$, то в выражении для плотности появляются эллиптические, гиперэллиптические и другие более сложные интегралы.

Если $T-1$ нечетное, то соответствующее распределение можно получить с помощью следствия 6.7.4 из распределения для четного $T-1$.

В случае серийного коэффициента корреляции (79) имеем $\lambda_t = \cos \pi t/T$, $t = 1, \dots, T-1$. Распределение величины r_1^* симметрично ввиду того, что $\lambda_{T-t} = -\lambda_t$, $t = 1, \dots, T-1$. Все корни образуют пары (положительное и отрицательное значения), если только T не является четным и $\lambda_{T/2} = 0$. Возможные значения коэффициента r_1^* заполняют интервал $(-\cos \pi/T, \cos \pi/T)$.

6.7.7. Моменты и производящие функции моментов

Если $r_1^* = Q_1^*/Q_0^*$, то

$$(84) \quad \mathbb{E} r_1^{*h} = \mathbb{E} Q_1^{*h} Q_0^{*-h}.$$

Моменты величин Q_1^* и Q_0^* могут быть найдены из характеристических функций этих величин или из производящих функций их моментов (моменты положительного порядка получаются дифференциро-

ванием, а моменты отрицательного порядка — интегрированием). Однако при нулевой гипотезе независимости ($\gamma_1 = 0$) более эффективным является вычисление этих моментов с использованием факта независимости r_1^* и Q_0^* (теорема 6.7.2).

ЛЕММА 6.7.8. Если $r_1^* = Q_1^*/Q_0^*$ и $\gamma_1 = 0$, то

$$(85) \quad \mathcal{E}r_1^{*h}Q_0^{*g} = \mathcal{E}r_1^{*h}\mathcal{E}Q_0^{*g}.$$

ТЕОРЕМА 6.7.9. Если $r_1^* = Q_1^*/Q_0^*$ и $\gamma_1 = 0$, то

$$(86) \quad \mathcal{E}r_1^{*h} = \frac{\mathcal{E}Q_1^{*h}}{\mathcal{E}Q_0^{*h}}.$$

ДОКАЗАТЕЛЬСТВО. Это следует непосредственно из леммы 6.7.8, если в последней положить $g = h$. ■

Если $r_1 = Q_1/Q_0$ и $\gamma_1 = 0$, то справедливо аналогичное соотношение $\mathcal{E}r_1^h = \mathcal{E}Q_1^h/\mathcal{E}Q_0^h$.

Обычно квадратичная форма Q_0^* имеет канонический вид $\sum_{i=1}^{T-1} z_i^2$. При этом она распределена по закону χ^2 с $T - 1$ степенями свободы. Аналогично, Q_0 имеет канонический вид $\sum_{i=1}^T z_i^2$ и распределена по закону χ^2 с T степенями свободы. Соответствующие моменты равны

$$(87) \quad \begin{aligned} \mathcal{E}Q_0^{*h} &= 2^h \frac{\Gamma\{(T-1)/2 + h\}}{\Gamma\{(T-1)/2\}}, & -\frac{(T-1)}{2} < h, \\ \mathcal{E}Q_0^h &= 2^h \frac{\Gamma\{T/2 + h\}}{\Gamma\{T/2\}}, & -\frac{T}{2} < h. \end{aligned}$$

Для отыскания $\mathcal{E}Q_1^h$ удобнее использовать производящие функции моментов. Если $Q_1 = \mathbf{y}'\mathbf{A}_1\mathbf{y}$, где $\mathbf{A}_1$ — произвольная симметрическая матрица с характеристическими корнями $\lambda_1, \dots, \lambda_T$ и $\mathbf{A}_0 = \mathbf{I}$, то из следствий 6.7.1 и 6.7.2 получаем для достаточно малых значений θ

$$(88) \quad \mathcal{E}e^{\theta Q_1} = |\mathbf{I} - 2\theta\mathbf{A}_1|^{-1/2} = \left(\prod_{i=1}^T (1 - 2\theta\lambda_i) \right)^{-1/2}.$$

(См. упр. 29.) Вычислим (88) для первых трех частных моделей, определенных в § 6.5.

ЛЕММА 6.7.9. Пусть в матрице $\mathbf{A}_1$ элементы, расположенные на диагоналях, смежных с главной, равны $1/2$, а остальные элементы

равны нулю. Тогда определитель $D_T = |\mathbf{I} - 2\theta\mathbf{A}_1|$ равен

$$(89) \quad D_T = \frac{1}{\sqrt{1-4\theta^2}} \left[\left(\frac{1 + \sqrt{1-4\theta^2}}{2} \right)^{T+1} - \left(\frac{1 - \sqrt{1-4\theta^2}}{2} \right)^{T+1} \right],$$

$$T = 1, 2, \dots$$

Доказательство. Разлагая определитель D_T по элементам первой строки, получаем

$$(90) \quad D_T = \begin{vmatrix} 1 - \theta & 0 & 0 & \dots & 0 & 0 \\ -\theta & 1 - \theta & 0 & \dots & 0 & 0 \\ 0 & -\theta & 1 - \theta & \dots & 0 & 0 \\ 0 & 0 & -\theta & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 - \theta \\ 0 & 0 & 0 & 0 & \dots & -\theta & 1 \end{vmatrix} =$$

$$= D_{T-1} + \theta \begin{vmatrix} -\theta & -\theta & 0 & \dots & 0 & 0 \\ 0 & 1 - \theta & \dots & 0 & 0 \\ 0 & -\theta & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 - \theta \\ 0 & 0 & 0 & \dots & -\theta & 1 \end{vmatrix} =$$

$$= D_{T-1} - \theta^2 D_{T-2}, \quad T = 3, 4, \dots,$$

где $D_1 = 1$. Таким образом, D_T удовлетворяет разностному уравнению

$$(91) \quad D_T - D_{T-1} + \theta^2 D_{T-2} = 0.$$

Характеристическое уравнение последнего

$$(92) \quad x^2 - x + \theta^2 = 0$$

имеет корни $(1 \pm \sqrt{1-4\theta^2})/2$. Поэтому решением (91) для $\theta^2 \neq 1/4$ будет

$$(93) \quad D_T = c_1 \left(\frac{1 + \sqrt{1-4\theta^2}}{2} \right)^T + c_2 \left(\frac{1 - \sqrt{1-4\theta^2}}{2} \right)^T, \quad T = 1, 2, \dots$$

Постоянные c_1 и c_2 определяются из условий $D_1 = 1$ и $D_2 = 1 - \theta^2$ для $\theta \neq 0$, $T = 1, 2, \dots$. Это и доказывает лемму, необходимую для изучения модели разд. 6.5.4. ■

ЛЕММА 6.7.10. Определитель $\Delta_T = |\mathbf{I} - 2\theta\mathbf{A}_1|$, где матрица $\mathbf{A}_1$ определяется сериальным циклическим коэффициентом корреляции

первого порядка, равен

$$(94) \quad \Delta_T = \left(\frac{1 + \sqrt{1 - 4\theta^2}}{2} \right)^T + \left(\frac{1 - \sqrt{1 - 4\theta^2}}{2} \right)^T - 2\theta^T =$$

$$= \left[\left(\frac{\sqrt{1 + 2\theta} + \sqrt{1 - 2\theta}}{2} \right)^T - \left(\frac{\sqrt{1 + 2\theta} - \sqrt{1 - 2\theta}}{2} \right)^T \right]^2,$$

$$T = 2, 3, \dots$$

Доказательство. Разлагая определитель Δ_T (сначала по элементам его первой строки, а затем по элементам первых столбцов полученных определителей), получаем

$$(95) \quad \Delta_T = \begin{vmatrix} 1 - \theta & 0 & 0 & \dots & 0 & 0 & -\theta \\ -\theta & 1 - \theta & 0 & \dots & 0 & 0 & 0 \\ 0 & -\theta & 1 - \theta & \dots & 0 & 0 & 0 \\ 0 & 0 & -\theta & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -\theta \\ 0 & 0 & 0 & 0 & \dots & -\theta & 1 \\ -\theta & 0 & 0 & 0 & \dots & 0 & -\theta & 1 \end{vmatrix} =$$

$$= D_{T-1} + \theta \begin{vmatrix} -\theta & -\theta & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 - \theta & \dots & 0 & 0 & 0 \\ 0 & -\theta & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -\theta & 0 \\ 0 & 0 & 0 & \dots & -\theta & 1 & -\theta \\ -\theta & 0 & 0 & \dots & 0 & -\theta & 1 \end{vmatrix} +$$

$$+ (-1)^T \theta \begin{vmatrix} -\theta & 1 - \theta & 0 & \dots & 0 & 0 \\ 0 & -\theta & 1 - \theta & \dots & 0 & 0 \\ 0 & 0 & -\theta & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -\theta \\ 0 & 0 & 0 & 0 & \dots & -\theta & 1 \\ -\theta & 0 & 0 & 0 & \dots & 0 & -\theta \end{vmatrix} =$$

$$= D_{T-1} - \theta^2 D_{T-2} + (-1)^{T-1} \theta^2 (-\theta)^{T-2} -$$

$$\begin{aligned}
 & -(-1)^T \theta^2 (-\theta)^{T-2} + (-1)^T \theta^2 (-1)^{T-1} D_{T-2} = \\
 & = D_{T-1} - 2\theta^2 D_{T-2} - 2\theta^T, \quad T = 3, 4, \dots
 \end{aligned}$$

Подстановка сюда выражения (89) приводит к (94). Для $T = 2$ имеем $a_{12} = a_{21} = 1$ и поэтому $\Delta_2 = 1 - 4\theta^2$. ■

ЛЕММА 6.7.11. *Определитель $C_T = |I - 2\theta A_1|$, где матрица A_1 соответствует модели с последовательными разностями, равен*

$$(96) \quad C_T = \frac{1-2\theta}{\sqrt{1-4\theta^2}} \left[\left(\frac{1+\sqrt{1-4\theta^2}}{2} \right)^T - \left(\frac{1-\sqrt{1-4\theta^2}}{2} \right)^T \right], \quad T = 2, 3, \dots$$

ДОКАЗАТЕЛЬСТВО. Разлагая определитель C_T (сначала по элементам его первой строки, а затем по элементам последней строки одного из получаемых определителей и по элементам первого столбца другого), имеем

$$\begin{aligned}
 (97) \quad C_T &= \begin{vmatrix} 1 & -\theta & -\theta & 0 & 0 & \dots & 0 & 0 & 0 \\ -\theta & 1 & -\theta & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & -\theta & 1 & -\theta & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & -\theta & 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -\theta & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & -\theta & 1 & -\theta & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & -\theta & 1 & -\theta \end{vmatrix} = \\
 &= (1-\theta) \begin{vmatrix} 1 & -\theta & 0 & \dots & 0 & 0 & 0 \\ -\theta & 1 & -\theta & \dots & 0 & 0 & 0 \\ 0 & -\theta & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -\theta & 0 \\ 0 & 0 & 0 & \dots & -\theta & 1 & -\theta \\ 0 & 0 & 0 & \dots & 0 & -\theta & 1 & -\theta \end{vmatrix} + \\
 &+ \theta \begin{vmatrix} -\theta & -\theta & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -\theta & \dots & 0 & 0 & 0 \\ 0 & -\theta & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -\theta & 0 \\ 0 & 0 & 0 & \dots & -\theta & 1 & -\theta \\ 0 & 0 & 0 & \dots & 0 & -\theta & 1 & -\theta \end{vmatrix} =
 \end{aligned}$$

$$\begin{aligned}
&= (1-\theta)^2 D_{T-2} + (1-\theta)\theta \begin{vmatrix} 1 & -\theta & 0 & \dots & 0 & 0 \\ -\theta & 1 & -\theta & \dots & 0 & 0 \\ 0 & -\theta & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -\theta & -\theta \end{vmatrix} - \\
&- \theta^2 \begin{vmatrix} 1 & -\theta & \dots & 0 & 0 & 0 \\ -\theta & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -\theta & 0 \\ 0 & 0 & \dots & -\theta & 1 & -\theta \\ 0 & 0 & \dots & 0 & -\theta & 1-\theta \end{vmatrix} = \\
&= (1-\theta)^2 D_{T-2} - 2(1-\theta)\theta^2 D_{T-3} - \\
&- \theta^3 \begin{vmatrix} 1 & -\theta & \dots & 0 & 0 \\ -\theta & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & -\theta & -\theta \end{vmatrix} = \\
&= (1-\theta)^2 D_{T-2} - 2(1-\theta)\theta^2 D_{T-3} + \theta^4 D_{T-4}, \quad T = 5, 6, \dots
\end{aligned}$$

Подстановка сюда выражения (89) дает

$$\begin{aligned}
(98) \quad C_T &= \frac{1-2\theta}{2\sqrt{1-4\theta^2}} \left\{ [(1-3\theta^2) + \right. \\
&+ (1-\theta^2)\sqrt{1-4\theta^2}] \left(\frac{1+\sqrt{1-4\theta^2}}{2} \right)^{T-3} - \\
&\left. - [(1-3\theta^2) - (1-\theta^2)\sqrt{1-4\theta^2}] \left(\frac{1-\sqrt{1-4\theta^2}}{2} \right)^{T-3} \right\}.
\end{aligned}$$

Перегруппировка приводит к (96) для $T = 5, 6, \dots$. В справедливости же (96) для $T = 2, 3, 4$ можно убедиться непосредственным подсчетом.

Если в (96) раскрыть по формуле бинома Ньютона обе входящие в это выражение T -е степени, получим

$$\begin{aligned}
(99) \quad C_T &= \frac{1}{2^T} \frac{1-2\theta}{\sqrt{1-4\theta^2}} \sum_{v=0}^T \binom{T}{v} (1-4\theta^2)^{v/2} [1 - (-1)^v] = \\
&= \frac{1-2\theta}{2^{T-1}} \sum_{\alpha=0}^H \binom{T}{2\alpha+1} (1-4\theta^2)^\alpha,
\end{aligned}$$

где $H = (T - 1)/2$, если T нечетное, и $H = T/2 - 1$, если T четное. ■

ЛЕММА 6.7.12. Если $Q_1^* = (y - \bar{y}e)' A_1 (y - \bar{y}e)$, $A_0 = I$, а e является характеристическим вектором матрицы A_1 , соответствующим характеристическому корню 1, и $\gamma_0 = 1$, $\gamma_1 = 0$, то

$$(100) \quad \xi e^{\theta Q_1^*} = \sqrt{\frac{1 - 2\theta}{1 - 2\theta A_1}}$$

для всех θ , достаточно малых по абсолютной величине.

ДОКАЗАТЕЛЬСТВО. Эта лемма является частным случаем следствия 6.7.2. ■

ТЕОРЕМА 6.7.10. Производящая функция моментов квадратичной формы Q_1^* , основанной на последовательных разностях, при $\gamma_0 = 1$ и $\gamma_1 = 0$ имеет вид

$$(101) \quad \xi e^{\theta Q_1^*} = \frac{\sqrt{1 - 2\theta}}{\left(\frac{\sqrt{1 + 2\theta} + \sqrt{1 - 2\theta}}{2} \right)^T - \left(\frac{\sqrt{1 + 2\theta} - \sqrt{1 - 2\theta}}{2} \right)^T},$$

$T = 2, 3, \dots,$

для всех θ , достаточно малых по абсолютной величине.

ТЕОРЕМА 6.7.11. Производящая функция моментов квадратичной формы Q_1^* , основанной на последовательных разностях, при $\gamma_0 = 1$ и $\gamma_1 = 0$ имеет вид

$$(102) \quad \xi e^{\theta Q_1^*} = \left[\frac{\sqrt{1 - 4\theta^2}}{\left(\frac{1 + \sqrt{1 - 4\theta^2}}{2} \right)^T - \left(\frac{1 - \sqrt{1 - 4\theta^2}}{2} \right)^T} \right]^{1/2},$$

$T = 2, 3, \dots,$

для всех θ , достаточно малых по абсолютной величине.

Следует отметить, что (102) совпадает с выражением $1/\sqrt{D_{T-1}}$, которое является производящей функцией моментов квадратичной формы $Q_1 = \sum_{t=2}^{T-1} y_t y_{t-1}$, определенной в разд. 6.5.4 для $T - 1$ наблюдений.

Моменты циклической квадратичной формы Q_1 можно найти из производящей функции моментов

$$(103) \quad \frac{1}{\sqrt{\Delta_T}} = \left[\left(\frac{1 + \sqrt{1 - 4\theta^2}}{2} \right)^{T/2} - \left(\frac{\sqrt{1 + 2\theta} - \sqrt{1 - 2\theta}}{2} \right)^T \right]^{-1} =$$

$$= \left(\frac{1 + \sqrt{1 - 4\theta^2}}{2} \right)^{-T/2} \left[1 - \left(\frac{\sqrt{1 + 2\theta} - \sqrt{1 - 2\theta}}{\sqrt{1 + 2\theta} + \sqrt{1 - 2\theta}} \right)^T \right]^{-1} =$$

$$= \left(\frac{1 + \sqrt{1 - 4\theta^2}}{2} \right)^{-T/2} \left[1 + \left(\frac{\sqrt{1 + 2\theta} - \sqrt{1 - 2\theta}}{\sqrt{1 + 2\theta} + \sqrt{1 - 2\theta}} \right)^T + \dots \right],$$

$$T = 2, 3, \dots,$$

для θ , достаточно малых по абсолютной величине. Лаплас (1829) показал, что если u является корнем $1 + \sqrt{1 - c}$ уравнения $v^2 - 2v + c = 0$, то

$$(104) \quad u^{-n} = \frac{1}{2^n} + \frac{n}{2^n} \left\{ \frac{c}{4} + \frac{n+3}{2!} \left(\frac{c}{4} \right)^2 + \frac{(n+4)(n+5)}{3!} \left(\frac{c}{4} \right)^3 + \dots \right.$$

$$\left. \dots + \frac{\Gamma(n+2k)}{k! \Gamma(n+k+1)} \left(\frac{c}{4} \right)^k + \dots \right\}.$$

Вывод этого разложения не будет здесь приводиться из-за его сложности. Если положить $c = 4\theta^2$ и $n = T/2$, получим

$$(105) \quad \left(\frac{1 + \sqrt{1 - 4\theta^2}}{2} \right)^{-T/2} = 1 + \frac{1}{2} T \left\{ \theta^2 + \frac{T/2 + 3}{2!} \theta^4 + \right.$$

$$\left. + \frac{(T/2 + 4)(T/2 + 5)}{3!} \theta^6 + \dots + \frac{\Gamma(T/2 + 2k)}{k! \Gamma(T/2 + k + 1)} \theta^{2k} + \dots \right\}.$$

Поскольку, далее,

$$(106) \quad \frac{\sqrt{1 + 2\theta} - \sqrt{1 - 2\theta}}{\sqrt{1 + 2\theta} + \sqrt{1 - 2\theta}} = \frac{\sqrt{1 + 2\theta} - \sqrt{1 - 2\theta}}{\sqrt{1 + 2\theta} + \sqrt{1 - 2\theta}} \cdot \frac{\sqrt{1 + 2\theta} + \sqrt{1 - 2\theta}}{\sqrt{1 + 2\theta} + \sqrt{1 - 2\theta}} =$$

$$= \frac{2\theta}{(1 + \sqrt{1 - 4\theta^2})} = \theta \left(\frac{1 + \sqrt{1 - 4\theta^2}}{2} \right)^{-1},$$

то второй член в квадратных скобках в (103) представляет собой степенной ряд по θ , начинающийся с θ^T как с самой низкой степени. Таким образом, выражение в квадратных скобках представляет собой ряд, состоящий из 1 и степеней θ , не меньших T . Моменты квадратичной формы Q_1 до порядка $T - 1$ получаются как

$$(107) \quad \mathbb{E} Q_1^n = \frac{d^n}{d\theta^n} \frac{1}{\sqrt{\Delta_T}} \Big|_{\theta=0} = \frac{d^n}{d\theta^n} \left(\frac{1 + \sqrt{1 - 4\theta^2}}{2} \right)^{-T/2} \Big|_{\theta=0},$$

$$n = 0, 1, \dots, T - 1$$

(См. упр. 51.) Из (105) и (107) находим, что

$$(108) \quad \mathbb{E} Q_1^{2k-1} = 0, \quad 1 \leq k \leq T/2, \quad T = 2, 3, \dots,$$

$$\mathbb{E} Q_1^{2k} = \frac{T}{2} \frac{\Gamma(T/2 + 2k) (2k)!}{\Gamma(T/2 + k + 1) k!}, \quad 0 \leq k \leq (T - 1)/2, \quad T = 2, 3, \dots$$

Из этого результата, соотношения (87), замечания, сделанного вслед за теоремой 6.7.9, и формулы удвоения для гамма-функции

$$(109) \quad \Gamma(2k+1) = \frac{2^{2k} \Gamma\left(k + \frac{1}{2}\right) \Gamma(k+1)}{\sqrt{\pi}}$$

получаем

$$(110) \quad g_{r1}^{2k-1} = 0, \quad 1 \leq k \leq T/2, \quad T = 2, 3, \dots,$$

$$g_{r1}^{2k} = \frac{\Gamma(T/2+1) \Gamma(k+1/2)}{\Gamma(T/2+k+1) \Gamma(1/2)}, \quad 0 \leq k \leq (T-1)/2, \quad T = 2, 3, \dots$$

Моменты циклического серийного коэффициента корреляции r_1^* , использующего остатки от среднего, равны отношениям соответствующих моментов квадратичных форм Q_1^* и Q_0^* . Моменты квадратичной формы Q_1^* можно найти, используя производящую функцию моментов, приведенную в теореме 6.7.10. При этом моменты до порядка $T-1$ можно получить, используя следующую аппроксимацию последней:

$$(111) \quad \left(\frac{1 + \sqrt{1-4\theta^2}}{2} \right)^{-T/2} \sqrt{1-2\theta}, \quad T = 2, 3, \dots$$

Дело в том, что первые $T-1$ производных функций (111) и (101) совпадают при $\theta = 0$.

ЛЕММА 6.7.13. Пусть

$$(112) \quad \varphi(T) = \left(\frac{1 + \sqrt{1-4\theta^2}}{2} \right)^{-T/2}, \quad T = \dots, -1, 0, 1, \dots,$$

и

$$(113) \quad \varphi^*(T) = \left(\frac{1 + \sqrt{1-4\theta^2}}{2} \right)^{-T/2} \sqrt{1-2\theta}, \quad T = \dots, -1, 0, 1, \dots$$

Тогда для действительных θ , $-1/2 \leq \theta \leq 1/2$,

$$(114) \quad \varphi^*(T) = \varphi(T-1) - \theta \varphi(T+1), \quad T = \dots, -1, 0, 1, \dots$$

ДОКАЗАТЕЛЬСТВО. Правая часть (114) равна

$$\begin{aligned} (115) \quad \varphi(T-1) - \theta \varphi(T+1) &= \\ &= \left(\frac{1 + \sqrt{1-4\theta^2}}{2} \right)^{-(T+1)/2} \left(\frac{1-2\theta + \sqrt{1-4\theta^2}}{2} \right) = \\ &= \left(\frac{1 + \sqrt{1-4\theta^2}}{2} \right)^{-T/2} \sqrt{1-2\theta} \frac{\sqrt{1-2\theta} + \sqrt{1+2\theta}}{\sqrt{2}(1 + \sqrt{1-4\theta^2})^{1/2}} = \\ &= \varphi^*(T) \frac{\sqrt{1-2\theta} + \sqrt{1+2\theta}}{\sqrt{2}(1 + \sqrt{1-4\theta^2})^{1/2}}. \end{aligned}$$

Но это и есть $\varphi^*(T)$, если θ действительное и $-1/2 \leq \theta \leq 1/2$. ■

Из леммы 6.7.13 и того факта, что $\varphi(T-1)$ и $\varphi(T+1)$ разлагаются в ряды по четным степеням θ , вытекают следующие соотношения. Во-первых,

$$(116) \quad \mathcal{E}[Q_1^*(T)]^{2k-1} = \frac{d^{2k-1}\varphi^*(T)}{d\theta^{2k-1}} \Big|_{\theta=0} = - \frac{d^{2k-1}\theta\varphi(T+1)}{d\theta^{2k-1}} \Big|_{\theta=0},$$

$$1 \leq k \leq T/2, \quad T = 2, 3, \dots,$$

что равно $-(2k-1) \mathcal{E}Q_1^{2k-2}(T+1)$. И, во-вторых,

$$(117) \quad \mathcal{E}[Q_1^*(T)]^{2k} = \frac{d^{2k}\varphi^*(T)}{2\theta^{2k}} \Big|_{\theta=0} = \frac{d^{2k}\varphi(T-1)}{d\theta^{2k}} \Big|_{\theta=0},$$

$$0 \leq k \leq (T-1)/2, \quad T = 2, 3, \dots,$$

в свою очередь равное $\mathcal{E}Q_1^{2k}(T-1)$ для $0 \leq k \leq (T-2)/2$; $T = 3, 4, \dots$. Здесь $Q_1^*(T)$ и $Q_1(T)$ обозначают соответственно квадратичные формы Q_1^* и Q_1 , построенные по T наблюдениям. Ограничения относительно значений k и T введены здесь для того, чтобы сохранить правые части в таком виде, в каком они определялись ранее. Таким образом,

$$(118) \quad \mathcal{E}Q_1^{*2k-1} = - \frac{T+1}{2} \cdot \frac{\Gamma(T/2 - 3/2 + 2k)(2k-1)!}{\Gamma(T/2 + 1/2 + k)(k-1)!},$$

$$1 \leq k \leq T/2, \quad T = 2, 3, \dots,$$

$$(119) \quad \mathcal{E}Q_1^{*2k} = \frac{T-1}{2} \cdot \frac{\Gamma(T/2 - 1/2 + 2k)(2k)!}{\Gamma(T/2 + 1/2 + k)k!},$$

$$0 \leq k \leq (T-1)/2, \quad T = 2, 3, \dots$$

Тогда

$$(120) \quad \mathcal{E}r_1^{*2k-1} = - \frac{T+1}{T-1} \cdot \frac{\Gamma(T/2 + 1/2) \Gamma(k + 1/2)}{\Gamma(T/2 + 1/2 + k) \Gamma(1/2)},$$

$$1 \leq k \leq T/2, \quad T = 2, 3, \dots,$$

$$(121) \quad \mathcal{E}r_1^{*2k} = \frac{\Gamma(T/2 + 1/2) \Gamma(k + 1/2)}{\Gamma(T/2 + 1/2 + k) \Gamma(1/2)},$$

$$0 \leq k \leq (T-1)/2, \quad T = 2, 3, \dots$$

Заметим, что $\mathcal{E}r_1^{*2k-1} = - \frac{T+1}{T-1} \mathcal{E}r_1^{*2k}$, $1 \leq k \leq (T-1)/2$, $T = 2, 3, \dots$

Кроме того, $\mathcal{E}r_1^{*2k}$ совпадает с $\mathcal{E}r_1^{2k}$, если r_1 вычислять по $T-1$ наблюдениям.

Найти интересные нас моменты можно и другим путем. Для этого заметим, что $Q_1 = Q_1^* + z_T^2$, причем Q_1^* и z_T независимы.

Тогда имеем

$$\begin{aligned}
 (122) \quad \mathbb{E} Q_1^h &= \mathbb{E} (Q_1^* + z_T^2)^h = \sum_{j=0}^h \binom{h}{j} \mathbb{E} (z_T^2)^j \mathbb{E} Q_1^{*h-j} = \\
 &= \sum_{j=0}^h \binom{h}{j} \frac{\Gamma(j+1/2)}{\Gamma(1/2)} 2^j \mathbb{E} Q_1^{*h-j}.
 \end{aligned}$$

Рассмотрим теперь квадратичную форму Q общего вида, имеющую характеристические корни $\lambda_1, \dots, \lambda_M$. Тогда при $\gamma_0 = 1$ и $\gamma_1 = 0$

$$\begin{aligned}
 (123) \quad \mathbb{E} e^{\theta Q} &= \prod_{t=1}^M (1 - 2\theta \lambda_t)^{-1/2} = \exp \left\{ -\frac{1}{2} \sum_{t=1}^M \log(1 - 2\theta \lambda_t) \right\} = \\
 &= \exp \left\{ \frac{1}{2} \sum_{t=1}^M \left[2\theta \lambda_t + \frac{(2\theta \lambda_t)^2}{2} + \frac{(2\theta \lambda_t)^3}{3} + \dots \right] \right\} = \\
 &= \exp \left\{ \frac{1}{2} \left[2\theta \sum_{t=1}^M \lambda_t + \frac{(2\theta)^2}{2} \sum_{t=1}^M \lambda_t^2 + \frac{(2\theta)^3}{3} \sum_{t=1}^M \lambda_t^3 + \dots \right] \right\}.
 \end{aligned}$$

При этом k -й семиинвариант равен $k!$, умноженному на величину

$$(124) \quad v_k = \frac{2^{k-1}}{k} \sum_{t=1}^M \lambda_t^k,$$

и (123) можно переписать в виде

$$\begin{aligned}
 (125) \quad \mathbb{E} e^{\theta Q} &= \exp(v_1 \theta + v_2 \theta^2 + v_3 \theta^3 + \dots) = \\
 &= 1 + \psi_1 \theta + \psi_2 \theta^2 + \psi_3 \theta^3 + \dots,
 \end{aligned}$$

где

$$\begin{aligned}
 (126) \quad \psi_1 &= v_1, \\
 \psi_2 &= v_2 + \frac{v_1^2}{2}, \\
 \psi_3 &= v_3 + v_1 v_2 + \frac{v_1^3}{6}, \\
 \psi_4 &= v_4 + \frac{v_2^2}{2} + v_1 v_3 + \frac{v_1^2 v_2}{2} + \frac{v_1^4}{24}.
 \end{aligned}$$

Другие ψ_k можно вычислить, воспользовавшись предыдущим тождеством. Из (125) получаем теперь

$$(127) \quad \mathbb{E} Q^h = h! \psi_h.$$

Если $\lambda_t = -\lambda_{M+1-t}$ (как это имеет место для квадратичной формы Q_1^* в случае использования последовательных разностей),

то $v_1 = v_3 = \dots = 0$ и $\psi_1 = \psi_3 = \dots = 0$. Тогда

$$\begin{aligned}
 \psi_2 &= v_2, \\
 \psi_4 &= v_4 + \frac{v_2^2}{2}, \\
 \psi_6 &= v_6 + v_2 v_4 + \frac{v_2^3}{6}, \\
 \psi_8 &= v_8 + \frac{v_4^2}{2} + v_2 v_6 + \frac{v_2^2 v_4}{2} + \frac{v_2^4}{24}.
 \end{aligned}
 \tag{128}$$

Если $M = T - 1$ и $\lambda_t = \cos \pi t/T$, то для четных k

$$\begin{aligned}
 v_k &= \frac{2^{k-1}}{k} \sum_{t=1}^{T-1} \left(\cos \frac{\pi t}{T} \right)^k = \frac{1}{2k} \sum_{t=1}^{T-1} (e^{i\pi t/T} + e^{-i\pi t/T})^k = \\
 &= \frac{1}{2k} \sum_{j=0}^k \binom{k}{j} \left[\sum_{t=0}^{T-1} e^{i2\pi t(j-k/2)/T} - 1 \right].
 \end{aligned}
 \tag{129}$$

Сумма, стоящая в квадратных скобках, равна T , если $j - k/2$ делится на T без остатка (включая значение $j - k/2 = 0$), и равна 0 в противном случае. Поэтому

$$v_k = \frac{T}{2k} \sum_j' \binom{k}{j} - \frac{2^{k-1}}{k},
 \tag{130}$$

где знак $\sum_j'$ указывает, что суммирование ведется по тем значениям j , $j = 0, 1, \dots, k$, для которых $j - k/2$ делится нацело на T . То есть это будут значения вида $j = k/2 \pm hT$, $h = 0, 1, 2, \dots$ ($0 \leq hT \leq k/2$). Поскольку

$$\binom{k}{k/2 + hT} = \binom{k}{k/2 - hT},
 \tag{131}$$

то $v_{2l-1} = 0$ и

$$v_{2l} = \frac{1}{4l} \left\{ T \left[\binom{2l}{l} + 2 \sum_{h=1}^{[l/T]} \binom{2l}{l-hT} \right] - 2^{2l} \right\}.
 \tag{132}$$

Число слагаемых под знаком суммы равно 0, если $l < T$, единице, если $T \leq l < 2T$, двум, если $2T \leq l < 3T$, и т. д. В частности,

$$v_{2l} = \frac{1}{4l} \left\{ \binom{2l}{l} T - 2^{2l} \right\}, \quad l < T,
 \tag{133}$$

так что при этом

$$\begin{aligned}
 (134) \quad v_2 &= \frac{T-2}{2}, & T=2, 3, \dots, \\
 v_4 &= \frac{3T-8}{4}, & T=3, 4, \dots, \\
 v_6 &= \frac{5T-16}{3}, & T=4, 5, \dots, \\
 v_8 &= \frac{35T-128}{8}, & T=5, 6, \dots, \\
 \psi_2 &= \frac{T-2}{2}, & T=2, 3, \dots, \\
 \psi_4 &= \frac{T^2+2T-12}{8}, & T=3, 4, \dots, \\
 (135) \quad \psi_6 &= \frac{T^3+12T^2+8T-168}{48}, & T=4, 5, \dots, \\
 \psi_8 &= \frac{T^4+28T^3+212T^2-64T-3696}{384}, & T=5, 6, \dots
 \end{aligned}$$

Как было отмечено выше, $\psi_k = 0$ для нечетных значений k .

Моменты сериального коэффициента корреляции r_1^* для модели с последовательными разностями равны, таким образом,

$$\begin{aligned}
 \mathcal{E}r_1^{*2} &= \frac{T-2}{(T-1)(T+1)} = \frac{T-2}{T^2-1}, & T=2, 3, \dots, \\
 \mathcal{E}r_1^{*4} &= \frac{3(T^2+2T-12)}{(T-1)(T+1)(T+3)(T+5)}, & T=3, 4, \dots, \\
 \mathcal{E}r_1^{*6} &= \frac{15(T^3+12T^2+8T-168)}{(T-1)(T+1)(T+3)(T+5)(T+7)(T+9)}, & T=4, 5, \dots \\
 \mathcal{E}r_1^{*8} &= \frac{105(T^4+28T^3+212T^2-64T-3696)}{(T-1)(T+1)(T+3)(T+5)(T+7)(T+9)(T+11)(T+13)}, & T=5, 6, \dots
 \end{aligned}$$

Разложение в ряд производящей функции моментов в (123) было дано Нейманом (1941). Оно может быть использовано и для других сериальных коэффициентов корреляции. Например, в случае циклического сериального коэффициента корреляции первого порядка, использующего остатки от выборочного среднего, коэффициенты в показателе экспоненты производящей функции моментов будут такими:

$$(137) \quad v_k = \frac{2^{k-1}}{k} \sum_{t=1}^{T-1} \left(\cos \frac{2\pi}{T} t \right)^k = \frac{1}{2k} \sum_{t=1}^{T-1} (e^{i2\pi t/T} + e^{-i2\pi t/T})^k =$$

$$\begin{aligned}
 &= \frac{1}{2k} \sum_{t=1}^{T-1} \sum_{j=0}^k \binom{k}{j} e^{i2\pi t(2j-k)/T} = \\
 &= \frac{1}{2k} \sum_{j=0}^k \binom{k}{j} \left[\sum_{t=0}^{T-1} e^{i2\pi t(2j-k)/T} - 1 \right].
 \end{aligned}$$

Сумма внутри квадратных скобок равна здесь T , если $2j - k$, $j = 0, 1, \dots, k$, делится на T без остатка (включая значение $2j - k = 0$), и 0 в противном случае. Поэтому

$$(138) \quad v_k = \frac{T}{2k} \sum_j' \binom{k}{j} - \frac{2^{k-1}}{k},$$

где знак $\sum_j'$ указывает, что суммирование ведется по тем из значений j , $j = 0, 1, \dots, k$, для которых $(2j - k)$ делится нацело на T . Эти значения получаются следующим образом. Предположим, что k четное, $k = 2l$. Тогда сумма берется только по тем значениям j , для которых $2(j - l)$ делится нацело на T . Если, далее, T четное, скажем $T = 2H$, то следует брать при этом только те j , для которых $(j - l)$ делится нацело на H . Но последние имеют вид $j = l \pm hH = k/2 \pm hT/2$, $h = 0, 1, 2, \dots$ ($0 \leq hT \leq k$). Если же T нечетное, то тогда надо суммировать по тем j , для которых $(j - l)$ делится нацело на T . Такие значения j имеют вид $j = l \pm hT$, $h = 0, 1, 2, \dots$ ($0 \leq hT \leq l = k/2$). Таким образом,

$$(139) \quad v_{2l} = \begin{cases} \frac{1}{4l} \left\{ T \left[\binom{2l}{l} + 2 \sum_{h=1}^{[2l/T]} \binom{2l}{l-hT/2} \right] - 2^{2l} \right\}, & T = 2, 4, \dots, \\ \frac{1}{4l} \left\{ T \left[\binom{2l}{l} + 2 \sum_{h=1}^{[l/T]} \binom{2l}{l-hT} \right] - 2^{2l} \right\}, & T = 3, 5, \dots \end{cases}$$

Пусть теперь k нечетное, скажем $k = 2l - 1$. Тогда суммирование ведется по тем значениям j , для которых $2j - 2l + 1 = 2(j - l) + 1$ делится без остатка на T . Если при этом T четное, то таких значений j не существует. Если же T нечетное, то $2(j - l) + 1 = \pm hT$ и отсюда h необходимо должно быть нечетным, скажем $h = 2g - 1$ ($g = 1, 2, \dots$). Таким образом, в этом случае $2(j - l) + 1 = \pm(2g - 1)T$ и

$$(140) \quad j = l \pm (gT - T/2) - \frac{1}{2}.$$

Если $T > 2l - 1$, то подходящих значений g не существует, а значит, не существует и подходящих значений j .

Итак,

$$(141) \quad v_{2l-1} = \begin{cases} -\frac{2^{2l-1}}{2(2l-1)}, & T = 2, 4, \dots, \\ \frac{1}{2(2l-1)} \left\{ 2T \sum_{g=1}^{[(l-1/2)/T+1/2]} \binom{2l-1}{l-1/2-(g-1/2)T} - 2^{2l-1} \right\}, & T = 3, 5, \dots \end{cases}$$

В частности,

$$(142) \quad v_{2l-1} = -\frac{2^{2l-1}}{2(2l-1)}, \quad 1 \leq l < (T+1)/2, \quad T = 2, 3, \dots,$$

$$(143) \quad v_{2l} = \frac{T}{2(2l)} \binom{2l}{l} - \frac{2^{2l}}{2(2l)}, \quad 1 \leq l < T, \quad T = 3, 5, \dots, \\ 1 \leq l < T/2, \quad T = 2, 4, \dots,$$

так что

$$v_1 = -1, \quad T = 2, 3, \dots,$$

$$(144) \quad v_2 = \frac{(T-2)}{2}, \quad T = 3, 4, \dots,$$

$$v_3 = -\frac{4}{3}, \quad T = 4, 5, \dots,$$

$$v_4 = \frac{3T-8}{4}, \quad T = 5, 6, \dots,$$

$$\psi_1 = -1, \quad T = 2, 3, \dots,$$

$$(145) \quad \psi_2 = \frac{T-1}{2}, \quad T = 3, 4, \dots,$$

$$\psi_3 = -\frac{T+1}{2}, \quad T = 4, 5, \dots,$$

$$\psi_4 = \frac{(T-1)(T+5)}{8}, \quad T = 5, 6, \dots$$

Первые четыре момента циклического коэффициента r_1^* равны, таким образом,

$$\mathcal{E}r_1^* = -\frac{1}{T-1}, \quad T = 2, 3, \dots,$$

$$(146) \quad \mathcal{E}r_1^{*2} = \frac{1}{T+1}, \quad T = 3, 4, \dots,$$

$$\mathcal{E}r_1^{*3} = -\frac{3}{(T-1)(T+3)}, \quad T = 4, 5, \dots,$$

$$\mathcal{E}r_1^{*4} = \frac{3}{(T+1)(T+3)}, \quad T = 5, 6, \dots$$

Они, разумеется, согласуются со (120) и (121).

С разложением Неймана связано использование для получения первых нескольких моментов не всей производящей функции, а лишь некоторой части ее аналитического выражения. Рассмотрим, например, случай сериального коэффициента корреляции r_1^* , использующего последовательные разности. Мы аппроксимируем $\log \mathcal{E} e^{\theta Q}$ суммой, в которой величины v_k для четных $k < 2T$ заменим величинами v_k^* , формально вычисляемыми по формуле (133) для всех l , а не только для $l < T$, т. е. суммой

$$(147) \quad \sum_{j=1}^{\infty} \theta^{2j} v_{2j}^* = \sum_{j=1}^{\infty} \frac{\theta^{2j}}{4j} \left\{ \binom{2j}{j} T - 2^{2j} \right\} = T \sum_{j=1}^{\infty} \frac{\theta^{2j}}{4j} \frac{(2j)!}{(j!)^2} - \sum_{j=1}^{\infty} \frac{(4\theta^2)^j}{4j} =$$

$$= T \sum_{j=1}^{\infty} \frac{\theta^{2j}}{4j} \frac{\Gamma(2j+1)}{j! \Gamma(j+1)} + \frac{1}{4} \log(1 - 4\theta^2).$$

Первую сумму в правой части последнего соотношения с помощью формулы удвоения для гамма-функции можно представить в виде

$$(148) \quad T \sum_{j=1}^{\infty} \frac{(4\theta^2)^j \Gamma(j + 1/2)}{4j \cdot j! \Gamma(1/2)}.$$

Из соотношения

$$(149) \quad \sum_{h=1}^{\infty} \frac{x^h \Gamma(h + 1/2)}{h! \Gamma(1/2)} = (1 - x)^{-1/2},$$

находим, что

$$(150) \quad \sum_{h=1}^{\infty} \frac{x^h \Gamma(h + 1/2)}{h! \Gamma(1/2)} = \frac{1}{x} [(1 - x)^{-1/2} - 1] = \frac{1 - \sqrt{1 - x}}{x \sqrt{1 - x}}.$$

Это приводит к тождеству

$$(151) \quad \sum_{h=1}^{\infty} \frac{x^h \Gamma(h + 1/2)}{h \cdot h! \Gamma(1/2)} = -2 \log \left(\frac{1 + \sqrt{1 - x}}{2} \right).$$

В справедливости последнего можно убедиться, дифференцируя обе части (151) и замечая, что они совпадают при $x = 0$. Суммируя полученные результаты, получаем

$$(152) \quad \exp \left(\sum_{j=1}^{\infty} \theta^{2j} v_{2j}^* \right) = \exp \left\{ \frac{T}{4} \left[-2 \log \left(\frac{1 + \sqrt{1 - 4\theta^2}}{2} \right) \right] + \right.$$

$$\left. + \frac{1}{4} \log(1 - 4\theta^2) \right\} =$$

$$= \frac{(1 - 4\theta^2)^{1/4}}{\left(\frac{1 + \sqrt{1 - 4\theta^2}}{2} \right)^{T/2}}.$$

Эта аппроксимация для $\mathcal{E} e^{\theta Q}$, по существу, сводится к опусканию члена $(1 - \sqrt{1 - 4\theta^2}/2)^T$ в выражении (102).

6.8. АППРОКСИМАЦИЯ РАСПРЕДЕЛЕНИЙ СЕРИАЛЬНЫХ КОЭФФИЦИЕНТОВ КОРРЕЛЯЦИИ

6.8.1. Аппроксимация распределений циклических серийных коэффициентов корреляции

Хотя распределение циклического серийного коэффициента корреляции, использующего остатки от выборочного среднего, является известным и довольно простым в случае независимых нормальных наблюдений и хотя оно частично затабулировано (табл. 6.1), тем не менее удобно использовать для определения процентных точек и оценки вариабельности r_1^* аппроксимирующие распределения. Мы рассмотрим также аппроксимации для распределения циклического серийного коэффициента корреляции r_1 , вычисляемого по наблюдениям, измеренным относительно их известного среднего значения. Эти аппроксимации достаточно просты из-за симметричности аппроксимирующих распределений.

Большинство аппроксимаций основано на подборе соответствующего бета-распределения с плотностью

$$(1) \quad f(x) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \frac{(x-a)^{p-1}(b-x)^{q-1}}{(b-a)^{p+q-1}}$$

для $a \leq x \leq b$, и $f(x) = 0$ в противном случае. Момент порядка h этой плотности есть

$$\begin{aligned} (2) \quad \int_a^b x^h f(x) dx &= \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \int_a^b \frac{(x-a)^{p-1}(b-x)^{q-1} x^h}{(b-a)^{p+q-1}} dx = \\ &= \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \int_0^1 z^{p-1} (1-z)^{q-1} [a + (b-a)z]^h dz = \\ &= \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \sum_{j=0}^h \binom{h}{j} (b-a)^j a^{h-j} \int_0^1 z^{p+j-1} (1-z)^{q-1} dz = \\ &= \frac{\Gamma(p+q)}{\Gamma(p)} \sum_{j=0}^h \binom{h}{j} (b-a)^j a^{h-j} \frac{\Gamma(p+j)}{\Gamma(p+q+j)}. \end{aligned}$$

В частности,

$$(3) \quad \mathbb{E}X = a + (b-a) \frac{p}{p+q},$$

$$(4) \quad \mathbb{E}X^2 = a^2 + 2a(b-a) \frac{p}{p+q} + (b-a)^2 \frac{p(p+1)}{(p+q)(p+q+1)},$$

$$(5) \quad \mathbb{E}X^3 = a^3 + 3a^2(b-a) \frac{p}{p+q} + 3a(b-a)^2 \frac{p(p+1)}{(p+q)(p+q+1)} + \\ + (b-a)^3 \frac{p(p+1)(p+2)}{(p+q)(p+q+1)(p+q+2)},$$

$$(6) \quad \mathbb{E}X^4 = a^4 + 4a^3(b-a) \frac{p}{p+q} + 6a^2(b-a)^2 \frac{p(p+1)}{(p+q)(p+q+1)} + \\ + 4a(b-a)^3 \frac{p(p+1)(p+2)}{(p+q)(p+q+1)(p+q+2)} + \\ + (b-a)^4 \frac{p(p+1)(p+2)(p+3)}{(p+q)(p+q+1)(p+q+2)(p+q+3)}.$$

Если $a = -b$ и $q = p$, то указанная плотность симметрична и имеет вид

$$(7) \quad f(x) = \frac{\Gamma(2p)}{\Gamma^2(p)} \cdot \frac{(b+x)^{p-1}(b-x)^{p-1}}{(2b)^{2p-1}} = \\ = \frac{\Gamma(2p)}{\Gamma^2(p)} \cdot \frac{(b^2 - x^2)^{p-1}}{(2b)^{2p-1}}, \quad -b \leq x \leq b.$$

Моменты нечетного порядка равны нулю, а моменты четных порядков

$$(8) \quad \int_{-b}^b x^{2k} f(x) dx = \frac{\Gamma(2p)}{\Gamma^2(p) (2b)^{2p-1}} \int_{-b}^b (b^2 - x^2)^{p-1} x^{2k} dx = \\ = \frac{\Gamma(2p) b^{2k}}{\Gamma^2(p) 2^{2p-1}} \int_0^1 (1-y)^{p-1} y^{k-1/2} dy = \\ = b^{2k} \frac{\Gamma(2p) \Gamma(k+1/2)}{2^{2p-1} \Gamma(p) \Gamma(p+k+1/2)}, \quad k=0, 1, \dots$$

Замечая, что при $k=0$ интеграл в левой части (8) равен 1, получаем отсюда формулу удвоения для гамма-функции

$$(9) \quad \Gamma(2p) = \frac{2^{2p-1} \Gamma(p) \Gamma(p+1/2)}{\Gamma(1/2)}.$$

Используя этот факт, выражения (7) и (8) для плотности и четных моментов можно записать соответственно в виде

$$(10) \quad f(x) = \frac{\Gamma(p+1/2)}{\Gamma(1/2) \Gamma(p)} \frac{(b^2 - x^2)^{p-1}}{b^{2p-1}}, \quad -b \leq x \leq b,$$

$$(11) \quad \mathbb{E}X^{2k} = b^{2k} \frac{\Gamma\left(p + \frac{1}{2}\right) \Gamma\left(k + \frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(p + k + \frac{1}{2}\right)}.$$

В частности, $\mathcal{E}X = \mathcal{E}X^3 = 0$,

$$(12) \quad \mathcal{E}X^2 = \frac{b^2}{2\rho + 1},$$

$$(13) \quad \mathcal{E}X^4 = \frac{3b^4}{(2\rho + 1)(2\rho + 3)}.$$

Подобного рода плотность имеет обычный пирсоновский коэффициент корреляции r , построенный по остаткам от выборочных средних N пар независимых наблюдений, выбираемых из двумерной нормальной совокупности с нулевой корреляцией. Плотность его распределения равна

$$(14) \quad \frac{\Gamma[(N-1)/2]}{\Gamma(1/2)\Gamma[(N-2)/2]} (1-r^2)^{(N-4)/2}, \quad -1 \leq r \leq 1,$$

а момент порядка $2k$ есть

$$(15) \quad \mathcal{E}r^{2k} = \frac{\Gamma\left[\frac{1}{2}(N-1)\right]\Gamma\left(k+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left[\frac{1}{2}(N-1)+k\right]}.$$

Заметим теперь, что моменты нечетного порядка у r_1 равны нулю. Само распределение для r_1 симметрично, если T четное. Четные моменты низших порядков коэффициента r_1 имеют вид

$$(16) \quad \mathcal{E}r_1^{2k} = \frac{\Gamma\left(\frac{1}{2}T+1\right)\Gamma\left(k+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}T+k+1\right)\Gamma\left(\frac{1}{2}\right)},$$

$$0 \leq 2k \leq T-1, \quad T=2, 3, \dots$$

Ввиду этого плотность r_1 можно аппроксимировать плотностью (14) пирсоновского коэффициента корреляции с $N = T + 3$. При этом точное и аппроксимирующее распределения будут иметь совпадающие моменты вплоть до моментов порядка $T-1$.

Один из путей получения этих моментов и распределений состоит в аппроксимации производящей функции моментов квадратичной формы Q_1 :

$$(17) \quad \mathcal{E}e^{\theta Q_1} = \prod_{t=1}^T \left(1 - 2\theta \cos \frac{2\pi t}{T}\right)^{-1/2} =$$

$$= \exp \left[-\frac{1}{2} \sum_{t=1}^T \log \left(1 - 2\theta \cos \frac{2\pi t}{T}\right) \right].$$

Показатель экспоненты в правой части (17), равный

$$(18) \quad -\frac{T}{4\pi} \sum_{t=1}^T \log \left(1 - 2\theta \cos \frac{2\pi t}{T}\right) \frac{2\pi}{T},$$

является интегральной суммой для интеграла

$$(19) \quad -\frac{T}{4\pi} \int_0^{2\pi} \log(1 - 2\theta \cos u) du = -\frac{T}{2} \log\left(\frac{1 + \sqrt{1 - 4\theta^2}}{2}\right).$$

(По поводу последнего равенства см. упр. 69.) Это приводит к аппроксимации производящей функции моментов выражением

$$(20) \quad \exp\left\{-\frac{T}{2} \log[(1 + \sqrt{1 - 4\theta^2})/2]\right\} = \left(\frac{1 + \sqrt{1 - 4\theta^2}}{2}\right)^{-T/2},$$

уже встречавшимся в разд. 6.7.7. Купменс (1942), используя эту идею, нашел выражение для аппроксимирующей плотности, а Рубин (1945) показал, что оно совпадает с (14) при $N = T + 3$.

Перейдем теперь к рассмотрению r_1^* — циклического сериального коэффициента корреляции, использующего остатки от среднего. Поскольку четные моменты r_1^* совпадают с соответствующими моментами r_1 , если брать для построения r_1 на единицу меньшее число наблюдений, то здесь для аппроксимации можно было бы попытаться использовать распределение пирсоновского коэффициента корреляции r с $N = T + 2$. При этом четные моменты точного и аппроксимирующего распределений будут совпадать до порядка $T - 1$, однако моменты нечетных порядков не будут совпадать.

Другой путь использования бета-распределений состоит в том, что полагают $b = -a = 1$ и подбирают p и q в (1) таким образом, чтобы выполнялись в (3) $\mathcal{E}r_1^* = -1/(T - 1)$ и в (4) $\mathcal{E}r_1^{*2} = 1/(T + 1)$, $T = 3, 4, \dots$ ($r_1^* \equiv -1$ для $T = 2$.) Тогда для p и q получаются следующие выражения:

$$(21) \quad p = \frac{(T-1)(T-2)}{2(T-3)}, \quad q = \frac{T(T-1)}{2(T-3)}.$$

Диксон (1944) сравнил такую аппроксимацию с точным распределением. (См. табл. 6.2.)

Еще один способ использования бета-распределения связан с учетом множества значений циклического сериального коэффициента корреляции. Приравнивая правые части (3) и (4) соответственно $\mathcal{E}r_1^*$ и $\mathcal{E}r_1^{*2}$, получаем

$$(22) \quad p + q = \frac{T+1}{T^2-3T} [-(T-1)a - 1][(T-1)b + 1] - 1,$$

$$(23) \quad p = \frac{-(T-1)a - 1}{(T-1)(b-a)} (p + q),$$

$$(24) \quad q = \frac{(T-1)b + 1}{(T-1)(b-a)} (p + q).$$

Таблица 6.2

ТОЧНОЕ И ПРИБЛИЖЕННОЕ РАСПРЕДЕЛЕНИЯ ЦИКЛИЧЕСКОГО СЕРИАЛЬНОГО КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ, СТРОЯЩЕГОСЯ ПО ОСТАТКАМ ОТ ВЫБОРОЧНОГО СРЕДНЕГО *)

T	0.01				0.05			
	точное	бета $p \neq q$	r	нормальное	точное	бета $p \neq q$	r	нормальное
5	-0.798	-0.858	-0.989	-1.000	-0.753	-0.742	-0.821	-0.781
10	-0.705	-0.702	-0.734	-0.763	-0.564	-0.562	-0.576	-0.572
15	-0.597	-0.596	-0.609	-0.629	-0.462	-0.461	-0.467	-0.466
20	-0.524	-0.524	-0.532	-0.545	-0.399	-0.399	-0.402	-0.401
25	-0.473	-0.473	-0.477	-0.487	-0.356	-0.356	-0.357	-0.357
30	-0.433	-0.433	-0.436	-0.444	-0.324	-0.324	-0.324	-0.324
45	-0.35	-0.356	-0.357	-0.362	-0.262	-0.262	-0.262	-0.262
75	-0.276	-0.276	-0.277	-0.278	-0.201	-0.201	-0.201	-0.201

T	0.95				0.99			
	точное	бета $p \neq q$	r	нормальное	точное	бета $p \neq q$	r	нормальное
5	0.253	0.317	0.421	0.281	0.297	0.527	0.589	0.501
10	0.360	0.362	0.376	0.350	0.525	0.533	0.534	0.541
15	0.328	0.329	0.333	0.323	0.475	0.477	0.476	0.486
20	0.299	0.299	0.302	0.296	0.432	0.433	0.432	0.440
25	0.276	0.276	0.277	0.274	0.398	0.398	0.397	0.404
30	0.257	0.257	0.258	0.255	0.370	0.371	0.370	0.375
45	0.218	0.218	0.218	0.217	0.313	0.313	0.313	0.316
75	0.174	0.174	0.175	0.174	0.250	0.250	0.250	0.251

*) Значения, соответствующие «бета $p \neq q$ », получены Диксоном (1944) для бета-распределения, подобранного по формуле (21). Значения, соответствующие «r», подсчитаны с использованием $r_1^* + 1/T$ как пирсоновского коэффициента корреляции с $N = T + 3$; см. (28).

Если T четное, то r_1^* изменяется в пределах $-1 \leq r_1^* \leq \cos 2\pi/T$, так, что можно взять $a = -1$ и $b = \cos 2\pi/T$. Если же T нечетное, то $\cos \pi (T-1)/T \leq r_1^* \leq \cos 2\pi/T$ и можно взять $a = -\cos \pi/T$, $b = \cos 2\pi/T$.

Хеннан (1960) предложил использовать для аппроксимации тот факт, что величина $r_1^* + 1/T$ распределена приближенно как пирсоновский коэффициент корреляции r , построенный по $T + 3$ наблю-

дениям. Действительно, если взять

$$(25) \quad a = -1 - \frac{1}{T-1} = -\frac{T}{T-1},$$

$$(26) \quad b = 1 - \frac{1}{T-1} = \frac{T-2}{T-1},$$

то

$$(27) \quad p = q = \frac{1}{2} \left(T + 1 + \frac{5T+1}{T^2-3T} \right).$$

Округляя все эти значения до значений

$$(28) \quad a = -1 - \frac{1}{T}, \quad b = 1 - \frac{1}{T}, \quad p = q = \frac{T+1}{2},$$

получаем, что вместо исследования r_1^* , построенного по T наблюдениям, можно исследовать r , построенный по $T+3$ наблюдениям.

Предельным распределением для

$$(29) \quad \left(r_1^* + \frac{1}{T-1} \right) \frac{(T-1)\sqrt{T+1}}{\sqrt{T^2-3T}}$$

является $N(0, 1)$. (Асимптотически r_1^* эквивалентен $-\hat{\beta}_1$ из гл. 5.)

Если использовать z -преобразование Фишера

$$(30) \quad z = \frac{1}{2} \log \frac{1 + (r_1^* + 1/T)}{1 - (r_1^* + 1/T)},$$

то величина $\sqrt{T} z$ асимптотически распределена по закону $N(0, 1)$. (Предложенная аппроксимация могла бы, по-видимому, быть обоснована заменой величины $r_1^* + 1/T$ величиной $f(T) [r_1^* + 1/T]$, где $f(T)$ — достаточно просто вычисляемая функция, незначительно превышающая единицу.)

Численное сравнение некоторых указанных аппроксимаций приведено в табл. 6.2. Следует отметить, что бета-аппроксимация оказывается здесь весьма точной для значений T , начиная с 15 или 20. Использование r дает значительно меньшую точность аппроксимации; для большинства целей ее можно считать хорошей лишь при $T \geq 30$.

6.8.2. Аппроксимация распределений сериальных корреляций, основанных на последовательных разностях

Точное распределение коэффициента r_1^* , основанного на последовательных разностях, было выписано в явном виде Хартом и фон Нейманом (1942) для $T = 3, 4$ и 5. Для значений $T = 6$ и 7 были численными методами найдены необходимые интегралы и получены

соответствующие таблицы. Для $T = 8, 9, \dots$ плотность распределения r_1^* аппроксимировалась несколькими членами ряда

$$(31) \quad \sum_{j=0}^{\infty} a_j \left(\cos^2 \frac{\pi}{T} - r_1^{*2} \right)^{T/2-2+j}.$$

Последний был выбран по той причине, что коэффициент r_1^* имеет симметричное распределение, принимает значения в интервале $(-\cos \pi/T, \cos \pi/T)$ и его плотность в точках $\pm \cos \pi/T$ имеет порядок убывания $T/2 - 2$. Из этого ряда были взяты первые четыре члена ($j = 0, 1, 2, 3$). При этом коэффициенты выбирались таким образом, чтобы с точными моментами r_1^* совпали первые три четных момента аппроксимирующего распределения. Для $T = 7$ получается хорошее согласие с точным распределением. На множестве значений функции распределения от 0.00678 до 0.07020 ошибка не превосходит 1.7%. Восьмые моменты точного и аппроксимирующего распределений равны соответственно 0.00413 и 0.00412, а десятые моменты — 0.00202 и 0.00201. Процентные точки приведены в табл. 6.3.

Рассмотрим аппроксимацию плотности r_1^* симметричным бета-распределением с $b = -a = \cos \pi/T$. При этом $\mathcal{E}r_1^{*2} = \mathcal{E}X^2$ для

$$(32) \quad p = q = \frac{1}{2} \left[\frac{T^2 - 1}{T - 2} \cos^2 \frac{\pi}{T} - 1 \right].$$

Если приблизить $\cos \pi/T$ величиной $1 - (\pi/T)^2/2$, $\cos^2 \pi/T$ — величиной $1 - (\pi/T)^2$ и последнюю — величиной $1 - 9/T^2$, то получим

$$(33) \quad p = q = \frac{1}{2} \left[\frac{T^2 - 1}{T - 2} \cdot \frac{T^2 - 9}{T^2} - 1 \right] = \frac{T + 1}{2} - \frac{3(2T^2 - 3)}{2(T^3 - 2T^2)}.$$

Янг (1941) рассмотрел подбор симметричного бета-распределения по второму и четвертому моментам r_1^* . В этом случае

$$(34) \quad p = q = \frac{T + 1}{2} - \frac{18}{T^3 - 13T + 24},$$

$$(35) \quad a^2 = b^2 = 1 - \frac{3T}{T^3 - 13T + 24},$$

$$(36) \quad b \sim 1 - \frac{3T}{2(T^3 - 13T + 24)}.$$

Следует отметить, что b соответствует здесь $\cos \pi/T \sim 1 - 9/(2T^2)$. Таблицы, насчитанные Янгом (с использованием для определения b выражения (35)) для 5- и 10-процентных точек, согласуются в пределах трех знаков после запятой со значениями, приведенными в табл. 6.3 (за исключением одного) для $T \geq 8$. Следующее правило

Таблица 6.3

ВЕРХНИЕ 100%-ПРОЦЕНТНЫЕ ТОЧКИ РАСПРЕДЕЛЕНИЯ СЕРИАЛЬНОГО
КОЭФФИЦИЕНТА КОРРЕЛЛЯЦИИ, ОСНОВАННОГО НА ПОСЛЕДОВАТЕЛЬНЫХ
РАЗНОСТЯХ *)

T	Уровни значимости α				T			
	0.05	0.01	0.001			0.05	0.01	0.001
4	0.610	0.687	0.705	32	0.282	0.391	0.504	
5	0.590	0.731	0.792	33	0.278	0.386	0.497	
				34	0.274	0.381	0.491	
6	0.555	0.719	0.818	35	0.271	0.376	0.485	
7	0.532	0.693	0.815					
8	0.509	0.669	0.798	36	0.267	0.371	0.479	
9	0.488	0.646	0.779	37	0.264	0.366	0.474	
10	0.469	0.624	0.759	38	0.260	0.362	0.468	
				39	0.257	0.358	0.463	
11	0.452	0.604	0.740	40	0.254	0.353	0.458	
12	0.436	0.586	0.722					
13	0.422	0.569	0.705	41	0.251	0.349	0.452	
14	0.409	0.553	0.689	42	0.248	0.345	0.448	
15	0.397	0.539	0.673	43	0.245	0.341	0.443	
				44	0.242	0.338	0.438	
16	0.386	0.525	0.659	45	0.240	0.334	0.434	
17	0.376	0.513	0.645					
18	0.367	0.501	0.632	46	0.237	0.331	0.430	
19	0.358	0.490	0.619	47	0.235	0.327	0.426	
20	0.350	0.480	0.607	48	0.232	0.324	0.422	
				49	0.230	0.321	0.418	
21	0.343	0.470	0.596	50	0.228	0.319	0.415	
22	0.335	0.461	0.586					
23	0.329	0.452	0.576	51	0.226	0.316	0.411	
24	0.322	0.444	0.567	52	0.224	0.313	0.408	
25	0.316	0.436	0.558	53	0.222	0.310	0.404	
				54	0.220	0.308	0.401	
26	0.311	0.429	0.549	55	0.218	0.305	0.398	
27	0.305	0.422	0.541					
28	0.300	0.415	0.533	56	0.216	0.303	0.395	
29	0.295	0.409	0.525	57	0.215	0.300	0.392	
30	0.291	0.402	0.518	58	0.213	0.298	0.389	
				59	0.211	0.295	0.386	
31	0.287	0.397	0.511	60	0.209	0.293	0.383	

*) Модификация Т. Андерсона (1948) таблицы Харта (1942), построенной по точному распределению для $T = 4, 5, 6, 7$ и приближенному распределению для $T > 8$.

аппроксимации вытекает из того, что вместо $r_1^*/[1 - 9/(2T^2)]$ можно рассматривать пирсоновский коэффициент корреляции r , построенный по $N = T + 3$ наблюдениям. Тогда, чтобы получить 100%-процентную точку для r_1^* , достаточно умножить 100%-процентную точку для r на $[1 - 9/(2T^2)]$. При $T \geq 15$ значения, получаемые с помощью этого правила, отличаются от значений, указанных в табл. 6.3 для уровней 0.05 и 0.1, лишь в третьем десятичном знаке. Замена $[1 - 9/(2T^2)]$ на $\cos \pi/T$ приводит к еще меньшему различию.

Еще одна аппроксимация состоит в том, что полагают $b = -a = 1$ и

$$(37) \quad p = q = \frac{T+1}{2} + \frac{3}{2(T-2)}.$$

Отсюда вытекает, что вместо r_1^* можно рассматривать пирсоновский коэффициент корреляции r , соответствующий $N = T + 3$ наблюдениям. О степени приближения можно в какой-то мере судить по табл. 6.4.

Таблица 6.4

ПРИБЛИЖЕНИЯ 100%-ПРОЦЕНТНЫХ ТОЧЕК РАСПРЕДЕЛЕНИЙ
СЕРИАЛЬНОГО КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ, ОСНОВАННОГО
НА ПОСЛЕДОВАТЕЛЬНЫХ РАЗНОСТЯХ

T	Уровни значимости α					
	Харт	0.05 r	br	Харт	0.01 r	br
10	0.469	0.476	0.455	0.624	0.634	0.605
15	0.397	0.400	0.392	0.539	0.542	0.532
20	0.350	0.352	0.348	0.480	0.482	0.477
25	0.316	0.317	0.315	0.436	0.437	0.434
30	0.291	0.291	0.290	0.402	0.403	0.401
45	0.240	0.240	0.239	0.334	0.334	0.334

Указанный серийный коэффициент корреляции также распределен асимптотически нормально. Следует отметить, что об использовании r здесь можно сказать то же, что и об использовании $r - 1/T$ в циклическом случае.

Колонки табл. 6.4, обозначенные «Харт», взяты из табл. 6.3. Значения в колонке, помеченной буквой r , подсчитаны заменой r_1^* пирсоновским коэффициентом корреляции с $N = T + 3$. Числа в колонке, обозначенной br , получены умножением соответствующих им чисел в соседней колонке (неокругленных) на $b = 1 - 9/(2T^2)$.

В качестве примера рассмотрим данные табл. 6.5, представляющие собой отношения общего числа голосов, поданных за кандида-

тов от республиканской партии, к общему числу голосов, поданных за кандидатов от демократической партии, на выборах в палату представителей в 1920—1954 гг.

Таблица 6,5

ОТНОШЕНИЕ ОБЩЕГО ЧИСЛА ГОЛОСОВ, ПОДАННЫХ ЗА РЕСПУБЛИКАНЦЕВ, К ОБЩЕМУ ЧИСЛУ ГОЛОСОВ, ПОДАННЫХ ЗА ДЕМОКРАТОВ, НА ВЫБОРАХ В ПАЛАТУ ПРЕДСТАВИТЕЛЕЙ

В 1920—1954 гг. *

Год	Отношение	Год	Отношение
1920	1.65	1938	0.97
1922	1.16	1940	0.89
1924	1.38	1942	1.10
1926	1.41	1944	0.93
1928	1.33	1946	1.21
1930	1.18	1948	0.88
1932	0.76	1950	1.00
1934	0.78	1952	1.00
1936	0.71	1954	0.90

* Statistical Abstract of the United States [Бюро переписей США (1955, табл. 390, стр. 330)].

Соответствующие статистики равны здесь

$$(38) \quad \bar{y} = 1.06 (8),$$

$$(39) \quad Q_0^* = 1.1029 (7),$$

$$(40) \quad Q_1^* = 0.67362 (7),$$

$$(41) \quad r_1^* = 0.6107.$$

Наблюдаемое значение r_1^* значительно отличается от 0, с уровнем значимости 0.01.

6.9. СОВМЕСТНЫЕ И УСЛОВНЫЕ РАСПРЕДЕЛЕНИЯ СЕРИАЛЬНЫХ КОЭФФИЦИЕНТОВ КОРРЕЛЯЦИИ

6.9.1. Совместные распределения

Совместное распределение p сериальных коэффициентов корреляции $r_1^*, \dots, r_p^*$ можно найти, в принципе, из характеристической функции. Как и в случае отдельного коэффициента, распределение

может быть найдено точно, если характеристические корни матрицы образуют пары, за исключением, возможно, одного корня.

Пусть

$$(1) \quad r_j^* = \frac{(y - \bar{y}e)' A_j (y - \bar{y}e)}{(y - \bar{y}e)' (y - \bar{y}e)}, \quad j = 1, \dots, p,$$

$$(2) \quad \bar{y} = \frac{y'e}{e'e},$$

где $e = (1, 1, \dots, 1)'$. Пусть характеристические корни A_j равны $\lambda_{j1}, \dots, \lambda_{jT}$. Допустим, что матрицы $A_1, \dots, A_p$ имеют одну и ту же совокупность характеристических векторов и что e является характеристическим вектором матрицы A_j , соответствующим корню $\lambda_{jT} = 1$, $j = 1, \dots, p$. Если вектор y распределен по закону $N(\mu e, \sigma^2 I)$, то $r_1^*, \dots, r_p^*$ распределены как отношения

$$(3) \quad r_j^* = \frac{\sum_{i=1}^{T-1} \lambda_{ji} z_i^2}{\sum_{i=1}^{T-1} z_i^2}, \quad j = 1, \dots, p,$$

где $z_1, \dots, z_{T-1}$ — независимые случайные величины, каждая из которых распределена по закону $N(0, 1)$. Пусть T — нечетное, скажем $T = 2H + 1$, и корни образуют пары. Положим

$$(4) \quad \lambda_{j,2h-1} = \lambda_{j,2h} = v_{jh}, \quad h = 1, \dots, H,$$

и

$$(5) \quad z_{2h-1}^2 + z_{2h}^2 = x_h, \quad h = 1, \dots, H.$$

Тогда $x_1, \dots, x_H$ — независимые случайные величины, каждая из которых имеет распределение χ^2 с двумя степенями свободы. Если зафиксировать условие $\sum_{h=1}^H x_h = 1$, то величины $x_1, \dots, x_H$ будут равномерно распределены на части этой гиперплоскости, расположенной в положительном ортанте:

$$(6) \quad \sum_{h=1}^H x_h = 1, \quad x_h \geq 0, \quad h = 1, \dots, H.$$

При этом распределение r_j^* совпадает с распределением

$$(7) \quad r_j^* = \sum_{h=1}^H v_{jh} x_h, \quad j = 1, \dots, p,$$

и

$$(8) \quad \Pr \{r_1^* \geq R_1^*, \dots, r_p^* \geq R_p^*\}$$

является отношением $(H - 1)$ -мерного объема пересечения (6) и

$$(9) \quad \sum_{h=1}^H v_{jh} x_h \geq R_j^*, \quad j = 1, \dots, p,$$

к $(H - 1)$ -мерному объему (6). Эта геометрическая задача подобна той, которая решалась при отыскании $\Pr \{r_1^* \geq R_1^*\}$. Множество возможных значений $(r_1^*, \dots, r_p^*)$ точно найти трудно. Оно включает точки $(v_{1h}, \dots, v_{ph})$, соответствующие $x_h = 1, x_j = 0, j \neq h$, для $h = 1, \dots, H$, и, кроме того, все линейные комбинации последних с неотрицательными весами. Таким образом, указанное множество во всяком случае должно быть выпуклым телом, содержащим перечисленные точки и $(0, \dots, 0)$.

Распределение (8) было дано Кенуем (1949) и Ватсоном (1956).

6.9.2. Условные распределения

Как было указано в § 6.3, для проверки гипотезы $\gamma_i = 0$, когда $\gamma_{i+1} = \dots = \gamma_p = 0$ и μ известно, используется распределение квадратичной формы Q_i при заданных значениях квадратичных форм $Q_0, \dots, Q_{i-1}$. Если верна нулевая гипотеза, то это равносильно использованию распределения коэффициента r_i при заданных значениях $r_1, \dots, r_{i-1}$. Отсюда вытекает, что для проверки гипотезы $\gamma_i = 0$ следует использовать условное распределение r_i при заданных значениях $r_1, \dots, r_{i-1}$ и $\gamma_i = 0$. Если используются остатки от среднего, то соответственно следует брать условное распределение r_i^* при заданных $r_1^*, \dots, r_{i-1}^*$.

Если T нечетное и все корни (отличные от 1) двойные, то $r_1^*, \dots, r_i^*$ можно представить в виде (7). Соответствующая условная вероятность

$$(10) \quad \Pr \{r_i^* \geq R_i^* \mid r_1^* = R_1^*, \dots, r_{i-1}^* = R_{i-1}^*\}$$

представляет собой отношение $(H - i)$ -мерного объема пересечения (6) с

$$(11) \quad \sum_{h=1}^H v_{ih} x_h \geq R_i^*,$$

$$(12) \quad \sum_{h=1}^H v_{jh} x_h = R_j^*, \quad j = 1, \dots, i-1,$$

к $(H - i)$ -мерному объему пересечения (6) и (12). И опять это наглядная, но сложная геометрическая задача.

Искомые распределения можно аппроксимировать. Пусть r_i^* представляет собой i -й частный циклический коэффициент корреляции, т. е. частный коэффициент корреляции между y_i и y_{i+i} при

фиксированных значениях $y_{t+1}, \dots, y_{t+i-1}$. Этот r'_i получается из $r_1^*, \dots, r_i^*$ как частная корреляция. [См., например, Т. Андерсон (1958, разд. 4.3.1).] Точнее говоря, пусть

$$(13) \quad \begin{bmatrix} 1 & r_1^* & r_2^* & \dots & r_{i-1}^* & r_i^* \\ r_1^* & 1 & r_1^* & \dots & r_{i-2}^* & r_{i-1}^* \\ r_2^* & r_1^* & 1 & \dots & r_{i-3}^* & r_{i-2}^* \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ r_{i-1}^* & r_{i-2}^* & r_{i-3}^* & \dots & 1 & r_1^* \\ r_i^* & r_{i-1}^* & r_{i-2}^* & \dots & r_1^* & 1 \end{bmatrix} = \begin{bmatrix} 1 & \tilde{r}^{*'} & \tilde{r}^{*'} \\ \mathbf{r}^* & \mathbf{R}^* & \tilde{\mathbf{r}}^* \\ \tilde{r}_i^* & \tilde{\mathbf{r}}^{*'} & 1 \end{bmatrix};$$

тогда

$$(14) \quad r'_i = \frac{r_i^* - \tilde{\mathbf{r}}^{*'} (\mathbf{R}^*)^{-1} \mathbf{r}^*}{(1 - \tilde{\mathbf{r}}^{*'} (\mathbf{R}^*)^{-1} \tilde{\mathbf{r}}^*)^{1/2} (1 - \mathbf{r}^{*'} (\mathbf{R}^*)^{-1} \mathbf{r}^*)^{1/2}} = \frac{r_i^* - \tilde{\mathbf{r}}^{*'} (\mathbf{R}^*)^{-1} \mathbf{r}^*}{1 - \tilde{\mathbf{r}}^{*'} (\mathbf{R}^*)^{-1} \tilde{\mathbf{r}}^*}.$$

В частности,

$$(15) \quad r'_1 = r_1^*,$$

$$(16) \quad r'_2 = \frac{r_2^* - r_1^{*2}}{1 - r_1^{*2}}.$$

Даниэлс (1956) (и Дженкинс (1956) для r'_2, r'_3 и r'_4) показал, что совместное распределение $r_1^*, \dots, r_i^*$ можно аппроксимировать некоторым совместным распределением, у которого частные сериальные корреляции независимы и имеют плотности

$$(17) \quad \left\{ \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left[\frac{1}{2}(T-1)\right]}{\Gamma\left(\frac{1}{2}T\right)} + \frac{\Gamma\left(\frac{3}{2}\right) \Gamma\left[\frac{1}{2}(T-1)\right]}{\Gamma\left(\frac{1}{2}T+1\right)} \right\}^{-1} \times \\ \times (1 - r'_{2j})^2 (1 - r_{2j}^{'2})^{(T-3)/2} = \\ = \frac{\Gamma(T+1)}{2^T \Gamma\left[\frac{1}{2}(T-1)\right] \Gamma\left[\frac{1}{2}(T+3)\right]} (1 + r'_{2j})^{(T-3)/2} (1 - r'_{2j})^{(T+1)/2},$$

для четных индексов и

$$(18) \quad \frac{\Gamma\left(\frac{1}{2}T\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left[\frac{1}{2}(T-1)\right]} (1 - r'_{2j-1}) (1 - r_{2j-1}^{'2})^{(T-3)/2} = \\ = \frac{\Gamma(T)}{2^{T-1} \Gamma\left[\frac{1}{2}(T-1)\right] \Gamma\left[\frac{1}{2}(T+1)\right]} (1 + r'_{2j-1})^{(T-3)/2} (1 - r'_{2j-1})^{(T-1)/2}$$

для нечетных. Хеннан (1960) предложил приближать $r'_{2j} + 2/T$ и $r'_{2j-1} + 1/T$ пирсоновскими коэффициентами корреляции, построенными по $T + 2$ наблюдениям. [См. также Кенуй (1949b).]

6.10. РАСПРЕДЕЛЕНИЯ ДЛЯ СЛУЧАЯ ЗАВИСИМЫХ НАБЛЮДЕНИЙ

Если наблюдения являются зависимыми, то распределения серийных коэффициентов корреляции в рассмотренных нами моделях легко вывести из соответствующих распределений для случая независимых наблюдений (приведенных, например, в разд. 6.7.5 и 6.7.6). [См. Лейпник (1947) и Мэдоу (1945).]

Рассмотрим плотность распределения величин $y_1, \dots, y_T$ в виде

$$(1) \quad K_i(\gamma_0, \dots, \gamma_i) \exp \left[-\frac{1}{2} (y - \mu \varepsilon)' (\gamma_0 A_0 + \gamma_1 A_1 + \dots + \gamma_i A_i) (y - \mu \varepsilon) \right] = K_i(\gamma_0, \dots, \gamma_i) \exp \left\{ -\frac{1}{2} [\gamma_0 Q_0^* + \dots + \gamma_i Q_i^* + T(\gamma_0 \lambda_0 + \gamma_1 \lambda_1 + \dots + \gamma_i \lambda_i) (\bar{y} - \mu)^2] \right\},$$

где $\varepsilon = (1, 1, \dots, 1)'$ — характеристический вектор матрицы A_i , соответствующий характеристическому корню λ_i ,

$$(2) \quad Q_i^* = (y - \bar{y} \varepsilon)' A_i (y - \bar{y} \varepsilon), \quad i = 0, 1, \dots, i,$$

числа $\gamma_0, \dots, \gamma_i$ таковы, что матрица $\gamma_0 A_0 + \gamma_1 A_1 + \dots + \gamma_i A_i$ положительно определена, $\gamma_0 > 0$, и

$$(3) \quad K_i(\gamma_0, \dots, \gamma_i) = | \gamma_0 A_0 + \gamma_1 A_1 + \dots + \gamma_i A_i |^{1/2} / (2\pi)^{T/2}.$$

Плотность совместного распределения квадратичных форм $Q_0^*, \dots, Q_i^*$ имеет вид

$$(4) \quad h_i^*(Q_0^*, \dots, Q_i^* | \gamma_0, \dots, \gamma_i) = K_i^*(\gamma_0, \dots, \gamma_i) \exp \left[-\frac{1}{2} (\gamma_0 Q_0^* + \dots + \gamma_i Q_i^*) \right] k_i^*(Q_0^*, \dots, Q_i^*),$$

на множестве допустимых значений $Q_0^*, \dots, Q_i^*$ и равна нулю вне этого множества; здесь

$$(5) \quad K_i^*(\gamma_0, \dots, \gamma_i) = \frac{\sqrt{2\pi} K_i(\gamma_0, \dots, \gamma_i)}{\sqrt{\gamma_0 \lambda_0 + \gamma_1 \lambda_1 + \dots + \gamma_i \lambda_i}}$$

(соответствующим образом нормировано и $k_i^*(Q_0^*, \dots, Q_i^*)$). Плотность совместного распределения величин $Q_0^*, r_1^* = Q_1^*/Q_0^*, \dots, r_i^* =$

$= Q_i^*/Q_0^*$ равна

$$(6) \quad h_i^*(Q_0^*, Q_0^*r_1^*, \dots, Q_0^*r_i^* | \gamma_0, \gamma_0\rho_1, \dots, \gamma_0\rho_i) = \\ = K_i^*(\gamma_0, \gamma_0\rho_1, \dots, \gamma_0\rho_i) \times \\ \times \exp\left[-\frac{1}{2} \gamma_0 Q_0^* \left(1 + \sum_{j=1}^i \rho_j r_j^*\right)\right] Q_0^{*i} k_i^*(Q_0^*, Q_0^*r_1^*, \dots, Q_0^*r_i^*)$$

на множестве допустимых значений $Q_0^*, r_1^*, \dots, r_i^*$ и нулю вне этого множества. Здесь $\rho_1 = \gamma_1/\gamma_0, \dots, \rho_i = \gamma_i/\gamma_0$. Плотность совместного распределения величин $r_1^*, \dots, r_i^*$ есть

$$(7) \quad g_i^*(r_1^*, \dots, r_i^* | \rho_1, \dots, \rho_i) = K_i^*(\gamma_0, \gamma_0\rho_1, \dots, \gamma_0\rho_i) \times \\ \times \int \exp\left[-\frac{1}{2} \gamma_0 Q_0^* \left(1 + \sum_{j=1}^i \rho_j r_j^*\right)\right] Q_0^{*i} k_i^*(Q_0^*, Q_0^*r_1^*, \dots, Q_0^*r_i^*) dQ_0^*,$$

где интегрирование производится по всем значениям Q_0^* , совместным с $r_1^*, \dots, r_i^*$. (Мы покажем далее, что область интегрирования есть интервал $(0, \infty)$ и что g_i^* не зависит от γ_0 .) При $\rho_1 = \dots = \rho_i = 0$ соотношение (7) равносильно соотношению

$$(8) \quad g_i^*(r_1^*, \dots, r_i^* | 0, \dots, 0) \gamma_0^{-(T-1)/2} \lambda_0^{1/2} |A_0|^{-1/2} (2\pi)^{(T-1)/2} = \\ = \int \exp\left(-\frac{1}{2} \gamma_0 Q_0^*\right) Q_0^{*i} k_i^*(Q_0^*, Q_0^*r_1^*, \dots, Q_0^*r_i^*) dQ_0^*,$$

тождественному по γ_0 . Поскольку в этом случае $r_1^*, \dots, r_i^*$ не зависят от Q_0^* , то область интегрирования по Q_0^* не должна зависеть от $r_1^*, \dots, r_i^*$, а это и означает, что она представляет собой интервал $(0, \infty)$. Частные случаи плотности $g_i^*(r_1^*, \dots, r_i^* | 0, \dots, 0)$ уже обсуждались в § 6.7 и 6.9. Используя тождественность соотношения (8) по γ_0 , интеграл в (7) можно найти, заменив γ_0 в (8) на $\gamma_0 (1 + \sum_{j=1}^i \rho_j r_j^*)$.

Так как здесь $\rho_1, \dots, \rho_i$ таковы, что матрица $A_0 + \rho_1 A_1 + \dots + \rho_i A_i$ положительно определена, имеем $1 + \rho_1 r_1^* + \dots + \rho_i r_i^* > 0$. В итоге получим

$$(9) \quad g_i^*(r_1^*, \dots, r_i^* | \rho_1, \dots, \rho_i) = \\ = \lambda_0^{1/2} |A_0|^{-1/2} (2\pi)^{1/2(T-1)} K_i^*(\gamma_0, \gamma_0\rho_1, \dots, \gamma_0\rho_i) \gamma_0^{-1/2(T-1)} \times \\ \times \left(1 + \sum_{j=1}^i \rho_j r_j^*\right)^{-1/2(T-1)} g_i^*(r_1^*, \dots, r_i^* | 0, \dots, 0) =$$

$$= \lambda_0^{1/2} |A_0|^{-1/2} (2\pi)^{1/2(T-1)} K_i^* (1, \rho_1, \dots, \rho_i) \times \\ \times \left(1 + \sum_{j=1}^i \rho_j r_j^*\right)^{-1/2(T-1)} g_i^* (r_1^*, \dots, r_i^* | 0, \dots, 0),$$

$$(10) \quad (2\pi)^{(T-1)/2} K_i^* (\gamma_0, \gamma_0 \rho_1, \dots, \gamma_0 \rho_i) = \\ = \frac{| \gamma_0 A_0 + \gamma_0 \rho_1 A_1 + \dots + \gamma_0 \rho_i A_i |^{1/2}}{(\gamma_0 \lambda_0 + \gamma_0 \rho_1 \lambda_1 + \dots + \gamma_0 \rho_i \lambda_i)^{1/2}} = \\ = \gamma_0^{(T-1)/2} \frac{| A_0 + \rho_1 A_1 + \dots + \rho_i A_i |^{1/2}}{(\lambda_0 + \rho_1 \lambda_1 + \dots + \rho_i \lambda_i)^{1/2}}.$$

Моменты могут быть найдены из соотношения (9) путем разложения $(1 + \sum_{j=1}^i \rho_j r_j^*)^{-(T-1)/2}$ для достаточно малых значений ρ_j в ряд по степеням $\sum_{j=1}^i \rho_j r_j^*$ и последующего интегрирования по переменным $r_1^*, \dots, r_i^*$.

Другой подход к получению распределений сериальных коэффициентов корреляции связан с использованием канонической формы. Пусть

$$(11) \quad r^* = \frac{\sum_{t=1}^{T-1} \lambda_t z_t^2}{\sum_{t=1}^{T-1} z_t^2},$$

где $z_1, \dots, z_{T-1}$ — независимые нормально распределенные случайные величины с нулевыми средними и дисперсиями $1/(\gamma_0 + \gamma_1 \lambda_1), \dots, 1/(\gamma_0 + \gamma_1 \lambda_{T-1})$ соответственно. Тогда

$$(12) \quad u_t = \sqrt{\gamma_0 + \gamma_1 \lambda_t} z_t$$

распределено по закону $N(0, 1)$. Запишем

$$(13) \quad \Pr \{r^* \geq R^*\} = \Pr \left\{ \sum_{t=1}^{T-1} \frac{\lambda_t u_t^2}{\gamma_0 + \gamma_1 \lambda_t} \geq R^* \sum_{t=1}^{T-1} \frac{u_t^2}{\gamma_0 + \gamma_1 \lambda_t} \right\} = \\ = \Pr \left\{ \sum_{t=1}^{T-1} \frac{\lambda_t - R^*}{\gamma_0 + \gamma_1 \lambda_t} u_t^2 \geq 0 \right\} = \\ = \Pr \left\{ \sum_{t=1}^{T-1} \frac{\lambda_t - R^*}{1 + \rho \lambda_t} u_t^2 \geq 0 \right\},$$

где $\rho = \gamma_1/\gamma_0$ в случае $r^* = r_1^*$. Здесь можно воспользоваться распределениями § 6.7 с заменой $\lambda_t - R$ на $(\lambda_t - R^*)/(1 + \rho \lambda_t)$. Например,

если $\lambda_h = \lambda_{T-h} = v_h$, $h = 1, \dots, H = (T-1)/2$, можно переписать (52) в виде

$$(14) \quad \text{Pr} \{r^* \geq R^*\} = \sum_{i=1}^m \left(\frac{v_i - R^*}{1 + \rho R^*} \right)^{H-1} \prod_{\substack{j=1 \\ j \neq i}}^H \frac{1 + \rho v_j}{v_i - v_j},$$

$$v_{m+1} \leq R^* \leq v_m, \quad m = 1, \dots, H-1.$$

6.11. ОЦЕНКИ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

6.11.1. Стационарные процессы первого порядка

Стационарный гауссовский процесс первого порядка удовлетворяет уравнению

$$(1) \quad y_t + \beta y_{t-1} = u_t, \quad t = \dots, -1, 0, 1, \dots,$$

в котором u_t — независимые случайные величины, каждая из которых распределена по закону $N(0, \sigma^2)$, а $|\beta| < 1$; причем $\mathbb{E} y_t = 0$, $\mathbb{E} y_t^2 = \sigma^2/(1 - \beta^2)$ и маргинальное распределение y_t есть $N[0, \sigma^2(1 - \beta^2)]$. Совместное распределение величин $y_1, \dots, y_T$ является нормальным $N(0, \Sigma)$; его ковариационная матрица Σ состоит из элементов $\sigma_{st} = \sigma^2(-\beta)^{|s-t|}/(1 - \beta^2)$, $s, t = 1, \dots, T$. Соответствующая плотность представляет частный случай (5) из § 6.2 и рассматривалась в разд. 6.3.3.

Найдем в этой модели оценку максимального правдоподобия для параметра β . (Эту задачу исследовал Купменс (1942).) Положим

$$(2) \quad P_0 = \sum_{t=2}^{T-1} y_t^2,$$

$$(3) \quad P'_0 = y_1^2 + y_T^2,$$

$$(4) \quad P_1 = \sum_{t=2}^T y_t y_{t-1}.$$

Тогда логарифм функции правдоподобия можно записать в виде

$$(5) \quad \log L = \frac{1}{2} \log(1 - \beta^2) - \frac{1}{2} T \log 2\pi - \frac{1}{2} T \log \sigma^2 - \\ - \frac{1}{2} [P'_0 + 2\beta P_1 + (1 + \beta^2) P_0]/\sigma^2.$$

Его производные по β и σ^2 имеют вид

$$(6) \quad \frac{\partial \log L}{\partial \beta} = -\frac{\beta}{1-\beta^2} - \frac{P_1}{\sigma^2} - \frac{\beta P_0}{\sigma^2},$$

$$(7) \quad \frac{\partial \log L}{\partial \sigma^2} = -\frac{T}{2} \cdot \frac{1}{\sigma^2} + \frac{P'_0 + 2\beta P_1 + (1 + \beta^2) P_0}{2\sigma^4}.$$

Приравнивая их нулю, получаем

$$(8) \quad \hat{\sigma}^2 = \frac{P'_0 + 2\hat{\beta}P_1 + (1 + \hat{\beta}^2) P_0}{T}$$

и

$$(9) \quad (1 - \hat{\beta}^2) (P_1 + \hat{\beta} P_0) + \hat{\beta} \left\{ \frac{P'_0 + 2\hat{\beta}P_1 + (1 + \hat{\beta}^2) P_0}{T} \right\} = 0,$$

т. е.

$$(10) \quad \hat{\beta}^3 \frac{T-1}{T} P_0 + \hat{\beta}^2 \frac{T-2}{T} P_1 - \hat{\beta} \left(\frac{T+1}{T} P_0 + \frac{1}{T} P'_0 \right) - P_1 = 0.$$

Левая часть последнего выражения положительна при $\hat{\beta} = -1$, отрицательна при $\hat{\beta} = 1$ и равна $-P_1$ при $\hat{\beta} = 0$. Отсюда следует, что последнее уравнение имеет один нулевой корень при $P_1 = 0$, один корень между -1 и 0 при $P_1 > 0$ и один корень между 0 и 1 при $P_1 < 0$. Кроме того, оно имеет один корень левее -1 и один — правее 1 . При $T \rightarrow \infty$ это уравнение сходится (по вероятности) к уравнению

$$(11) \quad \hat{\beta}^3 + \frac{P_1}{P_0} \hat{\beta}^2 - \hat{\beta} - \frac{P_1}{P_0} = 0.$$

Последнее имеет корни $\hat{\beta} = \pm 1$ и $-P_1/P_0$. Более точным приближением к соответствующему решению уравнения (10) является $-(P_1/P_0) (1 - 1/T)$. (См. упр. 78.)

Из изложенного вытекает, что в качестве оценки для параметра β можно использовать величины $-P_1/P_0$ или $-P_1/(P'_0 + P_0)$. Следует отметить, что кубическое уравнение (11) не имеет простого решения, например в виде отношения полиномов от P_0 , P_1 и P'_0 . Но численными методами его решение можно получить сколь угодно точно.

6.11.2. Циклическая модель

Пусть

$$(12) \quad y_t + \beta y_{t-1} = u_t, \quad t = 1, \dots, T,$$

где $y_0 \equiv y_T$, а случайные величины $u_1, \dots, u_T$ независимы и каждая из них распределена по закону $N(0, \sigma^2)$. Плотность совместного

распределения величин $y_1, \dots, y_T \equiv y_0$ равна тогда

$$(13) \quad \frac{1 - (-\beta)^T}{(\sqrt{2\pi\sigma^2})^T} \exp \left[-\frac{1}{2\sigma^2} \sum_{t=1}^T (y_t + \beta y_{t-1})^2 \right],$$

поскольку при переходе от переменных u_t к переменным y_t по формуле (12) соответствующий якобиан преобразования равен $1 - (-\beta)^T$. Матрица квадратичной формы в (13) представляет собой умноженную на $-1/2$ матрицу $\gamma_0 \mathbf{A}_0 + \gamma_1 \mathbf{A}_1$ с $\mathbf{A}_1$ из разд. 6.5.2 и $\gamma_0 = (1 + \beta^2)/\sigma^2$, $\gamma_1 = 2\beta/\sigma^2$. Производные логарифма функции правдоподобия равны здесь

$$(14) \quad \frac{\partial \log L}{\partial \beta} = \frac{T(-\beta)^{T-1}}{1 - (-\beta)^T} - \frac{1}{\sigma^2} \sum_{t=1}^T (y_t + \beta y_{t-1}) y_{t-1},$$

$$(15) \quad \frac{\partial \log L}{\partial \sigma^2} = -\frac{T}{2} \cdot \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{t=1}^T (y_t + \beta y_{t-1})^2.$$

Приравнявая их нулю, получаем следующие уравнения для оценок максимального правдоподобия:

$$(16) \quad \hat{\sigma}^2 = \frac{1}{T} \left[(1 + \hat{\beta}^2) \sum_{t=1}^T y_t^2 + 2\hat{\beta} \sum_{t=1}^T y_t y_{t-1} \right],$$

$$(17) \quad (-\hat{\beta})^T \sum_{t=1}^T y_t y_{t-1} - (-\hat{\beta})^{T-1} \sum_{t=1}^T y_t^2 + \hat{\beta} \sum_{t=1}^T y_t^2 + \sum_{t=1}^T y_t y_{t-1} = 0.$$

Если обозначить через r циклический сериальный коэффициент корреляции

$$(18) \quad r = \frac{\sum_{t=1}^T y_t y_{t-1}}{\sum_{t=1}^T y_t^2},$$

то эти уравнения можно переписать в виде

$$(19) \quad \hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^T y_t^2 [1 + \hat{\beta}^2 + 2\hat{\beta}r],$$

$$(20) \quad r(-\hat{\beta})^T - (-\hat{\beta})^{T-1} + \hat{\beta} + r = 0.$$

Если $r > 0$, то уравнение (20) имеет два действительных корня $\hat{\beta} = -a$ и $\hat{\beta} = -1/a$, $a > 0$, если T четное, и три действительных корня $-a$, $-1/a$, 1, если T нечетное. Если $r = 0$, то при четном T действительные корни равны нулю, а при нечетном T — нулю и единице. Наконец, если $r < 0$, то при четном T действительные корни равны a и $1/a$, $a > 0$, а при нечетном T и при $-(1 - 2/T) < r$, они равны a , $1/a$ и 1, но лишь 1 при $-\cos \pi/T < r \leq -(1 - 2/T)$. По-

сколько взаимно обратные величины $\hat{\beta}$ (при соответствующим образом скорректированных σ^2) определяют один и тот же процесс, можно ограничиться отысканием корня, лежащего между -1 и 1 . (Корень, равный 1 , не дает абсолютного максимума L .) Можно показать, что асимптотически корнем второго уравнения является $\hat{\beta} = -r$ (ввиду того что $rT \rightarrow 0$ и $\hat{\beta}T \rightarrow 0$ по вероятности). (См. упр. 85.) Можно также показать, что критерий отношения правдоподобия для проверки гипотезы $\beta = 0$ основывается на статистике

$$(21) \quad \frac{[1 + \hat{\beta}^2 + 2\hat{\beta}r]^{T/2}}{1 - (-\hat{\beta})^T}.$$

Указанная гипотеза отвергается, если (21) меньше надлежащим образом выбранной постоянной.

Отметим, что каждый корень уравнения (20) как функция от r не является рациональной функцией. Метод, основанный на максимуме правдоподобия (использующий $\hat{\beta}$), можно аппроксимировать приближением, описанным в предыдущих параграфах (использующим r). (Эти задачи исследованы Диксоном (1944).)

Уравнение (12) можно переписать в виде $(I + \beta V') u = u$, эквивалентном представлению

$$(22) \quad u = (I + \beta V')^{-1} u.$$

Поскольку $V^T = I$, то

$$(23) \quad (I + \beta V')^{-1} = \sum_{s=0}^{T-1} (-\beta)^s V'^s + (-\beta)^T (I + \beta V')^{-1}$$

или

$$(24) \quad (I + \beta V')^{-1} = \frac{1}{1 - (-\beta)^T} \sum_{s=0}^{T-1} (-\beta)^s V'^s.$$

Поэтому ковариационная матрица для u имеет вид

$$(25) \quad \begin{aligned} \mathfrak{L}_{uu'} &= (I + \beta V')^{-1} \mathfrak{L}_{uu'} (I + \beta V)^{-1} = \\ &= \frac{\sigma^2}{[1 - (-\beta)^T]^2} \sum_{s,t=0}^{T-1} (-\beta)^{s+t} V'^{t-s} = \\ &= \frac{\sigma^2}{[1 - (-\beta)^T]^2} \sum_{r=-(T-1)}^{T-1} \sum_{s \in S_r} (-\beta)^{2s+r} V^r, \end{aligned}$$

где $S_r = \{0, 1, \dots, T-1-r\}$ для $r \geq 0$ и $S_r = \{-r, -r+1, \dots, T-1\}$ для $r \leq 0$. Здесь использован тот факт, что $V'^q = V^{-q}$. Замечая еще, что $V^{-q} = V^{T-q}$, получаем

$$\begin{aligned}
 (26) \quad \mathcal{E}yy' &= \frac{\sigma^2}{[1 - (-\beta)^T]^2} \left\{ \sum_{s=0}^{T-1} (-\beta)^{2s} \mathbf{I} + \sum_{q=1}^{T-1} \left[\sum_{s=0}^{T-1-q} (-\beta)^{2s+q} \mathbf{B}^q + \right. \right. \\
 &\quad \left. \left. + \sum_{s=q}^{T-1} (-\beta)^{2s-q} \mathbf{B}^{-q} \right] \right\} = \\
 &= \frac{\sigma^2}{[1 - (-\beta)^T]^2 (1 - \beta^2)} \times \\
 &\quad \times \left\{ (1 - \beta^{2T}) \mathbf{I} + \sum_{q=1}^{T-1} [(-\beta)^q - (-\beta)^{2T-q}] (\mathbf{B}^q + \mathbf{B}^{-q}) \right\} = \\
 &= \frac{\sigma^2}{[1 - (-\beta)^T]^2 (1 - \beta^2)} \left\{ (1 - \beta^{2T}) \mathbf{I} + \sum_{q=1}^{T-1} (-\beta)^q [1 + \right. \\
 &\quad \left. + (-\beta)^{T-2q} - (-\beta)^T - (-\beta)^{2T-2q}] \mathbf{B}^q \right\} = \\
 &= \frac{\sigma^2}{[1 - (-\beta)^T]^2 (1 - \beta^2)} \left\{ [1 + (-\beta)^T] \mathbf{I} + \right. \\
 &\quad \left. + \sum_{q=1}^{T-1} [(-\beta)^q + (-\beta)^{T-q}] \mathbf{B}^q \right\}.
 \end{aligned}$$

Ковариация между y_t и y_{t+h} равна

$$(27) \quad \sigma(h) = \mathcal{E}y_t y_{t+h} = \frac{(-\beta)^h + (-\beta)^{T-h}}{[1 - (-\beta)^T] (1 - \beta^2)} \sigma^2, \quad h = 0, 1, \dots, T-1.$$

Соответствующая корреляция $[(-\beta)^h + (-\beta)^{T-h}] / [1 + (-\beta)^T]$ складывается из двух членов, один из которых зависит от интервала времени между t и $t+h$, а другой — от интервала времени между t и $t+T-h$. Матрица (26) является обратной для $[(1 + \beta^2)\mathbf{I} + \beta(\mathbf{B} + \mathbf{B}')] / \sigma^2$.

6.12. ОБСУЖДЕНИЕ

На основе сериальных коэффициентов корреляции первого порядка можно построить процедуры проверки нулевой гипотезы о независимости $y_1, \dots, y_T$ против альтернативы, состоящей в том, что эти величины коррелированы во времени. В предположении нормальности, проверка такой гипотезы может производиться с заранее выбранным уровнем значимости. Для проверки гипотез о зависимости более высокого порядка могут быть использованы частные сериальные корреляции.

В гл. 10 мы обсудим критерии для проверки независимости и критерии для проверки гипотез о порядке зависимости в случае, когда $\mathcal{E}y_t = f(t)$ является нетривиальным трендом.

ЛИТЕРАТУРА

- § 6.2. Т. Андерсон (1963).
 § 6.3. Т. Андерсон (1948), (1963), Кендалл и Стьюарт (1961), Леман (1959).
 § 6.4. Т. Андерсон (1963).
 § 6.5. Т. Андерсон (1963), Ватсон и Дурбин (1951).
 § 6.6. Р. Андерсон и Т. Андерсон (1950).
 § 6.7. Р. Андерсон (1941), (1942), Диксон (1944), Купменс (1942), Лаплас (1829), фон Нейман (1941).
 § 6.8. Т. Андерсон (1948), Бюро переписей США (1955), Диксон (1944), Купменс (1942), Рубин (1945), Харт (1942), Харт и фон Нейман (1942), Хеннан (1960), Янг (1941).
 § 6.9. Т. Андерсон (1958), Ватсон (1956), Даниэлс (1956), Хеннан (1960). Дженкинс (1956), Кенуй (1949а), (1949б).
 § 6.10. Лейпник (1947), Мэдоу (1945).
 § 6.11. Диксон (1944), Купменс (1942).

УПРАЖНЕНИЯ

1. (§ 6.2) Пусть $y_1, \dots, y_T$ являются последовательными элементами реализации стационарного гауссовского процесса порядка p . Покажите, что достаточное множество статистик образуют здесь следующие функции от наблюдений:

$$\begin{aligned} & \sum_{t=p+1}^{T-p} y_t^2, y_1^2 + y_T^2, y_2^2 + y_{T-1}^2, \dots, y_p^2 + y_{T-p+1}^2, \\ & \sum_{t=p+1}^{T-p+1} y_t y_{t-1}, y_1 y_2 + y_{T-1} y_T, \dots, y_{p-1} y_p + y_{T-p+2} y_{T-p+1}, \\ & \vdots \\ & \sum_{t=p+1}^T y_t y_{t-p}. \end{aligned}$$

2. (§ 6.2) Пусть $y_1, \dots, y_T$ — последовательные T элементов (с ковариационной матрицей $\Sigma_T = [\sigma(i-j)]$) реализации стационарного процесса, порожденного скользящим средним $y_t = v_t + \alpha_1 v_{t-1}$, где $\mathbb{E} v_t = 0$, $\mathbb{E} v_t^2 = \sigma^2$ и $\mathbb{E} v_t v_s = 0$, $t \neq s$. Тогда $\sigma(0) = \sigma^2 (1 + \alpha_1^2)$, $\sigma(1) = \sigma^2 \alpha_1 = \sigma(0) \rho$ (с некоторым ρ) и $\sigma(h) = 0$, $h > 1$. Покажите, что

$$\Sigma_3^{-1} = \frac{1}{(1 - 2\rho^2) \sigma(0)} \begin{pmatrix} 1 - \rho^2 & -\rho & \rho^2 \\ -\rho & 1 & -\rho \\ \rho^2 & -\rho & 1 - \rho^2 \end{pmatrix},$$

$$\Sigma_4^{-1} =$$

$$= \frac{1}{(1 - 3\rho^2 + \rho^4) \sigma(0)} \begin{pmatrix} 1 - 2\rho^2 & -\rho(1 - \rho^2) & \rho^2 & -\rho^3 \\ -\rho(1 - \rho^2) & 1 - \rho^2 & -\rho & \rho^2 \\ \rho^2 & -\rho & 1 - \rho^2 & -\rho(1 - \rho^2) \\ -\rho^3 & \rho^2 & -\rho(1 - \rho^2) & 1 - 2\rho^2 \end{pmatrix}.$$

3. (§ 6.2) Покажите, что матрица Σ_T^{-1} , определенная в упр. 2, является произведением $\sigma T^{-1}(0) / |\Sigma_T|$ на матрицу, элементы которой являются полиномами от ρ .

Покажите, что в этих полиномах встречаются все степени ρ от 0 до $T-1$ и что существует T линейно независимых полиномов. (Указание. Элементы обратной матрицы являются отношениями соответствующих алгебраических дополнений и определителю.)

4. (§ 6.2) Предположим, что $y_1, \dots, y_T$ являются последовательными элементами реализации процесса, определенного в упр. 2. Покажите, что размерность минимального достаточного множества статистик для σ (0) и ρ равна T . (Указание. Использовать результат упр. 3 и с его помощью убедиться, что показатель экспоненты в соответствующей плотности пропорционален $\sum_{j=0}^{T-1} \rho^j P_j$, где $P_j, j =$

$= 0, \dots, T-1$, — линейно независимые квадратичные формы. Если $Q_0, \dots, Q_S$ ($S \leq T-1$) — другое достаточное множество статистик, то соответствующий показатель экспоненты пропорционален $\sum_{j=0}^S f_j(\rho) R_j$, где $R_j, j = 0, \dots, S$, являются функциями от $Q_j, j = 1, \dots, T-1$, и обе экспоненты совпадают.)

5. (§ 6.2) Пусть Σ_T — матрица, определенная в упр. 2, и

$$a_{ij} = [-\rho \sigma(0)]^{j-i} |\Sigma_{i-1}| \cdot |\Sigma_{T-j}| / |\Sigma_T|, \quad i \leq j.$$

Покажите, что $A = [a_{ij}] = \Sigma_T^{-1}$. (Указание. Вычислить алгебраическое дополнение к элементу с индексами i, j с помощью разложения Лапласа по первым $i-1$ и затем по последним $T-j$ столбцам. Другой метод состоит в том, что с помощью леммы 6.7.9 показывают, что $|\Sigma_i|$ удовлетворяет некоторому разностному уравнению второго порядка, и убеждаются в том, что $\Sigma_T A = I$. Для проверки диагональных членов (за исключением первого и последнего) используется решение для $|\Sigma_i|$, а для проверки недиагональных элементов используется непосредственно само упомянутое разностное уравнение второго порядка.)

6. (§ 6.2) Покажите, что решением (13) для $\rho = 1$ является $\sigma^2 = 2(\gamma_0 \pm \sqrt{\gamma_0^2 - \gamma_1^2})/\gamma_1^2$ и $\beta_1 = (\gamma_0 \pm \sqrt{\gamma_0^2 - \gamma_1^2})/\gamma_1$.

7. (Разд. 6.3.1) Докажите теорему 6.3.2, т. е. что если

$$(i) \quad \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h_i(Q_0, \dots, Q_i | \gamma_0, \dots, \gamma_i) g(Q_0, \dots, Q_i) dQ_0 \dots dQ_i = 0$$

тождественно по $\gamma_0, \dots, \gamma_i$ (для которых $\sum_{j=0}^i \gamma_j Q_j$ положительно определена), то

$$(ii) \quad g(Q_0, \dots, Q_i) = 0$$

почти всюду по отношению к плотности $h_i(Q_0, \dots, Q_i | \gamma_0, \dots, \gamma_i)$. (Указание. Пусть $g(Q_0, \dots, Q_i) = g^+(Q_0, \dots, Q_i) - g^-(Q_0, \dots, Q_i)$, где g^+ и g^- — неотрицательные функции. Показать, что

$$(iii) \quad \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} \sum_{j=0}^i \gamma_j Q_j \right] g^+(Q_0, \dots, Q_i) h_i(Q_0, \dots, Q_i) dQ_0 \dots dQ_i \equiv \\ \equiv \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} \sum_{j=0}^i \gamma_j Q_j \right] g^-(Q_0, \dots, Q_i) h_i(Q_0, \dots, Q_i) dQ_0 \dots dQ_i.$$

Вывести (ii) из (iii) с помощью характеристических функций.)

8. (Разд. 6.3.1) Покажите, что из (4) следует (3).

9. (Разд. 6.3.1) Покажите, что из (3) следует (4) почти всюду. (Указание. Использовать свойство полноты, доказывавшееся в упр. 7 с

$$g(Q_0, \dots, Q_{i-1}) = \frac{\Pr \{S_i | Q_0, \dots, Q_{i-1}\}}{e_i} - 1.)$$

10. (Разд. 6.3.1) *Фундаментальная лемма Неймана — Пирсона*. Пусть $f_0(x)$ и $f_1(x)$ — функции плотности на некотором евклидовом пространстве, а R — произвольное (измеримое) множество, такое, что

$$(i) \quad \int_R f_0(x) dx = \varepsilon,$$

где dx указывает на интегрирование по Лебегу в этом пространстве и $0 \leq \varepsilon \leq 1$. Докажите, что если

$$(ii) \quad R^* = \{x | f_1(x) > k f_0(x)\}$$

для некоторого k и

$$\int_{R^*} f_0(x) dx = \varepsilon,$$

то

$$(iii) \quad \int_{R^*} f_1(x) dx \geq \int_R f_1(x) dx.$$

(Указание. В обе части неравенства (iii) интегралы по $R^* \cap R$ вносят одинаковый вклад. Сравните интегралы от $f_1(x)$ по $R^* \cap \bar{R}$ и $\bar{R}^* \cap R$ с соответствующими интегралами от $f_0(x)$.)

11. (Разд. 6.3.1) Пусть

$$y_{ij} = a_i + e_{ij}, \quad i = 1, \dots, M, \quad j = 1, \dots, N,$$

где случайные величины a_i и e_{ij} , $i = 1, \dots, M$, $j = 1, \dots, N$, независимы, распределение a_i есть $N(0, \sigma_a^2)$, а распределение величины e_{ij} есть $N(0, \sigma_e^2)$. Представьте плотность совместного распределения величин y_{ij} в виде (1). Найдите равномерно наиболее мощный подобный критерий для проверки гипотезы $\sigma_a^2 = 0$ против альтернатив $\sigma_a^2 > 0$ с уровнем значимости α . (Указание. Ковариационная матрица случайных величин $y_{11}, \dots, y_{1N}, y_{21}, \dots, y_{2N}, \dots, y_{M1}, \dots, y_{MN}$ равна $\sigma_e^2 I_{MN} + \sigma_a^2 I_M \otimes \mathbf{e}\mathbf{e}'$, где вектор $\mathbf{e} = (1, 1, \dots, 1)'$ состоит из N компонент, а символ $\otimes$ означает кронекерово произведение. Обратная матрица имеет вид $b I_{MN} + c I_M \otimes \mathbf{e}\mathbf{e}'$.)

12. (Разд. 6.3.2) *Обобщенная фундаментальная лемма Неймана — Пирсона*. Пусть $f_0(x)$ и $f_2(x)$ — функции плотности на некотором евклидовом пространстве, а $f_1(x)$ — другая функция, и пусть R — произвольное измеримое множество, такое, что

$$(i) \quad \int_R f_0(x) dx = \varepsilon,$$

$$(ii) \quad \int_R f_1(x) dx = c,$$

где dx указывает на то, что интегрирование производится по Лебегу, $0 \leq \varepsilon \leq 1$, а c — некоторая постоянная. Докажите, что если

$$(iii) \quad R^* = \{x \mid f_2(x) > k_0 f_0(x) + k_1 f_1(x)\}$$

для некоторых k_0, k_1 и на этом множестве выполняются соотношения (i), (ii), то

$$(iv) \quad \int_{R^*} f_2(x) dx \geq \int_R f_2(x) dx.$$

(Указание. Вклад интегралов по $R^* \cap R$ в обеих частях (iv) одинаков. Сравнить интегралы от $f_2(x)$ по $R^* \cap \bar{R}$ и $\bar{R}^* \cap R$ с соответствующими интегралами от $k_0 f_0(x) + k_1 f_1(x)$.)

13. (Разд. 6.3.2) Докажите, что область $R = (-\infty, c)$ не может удовлетворять (16) и (20). (Указание. Показать, что пара соотношений (16) и (20) была бы равносильна в противном случае соотношению

$$(i) \quad \int_{-\infty}^c (Q_i - \mu) h_i(Q_i \mid Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) dQ_i = 0,$$

где

$$(ii) \quad \mu = \int_{-\infty}^{\infty} Q_i h_i(Q_i \mid Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) dQ_i.$$

14. (Разд. 6.3.2) Пусть для некоторого стационарного процесса $\sigma(h) = \sigma(0) \rho_h$. Покажите, что частная корреляция между y_t и y_{t+2} при фиксированном y_{t+1} равна

$$\frac{\rho_2 - \rho_1^2}{1 - \rho_1^2},$$

а частная корреляция между y_t и y_{t+3} при фиксированных y_{t+1} и y_{t+2} равна

$$\frac{\rho_3 - \rho_1^2 \rho_3 + \rho_1^3 - 2\rho_1 \rho_2 + \rho_1 \rho_2^2}{1 - 2\rho_1^2 + 2\rho_1^2 \rho_2 - \rho_2^2}.$$

15. (Разд. 6.5.2) Докажите, что все элементы матрицы B^j равны нулю, за исключением элементов, принадлежащих диагоналям, расположенным на j единиц выше и на $T - j$ единиц ниже главной диагонали, и равных 1, $j = 0, 1, \dots, T - 1$.

16. (Разд. 6.5.2) Докажите, что $B^T = I$.

17. (Разд. 6.5.2) Докажите теорему 6.5.2 методом, используемым при доказательстве теоремы 6.5.4.

18. (Разд. 6.5.3) Используя матрицу A_1 , определенную соотношением (41), найдите по формулам (18) матрицы A_4, A_5 и A_6 .

19. (Разд. 6.5.3) Убедитесь в ортогональности матрицы (55), используя суммы, вычисленные в гл. 4.

20. (Разд. 6.5.3) Докажите, что матрицы C_r и C'_t отличаются только теми элементами, для которых $s, t \leq r$ и $(T - s + 1), (T - t + 1) \leq r$.

21. (Разд. 6.5.3) Выпишите матрицу C_1^5 в явном виде.

22. (Разд. 6.5.3) Докажите соотношение (61). (Указание. $G_1 = 2(I - A_1)$, а A_j является полиномом от A_1 степени j .)

23. (Разд. 6.5.4) Докажите следствие 6.5.4.

24. (Разд. 6.6.1) Пусть вектор x распределен по закону $N(\mu, \Sigma)$. Покажите, что любая несмещенная линейная оценка $m = \gamma'x$ параметра μ должна удовлетворять соотношению $\gamma'e = 1$. Покажите, что для того, чтобы такая оценка не зависела от остатков $x - m$, она должна быть марковской, т. е.

$$\gamma = \frac{1}{e'\Sigma^{-1}e} \Sigma^{-1}e.$$

25. (Разд. 6.6.1) Покажите, что если вектор x распределен по закону $N(\mu, \Sigma)$, то для независимости x и $x - x\bar{x}$ необходимо и достаточно, чтобы $\bar{x}$ являлся характеристическим вектором матрицы Σ .

26. (Разд. 6.6.1) Сформулируйте теоремы 6.4.1 и 6.4.2 для случая, когда $\xi_{y_i} = \mu$ и e — характеристический вектор матриц $A_0, \dots, A_p$.

27. (Разд. 6.7.1) Пусть вектор y при $T = 4$ имеет плотность

$$(i) \quad K(\gamma_0, \gamma_1) e^{-(\gamma_0 Q_0 + \gamma_1 Q_1)/2},$$

где $Q_0 = \sum_{i=1}^4 y_i^2$, $Q_1 = y_1^2 + y_2^2 - (y_3^2 + y_4^2)$, $K(\gamma_0, \gamma_1) = (\gamma_0^2 - \gamma_1^2)/(2\pi)^2$ и $|\gamma_1| < \gamma_0$. (а) Покажите, что плотность совместного распределения величин Q_0 и Q_1 равна

$$(ii) \quad \frac{\gamma_0^2 - \gamma_1^2}{8} e^{-(\gamma_0 Q_0 + \gamma_1 Q_1)/2}, \quad 0 \leq |Q_1| \leq Q_0.$$

(Указание. $(\gamma_0 + \gamma_1)(y_1^2 + y_2^2)$ и $(\gamma_0 - \gamma_1)(y_3^2 + y_4^2)$ независимы и имеют распределение χ^2 с двумя степенями свободы каждая.) (б) Покажите, что плотность совместного распределения величин Q_0 и $r = Q_1/Q_0$ равна

$$(iii) \quad \frac{\gamma_0^2 - \gamma_1^2}{8} e^{-(\gamma_0 + \gamma_1 r)Q_0/2}, \quad |r| \leq 1.$$

(с) Покажите, что маргинальная плотность для Q_0 при $\gamma_1 \neq 0$ равна

$$(iv) \quad \frac{\gamma_0^2 - \gamma_1^2}{4\gamma_1} [e^{-(\gamma_0 - \gamma_1)Q_0/2} - e^{-(\gamma_0 + \gamma_1)Q_0/2}], \quad Q_0 \geq 0.$$

(д) Найдите условную плотность для r при заданном значении Q_0 . (е) Найдите маргинальную плотность для r .

28. (Разд. 6.7.1) Пусть вектор x имеет плотность $f(x, \theta)$, где $\theta \in \Omega$. Предположим, что Q — достаточная статистика для θ и что существует группа преобразований g векторов x , порождающая соответствующую группу преобразований $\bar{g}$ статистик Q . Предположим также существование такого значения q_0 статистики Q , что для любого значения q_1 этой статистики найдется такое преобразование $\bar{g}$, при котором $\bar{g}q_1 = q_0$. Тогда распределение всякой статистики g , инвариантной относительно указанной группы преобразований, не зависит от Q и θ .

29. (Разд. 6.7.2) Выведите соотношения (10) для чисто мнимых значений $t_0, \dots, t_p$, используя тот факт, что

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2} \mathbf{x}' \mathbf{B} \mathbf{x}} d\mathbf{x}_1 \dots d\mathbf{x}_T = \frac{(2\pi)^{T/2}}{|\mathbf{B}|^{1/2}}.$$

30. (Разд. 6.7.2) Покажите, что если матрица $\mathbf{A}$ положительно определена, а матрица $\mathbf{B}$ симметрична, то существует невырожденная матрица $\mathbf{C}$, такая, что

$$\mathbf{C}'\mathbf{A}\mathbf{C} = \mathbf{I}, \quad \mathbf{C}'\mathbf{B}\mathbf{C} = \mathbf{D},$$

где матрица $\mathbf{D}$ диагональна и на ее диагонали стоят элементы, являющиеся корнями уравнения $|\mathbf{B} - d\mathbf{A}| = 0$. Покажите, что столбцы матрицы $\mathbf{C}$ являются характеристическими векторами матрицы $\mathbf{A}^{-1}\mathbf{B}$. (Указание. Пусть матрица $\mathbf{E}$ такова, что $\mathbf{E}'\mathbf{A}\mathbf{E} = \mathbf{I}$; в качестве $\mathbf{C}$ взять матрицу $\mathbf{E}\mathbf{\Theta}$, где $\mathbf{\Theta}$ — ортогональная матрица, приводящая $\mathbf{E}'\mathbf{B}\mathbf{E}$ к диагональному виду.)

31. (Разд. 6.7.3) Покажите, что если матрицы $\mathbf{A}_1, \dots, \mathbf{A}_p$ симметричны, то условия $\mathbf{A}_j\mathbf{A}_k = \mathbf{A}_k\mathbf{A}_j, j, k = 1, \dots, p$, необходимы и достаточны для существования ортогональной матрицы $\mathbf{Z}$, такой, что $\mathbf{Z}'\mathbf{A}_j\mathbf{Z} = \mathbf{\Lambda}_j$ — диагональная матрица, $j = 1, \dots, p$. (Указание. Соотношение $\mathbf{Z}'\mathbf{A}_j\mathbf{Z} = \mathbf{\Lambda}_j$ равносильно соотношению $\mathbf{A}_j = \mathbf{Z}\mathbf{\Lambda}_j\mathbf{Z}'$. Для доказательства достаточности следует предположить, что $\mathbf{A}_1\mathbf{A}_k = \mathbf{A}_k\mathbf{A}_1$, и использовать соотношение $\mathbf{A}_1 = \mathbf{Z}\mathbf{\Lambda}_1\mathbf{Z}'$.)

32. (Разд. 6.7.3) Пусть $\mathbf{A}_0$ — произвольная положительно определенная матрица, а $\mathbf{A}_1, \dots, \mathbf{A}_p$ — симметрические матрицы. Покажите, что если существует невырожденная матрица $\mathbf{C}$, столбцы которой являются надлежащим образом нормированными характеристическими векторами матриц $\mathbf{A}_0^{-1}\mathbf{A}_1, \dots, \mathbf{A}_0^{-1}\mathbf{A}_p$, то

$$\mathbf{C}'\mathbf{A}_0\mathbf{C} = \mathbf{I}, \quad \mathbf{C}'\mathbf{A}_j\mathbf{C} = \mathbf{\Lambda}_j, \quad j = 1, \dots, p,$$

где $\mathbf{\Lambda}_j$ — диагональная матрица, на диагонали которой стоят корни уравнения $|\mathbf{A}_j - \lambda\mathbf{A}_0| = 0$. Докажите, что необходимое и достаточное условие для этого состоит в том, что

$$\mathbf{A}_j\mathbf{A}_0^{-1}\mathbf{A}_k = \mathbf{A}_k\mathbf{A}_0^{-1}\mathbf{A}_j, \quad j, k = 1, \dots, p.$$

(Указание. См. упр. 30 и 31.)

33. (Разд. 6.7.4) Докажите, что $h^*(y_1, \dots, y_{H-2}) = k_2^* h(y_1, \dots, y_{H-2})$. (Указание: Пусть $\mathbf{Z}_1, \mathbf{Z}_2$ и $\mathbf{Z}_2^*$ — точки пересечения прямой, проходящей через точку $(y_1, \dots, y_{H-2})$ параллельно $\mathbf{V}_1\mathbf{V}_2$, с гранями $\mathbf{V}_1\mathbf{V}_3 \dots \mathbf{V}_H, \mathbf{V}_2\mathbf{V}_3 \dots \mathbf{V}_H$ и $\mathbf{V}_2^*\mathbf{V}_3 \dots \mathbf{V}_H$ соответственно, и пусть $\mathbf{W}$ — точка пересечения прямой $\mathbf{V}_3\mathbf{V}_4 \dots \mathbf{V}_H$ с плоскостью, проходящей через эту прямую и $\mathbf{V}_1\mathbf{V}_2$. Рассмотреть треугольники $\mathbf{V}_1\mathbf{V}_2\mathbf{W}, \mathbf{V}_1\mathbf{V}_2^*\mathbf{W}, \mathbf{Z}_1\mathbf{Z}_2\mathbf{W}$ и $\mathbf{Z}_1\mathbf{Z}_2^*\mathbf{W}$.)

34. (Разд. 6.7.4) Укажите множества U_1, U_2 и U_3 для $R < v_3$ (в барицентрических координатах с $H = 3$).

35. (Разд. 6.7.4) Покажите, что теорему 6.7.4 можно сформулировать в виде

$$\Pr \{r < R\} = \sum_{i=m+1}^H \frac{(v_i - R)^{H-1}}{\prod_{\substack{j=1 \\ j \neq i}}^H (v_i - v_j)}, \quad v_{m+1} \leq R \leq v_m, \quad m = 1, \dots, H-1.$$

(Указание. Использовать лемму 6.7.7.)

36. (Разд. 6.7.4) Докажите теорему 6.7.4 по индукции, применяя теорему 6.7.5. (Указание. Предположив, что (52) выполнено для $H = H^* - 1$, и используя теорему 6.7.5 с $M = 2(H^* - 1)$ и $L = 2$, убедиться в справедливости (52) для $H = H^*$.)

37. (Разд. 6.7.4) Докажите теорему 6.7.4 путем обращения (69). (Указание. Показать, что $u = \sum_{j=1}^H (v_j - R) x_j$ имеет (при $u \geq 0$) плотность

$$f(u) = \frac{1}{2} \sum_{j=1}^m e^{-\frac{1}{2} u/(v_j - R)} \frac{(v_j - R)^{H-2}}{\prod_{\substack{k=1 \\ k \neq j}}^H (v_j - v_k)}, \quad v_{m+1} \leq R \leq v_m, \quad m=1, \dots, H-1.$$

интегрируя умноженное на $e^{-itu}/(2\pi)$ выражение (69) вдоль замкнутого контура, содержащего полюсы, расположенные в нижней половине комплексной плоскости. Затем проинтегрировать эту плотность в пределах от 0 до ∞ ; см. (68).)

38. (Разд. 6.7.5) Покажите, что в циклической модели

(a) $r_1^* = -1$ для $T = 2$,

(b) $r_1^* = -1/2$ для $T = 3$,

(c) $-2r_1^*/(1 + r_1^*)$ имеет F -распределение с 1 и 2 степенями свободы (при нулевой гипотезе) для $T = 4$,

(d) если $c_1 = \cos 2\pi/5$, $c_2 = \cos 4\pi/5$, $T = 5$, то отношение $(c_1 - r_1^*)/(r_1^* - c_2)$ имеет (при нулевой гипотезе) F -распределение с 2 и 2 степенями свободы.

39. (Разд. 6.7.5) Покажите, что в циклической модели при нулевой гипотезе (a) отношение $(1 + r_1)/(1 - r_1)$ имеет при $T = 2$ F -распределение с 1 и 1 степенями свободы,

(b) отношение $(1 + 2r_1)/(1 - r_1)$ имеет при $T = 3$ F -распределение с 1 и 2 степенями свободы,

(c) распределение коэффициента r_1 при $T = 4$ совпадает с распределением отношения

$$\frac{F_{1,1}-1}{F_{1,1}+1} / (F_{2,2}+1),$$

в котором $F_{1,1}$ и $F_{2,2}$ — независимые случайные величины, имеющие распределения F с 1 и 1 степенями свободы и с 2 и 2 степенями свободы соответственно.

40. (Разд. 6.7.5) За период с 1858 по 1869 г. элементы ряда Вольфа чисел солнечной активности равны соответственно: 55, 94, 96, 77, 59, 44, 47, 30, 16, 7, 37, 74. Используя циклическую модель, вычислите Q_0^* , Q_1^* и r_1^* . Проверьте нулевую гипотезу с уровнем значимости 1% против альтернатив положительной корреляции первого порядка.

41. (Разд. 6.7.6) Покажите, что если случайные величины u , v и w независимы и имеют χ^2 -распределения с l , m и n степенями свободы соответственно, то распределение отношения $(\alpha u + \beta v + \gamma w)/(u + v + w)$ совпадает с распределением отношения

$$\left(\frac{\alpha x + \beta}{x + 1} + \gamma y \right) / (1 + y),$$

в котором x и y — независимые случайные величины, распределенные соответственно, как $lF_{l,m}/m$ и $nF_{n,l+m}/(l+m)$.

42. (Разд. 6.7.6) Найдите плотность сериального коэффициента корреляции r_1^* для модели с последовательными разностями при $T = 3$. Для этого получите плотность совместного распределения $r_1^* = (\lambda_1 u_1 + \lambda_2 + u_2) / (u_1 + u_2)$ и $v = u_1 + u_2$, где u_1 и u_2 — независимые случайные величины, каждая из которых имеет χ^2 -распределение с одной степенью свободы.

43. (Разд. 6.7.6) Покажите, что распределение коэффициента r_2^* для модели с последовательными разностями совпадает с распределением циклического коэффициента r_1^* . (Указание. Показать, что оба множества корней совпадают.)

44. (Разд. 6.7.7) Покажите, что если $\gamma_1 \neq 0$, то производящая функция моментов Q_1 получается из производящей функции моментов Q_1 для случая $\gamma_1 = 0$ заменой (-2θ) на $(\gamma_1 - 2\theta)$ и умножением последней на $K_1(\gamma_0, \gamma_1)/K_0(\gamma_0)$.

45. (Разд. 6.7.7) Покажите, что в лемме 6.7.9

$$D_3 = 1 - 2\theta^2,$$

$$D_4 = 1 - 3\theta^2 + \theta^4.$$

46. (Разд. 6.7.7) Покажите, что в лемме 6.7.10

$$\Delta_3 = 1 - 3\theta^2 - 2\theta^3 = (1 - 2\theta)(1 + \theta)^2,$$

$$\Delta_4 = 1 - 4\theta^2 = (1 - 2\theta)(1 + 2\theta).$$

47. (Разд. 6.7.7) Покажите, что в лемме 6.7.11

$$C_2 = 1 - 2\theta$$

$$C_3 = 1 - 2\theta - \theta^2 + 2\theta^3 = (1 - 2\theta)(1 - \theta^2),$$

$$C_4 = 1 - 2\theta - 2\theta^2 + 4\theta^3 = (1 - 2\theta)(1 - 2\theta^2),$$

$$C_5 = 1 - 2\theta - 3\theta^2 + 6\theta^3 + \theta^4 - 2\theta^5 = (1 - 2\theta)(1 - 3\theta^2 + \theta^4).$$

48. (Разд. 6.7.7) Покажите, что если $\gamma_1 \neq 0$ и $\mathbf{e} = (1, 1, \dots, 1)'$ является характеристическим вектором матриц $\mathbf{A}_0$ и $\mathbf{A}_1$, соответствующим характеристическому корню 1, то производящая функция моментов Q_1^* получается из производящей функции моментов Q_1^* для случая $\gamma_1 = 0$ заменой -2θ на $\gamma_1 - 2\theta$ и умножением последней на $K_1(\gamma_0, \gamma_1)/K_0(\gamma_0)$ и $\sqrt{\gamma_0/(\gamma_0 + \gamma_1)}$.

49. (Разд. 6.7.7) Покажите, что

$$\frac{d^n}{d\theta^n} F(\theta) G(\theta) = \sum_{j=0}^n \binom{n}{j} F^{(j)}(\theta) G^{(n-j)}(\theta),$$

где $F^{(j)}(\theta) = d^j F(\theta)/d\theta^j$ и $G^{(j)}(\theta) = d^j G(\theta)/d\theta^j$. (Указание. Провести доказательство по индукции (по аналогии с теоремой о биноме).)

50. (Разд. 6.7.7) Покажите, что если $f(0) = 0$, а $f^{(n)}(0)$ существует и конечно, $n = 1, \dots, H-1$, то

$$\left. \frac{d^n f^{(H)}(\theta)}{d\theta^n} \right|_{\theta=0} = 0, \quad n = 0, 1, \dots, H-1.$$

(Указание. Использовать предыдущее упражнение, полагая для доказательства по индукции $F(\theta) = f^{(H-1)}(\theta)$, $G(\theta) = f(\theta)$.)

51. (Разд. 6.7.7) Покажите, что если $G^{(n)} = 0$, $n = 0, 1, \dots, T-1$, и $\varphi(\theta) = F(\theta) [1 + G(\theta)]$, то

$$\varphi^{(n)}(0) = F^{(n)}(0), \quad n = 0, 1, \dots, T-1.$$

(Указание. Использовать упр. 49.)

52. (Разд. 6.7.7) Используя упр. 49, 50 и 51, докажите (107).

53. (Разд. 6.7.7) Покажите, что функция

$$h(\theta) = \left(\frac{1 - \sqrt{1 - 4\theta^2}}{1 + \sqrt{1 - 4\theta^2}} \right)^{1/2}$$

не имеет производной в точке $\theta = 0$.

54. (Разд. 6.7.7) Покажите, что при $T = 4$ производящая функция моментов Q_1 в случае циклического определения Q_1 имеет вид

$$\Delta_4^{-1/2} = (1 - 4\theta^2)^{-1/2} = 1 + 2\theta^2 + 6\theta^4 + \dots$$

55. (Разд. 6.7.7) Покажите, что

$$\left(\frac{1 + \sqrt{1 - 4\theta^2}}{2} \right)^{-2} = 1 + 2\theta^2 + 5\theta^4 + \dots$$

56. (Разд. 6.7.7) Убедитесь в справедливости (114) для $T = 3$ непосредственным подсчетом.

57. (Разд. 6.7.7) (а) Покажите, что

$$\varphi(2) = 1 + \theta^2 + \dots,$$

$$\varphi(4) = 1 + \dots,$$

$$\varphi^*(3) = 1 - \theta + \theta^2 - \dots$$

(б) Проверьте (116) и (117) для $T = 3$.

58. (Разд. 6.7.7) Выведите в циклическом случае выражения для $\mathcal{Q}_1^{*h}$, $h = 1, 2, 3, 4$, из соотношения $Q_1 = Q_1^* + z_T^2$.

59. (Разд. 6.7.7) Выразите ψ_5 , ψ_6 , ψ_7 и ψ_8 с помощью $v_1, \dots, v_8$, не предполагая выполнения условия $v_{2l-1} = 0$.

60. (Разд. 6.7.7) Для случая, когда Q_1^* основывается на сумме квадратов последовательных разностей, покажите, используя (129), что $v_k = 0$, $k = 1, 2, \dots$, для $T = 1$ и 2, что $v_{2l} = 1/(2l)$, $l = 1, 2, \dots$, для $T = 3$ и что $v_{2l} = 2^{l-1}/l$, $l = 1, 2, \dots$, для $T = 4$.

61. (Разд. 6.7.7) Для случая, когда Q_1^* основывается на сумме квадратов последовательных разностей, покажите, что $\psi_k = 0$, $k = 1, 2, \dots$, для $T = 1$ и 2, что $\psi_3 = 5/16$ и $\psi_8 = 35/128$ для $T = 3$ и что $\psi_8 = 35/8$ для $T = 4$. Покажите также, что $\mathcal{Q}_1^{*6} = 5/1024$ и $\mathcal{Q}_1^{*8} = 35/32768$ для $T = 3$ и что $\mathcal{Q}_1^{*8} = 112/21879$ для $T = 4$.

62. (Разд. 6.7.7) Найдите v_{10} , ψ_{10} , $\mathcal{Q}_1^{*10}$ и $\mathcal{Q}_1^{*10}$ для модели с последовательными разностями.

63. (Разд. 6.7.7) Для Q_1^* в случае циклического сериального коэффициента корреляции покажите, что $v_k = 0$, $k = 1, 2, \dots$, для $T = 1$, $v_k = (-1)^k 2^{k-1}/k$ для $T = 2$, $v_k = (-1)^k/k$ для $T = 3$ и $v_k = (-1)^k 2^{k-1}/k$ для $T = 4$. (Заметьте, что для $T = 2$ и $T = 4$ соответствующие моменты совпадают.)

64. (Разд. 6.7.7) Для Q_1^* в случае циклического сериального коэффициента корреляции покажите, что $\psi_k = 0$, $k = 1, 2, \dots$, для $T = 1$, $\psi_2 = 3/2$, $\psi_3 = -5/2$ и $\psi_4 = 35/8$ для $T = 2$, $\psi_3 = -1$, $\psi_4 = 1$ для $T = 3$ и $\psi_4 = 35/8$ для $T = 4$. Покажите, что $\mathfrak{g}_{r_1}^{*2} = 1$, $\mathfrak{g}_{r_1}^{*3} = -1$ и $\mathfrak{g}_{r_1}^{*4} = 1$ для $T = 2$, $\mathfrak{g}_{r_1}^{*3} = -1/8$ и $\mathfrak{g}_{r_1}^{*4} = 1/16$ для $T = 3$ и $\mathfrak{g}_{r_1}^{*4} = 1/9$ для $T = 4$.

65. (Разд. 6.7.7) Найдите v_5 , ψ_5 , $\mathfrak{g}_{r_1}^{*5}$ и $\mathfrak{g}_{r_1}^{*5}$ для циклического коэффициента.

66. (Разд. 6.7.7) Пусть $\lambda_t^* = \cos \pi t/T^*$, $t = 1, \dots, T^* - 1$, являются характеристическими корнями, связанными с коэффициентом r_1^* в модели с последовательными разностями, а $\lambda_t = \cos \pi t/T$, $t = 1, \dots, T - 1$, характеристическими корнями, связанными с циклическим коэффициентом r_1^* .

(а) Покажите, что если $T = 2T^*$, то

$$\sum_{t=1}^{T-1} \lambda_t^k = 2 \sum_{t=1}^{T^*-1} \lambda_t^{*k} + (-1)^k, \quad k = 0, 1, \dots$$

(б) Покажите, что если $T = T^*$ нечетное, то

$$\sum_{t=1}^{T-1} \lambda_t^{2l} = \sum_{t=1}^{T^*-1} \lambda_t^{*2l}, \quad l = 0, 1, \dots$$

67. (Разд. 6.7.7) Интегрируя выражение

$$(1-x)^{-1} = \sum_{j=1}^{\infty} x^{j-1},$$

покажите, что

$$-\log(1-x) = \sum_{j=1}^{\infty} \frac{x^j}{j}.$$

68. (Разд. 6.7.7) Пусть случайные величины $x_1, y_1, \dots, x_N, y_N$ независимы, и нормально распределены с нулевыми средними и единичными дисперсиями.

Пусть $r = A/\sqrt{BC}$, где $A = \sum_{i=1}^N x_i y_i$, $B = \sum_{i=1}^N x_i^2$ и $C = \sum_{i=1}^N y_i^2$. (а) Покажите, что распределение r не зависит от B и C . (б) Покажите, что B и C независимы и имеют χ^2 -распределение с N степенями свободы. (с) Покажите, что производящая функция моментов A равна

$$(I) \quad \mathfrak{g}_e^{\theta A} = (1 - \theta^2)^{-N/2} = \sum_{h=0}^{\infty} \frac{\theta^{2h} \Gamma\left(\frac{1}{2}N + h\right)}{h! \Gamma\left(\frac{1}{2}N\right)}.$$

(д) Покажите, что

$$(II) \quad \mathfrak{g}_A^{2k-1} = 0, \quad k = 1, 2, \dots,$$

$$\begin{aligned}
 \text{(iii)} \quad g_A^{2k} &= \frac{(2k)! \Gamma\left(\frac{1}{2}N + k\right)}{k! \Gamma\left(\frac{1}{2}N\right)} = \\
 &= \frac{2^{2k} \Gamma\left(k + \frac{1}{2}\right) \Gamma\left(\frac{1}{2}N + k\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{2}N\right)}, \quad k = 0, 1, \dots,
 \end{aligned}$$

$$\text{(iv)} \quad g_r^{2k-1} = 0, \quad k = 1, 2, \dots,$$

$$\text{(v)} \quad g_r^{2k} = \frac{\Gamma\left(k + \frac{1}{2}\right) \Gamma\left(\frac{1}{2}N\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{2}N + k\right)}, \quad k = 0, 1, \dots.$$

69. (Разд. 6.8.1) Покажите, что

$$\begin{aligned}
 \int_0^{2\pi} \log(a + b \cos x) dx &= 2 \int_0^{\pi} \log(a + b \cos x) dx = \\
 &= 2\pi \log \frac{1}{2} (a + \sqrt{a^2 - b^2}), \quad a > b \geq 0.
 \end{aligned}$$

(Указание. Дифференцированием по a убедитесь в справедливости второго равенства. Получающийся при этом в левой части интеграл можно вычислить, представляя $\cos x = \cos^2 \frac{1}{2}x - \sin^2 \frac{1}{2}x$ и полагая затем $y = (a - b) \operatorname{tg} \frac{1}{2}x / \sqrt{a^2 - b^2}$.)

70. (Разд. 6.8.1) Докажите, что

$$\frac{2\pi}{T} \sum_{i=1}^T (e^{i2\pi i/T} + e^{-i2\pi i/T})^k = \int_0^{2\pi} (e^{iu} + e^{-iu})^k du, \quad k = 0, 1, \dots, T-1.$$

71. (Разд. 6.8.1) Покажите, что вплоть до порядка $T-1$ семинварианты Q_1 для циклического случая совпадают с соответствующими семинвариантами случайной величины, имеющей производящую функцию моментов (20), сравнивая разложения (18) и (19) в ряды по степеням θ . (Указание. Использовать упр. 70.)

72. (Разд. 6.8.1) Покажите, что если X — случайная величина с нулевым средним, дисперсией σ^2 , значения которой с вероятностью 1 сосредоточены в интервале (a, b) , то $\sigma^2 < -ab$. (Указание. Показать, что

$$\int_a^b (x-a)(b-x)f(x)dx > 0,$$

где $f(x)$ — плотность распределения случайной величины X .)

73. (Разд. 6.8.1) Покажите, что нечетные моменты r_1^* до порядка $T-2$ включительно и четные моменты до порядка $T-1$ включительно совпадают с соот-

ветствующими моментами функции

$$\frac{\Gamma\left(\frac{1}{2}T + \frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}T\right)\Gamma\left(\frac{1}{2}\right)} (1 - r_1^{*2})^{(T-2)/2} \left(1 - \frac{T+1}{T-1} r_1^*\right), \quad -1 \leq r_1^* \leq 1.$$

(Заметьте, что эта функция не является плотностью, поскольку она отрицательна при $1 - 2/(T+1) < r_1^* \leq 1$.)

74. (Разд. 6.9.2) Покажите, что

$$r_3' = \frac{r_3^* - r_1^{*2} r_3^* + r_1^{*3} - 2r_1^* r_2^* + r_1^* r_2^{*2}}{1 - 2r_1^{*2} + 2r_1^{*2} r_2^* - r_2^{*2}}.$$

75. (§ 6.10) Покажите, что производная вероятности, дополнительной (до 1) к (14), по переменной R^* дается выражением (9) с $g_1^*(r_1^* | 0)$, задаваемой выражением (53) § 6.7, с заменой r и r_1^* на R^* и заменой ρ_1 на ρ .

76. (Разд. 6.11.1) Покажите, что матрица квадратичной формы в показателе экспоненты равна $-\frac{1}{2} \Sigma^{-1} = -1/(2\sigma^2) \mathbf{T}'\mathbf{T}$, где

$$\mathbf{T} = \begin{bmatrix} \sqrt{1-\beta^2} & 0 & 0 & \dots & 0 & 0 \\ \beta & 1 & 0 & \dots & 0 & 0 \\ 0 & \beta & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & \beta & 1 \end{bmatrix}.$$

77. (Разд. 6.11.1) Убедитесь в том, что решения кубического уравнения (10) не являются рациональными функциями полиномов P_0' , P_1 и P_0 .

78. (Разд. 6.11.1) Покажите, что

$$\hat{\beta} = -\left(1 - \frac{1}{T}\right) \frac{P_1}{P_0}$$

является решением уравнения (10) с точностью до $o(1/T)$. (Указание. Положить в (10) $\hat{\beta} = -P_1/P_0 + v/T$. Отбросить в полученном уравнении члены, имеющие порядок малости выше $1/T$, и решить его относительно v . Заметьте, что при $T \rightarrow \infty$ отношение P_0'/P_0 имеет (по вероятности) порядок $1/T$.)

79. (Разд. 6.11.1) Получите уравнения для оценок максимального правдоподобия, когда $\xi y_t = \mu$.

80. (Разд. 6.11.1) Покажите для случая стационарного гауссовского процесса первого порядка ($\xi y_t = \mu$, $\xi(y_t - \mu)(y_s - \mu) = \sigma(0)\rho^{|t-s|}$), что оценкой максимального правдоподобия для μ при известном ρ ($0 < \rho < 1$) является

$$\hat{\mu} = \frac{y_1 + (1-\rho) \sum_{t=2}^{T-1} y_t + y_T}{(T-2)(1-\rho) + 2},$$

и проверьте, что ее дисперсия равна

$$\text{Var } \hat{\mu} = \sigma(0) \frac{1 + \rho}{(T-2)(1-\rho) + 2}.$$

(Указание. Использовать результаты § 2.4.)

81. (Разд. 6.11.1) Найдите дисперсию величины $\bar{y}$ для стационарного гауссовского процесса первого порядка и сравните ее с дисперсией для $\hat{\mu}$.

82. (Разд. 6.11.2) Пусть у стационарного гауссовского процесса первого порядка $\rho = e^{-k/T}$ ($0 < k$). Покажите, что при $T \rightarrow \infty$

$$\lim_{T \rightarrow \infty} \text{Var } \hat{\mu} = \sigma(0) \frac{2}{2+k},$$

$$\lim_{T \rightarrow \infty} \text{Var } \bar{y} = \sigma(0) 2 \frac{e^{-k} - 1 + k}{k^2}.$$

Используя эти выражения, убедитесь в том, что

$$\lim_{T \rightarrow \infty} \text{Var } \hat{\mu} < \lim_{T \rightarrow \infty} \text{Var } \bar{y} < \sigma(0).$$

83. (Разд. 6.11.2) Проверьте, что якобиан преобразования (12) действительно равен $1 - (-\beta)^T$.

84. (Разд. 6.11.2) Существует ли решение $(-\hat{\beta})$ уравнения (20), являющееся монотонно возрастающей функцией от g ?

85. (Разд. 6.11.2) Пусть $(-\hat{\beta}_T)$ — корень уравнения (20), расположенный между -1 и 1 , и пусть $r = r_T$. Покажите, что $T^k (r_T + \hat{\beta}_T) \rightarrow 0$ по вероятности при $T \rightarrow \infty$ для любого $k > 0$. (Указание. Покажите, что для достаточно больших T со сколь угодно большой вероятностью выполняются неравенства $|r_T + \hat{\beta}_T| < \varepsilon$ и $a < -\hat{\beta}_T < b$ для любых заранее заданных значений ε , a и b , $-1 < a < b < 1$. Затем покажите, что $T^k (-\hat{\beta}_T)^{T-1} \rightarrow 0$ по вероятности.)

86. (Разд. 6.11.2) Убедитесь в том, что (21) соответствует критерию отношения правдоподобия для проверки гипотезы $\beta = 0$.

87. (Разд. 6.11.2) Соответствует ли указанный критерий отношения правдоподобия отвержению гипотезы $\beta = 0$ при $r > r_1$ и $r < r_2$?

88. (Разд. 6.11.2) Найдите уравнения максимального правдоподобия, когда $\xi_{y_i} = \mu$.

89. (Разд. 6.11.2) Найдите уравнения максимального правдоподобия для $\hat{\sigma}^2$ и $\hat{\beta}$ в модели с последовательными разностями, т. е. в модели разд. 6.5.3 с $\gamma_0 = (1 + \beta^2)/\sigma^2$ и $\gamma_1 = 2\beta/\sigma^2$. (Указание. Использовать лемму 6.7.11 с $\theta = -\beta/(1 + \beta^2)$ для вычисления определителя в соответствующем распределении.)

90. (Разд. 6.11.2) Выпишите уравнения максимального правдоподобия для оценок параметров σ^2 , β и μ модели с последовательными разностями, когда $\gamma_0 = (1 + \beta^2)/\sigma^2$, $\gamma_1 = 2\beta/\sigma^2$ и $\xi_{y_i} = \mu$.

СТАЦИОНАРНЫЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ

7.1. ВВЕДЕНИЕ

В статистическом анализе последовательность T наблюдений, образующих временной ряд, часто рассматривают как выборку T последовательных наблюдений через равные промежутки времени из существенно более продолжительной (генеральной) последовательности случайных величин. При этом статистические выводы делаются относительно вероятностной структуры генеральной последовательности. Такую последовательность удобно считать простирающейся неограниченно в будущее и, возможно, в прошлое. Последовательность случайных величин $y_1, y_2, \dots$ или $\dots, y_{-1}, y_0, y_1, \dots$ называется *случайным процессом с дискретным параметром времени*.

Несмотря на полную произвольность вероятностных моделей последовательностей случайных величин, полезно отличать случайные процессы от множества случайных величин этого процесса, учитывая понятие времени. Грубо говоря, в случайном процессе наблюдения, разделенные небольшими промежутками времени, близки по значениям в отличие от наблюдений, далеко отстоящих друг от друга по времени. Более того, модель значительно упрощается после расширения конечной последовательности наблюдений до бесконечной.

Одним из таких упрощений является свойство *стационарности*. Будем считать, что поведение множества случайных величин с вероятностной точки зрения не зависит от времени. Это свойство обсуждалось выше при рассмотрении процессов скользящего среднего или процессов, порождаемых стохастическими разностными уравнениями.

В гл. 5 нас интересовали процессы, описываемые конечным числом параметров. Теперь мы будем рассматривать процессы, в кото-

рых число параметров может быть бесконечным. Например, гауссовский (нормальный) стационарный процесс имеет конкретные среднее значение и дисперсию, но бесконечное число коэффициентов ковариации. Поставим вопрос: какую информацию о них можно получить исходя из конечного числа наблюдений?

Случайный процесс, как бесконечная последовательность случайных величин, связан с понятием вероятностной меры, определенной в бесконечномерном пространстве. Некоторые интересные вопросы, в частности о пределах функций последовательностей, касаются вероятностей множеств в бесконечномерном пространстве. Но, поскольку мы интересуемся выводами из конечного (хотя, возможно, большого) числа наблюдений, вероятности таких множеств не понадобятся. В этой главе математическая теория случайных процессов обсуждается лишь на уровне идей, поскольку детальное ее изучение на этом этапе не требуется.

Случайный процесс $y(t)$ с непрерывным параметром времени можно определить для $0 \leq t < \infty$ или $-\infty < t < \infty$ и рассматривать с привлечением вероятностной меры на пространстве функций $y(t)$. Выборка из такого процесса состоит из наблюдений в конечном числе точек времени, или из непрерывных наблюдений в интервале времени. Пусть, например, процесс представляет собой теоретическую модель наблюдений температуры в некоторой местности. Тогда последовательность ежечасных отсчетов или диаграмма непрерывных показаний является выборкой из случайного процесса. Случайный процесс с дискретным временем часто понимают как выборку через равные промежутки времени из случайного процесса с непрерывным временем.

7.2. СТАЦИОНАРНЫЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ, ОПРЕДЕЛЕНИЯ И ПРИМЕРЫ

7.2.1. Определения

Случайным процессом с дискретным временем называется бесконечная последовательность случайных величин $y_1, y_2, \dots$ или $\dots, y_{-1}, y_0, y_1, \dots$. Наблюдение процесса, часто называемое *реализацией*, есть точка в соответствующем бесконечномерном пространстве, где определена вероятностная мера. Вероятность определяется на некоторых множествах, называемых измеримыми. Этот класс множеств включает вместе с любым множеством его дополнение, а также объединение и пересечение счетного числа множеств этого класса; вероятностная мера на этом классе множеств определяется таким образом, что вероятность объединения непересекающихся множеств равна сумме вероятностей отдельных множеств.

Практически мы интересуемся вероятностями, которые связаны с конечным числом случайных величин. Эти вероятности включают в себя функцию совместного распределения $\text{Pr} \{y_1 \leq b_1, \dots, y_t \leq b_t\}$, а также вероятности, полученные из нее. Потребуем, чтобы в класс измеримых множеств входили множества $\{y_1 \leq b_1, \dots, y_t \leq b_t\}$. Множества, которые определяются в терминах конечного числа координат $y_{i_1}, \dots, y_{i_n}$, называются *цилиндрическими множествами*. Таким образом, вероятностная мера случайного процесса всегда определяет вероятности цилиндрических (измеримых) множеств, в частности считаются известными функции совместного распределения.

Наоборот, пусть заданы конечномерные функции совместного распределения (ф.с.р.), тогда на бесконечномерном пространстве вероятностную меру можно определить таким образом, что вероятностная мера цилиндрических множеств согласуется с ф.с.р. в том смысле, что может быть вычислена по соответствующим ф.с.р. Это расширение единственно. *Согласованность* семейства ф.с.р. означает: (i) любая маргинальная (совместная) ф.с.р., полученная из заданной ф.с.р., приводит к соответствующей ф.с.р. меньшего числа переменных, (ii) ф.с.р., соответствующие наборам точек параметра t , написанным в различном порядке, согласованны. Точнее, пусть $F(a_1, \dots, a_n; t_1, \dots, t_n)$ означает ф.с.р. от $a_1, \dots, a_n$, любых $t_1, \dots, t_n$ (из множества $\dots, -1, 0, 1, \dots$ или $1, 2, \dots$) и $n = 1, 2, \dots$. Потребуем выполнения равенств

$$(1) \quad F(a_1, \dots, a_m, \infty, \dots, \infty; t_1, \dots, t_m, t_{m+1}, \dots, t_n) = \\ = F(a_1, \dots, a_m; t_1, \dots, t_m), \quad m = 1, \dots, n,$$

$$(2) \quad F(a_1, \dots, a_n; t_1, \dots, t_n) = F(a_{i_1}, \dots, a_{i_n}; t_{i_1}, \dots, t_{i_n}),$$

где $(i_1, \dots, i_n)$ есть произвольная перестановка чисел $(1, \dots, n)$. Согласованное семейство ф.с.р. определяет случайный процесс y_t в том смысле, что $F(a_1, \dots, a_n; t_1, \dots, t_n) = \text{Pr} \{y_{t_1} \leq a_1, \dots, y_{t_n} \leq a_n\}$ для всех a_i, t_i и n . Это утверждение принадлежит Колмогорову (1933), доказательства его мы не приводим.

В качестве примера зададим ф.с.р. как ф.с.р. многомерных нормальных распределений, определив средние значения и дисперсии следующим образом:

$$(3) \quad \mathbb{E} y_t = m(t),$$

$$(4) \quad \mathbb{E} [y_t - m(t)] [y_s - m(s)] = \text{Cov}(y_t, y_s) = \sigma(t, s).$$

Последовательность $m(t)$ произвольна. Двойная последовательность $\sigma(t, s)$ должна удовлетворять обычным свойствам ковариаций, а именно $\sigma(t, s) = \sigma(s, t)$ для каждой пары s, t и любая матрица

$[\sigma(t_i, t_j)], i, j = 1, \dots, n$, положительно полуопределена (или, эквивалентно, неотрицательно определена.— *Перев.*). Эти условия определяют случайный процесс. Такой процесс называется *гауссовским*. Случайный процесс с дискретным временем называется *стационарным* (или *стационарным в узком смысле*), если распределение величин $y_t, \dots, y_{t_n}$ совпадает с распределением $y_{t_1+t}, \dots, y_{t_n+t}$ для любого конечного множества целых чисел $\{t_1, \dots, t_n\}$ и любого целого t . Это определение эквивалентно требованию того, что вероятностные меры последовательностей $\{y_s\}$ и $\{y_{s+t}\}$ совпадают для любого целого t .

Пусть существуют первые моменты, тогда из стационарности следует

$$(5) \quad \mathbb{E}y_s = \mathbb{E}y_{s+t}, \quad s, t = \dots, -1, 0, 1, \dots,$$

или

$$(6) \quad m(s) = m(s+t) = m$$

для любых s, t . Поскольку распределения (y_t, y_{t_1}) и (y_{t_1+t}, y_{t_1+t}) совпадают, то из существования моментов второго порядка и стационарности следует

$$(7) \quad \sigma(t_1, t_2) = \sigma(t_1+t, t_2+t).$$

Подставляя $t = -t_2$, получаем

$$(8) \quad \sigma(t_1, t_2) = \sigma(t_1 - t_2, 0) = \sigma(t_1, -t_2).$$

В случае нормальных ф. с. р. свойства (5) и (7) эквивалентны стационарности случайного процесса.

Случайный процесс называется *стационарным в широком смысле*, или *слабо стационарным*, или *стационарным второго порядка*, если функция средних $m(t)$ и ковариационная функция $\sigma(t, s)$, определенные в (3) и (4), существуют и удовлетворяют соотношениям (5) и (7), т. е. средние постоянны и не зависят от времени, а ковариация любых двух значений зависит только от их расстояния во времени. Очевидно, любой процесс, стационарный в узком смысле и имеющий конечную дисперсию, является также стационарным и в широком смысле. (В рассмотренном выше нормальном случае стационарности в узком и широком смысле эквивалентны.)

7.2.2. Примеры стационарных случайных процессов с дискретным параметром

Рассмотрим несколько примеров стационарных случайных процессов.

Пример 7.1. Пусть y_t независимы и одинаково распределены с

$$(9) \quad \mathbb{E}y_t = m, \quad \text{Var } y_t = \sigma^2.$$

Тогда

$$(10) \quad \sigma(t, s) = \begin{cases} \sigma^2 & \text{при } t = s, \\ 0 & \text{при } t \neq s. \end{cases}$$

Этот процесс стационарен в узком смысле. Если отбросить условие одинаковой распределенности величин y_t (в то время как (9) и (10) сохраняются), то результирующий процесс будет все же стационарен в широком смысле.

Пример 7.2. Пусть все y_t тождественно равны случайной величине y . Тогда, если существуют два первых момента y ,

$$(11) \quad \mathbb{E}y_t = \mathbb{E}y = m,$$

$$(12) \quad \sigma(t, s) = \text{Var } y = \sigma^2.$$

Этот процесс стационарен в узком смысле.

Пример 7.3. Определим последовательность $\{y_t\}$ следующим образом:

$$(13) \quad y_t = \sum_{j=1}^q (A_j \cos \lambda_j t + B_j \sin \lambda_j t), \quad t = \dots, -1, 0, 1, \dots,$$

где λ_j — постоянные, а $A_1, \dots, A_q, B_1, \dots, B_q$ — случайные величины, такие, что

$$(14) \quad \mathbb{E}A_j = \mathbb{E}B_j = 0, \quad j = 1, \dots, q,$$

$$(15) \quad \mathbb{E}A_j^2 = \mathbb{E}B_j^2 = \sigma_j^2, \quad j = 1, \dots, q,$$

$$(16) \quad \mathbb{E}A_i A_j = \mathbb{E}B_i B_j = 0, \quad i \neq j, \quad i, j = 1, \dots, q,$$

$$(17) \quad \mathbb{E}A_i B_j = 0, \quad i, j = 1, \dots, q.$$

Тогда

$$(18) \quad \mathbb{E}y_t = 0,$$

$$\begin{aligned} (19) \quad \mathbb{E}y_t y_s &= \sum_{i,j=1}^q \mathbb{E} (A_i \cos \lambda_i t + B_i \sin \lambda_i t) (A_j \cos \lambda_j s + B_j \sin \lambda_j s) = \\ &= \sum_{j=1}^q [\mathbb{E}A_j^2 \cos \lambda_j t \cos \lambda_j s + \mathbb{E}B_j^2 \sin \lambda_j t \sin \lambda_j s] = \\ &= \sum_{j=1}^q \sigma_j^2 [\cos \lambda_j t \cos \lambda_j s + \sin \lambda_j t \sin \lambda_j s] = \\ &= \sum_{j=1}^q \sigma_j^2 \cos \lambda_j (t - s). \end{aligned}$$

Следовательно, процесс стационарен в широком смысле. Если A_j и B_j нормально распределены, то y_t будут также распределены нор-

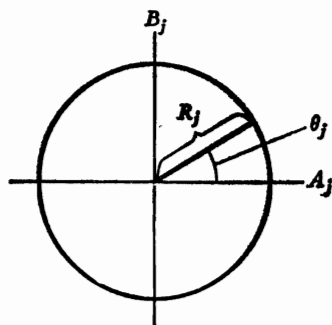


Рис. 7.1.

Определение параметров θ_j и R_j .

мально, так как они являются линейными комбинациями величин A_j и B_j . В этом случае процесс будет стационарным в узком смысле.

Этот процесс можно также представить следующим образом:

$$(20) \quad y_t = \sum_{j=1}^q R_j \cos(\lambda_j t - \theta_j), \quad t = \dots, -1, 0, 1, \dots,$$

где (как на рис. 7.1) $R_j \geq 0$,

$$(21) \quad R_j^2 = A_j^2 + B_j^2, \quad j = 1, \dots, q,$$

$$(22) \quad \operatorname{tg} \theta_j = B_j/A_j, \quad j = 1, \dots, q,$$

$0 < \theta_j < \pi$, если $B_j > 0$, и $\pi < \theta_j < 2\pi$, если $B_j < 0$.

Если A_j и B_j распределены нормально, то R_j^2 пропорционально χ^2 с двумя степенями свободы и θ_j равномерно распределено между 0 и 2π (ввиду симметрии нормального распределения) и не зависит от R_j^2 . Для доказательства стационарности процесса заметим, что

$$(23) \quad y_{t+s} = \sum_{j=1}^q R_j \cos[\lambda_j t - (\theta_j - \lambda_j s)],$$

где $(\theta_j - \lambda_j s)$ равномерно распределено на интервале длины 2π . Таким образом, y_t и y_{t+s} имеют одно и то же распределение. Если A_j и B_j имеют распределение, отличное от нормального, то последовательность $\{y_t\}$ не обязательно стационарна в строгом смысле.

Этот пример важен потому, что в известном смысле каждый слабо стационарный случайный процесс с конечной дисперсией можно аппроксимировать линейными комбинациями, аналогичными правой части (13).

Пример 7.4. Процесс скользящего среднего.

Пусть $\dots, v_{-1}, v_0, v_1, \dots$ — последовательность независимых и одинаково распределенных случайных величин, а $\alpha_0, \alpha_1, \dots, \alpha_q$ —

последовательность $q + 1$ чисел. Определим процесс скользящего среднего

$$(24) \quad y_t = \alpha_0 v_t + \alpha_1 v_{t-1} + \dots + \alpha_q v_{t-q}, \quad t = \dots, -1, 0, 1, \dots$$

Это стационарный случайный процесс. Если $\mathbb{E} v_t = v$ и $\text{Var } v_t = \sigma^2$, то

$$(25) \quad \mathbb{E} y_t = v(\alpha_0 + \alpha_1 + \dots + \alpha_q),$$

$$(26) \quad \text{Cov}(y_t, y_{t+s}) = \begin{cases} \sigma^2(\alpha_0 \alpha_s + \dots + \alpha_{q-s} \alpha_q), & s = 0, \dots, q, \\ 0, & s = q+1, \dots, \end{cases}$$

и последовательность $\{y_t\}$ стационарна в широком смысле. Для того чтобы последовательность $\{y_t\}$, определенная формулой (24), была стационарна в широком смысле, необходимо только, чтобы величины v_t имели одни и те же среднее значение и дисперсию и были взаимно не коррелированы.

Мы можем также определить процесс скользящего среднего с бесконечными пределами суммирования. Бесконечная сумма

$$(27) \quad \sum_{s=0}^{\infty} \alpha_s v_{t-s}$$

означает случайную величину y_t , если она существует, такую, что

$$(28) \quad \lim_{n \rightarrow \infty} \mathbb{E} \left(y_t - \sum_{s=0}^n \alpha_s v_{t-s} \right)^2 = 0.$$

Если выполнено

$$(29) \quad \sum_{s=0}^{\infty} \alpha_s^2 < \infty,$$

такая случайная величина существует (следствие 7.6.1). Говорят, что сумма (27) сходится в среднем или среднеквадратичном. Бесконечная сумма

$$(30) \quad \sum_{s=-\infty}^{\infty} \alpha_s v_{t-s}$$

определяется как случайная величина y_t , такая, что

$$(31) \quad \lim_{n, m \rightarrow \infty} \mathbb{E} \left(y_t - \sum_{s=-m}^n \alpha_s v_{t-s} \right)^2 = 0.$$

Достаточным условием существования случайной величины y_t является

$$(32) \quad \sum_{s=-\infty}^{\infty} \alpha_s^2 < \infty.$$

Условия (29) и (32) являются достаточными, если v_s не коррелированы и имеют общие среднее и дисперсию. Но, вообще говоря, эти случайные величины могут быть зависимыми.

Пример 7.5. Процесс авторегрессии. В гл. 5 мы рассматривали процесс, удовлетворяющий стохастическому разностному уравнению

$$(33) \quad \sum_{r=0}^p \beta_r y_{t-r} = u_t, \quad t = \dots, -1, 0, 1, \dots,$$

где u_t независимы и одинаково распределены со средним нулевым значением. Если корни связанного с данным процессом характеристического уравнения

$$(34) \quad \sum_{r=0}^p \beta_r x^{p-r} = 0$$

меньше 1 по абсолютной величине, то (33) можно переписать следующим образом:

$$(35) \quad y_t = \sum_{s=0}^{\infty} \delta_s u_{t-s}, \quad t = \dots, -1, 0, 1, \dots,$$

где правая часть сходится в среднем. Таким образом, последовательность $\{y_t\}$ стационарна в узком смысле. Если u_t не коррелированы и имеют общие среднее значение и дисперсию (u_t могут быть зависимыми и неодинаково распределенными), то процесс стационарен в широком смысле. В § 5.2 было показано, что ковариационная последовательность удовлетворяет разностному уравнению

$$(36) \quad \sum_{r=0}^p \beta_r \sigma(s-r) = 0, \quad s = 1, 2, \dots$$

Если корни характеристического уравнения (34) различны, скажем $x_1, \dots, x_p$, то

$$(37) \quad \sigma(h) = \sum_{j=1}^p c_j x_j^h, \quad h = 1-p, 2-p, \dots, -1, 0, 1, \dots,$$

и постоянные $c_1, \dots, c_p$ определяются из условий

$$(38) \quad \sigma(-h) = \sigma(h), \quad h = 1, \dots, p-1,$$

($\beta_p \neq 0$) и

$$(39) \quad \sum_{r=0}^p \beta_r \sigma(r) = \sigma^2,$$

где $\mathcal{E}u_t^2 = \sigma^2$.

Пример 7.6. Процесс авторегрессии со скользящим усреднением остатков. Две предыдущие модели могут быть объединены, т. е. в качестве возмущения u_t стохастического разностного уравнения

можно взять процесс скользящего среднего. Таким образом,

$$(40) \quad \sum_{r=0}^p \beta_r y_{t-r} = \sum_{j=0}^q \alpha_j v_{t-j}, \quad t = \dots, -1, 0, 1, \dots,$$

где v_t независимы и одинаково распределены. Тогда процесс $\{y_t\}$ стационарен в узком смысле. Если v_t не коррелированы и имеют общие среднее значение и дисперсию, то $\{y_t\}$ стационарен в широком смысле.

Пример 7.7. Процесс авторегрессии с ошибкой. Пусть y_t удовлетворяет стохастическому разностному уравнению (33); определим

$$(41) \quad z_t = y_t + w_t,$$

где w_t независимы, одинаково распределены и не зависят от y_t . Если процесс $\{y_t\}$ стационарен в узком смысле, то и $\{z_t\}$ стационарен в узком смысле. Если $\{y_t\}$ стационарен в широком смысле и w_t не коррелированы, имея общие среднее и дисперсию, то процесс $\{z_t\}$ также стационарен в широком смысле.

Пример 7.8. Это пример процесса, стационарного в широком, но не стационарного в узком смысле. Пусть

$$(42) \quad y_t = \cos t\omega, \quad t = 1, 2, \dots,$$

где ω равномерно распределено на интервале $(0, 2\pi)$. Тогда $\mathcal{E}y_t = 0$, $\mathcal{E}y_t^2 = \frac{1}{2}$ и $\mathcal{E}y_t y_s = 0$, $t \neq s$. Эти случайные величины некоррелированы, несмотря на функциональную и статистическую зависимость.

Пример 7.9. Рассмотрим еще один стационарный в широком смысле процесс, связанный с примером 7.3,

$$(43) \quad y_t = \sum_{j=1}^q \sigma_j \sqrt{2} \cos(\lambda_j t - v_j),$$

где $v_1, \dots, v_q$ независимы и равномерно распределены на интервале $(0, 2\pi)$. Тогда $\mathcal{E}y_t = 0$ и

$$(44) \quad \begin{aligned} \mathcal{E}y_t y_s &= \sum_{i,j=1}^q 2\sigma_i \sigma_j \mathcal{E} \cos(\lambda_i t - v_i) \cos(\lambda_j s - v_j) = \\ &= 2 \sum_{j=1}^q \sigma_j^2 \int_0^{2\pi} \cos(\lambda_j t - v_j) \cos(\lambda_j s - v_j) dv_j / (2\pi) = \end{aligned}$$

$$\begin{aligned}
&= 2 \sum_{j=1}^q \sigma_j^2 \int_0^{2\pi} (\cos \lambda_j t \cos v_j + \sin \lambda_j t \sin v_j) \times \\
&\quad \times (\cos \lambda_j s \cos v_j + \sin \lambda_j s \sin v_j) d v_j / (2\pi) = \\
&= \sum_{j=1}^q \sigma_j^2 (\cos \lambda_j t \cos \lambda_j s + \sin \lambda_j t \sin \lambda_j s) = \sum_{j=1}^q \sigma_j^2 \cos \lambda_j (t - s),
\end{aligned}$$

что совпадает с (19).

7.2.3. Случайный процесс с непрерывным параметром

Определение случайного процесса $y(t)$ с непрерывным параметром, где t — действительное число на некотором конечном или бесконечном интервале $0 \leq t < \infty$, или $-\infty < t < \infty$, основано на задании вероятностной меры в пространстве функций от t . Вероятностная мера определяет вероятности цилиндрических множеств, в частности функцию совместного распределения (ф.с.р.) каждого конечного множества переменных величин $y(t_1), \dots, y(t_n)$, где $t_1, \dots, t_n$ — любые n действительных чисел в области значений t . Обратно, совокупность согласованно определенных ф.с.р. $F(a_1, \dots, a_n; t_1, \dots, t_n)$ для каждого действительного набора $t_1, \dots, t_n$ в области значений t и каждого n может быть однозначно продолжена на вероятностную меру функционального пространства.

Определение стационарности в узком смысле и стационарности в широком смысле аналогично определению этих понятий для процессов с дискретным параметром.

7.3. СПЕКТРАЛЬНАЯ ПЛОТНОСТЬ И СПЕКТРАЛЬНАЯ ФУНКЦИЯ

Стационарный в широком смысле случайный процесс с дискретным временем определяет последовательность ковариаций $\sigma(0)$, $\sigma(1)$, ... Преобразование Фурье этой последовательности есть

$$\begin{aligned}
(1) \quad f(\lambda) &= \frac{1}{2\pi} \sigma(0) + \frac{1}{\pi} \sum_{h=1}^{\infty} \sigma(h) \cos \lambda h = \\
&= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \sigma(h) \cos \lambda h, \quad -\pi \leq \lambda \leq \pi.
\end{aligned}$$

Особый интерес представляет тот случай, когда этот ряд сходится. Если

$$(2) \quad \sigma(0) + 2 \sum_{h=1}^{\infty} |\sigma(h)| = \sum_{h=-\infty}^{\infty} |\sigma(h)| < \infty,$$

то (1) сходится равномерно. Умножая (1) на $\cos \lambda k$ и интегрируя, получаем

$$(3) \quad \sigma(k) = \int_{-\pi}^{\pi} \cos \lambda k f(\lambda) d\lambda.$$

(О рядах Фурье см. разд. 4.2.4.) Таким образом, ковариационная последовательность определяет функцию $f(\lambda)$ и обратно. Функция $f(\lambda)$ (если она существует) называется *спектральной плотностью*. Заметим, что $f(\lambda)$ — четная функция, т. е. $f(\lambda) = f(-\lambda)$. Иногда удобно ввести функцию $g(\lambda) = 2f(\lambda)$, $0 \leq \lambda \leq \pi$, т. е.

$$(4) \quad g(\lambda) = \frac{1}{\pi} \sigma(0) + \frac{2}{\pi} \sum_{h=1}^{\infty} \sigma(h) \cos \lambda h = \\ = \frac{1}{\pi} \sum_{h=-\infty}^{\infty} \sigma(h) \cos \lambda h, \quad 0 \leq \lambda \leq \pi.$$

Иногда $g(\lambda)$ также называют спектральной плотностью. Тогда

$$(5) \quad \sigma(k) = \int_0^{\pi} \cos \lambda k g(\lambda) d\lambda.$$

Определим *спектральную функцию* (если $f(\lambda)$ определена) следующим образом:

$$(6) \quad F(\lambda) = \int_{-\pi}^{\lambda} f(v) dv.$$

Пусть (2) имеет место; интегрируя (1) почленно, получаем

$$(7) \quad F(\lambda) = \frac{\sigma(0)}{2\pi} (\lambda + \pi) + \frac{1}{\pi} \sum_{h=1}^{\infty} \frac{\sigma(h) \sin \lambda h}{h}.$$

Отсюда (3) может быть представлено в виде

$$(8) \quad \sigma(k) = \int_{-\pi}^{\pi} \cos \lambda k dF(\lambda).$$

Покажем, что для любой ковариационной последовательности $\{\sigma(k)\}$ существует функция $F(\lambda)$, такая, что верна формула (8), и функция $F(\lambda)$ монотонно не убывает. Последнее свойство соответствует тому, что $f(\lambda) \geq 0$ (в случае когда $F(\lambda)$ абсолютно непрерывна).

Определим периодограммы по выборочным значениям, как в разд. 4.4.1:

$$(9) \quad R^2(\lambda) = A^2(\lambda) + B^2(\lambda), \quad -\pi \leq \lambda \leq \pi,$$

где

$$(10) \quad A(\lambda) = \frac{2}{T} \sum_{t=1}^T y_t \cos \lambda t, \quad -\pi \leq \lambda \leq \pi,$$

$$(11) \quad B(\lambda) = \frac{2}{T} \sum_{t=1}^T y_t \sin \lambda t, \quad -\pi \leq \lambda \leq \pi.$$

Здесь $y_1, \dots, y_T$ суть T последовательных случайных величин стационарного в широком смысле процесса, имеющего среднее нулевое значение. Пусть

$$\begin{aligned} (12) \quad I_T(\lambda) &= \frac{T}{8\pi} R^2(\lambda) = \\ &= \frac{1}{2\pi T} \left[\sum_{s,t=1}^T y_s y_t \cos \lambda s \cos \lambda t + \sum_{s,t=1}^T y_s y_t \sin \lambda s \sin \lambda t \right] = \\ &= \frac{1}{2\pi T} \sum_{s,t=1}^T y_s y_t \cos \lambda (t-s) = \\ &= \frac{1}{2\pi T} \sum_{s,t=1}^T y_s y_t e^{i\lambda(t-s)} = \frac{1}{2\pi T} \left| \sum_{t=1}^T y_t e^{i\lambda t} \right|^2. \end{aligned}$$

Тогда

$$\begin{aligned} (13) \quad \mathcal{E} I_T(\lambda) &= \frac{1}{2\pi T} \sum_{s,t=1}^T \sigma(t-s) \cos \lambda (t-s) = \\ &= \frac{1}{2\pi} \sum_{h=-(T-1)}^{T-1} \left(1 - \frac{|h|}{T} \right) \sigma(h) \cos \lambda h = f_T(\lambda), \end{aligned}$$

так как $T - |h|$ есть число пар $s, t, s, t = 1, \dots, T$, таких, что $t - s = h, h = -(T-1), \dots, T-1$. Так как $R^2(\lambda) \geq 0$, то

$$(14) \quad f_T(\lambda) \geq 0,$$

и

$$\begin{aligned} (15) \quad \sigma(k) \left(1 - \frac{|k|}{T} \right) &= \int_{-\pi}^{\pi} \cos \lambda k f_T(\lambda) d\lambda, \quad k = -(T-1), \dots, T-1, \\ 0 &= \int_{-\pi}^{\pi} \cos \lambda k f_T(\lambda) d\lambda, \quad k = \pm T, \pm (T+1), \dots \end{aligned}$$

Отсюда получаем, что функция

$$\begin{aligned} (16) \quad F_T(\lambda) &= \int_{-\pi}^{\lambda} f_T(v) dv = \\ &= \frac{\sigma(0)}{2\pi} (\lambda + \pi) + \frac{1}{\pi} \sum_{h=1}^{T-1} \left(1 - \frac{h}{T} \right) \frac{\sigma(h) \sin \lambda h}{h} \end{aligned}$$

монотонно не убывает, $F_T(-\pi) = 0$ и $F_T(\pi) = \sigma(0)$. Из теоремы о слабой компактности [Лозэ (1963, § 11.2)] вытекает, что существует последовательность переменных $T_1, T_2, \dots$, такая, что подпоследовательность $F_{T_j}(\lambda)$ сходится. Предельная функция $F(\lambda)$ монотонно не убывает, и, кроме того, $F(-\pi) = 0$ и $F(\pi) = \sigma(0)$. Подпоследовательность сходится в каждой точке непрерывности функции $F(\lambda)$, и (8) можно рассматривать как предел подпоследовательности (15) [Лемма Хелли — Брея, Лозэ (1963, § 11.3)]. Равенство (7) выполняется в каждой точке непрерывности $F(\lambda)$. Удобно определить $F(\lambda)$, как непрерывную справа функцию, исключая $\lambda = -\pi$, т. е.

$$(17) \quad F(\lambda +) = F(\lambda), \quad -\pi < \lambda < \pi.$$

ТЕОРЕМА 7.3.1. *Для любой ковариационной последовательности $\{\sigma(h)\}$ существует однозначно определенная монотонно неубывающая функция $F(\lambda)$, $-\pi \leq \lambda \leq \pi$, с симметричными приращениями, непрерывная справа на интервале $(-\pi < \lambda < \pi)$, $F(-\pi) = 0$, такая, что*

$$(18) \quad \sigma(k) = \int_{-\pi}^{\pi} \cos \lambda k dF(\lambda), \quad k = 0, \pm 1, \dots$$

Положим

$$(19) \quad G(\lambda) = 2F(\lambda) - F(0) - F(0-), \quad 0 < \lambda \leq \pi,$$

и $G(0) = 0$. Тогда G непрерывна справа, за исключением значения $\lambda = 0$.

Полезно рассмотреть спектральные функции для нескольких примеров, приведенных выше в разд. 7.2.2.

Пример 7.1. Пусть y_t независимы с $\text{Var } y_t = \sigma^2$. Тогда ковариационной последовательности

$$(20) \quad \sigma(0) = \sigma^2, \quad \sigma(h) = 0, \quad h = \pm 1, \pm 2, \dots$$

соответствует

$$(21) \quad F(\lambda) = \frac{\sigma^2}{2\pi} (\lambda + \pi), \quad f(\lambda) = \frac{\sigma^2}{2\pi}, \quad -\pi \leq \lambda \leq \pi.$$

В самом деле,

$$(22) \quad \sigma(0) = \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} d\lambda = \sigma^2,$$

$$(23) \quad \sigma(h) = \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} \cos \lambda h d\lambda = 0, \quad h = \pm 1, \pm 2, \dots$$

Пример 7.2. Пусть $y_t \equiv y$ с дисперсией $\text{Var } y = \sigma^2$. Тогда ковариационной последовательности

$$(24) \quad \sigma(h) = \sigma^2, \quad h = 0, \pm 1, \dots$$

соответствует

$$(25) \quad F(\lambda) = \begin{cases} 0, & -\pi \leq \lambda < 0, \\ \sigma^2, & 0 \leq \lambda \leq \pi. \end{cases}$$

Это следует из соотношения

$$(26) \quad \sigma(h) = \int_{-\pi}^{\pi} \cos \lambda h dF(\lambda) = \sigma^2 \cos 0 = \sigma^2, \quad h = 0, \pm 1, \dots$$

Пример 7.3. Пусть

$$(27) \quad y_t = \sum_{j=1}^q (A_j \cos \lambda_j t + B_j \sin \lambda_j t), \quad t = \dots, -1, 0, 1, \dots,$$

где $\mathbb{E}A_j = \mathbb{E}B_j = 0$, $\mathbb{E}A_j^2 = \mathbb{E}B_j^2 = \sigma_j^2$, а A_j и B_j взаимно не коррелированы. Предположим, что λ_j упорядочены следующим образом: $0 \leq \lambda_1 < \lambda_2 < \dots < \lambda_q \leq \pi$. Функцию $F(\lambda)$ определим так, чтобы

$$(28) \quad \sigma(h) = \sum_{j=1}^q \sigma_j^2 \cos \lambda_j h, \quad h = \dots, -1, 0, 1, \dots$$

Эта формула выполняется в том случае, когда $F(\lambda)$ есть ступенчатая функция со скачками $\frac{1}{2}\sigma_j^2$ в каждой точке $\lambda = \pm\lambda_j$, $j = 1, \dots, q$ ($\lambda_1 \neq 0$ или σ_1^2 в точке $\lambda = \lambda_1 = 0$), как видно из рис. 7.2. В этом случае мы имеем

$$(29) \quad \sigma(0) = F(\pi) = \sum_{j=1}^q \sigma_j^2.$$

Процесс y_t представляется в виде суммы q косинусов и q синусов, если $\lambda_1 > 0$. Веса j -х косинуса и синуса есть пара некоррелированных выборочных значений из совокупности с дисперсией σ_j^2 . Здесь случайность возникает при введении случайных весовых коэффициентов.

Если A_j и B_j нормально распределены, то процесс может быть представлен в виде

$$(30) \quad y_t = \sum_{j=1}^q \sigma_j \sqrt{U_j} \cos(\lambda_j t - \theta_j), \quad t = \dots, -1, 0, 1, \dots,$$

где $U_1, \dots, U_q$ — независимые величины, имеющие χ^2 -распределение с двумя степенями свободы, а величины $\theta_1, \dots, \theta_q$, равномерно

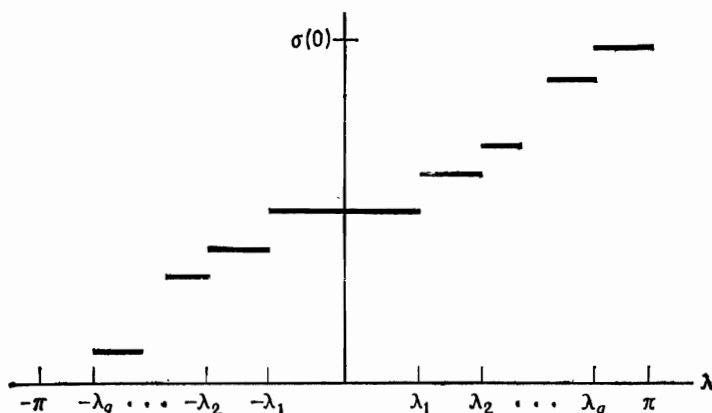


Рис. 7.2.

Дискретная спектральная функция распределения.

распределенные на отрезке $[0, 2\pi]$ тоже независимы. Реализация процесса представляет собой сумму q косинусов заданных частот. Фаза каждой частоты распределена равномерно, а амплитуды распределены одинаково с точностью до коэффициента пропорциональности σ_j .

Для любого стационарного случайного процесса можно построить процесс типа (27), спектральное распределение и ковариационные функции которого аппроксимируют те же характеристики исходного процесса посредством выбора $\sigma_1^2, \sigma_2^2, \dots, \sigma_q^2$. Аппроксимирующий процесс строится так, чтобы его ступенчатая спектральная функция была близка к функции $F(\lambda)$ исходного процесса. Выбирая q достаточно большим, можно получить сколь угодно хорошую аппроксимацию. В § 7.4 мы рассмотрим этот вопрос более подробно и покажем, что процесс $\{y_t\}$ можно представить в виде суммы двух интегралов.

Аппроксимация случайного процесса суммами или интегралами, подобными (30), позволяет представить процесс как сумму или интеграл косинусов со случайными фазами и амплитудами. Фазы часто не представляют никакого интереса, в то время как изучение амплитудных характеристик важно. Сумма средних значений квадратов амплитуд для заданных в некотором интервале частот представляет собой приращение спектральной функции на данном интервале.

Во многих случаях спектральное представление проще, более естественно и информативно, чем последовательность наблюдений. Это верно для многих природных явлений, таких, как звук, свет, электричество, волновой характер которых выражается суммой тригонометрических функций. Например, чистый звуковой тон описы-

вается тригонометрической функцией, где частота соответствует высоте, а амплитуда — силе звука; естественный рецептор, ухо, воспринимает высоту и силу звука. Спектр света дает наглядное представление о долях энергии на разных частотах. Такой спектральный подход очень важен в теории связи. В каждом из этих примеров представление тригонометрическими функциями дает основной способ описания наипростейших случаев и позволяет делать обобщения для более сложных ситуаций. Следует заметить, что спектральный анализ использует язык тригонометрических функций. Если бы «естественное» описание некоторых периодических явлений давалось на языке других периодических функций (например, пилообразная или ступенчатая функция), то приведенный спектральный анализ не был бы удачным. Функции косинус и синус появляются как решения дифференциальных или разностных уравнений второго порядка. Там, где явление описывается такими уравнениями, предпочтительны тригонометрические функции.

Значимость выборочных периодограммы и спектральной плотности (когда она существует) можно задать по-другому. Вычислим выборочную ковариацию между y_t и $z_t = \cos(\lambda t - \theta)$ (без коррекции для среднего значения):

$$\begin{aligned} (31) \quad \sum_{t=1}^T y_t \cos(\lambda t - \theta) &= \cos \theta \sum_{t=1}^T y_t \cos \lambda t + \\ &+ \sin \theta \sum_{t=1}^T y_t \sin \lambda t = \frac{T}{2} [\cos \theta A(\lambda) + \sin \theta B(\lambda)]. \end{aligned}$$

Найдем максимум по θ :

$$(32) \quad \max_{\theta} \sum_{t=1}^T y_t \cos(\lambda t - \theta) = \frac{T}{2} R(\lambda).$$

Таким образом, для рассматриваемого временного ряда $R(\lambda)$ пропорционально ковариации между наблюдениями временного ряда и максимумом по фазе косинуса частоты $\lambda/(2\pi)$. Исключая тот случай, когда сумма квадратов $\cos(\lambda t - \theta)$ зависит от θ , получаем, что $R(\lambda)$ пропорционально соответствующей корреляции. Грубо говоря, $R(\lambda)$ пропорционально коэффициенту множественной корреляции между $\{y_t\}$ и $\{(\cos \lambda t, \sin \lambda t)\}$. В этом смысле $R(\lambda)$ измеряет степень соответствия между наблюдениями временного ряда и тригонометрической функцией частоты $\lambda/(2\pi)$ (см. разд. 4.3.2); спектральную плотность можно интерпретировать как среднее некоторой совокупности.

Спектральная функция $F(\lambda)$ аналогична функции совместного распределения на отрезке $[-\pi, \pi]$, за исключением того факта, что $F(\pi) = \sigma(0)$ может отличаться от 1. Спектральная функция случайного стационарного процесса может быть представлена в виде

суммы

$$(33) \quad F(\lambda) = F_1(\lambda) + F_2(\lambda) + F_3(\lambda),$$

где $F_1(\lambda)$, $F_2(\lambda)$ и $F_3(\lambda)$ — неубывающие функции. Функция $F_1(\lambda)$ — ступенчатая, $F_2(\lambda)$ абсолютно непрерывна,

$$(34) \quad F_2(\lambda) = \int_{-\pi}^{\lambda} F_2'(\mu) d\mu,$$

и $F_3(\lambda)$ — сингулярная функция, т. е. $F_3(\lambda)$ непрерывна и может быть возрастающей, хотя $F_3'(\lambda) = 0$ почти всюду.

Процесс $y(t)$ с непрерывным параметром t , $-\infty < t < \infty$, стационарный в широком смысле, со средним нулевым значением и ковариационной функцией

$$(35) \quad \mathcal{E}y(t)y(s) = \sigma(t-s),$$

имеет спектральную функцию $F(\lambda)$, $-\infty < \lambda < \infty$, которая монотонно не убывает, имеет симметричные приращения и ограничена. Отсюда

$$(36) \quad \sigma(t) = \int_{-\infty}^{\infty} \cos \lambda t dF(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda t} dF(\lambda).$$

Если положить $G(\lambda) = 2F(\lambda) - F(0) - F(0-)$, $0 < \lambda < \infty$ и $G(0) = 0$, то (36) можно переписать:

$$(37) \quad \sigma(t) = \int_0^{\infty} \cos \lambda t dG(\lambda).$$

С точностью до постоянного множителя $F(\lambda)$ является функцией распределения, а $\sigma(t)$ — ее характеристической функцией.

Часто последовательность наблюдений y_t является выборкой через равные промежутки времени из процесса с непрерывным параметром, например ежечасные отсчеты температуры или ежедневные отсчеты уровня воды. Таким образом, $y_t = y(kt)$, $t = \dots, -1, 0, 1, \dots$, где k — интервал времени между наблюдениями. Ковариационная функция процесса с дискретным параметром есть

$$\begin{aligned} (38) \quad \mathcal{E}y_t y_{t+h} &= \mathcal{E}y(tk) y(tk + hk) = \sigma(hk) = \\ &= \int_{-\infty}^{\infty} \cos \lambda h k dF(\lambda) = \sum_{i=-\infty}^{\infty} \int_{(2j-1)\pi/k}^{(2j+1)\pi/k} \cos \lambda h k dF(\lambda) = \\ &= \sum_{j=-\infty}^{\infty} \int_{-\pi}^{\pi} \cos(v + 2\pi j) h dF\left(\frac{v + 2\pi j}{k}\right) = \\ &= \int_{-\pi}^{\pi} \cos v h \left\{ \sum_{j=-\infty}^{\infty} dF\left(\frac{v + 2\pi j}{k}\right) \right\}, \end{aligned}$$

при подстановке $\lambda = (\nu + 2\pi j)/k$. Предположим, что $F(\lambda)$ абсолютно непрерывна с плотностью $f(\lambda)$. Тогда

$$\begin{aligned}
 (39) \quad f_k(\nu) &= \frac{1}{k} \sum_{j=-\infty}^{\infty} f\left(\frac{\nu + 2\pi j}{k}\right) = \\
 &= \frac{1}{k} f\left(\frac{\nu}{k}\right) + \frac{1}{k} \sum_{j=1}^{\infty} \left[f\left(\frac{\nu + 2\pi j}{k}\right) + f\left(\frac{\nu - 2\pi j}{k}\right) \right] = \\
 &= \frac{1}{k} f\left(\frac{\nu}{k}\right) + \frac{1}{k} \sum_{j=1}^{\infty} \left[f\left(\frac{2\pi j + \nu}{k}\right) + f\left(\frac{2\pi j - \nu}{k}\right) \right], \\
 &\qquad\qquad\qquad -\pi \leq \nu \leq \pi
 \end{aligned}$$

есть не что иное, как спектральная плотность процесса $\{y_t\}$. Вес частоты $\nu/(2\pi)$ в процессе с дискретным параметром представляет собой сумму весов частот $\nu/(2\pi k)$, $(2\pi - \nu)/(2\pi k) = 1/k - \nu/(2\pi k)$, $(2\pi + \nu)/(2\pi k) = 1/k + \nu/(2\pi k)$, ... процесса с непрерывным временем. Предположим, например, что температура измеряется ежечасно; $k = 1$, если единица непрерывного параметра времени t есть один час. Период в 4 часа соответствует частоте $1/4$ и $\nu = \pi/2$. Для непрерывного процесса получим выборку частот $\frac{1}{4}$, $1 - \frac{1}{4} = \frac{3}{4}$, $1 + \frac{1}{4} = \frac{5}{4}$, $2 - \frac{1}{4} = \frac{7}{4}$, $2 + \frac{1}{4} = \frac{9}{4}$, ..., которая соответствует периодам 4, 4/3, 4/5, 4/7, 4/9, ... Этот эффект называется *подменой частот* или *свертыванием спектра*, а величина $1/(2k)$ называется *частотой Найквиста*.

Если выполняется (2), то ряд (1) для $f(\lambda)$ сходится абсолютно и равномерно, следовательно, $f(\lambda)$ непрерывна. Существует много различных условий на $f(\lambda)$, чтобы выполнялось (2), т. е. чтобы сумма абсолютных величин коэффициентов Фурье была сходящейся. Достаточным условием этого является

$$(40) \quad |f(\lambda_1) - f(\lambda_2)| \leq K |\lambda_1 - \lambda_2|^\alpha,$$

верное для некоторых $K > 0$ и α , $\frac{1}{2} < \alpha < 1$ [Хобсон (1907, разд. 359)].

Так как $f(\lambda)$ и $\cos \lambda h$ — четные функции, а $\sin \lambda h$ — нечетная, то (1) и (3) можно представить в следующем виде:

$$(41) \quad f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \sigma(h) e^{i\lambda h},$$

$$(42) \quad \sigma(k) = \int_{-\pi}^{\pi} e^{i\lambda k} f(\lambda) d\lambda.$$

Нормированная спектральная плотность $\bar{f}(\lambda) = f(\lambda)/\sigma(0)$ получается из (1) или (41) подстановкой $\rho_h = \sigma(h)/\sigma(0)$ вместо $\sigma(h)$. Интегралы Фурье функции $\bar{f}(\lambda)$ равны ρ_h ; в частности, интеграл от функции $\bar{f}(\lambda)$ на $[-\pi, \pi]$ есть $\rho_0 = 1$.

Перейдем теперь к спектральной теории случайных величин с комплексными значениями. Пусть $y_t = u_t + iv_t$, где $\{u_t, v_t\}$ — двумерный действительный случайный процесс, стационарный в широком смысле с $\mathbb{E}u_t = \mathbb{E}v_t = 0$ и

$$(43) \quad \begin{aligned} \sigma(h) &= \mathbb{E}y_{t+h}\bar{y}_t = \mathbb{E}(u_{t+h} + iv_{t+h})(u_t - iv_t) = \\ &= \mathbb{E}u_{t+h}u_t + \mathbb{E}v_{t+h}v_t + i(\mathbb{E}v_{t+h}u_t - \mathbb{E}u_{t+h}v_t), \\ h &= \dots, -1, 0, 1, \dots \end{aligned}$$

Заметим, что $\{\sigma(h)\}$ есть последовательность комплексных чисел, таких, что $\sigma(-h) = \bar{\sigma}(h)$, где $\bar{\sigma}(h)$ означает комплексное число, сопряженное с $\sigma(h)$. Если $\mathbb{E}u_{t+h}u_t = \mathbb{E}v_{t+h}v_t$ и $\mathbb{E}u_{t+h}v_t = -\mathbb{E}v_{t+h}u_t$, то $\{\sigma(h)\}$ определяет все ковариации процесса $\{u_t, v_t\}$.

Определим $I_T(\lambda)$ следующим образом:

$$(44) \quad I_T(\lambda) = \frac{1}{2\pi T} \left| \sum_{t=1}^T y_t e^{-i\lambda t} \right|^2 = \frac{1}{2\pi T} \sum_{t,s=1}^T y_t \bar{y}_s e^{-i\lambda(t-s)}.$$

Тогда

$$(45) \quad \begin{aligned} \mathbb{E}I_T(\lambda) &= \frac{1}{2\pi T} \sum_{t,s=1}^T \sigma(t-s) e^{-i\lambda(t-s)} = \\ &= \frac{1}{2\pi} \sum_{h=-(T-1)}^{T-1} \left(1 - \frac{|h|}{T}\right) \sigma(h) e^{-i\lambda h} = f_T(\lambda) \end{aligned}$$

неотрицательно согласно (44). [$f_T(\lambda)$, вообще говоря, не равна $f_T(-\lambda)$.] Имеем

$$(46) \quad \begin{aligned} \sigma(k) \left(1 - \frac{|k|}{T}\right) &= \int_{-\pi}^{\pi} e^{i\lambda k} f_T(\lambda) d\lambda, \\ k &= -(T-1), \dots, (T-1), \\ 0 &= \int_{-\pi}^{\pi} e^{i\lambda k} f_T(\lambda) d\lambda, \\ k &= \pm T, \pm(T+1), \dots \end{aligned}$$

Отсюда $F_T(\lambda) = \int_{-\pi}^{\lambda} f_T(v) dv$, согласно (16), монотонно не убывает с $F_T(-\pi) = 0$ и $F_T(\pi) = \sigma(0)$. Аналогичные рассуждения приводят к тому, что предел подпоследовательности $F(\lambda)$ монотонно не убывает.

вает, непрерывен справа, $F(-\pi) = 0$, $F(\pi) = \sigma(0)$ и

$$(47) \quad \sigma(k) = \int_{-\pi}^{\pi} e^{i\lambda k} dF(\lambda).$$

[Для комплексного процесса в формуле (41) нужно заменить $e^{i\lambda h}$ на $e^{-i\lambda h}$.]

Рассмотрим векторный случайный процесс $\{y_t\}$, стационарный в широком смысле, с p комплексными компонентами. Пусть

$$(48) \quad \Sigma(h) = \mathcal{E} y_{t+h} y_t^*,$$

где y_t^* есть транспонированный вектор $\bar{y}_t$, компоненты которого соответственно комплексно сопряжены компонентам y_t . Отметим, что

$$(49) \quad \bar{\Sigma}(h) = \mathcal{E} \bar{y}_{t+h} y_t' = (\mathcal{E} y_{t+h} y_t^*)' = \Sigma'(-h).$$

Для любого фиксированного вектора c случайная последовательность $\{c'y_t\}$ является комплексным случайным процессом, стационарным в широком смысле с математическим ожиданием $\mathcal{E} c'y_t = 0$ и ковариационной последовательностью

$$(50) \quad \mathcal{E} c'y_{t+h} y_t' \bar{c} = c' \Sigma(h) \bar{c} = \int_{-\pi}^{\pi} e^{i\lambda h} dF_c(\lambda),$$

где $F_c(\lambda)$ — действительная, монотонно неубывающая функция, непрерывная справа, такая, что $F_c(-\pi) = 0$ и $F_c(\pi) = c' \Sigma(0) \bar{c}$. Для вектора c , j -я компонента которого равна 1, а остальные 0, имеем

$$(51) \quad \sigma_{jj}(h) = \int_{-\pi}^{\pi} e^{i\lambda h} dF_{jj}(\lambda),$$

где $F_c(\lambda)$ для таких c записана как $F_{jj}(\lambda)$, $j = 1, \dots, p$. Для вектора c , j -я и k -я компоненты которого равны 1, а остальные 0, имеем

$$(52) \quad \sigma_{jj}(h) + \sigma_{kk}(h) + \sigma_{jk}(h) + \sigma_{kj}(h) = \int_{-\pi}^{\pi} e^{i\lambda h} dF_I(\lambda).$$

Для вектора c , j -я компонента которого равна 1, k -я равна i , а остальные — нулю, имеем

$$(53) \quad \sigma_{jj}(h) + \sigma_{kk}(h) - i\sigma_{jk}(h) + i\sigma_{kj}(h) = \int_{-\pi}^{\pi} e^{i\lambda h} dF_{II}(\lambda).$$

Таким образом,

$$(54) \quad \sigma_{jk}(h) = \int_{-\pi}^{\pi} e^{i\lambda h} dF_{jk}(\lambda), \quad j \neq k,$$

где

$$(55) \quad F_{jk}(\lambda) = \frac{1}{2} [F_{I}(\lambda) - F_{ij}(\lambda) - F_{kk}(\lambda)] + \\ + \frac{1}{2} i [F_{Ii}(\lambda) - F_{ji}(\lambda) - F_{kk}(\lambda)], \quad j \neq k.$$

Заметим, что $F_{jk}(\lambda)$, $j \neq k$, вообще говоря, не является действительной функцией. Выражение

$$(56) \quad \sigma_{jk}(h) = \int_{-\pi}^{\pi} e^{i\lambda h} dF_{jk}(\lambda), \quad j, k = 1, \dots, p,$$

можно переписать в виде

$$(57) \quad \Sigma(h) = \int_{-\pi}^{\pi} e^{i\lambda h} d\mathbf{F}(\lambda).$$

Так как (50) определяет функцию $F_c(\lambda)$ однозначно, то

$$(58) \quad F_c(\lambda) = \mathbf{c}' \mathbf{F}(\lambda) \bar{\mathbf{c}}.$$

Так как $F_c(\lambda)$ действительна и монотонно не убывает, то

$$(59) \quad [\mathbf{F}(\lambda_2) - \mathbf{F}(\lambda_1)] \bar{\mathbf{c}} \geq 0, \quad \lambda_1 < \lambda_2.$$

Таким образом, матрица $\mathbf{F}(\lambda)$ и ее приращения эрмитовы $[\mathbf{F}(\lambda) = \mathbf{F}^*(\lambda)]$ и положительно полуопределены.

Теперь предположим, что процесс $\{\mathbf{y}_t\}$ действителен. Тогда матрица $\Sigma(h)$ тоже действительна. Мнимая часть в формуле

$$(60) \quad \Sigma(h) = \int_{-\pi}^{\pi} [\cos \lambda h + i \sin \lambda h] [d\mathcal{R}\mathbf{F}(\lambda) + id\mathcal{I}\mathbf{F}(\lambda)] = \\ = \int_{-\pi}^{\pi} \{ \cos \lambda h d\mathcal{R}\mathbf{F}(\lambda) - \sin \lambda h d\mathcal{I}\mathbf{F}(\lambda) + \\ + i [\cos \lambda h d\mathcal{I}\mathbf{F}(\lambda) + \sin \lambda h d\mathcal{R}\mathbf{F}(\lambda)] \}$$

должна равняться 0 для каждого h . Если $\mathbf{F}(\lambda)$ абсолютно непрерывна и имеет плотность $\mathbf{f}(\lambda)$, то действительная часть функции $f_{jk}(\lambda)$ называется *коспектральной* плотностью, а мнимая часть — *квадратурной спектральной* плотностью для $j \neq k$. Отсюда

$$(61) \quad \mathcal{R}f_{jk}(-\lambda) = \mathcal{R}f_{jk}(\lambda), \quad j \neq k,$$

$$(62) \quad \mathcal{I}f_{jk}(-\lambda) = -\mathcal{I}f_{jk}(\lambda), \quad j \neq k,$$

а $f_{jj}(\lambda)$ действительна и симметрична. Используя обозначения спектральной плотности, (60) можно переписать в виде

$$(63) \quad \sigma_{jj}(h) = \int_{-\pi}^{\pi} \cos \lambda h f_{jj}(\lambda) d\lambda = 2 \int_0^{\pi} \cos \lambda h f_{jj}(\lambda) d\lambda,$$

$$(64) \quad \begin{aligned} \sigma_{jk}(h) &= \int_{-\pi}^{\pi} [\cos \lambda h \mathcal{R}f_{jk}(\lambda) - \sin \lambda h \mathcal{I}f_{jk}(\lambda)] d\lambda = \\ &= 2 \int_0^{\pi} [\cos \lambda h \mathcal{R}f_{jk}(\lambda) - \sin \lambda h \mathcal{I}f_{jk}(\lambda)] d\lambda, \quad j \neq k. \end{aligned}$$

В терминах спектральных функций формулам (61) и (62) соответствует $\mathbf{F}(\lambda_2) - \mathbf{F}(\lambda_1) = \mathbf{F}(-\lambda_1) - \mathbf{F}(-\lambda_2)$ для $0 < \lambda_1 < \lambda_2$.

7.4. СПЕКТРАЛЬНОЕ ПРЕДСТАВЛЕНИЕ СТАЦИОНАРНОГО СЛУЧАЙНОГО ПРОЦЕССА

7.4.1. Стохастические интегралы

Процесс $\{y_t\}$ с непрерывным параметром t можно представить в спектральной форме, используя понятие стохастического интеграла. Стохастические интегралы случайных процессов связаны с интегралами от спектральных функций этих процессов.

Спектральное представление случайного стационарного процесса позволяет найти связь между спектральными функциями и дисперсиями случайных амплитуд тригонометрических функций, составляющих этот процесс.

Мы не будем излагать теорию спектрального представления стационарных процессов с полными выводами ввиду того, что никакие дальнейшие математические построения на ней не основываются. Эта теория нам нужна только для того, чтобы сделать ясными основные статистические идеи.

Рассмотрим случайный процесс $C(\lambda)$, $0 \leq \lambda \leq \pi$. Будем считать известными совместные распределения конечных совокупностей $C(\lambda_1)$, ..., $C(\lambda_n)$. Процесс $C(\lambda)$ назовем *процессом с некоррелированными приращениями*, если

$$(1) \quad \mathcal{E}[C(\lambda_2) - C(\lambda_1)][C(\lambda_4) - C(\lambda_3)] = 0, \\ 0 \leq \lambda_1 < \lambda_2 \leq \lambda_3 < \lambda_4 \leq \pi,$$

где $\mathcal{E}C(\lambda) = 0$, $0 \leq \lambda \leq \pi$. То есть приращения процесса $C(\lambda)$ по двум непересекающимся интервалам некоррелированы. Для того чтобы (1) имело смысл, предположим, что

$$(2) \quad \mathcal{E}C^2(\lambda) = H(\lambda) < \infty, \quad 0 \leq \lambda \leq \pi.$$

Потребуем, чтобы $C(0)$ было не коррелировано с $C(\lambda_2) - C(\lambda_1)$, $0 \leq \lambda_1 < \lambda_2 \leq \pi$. Для $\lambda_1 < \lambda_2$ из (1) вытекает

$$(3) \quad H(\lambda_2) = \mathcal{E} C^2(\lambda_2) = \mathcal{E} \{C(\lambda_1) + [C(\lambda_2) - C(\lambda_1)]\}^2 = \\ = H(\lambda_1) + \mathcal{E} [C(\lambda_2) - C(\lambda_1)]^2.$$

Таким образом,

$$(4) \quad H(\lambda_2) \geq H(\lambda_1), \quad \lambda_2 > \lambda_1,$$

т. е. $H(\lambda)$ должна быть монотонно неубывающей. Более того, для $\lambda_1 < \lambda_2$

$$(5) \quad \mathcal{E} C(\lambda_1) C(\lambda_2) = \mathcal{E} C(\lambda_1) \{C(\lambda_1) + [C(\lambda_2) - C(\lambda_1)]\} = \\ = \mathcal{E} C^2(\lambda_1) = H(\lambda_1), \quad \lambda_1 < \lambda_2.$$

Говорят, что процесс $C(\lambda)$ является *процессом с независимыми приращениями*, если его приращения по каждому конечному набору непересекающихся интервалов взаимно независимы. В таком случае $C(\lambda)$ для любого λ представляет собой сумму независимых случайных величин.

Теперь мы попытаемся определить

$$(6) \quad \int_0^\pi h(\lambda) dC(\lambda)$$

по аналогии с интегралом Римана — Стильтьеса

$$(7) \quad \int_0^\pi h(\lambda) dH(\lambda),$$

где $h(\lambda)$ непрерывна на отрезке $[0, \pi]$, а функция $H(\lambda)$ — монотонно неубывающая. Предположим, что $H(\lambda)$ непрерывна в точках 0 и π . Тогда (7) аппроксимируется выражением

$$(8) \quad S = \sum_{j=1}^n h(\lambda'_j) [H(\lambda_j) - H(\lambda_{j-1})],$$

где

$$(9) \quad 0 = \lambda_0 < \lambda_1 < \dots < \lambda_{n-1} < \lambda_n = \pi, \\ \lambda_0 \leq \lambda'_1 \leq \lambda_1, \dots, \lambda_{n-1} \leq \lambda'_n \leq \lambda_n.$$

Мы хотим доказать, что если $\max_{0 \leq j < n} |\lambda_{j+1} - \lambda_j|$ достаточно мал, то сумма (8) слабо зависит от используемого разбиения.

Так как функция $h(\lambda)$ непрерывна на отрезке $[0, \pi]$, то она и равномерно непрерывна на этом же отрезке. То есть для любого $\varepsilon > 0$ существует такое δ , что

$$(10) \quad |h(v) - h(\mu)| < \varepsilon, \quad |v - \mu| < \delta.$$

ЛЕММА 7.4.1. Если S определено по формуле (8) и

$$(11) \quad R = \sum_{i=1}^m h(v_i') [H(v_i) - H(v_{i-1})],$$

где $v_0, v_1, \dots, v_m, v_1', \dots, v_m'$ — некоторое новое разбиение отрезка $[0, \pi]$, и если

$$(12) \quad \max_{i=1, \dots, n} (\lambda_i - \lambda_{i-1}) < \frac{1}{2} \delta, \quad \max_{i=1, \dots, m} (v_i - v_{i-1}) < \frac{1}{2} \delta$$

для некоторого δ , такого, что (10) имеет место для заданного ε , то

$$(13) \quad |S - R| < \varepsilon [H(\pi) - H(0)].$$

ДОКАЗАТЕЛЬСТВО. Пусть $0 = \mu_0 < \mu_1 < \dots < \mu_l = \pi$ отличны от $\lambda_0, \dots, \lambda_n, v_0, \dots, v_m$. Пусть λ_k^* равно λ_j , если $(\mu_{k-1}, \mu_k) \subset (\lambda_{j-1}, \lambda_j)$, и v_k равно v_j , если $(\mu_{k-1}, \mu_k) \subset (v_{j-1}, v_j)$. Тогда имеет место соотношение

$$(14) \quad |S - R| = \left| \sum_{k=1}^l h(\lambda_k^*) [H(\mu_k) - H(\mu_{k-1})] - \sum_{k=1}^l h(v_k^*) [H(\mu_k) - H(\mu_{k-1})] \right| \leq \\ \leq \sum_{k=1}^l |h(\lambda_k^*) - h(v_k^*)| [H(\mu_k) - H(\mu_{k-1})] \leq \varepsilon [H(\pi) - H(0)],$$

ввиду того что $|\lambda_k^* - v_k^*| < \delta$. ■

Если мы возьмем последовательность разбиений $\{\lambda_j^{(n)}, \lambda_j'^{(n)}\}_r$ таких, что $\max_j |\lambda_j^{(n)} - \lambda_{j-1}^{(n)}| \rightarrow 0$, то совместные суммы $S^{(n)}$ сходятся и их пределом является интеграл (7). По лемме 7.4.1 частичные суммы $R^{(m)}$, определенные любой другой последовательностью разбиений $\{v_j^{(m)}, v_j'^{(m)}\}$, сходятся к той же самой величине.

Определим стохастический интеграл (6), как предел в среднем¹⁾ аппроксимирующих сумм

$$(15) \quad S = \sum_{j=1}^n h(\lambda_j') [C(\lambda_j) - C(\lambda_{j-1})].$$

Заметим, что сумма является линейной комбинацией некоррелированных случайных величин. Математическое ожидание аппроксимирующей суммы (15) равно нулю, а дисперсия равна

$$(16) \quad \mathbb{E} \left\{ \sum_{j=1}^n h(\lambda_j') [C(\lambda_j) - C(\lambda_{j-1})] \right\}^2 = \sum_{j=1}^n h^2(\lambda_j') [H(\lambda_j) - H(\lambda_{j-1})].$$

¹⁾ Напомним, что так автор называет сходимость в среднеквадратичном.
— Прим. ред.

Покажем, что различие между S и выражением

$$(17) \quad R = \sum_{j=1}^m h(v_j') [C(v_j) - C(v_{j-1})]$$

мало в том смысле, что математическое ожидание $\mathcal{E}(S - R)^2$ мало.

ЛЕММА 7.4.2. Если S и R определены соответственно формулами (15) и (17), $H(\lambda)$ — ковариационная функция $C(\lambda)$ и (12) выполняется для некоторого δ , такого, что (10) имеет место для заданного ε , то

$$(18) \quad \mathcal{E}(S - R)^2 < \varepsilon^2 [H(\pi) - H(0)].$$

ДОКАЗАТЕЛЬСТВО. Пусть μ_k , λ_k'' и v_k'' , $k = 1, \dots, l$ определены как и раньше. Тогда

$$(19) \quad \begin{aligned} \mathcal{E}(S - R)^2 &= \mathcal{E} \left\{ \sum_{k=1}^l [h(\lambda_k'') - h(v_k'')] [C(\lambda_k) - C(\lambda_{k-1})] \right\}^2 = \\ &= \sum_{k=1}^l [h(\lambda_k'') - h(v_k'')]^2 [H(\lambda_k) - H(\lambda_{k-1})] < \\ &< \varepsilon^2 [H(\pi) - H(0)]. \blacksquare \end{aligned}$$

Последовательность сумм $S^{(n)}$, определенная последовательностью разбиений $\{\lambda_j^{(n)}, \lambda_j'^{(n)}\}$, таких, что $\max_j [\lambda_j^{(n)} - \lambda_{j-1}^{(n)}] \rightarrow 0$, сходится в среднем к случайной величине, которая записывается в виде интеграла (6) (теорема 7.6.1). Из леммы 7.4.2 следует, что любая другая последовательность сумм $R^{(n)}$ сходится в среднем к той же самой случайной величине.

Так как $\mathcal{E}C(\lambda) = 0$, то математическое ожидание интеграла (6) равно нулю. Из (16) получаем

$$(20) \quad \mathcal{E} \left\{ \int_0^\pi h(\lambda) dC(\lambda) \right\}^2 = \lim_{n \rightarrow \infty} \sum_{j=1}^n h^2(\lambda_j'^{(n)}) [H(\lambda_j^{(n)}) - H(\lambda_{j-1}^{(n)})] = \\ = \int_0^\pi h^2(\lambda) dH(\lambda).$$

Если $k(\lambda)$ — другая непрерывная функция на $[0, \pi]$, то

$$(21) \quad \mathcal{E} \int_0^\pi h(\lambda) dC(\lambda) \int_0^\pi k(\lambda) dC(\lambda) = \int_0^\pi h(\lambda) k(\lambda) dH(\lambda).$$

Это следует из теоремы 7.6.4.

Мы предположили, что $H(\lambda)$ непрерывна в 0 и π . Если это не так, то интеграл (6) будет включать $h(0) [C(0+) - C(0)]$ и $h(\pi) \times [C(\pi) - C(\pi-)]$, или любое из этих выражений в отдельности.

Для более глубокого изучения стохастических интегралов можно обратиться к книге Дуба (1953, гл. IX, разд. 2).

7.4.2. Спектральное представление

Пусть $C(\lambda)$ и $S(\lambda)$ — некоррелированные случайные процессы с непрерывным параметром λ , $0 \leq \lambda \leq \pi$, и некоррелированными приращениями. Предположим, что $\mathcal{E}C(\lambda) = \mathcal{E}S(\lambda) = 0$ и

$$(22) \quad \mathcal{E}[C(\lambda_2) - C(\lambda_1)]^2 = \mathcal{E}[S(\lambda_2) - S(\lambda_1)]^2 = G(\lambda_2) - G(\lambda_1),$$

$$0 \leq \lambda_1 < \lambda_2 \leq \pi$$

для некоторой монотонно неубывающей функции $G(\lambda)$. Определим процесс $\{y_t\}$, как

$$(23) \quad y_t = \int_0^\pi \cos \lambda t dC(\lambda) + \int_0^\pi \sin \lambda t dS(\lambda), \quad t = \dots, -1, 0, 1, \dots$$

Согласно разд. 7.4.1, $\mathcal{E}y_t = 0$ и

$$(24) \quad \mathcal{E}y_t y_s = \int_0^\pi \cos \lambda t \cos \lambda s dG(\lambda) + \int_0^\pi \sin \lambda t \sin \lambda s dG(\lambda) =$$

$$= \int_0^\pi \cos \lambda (t - s) dG(\lambda).$$

Таким образом, $G(\lambda)$ является спектральной функцией (на $[0, \pi]$) процесса, определенного формулой (23).

Обратно, пусть дан случайный процесс $\{y_t^*\}$, стационарный в широком смысле с функцией распределения $G^*(\lambda)$ (на $[0, \pi]$). Мы можем построить два некоррелированных процесса с некоррелированными приращениями $C^*(\lambda)$ и $S^*(\lambda)$, таких, что $\{y_t^*\}$ может быть представлен как интеграл (23) и

$$(25) \quad \mathcal{E}[C^*(\lambda_2) - C^*(\lambda_1)]^2 = \mathcal{E}[S^*(\lambda_2) - S^*(\lambda_1)]^2 =$$

$$= G^*(\lambda_2) - G^*(\lambda_1), \quad 0 \leq \lambda_1 < \lambda_2 \leq \pi.$$

Эти процессы по существу являются пределами интегралов преобразований $\sum_t y_t^* \cos \lambda t$ и $\sum_t y_t^* \sin \lambda t$. Исходя из наблюдений $y_{-T}^*, \dots, y_0^*, \dots, y_t^*, \dots$, имеем

$$(26) \quad c_T(v) = \frac{1}{\pi} \sum_{t=-T}^T y_t^* \cos vt, \quad 0 \leq v \leq \pi,$$

$$(27) \quad s_T(v) = \frac{1}{\pi} \sum_{t=-T}^T y_t^* \sin vt, \quad 0 \leq v \leq \pi,$$

$$(28) \quad C_T(\lambda) = \int_0^\lambda c_T(v) dv = \sum_{t=-T}^T \alpha_{\lambda,t} y_t^*,$$

$$(29) \quad S_T(\lambda) = \int_0^\lambda s_T(v) dv = \sum_{t=-T}^T \beta_{\lambda,t} y_t^*,$$

где

$$(30) \quad \alpha_{\lambda,t} = \frac{1}{\pi} \int_0^\lambda \cos vt dv = \begin{cases} \frac{\sin \lambda t}{\pi t}, & t = \pm 1, \pm 2, \dots, \\ \frac{\lambda}{\pi}, & t = 0, \end{cases}$$

$$(31) \quad \beta_{\lambda,t} = \frac{1}{\pi} \int_0^\lambda \sin vt dv = \begin{cases} \frac{1 - \cos \lambda t}{\pi t}, & t = \pm 1, \pm 2, \dots, \\ 0, & t = 0. \end{cases}$$

Эти коэффициенты являются коэффициентами Фурье ступенчатых функций

$$(32) \quad \alpha_\lambda(v) = \begin{cases} 1, & -\lambda < v < \lambda, \\ 0, & \lambda \leq |v|, \end{cases}$$

$$(33) \quad \beta_\lambda(v) = \begin{cases} -1, & -\lambda < v < 0, \\ 1, & 0 < v < \lambda, \\ 0, & \lambda \leq |v|. \end{cases}$$

Ряды Фурье

$$(34) \quad \alpha_\lambda^*(v) = \sum_{t=-\infty}^{\infty} \alpha_{\lambda,t} \cos vt,$$

$$(35) \quad \beta_\lambda^*(v) = \sum_{t=-\infty}^{\infty} \beta_{\lambda,t} \sin vt$$

сходятся поточечно, за исключением точек $v = \pm \lambda$, и сходятся в среднеквадратичном. Определим

$$(36) \quad C^*(\lambda) = \sum_{t=-\infty}^{\infty} \alpha_{\lambda,t} y_t^*,$$

$$(37) \quad S^*(\lambda) = \sum_{t=-\infty}^{\infty} \beta_{\lambda,t} y_t^*$$

как пределы в среднем соответственно $C_T(\lambda)$ и $S_T(\lambda)$. Эти случайные величины существуют ввиду существования и конечности пределов $\mathcal{E}C_T(\lambda)$, $C_{T'}(\lambda)$ и $\mathcal{E}S_T(\lambda)$, $S_{T'}(\lambda)$ (следствие 7.6.1).

Вычислим следующие величины:

$$(38) \quad \mathcal{E}C^*(\lambda) C^*(\mu) = \mathcal{E} \sum_{s,t=-\infty}^{\infty} \alpha_{\lambda,s} \alpha_{\mu,t} y_s^* y_t^* = \\ = \sum_{s,t=-\infty}^{\infty} \alpha_{\lambda,s} \alpha_{\mu,t} \int_0^\pi \cos v(s-t) dG^*(v) =$$

$$\begin{aligned}
&= \int_0^\pi \sum_{s,t=-\infty}^{\infty} \alpha_{\lambda,s} \alpha_{\mu,t} (\cos vs \cos vt + \sin vs \sin vt) dG^*(v) = \\
&= \int_0^\pi \left(\sum_{s=-\infty}^{\infty} \alpha_{\lambda,s} \cos vs \sum_{t=-\infty}^{\infty} \alpha_{\mu,t} \cos vt + \right. \\
&\quad \left. + \sum_{s=-\infty}^{\infty} \alpha_{\lambda,s} \sin vs \sum_{t=-\infty}^{\infty} \alpha_{\mu,t} \sin vt \right) dG^*(v) = \\
&= \int_0^\pi \alpha_\lambda(v) \alpha_\mu(v) dG^*(v) = \min [G^*(\lambda), G^*(\mu)],
\end{aligned}$$

$$\begin{aligned}
(39) \quad \mathcal{G}S^*(\lambda) S^*(\mu) &= \int_0^\pi \left(\sum_{s=-\infty}^{\infty} \beta_{\lambda,s} \cos vs \sum_{t=-\infty}^{\infty} \beta_{\mu,t} \cos vt + \right. \\
&\quad \left. + \sum_{s=-\infty}^{\infty} \beta_{\lambda,s} \sin vs \sum_{t=-\infty}^{\infty} \beta_{\mu,t} \sin vt \right) dG^*(v) = \\
&= \int_0^\pi \beta_\lambda(v) \beta_\mu(v) dG^*(v) = \min [G^*(\lambda), G^*(\mu)],
\end{aligned}$$

$$\begin{aligned}
(40) \quad \mathcal{G}C^*(\lambda) S^*(\mu) &= \int_0^\pi \left(\sum_{s=-\infty}^{\infty} \alpha_{\lambda,s} \cos vs \sum_{t=-\infty}^{\infty} \beta_{\mu,t} \cos vt + \right. \\
&\quad \left. + \sum_{s=-\infty}^{\infty} \alpha_{\lambda,s} \sin vs \sum_{t=-\infty}^{\infty} \beta_{\mu,t} \sin vt \right) dG^*(v) = \\
&= 0,
\end{aligned}$$

так как

$$(41) \quad \sum_{t=-\infty}^{\infty} \alpha_{\mu,t} \sin vt = 0,$$

$$(42) \quad \sum_{t=-\infty}^{\infty} \beta_{\mu,t} \cos vt = 0,$$

в силу того, что $\alpha_{\mu,t} = \alpha_{\mu,-t}$, $\sin vt = -\sin(-vt)$, $\beta_{\mu,t} = -\beta_{\mu,-t}$ и $\cos vt = \cos(-vt)$. Так как ряды Фурье для $\alpha_\lambda(v)$ и $\beta_\lambda(v)$ сходятся и непрерывны, за исключением точек $-\lambda, 0, \lambda$, то приведенные рассуждения верны, если $0, \lambda$ и μ являются точками непрерывности $G^*(v)$. Для того чтобы сделать доказательства полными, нужно подробнее рассмотреть пределы под знаком математического ожидания. Теорема 7.6.4 позволяет вычислить дисперсии и ковариации [см. Дуб (1953, гл. X, разд. 4)].

Предшествующие рассуждения показывают, каким образом можно построить процессы $C^*(\lambda)$ и $S^*(\lambda)$, чтобы процесс $\{y_t^*\}$, стационарный,

нарный в широком смысле, представлялся в спектральной форме. Процессы $C^*(\lambda)$ и $S^*(\lambda)$ определяются единственным образом.

В приведенных выше рассуждениях мы показали, что любой стационарный (в широком смысле) случайный процесс можно рассматривать как взвешенную сумму или интеграл от тригонометрических функций времени со случайными весами. Эффект этих весов в среднем определяется их дисперсиями.

7.5. ЛИНЕЙНЫЕ ОПЕРАЦИИ НАД СТАЦИОНАРНЫМИ ПРОЦЕССАМИ

7.5.1. Ковариационные и спектральные функции процессов, полученных линейными операциями над стационарными процессами

Пусть $\{y_t\}$ — случайный процесс со средним $\mathbb{E}y_t = 0$, ковариационной функцией $\mathbb{E}y_t y_s = \sigma(s - t)$ и спектральной функцией $F(\lambda)$. (Процесс $\{y_t\}$ стационарен в широком смысле.) Новый процесс $\{z_t\}$ может быть получен с помощью некоторой линейной операции над процессом $\{y_t\}$:

$$(1) \quad z_t = \sum_r c_r y_{t-r}, \quad t = \dots, -1, 0, 1, \dots,$$

где $\{c_r\}$ — последовательность постоянных величин. Если ряд в формуле (1) бесконечен, определим z_t как предел в среднем (предполагая, что он существует). Эта операция иногда называется *линейным фильтром*.

Среднее значение $\mathbb{E}z_t$ полученного процесса равно 0, а

$$\begin{aligned} (2) \quad \mathbb{E}z_t z_s &= \mathbb{E} \sum_{r,q} c_r y_{t-r} c_q y_{s-q} = \\ &= \sum_{r,q} c_r c_q \sigma[(t-r) - (s-q)] = \\ &= \sum_{r,q} c_r c_q \sigma[(t-s) - (r-q)] \end{aligned}$$

является его ковариационной функцией.

Случайная величина z_t существует как предел в среднем (когда ряд (1) бесконечен) тогда и только тогда, когда правая часть в (2) сходится при $s = t$ (следствие 7.6.1). В этом случае правая часть есть не что иное, как $\mathbb{E}z_t z_s$ (теорема 7.6.4). Равенство (2) показывает, что процесс $\{z_t\}$ стационарен в широком смысле.

Спектральная функция процесса $\{z_t\}$ получается из формулы (2), которая может быть записана в виде

$$\begin{aligned}
 (3) \quad \mathcal{E}z_t z_s &= \sum_{r,q} c_r c_q \int_{-\pi}^{\pi} \cos \lambda [(t-s) - (r-q)] dF(\lambda) = \\
 &= \int_{-\pi}^{\pi} \mathcal{R} \sum_{r,q} c_r c_q e^{i\lambda[(t-s)+(q-r)]} dF(\lambda) = \\
 &= \int_{-\pi}^{\pi} \mathcal{R} e^{i\lambda(t-s)} \sum_{r,q} c_q e^{i\lambda q} c_r e^{-i\lambda r} dF(\lambda) = \\
 &= \int_{-\pi}^{\pi} \cos \lambda (t-s) \left| \sum_r c_r e^{i\lambda r} \right|^2 dF(\lambda).
 \end{aligned}$$

Таким образом, спектральная функция процесса $\{z_t\}$ имеет вид

$$(4) \quad H(\lambda) = \int_{-\pi}^{\lambda} \left| \sum_r c_r e^{i\lambda r} \right|^2 dF(v)$$

и

$$(5) \quad \mathcal{E}z_t z_s = \int_{-\pi}^{\pi} \cos \lambda (t-s) dH(\lambda).$$

Если $\{y_t\}$ имеет спектральную плотность $F'(\lambda) = f(\lambda)$, то $\{z_t\}$ имеет спектральную плотность

$$(6) \quad H'(\lambda) = h(\lambda) = \left| \sum_r c_r e^{i\lambda r} \right|^2 f(\lambda).$$

Функцию $\sum_r c_r e^{i\lambda r}$ иногда называют *частотной характеристикой* линейного фильтра или *передаточной функцией*, а $\left| \sum_r c_r e^{i\lambda r} \right|^2$ называют *передаточной функцией мощности*.

ТЕОРЕМА 7.5.1. Если $\{y_t\}$ имеет спектральную функцию $F(\lambda)$, то процесс $\{z_t\}$, определенный формулой (1), имеет спектральную функцию (4). Если $\{y_t\}$ имеет спектральную плотность $f(\lambda)$, то спектральной плотностью процесса $\{z_t\}$ является (6).

7.5.2. Процессы скользящего среднего

Предположим, что $\{y_t\}$ — процесс скользящего среднего некоррелированных случайных величин

$$(7) \quad y_t = \sum_r \gamma_r v_{t-r},$$

где $\mathbb{E}v_t = 0$, $\mathbb{E}v_t^2 = 1$ и $\mathbb{E}v_t v_s = 0$, $t \neq s$. Для сходимости в среднем ряда (7) необходимо и достаточно, чтобы

$$(8) \quad \sum_r \gamma_r^2 < \infty$$

(следствие 7.6.1). Процесс $\{v_t\}$ имеет спектральную плотность, равную $1/(2\pi)$, а спектральной плотностью процесса $\{y_t\}$ является

$$(9) \quad f(\lambda) = \frac{1}{2\pi} \left| \sum_r \gamma_r e^{i\lambda r} \right|^2.$$

Процесс $\{y_t\}$ называется *процессом скользящего среднего*. Ковариационной функцией для $\{y_t\}$ является величина

$$(10) \quad \begin{aligned} \mathbb{E}y_t y_{t+h} &= \mathbb{E} \sum_{r,s} \gamma_r \gamma_s v_{t-r} v_{t+h-s} = \\ &= \sum_r \gamma_r \gamma_{r+h}, \end{aligned}$$

так как $\mathbb{E}v_{t-r} v_{t+h-s} = 1$, если $t-r = t+h-s$, и равно 0 в остальных случаях.

Обратно, если стационарный процесс $\{y_t\}$ имеет спектральную плотность $f(\lambda)$, его можно представить в виде (7). Квадратный корень из $\pi g(\lambda) = 2\pi f(\lambda)$ можно представить следующим образом:

$$(11) \quad \begin{aligned} \sqrt{\pi g(\lambda)} &= \sum_{r=-\infty}^{\infty} \gamma_r^* e^{i\lambda r} = \\ &= a_0 + \sum_{r=1}^{\infty} (a_r \cos \lambda r + b_r \sin \lambda r), \end{aligned}$$

где $a_0 = \gamma_0^*$, $a_r = \gamma_r^* + \gamma_{-r}^*$ и $b_r = i(\gamma_r^* - \gamma_{-r}^*)$, $r = 1, 2, \dots$. Так как $g(\lambda) = 2f(\lambda)$ — четная функция, то $b_r = 0$, $\gamma_r^* = \gamma_{-r}^*$ и $\gamma_r^* = \frac{1}{2} a_r$ ($r \neq 0$) вещественные. [Заметим, что в том случае, когда $\gamma_r = \gamma_{-r}$ не выполняется, а y_t определяется (7), $\sum_r \gamma_r e^{i\lambda r}$ является комплексной величиной.] Существует последовательность некоррелированных случайных величин $\{V_t\}$, такая, что y_t могут быть представлены как $\sum_{r=-\infty}^{\infty} \gamma_r^* V_{t-r}$, $\mathbb{E}V_t = 0$ и $\mathbb{E}V_t^2 = 1$.

Определим $\{V_t\}$, используя спектральное представление процесса $\{y_t\}$. Предположим, что $g(\lambda) > 0$, $0 \leq \lambda \leq \pi$. Пусть

$$(12) \quad V_t = \int_0^\pi \frac{\cos \lambda t}{\sqrt{\pi g(\lambda)}} dC(\lambda) + \int_0^\pi \frac{\sin \lambda t}{\sqrt{\pi g(\lambda)}} dS(\lambda).$$

Тогда

$$\begin{aligned}
 (13) \quad \sum_{r=-\infty}^{\infty} \gamma_r^* V_{t-r} &= \int_0^{\pi} \frac{\sum_{r=-\infty}^{\infty} \gamma_r^* \cos \lambda(t-r)}{\sqrt{\pi g(\lambda)}} dC(\lambda) + \\
 &+ \int_0^{\pi} \frac{\sum_{r=-\infty}^{\infty} \gamma_r^* \sin \lambda(t-r)}{\sqrt{\pi g(\lambda)}} dS(\lambda) = \\
 &= \int_0^{\pi} \frac{\sum_{r=-\infty}^{\infty} \gamma_r^* (\cos \lambda t \cos \lambda r + \sin \lambda t \sin \lambda r)}{\sqrt{\pi g(\lambda)}} dC(\lambda) + \\
 &+ \int_0^{\pi} \frac{\sum_{r=-\infty}^{\infty} \gamma_r^* (\sin \lambda t \cos \lambda r - \cos \lambda t \sin \lambda r)}{\sqrt{\pi g(\lambda)}} dS(\lambda) = \\
 &= \int_0^{\pi} \cos \lambda t dC(\lambda) + \int_0^{\pi} \sin \lambda t dS(\lambda) = y_t,
 \end{aligned}$$

так как

$$(14) \quad \sum_{r=-\infty}^{\infty} \gamma_r^* \cos \lambda r = \sqrt{\pi g(\lambda)}, \quad \sum_{r=-\infty}^{\infty} \gamma_r^* \sin \lambda r = 0$$

ввиду $\gamma_r^* = \gamma_{-r}^*$. Таким образом, из формулы (12) получаем требуемое равенство. Более подробное изложение см. в книге Дуба (1953, гл. X, разд. 8).

Если $\{y_t\}$ имеет спектральную плотность $f(\lambda)$, $f(\lambda) > 0$ почти везде на $[0, \pi]$ и

$$(15) \quad \int_{-\pi}^{\pi} \log f(\lambda) d\lambda > -\infty,$$

то существуют (действительные) постоянные $\{\gamma_r\}$ и последовательность случайных величин $\{v_t\}$, такие, что

$$(16) \quad y_t = \sum_{r=0}^{\infty} \gamma_r v_{t-r}, \quad t = \dots, -1, 0, 1, \dots$$

Сумма (16), вообще говоря, бесконечна (см. разд. 7.6.3).

Теперь рассмотрим процесс скользящего среднего с конечными пределами суммирования ¹⁾

$$(17) \quad y_t = \sum_{r=0}^q \alpha_r v_{t-r},$$

где $\alpha_0 = 1$ и $\mathcal{E}v_t^2 = \sigma^2$. Тогда

$$(18) \quad \begin{aligned} f(\lambda) &= \frac{\sigma^2}{2\pi} \left| \sum_{r=0}^q \alpha_r e^{i\lambda r} \right|^2 = \frac{\sigma^2}{2\pi} \left| \sum_{r=0}^q \alpha_r e^{-i\lambda r} \right|^2 = \\ &= \frac{\sigma^2}{2\pi} \left| \sum_{r=0}^q \alpha_r e^{i\lambda(q-r)} \right|^2 = \frac{\sigma^2}{2\pi} \prod_{i=1}^q |e^{i\lambda} - z_i|^2, \end{aligned}$$

где $z_1, \dots, z_q$ — корни уравнения

$$(19) \quad \sum_{r=0}^q \alpha_r z^{q-r} = 0.$$

Если $\alpha_q \neq 0$, то все q корней отличны от нуля. Для конечного q будем использовать $\left| \sum_{r=0}^q \alpha_r e^{i\lambda(q-r)} \right|^2$, $\alpha_0 = 1$, как стандартную форму в $f(\lambda)$.

Как было показано в разд. 5.7.1, любой процесс с конечным числом отличных от нуля ковариаций имеет ту же самую последовательность ковариаций, как выбранный соответствующим образом конечный процесс скользящего среднего. Тогда спектральную плотность можно записать в виде (18).

Если $q = 1$,

$$(20) \quad y_t = v_t + \alpha_1 v_{t-1},$$

$$(21) \quad \begin{aligned} f(\lambda) &= \frac{\sigma^2}{2\pi} |e^{i\lambda} + \alpha_1|^2 = \frac{\sigma^2}{2\pi} (e^{i\lambda} + \alpha_1)(e^{-i\lambda} + \alpha_1) = \\ &= \frac{\sigma^2}{2\pi} (1 + \alpha_1^2 + 2\alpha_1 \cos \lambda). \end{aligned}$$

Так как $\cos \lambda$ монотонно изменяется от -1 до 1 на отрезке $[-\pi, 0]$ и от 1 до -1 на отрезке $[0, \pi]$, то функция $f(\lambda)$ возрастает от $\sigma^2(1 - \alpha_1^2)/(2\pi)$ до $\sigma^2(1 + \alpha_1^2)/(2\pi)$ на отрезке $[-\pi, 0]$ и убывает до $\sigma^2(1 - \alpha_1^2)/(2\pi)$ на отрезке $[0, \pi]$ для $\alpha_1 > 0$; если же $\alpha_1 < 0$, то $f(\lambda)$ убывает на отрезке $[-\pi, 0]$ и возрастает на $[0, \pi]$. Таким образом, если $\alpha_1 > 0$, то большую плотность имеют нижние частоты, если $\alpha_1 < 0$ — верхние. Ввиду того что $|e^{i\lambda}| = 1$, спектральную

¹⁾ В дальнейшем вместо процесса скользящего среднего с конечными пределами суммирования будем писать просто конечный процесс скользящего среднего. — *Прим. перев.*

плотность можно записать следующим образом:

$$(22) \quad f(\lambda) = \frac{\sigma^2}{2\pi} |e^{i\lambda} + \alpha_1|^2 = \frac{\sigma^2}{2\pi} |1 + \alpha_1 e^{-i\lambda}|^2 = \\ = \frac{\sigma^2 \alpha_1^2}{2\pi} \left| \frac{1}{\alpha_1} + e^{-i\lambda} \right|^2 = \frac{\sigma^2 \alpha_1^2}{2\pi} \left| e^{i\lambda} + \frac{1}{\alpha_1} \right|^2.$$

Последняя форма соответствует процессу $\tilde{v}_t + (1/\alpha_1) \tilde{v}_{t-1}$, $\tilde{v}_t$ имеет дисперсию $\sigma^2 \alpha_1^2$. Ковариационные функции этого процесса и процесса, определенного формулой (20), совпадают. Если $\alpha_1 \neq \pm 1$, то последний процесс скользящего среднего отличается от предыдущего. [Спектральную плотность, соответствующую $\alpha_1 \tilde{v}_t + \tilde{v}_{t-1}$, можно записать в виде $\sigma^2 |\alpha_1 e^{i\lambda} + 1|^2 / (2\pi)$, где $\tilde{v}_t$ имеет дисперсию σ^2 .]

Если $q = 2$, то

$$(23) \quad y_t = v_t + \alpha_1 v_{t-1} + \alpha_2 v_{t-2},$$

$$(24) \quad f(\lambda) = \frac{\sigma^2}{2\pi} |e^{i2\lambda} + \alpha_1 e^{i\lambda} + \alpha_2|^2 = \\ = \frac{\sigma^2}{2\pi} (e^{i2\lambda} + \alpha_1 e^{i\lambda} + \alpha_2)(e^{-i2\lambda} + \alpha_1 e^{-i\lambda} + \alpha_2) = \\ = \frac{\sigma^2}{2\pi} [1 + \alpha_1^2 + \alpha_2^2 + 2\alpha_1(1 + \alpha_2) \cos \lambda + 2\alpha_2 \cos 2\lambda] = \\ = \frac{\sigma^2}{2\pi} [\alpha_1^2 + (1 - \alpha_2)^2 + 2\alpha_1(1 + \alpha_2) \cos \lambda + 4\alpha_2 \cos^2 \lambda] = \\ = \frac{\sigma^2}{2\pi} \left\{ 4\alpha_2 \left[\cos \lambda + \frac{\alpha_1(1 + \alpha_2)}{4\alpha_2} \right]^2 + \frac{(1 - \alpha_2)^2 (4\alpha_2 - \alpha_1^2)}{4\alpha_2} \right\}.$$

Если $\alpha_1(1 + \alpha_2) > 4|\alpha_2|$, то $f(0) = \sigma^2(1 + \alpha_1 + \alpha_2)^2 / (2\pi)$ есть максимум функции $f(\lambda)$, а $f(\pm\pi) = \sigma^2(1 - \alpha_1 + \alpha_2)^2 / (2\pi)$ — минимум; если $\alpha_1(1 + \alpha_2) < -4|\alpha_2|$, то $f(0)$ — минимум, а $f(\pm\pi)$ — максимум. Если $|\alpha_1(1 + \alpha_2)| < 4|\alpha_2|$, то $\cos \lambda = -\alpha_1(1 + \alpha_2)/(4\alpha_2)$ для значения λ на отрезке $[0, \pi]$ и в точке $\lambda = -\lambda_0$; если $\alpha_2 > 0$ (подразумевая, что $\alpha_1^2 < 4\alpha_2$ и корни соответствующего полинома комплексны), то $f(\pm\lambda_0)$ есть минимум, а $f(0)$ и $f(\pm\pi)$ — относительный максимум; если $\alpha_2 < 0$ (подразумевая, что $4\alpha_2 < \alpha_1^2$ и корни действительны), то $f(\pm\lambda_0)$ есть максимум, а $f(0)$ и $f(\pm\pi)$ — относительный минимум. Если z_1 и z_2 — корни уравнения

$$(25) \quad z^2 + \alpha_1 z + \alpha_2 = 0,$$

то

$$(26) \quad f(\lambda) = \frac{\sigma^2}{2\pi} |e^{i\lambda} - z_1|^2 |e^{i\lambda} - z_2|^2.$$

Так как $|e^{i\lambda}| = 1$, то множитель $|e^{i\lambda} - z_1|$ можно заменить на $|1 - z_1 e^{-i\lambda}| = |1 - \bar{z}_1 e^{i\lambda}| = |z_1| |e^{i\lambda} - 1/\bar{z}_1|$, где $\bar{z}_1$ сопряжено с z_1 , а $|e^{i\lambda} - z_2|$ можно заменить на $|z_2| |e^{i\lambda} - 1/\bar{z}_2|$. Таким образом, для функции $f(\lambda)$ верно любое из следующих выражений:

$$\begin{aligned}
 f(\lambda) &= \frac{\sigma^2 |z_1|^2}{2\pi} \left| e^{i\lambda} - \frac{1}{\bar{z}_1} \right|^2 |e^{i\lambda} - z_2|^2 = \\
 &= \frac{\sigma^2 |z_1|^3}{2\pi} \left| e^{i2\lambda} - \left(\frac{1}{\bar{z}_1} + z_2 \right) e^{i\lambda} + \frac{z_2}{\bar{z}_1} \right|^2, \\
 (27) \quad f(\lambda) &= \frac{\sigma^2 |z_2|^2}{2\pi} |e^{i\lambda} - z_1|^2 \left| e^{i\lambda} - \frac{1}{\bar{z}_2} \right|^2 = \\
 &= \frac{\sigma^2 |z_2|^3}{2\pi} \left| e^{i2\lambda} - \left(z_1 + \frac{1}{\bar{z}_2} \right) e^{i\lambda} + \frac{z_1}{\bar{z}_2} \right|^2, \\
 f(\lambda) &= \frac{\sigma^2 |z_1|^2 |z_2|^2}{2\pi} \left| e^{i\lambda} - \frac{1}{\bar{z}_1} \right|^2 \left| e^{i\lambda} - \frac{1}{\bar{z}_2} \right|^2 = \\
 &= \frac{\sigma^2 \alpha_2^2}{2\pi} \left| e^{i2\lambda} + \frac{\alpha_1}{\alpha_2} e^{i\lambda} + \frac{1}{\alpha_2} \right|^2.
 \end{aligned}$$

Если z_1 и z_2 действительны, то каждое из приведенных выражений $f(\lambda)$ соответствует спектральной плотности процесса скользящего среднего. Четыре процесса скользящего среднего различны, если $z_1 \neq z_2$ и $z_i \neq \pm 1$; три процесса скользящего среднего различны, если $z_1 = z_2 \neq \pm 1$; два процесса скользящего среднего различны, если $z_1 = \pm 1$ и $z_2 \neq \pm 1$ или если $z_1 \neq \pm 1$ и $z_2 = \pm 1$; наконец, существует только один процесс скользящего среднего, если $z_1 = \pm 1$ и $z_2 = \pm 1$. Спектральная плотность является произведением двух плотностей указанного типа для $q = 1$.

Если z_1 и z_2 комплексно сопряжены, скажем $z_1 = \gamma e^{i\theta}$ и $z_2 = \gamma e^{-i\theta}$ ($0 < \theta < \pi$), то $z_2 + 1/\bar{z}_1$ и $z_1 + 1/\bar{z}_2$ не действительны, если $\gamma \neq 1$, и первые два выражения функции $f(\lambda)$ в (27) не могут соответствовать процессу скользящего среднего с действительными коэффициентами. Два процесса скользящего среднего с действительными коэффициентами различны. Все выражения для $f(\lambda)$ и процессы скользящего среднего совпадают, если $\gamma = 1$ (т. е. $\alpha_2 = 1$). Когда корни комплексно сопряжены,

$$\begin{aligned}
 (28) \quad f(\lambda) &= \frac{\sigma^2}{2\pi} |e^{i\lambda} - \gamma e^{i\theta}|^2 |e^{i\lambda} - \gamma e^{-i\theta}|^2 = \\
 &= \frac{\sigma^2}{2\pi} |e^{i2\lambda} - 2\gamma e^{i\lambda} \cos \theta + \gamma^2|^2 = \\
 &= \frac{\sigma^2}{2\pi} |(1 - \gamma^2)^2 + 4\gamma^2 (\cos \theta - \cos \lambda)^2 -
 \end{aligned}$$

$$-4\gamma(1-\gamma)^2 \cos \theta \cos \lambda] = \\ = \frac{\sigma^2}{2\pi} \left[4\gamma^2 \left(\cos \lambda - \frac{1+\gamma^2}{2\gamma} \cos \theta \right)^2 + (1-\gamma^2)^2 \sin^2 \theta \right].$$

Если γ близко к 1, то минимальное значение $f(\lambda)$ достигается для значений λ , близких $\pm \theta$. Действительно, минимум функции (28) достигается при $\cos \lambda = (1 + \gamma^2) \cos \theta / (2\gamma)$, если последнее выражение меньше 1 по абсолютной величине.

Для произвольного q спектральная плотность есть произведение, аналогичное формулам (21) и (28). Пусть $z_j = \gamma_j e^{i\theta_j}$. (Если $0 < \theta_j < \pi$, то $\theta_j = -\theta_k$ и $\gamma_j = \gamma_k$ для некоторых k .) Тогда

$$(29) \quad \sum_{r=0}^q \alpha_r \gamma_j^{q-r} e^{i\theta_j(q-r)} = 0.$$

Если γ_j близко к 1 (т. е. если z_j лежит близко к единичному кругу в комплексной плоскости), то

$$(30) \quad f(\theta_j) = \frac{\sigma^2}{2\pi} \left| \sum_{r=0}^q \alpha_r e^{i\theta_j(q-r)} \right|^2$$

будет близко к 0. Таким образом, частоты вблизи θ_j будут иметь малую интенсивность.

В общем случае множитель $|e^{i\lambda} - z_j|$ в $f(\lambda)$, как показано в формуле (18), можно переписать так:

$$(31) \quad |z_j| \cdot |1 - e^{i\lambda}/z_j| = |z_j| \cdot |e^{i\lambda} - 1/\bar{z}_j|,$$

где $\bar{z}_j$ комплексно сопряжено z_j . Если все корни действительны, различны и отличны от ± 1 , то существует 2^q различных представлений функции $f(\lambda)$, соответствующих различным процессам скользящего среднего. Число различных процессов скользящего среднего в общем случае зависит от числа корней, абсолютные значения которых равны 1, а также от кратности различных корней и числа комплексно сопряженных корней. Мы не будем перечислять все возможности для случая $q > 2$.

Нам будет удобно представить процесс скользящего среднего в таком виде, чтобы ни один корень формулы (19) не был больше единицы по абсолютной величине. (Заметим, что корень, абсолютное значение которого есть 1, допускается для процесса скользящего среднего.)

Процесс скользящего среднего (17) можно записать в виде

$$(32) \quad y_t = \left(\sum_{r=0}^q \alpha_r \mathcal{L}^r \right) v_t = \left(\sum_{r=0}^q \alpha_r \mathcal{P}^{q-r} \right) \mathcal{L}^q v_t = \\ = \left(\sum_{r=0}^q \alpha_r \mathcal{P}^{q-r} \right) v_{t-q},$$

где операторы $\mathcal{L}$ и $\mathcal{P}$ определены так, что $\mathcal{L}v_t = v_{t-1}$ и $\mathcal{P}v_t = v_{t+1}$. Если корни в (19) меньше 1 по абсолютной величине, то из (32) следует

$$(33) \quad v_t = \left(\sum_{r=0}^q \alpha_r \mathcal{L}^r \right)^{-1} y_t = \prod_{i=1}^q (1 - z_i \mathcal{L})^{-1} y_t = \\ = \prod_{i=1}^q \sum_{r=0}^{\infty} (z_i \mathcal{L})^r y_t.$$

Если

$$(34) \quad \left(\sum_{r=0}^q \alpha_r z^r \right)^{-1} = \sum_{r=0}^{\infty} \gamma_r z^r,$$

то (33) перепишем в виде

$$(35) \quad v_t = \sum_{r=0}^{\infty} \gamma_r y_{t-r}$$

или для $\gamma_0 = 1$

$$(36) \quad y_t = v_t - \sum_{r=1}^{\infty} \gamma_r y_{t-r}.$$

Таким образом,

$$(37) \quad \mathcal{E}(y_t | y_{t-1}, y_{t-2}, \dots) = - \sum_{r=1}^{\infty} \gamma_r y_{t-r}$$

является наилучшим прогнозом величин y_t по значениям $y_{t-1}, y_{t-2}, \dots$ в том смысле, что минимизируется среднеквадратичная ошибка.

7.5.3. Процессы авторегрессии

В гл. 5 мы изучали стационарный процесс $\{y_t\}$, который удовлетворяет уравнениям

$$(38) \quad \sum_{r=0}^p \beta_r y_{t-r} = u_t,$$

где $\{u_t\}$ — процесс некоррелированных случайных величин с дисперсией σ^2 . (Для рассмотрения свойств, определяемых вторыми моментами, нет необходимости предполагать u_t независимыми и одинаково распределенными величинами.) Для удобства допустим, что $\mathcal{E}u_t = 0 = \mathcal{E}y_t$ и $\beta_0 = 1$. Если все p корней $x_1, \dots, x_p$ уравнения

$$(39) \quad \sum_{r=0}^p \beta_r x^{p-r} = 0$$

по абсолютной величине меньше 1, то, решая уравнение (38), получаем

$$(40) \quad y_t = \sum_{r=0}^{\infty} \delta_r u_{t-r},$$

что является процессом скользящего среднего. Если $f(\lambda)$ — спектральная плотность процесса $\{y_t\}$, то

$$(41) \quad \left| \sum_{r=0}^p \beta_r e^{i\lambda r} \right|^2 f(\lambda) = \frac{\sigma^2}{2\pi},$$

так как правая часть (41) — спектральная плотность процесса $\{u_t\}$. Таким образом,

$$(42) \quad f(\lambda) = \frac{\sigma^2}{2\pi \left| \sum_{r=0}^p \beta_r e^{i\lambda r} \right|^2} = \frac{\sigma^2}{2\pi \left| \sum_{r=0}^p \beta_r e^{i\lambda(p-r)} \right|^2}.$$

Как показано в разд. 5.2.1, из стационарности процесса $\{y_t\}$ и $\sigma^2 > 0$ следует, что ни один корень уравнения (39) не может иметь абсолютное значение, равное 1. Это эквивалентно тому, что знаменатель в выражении (42) не обращается в нуль, и, следовательно, $f(\lambda)$ интегрируема.

Если $p = 1$,

$$(43) \quad y_t + \beta_1 y_{t-1} = u_t,$$

$$(44) \quad f(\lambda) = \frac{\sigma^2}{2\pi |e^{i\lambda} + \beta_1|^2} = \frac{\sigma^2}{2\pi (1 + \beta_1^2 + 2\beta_1 \cos \lambda)}.$$

Если $\beta_1 > 0$, то $f(\lambda)$ достигает минимума, равного $\sigma^2/[2\pi (1 + \beta_1^2)]$, в точке $\lambda = 0$, и максимума, равного $\sigma^2/[2\pi (1 - \beta_1^2)]$, в точках $\lambda = \pm \pi$. Функция $f(\lambda)$ убывает на отрезке $[-\pi, 0]$ и возрастает на отрезке $[0, \pi]$. Основную роль в спектре играют верхние частоты. Если $\beta_1 < 0$, то максимум достигается в точке $\lambda = 0$, а минимум — в точках $\lambda = \pm \pi$; в этом случае основную роль в спектре играют нижние частоты.

Если $p = 2$, то

$$(45) \quad y_t + \beta_1 y_{t-1} + \beta_2 y_{t-2} = u_t,$$

$$\begin{aligned} (46) \quad f(\lambda) &= \frac{\sigma^2}{2\pi |e^{i2\lambda} + \beta_1 e^{i\lambda} + \beta_2|^2} = \\ &= \frac{\sigma^2}{2\pi [1 + \beta_1^2 + \beta_2^2 + 2\beta_1 (1 + \beta_2) \cos \lambda + 2\beta_2 \cos 2\lambda]} = \\ &= \frac{\sigma^2}{2\pi [\beta_1^2 + (1 - \beta_2)^2 + 2\beta_1 (1 + \beta_2) \cos \lambda + 4\beta_2 \cos^2 \lambda]} = \\ &= \frac{\sigma^2}{2\pi \left\{ 4\beta_2 \left[\cos \lambda + \frac{\beta_1 (1 + \beta_2)}{4\beta_2} \right]^2 + \frac{(1 - \beta_2)^2 (4\beta_2 - \beta_1^2)}{4\beta_2} \right\}}. \end{aligned}$$

Если $\beta_1 (1 + \beta_2) > 4 |\beta_2|$, то $f(0) = \sigma^2 / [2\pi (1 + \beta_1 + \beta_2)^2]$ есть минимум функции $f(\lambda)$, а $f(\pm \pi) = \sigma^2 / [2\pi (1 - \beta_1 + \beta_2)^2]$ — максимум; если $\beta_1 (1 + \beta_2) < -4 |\beta_2|$, то $f(0)$ — максимум, а $f(\pm \pi)$ — минимум. Если $|\beta_1 (1 + \beta_2)| < 4 |\beta_2|$, то $\cos \lambda = -\beta_1 (1 + \beta_2) / (4\beta_2)$ для значения $\lambda = \lambda_0$ на $[0, \pi]$ и в точке $\lambda = -\lambda_0$; если $\beta_2 > 0$ (при $\beta_1^2 < 4\beta_2$ корни комплексные), то $f(\pm \lambda_0)$ — максимум, а $f(\pm \pi)$ и $f(0)$ — относительные минимумы; если $\beta_2 < 0$ (при $4\beta_2 < \beta_1^2$ корни действительные), то $f(\pm \lambda_0)$ — минимум, а $f(\pm \pi)$ и $f(0)$ — относительные максимумы. Если корни комплексно сопряжены, скажем $\rho e^{\pm i\psi}$, то

$$(47) \quad f(\lambda) = \frac{\sigma^2}{2\pi \left[4\rho^2 \left(\cos \lambda - \frac{1+\rho^2}{2\rho} \cos \psi \right)^2 + (1-\rho^2)^2 \sin^2 \psi \right]} =$$

$$= \frac{\sigma^2}{2\pi [(1-\rho^2)^2 + 4\rho^2 (\cos \psi - \cos \lambda)^2 - 4\rho (1-\rho^2) \cos \psi \cos \lambda]}.$$

Если ρ близко к единице, то $f(\lambda)$ принимает максимальное значение в окрестности $\lambda = \pm \psi$.

В общем случае спектральная плотность есть произведение выражений вида (44) и (47). Если корень $x_j = \rho_j e^{i\psi_j}$, причем ρ_j близко к 1, то функция $f(\lambda)$ принимает максимум при значении λ из окрестностей точек $\pm \psi_j$.

Спектральную плотность (42) можно записать в виде

$$(48) \quad f(\lambda) = \frac{\sigma^2}{2\pi \left| \sum_{r=0}^p \beta_r e^{i\lambda(p-r)} \right|^2} = \frac{\sigma^2}{2\pi \prod_{j=1}^p |e^{i\lambda} - x_j|^2}.$$

Каждый множитель $|e^{i\lambda} - x_j|$ в выражении (48) можно заменить на $|x_j| \cdot |1 - e^{i\lambda}/x_j| = |x_j| \cdot |e^{i\lambda} - 1/x_j|$. Существуют различные способы представления $f(\lambda)$, как и для процесса скользящего среднего; здесь, однако, нет корней, равных по абсолютной величине 1. Если все корни выражения (39) по абсолютной величине меньше 1, $|1/\bar{x}_j| > 1$, тогда только формула (48) является выражением, которому соответствует уравнение со всеми корнями, меньшими единицы по абсолютной величине.

7.5.4. Процессы авторегрессии с остатками в форме скользящего среднего

Предположим, что возмущение в стохастическом разностном уравнении является процессом скользящего среднего, а именно

$$(49) \quad \sum_{r=0}^p \beta_r y_{t-r} = \sum_{s=0}^q \alpha_s v_{t-s},$$

где $\mathbb{E}v_t = 0$ (для удобства), $\mathbb{E}v_t^2 = \sigma^2$, $\mathbb{E}v_t v_s = 0$, $t \neq s$, $\beta_0 = \alpha_0 = 1$. Если $f(\lambda)$ — спектральная плотность процесса $\{y_t\}$, то

$$(50) \quad \left| \sum_{r=0}^p \beta_r e^{i\lambda r} \right|^2 f(\lambda) = \sigma^2 \left| \sum_{s=0}^q \alpha_s e^{i\lambda s} \right|^2 / (2\pi),$$

так как $\sigma^2 / (2\pi)$ — спектральная плотность процесса $\{v_t\}$. Тогда спектральная плотность для $\{y_t\}$ представляется в виде

$$(51) \quad f(\lambda) = \sigma^2 \left| \sum_{s=0}^q \alpha_s e^{i\lambda(q-s)} \right|^2 \left/ \left(2\pi \left| \sum_{r=0}^p \beta_r e^{i\lambda(p-r)} \right|^2 \right) \right|.$$

Это выражение является рациональной функцией от $e^{i\lambda}$. Числитель и знаменатель можно записывать различными способами. Потребуем, чтобы корни полинома, соответствующего знаменателю, были меньше 1 по абсолютной величине, а корни полинома, соответствующего числителю, были не больше 1 по абсолютной величине. (Например, 1 может быть корнем числителя, но не знаменателя, если $\sigma^2 > 0$; см. теорему 5.2.2.) Мы также предположим, что эти два полинома не имеют общих корней. (Если существуют общие корни, то числитель и знаменатель можно сократить на соответствующие множители.)

Произвольную непрерывную спектральную плотность можно аппроксимировать рациональной спектральной плотностью. Действительно, любую непрерывную спектральную плотность можно аппроксимировать спектральной плотностью процесса скользящего среднего. Это следует из теоремы Вейерштрасса об аппроксимации тригонометрическими функциями. Эта теорема утверждает, что если $k(\lambda)$ — непрерывная функция на отрезке $[-\pi, \pi]$, причем $k(\pi) = k(-\pi)$, то для любого $\varepsilon > 0$ существует тригонометрический полином с действительными значениями

$$(52) \quad h_m(\lambda) = \sum_{j=-m}^m c_j e^{i\lambda j} = a_0 + \sum_{j=1}^m (a_j \cos \lambda j + b_j \sin \lambda j),$$

где коэффициенты $a_0, a_1, \dots, a_m, b_1, \dots, b_m$ действительны, такой, что

$$(53) \quad |k(\lambda) - h_m(\lambda)| < \varepsilon, \quad -\pi \leq \lambda \leq \pi.$$

[См., например, Лукач (1960, Приложение С).] Если $k(\lambda) = k(-\lambda)$, то $h_m(\lambda)$ можно считать симметричным (так, если бы аппроксимирующий полином $h_m^*(\lambda)$ не был симметричен, то можно было выбрать $h_m(\lambda) = \frac{1}{2} [h_m^*(\lambda) + h_m^*(-\lambda)]$). Тогда $b_1 = \dots = b_m = 0$, а $c_j = c_{-j}$ действительны, $j = 0, 1, 2, \dots, m$.

ТЕОРЕМА 7.5.2. Если спектральная плотность $f(\lambda)$ непрерывна, то для произвольного $\varepsilon > 0$ существует спектральная плотность

$$(54) \quad h_m(\lambda) = \sum_{j=-m}^m c_j e^{i\lambda j} = \sum_{j=-m}^m c_j \cos \lambda j,$$

где $c_j = c_{-j}$ действительны, $h_m(\lambda) > \varepsilon/2$ и $|f(\lambda) - h_m(\lambda)| < \varepsilon$, $-\pi \leq \lambda \leq \pi$.

Доказательство. Пусть $k(\lambda) = k(-\lambda) = f(\lambda) + 3\varepsilon/4$ и пусть $h_m(\lambda) = h_m(-\lambda)$ удовлетворяет теореме Вейерштрасса об аппроксимации с заменой в (53) ε на $\varepsilon/4$. Тогда

$$(55) \quad |f(\lambda) - h_m(\lambda)| \leq |f(\lambda) - k(\lambda)| + |k(\lambda) - h_m(\lambda)| < < \frac{3}{4}\varepsilon + \frac{1}{4}\varepsilon = \varepsilon, \quad -\pi \leq \lambda \leq \pi.$$

Следствие 7.5.1. Если спектральная плотность $f(\lambda)$ непрерывна, то для произвольного $\varepsilon > 0$ существует такой конечный процесс скользящего среднего (17) с положительной спектральной плотностью (18), обозначаемой $h_q(\lambda)$, что будет выполняться условие $|f(\lambda) - h_q(\lambda)| < \varepsilon$ для $-\pi \leq \lambda \leq \pi$.

Доказательство. В разд. 7.5.1 было показано, что произвольная ковариационная последовательность, которая имеет конечное число ненулевых ковариаций, является ковариационной последовательностью конечного процесса скользящего среднего. Используя теорему 7.5.2, произвольную непрерывную спектральную плотность можно аппроксимировать положительной спектральной плотностью (54). ■

Следствие 7.5.2. Если спектральная плотность $f(\lambda)$ непрерывна, то для произвольного $\varepsilon > 0$ существует такой процесс авторегрессии $\{y_t\}$, удовлетворяющий (38), со спектральной плотностью (42), обозначаемой $g_p(\lambda)$, что выполняется условие $|f(\lambda) - g_p(\lambda)| < \varepsilon$, когда $-\pi \leq \lambda \leq \pi$.

Доказательство. Пусть $f_c(\lambda) = f(\lambda)$, если $f(\lambda) \geq c$, и $f_c(\lambda) = c$, если $f(\lambda) < c$ ($-\pi \leq \lambda \leq \pi$), где $0 < c \leq \varepsilon/2$. Тогда $|f(\lambda) - f_c(\lambda)| \leq \varepsilon/2$. Так как плотность $f_c(\lambda)$ непрерывна и положительна на отрезке $[-\pi, \pi]$, то $1/f_c(\lambda) = g(\lambda)$ тоже непрерывна, положительна и ограничена и представляет собой спектральную плотность стационарного процесса. Из следствия 7.5.1 вытекает существование такой положительной спектральной плотности процесса скользящего среднего, обозначаемой $k_q(\lambda)$, что выполняется условие $|g(\lambda) - k_q(\lambda)| < \varepsilon'$ для $-\pi \leq \lambda \leq \pi$. Тогда выражение $1/k_q(\lambda) = g_p(\lambda)$ представляет собой спектральную плотность процесса авторегрессии с $p = q$ и

$$(56) \quad |f_c(\lambda) - g_p(\lambda)| = \left| \frac{1}{g(\lambda)} - \frac{1}{k_q(\lambda)} \right| = \frac{|g(\lambda) - k_q(\lambda)|}{g(\lambda) k_q(\lambda)} < \frac{1}{2} \varepsilon,$$

если

$$(57) \quad \varepsilon' \leq \frac{\frac{1}{2} \varepsilon \min g(\lambda)}{\max f(\lambda) + \frac{1}{2} \varepsilon}.$$

Это доказывает следствие.

Таким образом, непрерывную спектральную плотность можно сколь угодно точно аппроксимировать спектральной плотностью процесса скользящего среднего или спектральной плотностью процесса авторегрессии. Аналогично соответствующую ковариационную последовательность можно также аппроксимировать сколь угодно точно (см. упр. 28).

7.5.5. Линейные операции

Линейные операции, такие, как дифференцирование и сглаживание, были рассмотрены в гл. 3. Влияние этих операций на ковариационную структуру процесса можно изучать с помощью спектральных плотностей.

Следствие 7.5.3. Если $P(z) = \sum_{r=0}^q c_r z^r$, то спектральная плотность $h(\lambda)$ процесса

$$(58) \quad z_t = P(\mathcal{L}) y_t$$

представляет собой выражение

$$(59) \quad h(\lambda) = |P(e^{i\lambda})|^2 f(\lambda),$$

где $f(\lambda)$ — спектральная плотность процесса $\{y_t\}$ и $\mathcal{L}$ — оператор сдвига.

Рассмотрим процесс, полученный с помощью r -кратного применения разностного оператора $\Delta = \mathcal{P} - 1$ к процессу $\{y_t\}$, т. е.

$$(60) \quad z_t = \Delta^r y_t = (1 - \mathcal{L})^r y_{t+r} = (1 - \mathcal{L})^r \mathcal{P}^r y_t.$$

Тогда если $h(\lambda)$ — спектральная плотность процесса $\{z_t\}$, то

$$(61) \quad \frac{h(\lambda)}{f(\lambda)} = |1 - e^{i\lambda}|^{2r} = |e^{i\lambda/2} - e^{-i\lambda/2}|^{2r} = \\ = 2^{2r} \sin^{2r} \frac{\lambda}{2} = |e^{i\lambda} - 2 + e^{-i\lambda}|^r = 2^r (1 - \cos \lambda)^r.$$

Вычитание приводит к поднятию высоких частот.

Добавление соседних членов представляет собой оператор $\mathcal{P} + 1$; пусть эта операция повторяется r раз.

$$(62) \quad x_t = (1 + \mathcal{L})^r y_{t+r} = (1 + \mathcal{L})^r \mathcal{P}^r y_t.$$

Если $k(\lambda)$ — спектральная плотность процесса $\{x_t\}$, то

$$(63) \quad \frac{k(\lambda)}{f(\lambda)} = |1 + e^{i\lambda}|^{2r} = |e^{i\lambda/2} + e^{-i\lambda/2}|^{2r} = \\ = 2^{2r} \cos^{2r} \frac{1}{2} \lambda = |e^{i\lambda} + 2 + e^{-i\lambda}|^r = 2^r (1 + \cos \lambda)^r.$$

Суммирование приводит к поднятию низких частот.

Предположим, что и вычитание, и сглаживание выполнены. Если

$$(64) \quad z_t = (1 - \mathcal{L})^r (1 + \mathcal{L})^s y_t,$$

то

$$(65) \quad \frac{h(\lambda)}{f(\lambda)} = 2^{r+s} (1 - \cos \lambda)^r (1 + \cos \lambda)^s.$$

Обозначим это через $\varphi(\lambda)$. Отсюда

$$(66) \quad \frac{d}{d\lambda} \varphi(\lambda) = 2^{r+s} (1 - \cos \lambda)^{r-1} (1 + \cos \lambda)^{s-1} \sin \lambda \times \\ \times [r(1 + \cos \lambda) - s(1 - \cos \lambda)], \quad r \geq 1, \quad s \geq 1,$$

$$(67) \quad \frac{d^2}{d\lambda^2} \varphi(\lambda) = \\ = 2^{r+s} (1 - \cos \lambda)^{r-2} (1 + \cos \lambda)^{s-2} \sin^2 \lambda \{ [r(1 + \cos \lambda) - \\ - s(1 - \cos \lambda)]^2 - \cos \lambda [r(1 + \cos \lambda) - \\ - s(1 - \cos \lambda)] - (r + s) \sin^2 \lambda \}, \quad r \geq 2, \quad s \geq 2.$$

Первая производная равна нулю, когда

$$(68) \quad \cos \lambda = \frac{s-r}{s+r} = \frac{1-r/s}{1+r/s},$$

при этом значении $\lambda = \lambda_0$ вторая производная принимает вид

$$(69) \quad -2^{r+s} (1 - \cos \lambda_0)^{r-2} (1 + \cos \lambda_0)^{s-2} (r + s) \sin^4 \lambda_0 = \\ = -\frac{4^{r+s} r^r s^s}{(r+s)^{r+s-1}}.$$

Таким образом, $h(\lambda)/f(\lambda)$ имеет максимум в точке λ_0 . Остальные нулевые значения первой производной есть нули выражения $h(\lambda)/f(\lambda)$. Выбирая должным образом r и s , мы можем сделать λ_0 произвольно близким к любой заданной частоте. Если r и s большие, то максимум (65) будет четко выраженным. Тогда $h(\lambda)$ будет иметь мак-

симум в окрестности точки λ_0 и соответствующий процесс будет преимущественно волной косинуса с этой частотой.

Рассмотрим следующий пример скользящего среднего

$$(70) \quad z_t = \frac{1}{2m+1} \sum_{s=-m}^m y_{t+s}.$$

Тогда

$$\begin{aligned} (71) \quad \frac{h(\lambda)}{f(\lambda)} &= \frac{1}{(2m+1)^2} \left| \sum_{s=-m}^m e^{i\lambda s} \right|^2 = \frac{1}{(2m+1)^2} \left| e^{-i\lambda m} \sum_{s=0}^{2m} e^{i\lambda s} \right|^2 = \\ &= \frac{1}{(2m+1)^2} \left| \frac{1 - e^{i\lambda(2m+1)}}{1 - e^{i\lambda}} \right|^2 = \\ &= \frac{1}{(2m+1)^2} \left| \frac{e^{i\lambda(2m+1)/2} - e^{-i\lambda(2m+1)/2}}{e^{i\lambda/2} - e^{-i\lambda/2}} \right|^2 = \\ &= \left[\frac{\sin \lambda(m+1/2)}{(2m+1) \sin \lambda/2} \right]^2. \end{aligned}$$

Сглаживание необходимо для увеличения спектральной плотности вблизи $\lambda = 0$ (низкие частоты) и для уменьшения ее вблизи $\lambda = k2\pi/(2m+1)$, $k = 1, 2, \dots$.

7.6. ГИЛЬБЕРТОВО ПРОСТРАНСТВО И ТЕОРИЯ ПРОГНОЗИРОВАНИЯ

7.6.1. Гильбертово пространство

Материал последних двух параграфов станет более прозрачным, если ввести в рассмотрение гильбертово пространство. Конечномерное евклидово пространство есть гильбертово пространство; общая теория гильбертова пространства переносит многие идеи и результаты евклидовой геометрии на бесконечномерные пространства.

Для определения гильбертова пространства введем бесконечную последовательность случайных величин $\{y_t\}$, $t = \dots, -1, 0, 1, \dots$. Эту последовательность случайных величин будем считать случайным стационарным процессом в широком смысле. Для удобства положим $\mathcal{E}y_t = 0$: нас будут интересовать только моменты второго порядка. Пусть $\mathcal{E}y_t y_s = \sigma(t-s)$. Будем считать случайные переменные y_t , $t = \dots, -1, 0, 1, \dots$, элементами или точками нашего гильбертова пространства. Конечные линейные комбинации $\sum_{i \in S} c_i y_i$, где S есть конечное множество (положительных или отрицательных) целых чисел, также являются точками этого пространства. Наконец

добавим все случайные величины, которые являются пределами в среднеквадратичном последовательностей случайных величин, уже принадлежащих пространству.

Если x и y — два элемента гильбертова пространства, то $\mathcal{E}xy$ назовем *скалярным произведением*, $\sqrt{\mathcal{E}x^2}$ — *нормой* x , а $\sqrt{\mathcal{E}(x-y)^2}$ — *расстоянием* между x и y . Если $\mathcal{E}(x-y)^2 = 0$, то будем считать x и y совпадающими. Последовательность элементов $\{x^{(n)}, n = 1, 2, \dots\}$ *сходится* тогда и только тогда, когда она удовлетворяет *критерию Коши*, а именно: для любого заданного $\varepsilon > 0$ существует $N(\varepsilon)$, такое, что для $m > N(\varepsilon)$ и $n > N(\varepsilon)$ выполняется неравенство

$$(1) \quad \sqrt{\mathcal{E}(x^{(n)} - x^{(m)})^2} < \varepsilon.$$

Гильбертово пространство, порожденное последовательностью $\{y_i\}$, является *полным* в том смысле, что любая последовательность Коши имеет предел, принадлежащий этому пространству. Мы будем писать $x^{(n)} \rightarrow x$, подразумевая, что $\mathcal{E}(x^{(n)} - x)^2 \rightarrow 0$, когда $n \rightarrow \infty$.

ТЕОРЕМА 7.6.1. Если $\mathcal{E}(x^{(n)} - x^{(m)})^2 \rightarrow 0$ при $n, m \rightarrow \infty$, то существует случайная величина x , такая, что $\mathcal{E}(x^{(n)} - x)^2 \rightarrow 0$ при $n \rightarrow \infty$.

Доказательство. Из неравенства Чебышева следует, что $x^{(n)} - x^{(m)}$ сходится к 0 по вероятности при $n, m \rightarrow \infty$. Тогда существует последовательность n_k и случайная величина x , такие, что $x^{(n_k)}$ сходится к x с вероятностью 1 [Лозв (1963, § 6.3)]. Для каждого фиксированного m разность $x^{(n_k)} - x^{(m)}$ сходится к $x - x^{(m)}$ с вероятностью 1. Тогда из уравнения теоремы Фату — Лебега [Лозв (1963, § 7.2)] и из условия $\mathcal{E}(x^{(n_k)} - x^{(m)})^2 \rightarrow 0$ при $n_k, m \rightarrow \infty$ следует

$$(2) \quad \mathcal{E}(x - x^{(m)})^2 \leq \lim_{k \rightarrow \infty} \mathcal{E}(x^{(n_k)} - x^{(m)})^2 \rightarrow 0,$$

когда $m \rightarrow \infty$. ■

Полезным критерием для сходимости является

ТЕОРЕМА 7.6.2. Если $\mathcal{E}x^{(n)}x^{(m)} \rightarrow \sigma^2 < \infty$ при $n, m \rightarrow \infty$, то $\mathcal{E}(x^{(n)} - x^{(m)})^2 \rightarrow 0$.

Доказательство.

$$(3) \quad \mathcal{E}(x^{(n)} - x^{(m)})^2 = \mathcal{E}x^{(n)^2} + \mathcal{E}x^{(m)^2} - 2\mathcal{E}x^{(n)}x^{(m)} \rightarrow 0. \blacksquare$$

Следствие 7.6.1. Если $\mathcal{E}x^{(n)}x^{(m)} \rightarrow \sigma^2 < \infty$ при $n, m \rightarrow \infty$, то существует случайная величина x , такая, что $\mathcal{E}(x^{(n)} - x)^2 \rightarrow 0$ при $n \rightarrow \infty$.

В общем случае гильбертово пространство есть полное линейное векторное пространство со скалярным произведением. *Линейное векторное пространство* состоит из элементов (иногда говорят точек, векторов) $x, y, z, \dots$, для которых определены операция сложения, ставящая в соответствие двум элементам x и y третий элемент $x + y$, и операция умножения на действительное число α ¹⁾; эти операции удовлетворяют следующим условиям:

$$(4) \quad \begin{aligned} x + y &= y + x, & x + (y + z) &= (x + y) + z, \\ \alpha(x + y) &= \alpha x + \alpha y, & (\alpha + \beta)x &= \alpha x + \beta x, \\ \alpha(\beta x) &= (\alpha\beta)x, & 1x &= x. \end{aligned}$$

Соотношения (4) выполняются для любых элементов x, y, z векторного пространства и для любых действительных чисел α и β . Кроме того, существует единственный элемент 0, называемый нулем (не путайте с числом 0!), такой, что для любого x из векторного пространства $x + 0 = x$ и $0 \cdot x = 0$.

Скалярное произведение есть действительная функция (x, y) двух элементов, для которой выполняются следующие условия:

$$(5) \quad \begin{aligned} (x, y) &= (y, x), & (x + y, z) &= (x, z) + (y, z), \\ (\alpha x, y) &= \alpha(x, y), & (x, x) &\geq 0, \\ (x, x) &> 0 \quad \text{при } x \neq 0. \end{aligned}$$

В евклидовом пространстве скалярное произведение есть $\mathbf{x}'\mathbf{y}$. Для конечных линейных комбинаций множества случайных величин и их пределов скалярное произведение есть $\mathfrak{E}xy$.

ТЕОРЕМА 7.6.3. (Неравенство Коши — Шварца.)

$$(6) \quad (x, y) \leq \sqrt{(x, x)} \sqrt{(y, y)}.$$

ДОКАЗАТЕЛЬСТВО. Теорема следует из неравенства

$$(7) \quad 0 \leq (\alpha x + \beta y, \alpha x + \beta y) = \alpha^2(x, x) + 2\alpha\beta(x, y) + \beta^2(y, y)$$

при $\alpha = \sqrt{(y, y)}$ и $\beta = -\sqrt{(x, x)}$. ■

Норма элемента определяется при помощи скалярного произведения как $\sqrt{(x, x)}$; часто ее записывают в виде $\|x\|$. Она удовлетворяет следующим условиям:

$$(8) \quad \begin{aligned} \|x\| &\geq 0, & \|x\| &> 0 \quad \text{при } x \neq 0, \\ \|\alpha x\| &= |\alpha| \|x\|, & \|x + y\| &\leq \|x\| + \|y\|. \end{aligned}$$

¹⁾ В комплексном гильбертовом пространстве — соответственно комплексное число.

Последнее неравенство (неравенство треугольника) следует из неравенства Коши — Шварца:

$$(9) \quad \begin{aligned} \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2(x, y) \leq \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = (\|x\| + \|y\|)^2. \end{aligned}$$

Норма $\|x - y\|$ задает *расстояние* между x и y и называется *метрикой* пространства.

Последовательность $\{x^{(n)}\}$ называется *последовательностью Коши*, если

$$(10) \quad \lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} \|x^{(n)} - x^{(m)}\| = 0.$$

Элемент x есть предел последовательности $\{x^{(n)}\}$, если

$$(11) \quad \lim_{n \rightarrow \infty} \|x^{(n)} - x\| = 0.$$

Нормированное линейное векторное пространство *полно*, если в нем существует предел любой последовательности Коши, принадлежащей пространству.

Для любого гильбертова пространства скалярное произведение двух элементов является непрерывной функцией этих элементов.

ТЕОРЕМА 7.6.4. Если $x^{(n)} \rightarrow x$ и $y^{(n)} \rightarrow y$ при $n \rightarrow \infty$, то

$$(12) \quad (x^{(n)}, y^{(n)}) \rightarrow (x, y).$$

ДОКАЗАТЕЛЬСТВО. Из неравенства Коши — Шварца имеем

$$(13) \quad \begin{aligned} |(x^{(n)}, y^{(n)}) - (x, y)| &= |(x^{(n)} - x, y^{(n)} - y) + \\ &+ (x, y^{(n)} - y) + (x^{(n)} - x, y)| \leq \\ &\leq |(x^{(n)} - x, y^{(n)} - y)| + \\ &+ |(x, y^{(n)} - y)| + |(x^{(n)} - x, y)| \leq \\ &\leq \|x^{(n)} - x\| \cdot \|y^{(n)} - y\| + \\ &+ \|x\| \cdot \|y^{(n)} - y\| + \|x^{(n)} - x\| \cdot \|y\|. \end{aligned}$$

Правая часть в (13) сходится к 0. ■

СЛЕДСТВИЕ 7.6.2. Если $x^{(n)} \rightarrow x$, то $\|x^{(n)}\| \rightarrow \|x\|$.

Если в гильбертовом пространстве, порожденном последовательностью $\{y_t\}$, скалярное произведение (x, y) определяется как $\mathfrak{E}xy$, то теорема 7.6.4 подтверждает существование пределов дисперсий и ковариаций в § 7.4 и 7.5.

7.6.2. Проекция и линейное прогнозирование

Мы будем рассматривать прогнозирование случайной величины y_t по известным $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ (части прошлого) или по $y_{t-1}, y_{t-2}, \dots$ (всему прошлому) для случайного процесса $\{y_t\}$ со средним значением $\mathbb{E}y_t = 0$ и ковариационной функцией $\mathbb{E}y_t y_s = \sigma(t-s)$. Пусть $\hat{y}_t$ — прогноз; в качестве критерия наилучшего прогноза возьмем $\mathbb{E}(y_t - \hat{y}_t)^2$. Среднеквадратичная ошибка прогноза будет минимальна, если $\hat{y}_t$ взято как условное математическое ожидание величины y_t при заданных $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ или $y_{t-1}, y_{t-2}, \dots$. Это условное математическое ожидание будет, конечно, зависеть от распределения величин $y_t, y_{t-1}, \dots, y_{t-p}$ или вероятностной меры для $y_t, y_{t-1}, \dots$. Если процесс гауссовский, то условное математическое ожидание будет линейной функцией от $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ или от $y_{t-1}, y_{t-2}, \dots$; в других примерах оно может и не быть линейной функцией.

Будем считать $\hat{y}_t$ линейной функцией, $\hat{y}_t = \sum_{s=1}^p c_s y_{t-s}$, в случае прогнозирования по $y_{t-1}, \dots, y_{t-p}$, или пределом $\hat{y}_t$ таких линейных функций. В таком случае $\mathbb{E}(y_t - \hat{y}_t)^2$ зависит только от $\{\sigma(h)\}$. Эту проблему удобно сформулировать в терминах геометрии гильбертова пространства.

Линейное многообразие есть непустое подмножество гильбертова пространства, такое, что $\alpha x + \beta y$ для каждого действительных α и β принадлежит подмножеству, если x и y принадлежат этому подмножеству. Линейное многообразие *замкнуто*, если оно содержит предел каждой последовательности Коши, содержащейся в нем. Конечное или бесконечное множество точек гильбертова пространства может порождать замкнутое линейное многообразие, которое состоит из всевозможных конечных линейных комбинаций точек этого множества и их пределов. Говорят, что замкнутое линейное многообразие *натяннуто* на исходное множество точек этого пространства. Замкнутое линейное многообразие является гильбертовым пространством и называется *подпространством*.

Так как математическое ожидание $\mathbb{E}(y_t - \hat{y}_t)^2$ есть $\|y_t - \hat{y}_t\|^2$, то нахождение наилучшего линейного прогноза для y_t сводится к нахождению точки линейного многообразия, натянутого на множество величин y_s ($s < t$), которая наиболее близка (по расстоянию, определяемому нормой) к точке y_t . В § 2.2 мы уже рассматривали этот вопрос для T -мерного евклидова пространства с нормой $\sqrt{x^T x}$. Вектор линейного многообразия, порожденного первыми p координатными векторами, ближайший к y , есть определяемый единственным образом вектор z , такой, что $y - z$ ортогонален этим

p координатным векторам (и, следовательно, ортогонален каждому вектору линейного многообразия). Покажем теперь, что это решение имеет место и в общем случае гильбертова пространства.

Используя понятие скалярного произведения (x, y) , два элемента x и y будем называть *ортогональными*, если $(x, y) = 0$. Это можно записать так: $x \perp y$.

ТЕОРЕМА 7.6.5. (Теорема о проекции.) *Если M — подпространство гильбертова пространства $\mathcal{H}$ и y — элемент в $\mathcal{H}$, то существует единственный элемент x в M , такой, что $\|y - x\| = \min_{z \in M} \|y - z\|$. Если $y \notin M$, то $y - x$ ортогонален каждому $z \in M$.*

Доказательство. Пусть $\min_{z \in M} \|y - z\| = d$ и $\{x^{(n)}\}$ — последовательность, удовлетворяющая условию $\|y - x^{(n)}\| \rightarrow d$. (Такая последовательность существует по определению минимума.) Доказав, что последовательность $\{x^{(n)}\}$ является последовательностью Коши, получим существование элемента x . Имеем

$$(14) \quad \|(y - x^{(n)}) - (y - x^{(m)})\|^2 + \|(y - x^{(n)}) + (y - x^{(m)})\|^2 = \\ = 2\|y - x^{(n)}\|^2 + 2\|y - x^{(m)}\|^2.$$

Перепишем левую часть:

$$(15) \quad \|x^{(m)} - x^{(n)}\|^2 + 4\left\|y - \frac{1}{2}(x^{(n)} + x^{(m)})\right\|^2;$$

так как $(x^{(n)} + x^{(m)})/2 \in M$, то второе слагаемое в (15) не менее $4d^2$. Таким образом, из (14) следует

$$(16) \quad \|x^{(m)} - x^{(n)}\|^2 \leq 2\|y - x^{(n)}\|^2 + 2\|y - x^{(m)}\|^2 - 4d^2,$$

а правая часть (16) стремится к 0. Значит, $\{x^{(n)}\}$ — последовательность Коши и существует предел x этой последовательности. По следствию 7.6.2 $\|y - x\| = d$.

Для любого элемента $z \in M$

$$(17) \quad d^2 \leq \|y - x - \alpha z\|^2 = \|y - x\|^2 - 2(y - x, \alpha z) + \|\alpha z\|^2 = \\ = d^2 - 2\alpha(y - x, z) + \alpha^2\|z\|^2.$$

Если $y \notin M$ и, следовательно, $y - x \neq 0$, то неравенство выполняется для любого α , если только $(y - x, z) = 0$, т. е. если $y - x$ ортогонален z .

Если существует другой элемент x' , такой, что $\|y - x'\|^2 = d^2$, то $(y - x', z) = 0$ для каждого $z \in M$. Для каждого $z = x - x'$ будем иметь $0 = (y - x', x - x') - (y - x, x - x') = (x - x', x - x') = \|x - x'\|^2$, откуда следует, что $x = x'$. ■

Точка x называется *проекцией* y на подпространство M , а $y - x$ называется его *остатком* по отношению к M .

Если M — подпространство гильбертова пространства $\mathcal{H}$, то множество всех элементов, ортогональных M , есть тоже подпространство, обозначаемое $M^\perp$. (Это подпространство замкнуто, что следует из непрерывности скалярного произведения и полноты $\mathcal{H}$.) Каждый элемент y из $\mathcal{H}$ может быть записан как $x + z$, где $x \in M$, а $z \in M^\perp$. Это можно записать еще как $\mathcal{H} = M \oplus M^\perp$.

В терминах гильбертова пространства линейный прогноз есть просто проекция. Гильбертово пространство $\mathcal{H}$ есть замкнутое многообразие, натянутое на $y_t, y_{t-1}, \dots$, а скалярное произведение есть $\mathcal{E}xy$. Пусть $M_{t-1,p}$ есть подпространство, натянутое на $y_{t-1}, \dots, y_{t-p}$. Проекция элемента y_t на $M_{t-1,p}$ является лучшим линейным прогнозом для y_t . Пусть это будет $\hat{y}_{t,p}$; оно минимизирует выражение

$$(18) \quad \mathcal{E} \left(y_t - \sum_{j=1}^p c_j y_{t-j} \right)^2 = \sigma(0) - 2 \sum_{j=1}^p c_j \sigma(j) + \sum_{i,j=1}^p c_i c_j \sigma(j-i).$$

Это аналогично рассмотренной в гл. 2 задаче. Решение дается нормальными уравнениями

$$(19) \quad \begin{bmatrix} \sigma(0) & \sigma(1) & \dots & \sigma(p-1) \\ \sigma(1) & \sigma(0) & \dots & \sigma(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma(p-1) & \sigma(p-2) & \dots & \sigma(0) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix} = \begin{bmatrix} \sigma(1) \\ \sigma(2) \\ \vdots \\ \sigma(p) \end{bmatrix}.$$

Пусть M_{t-1} — замкнутое линейное многообразие, натянутое на $y_{t-1}, y_{t-2}, \dots$. Наилучший линейный прогноз элемента y_t по $y_{t-1}, y_{t-2}, \dots$ есть его проекция на M_{t-1} , которую мы обозначим $\hat{y}_t$.

ТЕОРЕМА 7.6.6.

$$(20) \quad \lim_{p \rightarrow \infty} \mathcal{E}(\hat{y}_t - \hat{y}_{t,p})^2 = 0.$$

ДОКАЗАТЕЛЬСТВО. Пусть $\mathcal{E}(y_t - \hat{y}_{t,p})^2 = d_p$, и пусть $\mathcal{E}(y_t - \hat{y}_t)^2 = d$. Так как $M_{t-1,p} \subset M_{t-1,p+1}$ и $M_{t-1} = \bigcup_{p=1}^{\infty} M_{t-1,p}$, то существует такая последовательность $\{x^{(p)}\}$ с $x^{(p)} \in M_{t-1,p}$, что $\mathcal{E}(x^{(p)} - \hat{y}_t)^2 \rightarrow 0$, когда $p \rightarrow \infty$. Так как $\mathcal{E}(x^{(p)} - y_t)^2 \geq d_p \geq d$ и

$$(21) \quad \mathcal{E}(x^{(p)} - y_t)^2 \leq \mathcal{E}(x^{(p)} - \hat{y}_t)^2 + \mathcal{E}(\hat{y}_t - y_t)^2,$$

а последнее стремится к d , то имеем

$$(22) \quad \lim_{p \rightarrow \infty} \mathcal{E}(x^{(p)} - y_t)^2 = \lim_{p \rightarrow \infty} d_p = d.$$

Можно также написать

$$(23) \quad \mathbb{E} (x^{(p)} - y_t)^2 = \mathbb{E} (x^{(p)} - \hat{y}_{t,p})^2 + \mathbb{E} (\hat{y}_{t,p} - y_t)^2,$$

потому что элемент $y_t - \hat{y}_{t,p}$ ортогонален каждому вектору в $M_{t,p}$. Из этого следует, что $\mathbb{E} (x^{(p)} - \hat{y}_{t,p})^2 \rightarrow 0$, в то время как другие два выражения в (23) стремятся к d . Теорема следует из

$$(24) \quad \mathbb{E} (\hat{y}_t - \hat{y}_{t,p})^2 \leq \mathbb{E} (\hat{y}_t - x^{(p)})^2 + \mathbb{E} (x^{(p)} - \hat{y}_{t,p})^2. \blacksquare$$

Эта теорема показывает, что прогноз по бесконечному прошлому можно аппроксимировать прогнозом по конечному отрезку прошлого. На практике статистики должны пользоваться столь большой выборкой, чтобы достаточно хорошо оценить ковариации процесса и на основании их дать наилучший прогноз.

Дисперсией остатка может быть любое неотрицательное число, не превосходящее $\sigma^2(0)$. Если $\mathbb{E} (y_t - \hat{y}_t)^2 = \sigma^2 = 0$, то процесс называется *детерминированным*. Это означает, что y_t приближается элементом из M_{t-1} . В большинстве случаев этот элемент представляется как бесконечная линейная комбинация $\sum_{s=1}^{\infty} c_s y_{t-s}$.

Если $\mathbb{E} (y_t - \hat{y}_t)^2 = \sigma^2 > 0$, то процесс называется *регулярным*.

Так как $y_t - \hat{y}_t$ ортогонален каждому вектору в M_{t-1} , то можно написать

$$(25) \quad y_t = \hat{y}_t + u_t,$$

где $u_t = y_t - \hat{y}_t$ некоррелировано с $y_{t-1}, y_{t-2}, \dots$. Случайную величину u_t можно назвать *обновлением* или *возмущением*.

7.6.3. Разложение Вольда

Так называемое вольдовское разложение [Вольд (1954)] проявляет структуру стационарного процесса.

ТЕОРЕМА 7.6.7. Если $\{y_t\}$ — регулярный стационарный случайный процесс с $\mathbb{E} y_t = 0$, его можно представить в виде

$$(26) \quad y_t = \sum_{s=0}^{\infty} \gamma_s u_{t-s} + v_t,$$

где $\sum_{s=0}^{\infty} \gamma_s^2 < \infty$, $\gamma_0 = 1$, $\mathbb{E} u_s = \mathbb{E} v_s = 0$, $\mathbb{E} u_s^2 = \sigma^2$, $\mathbb{E} u_s u_t = 0$, $s \neq t$,

и $\mathbb{E} u_s v_t = 0$ для $u_s \in M_s$ и $v_t \in M_{-\infty} = \bigcap_{s=0}^{\infty} M_{t-s}$. Последовательности $\{\gamma_s\}$ и $\{u_s\}$ определяются единственным образом.

Доказательство. Пусть $u_s = y_s - \hat{y}_s$, $s = t, t-1, \dots$, $\sigma^2 = \mathcal{E}u_s^2$. Так как элемент u_s ортогонален $\mathcal{M}_{s-1}$, то $\mathcal{E}u_s u_r = 0$, $r < s$. Пусть $\gamma_s = \mathcal{E}y_t u_{t-s} / \sigma^2$, $s = 1, 2, \dots$. Тогда

$$(27) \quad 0 \leq \mathcal{E} \left(y_t - \sum_{s=0}^m \gamma_s u_{t-s} \right)^2 = \mathcal{E}y_t^2 - \sigma^2 \sum_{s=0}^m \gamma_s^2,$$

откуда $\sum_{s=0}^{\infty} \gamma_s^2 < \infty$. Поэтому $\sum_{s=0}^{\infty} \gamma_s u_{t-s}$ сходится в среднем и лежит внутри подпространства, порожденного $u_t, u_{t-1}, \dots$. Определим v_t следующим образом:

$$(28) \quad v_t = y_t - \sum_{s=0}^{\infty} \gamma_s u_{t-s}.$$

Тогда $\mathcal{E}v_t u_r = \mathcal{E}y_t u_r - \sigma^2 \gamma_{t-r} = 0$ для $r \leq t$; $\mathcal{E}v_t u_r = 0$ для $r > t$, потому что u_r ортогонален всем элементам из $\mathcal{M}_t$ и $v_t \in \mathcal{M}_t$. Так как элемент v_t ортогонален u_t , то он принадлежит $\mathcal{M}_{t-1}$. По индукции $v_t \in \mathcal{M}_s$ для $s \leq t$ и, следовательно, v_t принадлежит $\mathcal{M}_{-\infty} = \bigcap_{s=0}^{\infty} \mathcal{M}_{t-s}$. Решение единственно, потому что $u_s \in \mathcal{M}_s$, элемент u_s ортогонален $\mathcal{M}_{s-1}$ и $\gamma_s = \mathcal{E}y_t u_{t-s} / \sigma^2$. ■

ТЕОРЕМА 7.6.8. Процесс $\{v_t\}$, определенный в теореме 7.6.7, детерминированный, а процесс $\{w_t\}$, где $w_t = \sum_{s=0}^{\infty} \gamma_s u_{t-s}$, регулярный.

Доказательство. Для доказательства детерминированности процесса $\{v_t\}$ нужно показать, что замкнутое линейное множество, которое натянуто на $v_t, v_{t-1}, \dots$ (обозначим его $\mathcal{M}_{t0}$) есть $\mathcal{M}_{-\infty}$, т. е. нужно показать, что $\mathcal{M}_{t0} = \mathcal{M}_{t-1,0} = \dots = \mathcal{M}_{-\infty}$. Тогда v_t есть линейная комбинация (может быть бесконечная) элементов $v_{t-1}, v_{t-2}, \dots$.

ЛЕММА 7.6.1.

$$(29) \quad \mathcal{M}_{-\infty} = \mathcal{M}_u^{\perp}.$$

Доказательство. Здесь $\mathcal{M}_u^{\perp}$ является подпространством, ортогональным $\mathcal{M}_u$, порожденному $\{u_t\}$. Если $x \in \mathcal{M}_{-\infty} = \bigcap_{t=-\infty}^{\infty} \mathcal{M}_t$, то $x \in \mathcal{M}_t$ и, следовательно, x ортогонален u_{t+1} для каждого t , т. е. $x \in \mathcal{M}_u^{\perp}$. Обратно, пусть $x \in \mathcal{M}_u^{\perp}$. Так как $\mathcal{H} = \bigcup_{t=-\infty}^{\infty} \mathcal{M}_t$, то $x \in \mathcal{M}_t$ хотя бы для одного t . Так как $x \perp u_t$, то $x \in \mathcal{M}_{t-1}$ и по

индукции $x \in \mathcal{M}_s$ для $s \leq t$. Более того, $x \in \mathcal{M}_s$ для $s > t$, так как $\mathcal{M}_t \subseteq \mathcal{M}_s$. Лемма доказана. ■

ЛЕММА 7.6.2.

$$(30) \quad \mathcal{M}_{tu} = \mathcal{M}_{tw}.$$

Доказательство. Так как $w_t = \sum_{s=0}^{\infty} \gamma_s u_{t-s}$, то подпространство $\mathcal{M}_{tw}$, натянутое на $w_t, w_{t-1}, \dots$, содержится в подпространстве $\mathcal{M}_{tu}$, натянутом на $u_t, u_{t-1}, \dots$. Обратно, $u_t \in \mathcal{M}_t = \mathcal{M}_{tw} \oplus \mathcal{M}_{tv}$ и $u_t \perp \mathcal{M}_{tv}$, поэтому $u_t \in \mathcal{M}_{tw}$. Это доказывает лемму. ■

Закончим теперь доказательство теоремы 7.6.8. Так как $v_s \in \mathcal{M}_{-\infty}$ для каждого s , то $\mathcal{M}_{tv} \subseteq \mathcal{M}_{-\infty}$. Поскольку $\mathcal{M}_{-\infty} = \bigcap_{t=-\infty}^{\infty} \mathcal{M}_t$, то из $x \in \mathcal{M}_{-\infty}$ следует, что $x \in \mathcal{M}_t$ ввиду того, что $x \perp \mathcal{M}_{tu} = \mathcal{M}_{tw}$, $x \in \mathcal{M}_{tv}$. Из этого следует, что $\mathcal{M}_{tv} = \mathcal{M}_{-\infty}$ и что $\{v_t\}$ — детерминированный процесс.

Так как $w_t = u_t + \sum_{s=1}^{\infty} \gamma_s u_{t-s}$ и $u_t \perp \sum_{s=1}^{\infty} \gamma_s u_{t-s} \in \mathcal{M}_{t-1,u}$, то $\mathcal{E}(w_t - \hat{w}_t)^2 = \sigma^2 > 0$. Таким образом, $\{w_t\}$ — регулярный процесс, что и требовалось доказать. ■

Процесс $\{v_t\}$ часто называют чисто детерминированным. Процесс $\{w_t\}$ называют *чисто недетерминированным*, так как он не имеет детерминированной компоненты.

Так как $y_t = u_t + \sum_{s=1}^{\infty} \gamma_s u_{t-s} + v_t$ и $u_t \perp \left(\sum_{s=1}^{\infty} \gamma_s u_{t-s} + v_t \right) \in \mathcal{M}_{t-1}$, то $\hat{y}_t = \sum_{s=1}^{\infty} \gamma_s u_{t-s} + v_t$ есть проекция на $\mathcal{M}_{t-1}$. Эта проекция есть

наилучший линейный прогноз для y_t , и $\mathcal{E}(y_t - \hat{y}_t)^2 = \mathcal{E}u_t^2 = \sigma^2$. По лемме 7.6.2 $\hat{y}_t \in \mathcal{M}_{t-1,w} \oplus \mathcal{M}_{t-1,v} = \mathcal{M}_{t-1}$ и, следовательно, $\hat{y}_t$ является линейной функцией от $y_{t-1}, y_{t-2}, \dots$ или их пределом.

Спектральная плотность процесса $\{w_t\}$ есть $\sigma^2 \left| \sum_{s=0}^{\infty} \gamma_s e^{i\lambda s} \right|^2 / (2\pi)$. Процесс $\{v_t\}$ не имеет абсолютно непрерывной компоненты, и его спектральная функция состоит из ступенчатой и сингулярной компонент. Мы не будем доказывать последнее утверждение. [См. Дуб (1953, гл. XII, разд. 4).]

Колмогоров (1941b) показал, что дисперсия величины u_t (ошибка линейного прогноза) есть

$$(31) \quad \sigma^2 = 2\pi \exp \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \log F'(\lambda) d\lambda \right].$$

Действительно, $\{y_t\}$ будет регулярным процессом тогда и только тогда, когда $F'(\lambda) > 0$ почти всюду ($-\pi \leq \lambda \leq \pi$) и

$$(32) \quad \int_{-\pi}^{\pi} \log F'(\lambda) d\lambda > -\infty.$$

Разложение функции $\frac{1}{2} \log F'(\lambda) = \log \sqrt{F'(\lambda)}$ в ряд Фурье

$\sum_{n=-\infty}^{\infty} a_n e^{i\lambda n}$, где $a_{-n} = a_n$, действительно (так как $\log F'(\lambda)$ есть действительная и симметричная функция); тогда

$$(33) \quad \Gamma(z) = \sum_{s=0}^{\infty} \gamma_s z^s = \exp \left[a_0 + 2 \sum_{n=1}^{\infty} a_n z^n \right]$$

и $F'(\lambda) = |\Gamma(e^{-i\lambda})|^2$.

Если процесс $\{y_t\}$ имеет спектральную плотность $f(\lambda) = F'(\lambda)$ и если (32) выполняется, то

$$(34) \quad y_t = \sum_{s=0}^{\infty} \gamma_s u_{t-s}$$

с $\gamma_0 = 1$ и $\sum_{s=0}^{\infty} \gamma_s^2 < \infty$. Тогда наилучшим линейным прогнозом для

y_t по $u_{t-1}, u_{t-2}, \dots$ будет $\sum_{s=1}^{\infty} \gamma_s u_{t-s}$. Выясним теперь, справедливо ли соотношение

$$(35) \quad u_t = \sum_{s=0}^{\infty} \beta_s y_{t-s}$$

с $\beta_0 = 1$. Если оно справедливо, то наилучшим линейным прогнозом для y_t будет $-\sum_{s=1}^{\infty} \beta_s y_{t-s}$. Если представление (35) существует, то можно записать

$$(36) \quad u_t = \sum_{s=0}^{\infty} \beta_s \sum_{r=0}^{\infty} \gamma_r u_{t-r-s} = \sum_{q=0}^{\infty} \sum_{r=0}^q \beta_{q-r} \gamma_r u_{t-q}.$$

Здесь $\beta_0 \gamma_0 = 1$ и

$$(37) \quad \sum_{r=0}^q \beta_{q-r} \gamma_r = 0, \quad q = 1, 2, \dots$$

Для большей точности предположим, что функция $\Gamma(z) = \sum_{s=0}^{\infty} \gamma_s z^s$ аналитическая и не равна нулю при $|z| \leq 1$. Тогда

$$(38) \quad \frac{1}{\Gamma(z)} = B(z) = \sum_{r=0}^{\infty} \beta_r z^r$$

есть аналитическая функция, не равная нулю при $|z| \leq 1$. Отсюда

$$(39) \quad \mathbb{E} \left(\sum_{s=0}^{\infty} \beta_s y_{t-s} \right)^2 = \int_{-\pi}^{\pi} |B(e^{-i\lambda})|^2 dF(\lambda)$$

существует и определено, так как $|B(e^{-i\lambda})|$ ограничено. Следовательно, (35) существует как предел в среднеквадратичном. Если

$$(40) \quad \begin{aligned} u_t &= \int_0^{\pi} \cos \lambda t dC_u(\lambda) + \int_0^{\pi} \sin \lambda t dS_u(\lambda) = \\ &= \mathcal{R} \int_0^{\pi} e^{i\lambda t} d[C_u(\lambda) - iS_u(\lambda)], \end{aligned}$$

где $\mathbb{E} G_u^2(\lambda) = \mathbb{E} S_u^2(\lambda) = \sigma^2 \lambda / \pi$, то пусть

$$(41) \quad y_t = \mathcal{R} \int_0^{\pi} e^{i\lambda t} B^{-1}(e^{-i\lambda}) d[C_u(\lambda) - iS_u(\lambda)].$$

Тогда

$$(42) \quad \begin{aligned} \sum_{s=0}^{\infty} \beta_s y_{t-s} &= \mathcal{R} \int_0^{\pi} \sum_{s=0}^{\infty} \beta_s e^{i\lambda(t-s)} B^{-1}(e^{-i\lambda}) d[C_u(\lambda) - iS_u(\lambda)] = \\ &= \mathcal{R} \int_0^{\pi} e^{i\lambda t} d[C_u(\lambda) - iS_u(\lambda)]. \end{aligned}$$

Остается показать, что спектральная плотность для (41) есть

$$(43) \quad \frac{\sigma^2}{2\pi} |B^{-1}(e^{-i\lambda})|^2 = f(\lambda).$$

В большинстве случаев задачу прогнозирования можно рассматривать как нахождение $b_1, b_2, \dots$, которые минимизируют

$$(44) \quad \mathbb{E} \left(y_t - \sum_{s=1}^{\infty} b_s y_{t-s} \right)^2.$$

Нормальные уравнения будут иметь вид

$$(45) \quad \sum_{s=1}^{\infty} \sigma(r-s) b_s = \sigma(r), \quad r = 1, 2, \dots$$

Они эквивалентны

$$(46) \quad \int_{-\pi}^{\pi} e^{i\lambda r} f(\lambda) \left(\sum_{s=1}^{\infty} b_s e^{-i\lambda s} - 1 \right) d\lambda = 0, \quad r = 1, 2, \dots$$

Таким образом, мы хотим найти такую функцию $B(e^{-i\lambda})$

$$(47) \quad B(z) = \sum_{s=1}^{\infty} b_s z^s,$$

что положительные коэффициенты Фурье выражения $f(\lambda) \times \times [B(e^{-i\lambda}) - 1]$ равны нулю.

Более определенные результаты получены Акутовицем (1957), их мы приводим без доказательства.

ТЕОРЕМА 7.6.9. *Если $f(\lambda)$ ограничена, то (35) справедливо с ограничением $\sum_{s=0}^{\infty} |\beta_s|^2 < \infty$ тогда и только тогда, когда функция $\Gamma(z) = \sum_{s=0}^{\infty} \gamma_s z^s$, $|z| < 1$, такова, что интеграл*

$$(48) \quad \int_0^{2\pi} \frac{1}{|\Gamma(\rho e^{i\lambda})|^2} d\lambda$$

ограничен при $\rho \rightarrow 1$.

Говорят, что функция $1/\Gamma(z)$, регулярная в круге $|z| < 1$, принадлежит классу Харди H^2 , если интеграл (48) ограничен при $\rho \rightarrow 1$. [См. Зигмунд, (1959, т. I, гл. VII, разд. 7); см. также Гренандер и Розенблатт (1957, стр. 288).]

ТЕОРЕМА 7.6.10. *В условиях теоремы 7.6.9 наилучшим линейным прогнозом будет*

$$(49) \quad - \sum_{r=0}^{\infty} \beta_r y_{t-r}.$$

Теория прогнозирования развивалась независимо Колмогоровым (1941a), (1941b) и Винером (1949) (первоначальная публикация ограниченного распространения — в 1942 г.). Парзен (1961a) использовал подход, основанный на теории гильбертовых пространств.

7.7. НЕКОТОРЫЕ ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

Распределения многих статистик, представляющих интерес для анализа временных рядов, слишком усложнены, чтобы ими можно было пользоваться непосредственно. Однако в ряде случаев можно получать предельные распределения, которые используются как аппроксимации к точным распределениям, когда число наблюдений T велико. В этом параграфе мы рассмотрим несколько предельных теорем довольно общего характера.

Метод нахождения предельного распределения последовательности случайных величин заключается в нахождении предельного распределения аппроксимирующей последовательности (как в § 5.5).

ТЕОРЕМА 7.7.1. Пусть

$$(1) \quad S_T = Z_{kT} + X_{kT}, \quad T = 1, 2, \dots, \quad k = 1, 2, \dots$$

Предположим, что для любых $\delta > 0$ и $\varepsilon > 0$ существует k_0 , такое, что при $k > k_0$ выполняется

$$(2) \quad \text{Pr} \{ |X_{kT}| > \delta \} < \varepsilon$$

для всех T , и пусть

$$(3) \quad \text{Pr} \{ Z_{kT} \leq z \} = F_{kT}(z) \rightarrow F_k(z)$$

при $T \rightarrow \infty$, а

$$(4) \quad \lim_{k \rightarrow \infty} F_k(z) = F(z)$$

в каждой точке непрерывности функции $F(z)$. Тогда

$$(5) \quad \lim_{T \rightarrow \infty} \text{Pr} \{ S_T \leq z \} = F(z)$$

в каждой точке непрерывности $F(z)$.

ДОКАЗАТЕЛЬСТВО. Из условия (2) следует, что X_{kT} сходится по вероятности к 0 равномерно по T при $k \rightarrow \infty$. Так как функция $F(z)$ непрерывна в точке z , то для любого $\varepsilon > 0$ существует $\delta > 0$, такое, что

$$(6) \quad |F(z \pm \delta) - F(z)| < \frac{1}{2} \varepsilon,$$

и такое, что $z - \delta$ и $z + \delta$ суть точки непрерывности $F(z)$. Тогда найдется такое k_1 , что для $k > k_1$

$$(7) \quad |F_k(z \pm \delta) - F(z)| < \varepsilon.$$

Выберем δ так, чтобы выполнялось (6), и возьмем k таким, чтобы $k > k_0$ и $k > k_1$; зафиксируем δ и k . Тогда для $T > T_k$ (зависящего от ε)

$$(8) \quad |F_{kT}(z) - F_k(z)| < \varepsilon,$$

$$(9) \quad |F_{kT}(z + \delta) - F_k(z + \delta)| < \varepsilon,$$

$$(10) \quad |F_{kT}(z - \delta) - F_k(z - \delta)| < \varepsilon.$$

Теперь

$$\begin{aligned} (11) \quad \text{Pr} \{ S_T \leq z \} &= \text{Pr} \{ Z_{kT} + X_{kT} \leq z \} \leq \\ &\leq \text{Pr} \{ Z_{kT} \leq z + \delta, |X_{kT}| \leq \delta \} + \text{Pr} \{ |X_{kT}| > \delta \} \leq \\ &\leq F_{kT}(z + \delta) + \varepsilon. \end{aligned}$$

Таким образом,

$$(12) \quad \text{Pr} \{S_T \leq z\} \leq F(z) + 3\varepsilon.$$

Аналогично получаем

$$(13) \quad \text{Pr} \{S_T \leq z\} \geq F(z) - 3\varepsilon.$$

Это и доказывает теорему. ■

Следствие 7.7.1. Пусть для последовательности $\{M_k\}$

$$(14) \quad \mathbb{E} X_{kT}^2 \leq M_k,$$

$$(15) \quad \lim_{k \rightarrow \infty} M_k = 0$$

и, кроме того, имеет место (3) и (4). Тогда выполняется соотношение (5).

ДОКАЗАТЕЛЬСТВО. Неравенство Чебышева влечет за собой соотношение (2). Тогда теорема 7.7.1 доказывает утверждение следствия. ■

Этот результат был получен Андерсоном (1959) и в несколько ином виде Дианандой (1953).

Приведем несколько примеров центральных предельных теорем. Теорема 7.7.4 следует из теоремы 7.7.2.

ТЕОРЕМА 7.7.2. (Центральная предельная теорема Линдеберга.) Пусть $\omega_1^T, \dots, \omega_T^T$ — последовательность независимых случайных величин, $T = 1, 2, \dots$. Пусть $F_i^T(\omega)$ есть распределение величины ω_i^T , $\mathbb{E} \omega_i^T = 0$, $i = 1, \dots, T$,

$$(16) \quad \sum_{i=1}^T \text{Var} \omega_i^T = 1.$$

Достаточным условием сходимости распределения суммы $\sum_{i=1}^T \omega_i^T$ к $N(0, 1)$ является

$$(17) \quad \sum_{i=1}^T \int_{|\omega| > \delta} \omega^2 dF_i^T(\omega) \rightarrow 0$$

при $T \rightarrow \infty$ для любого $\delta > 0$.

Условие (17) является также и необходимым для сходимости распределения суммы $\sum_{i=1}^T \omega_i^T$ к $N(0, 1)$, если $\max_{i=1, \dots, T} \mathbb{E} (\omega_i^T)^2 \rightarrow 0$. [См. Лоэв (1963, § 21.2).]

ТЕОРЕМА 7.7.3. (Центральная предельная теорема Ляпунова.) Пусть $\omega_1^T, \dots, \omega_T^T$, $T = 1, 2, \dots$, — последовательность независимых

случайных величин, такая, что $\mathbb{E}w_t^T = 0$, $t = 1, \dots, T$, и выполняется (16). Достаточным условием сходимости распределения суммы $\sum_{t=1}^T w_t^T$ к $N(0, 1)$ является

$$(18) \quad \sum_{t=1}^T \mathbb{E}|w_t^T|^{2+\delta} \rightarrow 0,$$

если оно выполняется для некоторого $\delta > 0$.

ТЕОРЕМА 7.7.4. (Одинаково распределенные случайные величины.) Если $u_1, u_2, \dots$ независимы и одинаково распределены, $\mathbb{E}u_t = 0$ и $\mathbb{E}u_t^2 = \sigma^2$, то распределение $\sum_{t=1}^T u_t/(\sigma \sqrt{T})$ сходится к $N(0, 1)$ при $T \rightarrow \infty$.

Из теоремы 7.7.1 и следствия 7.7.1 можно получить центральные предельные теоремы для некоторых случайных стационарных процессов. Представляет интерес случай конечной зависимости, в котором переменные независимы, если разность индексов достаточно большая. Конечный процесс скользящего среднего $y_t = \sum_{s=0}^q \alpha_s v_{t-s}$, где v_t независимы и одинаково распределены, обладает этим свойством.

ТЕОРЕМА 7.7.5. Пусть $y_1, y_2, \dots$ — стационарный случайный процесс, такой, что для каждого целого n и целых $t_1, \dots, t_n$ ($0 < t_1 < \dots < t_n$) $y_{t_1}, \dots, y_{t_n}$ распределены независимо от $y_1, \dots, y_{t_1-m-1}$ и $y_{t_n+m+1}, \dots$. Если $\mathbb{E}y_t = 0$ и $\mathbb{E}y_t^2 < \infty$, то $\sum_{t=1}^T y_t/\sqrt{T}$ имеет предельное нормальное распределение со средним значением 0 и дисперсией

$$(19) \quad \mathbb{E}y_1^2 + 2\mathbb{E}y_1y_2 + \dots + 2\mathbb{E}y_1y_{m+1}.$$

ДОКАЗАТЕЛЬСТВО. Пусть

$$(20) \quad Y_T = \frac{1}{\sqrt{T}} \sum_{t=1}^T y_t.$$

Тогда

$$(21) \quad \mathbb{E}Y_T = 0,$$

$$(22) \quad \mathbb{E}Y_T^2 = \frac{1}{T} \sum_{t,s=1}^T \mathbb{E}y_t y_s = \frac{1}{T} [T\mathbb{E}y_1^2 + 2(T-1)\mathbb{E}y_1y_2 + \\ + 2(T-2)\mathbb{E}y_1y_3 + \dots + 2(T-m)\mathbb{E}y_1y_{m+1},$$

и

$$(23) \quad \lim_{T \rightarrow \infty} \mathbb{E} Y_T^2 = \mathbb{E} y_1^2 + 2\mathbb{E} y_1 y_2 + \dots + 2\mathbb{E} y_1 y_{m+1} = W.$$

Для данного целого $k > 2m$ пусть $T = kN + r$, где $0 \leq r < k$, т. е. $N = [T/k]$. Определим

$$(24) \quad U_{kT} = \frac{1}{\sqrt{T}} [(y_1 + \dots + y_{k-m}) + (y_{k+1} + \dots + y_{2k-m}) + \dots \\ \dots + (y_{k(N-1)+1} + \dots + y_{kN-m})],$$

$$(25) \quad X_{kT} = \frac{1}{\sqrt{T}} [(y_{k-m+1} + \dots + y_k) + (y_{2k-m+1} + \dots + y_{2k}) + \dots \\ \dots + (y_{kN-m+1} + \dots + y_{kN})],$$

$$(26) \quad R_{kT} = \frac{1}{\sqrt{T}} (y_{kN+1} + \dots + y_T)$$

для $k \leq T$. Пусть $U_{kT} = Y_T$, $X_{kT} = R_{kT} = 0$ для $k > T$.

Величина $U_{kT} \sqrt{T}$ есть сумма N независимо и одинаково распределенных слагаемых. Для j -го слагаемого имеем

$$(27) \quad \mathbb{E} (y_{k(j-1)+1} + \dots + y_{kj-m}) = 0,$$

$$(28) \quad \begin{aligned} \mathbb{E} (y_{k(j-1)+1} + \dots + y_{kj-m})^2 &= \\ &= \mathbb{E} (y_1 + \dots + y_{k-m})^2 = \\ &= (k-m) \mathbb{E} y_1^2 + 2(k-m-1) \mathbb{E} y_1 y_2 + \dots \\ &\quad \dots + 2(k-2m) \mathbb{E} y_1 y_{m+1} = kV_k. \end{aligned}$$

По теореме 7.7.4 сумма всех этих слагаемых, деленная на $\sqrt{N}$ (т. е. $U_{kT} \sqrt{T/N}$), имеет предельное нормальное распределение со средним значением 0 и дисперсией kV_k при $T \rightarrow \infty$ и $N \rightarrow \infty$. Таким образом, $U_{kT} \sqrt{T/(Nk)}$ имеет предельное нормальное распределение со средним значением 0 и дисперсией V_k .

Среднеквадратичное значение величины R_{kT} ограничено, так как

$$(29) \quad \mathbb{E} R_{kT}^2 \leq \frac{r^2}{T} \mathbb{E} y_1^2 \leq \frac{k^2}{T} \mathbb{E} y_1^2.$$

Мы использовали тот факт, что $|\mathbb{E} y_i y_s| \leq \mathbb{E} y_i^2$. Когда $T \rightarrow \infty$, (29) стремится к 0 и R_{kT} сходится по вероятности к 0. Следовательно, $U_{kT} + R_{kT}$ имеет предельное нормальное распределение со средним значением 0 и дисперсией V_k .

Для $k \leq T$ справедливо

$$(30) \quad \mathbb{E} X_{kT}^2 = \frac{N}{T} \mathbb{E} (y_1 + \dots + y_m)^2 \leq \frac{1}{k} \mathbb{E} (y_1 + \dots + y_m)^2,$$

а $\mathcal{E}X_{kT}^2 = 0$ для $k > T$. Поэтому для k достаточно большого $\mathcal{E}X_{kT}^2$ есть величина сколь угодно малая, равномерно по T . Заметим, что $V_k \rightarrow W$ при $k \rightarrow \infty$. Применяя следствие 7.7.1, получим утверждение теоремы.

ТЕОРЕМА 7.7.6. Пусть $y_1, y_2, \dots$ — последовательность векторных случайных величин, удовлетворяющая условиям теоремы 7.7.5, где условие $\mathcal{E}y_t^2 < \infty$ заменено на $\mathcal{E}y_t' y_t < \infty$. Тогда $(1/\sqrt{T}) \sum_{t=1}^T y_t$ имеет предельное нормальное распределение со средним значением 0 и ковариационной матрицей

$$(31) \quad \mathcal{E}y_1 y_1' + \mathcal{E}y_1 y_2' + \mathcal{E}y_2 y_1' + \dots + \mathcal{E}y_1 y_{m+1}' + \mathcal{E}y_{m+1} y_1'.$$

ДОКАЗАТЕЛЬСТВО. Это утверждение следует из теоремы 7.7.5 и следующей теоремы.

ТЕОРЕМА 7.7.7. Если последовательность случайных матриц W_T , $T = 1, 2, \dots$, такова, что каждая линейная комбинация элементов $\text{tr } \Psi W_T$ имеет предельное нормальное распределение со средним значением 0 и дисперсией $\text{tr } \Sigma \Psi A \Psi'$, то W_T имеет многомерное предельное нормальное распределение со средним значением 0 и ковариациями

$$(32) \quad \mathcal{E}w_{gi} w_{hj} = a_{gh} \sigma_{ij}.$$

ДОКАЗАТЕЛЬСТВО. Из условия теоремы следует, что

$$(33) \quad \lim_{T \rightarrow \infty} \mathcal{E} e^{it \text{tr } \Psi W_T} = e^{- (t^2/2) \text{tr } \Sigma \Psi A \Psi'}$$

для каждого действительного t ; в частности, для $t = 1$ получим

$$(34) \quad \lim_{T \rightarrow \infty} \mathcal{E} \exp \left[i \sum_{j,g} \psi_{jg} w_{gjT} \right] = \exp \left[- \frac{1}{2} \sum_{k,j,g,h} \sigma_{kj} \psi_{jg} a_{gh} \psi_{kh} \right],$$

а это есть характеристическая функция многомерного нормального распределения. Так как характеристическая функция непрерывна в $\Psi = 0$, то из так называемой теоремы непрерывности [см., например, Крамер (1946, стр. 112)] следует утверждение теоремы. ■

Эти теоремы можно обобщить и на линейные процессы. [См. также Марсалья (1954) и Парзен (1957а).]

ТЕОРЕМА 7.7.8. Если

$$(35) \quad y_t = \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s},$$

где $\{v_i\}$ — последовательность независимых одинаково распределенных случайных величин, $\mathcal{E}v_i = 0$, $\mathcal{E}v_i^2 = \sigma^2$ и $\sum_{s=-\infty}^{\infty} |\gamma_s| < \infty$,

то $\sum_{t=1}^T y_t / \sqrt{T}$ имеет предельное нормальное распределение со средним значением 0 и дисперсией

$$(36) \quad \sum_{r=-\infty}^{\infty} \sigma(r),$$

где

$$(37) \quad \sigma(r) = \sigma^2 \sum_{s=-\infty}^{\infty} \gamma_s \gamma_{s+r}.$$

ДОКАЗАТЕЛЬСТВО. Пусть

$$(38) \quad y_{t,k} = \sum_{s=-k}^k \gamma_s u_{t-s},$$

$$(39) \quad u_{t,k} = \sum_{s=k+1}^{\infty} \gamma_s u_{t-s},$$

$$(40) \quad u_{t,k}^* = \sum_{s=-\infty}^{-(k+1)} \gamma_s u_{t-s}.$$

Тогда $y_t = y_{t,k} + u_{t,k} + u_{t,k}^*$. Пусть

$$(41) \quad Z_{kT} = \frac{1}{\sqrt{T}} \sum_{t=1}^T y_{t,k},$$

$$(42) \quad X_{kT} = \frac{1}{\sqrt{T}} \sum_{t=1}^T u_{t,k},$$

$$(43) \quad X_{kT}^* = \frac{1}{\sqrt{T}} \sum_{t=1}^T u_{t,k}^*.$$

Тогда $\sum_{t=1}^T y_t / \sqrt{T} = Z_{kT} + X_{kT} + X_{kT}^*$. Имеем

$$\begin{aligned} (44) \quad \mathbb{E} X_{kT}^2 &= \frac{1}{T} \mathbb{E} \left(\sum_{t=1}^T u_{t,k} \right)^2 = \frac{1}{T} \sum_{t,s=1}^T \mathbb{E} u_{t,k} u_{s,k} = \\ &= \frac{1}{T} \sum_{t,s=1}^T \sum_{p,r=k+1}^{\infty} \gamma_p \gamma_r \mathbb{E} u_{t-p} u_{s-r} = \\ &= \frac{\sigma^2}{T} \sum_{t,s=1}^T \sum_{p=\max(0,t-s)+k+1}^{\infty} \gamma_p \gamma_{p+s-t} \leq \\ &\leq \frac{\sigma^2}{T} \sum_{t,s=1}^T \sum_{p=\max(0,t-s)+k+1}^{\infty} |\gamma_p| \cdot |\gamma_{p+s-t}| \leq \end{aligned}$$

$$\begin{aligned} &\leq \sigma^2 \sum_{q=k+1}^{\infty} \sum_{p=k+1}^{\infty} |\gamma_p| \cdot |\gamma_q| = \\ &= \sigma^2 \left(\sum_{p=k+1}^{\infty} |\gamma_p| \right)^2. \end{aligned}$$

Так как $\sum_{p=1}^{\infty} |\gamma_p| < \infty$, то (44) сходится к 0 при $k \rightarrow \infty$. Подобным же образом

$$(45) \quad \mathbb{E} X_{kT}^{*2} \leq \sigma^2 \left(\sum_{p=-\infty}^{-(k+1)} |\gamma_p| \right)^2,$$

что сходится к 0 при $k \rightarrow \infty$. Тогда

$$(46) \quad \begin{aligned} \mathbb{E} (X_{kT} + X_{kT}^*)^2 &\leq 2 [\mathbb{E} X_{kT}^2 + \mathbb{E} X_{kT}^{*2}] \leq \\ &\leq 2\sigma^2 \left[\left(\sum_{p=k+1}^{\infty} |\gamma_p| \right)^2 + \left(\sum_{p=-\infty}^{-(k+1)} |\gamma_p| \right)^2 \right], \end{aligned}$$

что сходится к 0 при $k \rightarrow \infty$. Утверждение теоремы 7.7.8 следует из следствия 7.7.1 и теоремы 7.7.5. ■

Подобные теоремы (для $y_t = \sum_{s=0}^{\infty} \gamma_s u_{t-s}$) были доказаны несколько по-иному Мораном (1947) и Дианандой (1953).

ТЕОРЕМА 7.7.9. Пусть $y_1, y_2, \dots$ — последовательность случайных величин, такая, что для каждого n и $t_1, \dots, t_n$ ($0 < t_1 < \dots < t_n$) совокупности $y_{t_1}, \dots, y_{t_n}$ и $y_1, \dots, y_{t_1-m-1}, y_{t_n+m+1}, \dots$ взаимно независимы. Пусть $\mathbb{E} y_t = 0$ и $\mathbb{E} |y_t|^{2+\delta} < M$, $t = 1, 2, \dots$, для некоторого $\delta > 0$ и некоторого M . Пусть $a_1, a_2, \dots$ — последовательность постоянных величин, такая, что $|a_t| < L$, $t = 1, 2, \dots$, для некоторого L , и такая, что существует предел

$$(47) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t,s=1}^T a_t a_s \mathbb{E} y_t y_s.$$

Тогда $\sum_{t=1}^T a_t y_t / \sqrt{T}$ имеет предельное нормальное распределение со средним значением 0 и дисперсией (47).

Доказательство. Пусть $y_t^* = a_t y_t$, $t = 1, 2, \dots$. Доказательство аналогично доказательству теоремы 7.7.5 с заменой y_t на y_t^* . Случайные величины $y_{k(j-1)+1}^* + \dots + y_{kj-m}^*$, $j = 1, \dots, N$, независимы, но не одинаково распределены. Однако их абсолютные моменты порядка $2 + \delta$ равномерно ограничены величиной $(k-m)^{2+\delta} L^{2+\delta} M$ (неравенство Минковского). Среднеквадратичное значение разнос-

ти между $\sum_{i=1}^T y_i^* / \sqrt{T}$ и U_{kT}^* [которое совпадает с (24) при замене y_t на y_t^*] ограничено величиной [приблизительно равной $m^2 L^2 (M + 1)/k$], которая стремится к 0 равномерно при $k \rightarrow \infty$; $\lim_{k \rightarrow \infty} \lim_{T \rightarrow \infty} \text{Var } U_{kT}^*$ есть (47). Доказательство завершается применением центральной предельной теоремы Ляпунова (теорема 7.7.3). ■

Следствие 7.7.2. Пусть $y_1, y_2, \dots$ — последовательность случайных величин, таких, что для каждого n и $t_1, \dots, t_n$ ($0 < t_1 < \dots < t_n$) совокупности $y_{t_1}, \dots, y_{t_n}$ и $y_1, \dots, y_{t_1-m-1}, y_{t_1+m+1}, \dots$ независимы. Пусть $\mathbb{E} y_t = 0$, $\mathbb{E} y_t y_s = \sigma(t-s)$ и $\mathbb{E} |y_t|^{2+\delta} < M$, $t = 1, 2, \dots$, для некоторого $\delta > 0$ и некоторого M . Тогда $\sum_{i=1}^T y_i / \sqrt{T}$ имеет предельное нормальное распределение со средним значением 0 и дисперсией $\sigma(0) + 2 \sum_{h=1}^m \sigma(h)$.

Доказательство несколько упростится, если положить $\delta = 1$. Эти результаты применяются для регрессионных оценок.

ЛИТЕРАТУРА

- Наиболее полно теория случайных процессов изложена в книге Дуба (1953). Достаточно полной является также книга Лозва (1963). Читатель может использовать книги Бартлетта (1966), Яглома (1962), Парзена (1962) и Розенблатта (1962).
- § 7.2. Колмогоров (1933).
- § 7.3. Лозв (1963), Хобсон (1907).
- § 7.4. Дуб (1953).
- § 7.5. Дуб (1953), Лукач (1960).
- § 7.6. Акутовиц (1957), Винер (1949), Вольд (1954), Гренандер и Розенблатт (1957), Дуб (1953), Зигмунд (1959), Колмогоров (1941a), (1941b), Лозв (1963), Парзен (1961a).
- § 7.7. Андерсон (1959), Диананда (1953), Крамер (1946), Лозв (1963), Марсалья (1954), Моран (1947), Парзен (1957a).

УПРАЖНЕНИЯ

1. (Разд. 7.2.1) Докажите, что процесс, стационарный в узком смысле и имеющий конечную дисперсию, стационарен в широком смысле.
2. (Разд. 7.2.1) Докажите, что если $\{y_t\}$ и $\{z_t\}$ — два независимых стационарных процесса, то $\{y_t + z_t\}$ стационарный процесс.
3. (Разд. 7.2.2) Пусть $y_t = (-1)^t y$, где y — случайная величина.
 - а) Каким должно быть необходимое и достаточное условие для распределения величины y , чтобы процесс $\{y_t\}$ был стационарным в узком смысле?
 - б) Каким должно быть необходимое и достаточное условие для распределения величины y , чтобы процесс $\{y_t\}$ был стационарным в широком смысле?

4. (Разд. 7.2.2) Покажите, что

$$y_t = \alpha \cos \lambda t + \beta \sin \lambda t + u_t,$$

где u_t независимы и одинаково распределены, не является стационарным процессом, если $\lambda \neq 0$.

5. (Разд. 7.2.2) Дайте пример таких A_j и B_j , $j = 1, \dots, q$, что (13) не стационарен в узком смысле.

6. (Разд. 7.2.2) Покажите, что

$$y_t = \rho \cos(\lambda t + v),$$

где v равномерно распределено на $(0, 2\pi)$, является стационарным случайным процессом.

7. (Разд. 7.2.2) Покажите, что процесс $\{y_t\}$, определяемый (43), стационарен в узком смысле.

8. (Разд. 7.2.2) Пусть

$$y_t = \sum_{j=1}^q D_j \cos(\lambda_j t - v_j),$$

где $D_1, \dots, D_q, v_1, \dots, v_q$ независимы, $D_j^2 = 2\sigma_j^2$ и v_j равномерно распределены на $(0, 2\pi)$. Покажите, что $\{y_t\}$ имеет ковариационную функцию (44).

9. (§ 7.3) Пусть $\{y_t\}$ и $\{z_t\}$ — независимые стационарные процессы со спектральными плотностями $f_y(\lambda)$ и $f_z(\lambda)$. Покажите, что процесс $\{x_t\} = \{y_t + z_t\}$ имеет спектральную плотность $f_y(\lambda) + f_z(\lambda)$. (См. упр. 2.)

10. (§ 7.3) Пусть $y_t = (-1)^t y$, где y — случайная величина, такая, что процесс $\{y_t\}$ стационарен в широком смысле. (См. упр. 3.) Найдите $\sigma(h)$ и $F(\lambda)$.

11. (§ 7.3) Нарисуйте графики функций $\cos 2\pi t$ и $\cos(4 \cdot 2\pi t)$ и покажите, что они совпадают в точках $t = 0, \pm 1/5, \pm 2/5, \dots$.

12. (§ 7.3) Для комплексного процесса положим $\sigma(k) = \alpha(k) + i\beta(k)$. Покажите, что $\alpha(-k) = \alpha(k)$ и $\beta(-k) = -\beta(k)$ и что

$$f(\lambda) = \frac{1}{2\pi} \left\{ \sigma(0) + 2 \sum_{k=1}^{\infty} [\alpha(k) \cos \lambda k + \beta(k) \sin \lambda k] \right\}.$$

13. (Разд. 7.4.1) Пусть

$$(i) \quad 0 = \lambda_0 < \lambda_1 < \dots < \lambda_n = \pi.$$

Если функция $h(\lambda)$ непрерывна на отрезке $[0, \pi]$, то она равномерно непрерывна и

$$(ii) \quad \max h(\lambda) - \min h(\lambda) < \varepsilon, \quad \lambda_{j-1} \leq \lambda \leq \lambda_j,$$

если $\lambda_j - \lambda_{j-1} < \delta$ (для соответственно выбранного δ). Покажите, что для этого δ

$$(iii) \quad \sum_{j=1}^n \max_{\lambda_{j-1} \leq \lambda \leq \lambda_j} h(\lambda) [H(\lambda_j) - H(\lambda_{j-1})] - \\ - \sum_{j=1}^n \min_{\lambda_{j-1} \leq \lambda \leq \lambda_j} h(\lambda) [H(\lambda_j) - H(\lambda_{j-1})] < \varepsilon [H(\pi) - H(0)].$$

Используя эти факты, докажите существование интеграла Римана — Стильеса

$$(iv) \quad \int_0^{\pi} h(\lambda) dH(\lambda).$$

14. (Разд. 7.4.1) Пусть

$$\begin{aligned} 0 &= \lambda_0 < \lambda_1 < \dots < \lambda_n = \pi, \\ \lambda_0 &\leq \lambda'_1 \leq \lambda_1, \dots, \lambda_{n-1} \leq \lambda'_n \leq \lambda_n, \\ \lambda_0 &\leq \lambda''_1 \leq \lambda_1, \dots, \lambda_{n-1} \leq \lambda''_n \leq \lambda_n. \end{aligned}$$

Пусть $C(\lambda)$, $0 \leq \lambda \leq \pi$, — случайный процесс с некоррелированными приращениями и ковариационной функцией $H(\lambda)$. Покажите, что

$$\begin{aligned} \mathbb{E} \left\{ \sum_{j=1}^n h(\lambda'_j) [C(\lambda_j) - C(\lambda_{j-1})] - \sum_{j=1}^n h(\lambda''_j) [C(\lambda_j) - C(\lambda_{j-1})] \right\}^2 = \\ = \sum_{j=1}^n [h(\lambda'_j) - h(\lambda''_j)]^2 [H(\lambda_j) - H(\lambda_{j-1})]. \end{aligned}$$

15. (Разд. 7.4.1) Пусть

$$\begin{aligned} 0 &= \lambda_0 < \lambda_1 < \dots < \lambda_{2n} = \pi. \\ S_n &= \sum_{j=1}^n h(\lambda_{2j}) [C(\lambda_{2j}) - C(\lambda_{2(j-1)})], \\ S_{2n} &= \sum_{i=1}^{2n} h(\lambda_i) [C(\lambda_i) - C(\lambda_{i-1})]. \end{aligned}$$

Покажите, что

$$\mathbb{E} (S_{2n} - S_n)^2 \leq \max_{j=1, \dots, n} [h(\lambda_{2j-1}) - h(\lambda_{2j})]^2 [H(\pi) - H(0)].$$

16. (Разд. 7.4.1) Дайте полное доказательство (21).

17. (Разд. 7.4.1) Покажите, что можно ввести определение стохастического интеграла Римана — Стильеса, если $h(\lambda)$ имеет конечное число разрывов и если $h(\lambda)$ и $H(\lambda)$ не имеют общих точек разрыва.

18. (Разд. 7.4.1) Пусть $C(\lambda)$, $0 \leq \lambda \leq \pi$ — случайный процесс с некоррелированными приращениями и ковариационной функцией $H(\lambda)$, а $k(\lambda)$ — ступенчатая функция

$$k(\lambda) = c_j, \quad a_{j-1} < \lambda \leq a_j, \quad j = 1, \dots, n,$$

где $0 = a_0 < a_1 < \dots < a_n = \pi$, $k(0) = c_1$. Допустим, что $H(\lambda)$ непрерывна в точках $a_0, a_1, \dots, a_n$. Покажите, что стохастический интеграл Римана — Стильеса (определенный, как в упр. 7) есть

$$\int_0^{\pi} k(\lambda) dC(\lambda) = \sum_{j=1}^n c_j [C(a_j) - C(a_{j-1})].$$

19. (Разд. 7.4.1) Пусть $h(\lambda)$ удовлетворяет неравенству

$$\int_0^\pi h^2(\lambda) dH(\lambda) < \infty.$$

Стохастический интеграл Лебега — Стильеса определяется как предел в среднем величин

$$\int_0^\pi h_n(\lambda) dC(\lambda),$$

где $\{h_n(\lambda)\}$ — последовательность ступенчатых функций, такая, что

$$\lim_{n \rightarrow \infty} \int_0^\pi |h_n(\lambda) - h(\lambda)|^2 dH(\lambda) = 0.$$

Покажите, что из существования стохастического интеграла Римана — Стильеса следует существование стохастического интеграла Лебега — Стильеса (с вероятностью 1).

20. (Разд. 7.5.1) Покажите, что неравенство

$$\sum_{r=-\infty}^{\infty} |C_r| < \infty$$

является достаточным условием сходимости правой части формулы (2).

21. (Разд. 7.5.2) Пусть $\rho e^{i\theta}$ и $\rho e^{-i\theta}$ — корни уравнения $z^2 + c_1 z + c_2 = 0$, и пусть

$$y_t = v_t + c_1 v_{t-1} + c_2 v_{t-2},$$

где $\xi v_t = 0$, $\xi v_t^2 = \sigma^2$ и $\xi v_t v_s = 0$, $t \neq s$. Покажите, что

$$\sigma(0) = (1 + 4\rho^2 \cos^2 \theta + \rho^4) \sigma^2,$$

$$\sigma(1) = -2\rho(1 + \rho^2) \sigma^2 \cos \theta,$$

$$\sigma(2) = \rho^2 \sigma^2.$$

22. (Разд. 7.5.3) Покажите, что если y_t удовлетворяет стохастическому разностному уравнению

$$\prod_{j=1}^p (1 - x_j \mathcal{L}) y_t = u_t$$

(где $|x_j| < 1$), то

$$\prod_{j=1}^r \left(1 - \frac{1}{x_j} \mathcal{L}\right) \prod_{j=1+r}^p (1 - x_j \mathcal{L}) y_t = u_t^*$$

— есть последовательность некоррелированных случайных величин с постоянной дисперсией; если x_j , $j = 1, \dots, r$, — комплексное число, то его сопряженное значение $x_i = x_j$, $i \neq j$, входит в число первых r корней.

23. (Разд. 7.5.3) Покажите, что

$$f(\lambda) = \sum_{i=1}^p \frac{c_i (1 - x_i^2)}{2\pi (1 - 2x_i \cos \lambda + x_i^2)},$$

где

$$\sigma(h) = \sum_{i=1}^p c_i x_i^h$$

и $|x_i| < 1$, $i = 1, \dots, p$.

24. (Разд. 7.5.3) Пусть

$$y_t - \frac{3}{4} y_{t-1} + \frac{1}{8} y_{t-2} = u_t$$

- a) Найдите корни x_1 и x_2 соответствующего характеристического уравнения.
 б) Выпишите ковариационную функцию $\sigma(h)$.
 в) Выпишите спектральную плотность $f(\lambda)$.
 г) Выразите спектральную плотность тремя другими способами (заменяя по крайней мере одно из x_1 и x_2 соответственно на $1/\bar{x}_1$ и $1/\bar{x}_2$).
 е) Выпишите стохастическое разностное уравнение и соответствующее характеристическое уравнение, относящиеся к каждому виду $f(\lambda)$.

25. (Разд. 7.5.3) Пусть

$$y_t - 1.4y_{t-1} + 0.98y_{t-2} = u_t.$$

Найдите: (а) корни x_1 и x_2 соответствующего характеристического уравнения; (б) $\sigma(h)$ и (в) $f(\lambda)$. (г) Выразите $f(\lambda)$ тремя другими способами (заменяя по крайней мере одно из x_1 и x_2 соответственно на $1/\bar{x}_1$ и $1/\bar{x}_2$). (е) Напишите стохастическое разностное уравнение и соответствующее характеристическое уравнение, относящиеся к каждому виду $f(\lambda)$, предполагая, что такие уравнения имеют действительные коэффициенты.

26. (Разд. 7.5.3) Напишите $\sigma(h)$ и $f(\lambda)$ для $y_t + y_{t-1}/2 = u_t$.27. (Разд. 7.5.3) Напишите $\sigma(h)$ и $f(\lambda)$ для $y_t - y_{t-1}/2 = u_t$.

28. (Разд. 7.5.4) Пусть $f(\lambda)$ — непрерывная спектральная плотность. Покажите, что для любого данного $\varepsilon > 0$ существует процесс $\sum_{j=0}^m \alpha_j v_{t-j}$ скользящего среднего со спектральной плотностью $f_m(\lambda)$, такой, что для соответствующих ковариационных последовательностей $\{\sigma(h)\}$ и $\{\sigma_m(h)\}$ выполняется условие

$$|\sigma(h) - \sigma_m(h)| < 2\pi\varepsilon.$$

Покажите, что существует такой процесс авторегрессии, что (i) выполняется для $f_m(\lambda)$ и $\{\sigma_m(h)\}$, соответствующих этому процессу.

29. (Разд. 7.5.4) Пусть

$$z_t = y_t + w_t,$$

где

$$y_t + \beta y_{t-1} = u_t$$

и u_t и w_t независимо распределены со средними значениями 0 и дисперсиями $\xi u_t^2 = \sigma^2$ и $\xi w_t^2 = \tau^2$. Покажите, что спектральная плотность процесса $\{z_t\}$ равна

$$\frac{\sigma^2 + (2\beta \cos \lambda + 1 + \beta^2) \tau^2}{2\pi (2\beta \cos \lambda + 1 + \beta^2)}.$$

Найдите процесс $\{z_t^*\}$ с той же спектральной плотностью, удовлетворяющий следующему уравнению:

$$z_t^* + \beta z_{t-1}^* = v_t + \alpha v_{t-1},$$

где $\mathbb{E}v_t = 0$, $\mathbb{E}v_t^2 = K$ и $\mathbb{E}v_t v_s = 0$, $t \neq s$.

30. (Разд. 7.5.5) Формула сглаживания, основанная на нахождении полинома второго порядка по пяти точкам, есть

$$z_t = -\frac{3}{35}y_{t-2} + \frac{12}{35}y_{t-1} + \frac{17}{35}y_t + \frac{12}{35}y_{t+1} - \frac{3}{35}y_{t+2},$$

где $\mathbb{E}y_t = 0$, $\mathbb{E}y_t^2 = \sigma^2$, $\mathbb{E}y_t y_s = 0$, $t \neq s$. Найдите спектральную плотность и ковариационную последовательность процесса $\{z_t\}$.

31. (Разд. 7.5.5) Коэффициенты формулы сглаживания, основанные на нахождении полинома степени $2k$ (или $2k+1$) по $2k+3$ точкам, определены в (19), (20) и (21) § 3.3. Покажите, что спектральная плотность остатков при сглаживании некоррелированной последовательности с дисперсией σ^2 равна

$$\frac{\sigma^2}{2\pi} C^2 \left(2 \sin \frac{1}{2} \lambda \right)^{4k+4},$$

где C определено формулой (21) § 3.3.

32. (Разд. 7.5.5) Дайте явную формулу спектральной плотности и ковариационной последовательности процесса $\{\Delta^2 y_t\}$, где $\mathbb{E}y_t = 0$, $\mathbb{E}y_t^2 = \sigma^2$, $\mathbb{E}y_t y_s = 0$, $t \neq s$, и начертите их графики.

33. (Разд. 7.5.5) Покажите, что спектральная плотность остатков при сглаживании некоррелированной последовательности $\{y_t\}$ с дисперсией σ^2 , суммой

$$\sum_{s=-m}^m y_{t+s} / (2m+1) \text{ равна}$$

$$\frac{\sigma^2}{2\pi} \left(1 - \frac{\sin \lambda \left(m + \frac{1}{2} \right)}{(2m+1) \sin \frac{1}{2} \lambda} \right)^2.$$

34. (Разд. 7.6.1) Докажите, что в линейном векторном пространстве каждой точке x соответствует единственная точка $-x$, такая, что $x + (-x) = 0$.

35. (Разд. 7.6.1) Докажите, что $(x, x) = 0$ тогда и только тогда, когда $x = 0$.

36. (Разд. 7.6.1) Покажите, что конечномерное евклидово пространство есть линейное векторное пространство. Покажите, что $x'y$ есть скалярное произведение.

37. (Разд. 7.6.1) Покажите, что множество всех конечных линейных комбинаций величин y_t , $t = \dots, -1, 0, 1, \dots$, составляет линейное векторное пространство. Покажите, что $\mathbb{E}xy$ есть скалярное произведение.

38. (Разд. 7.6.1) Покажите, что (29) и (32) в § 7.2 удовлетворяют условию следствия 7.6.1.

39. (Разд. 7.6.2) Проверьте, что $M^\perp$ и $M \oplus M^\perp$ являются подпространствами, если M есть подпространство гильбертова пространства.

40. (Разд. 7.6.3) Проверьте, что спектральная плотность формулы (41) определяется формулой (43).

ВЫБОРОЧНЫЕ СРЕДНЕЕ, КОВАРИАЦИИ И СПЕКТРАЛЬНАЯ ПЛОТНОСТЬ

8.1. ВВЕДЕНИЕ

Стационарный гауссовский процесс полностью описывается своим средним значением $\mathbb{E}y_t = \mu$ и ковариационной последовательностью $\mathbb{E}(y_t - \mu)(y_{t+h} - \mu) = \sigma(h)$, $h = 0, 1, \dots$, или средним значением и спектральной функцией $F(\lambda)$, или средним значением и спектральной плотностью, когда последняя существует. Все эти описания процесса эквивалентны. Если процесс не гауссовский, указанные величины также немаловажны, хотя они не определяют процесс однозначно. В настоящей главе мы рассмотрим выборочные аналоги этих величин. Выборочное среднее величин $y_1, \dots, y_T$, обозначаемое $\bar{y}$, есть несмещенная оценка среднего значения процесса μ . Если математическое ожидание известно, то выборочная ковариация определяется как сумма произведений $\sum_{t=1}^{T-h} (y_t - \mu)(y_{t+h} - \mu)$, деленная на число наблюдений T или на число членов в сумме $T - h$. Если μ неизвестно, то выборочную ковариацию можно определить, используя отклонения от средних значений $y_t - \bar{y}$. Мы также рассмотрим и другие определения.

Тригонометрические коэффициенты определяются в виде $A(\lambda) = (2/T) \sum_{t=1}^T (y_t - \mu) \cos \lambda t$ и $B(\lambda) = (2/T) \sum_{t=1}^T (y_t - \mu) \sin \lambda t$, где μ известно; μ можно заменить на $\bar{y}$, когда μ неизвестно. Выборочная спектральная плотность $I(\lambda)$ есть число, кратное выборочному квадрату амплитуды $R^2(\lambda) = A^2(\lambda) + B^2(\lambda)$, точнее $I(\lambda) = [T/(8\pi)]R^2(\lambda)$.

Найдем первый и второй моменты этих выборочных величин. Большинство этих моментов являются сложными функциями первого, второго и четвертого моментов процесса. В § 8.3 мы найдем

пределы соответствующим образом нормированных моментов. Выборочные значения среднего, ковариации, корреляции и тригонометрические коэффициенты асимптотически нормально распределены при соответствующих условиях (§ 8.4).

Выборочная спектральная плотность $I(\lambda)$ распределена в пределе как χ^2 с двумя степенями свободы, умноженная на $f(\lambda)/2$. Таким образом, $I(\lambda)$ не является состоятельной оценкой для $f(\lambda)$. В гл. 9 мы рассмотрим оценку спектральной плотности $f(\lambda)$ в более общих условиях.

8.2. ОПРЕДЕЛЕНИЯ ВЫБОРОЧНЫХ СРЕДНЕГО, КОВАРИАЦИЙ, СПЕКТРАЛЬНОЙ ПЛОТНОСТИ И ИХ МОМЕНТОВ

8.2.1. Определения и некоторые соотношения

Пусть $y_1, \dots, y_T$ — T последовательных наблюдений случайного процесса $\{y_t\}$, $t = \dots, -1, 0, 1, \dots$, стационарного в широком смысле. Среднее значение и ковариационная последовательность равны

$$(1) \quad \mathbb{E} y_t = \mu, \quad t = \dots, -1, 0, 1, \dots,$$

$$(2) \quad \begin{aligned} \text{Cov}(y_t, y_{t+h}) &= \mathbb{E}(y_t - \mu)(y_{t+h} - \mu) = \\ &= \sigma(h), \quad t = \dots, -1, 0, 1, \dots, \\ & \quad h = 0, \pm 1, \pm 2, \dots, \end{aligned}$$

с $\sigma(h) = \sigma(-h)$. Допустим, что спектральная функция абсолютно непрерывна, а ее спектральная плотность есть $f(\lambda)$, $-\pi \leq \lambda \leq \pi$. Тогда

$$(3) \quad \begin{aligned} \sigma(h) &= \int_{-\pi}^{\pi} \cos \lambda h f(\lambda) d\lambda = \int_{-\pi}^{\pi} e^{i\lambda h} f(\lambda) d\lambda, \\ & \quad h = 0, \pm 1, \pm 2, \dots \end{aligned}$$

Если

$$(4) \quad \sum_{h=-\infty}^{\infty} |\sigma(h)| = \sigma(0) + 2 \sum_{h=1}^{\infty} |\sigma(h)| < \infty,$$

то спектральная плотность непрерывна и

$$(5) \quad f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \sigma(h) \cos \lambda h = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \sigma(h) e^{i\lambda h}.$$

Несмещенной оценкой для μ является

$$(6) \quad \bar{y} = \frac{1}{T} \sum_{t=1}^T y_t.$$

Если μ известно, то несмещенная оценка для $\sigma^2(h)$ есть

$$(7) \quad C_h = C_{-h} = \frac{1}{T-h} \sum_{t=1}^{T-h} (y_t - \mu)(y_{t+h} - \mu), \quad h = 0, 1, \dots, T-1.$$

Если μ неизвестно, то по аналогии можно построить следующую оценку:

$$(8) \quad C_h^* = C_{-h}^* = \frac{1}{T-h} \sum_{t=1}^{T-h} (y_t - \bar{y})(y_{t+h} - \bar{y}), \quad h = 0, 1, \dots, T-1.$$

Возможны еще и другие оценки, такие, как

$$(9) \quad \begin{aligned} \hat{C}_h = \hat{C}_{-h} &= \frac{1}{T-h} \sum_{t=1}^{T-h} (y_t - \bar{y}_h)(y_{t+h} - \bar{y}_{h+}) = \\ &= \frac{1}{T-h} \left[\sum_{t=1}^{T-h} y_t y_{t+h} - (T-h) \bar{y}_h \bar{y}_{h+} \right], \quad h = 0, 1, \dots, T-2, \end{aligned}$$

где

$$(10) \quad \bar{y}_h = \frac{1}{T-h} \sum_{t=1}^{T-h} y_t, \quad h = 0, 1, \dots, T-2,$$

$$(11) \quad \bar{y}_{h+} = \frac{1}{T-h} \sum_{t=1}^{T-h} y_{t+h}, \quad h = 0, 1, \dots, T-2.$$

Удобно рассмотреть также

$$(12) \quad c_h = c_{-h} = \frac{1}{T} \sum_{t=1}^{T-h} (y_t - \mu)(y_{t+h} - \mu) = \frac{T-h}{T} C_h, \\ h = 0, 1, \dots, T-1,$$

и

$$(13) \quad c_h^* = c_{-h}^* = \frac{1}{T} \sum_{t=1}^{T-h} (y_t - \bar{y})(y_{t+h} - \bar{y}) = \frac{T-h}{T} C_h^*, \\ h = 0, 1, \dots, T-1.$$

Согласно определениям гл. 4 положим

$$(14) \quad \begin{aligned} A(\lambda) &= \frac{2}{T} \sum_{t=1}^T (y_t - \mu) \cos \lambda t = \\ &= \frac{2}{T} \sum_{t=1}^T y_t \cos \lambda t - \mu \frac{2}{T} \sum_{t=1}^T \cos \lambda t, \quad -\pi \leq \lambda \leq \pi, \end{aligned}$$

$$\begin{aligned}
 (15) \quad B(\lambda) &= \frac{2}{T} \sum_{t=1}^T (y_t - \mu) \sin \lambda t = \\
 &= \frac{2}{T} \sum_{t=1}^T y_t \sin \lambda t - \mu \frac{2}{T} \sum_{t=1}^T \sin \lambda t, \quad -\pi \leq \lambda \leq \pi,
 \end{aligned}$$

$$(16) \quad R^2(\lambda) = A^2(\lambda) + B^2(\lambda), \quad -\pi \leq \lambda \leq \pi.$$

Последнее выражение для $R^2(\lambda)$ есть выборочная спектрограмма. Заметим, что $A(\lambda) = A(-\lambda)$, $B(\lambda) = -B(-\lambda)$ и $R^2(\lambda) = R^2(-\lambda)$. Мы можем определить также

$$\begin{aligned}
 (17) \quad A^*(\lambda) &= \frac{2}{T} \sum_{t=1}^T (y_t - \bar{y}) \cos \lambda t = \\
 &= \frac{2}{T} \sum_{t=1}^T y_t \cos \lambda t - \bar{y} \frac{2}{T} \sum_{t=1}^T \cos \lambda t, \quad -\pi \leq \lambda \leq \pi,
 \end{aligned}$$

$$\begin{aligned}
 (18) \quad B^*(\lambda) &= \frac{2}{T} \sum_{t=1}^T (y_t - \bar{y}) \sin \lambda t = \\
 &= \frac{2}{T} \sum_{t=1}^T y_t \sin \lambda t - \bar{y} \frac{2}{T} \sum_{t=1}^T \sin \lambda t, \quad -\pi \leq \lambda \leq \pi,
 \end{aligned}$$

$$(19) \quad R^{*2}(\lambda) = A^{*2}(\lambda) + B^{*2}(\lambda), \quad -\pi \leq \lambda \leq \pi.$$

Из свойств ортогональности $\cos 2\pi jt/T$ и $\sin 2\pi jt/T$, $t = 1, \dots, T$, следует

$$\begin{aligned}
 A\left(\frac{2\pi j}{T}\right) &= A^*\left(\frac{2\pi j}{T}\right), \\
 (20) \quad B\left(\frac{2\pi j}{T}\right) &= B^*\left(\frac{2\pi j}{T}\right), \\
 R^2\left(\frac{2\pi j}{T}\right) &= R^{*2}\left(\frac{2\pi j}{T}\right), \quad j = 1, \dots, \left[\frac{1}{2}T\right],
 \end{aligned}$$

и $A^*(0) = B(0) = B^*(0) = 0$, $A(0) = 2(\bar{y} - \mu)$.

Выборочная спектральная плотность равна

$$(21) \quad I(\lambda) = \frac{T}{8\pi} R^2(\lambda) = \frac{1}{2\pi T} \left| \sum_{t=1}^T (y_t - \mu) e^{i\lambda t} \right|^2, \quad -\pi \leq \lambda \leq \pi,$$

когда μ известно. Если μ неизвестно, то

$$(22) \quad I^*(\lambda) = \frac{T}{8\pi} R^{*2}(\lambda) = \frac{1}{2\pi T} \left| \sum_{t=1}^T (y_t - \bar{y}) e^{i\lambda t} \right|^2, \quad -\pi \leq \lambda \leq \pi,$$

$[I^*(\lambda) = I^*(-\lambda), -\pi \leq \lambda \leq \pi]$, хотя можно рассмотреть и другие определения [соответствующие (9) для ковариационной последовательности]. Функции $I(\lambda)$ и $I^*(\lambda)$ следует рассматривать как оценки для $f(\lambda)$.

ТЕОРЕМА 8.2.1.

$$(23) \quad I(\lambda) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} c_r \cos \lambda r, \quad -\pi \leq \lambda \leq \pi,$$

$$(24) \quad I^*(\lambda) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} c_r^* \cos \lambda r, \quad -\pi \leq \lambda \leq \pi,$$

ДОКАЗАТЕЛЬСТВО. Имеем

$$\begin{aligned} (25) \quad I(\lambda) &= \frac{T}{8\pi} \left(\frac{2}{T} \right)^2 \left[\sum_{t,s=1}^T (y_t - \mu)(y_s - \mu) \cos \lambda t \cos \lambda s + \right. \\ &\quad \left. + \sum_{t,s=1}^T (y_t - \mu)(y_s - \mu) \sin \lambda t \sin \lambda s \right] = \\ &= \frac{1}{2\pi T} \sum_{t,s=1}^T (y_t - \mu)(y_s - \mu) \cos \lambda(t - s) = \\ &= \frac{1}{2\pi T} \sum_{r=-(T-1)}^{T-1} \sum_{s \in S_r} (y_{s+r} - \mu)(y_s - \mu) \cos \lambda r, \end{aligned}$$

где $S_r = \{1, 2, \dots, T - r\}$, $r \geq 0$ и $S_r = \{1 - r, 2 - r, \dots, T\}$, $r \leq 0$. Отсюда следует (23). Заменяя μ на $\bar{y}$, получаем (24). ■

ТЕОРЕМА 8.2.2.

$$(26) \quad \int_{-\pi}^{\pi} \cos \lambda h I(\lambda) d\lambda = \begin{cases} c_h, & h = 0, \pm 1, \dots, \pm(T-1), \\ 0, & h = \pm T, \pm(T+1), \dots, \end{cases}$$

$$(27) \quad \int_{-\pi}^{\pi} \cos \lambda h I^*(\lambda) d\lambda = \begin{cases} c_h^*, & h = 0, \pm 1, \dots, \pm(T-1), \\ 0, & h = \pm T, \pm(T+1), \dots. \end{cases}$$

ДОКАЗАТЕЛЬСТВО. Эти равенства следуют из свойств ортогональности $\cos \lambda h$ и теоремы 8.2.1. ■

Выборочные ковариации и спектральные плотности связаны точно так же, как и в исходном процессе. Формулы (23) и (24) аналогичны (5), а (26) и (27) соответствуют (3).

8.2.2. Моменты выборочного среднего и выборочных ковариаций

Среднее значение и дисперсия произвольной линейной комбинации

$$(28) \quad L = \sum_{t=1}^T k_t y_t$$

равны

$$(29) \quad \mathbb{E}L = \mu \sum_{t=1}^T k_t,$$

$$(30) \quad \begin{aligned} \text{Var } L &= \mathbb{E} \left[\sum_{t=1}^T k_t (y_t - \mu) \right]^2 = \sum_{t,s=1}^T \sigma(t-s) k_t k_s = \\ &= \sum_{r=-(T-1)}^{T-1} \sigma(r) \sum_{s \in S_r} k_{s+r} k_s = \int_{-\pi}^{\pi} \left| \sum_{t=1}^T k_t e^{i\lambda t} \right|^2 f(\lambda) d\lambda, \end{aligned}$$

где S_r определено при доказательстве теоремы 8.2.1. Для того чтобы L была несмещенной оценкой среднего μ , должно выполняться равенство

$$(31) \quad \sum_{t=1}^T k_t = 1.$$

Для данной ковариационной матрицы $\Sigma = [\sigma(t-s)]$ дисперсия величины L , удовлетворяющей (31), минимизируется величиной $\mathbf{k} = (k_1, \dots, k_T)'$, такой, что $\mathbf{k} = [1/\mathbf{e}'\Sigma^{-1}\mathbf{e}] \Sigma^{-1}\mathbf{e}$, где $\mathbf{e} = (1, \dots, 1)'$. Это определяет среднее значение, если $\mathbf{k} = (1/T)\mathbf{e}$, т. е. если $\mathbf{e}$ — характеристический вектор матрицы Σ , т. е. если $\sum_{s=1}^T \sigma(t-s)$ не зависит от t . (См. упр. 24 и 25 гл. 6.) Однако это выполняется не всегда.

ТЕОРЕМА 8.2.3. Дисперсия среднего значения $\bar{y}$ есть

$$(32) \quad \begin{aligned} \text{Var } \bar{y} &= \frac{1}{T^2} \sum_{t,s=1}^T \sigma(t-s) = \frac{1}{T} \sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T} \right) \sigma(r) = \\ &= \int_{-\pi}^{\pi} \left| \frac{1}{T} \sum_{t=1}^T e^{i\lambda t} \right|^2 f(\lambda) d\lambda = \int_{-\pi}^{\pi} \left(\frac{\sin \frac{1}{2} \lambda T}{T \sin \frac{1}{2} \lambda} \right)^2 f(\lambda) d\lambda. \end{aligned}$$

Для доказательства нужно использовать следующее: $\sum_{t=1}^T e^{i\lambda t} = e^{i\lambda} (1 - e^{i\lambda T}) / (1 - e^{i\lambda}) = e^{i\frac{1}{2}\lambda(T+1)} \sin \frac{1}{2} \lambda T / \sin \frac{1}{2} \lambda$.

Оценка ковариации или спектральной плотности в данной точке является обычно квадратичной формой

$$(33) \quad Q = \sum_{s,t=1}^T a_{st} (y_s - \mu) (y_t - \mu).$$

Если μ неизвестно, то матрицу (a_{st}) выбирают так, чтобы Q не зависело от μ (т. е. $\sum_{t=1}^T a_{st} = 0, s = 1, \dots, T$).

Дисперсия квадратичной формы будет включать моменты четвертого порядка. Допустим, что моменты четвертого порядка существуют, т. е. $\mathbb{E}(y_t - \mu)^4 < \infty$. Допустим также, что моменты четвертого порядка соответствуют стационарности, т. е. что

$$(34) \quad \mathbb{E}(y_t - \mu)(y_{t+h} - \mu)(y_{t+r} - \mu)(y_{t+s} - \mu) = \\ = \mathbb{E}(y_0 - \mu)(y_h - \mu)(y_r - \mu)(y_s - \mu), \quad t = \dots, -1, 0, 1, \dots$$

Для (34) введем обозначение $v(h, r, s)$. Если $\{y_t\}$ есть гауссовский процесс, то этот момент четвертого порядка будет равен

$$(35) \quad \sigma(h)\sigma(r-s) + \sigma(r)\sigma(h-s) + \sigma(s)\sigma(h-r).$$

В общем случае пусть

$$(36) \quad \kappa(h, r, s) = v(h, r, s) - [\sigma(h)\sigma(r-s) + \sigma(r)\sigma(h-s) + \\ + \sigma(s)\sigma(h-r)] = \\ = \mathbb{E}(y_t - \mu)(y_{t+h} - \mu)(y_{t+r} - \mu)(y_{t+s} - \mu) - \\ - [\sigma(h)\sigma(r-s) + \sigma(r)\sigma(h-s) + \sigma(s)\sigma(h-r)]. \\ t = \dots, -1, 0, 1, \dots,$$

что является семиинвариантом четвертого порядка. [См. упр. 2 относительно симметрии $v(h, r, s)$ и $\kappa(h, r, s)$.]

ТЕОРЕМА 8.2.4. Среднее значение и дисперсия квадратичной формы Q , определенной в (33), выражаются соотношениями

$$(37) \quad \mathbb{E}Q = \sum_{s,t=1}^T a_{st}\sigma(t-s) = \sum_{r=-(T-1)}^{T-1} \sigma(r) \sum_{s \in S_r} a_{s,s+r} = \\ = \int_{-\pi}^{\pi} \sum_{s,t=1}^T a_{st} e^{i\lambda(t-s)} f(\lambda) d\lambda = \int_{-\pi}^{\pi} A(\lambda, \lambda) f(\lambda) d\lambda,$$

$$(38) \quad \text{Var } Q = \sum_{s,t,s',t'=1}^T a_{st}a_{s't'} [2\sigma(s-s')\sigma(t-t') + \\ + \kappa(t-s, s'-s, t'-s)] = \\ = 2 \sum_{r,p=-(T-1)}^{T-1} \sigma(r)\sigma(p) \sum_{s \in S_r} \sum_{t \in S_p} a_{st}a_{s+r,t+p} + \\ + \sum_{r,p=-(T-1)}^{T-1} \sum_{s \in S_r} \sum_{t \in S_p} a_{st}a_{s+r,t+p} \kappa(t-s, r, t-s+p) =$$

$$\begin{aligned}
&= 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left| \sum_{s,t=1}^T a_{st} e^{i(\lambda s - \nu t)} \right|^2 f(\lambda) f(\nu) d\lambda d\nu + \\
&\quad + \sum_{s,t,s',t'=1}^T a_{st} a_{s't'} \kappa(t-s, s'-s, t'-s) = \\
&= 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |A(\lambda, \nu)|^2 f(\lambda) f(\nu) d\lambda d\nu + \\
&\quad + \sum_{s,t,s',t'=1}^T a_{st} a_{s't'} \kappa(t-s, s'-s, t'-s),
\end{aligned}$$

где

$$(39) \quad A(\lambda, \nu) = \sum_{s,t=1}^T a_{st} e^{i(\lambda s - \nu t)}$$

$u \ S_r = \{1, 2, \dots, T-r\}$, $r \geq 0$ и $S_r = \{1-r, 2-r, \dots, T\}$, $r \leq 0$.

ДОКАЗАТЕЛЬСТВО. Искомые выражения для $\mathbb{E}Q$ получаются непосредственно. Математическое ожидание Q^2 равно

$$\begin{aligned}
(40) \quad \mathbb{E}Q^2 &= \sum_{s,t,s',t'=1}^T a_{st} a_{s't'} \mathbb{E}(y_s - \mu)(y_t - \mu)(y_{s'} - \mu)(y_{t'} - \mu) = \\
&= \sum_{s,t,s',t'=1}^T a_{st} a_{s't'} [\sigma(s-t)\sigma(s'-t') + \sigma(s-s')\sigma(t-t') + \\
&\quad + \sigma(s-t')\sigma(t-s') + \kappa(t-s, s'-s, t'-s)].
\end{aligned}$$

Заметим, что

$$(41) \quad \sum_{s,t,s',t'=1}^T a_{st} a_{s't'} \sigma(s-s')\sigma(t-t') = \sum_{s,t,s',t'=1}^T a_{st} a_{s't'} \sigma(s-t')\sigma(t-s'),$$

где $a_{s't'} = a_{t's'}$. Два первых выражения для $\text{Var } Q$ получаются непосредственно. Спектральное представление функции $\sigma(h)$ дает

$$\begin{aligned}
(42) \quad \text{Var } Q &= 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{s,t,s',t'=1}^T a_{st} a_{s't'} e^{i[\lambda(s-s') - \nu(t-t')]} f(\lambda) f(\nu) d\lambda d\nu + \\
&\quad + \sum_{s,t,s',t'=1}^T a_{st} a_{s't'} \kappa(t-s, s'-s, t'-s),
\end{aligned}$$

так как $f(\nu) = f(-\nu)$. Отсюда вытекают два последних выражения для $\text{Var } Q$. ■

Пусть

$$(43) \quad P = \sum_{s,t=1}^T b_{st} (y_s - \mu)(y_t - \mu),$$

ТЕОРЕМА 8.2.5.

$$\begin{aligned}
 (44) \quad \text{Cov}(Q, P) &= \sum_{s, t, s', t'=1}^T a_{st} b_{s't'} [2\sigma(s-s')\sigma(t-t') + \\
 &\quad + \kappa(t-s, s'-s, t'-s)] = \\
 &= 2 \sum_{r, p=-(T-1)}^{T-1} \sigma(r)\sigma(p) \sum_{s \in S_r} \sum_{t \in S_p} a_{st} b_{s+r, t+p} + \\
 &\quad + \sum_{r, p=-(T-1)}^{T-1} \sum_{s \in S_r} \sum_{t \in S_p} a_{st} b_{s+r, t+p} \kappa(t-s, r, t-s+p) = \\
 &= 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{s, t=1}^T a_{st} e^{i(\lambda s - \nu t)} \sum_{s', t'=1}^T b_{s't'} e^{-i(\lambda s' - \nu t')} f(\lambda) \times \\
 &\quad \times f(\nu) d\lambda d\nu + \sum_{s, t, s', t'=1}^T a_{st} b_{s't'} \kappa(t-s, s'-s, t'-s) = \\
 &= 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} A(\lambda, \nu) \bar{B}(\lambda, \nu) f(\lambda) f(\nu) d\lambda d\nu + \\
 &\quad + \sum_{s, t, s', t'=1}^T a_{st} b_{s't'} \kappa(t-s, s'-s, t'-s),
 \end{aligned}$$

где

$$(45) \quad \bar{B}(\lambda, \nu) = \sum_{s', t'=1}^T b_{s't'} e^{-i(\lambda s' - \nu t')}.$$

ДОКАЗАТЕЛЬСТВО.

$$\begin{aligned}
 (46) \quad \mathbb{E}QP &= \sum_{s, t, s', t'=1}^T a_{st} b_{s't'} \mathbb{E}(y_s - \mu)(y_t - \mu)(y_{s'} - \mu)(y_{t'} - \mu) = \\
 &= \sum_{s, t, s', t'=1}^T a_{st} b_{s't'} [\sigma(s-t)\sigma(s'-t') + \sigma(s-s')\sigma(t-t') + \\
 &\quad + \sigma(s-t')\sigma(t-s') + \kappa(t-s, s'-s, t'-s)].
 \end{aligned}$$

Из этой формулы следуют два первых выражения для $\text{Cov}(Q, P)$. Кроме того,

$$\begin{aligned}
 (47) \quad \text{Cov}(Q, P) &= 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{s, t, s', t'=1}^T a_{st} b_{s't'} e^{i[\lambda(s-s') - \nu(t-t')]} \times \\
 &\quad \times f(\lambda) f(\nu) d\lambda d\nu + \sum_{s, t, s', t'=1}^T a_{st} b_{s't'} \kappa(t-s, s'-s, t'-s).
 \end{aligned}$$

Это дает нам два последних выражения для $\text{Cov}(Q, P)$. Заметим, что теоремы 8.2.4 и 8.2.5 являются обобщениями лемм 3.4.3, 3.4.4 и 3.4.5. ■

Если процесс гауссовский, то семиинварианты четвертого порядка $\kappa(h, r, s) = 0$. (Они могут быть нулями и в случае негауссовского процесса.)

Вернемся теперь к моментам первого и второго порядков оценок величины $\sigma(h)$. Мы не будем пользоваться теоремами 8.2.4 и 8.2.5, а вычислим средние значения, дисперсии и ковариации непосредственно (поскольку каждую ковариацию можно записать одинарной суммой вместо двойной).

Если μ известно, оценка C_h величины $\sigma(h)$ является несмещенной,

$$(48) \quad \mathbb{E}C_h = \frac{1}{T-h} \sum_{t=1}^{T-h} \mathbb{E}(y_t - \mu)(y_{t+h} - \mu) = \sigma(h),$$

$$h = 0, 1, \dots, T-1,$$

и, следовательно,

$$(49) \quad \mathbb{E}c_h = \frac{T-h}{T} \sigma(h) = \left(1 - \frac{h}{T}\right) \sigma(h), \quad h = 0, 1, \dots, T-1.$$

Смещение c_h мало для малого h и относительно велико для большого h (когда относительная выборочная вариабельность велика).

Математическое ожидание величины C_h^* ($h \geq 0$) равно

$$(50) \quad \mathbb{E}C_h^* = \frac{1}{T-h} \mathbb{E} \left\{ \sum_{t=1}^{T-h} (y_t - \mu)(y_{t+h} - \mu) - \frac{1}{T} \sum_{t=1}^{T-h} \sum_{s=1}^T (y_t - \mu) \times \right. \\ \times (y_s - \mu) - \frac{1}{T} \sum_{t=1}^{T-h} \sum_{s=1}^T (y_{t+h} - \mu)(y_s - \mu) + \\ \left. + \frac{T-h}{T^2} \sum_{t,s=1}^T (y_t - \mu)(y_s - \mu) \right\} = \\ = \sigma(h) - \frac{1}{T(T-h)} \sum_{t=1}^{T-h} \sum_{s=1}^T [\sigma(t-s) + \sigma(t+h-s)] + \\ + \frac{1}{T^2} \sum_{t,s=1}^T \sigma(t-s).$$

После очевидных, но весьма утомительных вычислений получаем следующие соотношения:

$$(51) \quad \mathbb{E}C_0^* = \sigma(0) - \frac{1}{T} \left\{ \sigma(0) + 2 \sum_{r=1}^{T-1} \left(1 - \frac{r}{T}\right) \sigma(r) \right\},$$

$$\begin{aligned}
 (52) \quad \mathbb{E}C_h^* = \sigma(h) - \frac{1}{T} & \left\{ \sigma(0) + 2 \sum_{r=1}^h \left[1 - \frac{rh}{T(T-h)} \right] \sigma(r) + \right. \\
 & + 2 \sum_{r=h+1}^{T-h-1} \left[1 - \frac{rh}{T(T-h)} - \frac{r-h}{T-h} \right] \sigma(r) + \\
 & \left. + 2 \sum_{r=T-h}^{T-1} \frac{(T-r)h}{T(T-h)} \sigma(r) \right\}, \quad 1 \leq h < T-h-1,
 \end{aligned}$$

$$\begin{aligned}
 (53) \quad \mathbb{E}C_h^* = \sigma(h) - \frac{1}{T} & \left\{ \sigma(0) + 2 \sum_{r=1}^h \left[1 - \frac{rh}{T(T-h)} \right] \sigma(r) + \right. \\
 & \left. + 2 \sum_{r=h+1}^{T-1} \frac{(T-r)h}{T(T-h)} \sigma(r) \right\}, \quad 1 \leq h = T-h-1,
 \end{aligned}$$

$$\begin{aligned}
 (54) \quad \mathbb{E}C_h^* = \sigma(h) - \frac{1}{T} & \left\{ \sigma(0) + 2 \sum_{r=1}^{T-h-1} \left[1 - \frac{rh}{T(T-h)} \right] \sigma(r) + \right. \\
 & + 2 \sum_{r=T-h}^h \frac{r}{T} \sigma(r) + 2 \sum_{r=h+1}^{T-1} \frac{(T-r)h}{T(T-h)} \sigma(r) \Big\}, \\
 & T-h-1 < h < T-1,
 \end{aligned}$$

$$(55) \quad \mathbb{E}C_{T-1}^* = \sigma(T-1) - \frac{1}{T} \left\{ \sigma(0) + 2 \sum_{r=1}^{T-1} \frac{r}{T} \sigma(r) \right\}.$$

Приведенные оценки являются смещенными, порядок смещения равен $1/T$. Это будет показано в § 8.3.

Математическое ожидание величины C_h^* можно также выразить с помощью спектральной плотности. Из (50) имеем

$$\begin{aligned}
 (56) \quad \mathbb{E}C_h^* &= \int_{-\pi}^{\pi} \left\{ \cos \lambda h - \frac{1}{T(T-h)} \sum_{t=1}^{T-h} \sum_{s=1}^t [e^{i\lambda(t-s)} + e^{i\lambda(t-s+h)}] + \right. \\
 & \quad \left. + \frac{1}{T^2} \sum_{s,t=1}^T e^{i\lambda(t-s)} \right\} f(\lambda) d\lambda = \\
 &= \int_{-\pi}^{\pi} \left\{ \cos \lambda h - 2 \frac{\sin \frac{1}{2} \lambda T \sin \frac{1}{2} \lambda (T-h)}{T \sin \frac{1}{2} \lambda (T-h) \sin \frac{1}{2} \lambda} \cos \frac{1}{2} \lambda h + \right. \\
 & \quad \left. + \left(\frac{\sin \frac{1}{2} \lambda T}{T \sin \frac{1}{2} \lambda} \right)^2 \right\} f(\lambda) d\lambda.
 \end{aligned}$$

Математическое ожидание величины $\tilde{C}_h$ равно

$$\begin{aligned}
 (57) \quad \mathbb{E}\tilde{C}_h &= \frac{1}{T-h} \mathbb{E} \sum_{t=1}^{T-h} [(y_t - \mu) - (\bar{y}_h - \mu)] [(y_{t+h} - \mu) - (\bar{y}_{h+} - \mu)] = \\
 &= \frac{1}{T-h} \mathbb{E} \left[\sum_{t=1}^{T-h} (y_t - \mu)(y_{t+h} - \mu) - (T-h) \times \right. \\
 &\quad \left. \times (\bar{y}_h - \mu)(\bar{y}_{h+} - \mu) \right] = \\
 &= \sigma(h) - \frac{1}{(T-h)^2} \mathbb{E} \sum_{t=1}^{T-h} \sum_{s=h+1}^T (y_t - \mu)(y_s - \mu) = \\
 &= \sigma(h) - \frac{1}{(T-h)^2} \sum_{t=1}^{T-h} \sum_{s=h+1}^T \sigma(t-s).
 \end{aligned}$$

Так как $\tilde{C}_0 = C_0^*$, то $\mathbb{E}\tilde{C}_0$ выразится формулой (51). Поэтому из (57) следует

$$\begin{aligned}
 (58) \quad \mathbb{E}\tilde{C}_h &= \sigma(h) - \frac{1}{T-h} \left\{ \left(1 - \frac{h}{T-h}\right) \sigma(0) + 2 \left(1 - \frac{h}{T-h}\right) \sum_{r=1}^h \sigma(r) + \right. \\
 &\quad \left. + \sum_{r=h+1}^{T-2h-1} 2 \left(1 - \frac{r}{T-h}\right) \sigma(r) + \sum_{r=T-2h}^{T-1} \left(1 - \frac{r-h}{T-h}\right) \sigma(r) \right\}, \\
 &\quad 1 \leq h < T-2h-1.
 \end{aligned}$$

Другие случаи рассматриваются в упр. 10.

Эти математические ожидания можно выразить также при помощи спектральной плотности, а именно

$$\begin{aligned}
 (59) \quad \mathbb{E}\tilde{C}_h &= \int_{-\pi}^{\pi} \left[\cos \lambda h - \frac{1}{(T-h)^2} \sum_{t=1}^{T-h} \sum_{s=h+1}^T e^{i\lambda(t-s)} \right] f(\lambda) d\lambda = \\
 &= \int_{-\pi}^{\pi} \cos \lambda h \left\{ 1 - \left[\frac{\sin \frac{1}{2} \lambda (T-h)}{(T-h) \sin \frac{1}{2} \lambda} \right]^2 \right\} f(\lambda) d\lambda.
 \end{aligned}$$

Вернемся теперь к вычислению дисперсий и ковариаций этих оценок для величин $\sigma(0)$, $\sigma(1)$, ..., $\sigma(T-1)$.

ТЕОРЕМА 8.2.6.

$$\begin{aligned}
 (60) \quad &(T-h)(T-g) \text{Cov}(C_h, C_g) = \\
 &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2}(\lambda - \nu)(T-h)}{\sin \frac{1}{2}(\lambda - \nu)} \frac{\sin \frac{1}{2}(\lambda - \nu)(T-g)}{\sin \frac{1}{2}(\lambda - \nu)} [e^{i \frac{1}{2}(\lambda + \nu)(g-h)} +
 \end{aligned}$$

$$\begin{aligned}
& + e^{-i \frac{1}{2} (\lambda + \nu) (g+h)} |f(\lambda) f(\nu) d\lambda d\nu + \sum_{t=1}^{T-h} \sum_{s=1}^{T-g} \kappa(h, s-t, s-t+g) = \\
& = 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2} (\lambda - \nu) (T-h) \sin \frac{1}{2} (\lambda - \nu) (T-g)}{\sin^2 \frac{1}{2} (\lambda - \nu)} \times \\
& \times \cos \frac{1}{2} (\lambda + \nu) h \cos \frac{1}{2} (\lambda + \nu) g f(\lambda) f(\nu) d\lambda d\nu + \\
& + \sum_{t=1}^{T-h} \sum_{s=1}^{T-g} \kappa(h, s-t, s-t+g) = \\
& = \sum_{r=\max(g-h, 0)+1}^{T-h-1} (T-h-r) [\sigma(r) \sigma(r+h-g) + \\
& + \sigma(r-g) \sigma(r+h) + \kappa(h, -r, g-r)] + \\
& + \sum_{r=\min(g-h, 0)}^{\max(g-h, 0)} [T - \max(g, h)] [\sigma(r) \sigma(r+h-g) + \\
& + \sigma(r-g) \sigma(r+h) + \kappa(h, -r, g-r)] + \\
& + \sum_{r=-(T-g-1)}^{\min(g-h, 0)-1} (T-g-|r|) [\sigma(r) \sigma(r+h-g) + \\
& + \sigma(r-g) \sigma(r+h) + \kappa(h, -r, g-r)], \\
& 0 \leq h < T, 0 \leq g < T.
\end{aligned}$$

ДОКАЗАТЕЛЬСТВО. Имеем

$$\begin{aligned}
(61) \quad & (T-h)(T-g) \mathbb{E} C_h C_g = \\
& = \mathbb{E} \sum_{t=1}^{T-h} \sum_{s=1}^{T-g} (y_t - \mu)(y_{t+h} - \mu)(y_s - \mu)(y_{s+g} - \mu) = \\
& = \sum_{t=1}^{T-h} \sum_{s=1}^{T-g} v(h, s-t, s-t+g) = \\
& = \sum_{t=1}^{T-h} \sum_{s=1}^{T-g} [\sigma(h) \sigma(g) + \sigma(t-s) \sigma(t+h-s-g) + \\
& + \sigma(t-s-g) \sigma(t+h-s) + \kappa(h, s-t, s-t+g)], \\
& 0 \leq h < T, 0 \leq g < T,
\end{aligned}$$

и

$$\begin{aligned}
(62) \quad & (T-h)(T-g) \text{Cov}(C_h, C_g) = \\
& = \sum_{t=1}^{T-h} \sum_{s=1}^{T-g} [\sigma(t-s) \sigma(t+h-s-g) + \sigma(t-s-g) \sigma(t+h-s) +
\end{aligned}$$

$$\begin{aligned}
& + \kappa(h, s-t, s-t+g)] = \\
& = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{t=1}^{T-h} \sum_{s=1}^{T-g} [e^{i\lambda(t-s)} - i\nu(t-s+h-g) + \\
& + e^{i\lambda(t-s-g)} - i\nu(t-s+h)}] f(\lambda) f(\nu) d\lambda d\nu + \sum_{t=1}^{T-h} \sum_{s=1}^{T-g} \kappa(h, s-t, s-t+g) = \\
& = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2}(\lambda-\nu)(T-h) \sin \frac{1}{2}(\lambda-\nu)(T-g)}{\sin \frac{1}{2}(\lambda-\nu) \sin \frac{1}{2}(\lambda-\nu)} [e^{i \frac{1}{2}(\lambda+\nu)(g-h)} + \\
& + e^{-i \frac{1}{2}(\lambda+\nu)(g+h)}] f(\lambda) f(\nu) d\lambda d\nu + \sum_{t=1}^{T-h} \sum_{s=1}^{T-g} \kappa(h, s-t, s-t+g).
\end{aligned}$$

Отсюда вытекают и все последующие равенства в (60). ■

В частности,

$$\begin{aligned}
(63) \quad (T-h) \text{Var } C_h &= \sum_{r=-(T-h-1)}^{T-h-1} \left(1 - \frac{|r|}{T-h}\right) \times \\
&\times [\sigma^2(r) + \sigma(r+h)\sigma(r-h) + \kappa(h, -r, h-r)] = \\
&= \frac{1}{T-h} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left[\frac{\sin \frac{1}{2}(\lambda-\nu)(T-h)}{\sin \frac{1}{2}(\lambda-\nu)} \right]^2 \times \\
&\times [1 + e^{-i(\nu+\lambda)h}] f(\lambda) f(\nu) d\lambda d\nu + \\
&+ \sum_{t,s=1}^{T-h} \kappa(h, s-t, s-t+h) = \\
&= 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{\sin^2 \frac{1}{2}(\lambda-\nu)(T-h)}{(T-h) \sin^2 \frac{1}{2}(\lambda-\nu)} \times \\
&\times \cos^2 \frac{1}{2}(\lambda+\nu) h f(\lambda) f(\nu) d\lambda d\nu + \\
&+ \sum_{t,s=1}^{T-h} \kappa(h, s-t, s-t+h), \\
&h = 0, 1, \dots, T-1.
\end{aligned}$$

Если процесс гауссовский, то семинварианты четвертого порядка в приведенных выше выражениях обращаются в нуль.

Вычисление ковариаций для выборочных ковариаций, когда имеются отклонения от выборочного среднего, более трудоемко. Рассмотрим

$$(64) \quad \begin{aligned} \tilde{C}_h &= \frac{1}{T-h} \sum_{t=1}^{T-h} (y_t - \bar{y}_h) (y_{t+h} - \bar{y}_{h+h}) = \\ &= \frac{1}{T-h} \left[\sum_{t=1}^{T-h} y_t y_{t+h} - \frac{1}{T-h} \sum_{t,s=1}^{T-h} y_t y_{s+h} \right]. \end{aligned}$$

Тогда

$$(65) \quad \begin{aligned} (T-h)(T-g) \operatorname{Cov}(\tilde{C}_h, \tilde{C}_g) &= \\ &= \sum_{t=1}^{T-h} \sum_{t'=1}^{T-g} [\sigma(t-t') \sigma(t+h-t'-g) + \\ &\quad + \sigma(t-t'-g) \sigma(t+h-t') + \kappa(h, t'-t, t'-t+g)] - \\ &\quad - \frac{1}{T-g} \sum_{t=1}^{T-h} \sum_{t',s'=1}^{T-g} [\sigma(t-t') \sigma(t+h-s'-g) + \\ &\quad + \sigma(t-s'-g) \sigma(t+h-t') + \kappa(h, t'-t, s'-t+g)] - \\ &\quad - \frac{1}{T-h} \sum_{t,s=1}^{T-h} \sum_{t'=1}^{T-g} [\sigma(t-t') \sigma(s+h-t'-g) + \\ &\quad + \sigma(t-t'-g) \sigma(s+h-t') + \\ &\quad + \kappa(s+h-t, t'-t, t'+g-t)] + \\ &\quad + \frac{1}{(T-h)(T-g)} \sum_{t,s=1}^{T-h} \sum_{t',s'=1}^{T-g} [\sigma(t-t') \sigma(s+h-s'-g) + \\ &\quad + \sigma(t-s'-g) \sigma(s+h-t') + \\ &\quad + \kappa(s+h-t, t'-t, s'+g-t)] = \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left[\frac{\sin \frac{1}{2}(\lambda + \nu)(T-h)}{\sin \frac{1}{2}(\lambda - \nu)} - \frac{\sin \frac{1}{2} \lambda (T-h) \sin \frac{1}{2} \nu (T-h)}{(T-h) \sin \frac{1}{2} \lambda \sin \frac{1}{2} \nu} \right] \times \\ &\quad \times \left[\frac{\sin \frac{1}{2}(\lambda - \nu)(T-g)}{\sin \frac{1}{2}(\lambda - \nu)} - \frac{\sin \frac{1}{2} \lambda (T-g) \sin \frac{1}{2} \nu (T-g)}{(T-g) \sin \frac{1}{2} \lambda \sin \frac{1}{2} \nu} \right] \times \\ &\quad \times [e^{i \frac{1}{2}(\lambda + \nu)(g-h)} + e^{-i \frac{1}{2}(\lambda + \nu)(g+h)}] f(\lambda) f(\nu) d\lambda d\nu + \end{aligned}$$

$$\begin{aligned}
& + \sum_{t=1}^{T-h} \sum_{t'=1}^{T-g} \left[\kappa(h, t' - t, t' - t + g) - \right. \\
& - \frac{1}{T-g} \sum_{s'=1}^{T-g} \kappa(h, t' - t, s' - t + g) - \\
& - \frac{1}{T-h} \sum_{s=1}^{T-h} \kappa(s + h - t, t' - t, t' + g - t) + \\
& \left. + \frac{1}{(T-h)(T-g)} \sum_{s=1}^{T-h} \sum_{s'=1}^{T-g} \kappa(s + h - t, t' - t, s' + g - t) \right], \\
& 0 \leq h < T, 0 \leq g < T.
\end{aligned}$$

(См. упр. 11.) Ковариации величин C_h^* могут быть найдены аналогичным образом.

8.2.3. Моменты выборочной спектральной плотности

Рассмотрим теперь первый и второй моменты выборочной спектральной плотности $I(\lambda)$, определяемой формулой (21), если среднее значение μ известно. Первый момент равен

$$\begin{aligned}
(66) \quad \mathcal{E}I(\lambda) &= \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) \sigma(r) \cos \lambda r = \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) \cos vr \cos \lambda r f(v) dv, \\
& \quad -\pi \leq \lambda \leq \pi,
\end{aligned}$$

так как $\mathcal{E}c_r = (T - |r|) \sigma(r)/T$. Его можно переписать в виде

$$\begin{aligned}
(67) \quad \mathcal{E}I(\lambda) &= \frac{1}{2\pi T} \int_{-\pi}^{\pi} \sum_{t,s=1}^T \cos v(t-s) \cos \lambda(t-s) f(v) dv = \\
&= \frac{1}{2\pi T} \int_{-\pi}^{\pi} \left| \sum_{t=1}^T e^{i(v-\lambda)t} \right|^2 f(v) dv = \\
&= \frac{1}{2\pi T} \int_{-\pi}^{\pi} \left[\frac{\sin \frac{1}{2}(v-\lambda)T}{\sin \frac{1}{2}(v-\lambda)} \right]^2 f(v) dv.
\end{aligned}$$

ТЕОРЕМА 8.2.7.

$$(68) \quad \mathcal{G}I(\lambda) = \int_{-\pi}^{\pi} k_T(v - \lambda) f(v) dv,$$

где

$$(69) \quad k_T(v) = \frac{\sin^2 \frac{1}{2} vT}{2\pi T \sin^2 \frac{1}{2} v}.$$

Ядро $k_T(v)$ симметрично, имеет максимум при $v = 0$ и периодически с периодом 2π ; $k_T(0) = T/(2\pi)$. Это ядро называется ядром Фейера [см., например, Ланцош (1956, гл. IV, разд. 2)]. Математическое ожидание величины $I(\lambda)$ не равно $f(\lambda)$, а является взвешенным интегралом функции $f(\lambda)$ [или $g(\lambda)$]. Так как $f(\lambda)$ есть бесконечный ряд с коэффициентами $\sigma(r)$, $r = 0, 1, \dots$, то, вообще говоря, $I(\lambda)$ не может быть несмещенной оценкой, так как $I(\lambda)$ включает только T величин c_r . Мы покажем в следующем параграфе, что $\lim_{T \rightarrow \infty} \mathcal{G}I(\lambda) = f(\lambda)$.

Заметим также, что

$$(70) \quad \begin{aligned} \mathcal{G}I^*(\lambda) &= \frac{1}{2\pi} \sum_{h=-(T-1)}^{T-1} \left(1 - \frac{|h|}{T}\right) \mathcal{G}C_h^* \cos \lambda h = \\ &= \frac{1}{2\pi} \sum_{h=-(T-1)}^{T-1} \left(1 - \frac{|h|}{T}\right) \sigma(h) \cos \lambda h - \\ &\quad - \frac{1}{2\pi T} \sum_{h=-(T-1)}^{T-1} \left(1 - \frac{|h|}{T}\right) \cos \lambda h \sum_{r=-(T-1)}^{T-1} C(h, r) \sigma(r), \end{aligned}$$

$-\pi \leq \lambda \leq \pi,$

где коэффициенты $C(h, r)$ даны в формулах (51) — (55).

Для вычисления этого математического ожидания в виде интеграла от спектральной плотности запишем

$$(71) \quad \begin{aligned} \mathcal{G}I^*(\lambda) &= \frac{1}{2\pi T} \mathcal{G} \sum_{s,t=1}^T (y_t y_s - y_t \bar{y} - \bar{y} y_s + \bar{y}^2) e^{i\lambda(t-s)} = \\ &= \frac{1}{2\pi T} \left[\sum_{s,t=1}^T \sigma(t-s) e^{i\lambda(t-s)} - \right. \\ &\quad \left. - \frac{1}{T} \sum_{t,r=1}^T \sigma(t-r) e^{i\lambda t} \frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} \lambda} e^{-i \frac{1}{2} \lambda (T+1)} - \right. \end{aligned}$$

$$\begin{aligned}
& - \frac{1}{T} \sum_{r,s=1}^T \sigma(r-s) e^{-i\lambda s} \frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} \lambda} e^{i \frac{1}{2} \lambda (T+1)} + \\
& + \text{Var } \bar{y} \left(\frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} \lambda} \right)^2 \Bigg] = \frac{1}{2\pi T} \int_{-\pi}^{\pi} \left[\sum_{s,t=1}^T e^{i(\lambda-\nu)(t-s)} - \right. \\
& - \frac{1}{T} \frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} \lambda} \left\{ \sum_{t,r=1}^T e^{i((\lambda-\nu)t + t\nu r - i \frac{1}{2} \lambda (T+1))} + \right. \\
& + \left. \sum_{r,s=1}^T e^{i(\nu-\lambda)s - t\nu r + i \frac{1}{2} \lambda (T+1)} \right\} + \\
& + \left. \left(\frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} \lambda} \right)^2 \left(\frac{\sin \frac{1}{2} \nu T}{T \sin \frac{1}{2} \nu} \right)^2 \right] f(\nu) d\nu = \\
& = \frac{1}{2\pi T} \int_{-\pi}^{\pi} \left\{ \frac{\sin \frac{1}{2} (\nu - \lambda) T}{\sin \frac{1}{2} (\nu - \lambda)} - \right. \\
& - \left. \frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} \lambda} \frac{\sin \frac{1}{2} \nu T}{T \sin \frac{1}{2} \nu} \right\}^2 f(\nu) d\nu, \quad \lambda \neq 0.
\end{aligned}$$

Теперь вернемся к ковариациям величины $I(\lambda)$.

ТЕОРЕМА 8.2.8.

(72) $\text{Cov}[I(\lambda), I(\lambda')] =$

$$\begin{aligned}
& = \left[\int_{-\pi}^{\pi} k_T(\nu + \lambda, \nu - \lambda') f(\nu) d\nu \right]^2 + \\
& + \left[\int_{-\pi}^{\pi} k_T(\nu + \lambda, \nu + \lambda') f(\nu) d\nu \right]^2 + \\
& + \frac{1}{4\pi^2 T^2} \sum_{t,s,t',s'=1}^T \cos \lambda(t-s) \cos \lambda'(t'-s') \kappa(s-t, t'-t, s'-t),
\end{aligned}$$

где

$$(73) \quad k_T(\lambda, \nu) = \frac{\sin \frac{1}{2} \lambda T \sin \frac{1}{2} \nu T}{2\pi T \sin \frac{1}{2} \lambda \sin \frac{1}{2} \nu}.$$

Доказательство. Так как (для $\mu = 0$)

$$(74) \quad 2\pi T I(\lambda) = \sum_{t,s=1}^T y_t y_s (\cos \lambda t \cos \lambda s + \sin \lambda t \sin \lambda s) = \\ = \sum_{t,s=1}^T y_t y_s \cos \lambda(t-s) = \sum_{t,s=1}^T y_t y_s e^{i\lambda(t-s)}$$

и

$$(75) \quad \sigma(p-r) = \int_{-\pi}^{\pi} \cos \nu(p-r) f(\nu) d\nu = \int_{-\pi}^{\pi} e^{i\nu(p-r)} f(\nu) d\nu,$$

то имеем

$$(76) \quad (2\pi T)^2 \text{Cov}[I(\lambda), I(\lambda')] = \\ = \sum_{t,s,t',s'=1}^T e^{i\lambda(t-s)} e^{i\lambda'(t'-s')} \text{Cov}(y_t y_s, y_{t'} y_{s'}) = \\ = \sum_{t,s,t',s'=1}^T e^{i\lambda(t-s) + i\lambda'(t'-s')} [\sigma(t-t') \sigma(s-s') + \\ + \sigma(t-s') \sigma(s-t') + \kappa(s-t, t'-t, s'-t)] = \\ = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{t,s,t',s'=1}^T e^{i\lambda(t-s) + i\lambda'(t'-s')} [e^{i\nu(t-t')} e^{i\mu(s-s')} + \\ + e^{i\nu(t-s')} e^{i\mu(s-t')}] f(\nu) f(\mu) d\nu d\mu + \\ + \sum_{t,s,t',s'=1}^T \cos \lambda(t-s) \cos \lambda'(t'-s') \kappa(s-t, t'-t, s'-t) = \\ = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left[\frac{\sin \frac{1}{2}(\lambda + \nu)T}{\sin \frac{1}{2}(\lambda + \nu)} \frac{\sin \frac{1}{2}(\mu - \lambda)T}{\sin \frac{1}{2}(\mu - \lambda)} \frac{\sin \frac{1}{2}(\lambda' - \nu)T}{\sin \frac{1}{2}(\lambda' - \nu)} \times \right. \\ \left. \times \frac{\sin \frac{1}{2}(\lambda' + \mu)T}{\sin \frac{1}{2}(\lambda' + \mu)} + \frac{\sin \frac{1}{2}(\lambda + \nu)T \sin \frac{1}{2}(\mu - \lambda)T}{\sin \frac{1}{2}(\lambda + \nu) \sin \frac{1}{2}(\mu - \lambda)} \times \right.$$

$$\begin{aligned}
& \times \frac{\sin \frac{1}{2} (\lambda' - \mu) T}{\sin \frac{1}{2} (\lambda' - \mu)} \frac{\sin \frac{1}{2} (\lambda' + \nu) T}{\sin \frac{1}{2} (\lambda' + \nu)} \Big] f(\nu) f(\mu) d\nu d\mu + \\
& + \sum_{t, s, t', s'=1}^T \cos \lambda (t - s) \cos \lambda' (t' - s') \kappa(s - t, t' - t, s' - t) = \\
& = \left\{ \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2} (\nu + \lambda) T}{\sin \frac{1}{2} (\nu + \lambda)} \frac{\sin \frac{1}{2} (\nu - \lambda') T}{\sin \frac{1}{2} (\nu - \lambda')} f(\nu) d\nu \right\}^2 + \\
& + \left\{ \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2} (\nu + \lambda) T}{\sin \frac{1}{2} (\nu + \lambda)} \frac{\sin \frac{1}{2} (\nu + \lambda') T}{\sin \frac{1}{2} (\nu + \lambda')} f(\nu) d\nu \right\}^2 + \\
& + \sum_{t, s, t', s'=1}^T \cos \lambda (t - s) \cos \lambda' (t' - s') \kappa(s - t, t' - t, s' - t),
\end{aligned}$$

откуда и следует (72). ■

Другая форма записи для ковариации имеет вид

$$\begin{aligned}
(77) \quad \text{Cov} [I(\lambda), I(\lambda')] &= \frac{1}{4\pi^2} \sum_{r, p=-(T-1)}^{T-1} \text{Cov}(c_r, c_p) \cos \lambda r \cos \lambda' p = \\
&= \frac{1}{4\pi^2 T^2} \sum_{r, p=-(T-1)}^{T-1} \sum_{t=1}^{T-|r|} \sum_{s=1}^{T-|p|} [\sigma(t-s) \sigma(t-s+|r|-|p|) + \\
&\quad + \sigma(t-s-|p|) \sigma(t-s+|r|) + \\
&\quad + \kappa(|r|, s-t, s-t+|p|)] \cos \lambda r \cos \lambda' p.
\end{aligned}$$

Так как $k_T(\lambda, \lambda) = k_T(\lambda)$, то (72) для $\lambda = \lambda'$ запишется следующим образом:

$$\begin{aligned}
(78) \quad \text{Var} [I(\lambda)] &= \left[\int_{-\pi}^{\pi} k_T(\nu + \lambda, \nu - \lambda) f(\nu) d\nu \right]^2 + \\
&+ \left[\int_{-\pi}^{\pi} k_T(\nu - \lambda) f(\nu) d\nu \right]^2 + \\
&+ \frac{1}{4\pi^2 T^2} \sum_{t, s, t', s'=1}^T \cos \lambda (t - s) \cos \lambda (t' - s') \times \\
&\times \kappa(s - t, t' - t, s' - t) = \\
&= \left[\int_{-\pi}^{\pi} k_T(\nu + \lambda, \nu - \lambda) f(\nu) d\nu \right]^2 + [\mathcal{G}I(\lambda)]^2 +
\end{aligned}$$

$$+ \frac{1}{4\pi^2 T^2} \sum_{t,s,t',s'=1}^T \cos \lambda (t-s) \cos \lambda (t'-s') \times \\ \times \kappa(s-t, t'-t, s'-t).$$

Нам будет полезно также знать ковариации косинус- и синус-преобразований $A(\lambda)$ и $B(\lambda)$.

ТЕОРЕМА 8.2.9.

$$(79) \quad \text{Cov}[A(\lambda), A(\lambda')] = \\ = \frac{4\pi}{T} \int_{-\pi}^{\pi} \left[k_T(v-\lambda', v+\lambda) \cos \frac{1}{2}(\lambda'+\lambda)(T+1) + \right. \\ \left. + k_T(v-\lambda', v-\lambda) \cos \frac{1}{2}(\lambda'-\lambda)(T+1) \right] f(v) dv,$$

$$(80) \quad \text{Cov}[A(\lambda), B(\lambda')] = \\ = \frac{4\pi}{T} \int_{-\pi}^{\pi} \left[k_T(v-\lambda', v+\lambda) \sin \frac{1}{2}(\lambda'+\lambda)(T+1) + \right. \\ \left. + k_T(v-\lambda', v-\lambda) \sin \frac{1}{2}(\lambda'-\lambda)(T+1) \right] f(v) dv,$$

$$(81) \quad \text{Cov}[B(\lambda), B(\lambda')] = \\ = -\frac{4\pi}{T} \int_{-\pi}^{\pi} \left[k_T(v+\lambda, v+\lambda') \cos \frac{1}{2}(\lambda-\lambda')(T+1) - \right. \\ \left. - k_T(v+\lambda, v-\lambda') \cos \frac{1}{2}(\lambda+\lambda')(T+1) \right] f(v) dv.$$

ДОКАЗАТЕЛЬСТВО. Имеем

$$(82) \quad \mathcal{G}A(\lambda) A(v) + i\mathcal{G}A(\lambda) B(v) = \mathcal{G}A(\lambda) [A(v) + iB(v)] = \\ = \frac{4}{T^2} \sum_{t,s=1}^T \sigma(t-s) \cos \lambda t e^{i\lambda s} = \\ = \frac{4}{T^2} \int_{-\pi}^{\pi} \sum_{t,s=1}^T e^{i\mu(t-s)+i\lambda s} \cos \lambda t f(\mu) d\mu = \\ = \frac{4}{T^2} \int_{-\pi}^{\pi} \sum_{t=1}^T e^{i\lambda t} \cos \lambda t \frac{\sin \frac{1}{2}(v-\mu)T}{\sin \frac{1}{2}(v-\mu)} e^{i\frac{1}{2}(v-\mu)(T+1)} f(\mu) d\mu =$$

$$\begin{aligned}
&= \frac{4}{T^2} \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2} (v - \mu) T}{\sin \frac{1}{2} (v - \mu)} e^{i \frac{1}{2} (v - \mu)(T+1)} \sum_{t=1}^T e^{i \mu t} \frac{1}{2} [e^{i \lambda t} + e^{-i \lambda t}] f(\mu) d\mu = \\
&= \frac{2}{T^2} \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2} (v - \mu) T}{\sin \frac{1}{2} (v - \mu)} e^{i \frac{1}{2} v(T+1)} \left[\frac{\sin \frac{1}{2} (\lambda + \mu) T}{\sin \frac{1}{2} (\lambda + \mu)} e^{i \frac{1}{2} \lambda(T+1)} + \right. \\
&\quad \left. + \frac{\sin \frac{1}{2} (\lambda - \mu) T}{\sin \frac{1}{2} (\lambda - \mu)} e^{-i \frac{1}{2} \lambda(T+1)} \right] f(\mu) d\mu = \\
&= \frac{2}{T^2} \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2} (v - \mu) T}{\sin \frac{1}{2} (v - \mu)} \left[\frac{\sin \frac{1}{2} (\lambda + \mu) T}{\sin \frac{1}{2} (\lambda + \mu)} e^{i \frac{1}{2} (v + \lambda)(T+1)} + \right. \\
&\quad \left. + \frac{\sin \frac{1}{2} (\lambda - \mu) T}{\sin \frac{1}{2} (\lambda - \mu)} e^{i \frac{1}{2} (v - \lambda)(T+1)} \right] f(\mu) d\mu,
\end{aligned}$$

и

$$\begin{aligned}
(83) \quad \mathfrak{E} A(\lambda) B(v) + i \mathfrak{E} B(\lambda) B(v) &= \mathfrak{E} [A(\lambda) + i B(\lambda)] B(v) = \\
&= \frac{4}{T^2} \sum_{t,s=1}^T \sigma(t-s) e^{i \lambda t} \sin v s = \frac{4}{T^2} \int_{-\pi}^{\pi} \sum_{t,s=1}^T e^{i \mu(t-s) + i \lambda t} \sin v s f(\mu) d\mu = \\
&= \frac{4}{T^2} \int_{-\pi}^{\pi} \sum_{s=1}^T e^{-i \mu s} \sin v s \frac{\sin \frac{1}{2} (\lambda + \mu) T}{\sin \frac{1}{2} (\lambda + \mu)} e^{i \frac{1}{2} (\lambda + \mu)(T+1)} f(\mu) d\mu = \\
&= \frac{4}{T^2} \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2} (\lambda + \mu) T}{\sin \frac{1}{2} (\lambda + \mu)} e^{i \frac{1}{2} (\lambda + \mu)(T+1)} \sum_{s=1}^T e^{-i \mu s} \frac{1}{2i} [e^{i v s} - e^{-i v s}] f(\mu) d\mu = \\
&= \frac{2}{iT^2} \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2} (\lambda + \mu) T}{\sin \frac{1}{2} (\lambda + \mu)} e^{i \frac{1}{2} \lambda(T+1)} \left[\frac{\sin \frac{1}{2} (v - \mu) T}{\sin \frac{1}{2} (v - \mu)} e^{i \frac{1}{2} v(T+1)} - \right.
\end{aligned}$$

$$\begin{aligned}
& - \frac{\sin \frac{1}{2} (v + \mu) T}{\sin \frac{1}{2} (v + \mu)} e^{-i \frac{1}{2} v(T+1)} \Big] f(\mu) d\mu = \\
& = \frac{2}{iT^2} \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2} (\lambda + \mu) T}{\sin \frac{1}{2} (\lambda + \mu)} \left[\frac{\sin \frac{1}{2} (v - \mu) T}{\sin \frac{1}{2} (v - \mu)} e^{i \frac{1}{2} (\lambda + v)(T+1)} - \right. \\
& \quad \left. - \frac{\sin \frac{1}{2} (v + \mu) T}{\sin \frac{1}{2} (v + \mu)} e^{i \frac{1}{2} (\lambda - v)(T+1)} \right] f(\mu) d\mu.
\end{aligned}$$

Теорема следует из рассмотрения действительной и мнимой частей (82) и (83). ■

8.3. АСИМПТОТИЧЕСКИЕ СРЕДНИЕ ЗНАЧЕНИЯ И КОВАРИАЦИИ ВЫБОРОЧНЫХ СРЕДНЕГО, КОВАРИАЦИЙ И СПЕКТРАЛЬНОЙ ПЛОТНОСТИ

8.3.1. Выборочное среднее

Выборочное среднее $\bar{y} = \sum_{t=1}^T y_t / T$ есть несмещенная оценка величины $\mu = \mathbb{E} y_t$ для процесса, стационарного в широком смысле, т. е. $\mathbb{E} \bar{y} = \mu$. Дисперсия величины $\bar{y}$ была приведена в разд. 8.2.2 в виде линейной комбинации величин $\sigma(0), \dots, \sigma(T-1)$ и в виде взвешенного интеграла функции $f(\lambda)$. Имеем

$$\begin{aligned}
(1) \quad T \text{Var } \bar{y} &= \sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) \sigma(r) = \\
&= \int_{-\pi}^{\pi} \frac{\sin^2 \frac{1}{2} \lambda T}{T \sin^2 \frac{1}{2} \lambda} f(\lambda) d\lambda.
\end{aligned}$$

ТЕОРЕМА 8.3.1. Если

$$(2) \quad \sum_{r=-\infty}^{\infty} \sigma(r) < \infty,$$

то

$$(3) \quad \lim_{T \rightarrow \infty} T \text{Var } \bar{y} = \sum_{r=-\infty}^{\infty} \sigma(r).$$

Если $f(\lambda)$ непрерывна при $\lambda = 0$, то

$$(4) \quad \lim_{T \rightarrow \infty} T \operatorname{Var} \bar{y} = 2\pi f(0).$$

Доказательство этой теоремы, использующее сумму в формуле (1), представляет собой частный случай утверждения, что суммирование по Чезаро сходящегося ряда не меняет его значения. Это эквивалентно следующей лемме.

ЛЕММА 8.3.1. Если $\sum_{r=1}^{\infty} a_r$ сходится, то

$$(5) \quad \lim_{T \rightarrow \infty} \sum_{r=1}^{T-1} \left(1 - \frac{r}{T}\right) a_r = \sum_{r=1}^{\infty} a_r.$$

Доказательство леммы. Из сходимости $\sum_{r=1}^{\infty} a_r$ следует, что для любого положительного ε существует число T_ε , такое, что для каждого $T \geq T_\varepsilon$

$$(6) \quad \left| \sum_{r=T}^{\infty} a_r \right| < \frac{1}{2} \varepsilon.$$

Из сходимости следует также, что существует число M , такое, что

$$(7) \quad \left| \sum_{r=T}^{\infty} a_r \right| < M, \quad T = 1, 2, \dots$$

Тогда для $T > T_\varepsilon$

$$\begin{aligned} (8) \quad \left| \sum_{r=1}^{\infty} a_r - \sum_{r=1}^{T-1} \left(1 - \frac{r}{T}\right) a_r \right| &= \left| \frac{T \sum_{r=T}^{\infty} a_r + \sum_{r=1}^{T-1} r a_r}{T} \right| = \\ &= \frac{\left| \sum_{r=1}^{T-1} \sum_{s=r+1}^{\infty} a_s + \sum_{s=1}^{\infty} a_s \right|}{T} \leq \\ &\leq \left\{ \sum_{r=1}^{T_\varepsilon-1} \left| \sum_{s=r+1}^{\infty} a_s \right| + \sum_{r=T_\varepsilon}^{T-1} \left| \sum_{s=r+1}^{\infty} a_s \right| + \left| \sum_{r=1}^{\infty} a_r \right| \right\} T^{-1} < \\ &< \frac{T_\varepsilon M}{T} + \frac{1}{2} \varepsilon. \end{aligned}$$

Для $T > T_\varepsilon$ и $T > T'_\varepsilon = 2T_\varepsilon M/\varepsilon$ формула (8) меньше ε . Это и доказывает лемму. ■

Заметим, что абсолютная сходимость (2) не является необходимой для (3).

Второе соотношение для предельной дисперсии следует из первого, если ряд Фурье для $f(\lambda)$ сходится к $f(0)$ при $\lambda = 0$. Тем не менее мы выведем этот результат непосредственно из интегрального представления в (1), так как такие методы доказательства будут полезны в дальнейшем. Взвешенный интеграл содержит ядро Фейера

$$(9) \quad k_T(\lambda) = \frac{\sin^2 \frac{1}{2} \lambda T}{2\pi T \sin^2 \frac{1}{2} \lambda}.$$

Интеграл от $k_T(\lambda)$ по $(-a, a)$ равен

$$\begin{aligned} (10) \quad \int_{-a}^a k_T(\lambda) d\lambda &= \int_{-a}^a \frac{\sin^2 \frac{1}{2} \lambda T}{2\pi T \sin^2 \frac{1}{2} \lambda} d\lambda = \int_{-a}^a \frac{1}{2\pi T} \sum_{s,t=1}^T e^{i\lambda(s-t)} d\lambda = \\ &= \int_{-a}^a \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) \cos \lambda r d\lambda = \\ &= \frac{1}{\pi} \left[a + 2 \sum_{r=1}^{T-1} \frac{T-r}{T} \sin ar \right]. \end{aligned}$$

В частности,

$$(11) \quad \int_{-\pi}^{\pi} k_T(\lambda) d\lambda = 1.$$

Поскольку $|\sin x| \leq 1$, то

$$(12) \quad k_T(\lambda) \leq \frac{1}{2\pi T \sin^2 \frac{1}{2} \lambda}$$

и отсюда

$$(13) \quad \lim_{T \rightarrow \infty} k_T(\lambda) = 0, \quad \lambda \neq 0.$$

Более того, так как $\sin \frac{1}{2} \lambda$ возрастает на $[0, \pi]$, то сходимость равномерна на любом отрезке $[a, \pi]$ для $a > 0$, т. е.

$$(14) \quad k_T(\lambda) \leq \frac{1}{2\pi T \sin^2 \frac{1}{2} a}, \quad 0 < a \leq |\lambda| \leq \pi.$$

Таким образом,

$$(15) \quad \int_a^\pi k_T(\lambda) d\lambda \leq \int_a^\pi \frac{d\lambda}{2\pi T \sin^2 \frac{1}{2} a} = \frac{\pi - a}{2\pi T \sin^2 \frac{1}{2} a} \rightarrow 0$$

при $T \rightarrow \infty$. Итак, доказана

ЛЕММА 8.3.2.

$$(16) \quad \lim_{T \rightarrow \infty} \int_{-a}^a k_T(\lambda) d\lambda = 1, \quad 0 < a \leq \pi,$$

$$(17) \quad \lim_{T \rightarrow \infty} \int_{-\pi}^{-a} k_T(\lambda) d\lambda = \lim_{T \rightarrow \infty} \int_a^\pi k_T(\lambda) d\lambda = 0, \quad 0 < a \leq \pi.$$

Одно из утверждений теоремы 8.3.1 вытекает из следующей леммы (формулировка которой дается в достаточно общем виде для того, чтобы ее можно было использовать и в других случаях).

ЛЕММА 8.3.3. Пусть выполняются следующие условия:
 $\int_{-\pi}^\pi |h(\lambda)| d\lambda < \infty$, функция $h(\lambda)$ непрерывна при $\lambda = \nu$ для некоторого T_0

$$(18) \quad \int_{-\pi}^\pi l_T(\lambda) d\lambda = 1, \quad T = T_0, \dots,$$

существуют число K и целое число T_1 , такие, что

$$(19) \quad \int_{-\pi}^\pi |l_T(\lambda)| d\lambda \leq K, \quad T = T_1, \dots,$$

и для некоторого T_2

$$(20) \quad |l_T(\lambda)| \leq m_T n(a), \quad |\lambda - \nu| \geq a, \quad T = T_2, \dots,$$

для каждого $a > 0$, а $n(a)$ — монотонно невозрастающая функция от a для $0 < a < \pi$, $m_T \rightarrow 0$ при $T \rightarrow \infty$.

Тогда

$$(21) \quad \lim_{T \rightarrow \infty} \int_{-\pi}^\pi h(\lambda) l_T(\lambda) d\lambda = h(\nu).$$

ДОКАЗАТЕЛЬСТВО. Предположим, что $-\pi < \nu < \pi$. Для заданного ε существует δ , такое, что $|h(\lambda) - h(\nu)| < \varepsilon$ при $|\lambda - \nu| < \delta$. Из (18) и (20) получаем по аналогии с доказательством леммы 8.3.2

$$(22) \quad \lim_{T \rightarrow \infty} \int_{\nu-\delta}^{\nu+\delta} l_T(\lambda) d\lambda = 1.$$

Тогда

$$(23) \quad \left| \int_{v-\delta}^{v+\delta} h(\lambda) l_T(\lambda) d\lambda - h(v) \int_{v-\delta}^{v+\delta} l_T(\lambda) d\lambda \right| = \left| \int_{v-\delta}^{v+\delta} [h(\lambda) - h(v)] \times \right. \\ \left. \times l_T(\lambda) d\lambda \right| \leq \int_{v-\delta}^{v+\delta} |h(\lambda) - h(v)| |l_T(\lambda)| d\lambda \leq K\epsilon, \quad T = T_1, \dots$$

Для $T \geq T_2$ имеем

$$(24) \quad \left| \int_{-\pi}^{v-\delta} h(\lambda) l_T(\lambda) d\lambda + \int_{v+\delta}^{\pi} h(\lambda) l_T(\lambda) d\lambda \right| \leq \\ \leq m_T \left[\int_{-\pi}^{v-\delta} |h(\lambda)| d\lambda + \int_{v+\delta}^{\pi} |h(\lambda)| d\lambda \right] n(\delta),$$

а последнее стремится к 0 при $T \rightarrow \infty$. Доказательства для случаев $v = -\pi$ и $v = \pi$ проводятся аналогично. Заметим, что из (18) следует (19) для $K = 1$, если $l_T(\lambda) \geq 0$. ■

Доказательство формулы (4) представляет собой частный случай леммы 8.3.3, когда $h(\lambda) = 2\pi f(\lambda)$ и $l_T(\lambda) = k_T(\lambda)$.

Оба условия теоремы 8.3.1 $\left[\sum_{r=-\infty}^{\infty} \sigma(r) < \infty \right]$ и непрерывность $f(\lambda)$ при $\lambda = 0$ вытекают из предположения

$$(25) \quad \sum_{r=-\infty}^{\infty} |\sigma(r)| < \infty.$$

8.3.2. Выборочные ковариации

Если среднее значение μ известно, то C_h , определенное формулой (7) § 8.2, является несмещенной оценкой $\sigma(h)$, $h = 0, 1, \dots, T-1$. Оценка $c_h = (T-|h|)C_h/T$ имеет смещение $-|h|\sigma(h)/T$, и для любого h предел смещения равен 0 при $T \rightarrow \infty$.

В разд. 8.2.2 было показано, что когда μ неизвестно, математические ожидания оценок C_h^* и $\tilde{C}_h$ являются линейными комбинациями величин $\sigma(0), \dots, \sigma(T-1)$. Отсюда, вообще говоря, следует, что они смещены. Покажем, что это смещение имеет порядок $1/T$. Действительно, смещение приблизительно в $1/T$ раз отличается от величины

$$(26) \quad -\frac{1}{T} \sum_{t,s=1}^T \sigma(t-s) = -\sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) \sigma(r),$$

предел которой равен $-\sum_{r=-\infty}^{\infty} \sigma(r)$, если последний ряд сходится.

Разность между $\sigma(h)$ и $\mathcal{E}C_h^*$, как видно из формулы (50) § 8.2, содержит три члена. Последний из них совпадает с выражением (26), умноженным на $-1/T$. Второй и третий члены дают в сумме

$$-\sum_{t=1}^{T-h} \sum_{s=1}^T [\sigma(t-s) + \sigma(t+h-s)]/[T(T-h)].$$

Поскольку

$$\begin{aligned} (27) \quad & \left| \sum_{t=1}^{T-h} \sum_{s=1}^T \sigma(t-s) - \sum_{t,s=1}^T \sigma(t-s) \right| = \\ & = \left| \sum_{t=T-h+1}^T \sum_{s=1}^T \sigma(t-s) \right| = \\ & = \left| \sum_{r=-(h-1)}^{-1} (h-|r|) \sigma(r) + h \sum_{r=0}^{T-h} \sigma(r) + \sum_{r=T-h+1}^T (T-r) \sigma(r) \right| \rightarrow \\ & \rightarrow \left| h \sum_{r=0}^{\infty} \sigma(r) + \sum_{r=1}^{h-1} (h-r) \sigma(r) \right|, \end{aligned}$$

когда $T \rightarrow \infty$, если $\sum_{r=0}^{\infty} \sigma(r) < \infty$, то смещение, умноженное на T ,

сходится к $-\sum_{r=-\infty}^{\infty} \sigma(r)$. Аналогично можно получить и другие оценки, такие, как $\tilde{C}_h$.

ТЕОРЕМА 8.3.2. Если $\sum_{r=-\infty}^{\infty} \sigma(r) < \infty$, то

$$(28) \quad \lim_{T \rightarrow \infty} T [\mathcal{E}C_h^* - \sigma(h)] = \lim_{T \rightarrow \infty} T [\mathcal{E}\tilde{C}_h - \sigma(h)] = -\sum_{r=-\infty}^{\infty} \sigma(r).$$

Если $f(\lambda)$ непрерывна при $\lambda = 0$, то

$$(29) \quad \lim_{T \rightarrow \infty} T [\mathcal{E}C_h^* - \sigma(h)] = \lim_{T \rightarrow \infty} T [\mathcal{E}\tilde{C}_h - \sigma(h)] = -2\pi f(0).$$

Вернемся теперь к ковариациям оценок. Большинство этих результатов было получено Бартлеттом (1935), (1946) и Парзенем (1957b). Из теоремы 8.2.6 имеем

$$\begin{aligned} (30) \quad & (T-h) \text{Cov}(C_h, C_g) = \\ & = \sum_{r=0}^{g-h} [\sigma(r) \sigma(r+h-g) + \sigma(r-g) \sigma(r+h)] + \end{aligned}$$

$$\begin{aligned}
& + \kappa(h, -r, g-r) + \sum_{r=g-h+1}^{T-h-1} \left(1 - \frac{r-(g-h)}{T-g}\right) \times \\
& \times [\sigma(r) \sigma(r+h-g) + \sigma(r-g) \sigma(r+h) + \\
& + \kappa(h, -r, g-r)] + \sum_{r=-(T-g-1)}^{-1} \left(1 - \frac{|r|}{T-g}\right) \times \\
& \times [\sigma(r) \sigma(r+h-g) + \sigma(r-g) \sigma(r+h) + \\
& + \kappa(h, -r, g-r)],
\end{aligned}$$

$$0 \leq h \leq g < T.$$

Предел выражения (30) равен

$$\begin{aligned}
(31) \quad \lim_{T \rightarrow \infty} (T-h) \text{Cov}(C_h, C_g) &= \\
&= \sum_{r=-\infty}^{\infty} [\sigma(r) \sigma(r+h-g) + \\
&+ \sigma(r-g) \sigma(r+h) + \kappa(h, -r, g-r)],
\end{aligned}$$

если ряд сходится. Так как правая часть в (31) симметрична относительно h и g (ввиду симметрии $\kappa(r, s, t)$, указанной в упр. 2), то ограничение $h \leq g$ можно отбросить.

Следующее выражение представляет собой интегральную форму ковариации, когда $\kappa(r, s, t) = 0$:

$$\begin{aligned}
(32) \quad (T-h) \text{Cov}(C_h, C_g) &= \\
&= 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2}(\lambda - \nu)(T-h) \sin \frac{1}{2}(\lambda - \nu)(T-g)}{(T-g) \sin^2 \frac{1}{2}(\lambda - \nu)} \times \\
&\times \cos \frac{1}{2}(\lambda + \nu) g \cos \frac{1}{2}(\lambda + \nu) h f(\lambda) f(\nu) d\lambda d\nu = \\
&= 2 \int_{-\pi}^{\pi} \int_{-\pi-\nu}^{\pi-\nu} \frac{\sin \frac{1}{2}\mu(T-h) \sin \frac{1}{2}\mu(T-g)}{(T-g) \sin^2 \frac{1}{2}\mu} \cos\left(\nu + \frac{1}{2}\mu\right) \times \\
&\times g \cos\left(\nu + \frac{1}{2}\mu\right) h f(\nu + \mu) f(\nu) d\mu d\nu = \\
&= 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2}\mu(T-h) \sin \frac{1}{2}\mu(T-g)}{(T-g) \sin^2 \frac{1}{2}\mu} \cos\left(\nu + \frac{1}{2}\mu\right) \times
\end{aligned}$$

$$\times g \cos \left(\nu + \frac{1}{2} \mu \right) h f(\nu + \mu) f(\nu) d\mu d\nu,$$

$$0 \leq h < T, \quad 0 \leq g < T,$$

потому что $f(\lambda \pm 2\pi) = f(\lambda)$, $\sin^{1/2}(\lambda \pm 2\pi)k = (-1)^k \sin^{1/2} \lambda k$ и $\cos^{1/2}(\lambda \pm 2\pi)k = (-1)^k \cos^{1/2} \lambda k$. [Во втором выражении в формуле (32) интегрирование от $-\pi - \nu$ до $-\pi$ заменяется интегрированием от $\pi - \nu$ до π при замене μ на $\mu - 2\pi$.] Если предположить, что функция $f(\lambda)$ непрерывна при $-\pi \leq \lambda \leq \pi$, то она равномерно непрерывна. Равномерно непрерывны также и функции $\cos \frac{1}{2} \lambda g$ и $\cos \frac{1}{2} \lambda h$. Тогда

$$(33) \quad 2 \int_{-\pi}^{\pi} \cos \left(\nu + \frac{1}{2} \mu \right) g \cos \left(\nu + \frac{1}{2} \mu \right) h f(\nu + \mu) f(\nu) d\nu = m(\mu)$$

есть равномерно непрерывная функция по μ . Интеграл (32) можно переписать следующим образом:

$$(34) \quad \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2} \mu (T-h) \sin \frac{1}{2} \mu (T-g)}{(T-g) \sin^2 \frac{1}{2} \mu} m(\mu) d\mu =$$

$$= \int_{-\pi}^{\pi} \frac{\sin^2 \frac{1}{2} \mu (T-g)}{(T-g) \sin^2 \frac{1}{2} \mu} \cos \frac{1}{2} \mu (g-h) m(\mu) d\mu +$$

$$+ \int_{-\pi}^{\pi} \frac{\sin \frac{1}{2} \mu (T-g) \sin \frac{1}{2} \mu (g-h)}{(T-g) \sin^2 \frac{1}{2} \mu} \cos \frac{1}{2} \mu (T-g) m(\mu) d\mu.$$

Первый интеграл в правой части (34) сходится к $2\pi m(0)$. Абсолютное значение второго интеграла меньше, чем умноженное на $\max_{-\pi \leq \mu \leq \pi} |m(\mu)|$ выражение

$$(35) \quad \int_{-\delta^*}^{\delta^*} \frac{\sin \frac{1}{2} \mu (T-g) \sin \frac{1}{2} \mu (g-h)}{(T-g) \sin^2 \frac{1}{2} \mu} d\mu + 2 \int_{\delta^*}^{\pi} \frac{1}{(T-g) \sin^2 \frac{1}{2} \mu} d\mu \leq$$

$$\leq 2|g-h|\delta^* + 2 \frac{\pi - \delta^*}{(T-g) \sin^2 \frac{1}{2} \delta^*}$$

для любого δ^* , $0 < \delta^* \leq \pi$, так как

$$(36) \quad \left| \frac{\sin \frac{1}{2} \mu k}{\sin \frac{1}{2} \mu} \right| = \left| \frac{e^{i \frac{1}{2} \mu k} - e^{-i \frac{1}{2} \mu k}}{e^{i \frac{1}{2} \mu} - e^{-i \frac{1}{2} \mu}} \right| = \left| \sum_{j=0}^{k-1} e^{i \mu j} \right| \leq k.$$

Правую часть (35) можно сделать сколь угодно малой, если δ^* выбрать достаточно малым, а T достаточно большим. Таким образом, предел второго слагаемого правой части (34) равен 0.

ТЕОРЕМА 8.3.3. Если $f(\lambda)$ непрерывна при $-\pi \leq \lambda \leq \pi$ и если выполняется условие $\left| \sum_{r=-\infty}^{\infty} \kappa(h, -r, g-r) \right| < \infty$, то

$$(37) \quad \lim_{T \rightarrow \infty} (T-h) \text{Cov}(C_h, C_g) = \\ = 4\pi \int_{-\pi}^{\pi} \cos \nu g \cos \nu h f^2(\nu) d\nu + \sum_{r=-\infty}^{\infty} \kappa(h, -r, g-r).$$

Если $\sum_{r=-\infty}^{\infty} \sigma^2(r) < \infty$ и $\left| \sum_{r=-\infty}^{\infty} \kappa(h, -r, g-r) \right| < \infty$, то

$$(38) \quad \lim_{T \rightarrow \infty} (T-h) \text{Cov}(C_h, C_g) = \\ = \sum_{r=-\infty}^{\infty} [\sigma(r) \sigma(r+h-g) + \sigma(r-g) \sigma(r+h) + \kappa(h, -r, g-r)].$$

Для гауссовского процесса $\kappa(r, s, t) = 0$. Если $\{y_t\}$ — процесс скользящего среднего, т. е.

$$(39) \quad y_t - \mu = \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}, \quad t = \dots, -1, 0, 1, \dots,$$

где $\mathbb{E} v_t = 0$, $\mathbb{E} v_t^2 = \sigma^2$ и $\mathbb{E} v_t v_s = 0$ при $t \neq s$ и $\sum_{s=-\infty}^{\infty} \gamma_s^2 < \infty$, и если

$\mathbb{E} v_t^4 < \infty$, то центральные моменты четвертого порядка процесса $\{y_t\}$ будут равны

$$(40) \quad \nu(h, g, k) = \sum_{n, s, r, p=-\infty}^{\infty} \gamma_n \gamma_s \gamma_r \gamma_p \mathbb{E} v_{t-n} v_{t+h-s} v_{t+g-r} v_{t+k-p}.$$

Если v_t независимы (или их моменты до четвертого порядка соответствуют моментам независимых величин), то каждое математическое ожидание приведенной выше суммы равно 0, если только ни

Одна из следующих трех пар соотношений не имеет места:

$$(41) \quad \begin{aligned} t-n &= t+h-s, & t+g-r &= t+k-p, \\ t-n &= t+g-r, & t+h-s &= t+k-p, \\ t-n &= t+k-p, & t+h-s &= t+g-r, \end{aligned}$$

или

$$(42) \quad \begin{aligned} n &= s-h, & r-g &= p-k, \\ n &= r-g, & s-h &= p-k, \\ n &= p-k, & s-h &= r-g. \end{aligned}$$

Тогда $v(h, g, k)$ равно

$$(43) \quad v(h, g, k) = \sum_{s=-\infty}^{\infty} \gamma_s \gamma_{s-h} \gamma_{s-h+g} \gamma_{s-h+k} \mathcal{E} v_s^4 + \sum_{\substack{s, r=-\infty \\ s-h \neq r-g}}^{\infty} \gamma_s \gamma_{s-h} \gamma_r \times \\ \times \gamma_{r-g+k} \sigma^4 + \sum_{\substack{s, r=-\infty \\ s-h \neq r-g}}^{\infty} \gamma_s \gamma_{s-h+k} \gamma_r \gamma_{r-g} \sigma^4 + \sum_{\substack{s, p=-\infty \\ s-h \neq p-k}}^{\infty} \gamma_s \gamma_{s-h+g} \gamma_p \gamma_{p-k} \sigma^4.$$

Если процесс $\{v_t\}$ стационарен, то $\mathcal{E} v_t^4$ не зависит от t . Имеем

$$(44) \quad \begin{aligned} \kappa(h, g, k) &= v(h, g, k) - [\sigma(h) \sigma(g-k) + \sigma(g) \sigma(h-k) + \\ &+ \sigma(k) \sigma(h-g)] = \\ &= \sum_{s=-\infty}^{\infty} \gamma_s \gamma_{s-h} \gamma_{s-h+g} \gamma_{s-h+k} (\mathcal{E} v_s^4 - 3\sigma^4) = \\ &= \sum_{s=-\infty}^{\infty} \gamma_s \gamma_{s-h} \gamma_{s-h+g} \gamma_{s-h+k} \kappa_4 = \kappa_4 \sum_{t=-\infty}^{\infty} \gamma_t \gamma_{t+h} \gamma_{t+g} \gamma_{t+k}, \end{aligned}$$

где $\kappa_4 = \mathcal{E} v_s^4 - 3\sigma^4$ есть семиинвариант четвертого порядка величины v_s и

$$(45) \quad \sigma(h) = \sigma^2 \sum_{t=-\infty}^{\infty} \gamma_t \gamma_{t+h}.$$

Тогда

$$(46) \quad \begin{aligned} \sum_{r=-\infty}^{\infty} \kappa(h, -r, g-r) &= \sum_{r=-\infty}^{\infty} \kappa(h, r, r-g) = \\ &= \kappa_4 \sum_{r, s=-\infty}^{\infty} \gamma_s \gamma_{s-h} \gamma_{s-h+r} \gamma_{s-h+r-g} = \\ &= \kappa_4 \sum_{s, t=-\infty}^{\infty} \gamma_s \gamma_{s-h} \gamma_t \gamma_{t-g} = \kappa_4 \frac{\sigma(h) \sigma(g)}{\sigma^4}. \end{aligned}$$

Следствие 8.3.1. Если процесс $\{y_t\}$ порожден соотношением (39), где $\mathcal{E} v_t = 0$, $\mathcal{E} v_t^2 = \sigma^2$, $\mathcal{E} v_t v_s = 0$ при $t \neq s$, $\mathcal{E} v_t v_s v_q = 0$ при $t \neq s$, $t \neq r$, $t \neq q$, $\mathcal{E} v_t^4 = 3\sigma^4 + \kappa_4 < \infty$, $\mathcal{E} v_t^2 v_s^2 = \sigma^4$ при $t \neq s$.

$$\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty, \text{ то}$$

$$\begin{aligned}
 (47) \quad & \lim_{T \rightarrow \infty} (T-h) \operatorname{Cov}(C_h, C_g) = \\
 & = \sum_{r=-\infty}^{\infty} [\sigma(r+h)\sigma(r+g) + \sigma(r-g)\sigma(r+h)] + \frac{\kappa_4}{\sigma^4} \sigma(h)\sigma(g) = \\
 & = 4\pi \int_{-\pi}^{\pi} \cos vg \cos vhf^2(v) dv + \frac{\kappa_4}{\sigma^4} \sigma(h)\sigma(g) = \\
 & = 4\pi \int_{-\pi}^{\pi} \cos vg \cos vhf^2(v) dv + \frac{\kappa_4}{\sigma^4} \int_{-\pi}^{\pi} \cos vgf(v) dv \int_{-\pi}^{\pi} \cos vhf(v) dv.
 \end{aligned}$$

Доказательство. Из условия $\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty$ следует, что

$$\sum_{t=-\infty}^{\infty} \gamma_t^2 < \infty, \quad \sum_{h=-\infty}^{\infty} \sigma^2(h) < \infty,$$

$$(48) \quad \sum_{r=-\infty}^{\infty} \sigma(r) = \sigma^2 \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} \gamma_s \gamma_{s+r} = \sigma^2 \left(\sum_{s=-\infty}^{\infty} \gamma_s \right)^2 < \infty,$$

и функция

$$(49) \quad f(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} \right|^2$$

непрерывна. Таким образом, следствие 8.3.1 вытекает из теоремы 8.3.3. Одновременно доказана и теорема 8.3.2. ■

Для любого процесса, такого, что для него $\sum_{r=-\infty}^{\infty} \sigma^2(r) < \infty$ и

$$\left| \sum_{r=-\infty}^{\infty} \kappa(h, -r, h-r) \right| < \infty, \text{ предельной дисперсией величины}$$

$\sqrt{T-h} C_h$ будет

$$\begin{aligned}
 (50) \quad & \lim_{T \rightarrow \infty} (T-h) \operatorname{Var} C_h = \\
 & = \sum_{r=-\infty}^{\infty} [\sigma^2(r) + \sigma(r)\sigma(r+2h)] + \sum_{r=-\infty}^{\infty} \kappa(h, -r, h-r) = \\
 & = 2\pi \int_{-\pi}^{\pi} [1 + e^{i2vh}] f^2(v) dv + \sum_{r=-\infty}^{\infty} \kappa(h, -r, h-r) = \\
 & = 4\pi \int_{-\pi}^{\pi} \cos^2 vhf^2(v) dv + \sum_{r=-\infty}^{\infty} \kappa(h, -r, h-r).
 \end{aligned}$$

Слагаемое $\sum_{r=-\infty}^{\infty} \sigma^2(r)$ появляется в каждой предельной дисперсии независимо от h . Второе слагаемое

$$(51) \quad \sum_{r=-\infty}^{\infty} \sigma(r-h) \sigma(r+h) = \sigma(-h) \sigma(h) + 2 \sum_{r=1}^{\infty} \sigma(r-h) \sigma(r+h)$$

по абсолютной величине меньше или равно

$$(52) \quad \sigma^2(h) + 2 \sum_{r=1}^{\infty} |\sigma(r-h) \sigma(r+h)| \leq \sigma^2(h) + 2\sigma(0) \sum_{s=h+1}^{\infty} |\sigma(s)|,$$

а последнее сходится к 0 при $h \rightarrow \infty$, если $\sum_{r=-\infty}^{\infty} |\sigma(r)| < \infty$.

Если $\kappa(r, s, t) = 0$, то предельная дисперсия стремится к $\sum_{r=-\infty}^{\infty} \sigma^2(r)$, когда $h \rightarrow \infty$.

Теперь мы обратимся к оценкам ковариаций, когда среднее значение неизвестно, таким, как

$$(53) \quad \begin{aligned} \tilde{C}_h &= \frac{1}{T-h} \sum_{t=1}^{T-h} y_t y_{t+h} - \bar{y}_h \bar{y}_{h+} = \\ &= C_h - \frac{1}{(T-h)^2} \sum_{s,t=1}^{T-h} (y_s - \mu)(y_{t+h} - \mu). \end{aligned}$$

Поскольку $\tilde{C}_h$ инвариантно относительно сдвига ($y_t - \mu \rightarrow y_t$), предположим, что $\mu = 0$. Из выражений (62) и (65) § 8.2 получаем

$$(54) \quad \begin{aligned} (T-g) [(T-h) \text{Cov}(C_g, C_h) - (T-h) \text{Cov}(\tilde{C}_g, \tilde{C}_h)] &= \\ &= \frac{1}{T-g} \sum_{t=1}^{T-h} \sum_{t',s'=1}^{T-g} [\sigma(t-t') \sigma(t+h-s'-g) + \\ &\quad + \sigma(t-s'-g) \sigma(t+h-t') + \kappa(h, t'-t, s'-t+g)] + \\ &\quad + \frac{1}{T-h} \sum_{t,s=1}^{T-h} \sum_{t'=1}^{T-g} [\sigma(t-t') \sigma(s+h-t'-g) + \sigma(t-t'-g) \times \\ &\quad \times \sigma(s+h-t') + \kappa(s+h-t, t'-t, t'+g-t)] - \\ &\quad - \frac{1}{(T-h)(T-g)} \sum_{t,s=1}^{T-h} \sum_{t',s'=1}^{T-g} [\sigma(t-t') \sigma(s+h-s'-g) + \\ &\quad + \sigma(t-s'-g) \sigma(s+h-t') + \kappa(s+h-t, t'-t, s'+g-t)] = \\ &= \frac{1}{T-g} \sum_{t=1}^{T-h} \sum_{q,r=1-t}^{T-g-t} [\sigma(r) \sigma(q+g-h) + \sigma(q+g) \sigma(r-h) + \end{aligned}$$

$$\begin{aligned}
& + \kappa(h, r, q + g)] + \frac{1}{T-h} \sum_{t'=1}^{T-g} \sum_{q, r=1-t'}^{T-h-t'} [\sigma(r) \sigma(q + h - g) + \\
& + \sigma(r - g) \sigma(q + h) + \kappa(q - r + h, -r, g - r)] - \\
& - \frac{1}{(T-h)(T-g)} \sum_{t, s=1}^{T-h} \left[\sum_{r=1-t}^{T-g-t} \sigma(r) \sum_{q=1-s}^{T-g-s} \sigma(q + g - h) + \right. \\
& \left. + \sum_{r=1-t}^{T-g-t} \sigma(r + g) \sum_{q=1-s}^{T-g-s} \sigma(q - h) + \sum_{q, r=1-t}^{T-g-t} \kappa(s + h - t, r, q + g) \right].
\end{aligned}$$

Если $\sum_{r=-\infty}^{\infty} \sigma(r) < \infty$, то для любого $\varepsilon > 0$ существует N , такое, что

$$2 \left| \sum_{r=k}^{\infty} \sigma(r) \right| < \varepsilon \text{ для всех } k > N. \text{ Тогда}$$

$$(55) \quad \left| \sum_{r=1-t}^{T-g-t} \sigma(r) - \sum_{r=-\infty}^{\infty} \sigma(r) \right| < \varepsilon,$$

если $1 - t < -N$ и $T - g - t > N$, т. е. если $N + 1 < t < T - g - N$. Таким образом, если T достаточно большое (с фиксированными g и h), то сумма $\sum_{k=1-t}^{T-g-t} \sigma(r)$ будет удовлетворять (55) для значений t в пределах от 1 до $T - h$, доля которых сколь угодно велика. Аналогично сумма $\sum_{q=1-t}^{T-g-t} \sigma(q + g - h)$ будет отличаться не более, чем на ε , от величины $\sum_{s=-\infty}^{\infty} \sigma(s)$ для сколь угодно большой доли значений t , если T достаточно велико. Поэтому первая тройная сумма правой части (54) приближается к $\left[\sum_{s=-\infty}^{\infty} \sigma(s) \right]^2$. Точно также каждая сумма, включающая произведения ковариаций, имеет тот же предел. Соответствующий анализ выражений семиинвариантов четвертого порядка показывает, что если

$$(56) \quad \left| \sum_{r, q=-\infty}^{\infty} \kappa(h, r, q) \right| < \infty,$$

то первая сумма, включающая семиинварианты четвертого порядка, имеет предел. Действительно, если $\{y_t\}$ — линейный процесс, определенный формулой (39) с $\sum_{s=-\infty}^{\infty} |\gamma_s| < \infty$, то сумма, включающая

$\kappa(h, r, q + g)$, имеет предел

$$(57) \quad \lim_{T \rightarrow \infty} \frac{\kappa_4}{T - g} \sum_{t=1}^{T-h} \sum_{q, r=1-t}^{T-g-t} \sum_{s=-\infty}^{\infty} \gamma_s \gamma_{s+h} \gamma_{s+r} \gamma_{s+q+g} = \\ = \kappa_4 \sum_{s=-\infty}^{\infty} \gamma_s \gamma_{s+h} \left(\sum_{r=-\infty}^{\infty} \gamma_r \right)^2.$$

Сумма, включающая $\kappa(q - r + h, -r, g - r)$, имеет предел

$$(58) \quad \lim_{T \rightarrow \infty} \frac{\kappa_4}{T - h} \sum_{t'=1}^{T-g} \sum_{q, r=1-t'}^{T-h-t'} \sum_{s=-\infty}^{\infty} \gamma_s \gamma_{s+g-r+h} \gamma_{s-r} \gamma_{s-r+g} = \\ = \kappa_4 \sum_{s=-\infty}^{\infty} \gamma_s \gamma_{s+g} \left(\sum_{r=-\infty}^{\infty} \gamma_r \right)^2.$$

Сумма, включающая $\kappa(s + h - t, r, q + g)$, имеет предел

$$(59) \quad \lim_{T \rightarrow \infty} \frac{\kappa_4}{(T - h)(T - g)} \sum_{t, s=1}^{T-h} \sum_{q, r=1-t}^{T-g-t} \sum_{p=-\infty}^{\infty} \gamma_p \gamma_{p+s-t+h} \gamma_{p+r} \gamma_{p+q+g} = 0.$$

Следствие 8.3.2. Если моменты процесса $\{y_t\}$ вплоть до четвертого порядка соответствуют стационарности и если $\sum_{r=-\infty}^{\infty} \sigma(r) < \infty$, то

$$(60) \quad \lim_{T \rightarrow \infty} (T - g) [(T - h) \text{Cov}(C_g, C_h) - (T - h) \text{Cov}(\tilde{C}_g, \tilde{C}_h)] = \\ = 2 \left[\sum_{r=-\infty}^{\infty} \sigma(r) \right]^2 + \sum_{q, r=-\infty}^{\infty} [\kappa(h, r, q) + \kappa(g, r, q)],$$

если $\left| \sum_{s, r, q=-\infty}^{\infty} \kappa(s, r, q) \right| < \infty$. Если $\{y_t\}$ есть линейный процесс с $\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty$ и $|\kappa_4| < \infty$, то последние две суммы в (60) выражаются соответственно формулами (57) и (58). Первый член в правой части (60) равен $8\pi^2 f^2(0)$.

Таким образом, разность между $(T - h) \text{Cov}(C_g, C_h)$ и $(T - h) \times \text{Cov}(\tilde{C}_g, \tilde{C}_h)$ имеет порядок $1/(T - g)$, если $\sum_{r=-\infty}^{\infty} \sigma(r) < \infty$, $\left| \sum_{s, r, q=-\infty}^{\infty} \kappa(s, r, q) \right| < \infty$, и если моменты процесса $\{y_t\}$ до четвертого порядка включительно стационарны. Для больших выборок мы можем использовать $\text{Cov}(C_g, C_h)$ как аппроксимацию для $\text{Cov}(\tilde{C}_g, \tilde{C}_h)$. $\text{Cov}(C_g^*, C_h^*)$ можно исследовать подобным же образом.

8.3.3. Выборочная спектральная плотность

Рассмотрим теперь асимптотические средние значения и дисперсии величин $I(\lambda)$ и $I^*(\lambda)$. Заметим, что $\sigma(0) = \int_{-\pi}^{\pi} f(\lambda) d\lambda < \infty$.

ТЕОРЕМА 8.3.4. Если $\sum_{r=-\infty}^{\infty} \sigma(r) \cos \lambda r$ сходится, то

$$(61) \quad \lim_{T \rightarrow \infty} \mathcal{E}I(\lambda) = \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} \sigma(r) \cos \lambda r.$$

Если $f(\nu)$ непрерывна при $\nu = \lambda$, то

$$(62) \quad \lim_{T \rightarrow \infty} \mathcal{E}I(\lambda) = f(\lambda).$$

ДОКАЗАТЕЛЬСТВО. Теорема следует из (66) § 8.2 и леммы 8.3.1 и из (67) § 8.2 и леммы 8.3.3. ■

ТЕОРЕМА 8.3.5. Если $\sum_{r=-\infty}^{\infty} |r\sigma(r)| < \infty$, то

$$(63) \quad \lim_{T \rightarrow \infty} T[\mathcal{E}I(\lambda) - f(\lambda)] = -\frac{1}{\pi} \sum_{r=1}^{\infty} r\sigma(r) \cos \lambda r.$$

ДОКАЗАТЕЛЬСТВО. Из (66) § 8.2 имеем

$$(64) \quad T[\mathcal{E}I(\lambda) - f(\lambda)] = -\frac{1}{\pi} \left[\sum_{r=1}^{T-1} r\sigma(r) \cos \lambda r + T \sum_{r=T}^{\infty} \sigma(r) \cos \lambda r \right].$$

Так как

$$(65) \quad \left| T \sum_{r=T}^{\infty} \sigma(r) \cos \lambda r \right| \leq \sum_{r=T}^{\infty} T |\sigma(r)| \leq \sum_{r=T}^{\infty} |r| \cdot |\sigma(r)|,$$

отсюда следует утверждение теоремы. ■

Для изучения асимптотического поведения $\mathcal{E}I^*(\lambda)$ докажем следующую лемму.

ЛЕММА 8.3.4. Если $f(\nu)$ ограничена на $(\lambda - \delta, \lambda + \delta)$ и $(\lambda' - \delta, \lambda' + \delta)$ для некоторого $\delta > 0$ и $\lambda \neq \lambda'$, то

$$(66) \quad \lim_{T \rightarrow \infty} \int_{-\pi}^{\pi} k_T(\nu - \lambda, \nu - \lambda') f(\nu) d\nu = 0, \quad \lambda \neq \lambda', \quad |\lambda - \lambda'| \neq 2\pi.$$

ДОКАЗАТЕЛЬСТВО. Имеем

$$(67) \quad |k_T(\nu - \lambda, \nu - \lambda')| \leq \frac{1}{2\pi T \sin^2 \frac{1}{2} \delta^*}, \quad \begin{aligned} \delta^* &\leq |\nu - \lambda| \leq 2\pi - \delta^*, \\ \delta^* &\leq |\nu - \lambda'| \leq 2\pi - \delta^*, \end{aligned}$$

для каждого $\delta^* > 0$. Если $\delta^* \leq |\lambda - \lambda'|/2$, то $|\nu - \lambda| < \delta^*$ влечет за собой $|\nu - \lambda'| > |\lambda - \lambda'|/2$ и

$$(68) \quad |k_T(\nu - \lambda, \nu - \lambda')| \leq \left| \frac{\sin \frac{1}{2}(\lambda - \nu)T}{T \sin \frac{1}{2}(\nu - \lambda)} \right| \frac{1}{2\pi \sin(|\lambda - \lambda'|/4)} \leq \\ \leq \frac{1}{2\pi \sin(|\lambda - \lambda'|/4)}.$$

Так как δ^* может быть выбрано произвольно малым, отсюда и следует утверждение леммы. ■

ТЕОРЕМА 8.3.6. Если $f(\nu)$ непрерывна при $\nu = 0$ и $\nu = \lambda$, то

$$(69) \quad \lim_{T \rightarrow \infty} T[\mathcal{E}I^*(\lambda) - \mathcal{E}I(\lambda)] = 0, \quad \lambda \neq 0.$$

ДОКАЗАТЕЛЬСТВО. Из (71) § 8.2 получаем, что

$$(70) \quad T[\mathcal{E}I^*(\lambda) - \mathcal{E}I(\lambda)] = \\ = \int_{-\pi}^{\pi} \left[-2 \frac{\sin \frac{1}{2} \lambda T}{\sin \frac{1}{2} \lambda} k_T(\nu - \lambda, \nu) + \frac{1}{T} \frac{\sin^2 \frac{1}{2} \lambda T}{\sin^2 \frac{1}{2} \lambda} k_T(\nu) \right] f(\nu) d\nu, \quad \lambda \neq 0$$

Из леммы 8.3.4 следует, что первое слагаемое в правой части имеет предел 0. Второе слагаемое есть умноженный на $(1/T) \times \times \left[\sin^2 \frac{1}{2} \lambda T / \sin^2 \frac{1}{2} \lambda \right]$ интеграл, сходящийся к $f(0)$. ■

Заметим, что $I^*(0) = 0$, в то время как $\mathcal{E}I(0) \rightarrow f(0)$.

Теперь, применяя лемму 8.3.4 к выражению

$$(71) \quad \text{Cov}[I(\lambda), I(\lambda')] = \\ = \left[\int_{-\pi}^{\pi} k_T(\nu + \lambda, \nu - \lambda') f(\nu) d\nu \right]^2 + \left[\int_{-\pi}^{\pi} k_T(\nu + \lambda, \nu + \lambda') f(\nu) d\nu \right]^2 + \\ + \frac{1}{4\pi^2 T^2} \sum_{t,s,t',s'=1}^T \cos \lambda(t-s) \cos \lambda'(t'-s') \kappa(s-t, t'-t, s'-t),$$

получаем следующую теорему.

ТЕОРЕМА 8.3.7.

$$(72) \quad \lim_{T \rightarrow \infty} \text{Var } I(0) = 2f^2(0) + \lim_{T \rightarrow \infty} \frac{1}{4\pi^2 T^2} \sum_{t,s,t',s'=1}^T \kappa(s-t, t'-t, s'-t),$$

если $f(v)$ непрерывна при $v = 0$ и предел справа существует;

$$(73) \quad \lim_{T \rightarrow \infty} \text{Var } I(\pi) = \lim_{T \rightarrow \infty} \text{Var } I(-\pi) = \\ = 2f^2(\pi) + \lim_{T \rightarrow \infty} \frac{1}{4\pi^2 T^2} \sum_{t,s,t',s'=1}^T (-1)^{t+s+t'+s'} \kappa(s-t, t'-t, s'-t),$$

если $f(v)$ непрерывна при $v = \pi$ и предел справа существует;

$$(74) \quad \lim_{T \rightarrow \infty} \text{Var } I(\lambda) = \\ = f^2(\lambda) + \lim_{T \rightarrow \infty} \frac{1}{4\pi^2 T^2} \sum_{s,t,t',s'=1}^T \cos \lambda(t-s) \cos \lambda(t'-s') \kappa(s-t, \\ t'-t, s'-t), \quad \lambda \neq 0, \pm \pi,$$

если $f(v)$ непрерывна при $v = \lambda$ и предел справа существует, и

$$(75) \quad \lim_{T \rightarrow \infty} \text{Cov } [I(\lambda), I(\lambda')] = \\ = \lim_{T \rightarrow \infty} \frac{1}{4\pi^2 T^2} \sum_{t,s,t',s'=1}^T \cos \lambda(t-s) \cos \lambda'(t'-s') \kappa(s-t, t'-t, s'-t), \\ \lambda \neq \pm \lambda',$$

если $f(v)$ ограничена на $(\lambda - \delta, \lambda + \delta)$ и $(\lambda' - \delta, \lambda' + \delta)$ для некоторого $\delta > 0$ и предел справа в (75) существует.

ЛЕММА 8.3.5. Если

$$(76) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{r,u,w=-(T-1)}^{T-1} |\kappa(r, u, w)| = 0,$$

то

$$(77) \quad \lim_{T \rightarrow \infty} \frac{1}{4\pi^2 T^2} \sum_{t,s,t',s'=1}^T \cos \lambda(t-s) \cos \lambda'(t'-s') \times \\ \times \kappa(s-t, t'-t, s'-t) = 0.$$

ДОКАЗАТЕЛЬСТВО. Сделав замену $s = t + r$, $t' = t + u$, $s' = t + w$, имеем

$$(78) \quad \frac{1}{T^2} \left| \sum_{t,s,t',s'=1}^T \cos \lambda(t-s) \cos \lambda'(t'-s') \kappa(s-t, t'-t, s'-t) \right| \leq \\ \leq \frac{1}{T^2} \sum_{t,s,t',s'=1}^T |\kappa(s-t, t'-t, s'-t)| \leq \\ \leq \frac{1}{T^2} \sum_{t=1}^T \sum_{r,u,w=-(T-1)}^{T-1} |\kappa(r, u, w)| = \frac{1}{T} \sum_{r,u,w=-(T-1)}^{T-1} |\kappa(r, u, w)|. \blacksquare$$

Следствие 8.3.3. Если (76) выполняется и $f(v)$ непрерывна при $v = 0$ и $v = \pi$, то

$$(79) \quad \lim_{T \rightarrow \infty} \text{Var } I(0) = 2f^2(0), \quad \lim_{T \rightarrow \infty} \text{Var } I(\pm \pi) = 2f^2(\pi).$$

Если $f(v)$ непрерывна при $v = \lambda$, то

$$(80) \quad \lim_{T \rightarrow \infty} \text{Var } I(\lambda) = f^2(\lambda), \quad \lambda \neq 0, \pm \pi.$$

Если $f(v)$ ограничена на $(\lambda - \delta, \lambda + \delta)$ и $(\lambda' - \delta, \lambda' + \delta)$ для некоторого $\delta > 0$, то

$$(81) \quad \lim_{T \rightarrow \infty} \text{Cov } [I(\lambda), I(\lambda')] = 0, \quad \lambda \neq \pm \lambda'.$$

Важно заметить, что дисперсия величины $I(\lambda)$ не стремится к 0 при $T \rightarrow \infty$. В действительности, как мы увидим позднее, $I(\lambda)$ не является состоятельной оценкой для $f(\lambda)$ [так как $2I(\lambda)/f(\lambda)$ имеет предельное χ^2 -распределение]. Более того, $I(\lambda)$ и $I(\lambda')$ для $\lambda \neq \pm \lambda'$ асимптотически некоррелированы. Это предполагает усреднение $I(\lambda)$ по некоторому интервалу для улучшения оценки величины $f(\lambda)$.

Семиинварианты четвертого порядка в выражениях ковариаций для $I(\lambda)$ и $I(\lambda')$ равны 0, если процесс гауссовский. Если $\{y_t - \mu\}$ — линейный процесс, то соответствующий член, умноженный на T , равен

$$\begin{aligned} (82) \quad & \frac{1}{4\pi^2 T} \sum_{t,s,t',s'=1}^T \cos \lambda(t-s) \cos \lambda'(t'-s') \kappa(s-t, t'-t, s'-t) = \\ & = \frac{\kappa_4}{4\pi^2 T} \sum_{t,s,t',s'=1}^T \cos \lambda(t-s) \cos \lambda'(t'-s') \sum_{p=-\infty}^{\infty} \gamma_{p+t} \gamma_{p+s} \gamma_{p+t'} \gamma_{p+s'} = \\ & = \frac{\kappa_4}{4\pi^2 T} \sum_{r,r'=-T+1}^{T-1} \cos \lambda r \cos \lambda' r' \sum_{s \in S_r} \sum_{s' \in S_{r'}} \sum_{p=-\infty}^{\infty} \gamma_{p+s+r} \gamma_{p+s} \gamma_{p+s'+r'} \gamma_{p+s'} \end{aligned}$$

где $S_r = \{1, 2, \dots, T-r\}$ для $r \geq 0$ и $S_r = \{1-r, 2-r, \dots, T\}$ для $r \leq 0$. Абсолютное значение выражения (82) не больше, чем

$$\begin{aligned} (83) \quad & \frac{|\kappa_4|}{4\pi^2 T} \sum_{t,s,t',s'=1}^T \sum_{p=-\infty}^{\infty} |\gamma_{p+t}| |\gamma_{p+s}| |\gamma_{p+t'}| |\gamma_{p+s'}| \leq \\ & \leq \frac{|\kappa_4|}{4\pi^2 T} \sum_{t=1}^T \sum_{s,t',s',p=-\infty}^{\infty} |\gamma_{p+t}| |\gamma_{p+s}| |\gamma_{p+t'}| |\gamma_{p+s'}| = \\ & = \frac{\kappa_4}{4\pi^2} \left(\sum_{p=-\infty}^{\infty} |\gamma_p| \right)^4, \end{aligned}$$

а последнее конечно, если $\sum_{p=-\infty}^{\infty} |\gamma_p| < \infty$ и $|\kappa_4| < \infty$. Таким образом абсолютное значение выражения (82) ограничено абсолютно сходящимся рядом.

Следствие 8.3.4. Если процесс $\{y_t\}$ порожден $y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, где $\mathbb{E} v_t = 0$, $\mathbb{E} v_t^2 = \sigma^2$, $\mathbb{E} v_t v_s = 0$ при $t \neq s$, $\mathbb{E} v_t v_s v_r v_q = 0$ при $t \neq s$, $t \neq r$, $t \neq q$, $\mathbb{E} v_t^4 = 3\sigma^4 + \kappa_4 < \infty$, $\mathbb{E} v_t^2 v_s^2 = \sigma^4$ при $t \neq s$ и $\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty$, тогда: если $f(v)$ непрерывна при $v = 0$ и $v = \pi$, то

$$(84) \quad \lim_{T \rightarrow \infty} \text{Var } I(0) = 2f^2(0), \quad \lim_{T \rightarrow \infty} \text{Var } I(\pm \pi) = 2f^2(\pi),$$

если $f(v)$ непрерывна при $v = \lambda$, то

$$(85) \quad \lim_{T \rightarrow \infty} \text{Var } I(\lambda) = f^2(\lambda), \quad \lambda \neq 0, \pm \pi,$$

и если $f(v)$ ограничена на $(\lambda - \delta, \lambda + \delta)$ и $(\lambda' - \delta, \lambda' + \delta)$ для некоторого $\delta > 0$, то

$$(86) \quad \lim_{T \rightarrow \infty} \text{Cov } [I(\lambda), I(\lambda')] = 0, \quad \lambda \neq \pm \lambda'.$$

Вычисление предела выражения (82) представляет особый интерес. Тройная сумма для $r, r' \geq 0$ в правой части (82), деленная на T , равна

$$(87) \quad \begin{aligned} \frac{1}{T} \sum_{s=1}^{T-r} \sum_{s'=1}^{T-r'} \sum_{p=-\infty}^{\infty} \gamma_{p+s+r} \gamma_{p+s} \gamma_{p+s'+r'} \gamma_{p+s'} &= \\ &= \frac{1}{T} \sum_{s=1}^{T-r} \sum_{s'=1}^{T-r'} \sum_{q=-\infty}^{\infty} \gamma_{q+r'} \gamma_q \gamma_{q+s-s'+r} \gamma_{q+s-s'} = \\ &= \sum_{q=-\infty}^{\infty} \gamma_{q+r'} \gamma_q \frac{1}{T} \sum_{s=1}^{T-r} \sum_{s'=1}^{T-r'} \gamma_{q+s-s'+r} \gamma_{q+s-s'}. \end{aligned}$$

Для данного q

$$(88) \quad \begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{s=1}^{T-r} \sum_{s'=1}^{T-r'} \gamma_{q+s-s'+r} \gamma_{q+s-s'} &= \\ &= \lim_{T \rightarrow \infty} \sum_{n=-(T-r'-1)}^{T-r-1} \left(1 - \frac{|n|}{T}\right) \gamma_{q+n+r} \gamma_{q+n} = \sum_{n=-\infty}^{\infty} \gamma_{q+n+r} \gamma_{q+n} = \\ &= \sum_{r=-\infty}^{\infty} \gamma_{t+r} \gamma_t = \frac{\sigma(r)}{\sigma^2}. \end{aligned}$$

Таким образом, предел (87) есть

$$(89) \quad \frac{\sigma(r)}{\sigma^2} \sum_{q=-\infty}^{\infty} \gamma_{q+r} \gamma_q = \frac{\sigma(r) \sigma(r')}{\sigma^4}.$$

Из соображений удобства вывод был проведен для $r, r' \geq 0$. Результат же имеет место для каждой пары целых чисел r, r' . Окончательно пределом (82) является выражение

$$(90) \quad \frac{\kappa_4}{4\pi^2\sigma^4} \lim_{T \rightarrow \infty} \sum_{r, r' = -(T-1)}^{T-1} \cos \lambda r \cos \lambda' r' \sigma(r) \sigma(r') = \\ = \frac{\kappa_4}{4\pi^2\sigma^4} \sum_{r, r' = -\infty}^{\infty} \cos \lambda r \cos \lambda' r' \sigma(r) \sigma(r') = \frac{\kappa_4}{\sigma^4} f(\lambda) f(\lambda').$$

Эти преобразования возможны ввиду сходимости ряда $\sum_{q=-\infty}^{\infty} |\gamma_q|$.

Рассмотрим теперь предельные ковариации функции

$$(91) \quad I^*(\lambda) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) C_r^* \cos \lambda r.$$

В этом случае

$$(92) \quad T \{ \text{Cov} [I(\lambda); I(\lambda')] - \text{Cov} [I^*(\lambda), I^*(\lambda')] \} = \\ = \frac{1}{4\pi^2} \sum_{r, r' = -(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) \times \\ \times \left(1 - \frac{|r'|}{T}\right) \cos \lambda r \cos \lambda' r' T [\text{Cov} (C_r, C_{r'}) - \text{Cov} (C_r^*, C_{r'}^*)].$$

Если $\kappa(g, h, r) = 0$ и если мы заменим C_h^* на $\tilde{C}_h$, то (92) аппроксимируется выражением

$$(93) \quad \frac{1}{4\pi^2 T} \sum_{r, r' = -(T-1)}^{T-1} \cos \lambda r \cos \lambda' r' 2 \left[\sum_{q=-\infty}^{\infty} \sigma(q) \right]^2 = \\ = \frac{1}{2\pi^2} \left[\sum_{q=-\infty}^{\infty} \sigma(q) \right]^2 \frac{1}{T} \sum_{r, r' = -(T-1)}^{T-1} e^{i\lambda r + i\lambda' r'} = \\ = \frac{1}{2\pi^2} \left[\sum_{q=-\infty}^{\infty} \sigma(q) \right]^2 \frac{1}{T} \frac{\sin \frac{1}{2} \lambda (2T-1)}{\sin \frac{1}{2} \lambda} \frac{\sin \frac{1}{2} \lambda' (2T-1)}{\sin \frac{1}{2} \lambda'} \rightarrow 0,$$

$\lambda, \lambda' \neq 0$.

Возможность аппроксимации следует из эвристических соображений.

Найдем предельные дисперсии и ковариации величин $\sqrt{T}A(\lambda)$ и $\sqrt{T}B(\lambda)$ из разд. 8.2.3.

ТЕОРЕМА 8.3.8. Если $f(v)$ непрерывна при $v = 0$, то

$$(94) \quad \lim_{T \rightarrow \infty} T \mathfrak{E} A^2(0) = 8\pi f(0),$$

$$(95) \quad \lim_{T \rightarrow \infty} T \mathfrak{E} B^2(0) = 0;$$

если $f(v)$ непрерывна при $v = \pi$, то

$$(96) \quad \lim_{T \rightarrow \infty} T \mathfrak{E} A^2(\pm \pi) = 8\pi f(\pi),$$

$$(97) \quad \lim_{T \rightarrow \infty} T \mathfrak{E} B^2(\pm \pi) = 0;$$

если $f(v)$ непрерывна при $v = \lambda$, то

$$(98) \quad \lim_{T \rightarrow \infty} T \mathfrak{E} A^2(\lambda) = 4\pi f(\lambda), \quad \lambda \neq 0, \pm \pi,$$

$$(99) \quad \lim_{T \rightarrow \infty} T \mathfrak{E} B^2(\lambda) = 4\pi f(\lambda), \quad \lambda \neq 0, \pm \pi;$$

если $f(v)$ непрерывна при $v = \lambda$ и $v = \lambda'$, то

$$(100) \quad \lim_{T \rightarrow \infty} T \mathfrak{E} A(\lambda) A(\lambda') = 0, \quad \lambda \neq \pm \lambda',$$

$$(101) \quad \lim_{T \rightarrow \infty} T \mathfrak{E} B(\lambda) B(\lambda') = 0, \quad \lambda \neq \pm \lambda',$$

$$(102) \quad \lim_{T \rightarrow \infty} T \mathfrak{E} A(\lambda) B(\lambda') = 0.$$

Если $\sqrt{T}A(\lambda)$ и $\sqrt{T}B(\lambda)$ имеют нормальное предельное распределение, то $\sqrt{T}A(\lambda)/\sqrt{4\pi f(\lambda)}$ и $\sqrt{T}B(\lambda)/\sqrt{4\pi f(\lambda)}$, $\lambda \neq 0, \pm \pi$ имеют нормальное предельное распределение со средним значением 0, дисперсией 1 и корреляцией 0, а выражение

$$(103) \quad T \frac{A^2(\lambda) + B^2(\lambda)}{4\pi f(\lambda)} = T \frac{R^2(\lambda)}{4\pi f(\lambda)} = \frac{2I(\lambda)}{f(\lambda)}$$

имеет предельное χ^2 -распределение с 2 степенями свободы.

8.4. АСИМПТОТИЧЕСКИЕ РАСПРЕДЕЛЕНИЯ ВЫБОРОЧНЫХ СРЕДНЕГО, КОВАРИАЦИЙ И СПЕКТРАЛЬНОЙ ПЛОТНОСТИ

8.4.1. Выборочное среднее

В § 7.7 была доказана центральная предельная теорема (теорема 7.7.8).

ТЕОРЕМА 8.4.1. Если $y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, где $\{v_t\}$ состоит из независимых и одинаково распределенных случайных величин с $\mathfrak{E} v_t =$

$= 0$ и $\mathbb{E}v_i^2 = \sigma^2$ и $\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty$, то $\sqrt{T}(\bar{y} - \mu)$ имеет нормальное предельное распределение со средним значением 0 и дисперсией

$$(1) \quad \sigma^2 \left(\sum_{s=-\infty}^{\infty} \gamma_s \right)^2 = \sum_{r=-\infty}^{\infty} \sigma(r).$$

Существуют и другие условия, при которых выборочное среднее асимптотически нормально. Одно из них будет приведено в разд. 8.4.3. Значение этого результата определяется тем, что он допускает использование теории нормальных распределений для выводов относительно μ при больших T .

8.4.2. Выборочные ковариации

Для больших выборок статистические выводы для ковариаций могут быть основаны на асимптотической нормальности выборочных ковариаций.

ТЕОРЕМА 8.4.2. Если $y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, где $\{v_t\}$ состоит из независимых и одинаково распределенных случайных величин с $\mathbb{E}v_t = 0$, $\mathbb{E}v_t^2 = \sigma^2$ и $\mathbb{E}v_t^4 = 3\sigma^4 + \kappa_4 < \infty$ и $\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty$, то $\sqrt{T}[C_0 - \sigma(0)], \dots, \sqrt{T}[C_n - \sigma(n)]$ имеет нормальное предельное распределение со средним значением 0 и ковариациями

$$(2) \quad \lim_{T \rightarrow \infty} T \text{Cov}(C_h, C_g) = \\ = \sum_{r=-\infty}^{\infty} [\sigma(r+h)\sigma(r+g) + \sigma(r-g)\sigma(r+h)] + \frac{\kappa_4}{\sigma^4} \sigma(h)\sigma(g) = \\ = 4\pi \int_{-\pi}^{\pi} \cos vg \cos vh f^2(v) dv + \frac{\kappa_4}{\sigma^4} \sigma(h)\sigma(g).$$

ДОКАЗАТЕЛЬСТВО. Положим $\mu = 0$. Пусть

$$(3) \quad y_{t,k} = \sum_{s=-k}^k \gamma_s v_{t-s},$$

$$(4) \quad C_{h,k} = \frac{1}{T-h} \sum_{t=1}^{T-h} y_{t,k} y_{t+h,k} = \\ = \frac{1}{T-h} \sum_{t=1}^{T-h} \sum_{s,s'=-k}^k \gamma_s \gamma_{s'} v_{t-s} v_{t+h-s'}, \quad h = 0, 1, \dots, T-1,$$

$$\begin{aligned}
 (5) \quad \sigma(h, k) &= \mathbb{E} y_{t,k} y_{t+h,k} = \sum_{s, s'=-k}^k \gamma_s \gamma_{s'} \mathbb{E} v_{t-s} v_{t+h-s'} = \\
 &= \sigma^2 \sum_{s=-k}^{k-h} \gamma_s \gamma_{s+h} \quad \text{при } h = 0, 1, \dots, 2k, \\
 \sigma(h, k) &= 0 \quad \text{при } h = 2k+1, \dots
 \end{aligned}$$

Тогда $\mathbb{E} C_{h,k} = \sigma(h, k)$ и

$$\begin{aligned}
 (6) \quad \lim_{T \rightarrow \infty} T \operatorname{Cov}(C_{h,k}, C_{g,k}) &= \\
 &= \sum_{r=-\infty}^{\infty} [\sigma(r+h, k) \sigma(r+g, k) + \sigma(r-g, k) \sigma(r+h, k)] + \\
 &\quad + \frac{\kappa_4}{\sigma^4} \sigma(h, k) \sigma(g, k).
 \end{aligned}$$

(Заметим, что сумма в правой части формулы имеет только конечное число ненулевых членов.) Так как $y_{t,k} y_{t+h,k}$ не зависит от $y_{s,k} y_{s+h,k}$, если t и s различаются более чем на $2k+h$, то $\{y_{t,k}, y_{t+h,k}\}$ является конечно-зависимым стационарным процессом со средним значением $\sigma(h, k)$ и дисперсией

$$\begin{aligned}
 (7) \quad \operatorname{Var} y_{t,k} y_{t+h,k} &= \\
 &= \begin{cases} \sigma^2(0, k) + \sigma^2(h, k) + \kappa_4 \sum_{s=-k}^{k-h} \gamma_s^2 \gamma_{s+h}^2 & \text{при } h = 0, 1, \dots, 2k, \\ \sigma^2(0, k) & \text{при } h = 2k+1, \dots \end{cases}
 \end{aligned}$$

По теореме 7.7.6 $\sqrt{T} [C_{0,k} - \sigma(0, k)], \dots, \sqrt{T} [C_{n,k} - \sigma(n, k)]$ имеет нормальное предельное распределение со средними значениями 0 и ковариациями, которые даются формулой (6).

Теперь рассмотрим

$$\begin{aligned}
 (8) \quad \sqrt{T} (C_h - C_{h,k}) &= \frac{\sqrt{T}}{T-h} \sum_{t=1}^{T-h} (y_t y_{t+h} - y_{t,k} y_{t+h,k}) = \\
 &= \frac{\sqrt{T}}{T-h} \sum_{t=1}^{T-h} (u_{t,k} y_{t+h,k} + y_{t,k} u_{t+h,k} + u_{t,k} u_{t+h,k}),
 \end{aligned}$$

где

$$(9) \quad u_{t,k} = y_t - y_{t,k} = \sum_{|s| > k} \gamma_s v_{t-s}.$$

Пусть

$$(10) \quad S_1 = \frac{\sqrt{T}}{T-h} \sum_{t=1}^{T-h} u_{t,k} y_{t+h,k},$$

$$(11) \quad S_2 = \frac{\sqrt{T}}{T-h} \sum_{t=1}^{T-h} y_{t,k} u_{t+h,k},$$

$$(12) \quad S_3 = \frac{\sqrt{T}}{T-h} \sum_{t=1}^{T-h} u_{t,k} u_{t+h,k}.$$

Положим $\gamma'_s = \gamma_s$ для $|s| \leq k$, $\gamma'_s = 0$ для $|s| > k$, и пусть $\gamma_s^* = 0$ для $|s| \leq k$ и $\gamma_s^* = \gamma_s$ для $|s| > k$. Тогда (для $0 \leq h \leq k$)

$$(13) \quad \mathbb{E} S_1 = \sqrt{T} \sigma^2 \sum_{s=-\infty}^{\infty} \gamma_s^* \gamma'_{s+h},$$

$$(14) \quad \mathbb{E} S_2 = \sqrt{T} \sigma^2 \sum_{s=-\infty}^{\infty} \gamma_s^* \gamma'_{s+h},$$

$$(15) \quad \mathbb{E} S_3 = \sqrt{T} \sigma^2 \sum_{s=-\infty}^{\infty} \gamma_s^* \gamma^*_{s+h}.$$

Заметим, что

$$(16) \quad \mathbb{E} S_1 + \mathbb{E} S_2 + \mathbb{E} S_3 = \sqrt{T} \mathbb{E} C_h - \sqrt{T} \mathbb{E} C_{h,k}.$$

Дисперсия величины S_3 есть

$$\begin{aligned} (17) \quad \text{Var } S_3 &= \frac{T}{(T-h)^2} \sum_{t,t'=1}^{T-h} \text{Cov}(u_{t,k} u_{t+h,k}, u_{t',k} u_{t'+h,k}) = \\ &= \frac{T}{(T-h)^2} \sum_{t,t'=1}^{T-h} \sum_{s,r,s',r'=-\infty}^{\infty} \gamma_s^* \gamma_r^* \gamma_{s'}^* \gamma_{r'}^* \times \\ &\times \text{Cov}(u_{t-s} u_{t+h-r}, u_{t'-s'} u_{t'+h-r'}) = \\ &= \frac{T}{(T-h)^2} \sum_{t,t'=1}^{T-h} \left[\sigma^4 \sum_{s,r=-\infty}^{\infty} \gamma_s^* \gamma_{s-t+t'}^* \gamma_r^* \gamma_{r-t+t'}^* + \right. \\ &+ \sigma^4 \sum_{s,r=-\infty}^{\infty} \gamma_s^* \gamma_{s-t+t'+h}^* \gamma_r^* \gamma_{r-t+t'-h}^* + \\ &\left. + \kappa_4 \sum_{s=-\infty}^{\infty} \gamma_s^* \gamma_{s-t+t'}^* \gamma_{s+h}^* \gamma_{s-t+t'+h}^* \right]. \end{aligned}$$

Таким образом, для $T \geq 2h$

$$\begin{aligned} (81) \text{Var } S_3 &\leq \frac{T}{T-h} \left[\sigma^4 \sum_{p,s,r=-\infty}^{\infty} |\gamma_s^*| |\gamma_{s+p}^*| |\gamma_r^*| |\gamma_{r+p}^*| + \right. \\ &+ \sigma^4 \sum_{p,s,r=-\infty}^{\infty} |\gamma_s^*| |\gamma_{s+p+h}^*| |\gamma_r^*| |\gamma_{r+p-h}^*| + \\ &\left. + \kappa_4 \sum_{p,s=-\infty}^{\infty} |\gamma_s^*| |\gamma_{s+p}^*| |\gamma_{s+h}^*| |\gamma_{s+p+h}^*| \right] \leq \end{aligned}$$

$$\begin{aligned}
&\leq 2 \left[2\sigma^4 \sum_{q,p,s,r=-\infty}^{\infty} |\gamma_s^*| |\gamma_p^*| |\gamma_r^*| |\gamma_q^*| + \right. \\
&\quad \left. + \kappa_4 \sum_{q,s=-\infty}^{\infty} |\gamma_s^*| |\gamma_{s+h}^*| |\gamma_q^*| |\gamma_{q+h}^*| \right] = \\
&= 2 \left[2\sigma^4 \left(\sum_{|s|>k} |\gamma_s| \right)^4 + \kappa_4 \left(\sum_{\substack{|s|>k \\ |s+h|>k}} |\gamma_s \gamma_{s+h}| \right)^2 \right],
\end{aligned}$$

которая не зависит от T и стремится к 0, когда $k \rightarrow \infty$. Аналогично для $T \geq 2h$

$$\begin{aligned}
(19) \quad \text{Var } S_1 &\leq 2 \left[2\sigma^4 \sum_{q,p,s,r=-\infty}^{\infty} |\gamma_q^*| |\gamma_p^*| |\gamma_s^*| |\gamma_r^*| + \right. \\
&\quad \left. + \kappa_4 \sum_{q,s=-\infty}^{\infty} |\gamma_s^*| |\gamma_{s+h}^*| |\gamma_q^*| |\gamma_{q+h}^*| \right] \leq \\
&\leq 2 \left[2\sigma^4 \left(\sum_{|p| \leq k} |\gamma_p| \right)^2 \left(\sum_{|s|>k} |\gamma_s| \right)^2 + \kappa_4 \left(\sum_{|s|>k} |\gamma_s \gamma_{s+h}| \right)^2 \right]
\end{aligned}$$

и $\text{Var } S_2$ имеет ту же самую границу. Тогда

$$(20) \quad \text{Var}(S_1 + S_2 + S_3) \leq 3(\text{Var } S_1 + \text{Var } S_2 + \text{Var } S_3).$$

По следствию 7.7.1 предельное распределение величины $\sqrt{T} [C_h - \sigma(h)]$ есть предел по k предельного распределения величины $\sqrt{T} [C_{h,k} - \sigma(h, k)]$. Это доказывает теорему. ■

Следствие 8.4.1. В условиях теоремы 8.4.2 предельные распределения величин $\sqrt{T} [C_0^* - \sigma(0)], \dots, \sqrt{T} [C_n^* - \sigma(n)], \sqrt{T} [\tilde{C}_0 - \sigma(0)], \dots, \sqrt{T} [\tilde{C}_n - \sigma(n)], \sqrt{T} [c_0 - \sigma(0)], \dots, \sqrt{T} [c_n - \sigma(n)], \sqrt{T} [c_0^* - \sigma(0)], \dots, \sqrt{T} [c_n^* - \sigma(n)]$ и $\sqrt{T} [\tilde{c}_0 - \sigma(0)], \dots, \sqrt{T} [\tilde{c}_n - \sigma(n)]$ нормальны со средними значениями 0 и ковариациями, которые даются формулой (2).

8.4.3. Тригонометрические коэффициенты

Тригонометрические коэффициенты

$$(21) \quad A(\lambda) = \frac{2}{T} \sum_{t=1}^T (y_t - \mu) \cos \lambda t, \quad -\pi \leq \lambda \leq \pi,$$

$$(22) \quad B(\lambda) = \frac{2}{T} \sum_{t=1}^T (y_t - \mu) \sin \lambda t, \quad -\pi \leq \lambda \leq \pi,$$

являются линейными функциями от $y_1 - \mu, \dots, y_T - \mu$ со средними значениями 0 и дисперсиями и ковариациями, указанными в разд. 8.2.3. Если $\{y_t\}$ — гауссовский процесс, то любое множество этих тригонометрических коэффициентов распределено нормально. (Если число этих коэффициентов больше T , то распределение вырождено.) В этом разделе мы покажем, что при точно установленных общих условиях $\sqrt{T}A(\lambda)$ и $\sqrt{T}B(\lambda)$ асимптотически нормально распределены со средними значениями 0, нулевыми ковариациями и дисперсиями, данными в разд. 8.3.3.

ТЕОРЕМА 8.4.3. Если $y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, где $\{v_t\}$ состоит из независимо распределенных случайных величин, причем v_t имеет среднее значение 0, дисперсию σ^2 и функцию распределения $F_t(v)$,

$$\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty, \text{ и если}$$

$$(23) \quad \sup_{t=1,2,\dots} \int_{|v|>c} v^2 dF_t(v) \rightarrow 0,$$

при $c \rightarrow \infty$, то $\sqrt{T}A(\lambda_1), \sqrt{T}B(\lambda_1), \dots, \sqrt{T}A(\lambda_n), \sqrt{T}B(\lambda_n), 0 < \lambda_j < \pi, \lambda_i \neq \lambda_j, i \neq j, i, j = 1, \dots, n$, имеют нормальное предельное распределение со средними значениями 0, нулевыми ковариациями и дисперсиями, соответственно равными $4\pi f(\lambda_1), 4\pi f(\lambda_1), \dots, 4\pi f(\lambda_n), 4\pi f(\lambda_n)$.

ДОКАЗАТЕЛЬСТВО. Предположим, что $\mu = 0$. Пусть

$$(24) \quad A_k(\lambda) = \frac{2}{T} \sum_{t=1}^T y_{t,k} \cos \lambda t.$$

Предельное распределение величины $\sqrt{T}A(\lambda)$ при $T \rightarrow \infty$ это то же самое, что и предел при $k \rightarrow \infty$ предельного распределения величины $\sqrt{T}A_k(\lambda)$ при $T \rightarrow \infty$, для

$$(25) \quad \begin{aligned} \mathbb{E}T[A(\lambda) - A_k(\lambda)]^2 &= \frac{4}{T} \mathbb{E} \left(\sum_{t=1}^T u_{t,k} \cos \lambda t \right)^2 = \\ &= \frac{4}{T} \mathbb{E} \sum_{t,t'=1}^T \cos \lambda t \cos \lambda t' \sum_{|s|,|s'|>k} \gamma_s \gamma_{s'} v_{t-s} v_{t'-s'} = \\ &= \frac{4}{T} \sum_{|s|,|s'|>k} \gamma_s \gamma_{s'} \sigma^2 \sum_t \cos \lambda t \cos \lambda (t-s+s'), \end{aligned}$$

где суммирование по t проводится для значений t и $t-s+s'$ между 1 и T . Перестановка суммирования по s и s' и математического ожидания оправдывается тем фактом, что

$$(26) \quad \mathbb{E} \left| \sum_{s,s'} \gamma_s \gamma_{s'} v_s v_{s'} \right| \leq \mathbb{E} \sum_{s,s'} |\gamma_s| |\gamma_{s'}| |v_s v_{s'}|,$$

все члены суммы в правой части неотрицательны и $\sum_{s=-\infty}^{\infty} |\gamma_s| < \infty$.

Тогда

$$(27) \quad \mathbb{E}T[A(\lambda) - A_k(\lambda)]^2 \leq 4\sigma^2 \max_{t=1, \dots, T} \cos^2 \lambda t \sum_{|s|, |s'| > k} |\gamma_s| |\gamma_{s'}| \leq \\ \leq 4\sigma^2 \left(\sum_{|s| > k} |\gamma_s| \right)^2,$$

что стремится к 0 при $k \rightarrow \infty$. Мы применим следствие 7.7.1.

Запишем

$$(28) \quad A_k(\lambda) = \frac{2}{T} \sum_{t=1}^T \sum_{s=-k}^k \gamma_s v_{t-s} \cos \lambda t = \\ = \frac{2}{T} \left\{ \sum_{r=k+1}^{T-k} \left[\sum_{s=-k}^k \gamma_s \cos \lambda(r+s) \right] v_r + \right. \\ \left. + \sum_{r=1-k}^k \left[\sum_{s=-r+1}^k \gamma_s \cos \lambda(r+s) \right] v_r + \right. \\ \left. + \sum_{r=T-k+1}^{T+k} \left[\sum_{s=-k}^{T-r} \gamma_s \cos \lambda(r+s) \right] v_r \right\}.$$

Сумма квадратов коэффициентов величин $v_{1-k}, \dots, v_k, v_{T-k+1}, \dots, v_{T+k}$, умноженная на T , сходится к 0, когда T возрастает. Таким образом, предельное распределение величин $\sqrt{T}A_k(\lambda)$ есть предельное распределение произведения величин

$$2\sigma \sqrt{\sum_{r=k+1}^{T-k} \left[\sum_{s=-k}^k \gamma_s \cos \lambda(r+s) \right]^2 / T} \text{ и } \sum_{r=k+1}^{T-k} w_r^T,$$

где

$$(29) \quad w_r^T = \frac{\sum_{s=-k}^k \gamma_s \cos \lambda(r+s)}{\sigma \sqrt{\sum_{r=k+1}^{T-k} \left[\sum_{s=-k}^k \gamma_s \cos \lambda(r+s) \right]^2}} v_r, \\ r = k+1, \dots, T-k; T = 2k+1, \dots$$

Тогда из центральной предельной теоремы Линдеберга (теорема 7.7.2) следует, что $\sum_{r=k+1}^{T-k} w_r^T$ имеет нормальное предельное распределение (согласно доказательству теоремы 2.6.1), а другой сомножитель сходится к постоянной. Затем применим следствие 7.7.1,

чтобы показать, что $\sqrt{T}A(\lambda)$ имеет нормальное предельное распределение. Теорема следует из рассмотрения произвольной линейной комбинации

$$(30) \quad \sqrt{T} [a_1 A(\lambda_1) + b_1 B(\lambda_1) + \dots + a_n A(\lambda_n) + b_n B(\lambda_n)] = \\ = \frac{2}{\sqrt{T}} \sum_{t=1}^T y_t [a_1 \cos \lambda_1 t + b_1 \sin \lambda_1 t + \dots \\ \dots + a_n \cos \lambda_n t + b_n \sin \lambda_n t],$$

так как характеристическая функция предельного распределения $2n$ переменных определяется характеристическими функциями всех линейных комбинаций (теорема 7.7.7).■

Следствие 8.4.2. Если $y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, где $\{v_t\}$ состоит из независимо распределенных величин, причем v_t имеет среднее значение 0, дисперсию σ^2 и функцию распределения $F_t(v)$, $\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty$, и если (23) выполняется при $c \rightarrow \infty$, то $\sqrt{T} A^*(\lambda_1)$, $\sqrt{T} B^*(\lambda_1)$, ..., ..., $\sqrt{T} A^*(\lambda_n)$, $\sqrt{T} B^*(\lambda_n)$, $0 < \lambda_j < \pi$, $\lambda_i \neq \lambda_j$, $i \neq j$, $i, j = 1, \dots, n$, имеет нормальное предельное распределение со средними значениями 0, нулевыми ковариациями и дисперсиями, соответственно равными $4\pi f(\lambda_1)$, $4\pi f(\lambda_1)$, ..., $4\pi f(\lambda_n)$, $4\pi f(\lambda_n)$.

Доказательство. Следствие вытекает из теоремы, так как

$$(31) \quad \sqrt{T} A^*(\lambda) = \sqrt{T} A(\lambda) - \sqrt{T} (\bar{y} - \mu) \left\{ \frac{2}{T} \sum_{t=1}^T \cos \lambda t \right\},$$

$$(32) \quad \sqrt{T} B^*(\lambda) = \sqrt{T} B(\lambda) - \sqrt{T} (\bar{y} - \mu) \left\{ \frac{2}{T} \sum_{t=1}^T \sin \lambda t \right\},$$

а выражения в фигурных скобках стремятся к 0 при $T \rightarrow \infty$.

Теорема и следствие могут быть доказаны при замене (23) на предположение о том, что $\mathbb{E}|v_t|^{2+\delta} < M$ для некоторого $\delta > 0$. По поводу более слабых условий см. Олшен (1967).■

Иногда нас будут интересовать выражения $\bar{A}(\lambda) = A(\lambda)/\sqrt{c_0}$, $\bar{B}(\lambda) = B(\lambda)/\sqrt{c_0}$ и $\bar{I}(\lambda) = I(\lambda)/c_0$. Так как $c_0 \rightarrow \sigma(0)$ по вероятности и $\mathbb{E}\bar{A}(\lambda) = \mathbb{E}\bar{B}(\lambda) = 0$, то предельное распределение множества $\sqrt{T}\bar{A}(\lambda_j)$, $\sqrt{T}\bar{B}(\lambda_j)$, $j = 1, \dots, n$, то же самое, что и для $\sqrt{T}A(\lambda_j)$, $\sqrt{T}B(\lambda_j)$, $j = 1, \dots, n$, за исключением тех случаев, когда $\bar{f}(\lambda_j)$ заменено на $f(\lambda_j)$ для каждого j . Аналогично предель-

ное распределение величин $\bar{I}(\lambda_1), \dots, \bar{I}(\lambda_n)$ есть не что иное, как предельное распределение $I(\lambda_1), \dots, I(\lambda_n)$ с заменой $f(\lambda_j)$ на $\bar{f}(\lambda_j)$ для каждого j .

8.4.4. Выборочная спектральная плотность

Так как

$$(33) \quad I(\lambda) = \frac{T}{8\pi} [A^2(\lambda) + B^2(\lambda)],$$

то мы можем получить предельное распределение выборочной спектральной плотности $I(\lambda)$ из предельного распределения тригонометрических коэффициентов $A(\lambda)$ и $B(\lambda)$.

ТЕОРЕМА 8.4.4. Если $y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, где $\{v_t\}$ состоит из независимо распределенных случайных величин, причем v_t имеет среднее значение 0, дисперсию σ^2 и функцию распределения $F_t(v)$,

$\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty$ и если (23) выполняется при $c \rightarrow \infty$, то $2I(\lambda_1)/f(\lambda_1), \dots, 2I(\lambda_n)/f(\lambda_n)$, $0 < \lambda_j < \pi$, $\lambda_i \neq \lambda_j$, $i \neq j$, $i, j = 1, \dots, n$, имеет предельное распределение, в котором эти n переменных величин независимы и каждая имеет χ^2 -распределение с двумя степенями свободы.

Теорема следует из общего результата о том, что если случайный вектор $\mathbf{X}_T$ есть векторная функция $\mathbf{h}(\mathbf{Z}_T)$ случайного вектора $\mathbf{Z}_T$, $\mathbf{h}(\mathbf{Z})$ непрерывна и предельное распределение вектора $\mathbf{Z}_T$ является распределением вектора $\mathbf{Z}$, то предельное распределение $\mathbf{X}_T$ является распределением для $\mathbf{h}(\mathbf{Z})$. [См., например, Манн и Вальд (1943а).]

Теорема 8.4.4 также выполняется, если $I(\lambda_1), \dots, I(\lambda_n)$ заменить на $I^*(\lambda_1), \dots, I^*(\lambda_n)$.

Для больших выборок мы можем сказать, что $I(\lambda)$ распределено приблизительно как $f(\lambda)\chi^2/2$, где χ^2 имеет χ^2 -распределение с 2 степенями свободы. Это показывает, конечно, что $I(\lambda)$ не является состоятельной оценкой функции $f(\lambda)$. Однако тот факт, что $I(\lambda)$ для различных значений λ асимптотически независимы, означает, что усреднение $I(\lambda)$ по значениям λ в некотором интервале даст оценку соответствующего среднего величины $f(\lambda)$ с малой дисперсией. Если $f(\lambda)$ мало изменяется в этом интервале значений λ , то оценка будет приемлемой. Эта идея будет более подробно развита в следующей главе.

Несколько иной подход был использован Бартлеттом (1966, разд. 9.2.2).

ТЕОРЕМА 8.4.5. Если $y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, где $\mathbb{E}v_t = 0$, $\mathbb{E}v_t^2 = \sigma^2$ и $\mathbb{E}v_t v_s = 0$, $t \neq s$, и $\sum_{t=-\infty}^{\infty} |\gamma_t| \sqrt{|t|} < \infty$, то абсолютная величина разности между $I(\lambda)$ и $\pi f(\lambda) I_v(\lambda)/\sigma^2$, где

$$(34) \quad I_v(\lambda) = \frac{1}{2\pi T} \left[\left(\sum_{t=1}^T v_t \cos \lambda t \right)^2 + \left(\sum_{t=1}^T v_t \sin \lambda t \right)^2 \right],$$

имеет математическое ожидание, которое меньше некоторой постоянной, умноженной на $1/\sqrt{T}$. Если v_t независимы и $\mathbb{E}v_t^4 = 3\sigma^4 + \kappa_4 < \infty$, то среднеквадратичное отклонение этой разности меньше постоянной, умноженной на $1/\sqrt{T}$.

ДОКАЗАТЕЛЬСТВО. Пусть

$$\begin{aligned} (35) \quad \sqrt{T}A_y(\lambda) + i\sqrt{T}B_y(\lambda) &= \frac{2}{\sqrt{T}} \sum_{t=1}^T (y_t - \mu) e^{i\lambda t} = \\ &= \frac{2}{\sqrt{T}} \sum_{t=1}^T \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s} e^{i\lambda t} = \\ &= \frac{2}{\sqrt{T}} \sum_{s=-\infty}^{\infty} \gamma_s \sum_{r=1}^{T-s} v_r e^{i\lambda(r+s)} = \\ &= \frac{2}{\sqrt{T}} \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} \left[\sum_{r=1}^T v_r e^{i\lambda r} - \right. \\ &\quad \left. - \sum_{r=1}^{-s} v_r e^{i\lambda r} + \sum_{r=T+1}^{T-s} v_r e^{i\lambda r} - \right. \\ &\quad \left. - \sum_{r=T-s+1}^T v_r e^{i\lambda r} + \sum_{r=1-s}^0 v_r e^{i\lambda r} \right], \end{aligned}$$

где каждая из последних пар сумм есть 0, если верхний индекс меньше нижнего (т. е. $\sum_{r=T-s+1}^T v_r e^{i\lambda r} \equiv \sum_{r=1-s}^0 v_r e^{i\lambda r} \equiv 0$, если $s \leq 0$

и $\sum_{r=1}^{-s} v_r e^{i\lambda r} \equiv \sum_{r=T+1}^{T-s} v_r e^{i\lambda r} \equiv 0$, если $s \geq 0$).

Тогда

$$\begin{aligned} (36) \quad \sqrt{T} \left[A_y(\lambda) - A_v(\lambda) \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} \right] + \\ + i\sqrt{T} \left[B_y(\lambda) - B_v(\lambda) \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} \right] = \end{aligned}$$

$$\begin{aligned}
&= \frac{2}{\sqrt{T}} \sum_{s=1}^{\infty} \gamma_s e^{i\lambda s} \left[\sum_{r=1-s}^0 v_r e^{i\lambda r} - \sum_{r=T-s+1}^T v_r e^{i\lambda r} \right] + \\
&\quad + \frac{2}{T} \sum_{s=-\infty}^{-1} \gamma_s e^{i\lambda s} \left[\sum_{r=T+1}^{T-s} v_r e^{i\lambda r} - \sum_{r=1}^{-s} v_r e^{i\lambda r} \right] = \frac{2}{\sqrt{T}} \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} g_s(\lambda).
\end{aligned}$$

Отсюда математическое ожидание абсолютной величины квадрата выражения (36) равно

$$\begin{aligned}
(37) \quad &\frac{4}{T} \mathbb{E} \sum_{s,t=-\infty}^{\infty} \gamma_s \gamma_t e^{i\lambda(s-t)} g_s(\lambda) \bar{g}_t(\lambda) \leq \\
&\leq \frac{4}{T} \sum_{s,t=-\infty}^{\infty} |\gamma_s| |\gamma_t| \mathbb{E} |g_s(\lambda) \bar{g}_t(\lambda)| \leq \\
&\leq \frac{4}{T} \sum_{s,t=-\infty}^{\infty} |\gamma_s| |\gamma_t| \sqrt{\mathbb{E} |g_s(\lambda)|^2} \sqrt{\mathbb{E} |g_t(\lambda)|^2} = \\
&= \frac{4}{T} \left(\sum_{s,t=-\infty}^{\infty} |\gamma_s| \sqrt{\mathbb{E} |g_s(\lambda)|^2} \right)^2 \leq \\
&\leq \frac{4}{T} \left(\sigma \sum_{s=-\infty}^{\infty} |\gamma_s| \sqrt{2|s|} \right)^2 = \frac{8\sigma^2}{T} \left(\sum_{s=-\infty}^{\infty} |\gamma_s| \sqrt{|s|} \right)^2,
\end{aligned}$$

так как

$$\begin{aligned}
(38) \quad &\mathbb{E} |g_s(\lambda)|^2 = \mathbb{E} \left| \sum_{r=1-s}^0 v_r e^{i\lambda r} - \sum_{r=T-s+1}^T v_r e^{i\lambda r} \right|^2 = \\
&= \begin{cases} 2\sigma^2 & \text{при } 0 < s < T, \\ 2T\sigma^2 & \text{при } T \leq s \end{cases}
\end{aligned}$$

$$\begin{aligned}
(39) \quad &\mathbb{E} |g_s(\lambda)|^2 = \mathbb{E} \left| \sum_{r=T+1}^{T-s} v_r e^{i\lambda r} - \sum_{r=1}^{-s} v_r e^{i\lambda r} \right|^2 = \\
&= \begin{cases} 2|s|\sigma^2 & \text{при } -T < s < 0, \\ 2T\sigma^2 & \text{при } s \leq -T. \end{cases}
\end{aligned}$$

Точно также математическое ожидание четвертой степени абсолютной величины выражения (36) равно (когда v_t независимы и $\mathbb{E} v_t^4 < \infty$)

$$\begin{aligned}
(40) \quad &\frac{16}{T^2} \mathbb{E} \sum_{s,s',t,t'=-\infty}^{\infty} \gamma_s \gamma_{s'} \gamma_t \gamma_{t'} e^{i\lambda(s+s'-t-t')} g_s(\lambda) g_{s'}(\lambda) \bar{g}_t(\lambda) \bar{g}_{t'}(\lambda) \leq \\
&\leq \frac{16}{T^2} \sum_{s,s',t,t'=-\infty}^{\infty} |\gamma_s| |\gamma_{s'}| |\gamma_t| |\gamma_{t'}| [\mathbb{E} |g_s(\lambda)|^4]^{1/4} \times \\
&\quad \times [\mathbb{E} |g_{s'}(\lambda)|^4]^{1/4} [\mathbb{E} |g_t(\lambda)|^4]^{1/4} [\mathbb{E} |g_{t'}(\lambda)|^4]^{1/4} =
\end{aligned}$$

$$= \frac{16}{T^2} \left(\sum_{s=-\infty}^{\infty} |\gamma_s| |\mathcal{G}| g_s(\lambda) |^4 \right)^4 \leq \\ \leq \frac{48}{T^2} \left[(\mathcal{G} \sigma_t^4)^{1/4} \sum_{s=-\infty}^{\infty} |\gamma_s| \sqrt{2|s|} \right]^4.$$

Правую часть в (36) обозначим U . Тогда

$$(41) \quad \sqrt{T} [A_y(\lambda) + iB_y(\lambda)] = \sqrt{T} [A_v(\lambda) + iB_v(\lambda)] \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} + U$$

и

$$(42) \quad T [A_y^2(\lambda) + B_y^2(\lambda)] = T [A_v^2(\lambda) + B_v^2(\lambda)] \left| \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} \right|^2 + \\ + \sqrt{T} \left[A_v(\lambda) \left(\overline{U} \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} + U \sum_{s=-\infty}^{\infty} \gamma_s e^{-i\lambda s} \right) + \right. \\ \left. + iB_v(\lambda) \left(\overline{U} \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} - U \sum_{s=-\infty}^{\infty} \gamma_s e^{-i\lambda s} \right) \right] + |U|^2.$$

Математическое ожидание величины $|U|^2$ было оценено в формуле (37). Математическое ожидание абсолютной величины второго члена правой части выражения (42) меньше квадратного корня математического ожидания величины $T [A_v^2(\lambda) + B_v^2(\lambda)]$, которое меньше $4\sigma^2$, умноженного на квадратный корень $4\mathcal{G} |U|^2 \left| \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} \right|^2$, что в свою очередь меньше правой части (37), умноженной на $8\pi f(\lambda)/\sigma^2$. Это доказывает первую часть теоремы.

Из (42) получаем

$$(43) \quad (8\pi)^2 [I(\lambda) - 2\pi f(\lambda) I_v(\lambda)/\sigma^2]^2 = \\ = \left\{ 2\sqrt{T} \left[A_v(\lambda) \mathcal{R} \left(\overline{U} \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} \right) - \right. \right. \\ \left. \left. - B_v(\lambda) \mathcal{I} \left(\overline{U} \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} \right) \right] + |U|^2 \right\}^2 \leq \\ \leq 4T [|A_v(\lambda)| + |B_v(\lambda)|]^2 |U|^2 \left| \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} \right|^2 + \\ + 4\sqrt{T} [|A_v(\lambda)| + |B_v(\lambda)|] |U|^3 \left| \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} \right| + |U|^4.$$

Вторая часть теоремы следует из предыдущего, упр. 33, формулы (40) и условия, что $\mathcal{G} [\sqrt{T} A_v(\lambda)]^4$ и $\mathcal{G} [\sqrt{T} B_v(\lambda)]^4$ ограничены.

А. М. Уолкер (1965) показал, что предел по вероятности разности между $I(\lambda)$ и $2\pi f(\lambda) I_y(\lambda)/\sigma^2$ есть 0, где $y_t = \sum_{s=0}^{\infty} \gamma_s y_{t-s}$ и $\sum_{s=0}^{\infty} |\gamma_s| < \infty$. Олшен (1967) получил в этом направлении более глубокие результаты.

8.4.5. Выборочные корреляции

Если выборочные ковариации имеют нормальное предельное распределение, то же самое выполняется и для выборочных корреляций. Однако существование нормального предельного распределения корреляций может быть доказано и при более слабых условиях. Мы сделаем это в данном разделе. Соответствующий результат был приведен в теореме 5.7.1 для конечного процесса скользящего среднего.

Если $r_h = c_h/c_0$ и $\rho_h = \sigma(h)/\sigma(0)$, то можно написать

$$(44) \quad \sqrt{T}(r_h - \rho_h) = \frac{\sqrt{T}(c_h - \rho_h c_0)}{c_0}.$$

Знаменатель ведет себя, как $\sigma(0)$, а числитель имеет предельную ковариацию

$$\begin{aligned} (45) \quad & \lim_{T \rightarrow \infty} T \operatorname{Cov}[(c_h - \rho_h c_0), (c_g - \rho_g c_0)] = \\ & = \lim_{T \rightarrow \infty} T \operatorname{Cov}(c_h, c_g) - \rho_h \lim_{T \rightarrow \infty} T \operatorname{Cov}(c_0, c_g) - \\ & \quad - \rho_g \lim_{T \rightarrow \infty} T \operatorname{Cov}(c_0, c_h) + \rho_g \rho_h \lim_{T \rightarrow \infty} T \operatorname{Var} c_0 = \\ & = \sum_{r=-\infty}^{\infty} [\sigma(r+g)\sigma(r+h) + \sigma(r-g)\sigma(r+h) - \\ & \quad - 2\rho_h \sigma(r)\sigma(r+g) - 2\rho_g \sigma(r)\sigma(r+h) + \\ & \quad + 2\rho_g \rho_h \sigma^2(r) + \kappa(h, -r, g-r) - \\ & \quad - \rho_h \kappa(0, -r, g-r) - \rho_g \kappa(h, -r, -r) + \\ & \quad + \rho_g \rho_h \kappa(0, -r, -r)] = \\ & = \sum_{r=-\infty}^{\infty} [\xi y_t y_{t+h} y_{t-r} y_{t+g-r} - \rho_h \xi y_t^2 y_{t-r} y_{t+g-r} - \\ & \quad - \rho_g \xi y_t y_{t+h} y_{t-r}^2 + \rho_g \rho_h \xi y_t^2 y_{t-r}^2] = \\ & = \sum_{r=-\infty}^{\infty} \xi y_t y_{t-r} (y_{t+h} - \rho_h y_t) (y_{t+g-r} - \rho_g y_{t-r}), \end{aligned}$$

поскольку $\mathfrak{E}y_t = 0$, если ряд сходится. (См. разд. 8.3.2.) Если $y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, $\mathfrak{E}v_t = 0$, $\mathfrak{E}v_t^2 = \sigma^2$ и $\mathfrak{E}v_t^4 = 3\sigma^4 + \kappa_4 < \infty$, v_t независимы и $\sum_{s=-\infty}^{\infty} \gamma_s^2 < \infty$, то

$$(46) \quad \sum_{r=-\infty}^{\infty} [\kappa(h, -r, g-r) - \rho_h \kappa(0, -r, g-r) - \\ - \rho_g \kappa(h, -r, -r) + \rho_g \rho_h \kappa(0, -r, -r)] = \\ = \frac{\kappa_4}{\sigma^4} [\sigma(h) \sigma(g) - \rho_h \sigma(0) \sigma(g) - \rho_g \sigma(0) \sigma(h) + \rho_g \rho_h \sigma^2(0)] = 0.$$

Поэтому предельная ковариация (45) не зависит от κ_4 . Когда величины $\sqrt{T} [c_h - \sigma(h)]$, $\sqrt{T} [c_g - \sigma(g)]$ и $\sqrt{T} [c_0 - \sigma(0)]$ имеют нормальное предельное распределение, тогда величины $\sqrt{T} (r_h - \rho_h)$ и $\sqrt{T} (r_g - \rho_g)$ имеют нормальное предельное распределение со средними значениями 0 и ковариациями

$$(47) \quad w_{gh} = \sum_{r=-\infty}^{\infty} (\rho_{r+g} \rho_{r+h} + \rho_{r-g} \rho_{r+h} - 2\rho_h \rho_r \rho_{r+g} - \\ - 2\rho_g \rho_r \rho_{r+h} - 2\rho_g \rho_h \rho_r^2) = \\ = \frac{4\pi}{\sigma^2(0)} \int_{-\pi}^{\pi} (\cos v h - \rho_h) (\cos v g - \rho_g) f^2(v) dv,$$

которые являются интегральным представлением предельных ковариаций. Так как предельное распределение не зависит от κ_4 , можно ожидать, что результат получится при ограничениях только на вторые моменты.

ТЕОРЕМА 8.4.6. Пусть $y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, где $\sum_{h=-\infty}^{\infty} |\gamma_h| < \infty$, $\sum_{h=-\infty}^{\infty} |h| \gamma_h^2 < \infty$ и $\{v_t\}$ состоит из независимых и одинаково распределенных случайных величин с $\mathfrak{E}v_t = 0$ и $\mathfrak{E}v_t^2 = \sigma^2 < \infty$. Пусть $r_h = c_h/c_0$, $h = 1, \dots, m$. Тогда совместное распределение величин $\sqrt{T} (r_1 - \rho_1), \dots, \sqrt{T} (r_m - \rho_m)$, где $\rho_h = \sigma(h)/\sigma(0) = \sum_{s=-\infty}^{\infty} \gamma_s \gamma_{s+h} / \sum_{s=-\infty}^{\infty} \gamma_s^2$, стремится к $N(0, \mathbf{W})$ при $T \rightarrow \infty$, где $\mathbf{W} = (w_{gh})$ дается формулой (47).

Следствие 8.4.3. В условиях теоремы 8.4.6 совместное распределение величин $\sqrt{T} (r_h^* - \rho_h)$, $h = 1, \dots, m$, где $r_h^* = c_h^*/c_0^*$, стремится к $N(0, W)$ при $T \rightarrow \infty$, где $W = (w_{gh})$ дается формулой (47).

Доказательство. Сначала мы докажем, что $\sqrt{T} (r_l - \rho_l)$ имеет своим предельным распределением $N(0, w_{ll})$. Это будет следовать из того, что предельное распределение

$$(48) \quad z_T^{(l)} = \frac{1}{\sqrt{T}} \left\{ \sum_{i=1}^{T-l} y_i y_{i+l} - \rho_l \sum_{i=1}^T y_i^2 \right\},$$

где мы полагаем $\mu = 0$, есть $N(0, \sigma^4 s_l)$ с

$$(49) \quad s_l = \left(\sum_{h=-\infty}^{\infty} \gamma_h^2 \right)^2 w_{ll}.$$

Пусть

$$(50) \quad z_{T,k}^{(l)} = \frac{1}{\sqrt{T}} \left\{ \sum_{i=1}^{T-l} y_{i,k} y_{i+l,k} - \rho_{l,k} \sum_{i=1}^T y_{i,k}^2 \right\},$$

где $y_{t,k}$ определены в (3) и

$$(51) \quad \rho_{l,k} = \sum_{h=-k}^{k-l} \gamma_h \gamma_{h+l} / \sum_{h=-k}^k \gamma_h^2.$$

Подставляя выражения для $y_{t,k}$ в (50), имеем

$$(52) \quad z_{T,k}^{(l)} = \frac{1}{\sqrt{T}} \left[\sum_{i=1}^{T-l} \sum_{h,g=-k}^k \gamma_h \gamma_g v_{i-h} v_{i+l-g} - \rho_{l,k} \sum_{i=1}^T \sum_{h,g=-k}^k \gamma_h \gamma_g v_{i-h} v_{i-g} \right].$$

Пусть $z_{T,k}^{(l)*}$ обозначает выражение, полученное из $z_{T,k}^{(l)}$ путем отбрасывания всех членов, содержащих v_s^2 ($1 - k \leq s \leq T + k$), а именно

$$(53) \quad z_{T,k}^{(l)*} = \frac{1}{\sqrt{T}} \left[\sum_{i=1}^{T-l} \sum_{h=-k}^k \sum_{\substack{g=k-l \\ h \neq g}}^{k-l} \gamma_h \gamma_g v_{i-h} v_{i+l-g} - \rho_{l,k} \sum_{i=1}^T \sum_{\substack{h,g=-k \\ h \neq g}}^k \gamma_h \gamma_g v_{i-h} v_{i-g} \right].$$

ЛЕММА 8.4.1. В условиях теоремы 8.4.6 предельное распределение величины $z_{T,k}^{(l)*}$ при $T \rightarrow \infty$ есть $N(0, \sigma^4 s_{l,k})$, где

$$(54) \quad s_{l,k} = \sum_{r=1}^{2k+l} \delta_{r,k}^{(l)2},$$

$$(55) \quad \delta_{r,k}^{(l)} = \sum_{h=-k}^k [\gamma'_h \gamma'_{h+l+r} + \gamma'_h \gamma'_{h+l-r} - \rho_{l,k} (\gamma'_h \gamma'_{h+r} + \gamma'_h \gamma'_{h-r})],$$

причем $\gamma'_h = \gamma_h$, когда $|h| \leq k$, и $\gamma'_h = 0$, когда $|h| > k$.

ДОКАЗАТЕЛЬСТВО. Это было доказано в разд. 5.7.3. ■

Пусть $z_T^{(l)*}$ обозначает выражение, полученное заменой y_t на $\sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$ в (48) и отбрасыванием всех членов, содержащих v_r^2 ($-\infty < r < \infty$), а именно

$$(56) \quad z_T^{(l)*} = \frac{1}{\sqrt{T}} \left[\sum_{t=1}^{T-l} \sum_{\substack{h,g=-\infty \\ h \neq g}}^{\infty} \gamma_h \gamma_g v_{t-h} v_{t-g} - \right. \\ \left. - \rho_l \sum_{t=1}^T \sum_{\substack{h,g=-\infty \\ h \neq g}}^{\infty} \gamma_h \gamma_g v_{t-h} v_{t-g} \right].$$

ЛЕММА 8.4.2. В условиях теоремы 8.4.6 предельное распределение величины $z_T^{(l)*}$ при $T \rightarrow \infty$ есть $N(0, \sigma^4 s_l)$, где s_l дается формулой (49).

ДОКАЗАТЕЛЬСТВО. Предел распределения $N(0, \sigma^4 s_{l,k})$ (которое является предельным распределением для $z_{T,k}^{(l)*}$ при $T \rightarrow \infty$) есть $N(0, \sigma^4 s_l)$, когда $k \rightarrow \infty$, так как

$$(57) \quad \lim_{k \rightarrow \infty} s_{l,k} = \frac{1}{2} \sum_{r=-\infty}^{\infty} \left\{ \sum_{h=-\infty}^{\infty} [\gamma_h \gamma_{h+l+r} + \gamma_h \gamma_{h+l-r} - \right. \\ \left. - \rho_l (\gamma_h \gamma_{h+r} + \gamma_h \gamma_{h-r})] \right\}^2,$$

что эквивалентно (49). Лемма 8.4.2 будет вытекать из следствия 7.7.1, если мы покажем, что

$$(58) \quad \mathbb{E} |R_{T,k}^{(l)*}|^2 \leq M_k,$$

$$(59) \quad \lim_{k \rightarrow \infty} M_k = 0,$$

где

$$(60) \quad R_{T,k}^{(l)*} = z_T^{(l)*} - z_{T,k}^{(l)*}.$$

Пусть

$$(61) \quad R_{T,k}^{(l)} = z_T^{(l)} - z_{T,k}^{(l)}.$$

Тогда

$$(62) \quad \begin{aligned} \sqrt{T} R_{T,k}^{(l)} &= \sum_{t=1}^{T-l} [(y_{t,k} + u_{t,k})(y_{t+l,k} + u_{t+l,k}) - y_{t,k} y_{t+l,k}] + \\ &+ \sum_{t=1}^T [-\rho_l (y_{t,k} + u_{t,k})^2 + \rho_{l,k} y_{t,k}^2] = \\ &= \sum_{t=1}^{T-l} [u_{t,k} y_{t+l,k} + y_{t,k} u_{t+l,k} + u_{t,k} u_{t+l,k}] + \\ &+ \sum_{t=1}^T [(\rho_{l,k} - \rho_l) y_{t,k}^2 - 2\rho_l y_{t,k} u_{t,k} - \rho_l u_{t,k}^2] = \sum_{r=1}^6 T_r, \end{aligned}$$

где

$$(63) \quad T_1 = \sum_{t=1}^{T-l} u_{t,k} y_{t+l,k}, \dots, \quad T_6 = -\rho_l \sum_{t=1}^T u_{t,k}^2.$$

Поэтому из (53), (56) и (60) будем иметь

$$(64) \quad \sqrt{T} R_{T,k}^{(l)*} = \sum_{r=1}^6 T_r^*,$$

где T_r^* получается из T_r путем представления последнего линейной комбинацией конечного или счетного числа членов вида $\sum_{t=1}^{T-l} v_{t-h} v_{t-g}$.

или $\sum_{t=1}^T v_{t-h} v_{t-g}$, исключая все члены, для которых $h = g$.

Теперь мы имеем

$$(65) \quad \frac{1}{T} \mathbb{E} T_r^{*2} \leq M_{r,k}, \quad \lim_{k \rightarrow \infty} M_{r,k} = 0, \quad r = 1, \dots, 6.$$

Например,

$$(66) \quad \mathbb{E} T_3^{*2} = \sum_{\substack{|h|, |g+l| > k \\ h \neq g}} \gamma_h \gamma_{g+l} \sum_{\substack{|h'|, |g'+l| > k \\ h' \neq g'}} \gamma_{h'} \gamma_{g'+l} \sum_{t, t'=1}^{T-l} \mathbb{E} v_{t-h} v_{t-g} v_{t'-h'} v_{t'-g'}.$$

Если

$$(67) \quad \mathbb{E} v_{t-h} v_{t-g} v_{t'-h'} v_{t'-g'} \neq 0,$$

то мы имеем или $t - h = t' - h'$, $t - g = t' - g'$ и тогда $t' - t = h' - h = g' - g$, или $t - h = t' - g'$, $t - g = t' - h'$ и тогда $t' - t = g' - h = h' - g$; причем значение t' для ненулевой компоненты в последней сумме (66) (если такое имеется) однозначно.

определяется значением t . Из этого следует, что

$$(68) \quad \mathbb{E} T_3^{*2} \leq (T-l) \sigma^4 \sum_{|h|, |h'|, |g+l|, |g'+l| > k} |\gamma_h| |\gamma_{h'}| |\gamma_{g+l}| |\gamma_{g'+l}| = \\ = (T-l) \sigma^4 \left\{ \sum_{|h| > k} |\gamma_h| \right\}^4,$$

приводит к

$$(69) \quad \frac{1}{T} \mathbb{E} T_3^{*2} \leq \sigma^4 \left\{ \sum_{|h| > k} |\gamma_h| \right\}^4,$$

которое показывает, что (65) имеет место для $r = 3$. Аналогичные рассуждения применяются для $r = 1, 2, 5, 6$, а для $r = 4$ мы пользуемся тем, что $\lim_{k \rightarrow \infty} \rho_{l,k} = \rho_l$. Тогда получим (58) из (65), используя неравенство

$$(70) \quad \mathbb{E} \left(\sum_{r=1}^6 T_r^* \right)^2 \leq 6 \sum_{r=1}^6 \mathbb{E} T_r^{*2}.$$

Это доказывает лемму 8.4.2. ■

Лемма 8.4.3. В условиях теоремы 8.4.6 предельное распределение для $z_T^{(l)}$ есть $N(0, \sigma^4 s_l)$, когда $T \rightarrow \infty$.

Доказательство. Имеем

$$(71) \quad z_T^{(l)} = z_T^{(l)*} = \frac{1}{\sqrt{T}} \left[\sum_{t=1}^{T-l} \sum_{h=-\infty}^{\infty} \gamma_h \gamma_{h+l} v_{t-h}^2 - \rho_l \sum_{t=1}^T \sum_{h=-\infty}^{\infty} \gamma_h^2 v_{t-h}^2 \right] = \\ = \frac{1}{\sqrt{T}} \left[\sum_{h=-\infty}^{\infty} \gamma_h \gamma_{h+l} \sum_{u=1-h}^{T-l-h} v_u^2 - \rho_l \sum_{h=-\infty}^{\infty} \gamma_h^2 \sum_{u=1-h}^{T-h} v_u^2 \right] = \\ = \frac{1}{\sqrt{T}} \left[\sum_{h=-\infty}^{\infty} \gamma_h \gamma_{h+l} \left(\sum_{t=1}^T v_t^2 + U_{T,h}^{(l)} \right) - \right. \\ \left. - \rho_l \sum_{h=-\infty}^{\infty} \gamma_h^2 \left(\sum_{t=1}^T v_t^2 + U_{T,h}^{(0)} \right) \right] = \\ = \frac{1}{\sqrt{T}} \left[\sum_{h=-\infty}^{\infty} \gamma_h \gamma_{h+l} U_{T,h}^{(l)} - \rho_l \sum_{h=-\infty}^{\infty} \gamma_h^2 U_{T,h}^{(0)} \right],$$

где

$$(72) \quad U_{T,h}^{(l)} = \sum_{t=1-h}^{T-l-h} v_t^2 - \sum_{t=1}^T v_t^2.$$

Тогда

$$(73) \quad \mathbb{E} |U_{T,h}^{(l)}| \leq (2|h| + l) \sigma^2$$

и

$$(74) \quad \mathbb{E} |\sqrt{T} z_T^{(l)} - z_T^{(l)*}| \leq \sigma^2 \left[\sum_{h=-\infty}^{\infty} |\gamma_h| |\gamma_{h+l}| (2|h| + 1) + \right. \\ \left. + |\rho_l| \sum_{h=-\infty}^{\infty} \gamma_h^2 2|h| \right] < \infty,$$

поскольку

$$(75) \quad \left[\sum_{h=-\infty}^{\infty} |h| |\gamma_h| |\gamma_{h+l}| \right]^2 \leq \sum_{h=-\infty}^{\infty} |h| \gamma_h^2 \sum_{h=-\infty}^{\infty} |h| \gamma_{h+l}^2 = \\ = \sum_{h=-\infty}^{\infty} |h| \gamma_h^2 \sum_{g=-\infty}^{\infty} |g-l| \gamma_g^2 < \infty.$$

Таким образом,

$$(76) \quad \text{plim}_{T \rightarrow \infty} [z_T^{(l)} - z_T^{(l)*}] = 0.$$

Это и доказывает лемму 8.4.3. ■

Лемма 8.4.4. В условиях теоремы 8.4.6 предельное распределение для $\sqrt{T} (r_l - \rho_l)$ при $T \rightarrow \infty$ есть $N(0, w_{ll})$.

Доказательство. Это следует из леммы 8.4.3 и условий

$$(77) \quad \sqrt{T} (r_l - \rho_l) = \sqrt{T} \frac{c_l - \rho_l c_0}{c_0}$$

и

$$(78) \quad \text{plim}_{T \rightarrow \infty} c_0 = \sigma(0) = \sigma^2 \sum_{h=-\infty}^{\infty} \gamma_h^2.$$

Последняя формула вытекает из следующей записи

$$(79) \quad \frac{1}{T} \sum_{t=1}^T y_t^2 = \frac{1}{T} \sum_{t=1}^T y_{t,k}^2 + \frac{2}{T} \sum_{t=1}^T y_{t,k} u_{t,k} + \frac{1}{T} \sum_{t=1}^T u_{t,k}^2$$

и использования

$$(80) \quad \frac{1}{T} \mathbb{E} \sum_{t=1}^T u_{t,k}^2 = \sigma \sum_{|h| > k} \gamma_h^2 \rightarrow 0,$$

когда $k \rightarrow \infty$, и

$$(81) \quad \frac{1}{T} \mathbb{E} \left| \sum_{t=1}^T y_{t,k} u_{t,k} \right| \leq \mathbb{E} |y_{t,k} u_{t,k}| \leq [\mathbb{E} y_{t,k}^2 \mathbb{E} u_{t,k}^2]^{1/2} = \\ = \sigma^2 \left\{ \sum_{|h| \leq k} \gamma_h^2 \sum_{|h| > k} \gamma_h^2 \right\}^{1/2} \rightarrow 0,$$

когда $k \rightarrow \infty$. В разд. 5.7.3 было показано, что

$$(82) \quad \text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{i=1}^T y_{i,k}^2 = \sigma^2 \sum_{h=-k}^k \gamma_h^2 \rightarrow \sigma(0),$$

когда $k \rightarrow \infty$. Это доказывает лемму 8.4.4. ■

Для полного доказательства теоремы 8.4.6 отметим, что аналогичным образом можно показать, что для $\sum_{l=1}^m k_l \sqrt{T} (r_l - \rho_l)$, где $\mathbf{k} = (k_1, \dots, k_m)'$ — произвольный набор постоянных, предельное распределение есть $N(0, \mathbf{k}' \mathbf{W} \mathbf{k})$. Тогда по теореме 7.7.7 следует, что совместное предельное распределение для $\sqrt{T} (r_l - \rho_l)$, $l = 1, \dots, m$, есть $N(0, \mathbf{W})$.

Следствие вытекает из теорем 8.4.6 и 8.4.1, поскольку $\sqrt{T} (c_l - c_l^*)$, а значит, и $\sqrt{T} (r_l - r_l^*)$ сходятся по вероятности к 0 при $T \rightarrow \infty$.

Теорема 8.4.6, которая предполагает существование только вторых моментов величин v_t , была доказана Т. Андерсоном и А. Уолкером (1964). В предыдущих работах существование старших моментов величин v_t всегда предполагалось. Например, Манн и Вальд (1943b) при использовании стохастических разностных уравнений полагали, что $\mathbb{E} |v_t|^r < \infty$ для всех $r > 0$. Диананда (1953) предполагал, что $\mathbb{E} v_t^4 < \infty$. Это является ослаблением условия $\mathbb{E} v_t^6 < \infty$, использованного ранее Хёффдингом и Роббинсом (1948). А. Уолкер (1954), обобщая метод Диананды для произвольного линейного процесса, сохраняет его условие $\mathbb{E} v_t^4 < \infty$ и требует, чтобы $\sum_{h=-\infty}^{\infty} |h \gamma_h| < \infty$. (Диананда и А. Уолкер рассматривали

обобщение линейного процесса, в котором $\{v_t\}$ есть просто стационарный процесс с конечной зависимостью.) [Тот факт, что члены, соответствующие $\mathbb{E} v_t^4$, в (45) не входят (это впервые заметил Бартлетт (1946, стр. 29)), делает правдоподобной гипотезу о том, что можно не предполагать конечности $\mathbb{E} v_t^4$.] Действительно, Т. Андерсон (1959, разд. 4) доказал результат, эквивалентный асимптотической нормальности величины r_1 для стохастического разностного уравнения первого порядка. Настоящее доказательство, по существу, представляет собой обобщение указанного метода.

8.5. ПРИМЕРЫ

Представление о поведении выборочной ковариационной последовательности или выборочной корреляционной последовательности простого стационарного случайного процесса можно получить,

рассмотрев искусственные ряды, смоделированные Вольдом (1965) с использованием случайных чисел, как это описано в § А.2. Выборочные корреляционные последовательности приведены там для случаев процесса авторегрессии второго порядка и изображены графически в § А.2. Следует заметить, сколь они нерегулярны. Теоретические спектральные плотности (и спектральные плотности, определенные по оцененным процессам авторегрессии второго порядка) представлены на рис. А.2.7, А.2.8 и А.2.9.

Выборочная спектральная плотность (умноженная на постоянную) ряда Бевериджа цен на пшеницу с выделенным трендом (описанного в § 4.5) табулируется в табл. А.1.3 и показана на рис. А.1.3. Они также обнаруживают значительную вариабельность.

Другой пример выборочной спектральной плотности соответствует данным о числе солнечных пятен (рассмотренным в § 5.9), представленным в табл. А.3.2 и на рис. А.3.1. Этот случай не похож на другие. Здесь имеется довольно хорошо выраженный пик на частоте около 0.09 (период около 11 лет).

8.6. ОБСУЖДЕНИЕ

Информация, полученная из наблюдаемых временных рядов порожденных случайными процессами, стационарными в широком смысле, может быть представлена выборочным средним и соответственно выборочной ковариационной последовательностью, выборочной дисперсией и выборочной корреляционной последовательностью, выборочной спектральной плотностью или выборочной дисперсией и нормализованной выборочной спектральной плотностью. Выборочные ковариации и спектральные плотности подвержены значительной выборочной вариабельности. Ковариации или корреляции для малых запаздываний дают подходящую информацию о зависимости в процессе, но для больших запаздываний они менее информативны; меньшая зависимость мало связана с выборочной вариабельностью.

Как мы увидим в следующей главе, выборочную спектральную плотность можно сгладить, чтобы получить оценку теоретической спектральной плотности с относительно малой изменчивостью.

Выборочная спектральная функция $\hat{F}_T(\nu) = \int_{-\pi}^{\nu} I(\lambda) d\lambda$ является состоятельной оценкой величины $F(\nu)$ в точке непрерывности и $\sqrt{T} [\hat{F}_T(\nu) - F(\nu)]$ имеет нормальное предельное распределение. (См. упр. 26, гл. 9.) Ковариационная функция предельной нормальной функции распределения (для неотрицательных значений ν) подобна эмпирическим функциям распределения ($0 \leq \nu \leq \pi$), за исключением того, что спектральная функция распределения не

обязательно равна 1 при $\nu = \pi$; можно использовать соответствующим образом измененные критерии согласия. [См. Гренандер и Розенблатт (1957, гл. 6).]

ЛИТЕРАТУРА

- § 8.2. Ланцош (1956).
 § 8.3. Бартлетт (1935), (1946), Парзен (1957b).
 § 8.4. Т. Андерсон (1959), Т. Андерсон и А. Уолкер (1964), Бартлетт (1946), (1966), Диананда (1953), Манн и Вальд (1943a), (1943b), Олшен (1967), А. Уолкер (1954), (1965), Хёфдинг и Роббинс (1948).
 § 8.5. Вольд (1965).
 § 8.6. Гренандер и Розенблатт (1957).

УПРАЖНЕНИЯ

1. (Разд. 8.2.2) Покажите, что если $\sum_{s=1}^T \sigma(t-s)$ не зависит от t , $t = 1, \dots, T$, то $\sigma(h) = \sigma(T-h)$, $h = 1, \dots, T-1$. [Указание.

$$\sum_{s=1}^T \sigma(t-s) = \sigma(0) + 2 \sum_{h=1}^{\min(t-1, T-t)} \sigma(h) + \sum_{h=\min(t, T-t+1)}^{\max(t-1, T-t)} \sigma(h),$$

$$t = 2, \dots, T-1,$$

$$\left. \sum_{s=1}^T \sigma(1-s) = \sum_{h=0}^{T-1} \sigma(h) = \sum_{s=1}^T \sigma(T-s). \right]$$

2. (Разд. 8.2.2) Найдите свойства симметрии семиинвариантов четвертого порядка.

а) Покажите, что

$$\kappa(h, r, s) = \kappa(h, s, r) = \kappa(r, h, s) = \kappa(r, s, h) = \kappa(s, h, r) = \kappa(s, r, h).$$

б) Покажите, что

$$\begin{aligned} \kappa(h, r, s) &= \kappa(-h, r-h, s-h) = \kappa(-r, h-r, s-r) = \\ &= \kappa(-s, h-s, r-s). \end{aligned}$$

с) Найдите 15 других форм $\kappa(h, r, s)$.

3. (Разд. 8.2.2) Найдите ξQ_1^* , определенное циклически, когда $\{y_t\}$ есть случайный процесс, стационарный в широком смысле.

4. (Разд. 8.2.2) Найдите ξQ_1^* , определенное по последовательным разностям, когда $\{y_t\}$ есть процесс, стационарный в широком смысле.

5. (Разд. 8.2.2) Пусть

$$Q = a_{11}y_1^2 + a_{22}y_2^2 + 2a_{12}y_1y_2,$$

$$\tilde{Q} = \frac{1}{2} (a_{11} + a_{22}) (y_1^2 + y_2^2) + 2a_{12}y_1y_2.$$

где $\mathbb{E}y_1 = \mathbb{E}y_2 = 0$, $\mathbb{E}y_1^2 = \mathbb{E}y_2^2 = \sigma(0)$, $\mathbb{E}y_1y_2 = \sigma(1)$, а y_1 и y_2 имеют двумерное нормальное распределение. Покажите, что

$$\mathbb{E}Q - \mathbb{E}\tilde{Q} = 0,$$

$$\text{Var } Q - \text{Var } \tilde{Q} = (a_{11} - a_{22})^2 [\sigma^2(0) - \sigma^2(1)].$$

6. (Разд. 8.2.2) Пусть

$$Q = 3y_1^2 + y_2^2 + 2y_3^2,$$

$$\tilde{Q} = 2y_1^2 + 2y_2^2 + 2y_3^2,$$

где $\mathbb{E}y_t = 0$, $\mathbb{E}y_t y_{t+h} = \sigma(h)$, $t = 1, 2, 3$, $h = 0, 1, 2$, а y_1, y_2, y_3 имеют совместное нормальное распределение. Покажите, что

$$\mathbb{E}Q - \mathbb{E}\tilde{Q} = 0,$$

$$\text{Var } Q - \text{Var } \tilde{Q} = 4 [\sigma^2(0) - 3\sigma^2(1) + 2\sigma^2(2)].$$

Покажите, что разность между дисперсиями может быть положительной, нулем или отрицательной. [Указание. Для вывода последнего утверждения использовать стационарный гауссовский процесс первого порядка.]

7. (Разд. 8.2.2) Выведите леммы 3.4.3, 3.4.4 и 3.4.5 как частные случаи теорем 8.2.4 и 8.2.5.

8. (Разд. 8.2.2) Проверьте соотношения (51) — (55).

9. (Разд. 8.2.2) Проверьте формулу (58).

10. (Разд. 8.2.2) Проверьте

$$\begin{aligned} \mathbb{E}\tilde{C}_h = \sigma(h) - \frac{1}{T-h} \left[\left(1 - \frac{h}{T-h}\right) \sigma(0) + 2 \left(1 - \frac{h}{T-h}\right) \sum_{r=1}^h \sigma(r) + \right. \\ \left. + \sum_{r=h+1}^{T-1} \left(1 - \frac{r-h}{T-h}\right) \sigma(r) \right], \quad 1 \leq h = T-2h-1, \end{aligned}$$

$$\begin{aligned} \mathbb{E}\tilde{C}_h = \sigma(h) - \frac{1}{T-h} \left[\left(1 - \frac{h}{T-h}\right) \sigma(0) + \left(2 - \frac{2h}{T-h}\right) \sum_{r=1}^{T-2h-1} \sigma(r) + \right. \\ \left. + \sum_{r=T-2h}^h \left(1 - \frac{h-r}{T-h}\right) \sigma(r) + \sum_{r=h+1}^{T-1} \left(1 - \frac{r-h}{T-h}\right) \sigma(r) \right], \\ 1 \leq T-2h-1 < h < T-h-1, \end{aligned}$$

$$\begin{aligned} \mathbb{E}\tilde{C}_h = \sigma(h) - \frac{1}{T-h} \left[\sum_{r=2h-T+1}^h \left(1 - \frac{h-r}{T-h}\right) \sigma(r) + \right. \\ \left. + \sum_{r=h+1}^{T-1} \left(1 - \frac{r-h}{T-h}\right) \sigma(r) \right], \end{aligned}$$

$$T-h-1 \leq h < T-1.$$

11. (Разд. 8.2.2) Выведите вторую часть формулы (65) из первой ее части.

12. (Разд. 8.2.2) Покажите, что если $y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, $t = \dots, -1,$

$0, 1, \dots$, с $\sum_{r=-\infty}^{\infty} |\gamma_r| < \infty$ и независимыми v_r , $\xi v_r = 0$ и $\xi v_r^3 = \kappa_3$, то

$$(i) \quad \text{Cov}(\bar{y}, C_0) = \frac{\kappa_3}{T} \sum_{h=-(T-1)}^{T-1} \left(1 - \frac{|h|}{T}\right) \sum_{r=-\infty}^{\infty} \gamma_r^2 \gamma_{r+h}.$$

Покажите, что

$$(ii) \quad \lim_{T \rightarrow \infty} \text{Cov}(\bar{y}, C_0) = 0.$$

13. (Разд. 8.2.3) Покажите, что

$$k'_T(v) = \frac{[T \cos(vT/2) \sin(v/2) - \sin(vT/2) \cos(v/2)] \sin(vT/2)}{2\pi T \sin^3(v/2)}.$$

14. (Разд. 8.2.3) Покажите, что

$$k_T(v \pm 2\pi n) = k_T(v), \quad n = 1, 2, \dots$$

15. (Разд. 8.2.3) Покажите, что

$$\frac{d}{d\lambda} \frac{\sin(\lambda T/2)}{\sin(\lambda/2)} = \frac{T \cos(\lambda T/2) \sin(\lambda/2) - \sin(\lambda T/2) \cos(\lambda/2)}{2 \sin^2(\lambda/2)}.$$

16. (Разд. 8.2.3) Покажите, что

$$\frac{\sin[(\lambda \pm 2\pi k) T/2]}{\sin[(\lambda \pm 2\pi k)/2]} = (-1)^{k(T-1)} \frac{\sin(\lambda T/2)}{\sin(\lambda/2)}, \quad k = 1, 2, \dots$$

17. (Разд. 8.2.3) Покажите, что если $\{y_t\}$ состоит из независимых случайных величин с $\xi y_t = 0$, $\xi y_t^2 = \sigma^2$ и $\xi y_t^4 = 3\sigma^4 + \kappa_4$, то

$$\text{Var } I(\lambda) = \frac{\sigma^4}{(2\pi)^2} + \frac{\kappa_4}{(2\pi)^2 T} + \frac{\sigma^4}{(2\pi)^2 T^2} \frac{\sin^2 \lambda T}{\sin^2 \lambda}, \quad \lambda \neq 0, \pm \pi.$$

$$\text{Var } I(0) = \text{Var } I(\pm \pi) = \frac{2\sigma^4}{(2\pi)^2} + \frac{\kappa_4}{(2\pi)^2 T},$$

$$\text{Cov}[I(\lambda), I(\lambda')] = \frac{\kappa_4}{(2\pi)^2 T} + \frac{\sigma^4}{(2\pi)^2 T^2} \left[\frac{\sin^2(\lambda + \lambda') T/2}{\sin^2(\lambda + \lambda')/2} + \frac{\sin^2(\lambda - \lambda') T/2}{\sin^2(\lambda - \lambda')/2} \right],$$

$\lambda \neq \pm \lambda'.$

18. (Разд. 8.3.1) Проверьте условия теоремы 8.3.1 для процесса авторегрессии первой степени, когда $|\beta_1| < 1$.

19. (Разд. 8.3.1) Проверьте условия теоремы 8.3.1 для стационарного в широком смысле процесса, порожденного стохастическим разностным уравнением.

20. (Разд. 8.3.1) Пусть $a_1, a_2, \dots$ — последовательность чисел и

$$s_n = a_1 + a_2 + \dots + a_n,$$

$$\bar{s}_n = \frac{s_1 + s_2 + \dots + s_n}{n}.$$

Покажите, что

$$\bar{s}_n = \sum_{r=0}^{n-1} \left(1 - \frac{r}{n}\right) a_{r+1}.$$

[Указание. Воспользоваться индукцией.]

21. (Разд. 8.3.1) Докажите, что если $h(\lambda)$ непрерывна при $\lambda = v$, $|h(\lambda)|$ ограничено,

$$\int_{-\pi}^{\pi} l_T(\lambda) d\lambda = 1, \quad T = T_0, \dots,$$

существует такое число K , что

$$\int_{-\pi}^{\pi} |l_T(\lambda)| d\lambda \leq K, \quad T = T_1, \dots,$$

и если

$$|l_T(\lambda)| \leq m_T n(\lambda), \quad \lambda \neq v, \quad T = T_2, \dots,$$

где

$$\int_{-\pi}^{v-a} n(\lambda) d\lambda + \int_{v+a}^{\pi} n(\lambda) d\lambda < \infty$$

для каждого $a > 0$ и $m_T \rightarrow 0$ при $T \rightarrow \infty$, то

$$\lim_{T \rightarrow \infty} \int_{-\pi}^{\pi} h(\lambda) l_T(\lambda) d\lambda = h(v).$$

22. (Разд. 8.3.2) Докажите формулу (28) для $\tilde{C}_h$.

23. (Разд. 8.3.2) Докажите формулу (29).

24. (Разд. 8.3.2) Покажите, что для четных $g - h (\geq 0)$

$$\int_{-\pi}^{\pi} \frac{\sin [\mu (T-h)/2] \sin [\mu (T-g)/2]}{(T-g) \sin^2 (\mu/2)} d\mu = 2\pi.$$

25. (Разд. 8.3.2) Докажите, что $\sum_{t=-\infty}^{\infty} \gamma_t^2 < \infty$ следует из $\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty$.

[Указание. Второе условие влечет $\lim_{t \rightarrow \pm \infty} |\gamma_t| = 0$.]

26. (Разд. 8.3.2) Докажите, что $\sum_{h=-\infty}^{\infty} \sigma^2(h) < \infty$, если $y_t = \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$,

$$\sum_{s=-\infty}^{\infty} |\gamma_s| < \infty, \quad \mathcal{E} v_t = 0, \quad \mathcal{E} v_t^2 = \sigma^2, \quad \mathcal{E} v_t v_s = 0, \quad t \neq s.$$

27. (Разд. 8.3.2) Докажите, что из условия $\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty$ следует, что

$$f(\lambda) = \sigma^2 \left| \sum_{s=-\infty}^{\infty} \gamma_s e^{i\lambda s} \right|^2 / (2\pi) \text{ непрерывна.}$$

28. (Разд. 8.3.2) Пусть $y_t = \sum_{s=1}^{\infty} \gamma_s v_{t-s}$, $t = \dots, -1, 0, 1, \dots$, где $\gamma_t = (-1)^t/t$, $\mathbb{E}v_t = 0$, $\mathbb{E}v_t^2 = \sigma^2$, $\mathbb{E}v_t v_s = 0$, $t \neq s$.

а) Покажите, что

$$\sigma(h) = (-1)^h \frac{\sigma^2}{h} \sum_{t=1}^h \frac{1}{t}, \quad h = 1, 2, \dots$$

б) Покажите, что $\sum_{h=-\infty}^{\infty} \sigma^2(h) < \infty$.

с) Покажите, что $\sum_{h=-\infty}^{\infty} \sigma(h) < \infty$.

29. (Разд. 8.3.2) Докажите, что из условия $\sum_{h=-\infty}^{\infty} \sigma(h) < \infty$ не следует, что

$$\sum_{h=-\infty}^{\infty} \sigma^2(h) < \infty.$$

30. (Разд. 8.3.2) Покажите, что из условия $\left| \sum_{s,r,q=-\infty}^{\infty} \kappa(s, r, q) \right| < \infty$ следует формула (56) для любого h . Покажите, что это условие влечет за собой существование пределов трех семинвариантных сумм четвертого порядка в формуле (54).

31. (Разд. 8.3.2) Найдите

$$\lim_{T \rightarrow \infty} (T - g) [(T - h) \text{Cov}(C_g, C_h) - (T - h) \text{Cov}(\tilde{C}_g, \tilde{C}_h)],$$

используя интегральные представления ковариаций, когда $\kappa(r, s, t) = 0$ и $f(\lambda)$ непрерывна при $-\pi \leq \lambda \leq \pi$.

32. (Разд. 8.4.3) Покажите, что $\sum_{t=1}^T a_t y_t / \sqrt{T}$ имеет нормальное предельное распределение, если $|a_t| \leq 1$ и $\{y_t\}$ состоит из независимых и одинаково распределенных случайных величин с $\mathbb{E}y_t = 0$ и $\mathbb{E}y_t^2 = \sigma^2 < \infty$. [Указание. Показать, что $\{a_t y_t\}$ удовлетворяет условию Линдеберга. Если $\sum_{t=1}^{\infty} a_t^2 < \infty$, то дисперсия нормального предельного распределения равна 0.]

33. (Разд. 8.4.4) Покажите, что

$$(\mathbb{E} |X_1 X_2 X_3 X_4|)^4 \leq \mathbb{E} X_1^4 \mathbb{E} X_2^4 \mathbb{E} X_3^4 \mathbb{E} X_4^4.$$

[Указание. Использовать условие, что $(\mathbb{E} |Y_1 Y_2|)^2 \leq \mathbb{E} Y_1^2 \mathbb{E} Y_2^2$, где $Y_1 = X_1 X_2$ и $Y_2 = X_3 X_4$.]

34. (Разд. 8.4.5) Пусть $y_t = \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, где величины v_t независимы и одинаково распределены с $\mathbb{E}v_t = 0$, $\mathbb{E}v_t^2 = \sigma^2$, $\mathbb{E}v_t^4 = 3\sigma^4 + \kappa_4 < \infty$. Пусть $\text{plim}_{T \rightarrow \infty} \sum_{t,s=1}^T a_{ts}^{(T)} y_t y_s / \sum_{t,s=1}^T b_{ts}^{(T)} y_t y_s = \alpha/\beta$. При каких условиях дисперсия асимптотического распределения этого отношения не зависит от κ_4 ?

ОЦЕНИВАНИЕ СПЕКТРАЛЬНОЙ ПЛОТНОСТИ

9.1. ВВЕДЕНИЕ

Если стационарный случайный процесс имеет абсолютно непрерывную спектральную функцию, то дисперсия и ковариации такого процесса однозначно выражаются через его спектральную плотность. Представляет интерес оценка спектральной плотности процесса по его наблюдениям $y_1, \dots, y_T$ в T последовательных моментах времени. В гл. 8 была подробно изучена выборочная спектральная плотность

$I(\lambda) = [1/(2\pi)] \sum_{r=-(T-1)}^{T-1} \cos \lambda r c_r$. Было показано, что $\mathcal{E}I(\lambda)$ сходится при $T \rightarrow \infty$ к теоретической спектральной плотности $f(\lambda)$. Однако дисперсия величины $I(\lambda)$ не стремится при этом к нулю, вследствие чего $I(\lambda)$ не может считаться удовлетворительной оценкой для $f(\lambda)$.

В настоящей главе будет рассмотрено оценивание значения $f(\lambda)$ в точке $\lambda = \nu$ с помощью величин вида

$$\hat{f}(\nu) = (2\pi)^{-1} \sum_{r=-(T-1)}^{T-1} w_r \cos \nu r C_r = (2\pi)^{-1} \sum_{r=-(T-1)}^{T-1} w_r^* \cos \nu r c_r,$$

где $w_r = (T - |r|) w_r^* / T$ суть надлежащим образом выбранные числа, зависящие от T . Такую оценку можно записать иначе в виде

$\int_{-\pi}^{\pi} w^*(\lambda | \nu) I(\lambda) d\lambda$, т. е. в виде взвешенного среднего величины

$I(\lambda)$. Весовая функция $w^*(\lambda | \nu) = (2\pi)^{-1} \sum_{r=-(T-1)}^{T-1} w_r^* \cos \lambda r \cos \nu r$

также может зависеть от T . Подобного рода оценки вводятся в 9.2. Там же приведен целый ряд примеров. Математическое ожидание

такой оценки равно $\int_{-\pi}^{\pi} w(\lambda | \nu) f(\lambda) d\lambda$, причем и здесь весовая функция $w(\lambda | \nu) = (2\pi)^{-1} \sum_{r=-(T-1)}^{T-1} w_r \cos \lambda r \cos \nu r$ может зависеть от T . Состоятельно оценить $\hat{f}(\nu)$ можно с помощью последовательности таких оценок. В разд. 9.3.2 указан метод построения подобной последовательности и вычислено асимптотическое смещение. В разд. 9.3.3 получены асимптотические дисперсии и ковариации указанных оценок. Эти оценки при соответствующих условиях оказываются (§ 9.4) асимптотически нормальными.

Всюду в этой главе предполагается, что $\sigma(0) = \int_{-\pi}^{\pi} f(\lambda) d\lambda < \infty$.

Поскольку $C_r = C_{-r}$, $\cos \lambda r = \cos(-\lambda r)$ и т. п., в дальнейшем мы всегда будем брать $w_r = w_{-r}$ и $w_r^* = w_{-r}^*$ (чем, не теряя общности, упростим вычисления).

Если $\mathcal{E}y_t = \mu$ неизвестно, то в указанной выше оценке $\hat{f}(\nu)$ следует вместо C_r и c_r брать C_r^* , c_r^* , $\tilde{C}_r$ или $\tilde{c}_r$. Тогда вместо $I(\lambda)$ взвешиваются уже $I^*(\lambda)$ или $\tilde{I}(\lambda)$. Асимптотические свойства этих величин одни и те же.

Если $\hat{f}(\nu) \geq 0$, $-\pi \leq \nu \leq \pi$, она представляет собой спектральную плотность некоторого конечного процесса скользящего среднего, поскольку в этом случае $w_r C_r = w_{-r} C_{-r}$ действительны, $r = 0, 1, \dots, T-1$. Кроме того, если $w_r = 0$, $r = K+1, \dots, T-1$, то соответствующий процесс скользящего среднего является усреднением $K+1$ некоррелированных переменных. Заметим, что если $w^*(\lambda | 0) \geq 0$, $-\pi \leq \lambda \leq \pi$, то и $\hat{f}(\nu) \geq 0$, $-\pi \leq \nu \leq \pi$.

9.2. ОЦЕНКИ, ОСНОВАННЫЕ НА ВЫБОРОЧНЫХ КОВАРИАЦИЯХ

9.2.1. Квадратичные формы в качестве оценок спектральной плотности

Как было показано ранее, значение теоретической спектральной плотности $f(\lambda)$ в точке $\lambda = \nu$ может быть представлено в виде линейной комбинации ковариаций процесса. В то же время каждая из этих ковариаций $\sigma(h)$ является математическим ожиданием некоторой квадратичной формы от $y_t - \mu$. Поэтому представляется естественным оценивать $\hat{f}(\nu)$ с помощью квадратичных форм от наблюдаемых значений ряда. Мы будем предполагать пока, что сред-

нее μ известно и равно нулю. Возьмем квадратичную форму в виде

$$(1) \quad \sum_{s,t=1}^T w_{st} y_s y_t.$$

[Если μ известно, но не равно 0, то следует y_t в (1) заменить на $y_t - \mu$.] Ее математическое ожидание есть

$$(2) \quad \mathbb{E} \sum_{s,t=1}^T w_{st} y_s y_t = \sum_{s,t=1}^T w_{st} \sigma(s-t) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r \sigma(r),$$

где

$$(3) \quad w_r = 2\pi \sum_{t=1}^{T-r} w_{t,t+r} = w_{-r}, \quad r = 0, 1, \dots, T-1,$$

и является, таким образом, линейной функцией теоретических ковариаций $\sigma(0), \sigma(1), \dots, \sigma(T-1)$, которые можно оценить по наблюдениям $y_1, \dots, y_T$. Подставляя в (2) вместо $\sigma(r)$ их выражения через спектральную плотность, представим это математическое ожидание в виде взвешенного среднего спектральной плотности:

$$(4) \quad \mathbb{E} \sum_{s,t=1}^T w_{st} y_s y_t = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r \int_{-\pi}^{\pi} \cos \lambda r f(\lambda) d\lambda = \\ = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{r=-(T-1)}^{T-1} w_r \cos \lambda r f(\lambda) d\lambda = \int_{-\pi}^{\pi} w(\lambda) f(\lambda) d\lambda.$$

Здесь

$$(5) \quad w(\lambda) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r \cos \lambda r = \sum_{s,t=1}^T w_{st} \cos \lambda(s-t).$$

Ясно, что если интересоваться лишь математическими ожиданиями, то оценку (1) без всякого ущерба можно заменить линейной комбинацией выборочных ковариаций C_r ,

$$(6) \quad W = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r C_r,$$

где

$$(7) \quad C_r = \frac{1}{T-r} \sum_{t=1}^{T-r} y_t y_{t+r} = C_{-r}, \quad r = 0, 1, \dots, T-1.$$

Что касается дисперсий, то дисперсия квадратичной функции (1), удовлетворяющей соотношениям (3) для $w_0, w_1, \dots, w_{T-1}$, может оказаться и меньше, и больше, и равной дисперсии (6). (По поводу примеров таких различных случаев см. упр. 5 и 6 гл. 8.) Есть, однако, основания полагать, что зачастую дисперсия оценки (6) будет меньше дисперсий других оценок вида (1), удовлетворяющих (3).

Впрочем, вопрос о том, когда это будет действительно так, остается открытым. Но как бы то ни было оценки (6) сравнительно легко вычисляются, поскольку для этого нужны лишь значения $C_0, C_1, \dots, C_{T-1}$.

В случае больших выборок можно доказать, что оценки (6) ничуть не хуже любых других оценок. Если взять последовательность оценок $\sum_{s, t=1}^T w_{st}^T y_s y_t$, $T = 1, 2, \dots$, то последовательность оценок $(2\pi)^{-1} \sum_{r=-(T-1)}^{T-1} w_r^T C_r$, где $w_r^T = w_{-r}^T = 2\pi \sum_{t=1}^{T-r} w_{t, t+r}^T$, будет иметь ту же последовательность математических ожиданий, что и первая последовательность, для любого стационарного в широком смысле случайного процесса. Гренандер и Розенблатт (1957, разд. 4.2), показали, что если

$$(8) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{s, t=1}^T w_{st}^T y_s y_t = f(v)$$

для некоторого v и всех спектральных плотностей $f(\lambda)$ из некоторого класса (такого, например, как класс всех линейных процессов), то вторая последовательность имеет в пределе дисперсию, не превосходящую предельной дисперсии первой последовательности,

для всякого линейного процесса $y_t = \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, такого, что

(а) первые четыре момента $\{v_t\}$ конечны и являются моментами некоторой стационарной последовательности независимых случайных величин, (б) для некоторого $\delta > 0$ выполняются

$$(9) \quad \lim_{s \rightarrow \infty} s^{2+\delta} \gamma_s = 0, \quad \lim_{s \rightarrow -\infty} |s|^{2+\delta} \gamma_s = 0,$$

и (с) $f(\lambda) > 0$ для всех λ , $-\pi \leq \lambda \leq \pi$. (Мы не будем доказывать этот результат.) Оценки, основанные на выборочных ковариациях, имеют, по крайней мере асимптотически, наименьшую дисперсию. Поэтому всюду в дальнейшем мы будем рассматривать только оценки, являющиеся линейными комбинациями выборочных ковариаций.

Эти оценки могут быть выражены также с помощью выборочной спектральной плотности, поскольку (теорема 8.2.2)

$$(10) \quad \int_{-\pi}^{\pi} \cos \lambda r I(\lambda) d\lambda = \begin{cases} c_r = \frac{T-|r|}{T} C_r, & r = 0, \pm 1, \dots, \pm(T-1), \\ 0, & r = \pm T, \pm(T+1), \dots \end{cases}$$

Используя (10), имеем

$$(11) \quad W = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r C_r = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r \frac{T}{T-|r|} c_r = \\ = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r^* c_r,$$

где

$$(12) \quad w_r^* = \frac{T}{T-|r|} w_r, \quad r = 0, \pm 1, \dots, \pm (T-1).$$

Отсюда получаем

$$(13) \quad W = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r^* \int_{-\pi}^{\pi} \cos \lambda r I(\lambda) d\lambda = \\ = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{r=-(T-1)}^{T-1} w_r^* \cos \lambda r I(\lambda) d\lambda = \int_{-\pi}^{\pi} w^*(\lambda) I(\lambda) d\lambda,$$

где

$$(14) \quad w^*(\lambda) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r^* \cos \lambda r = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r \frac{T}{T-|r|} \cos \lambda r.$$

Отметим, что $w^*(\lambda) \equiv w(\lambda)$ только тогда, когда $w_r = w_r^* = 0$, $r \neq 0$. Весовая функция $w^*(\lambda)$ более удобна для определения указанной оценки как взвешенного среднего выборочной спектральной плотности, а весовая функция $w(\lambda)$ — для представления математического ожидания этой оценки в виде взвешенного среднего теоретической спектральной плотности.

Дисперсию оценок, подобных (11), можно находить, используя результаты § 8.2. Пусть $G = (2\pi)^{-1} \sum_{r=-(T-1)}^{T-1} g_r C_r$, какая-то другая оценка такого рода. Тогда

$$(15) \quad \text{Cov}(G, W) = \frac{1}{(2\pi)^2} \text{Cov} \left(g_0 C_0 + 2 \sum_{r=1}^{T-1} g_r C_r, w_0 C_0 + 2 \sum_{s=1}^{T-1} w_s C_s \right) = \\ = \frac{1}{(2\pi)^2} \left[g_0 w_0 \text{Var } C_0 + 2 w_0 \sum_{r=1}^{T-1} g_r \text{Cov}(C_r, C_0) + \right. \\ \left. + 2 g_0 \sum_{s=1}^{T-1} w_s \text{Cov}(C_0, C_s) + 4 \sum_{r,s=1}^{T-1} g_r w_s \text{Cov}(C_r, C_s) \right].$$

Если процесс гауссовский, то

$$\begin{aligned}
 (16) \quad \text{Cov}(G, W) &= \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left\{ 2g_0 w_0 \frac{\sin^2 \frac{1}{2} (v - v') T}{T^2 \sin^2 \frac{1}{2} (v - v')} + \right. \\
 &+ 2w_0 \sum_{r=1}^{T-1} g_r \frac{\sin \frac{1}{2} (v - v') (T - r) \sin \frac{1}{2} (v - v') T}{(T - r) T \sin^2 \frac{1}{2} (v - v')} 2e^{-i(v+v')r/2} + \\
 &+ 2g_0 \sum_{s=1}^{T-1} w_s \frac{\sin \frac{1}{2} (v - v') T \sin \frac{1}{2} (v - v') (T - s)}{T (T - s) \sin^2 \frac{1}{2} (v - v')} \times \\
 &\quad \times [e^{i(v+v')s/2} + e^{-i(v+v')s/2}] + \\
 &+ 4 \sum_{r,s=1}^{T-1} g_r w_s \frac{\sin \frac{1}{2} (v - v') (T - r) \sin \frac{1}{2} (v - v') (T - s)}{(T - r) (T - s) \sin^2 \frac{1}{2} (v - v')} \times \\
 &\quad \times [e^{i(v+v')(s-r)/2} + e^{-i(v+v')(s+r)/2}] \Big\} f(v) f(v') dv dv' = \\
 &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} g^*(\lambda) w^*(\lambda') \left\{ \left[\int_{-\pi}^{\pi} k_T(v + \lambda, v - \lambda') f(v) dv \right]^2 + \right. \\
 &\quad \left. + \left[\int_{-\pi}^{\pi} k_T(v + \lambda, v + \lambda') f(v) dv \right]^2 \right\} d\lambda d\lambda' = \\
 &= 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} g^*(\lambda) w^*(\lambda') k_T(v + \lambda, v - \lambda') \times \\
 &\quad \times k_T(v' + \lambda, v' - \lambda') f(v) f(v') d\lambda d\lambda' dv dv' = \\
 &= 2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} G_T^*(v, v') W_T^*(v, v') f(v) f(v') dv dv',
 \end{aligned}$$

где

$$(17) \quad G_T^*(v, v') = \int_{-\pi}^{\pi} g^*(\lambda) k_T(v + \lambda, v' + \lambda) d\lambda,$$

$$(18) \quad W_T^*(v, v') = \int_{-\pi}^{\pi} w^*(\lambda') k_T(v + \lambda', v' + \lambda') d\lambda',$$

а $k_T(\lambda, \nu)$ определяется соотношением (73) § 8.2. (Первое выражение для $\text{Cov}(G, W)$ в (16) получается из теоремы 8.2.6, а второе — из теоремы 8.2.8 и соотношения (13) настоящего раздела.) Негауссовская компонента $\text{Cov}(G, W)$ имеет вид

$$\begin{aligned}
 (19) \quad & \frac{1}{(2\pi)^2} \left\{ \frac{g_0 w_0}{T^2} \sum_{t,p=1}^T \kappa(0, p-t, p-t) + \right. \\
 & + 2 \frac{w_0}{T} \sum_{r=1}^{T-1} \frac{g_r}{T-r} \sum_{t=1}^{T-1} \sum_{p=1}^1 \kappa(r, p-t, p-t) + \\
 & + 2 \frac{g_0}{T} \sum_{s=1}^{T-1} \frac{w_s}{T-s} \sum_{t=1}^1 \sum_{p=1}^{T-s} \kappa(0, p-t, p-t+s) + \\
 & + 4 \sum_{r,s=1}^{T-1} \frac{g_r w_s}{(T-r)(T-s)} \times \\
 & \times \left. \sum_{t=1}^{T-r} \sum_{p=1}^{T-s} \kappa(r, p-t, p-t+s) \right\} = \\
 & = \frac{1}{4\pi^2 T^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} g^*(\lambda) w^*(\lambda') \times \\
 & \times \sum_{t,s,t',s'=1}^T \cos \lambda(t-s) \cos \lambda'(t'-s') \times \\
 & \times \kappa(s-t, t'-t, s'-t) d\lambda d\lambda'.
 \end{aligned}$$

9.2.2. Особе ности оценивания спектральной плотности

Если в качестве параметра положения брать математическое ожидание некоторой статистики и если интересоваться оценкой значения теоретической спектральной плотности $f(\lambda)$, в точке $\lambda = \nu$, то следовало бы использовать такую линейную комбинацию W выборочных ковариаций, математическое ожидание которой, являющееся одновременно взвешенным средним выборочной спектральной плотности, было бы по возможности более близким к $f(\nu)$. Свойства этого математического ожидания можно рассматривать либо с помощью коэффициентов $\{w_r\}$, либо с помощью весовой функции $w(\lambda)$. Если мы оцениваем значение $f(\nu)$, то при этом желательно, чтобы функция $w(\lambda)$ имела пик в точке $\lambda = \nu$. В то же время мы заинтересованы в том, чтобы дисперсия оценки была по возможности меньшей. Эти два условия в известном смысле

противоречивы. В § 9.3 будет рассмотрено асимптотическое смещение и дисперсия последовательности такого рода оценок.

Используемая нами оценка является линейной комбинацией выборочных ковариаций. Число их в оценке $I(v)$ равно T . При увеличении T дисперсия каждой выборочной ковариации уменьшается. Однако этот эффект погашается возрастанием числа входящих в оценку выборочных ковариаций. Мы будем изучать оценки, в которых число выборочных ковариаций растет не слишком быстро и в то же время дисперсия сравнительно велика по сравнению со средними значениями.

Указанные оценки можно рассматривать также как взвешенные средние выборочной спектральной плотности. Поскольку значения выборочной спектральной плотности в различных точках асимптотически некоррелированы, то следует ожидать, что взвешенное среднее этих значений в различных точках будет иметь малую асимптотическую дисперсию. Для уменьшения асимптотической дисперсии последовательность весовых функций $w^*(\lambda)$ не должна в точке, где производится оценка спектральной плотности, иметь слишком быстро возрастающий пик. Это соответствует случаю, когда число косинус-функций, входящих в весовую функцию, растет с ростом T медленно.

Если оценивать $f(0)$ величиной

$$(20) \quad \hat{f}(0) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r C_r,$$

($w_r = w_{-r}$), то соответствующей оценкой для $f(v)$ будет

$$(21) \quad \hat{f}(v) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r \cos vr C_r.$$

Весовыми функциями для $\hat{f}(0)$ являются

$$(22) \quad w(\lambda|0) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r \cos \lambda r,$$

$$(23) \quad w^*(\lambda|0) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r^* \cos \lambda r,$$

а для $\hat{f}(v)$ —

$$(24) \quad \begin{aligned} w(\lambda|v) &= \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r \cos \lambda r \cos vr = \\ &= \frac{1}{2} [w(\lambda - v|0) + w(\lambda + v|0)], \end{aligned}$$

$$\begin{aligned}
 (25) \quad w^*(\lambda | \nu) &= \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_r^* \cos \lambda r \cos \nu r = \\
 &= \frac{1}{2} [w^*(\lambda - \nu | 0) + w^*(\lambda + \nu | 0)].
 \end{aligned}$$

Таким образом, задание оценки для $\hat{f}(0)$ приводит к соответствующей оценке функции $f(\nu)$ для каждого ν . Отметим, что $\hat{f}(\nu) = \hat{f}(-\nu)$, $w(\lambda | \nu) = w(\lambda | -\nu)$ и $w^*(\lambda | \nu) = w^*(\lambda | -\nu)$. Поскольку $I(\lambda)$, $f(\lambda)$, $w(\lambda | \nu)$ и $w^*(\lambda | \nu)$ для всех ν являются четными периодическими функциями с периодом 2π , то

$$(26) \quad \int_{-\pi}^{\pi} w(\lambda | \nu) f(\lambda) d\lambda = \int_{-\pi}^{\pi} w(\lambda - \nu | 0) f(\lambda) d\lambda,$$

$$(27) \quad \int_{-\pi}^{\pi} w^*(\lambda | \nu) I(\lambda) d\lambda = \int_{-\pi}^{\pi} w^*(\lambda - \nu | 0) I(\lambda) d\lambda.$$

Мы обычно будем требовать, чтобы для каждого ν выполнялись условия $\int_{-\pi}^{\pi} w(\lambda | \nu) d\lambda = 1$ и $\int_{-\pi}^{\pi} w^*(\lambda | \nu) d\lambda = 1$. Отсюда, в частности, следует что $w_0 = w_0^* = 1$.

Функции $w(\lambda | \nu)$ и $w^*(\lambda | \nu)$ называются *окнами*, поскольку в некотором смысле они определяют части $f(\lambda)$ и $I(\lambda)$, которые «просматриваются» математическим ожиданием оценки и самой оценкой соответственно. Для весовой функции типично наличие одного большого пика с центром в точке ν , окруженного более мелкими лепестками.

9.2.3. Примеры оценок спектральной плотности

Как уже отмечалось, в качестве оценок спектральной плотности можно брать линейные комбинации выборочных ковариаций $C_0, C_1, \dots, C_{T-1}$. Это приводит к спектральным окнам, являющимся полиномами степени не выше $T - 1$. Ниже будут рассмотрены различные варианты подобных оценок. Во многие ядра будут входить функции $k_T(\lambda) = \sin^2 \frac{1}{2} \lambda T / \left[2\pi T \sin^2 \frac{1}{2} \lambda \right]$ и $h_{2T-1}(\lambda) = \sin \frac{1}{2} \lambda (2T - 1) / \left[2\pi \sin \frac{1}{2} \lambda \right]$. Мы рекомендуем читателю в каждом из рассматриваемых ниже случаев строить графики последовательностей $\{w_r\}$, $\{w_r^*\}$ и функций $w(\lambda | 0)$, $w^*(\lambda | 0)$.

А. Выборочная спектральная плотность. Выборочная спектральная плотность имеет указанный выше вид с $w_r = 1 - |r|/T$ для

$r = 0, \pm 1, \dots, \pm (T - 1)$ и $w_r^* = 1$ для $r = 0, \pm 1, \dots, \pm (T - 1)$.
Окна здесь таковы:

$$(28) \quad w(\lambda | 0) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) \cos \lambda r =$$

$$= k_T(\lambda) = \frac{\sin^2 \frac{1}{2} \lambda T}{2\pi T \sin^2 \frac{1}{2} \lambda},$$

$$(29) \quad w^*(\lambda | 0) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} \cos \lambda r =$$

$$= h_{2T-1}(\lambda) = \frac{\sin \frac{1}{2} \lambda (2T - 1)}{2\pi \sin \frac{1}{2} \lambda}.$$

Отметим, что $w^*(\lambda | 0)$ не совпадает с дельта-функцией Дирака. Дело в том, что $I(v) = \hat{f}(v)$ является линейной комбинацией толь-

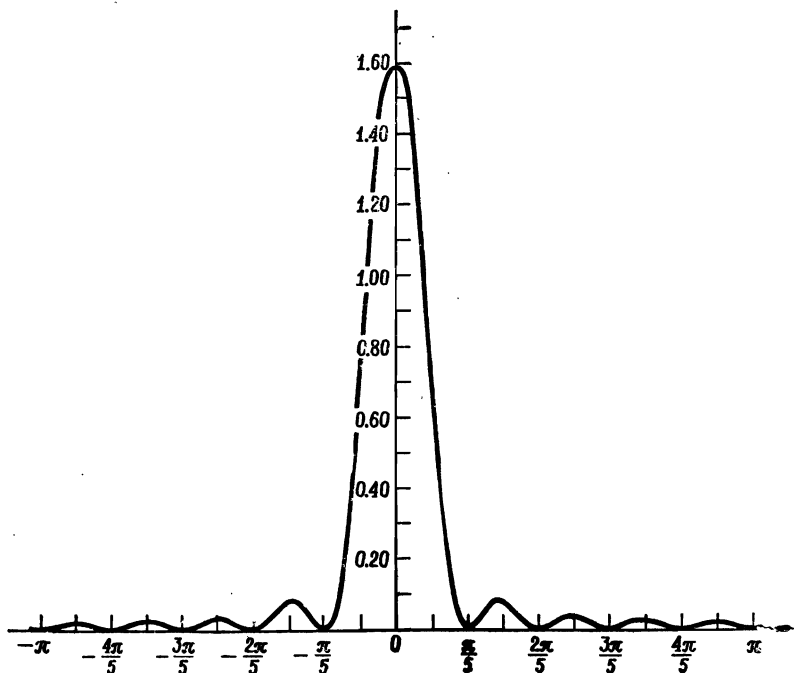


Рис. 9.1.

Функция $k_T(\lambda) = \sin^2(\lambda/2)T/[2\pi T \sin^2(\lambda/2)]$ для $T = 10$.

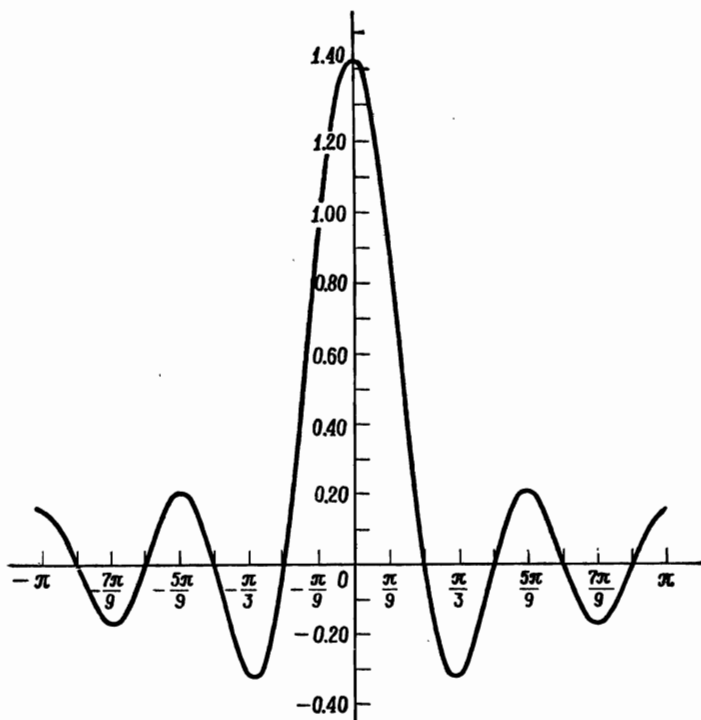


Рис. 9.2.

Функция $h_{2T-1}(\lambda) = \sin(\lambda/2) (2T-1) / [2\pi \sin(\lambda/2)]$ для $T=5$.

ко T косинус-функций.] Окно $w(\lambda | 0) = k_T(\lambda)$ неотрицательно и имеет максимум, равный $T/(2\pi)$, в точке $\lambda = 0$ и минимум, равный 0, в точках $\lambda = \pm k2\pi/T$, $k = 1, \dots, (T-1)/2$ или $T/2$. Между его нулями, в точках, являющихся корнями ($\lambda \neq 0$) уравнения $\lambda T/2 = T \operatorname{tg}(\lambda/2)$, расположены локальные максимумы. Указанные корни приближенно равны $\pm (2k+1)\pi/T$ для $k = 1, \dots, (T-2)/2$ или $(T-1)/2$. (Если T нечетное, то локальные максимумы расположены в точках $\lambda = \pm\pi$.) Окно $w^*(\lambda | 0) = \sin(\lambda/2) (2T-1) / [2\pi \sin(\lambda/2)]$ имеет максимум $(2T-1)/(2\pi)$ в точке $\lambda = 0$ и нули в точках $\lambda = \pm k2\pi/(2T-1)$ для $k = 1, \dots, T-1$. Его локальные максимумы и локальные минимумы располагаются поочередно между нулями, в точках, являющихся корнями ($\lambda \neq 0$) уравнения $\lambda (2T-1)/2 = (2T-1) \operatorname{tg}(\lambda/2)$. (Если T нечетное, то в точках $\lambda = \pm\pi$ будет относительный максимум; при четном T в этих точках будет относительный минимум.) Графики этих окон представлены на рис. 9.1 и 9.2.

Окна, представленные на рис. 9.1 и 9.2, обладают чертами, характерными для весовых функций. У них имеется ярко выраженный пик с центром в нуле и боковые лепестки, высота которых уменьшается по мере того, как их центры удаляются от нуля. Если ядро неотрицательное, такое, как $k_T(\lambda)$, то это обеспечивает неотрицательность взвешенного с этим ядром интеграла от неотрицательной функции [т. е. $f(\lambda)$ или $I(\lambda)$]. При ядре, принимающем и положительные и отрицательные значения, такое, как представленное на рис. 9.2, взвешенный интеграл может быть и отрицательным. Получающиеся при этом оценки и математические ожидания оценок следует, конечно, считать неудовлетворительными, поскольку сами $f(\lambda)$ и $I(\lambda)$ неотрицательны. Ввиду того что $w_0 = w_0^* =$

$$= 1, \text{ выполняются равенства } \int_{-\pi}^{\pi} w(\lambda | 0) d\lambda = \int_{-\pi}^{\pi} k_T(\lambda) d\lambda = 1 \text{ и } \int_{-\pi}^{\pi} w^*(\lambda | 0) d\lambda = \int_{-\pi}^{\pi} h_{2T-1}(\lambda) d\lambda = 1.$$

В. Усеченная оценка. Усечение рядов, входящих в выражение выборочной спектральной плотности, дает оценку в виде

$$(30) \quad \hat{f}(v) = \frac{1}{2\pi} \sum_{r=-K}^K \cos vr c_r = \frac{1}{2\pi} \sum_{r=-K}^K \left(1 - \frac{|r|}{T}\right) \cos vr C_r,$$

где K — некоторое целое число, $K \leq T-1$. Выборочные ковариации $C_{K+1}, \dots, C_{T-1}$ в этой оценке опускаются. Дело в том, что последние основываются на относительно малом числе произведений и вносят сравнительно небольшую информацию. Соответствующими коэффициентами здесь являются $w_r = 1 - |r|/T$ для $r = 0, \pm 1, \dots, \pm K$ и $w_r = 0$ для $r = \pm(K+1), \dots, \pm(T-1)$. Остальные коэффициенты таковы: $w^* = 1$ для $r = 0, \pm 1, \dots, \pm K$ и $w^* = 0$ для $r = \pm(K+1), \dots, \pm(T-1)$. Окна имеют вид

$$(31) \quad w(\lambda | 0) = \frac{K+1}{T} k_{K+1}(\lambda) + \left(1 - \frac{K+1}{T}\right) h_{2K+1}(\lambda),$$

$$(32) \quad w^*(\lambda | 0) = h_{2K+1}(\lambda).$$

Максимум окна $w(\lambda | 0)$ расположен в точке $\lambda = 0$ и равен

$$(33) \quad \frac{2K+1}{2\pi} - \frac{K(K+1)}{2\pi T}.$$

[Для $K = T/2$ последняя величина равна, например, $(3/4 T + 1/2)/(2\pi)$.] Это окно принимает и положительные и отрицательные значения. Окно $w^*(\lambda | 0)$ принимает максимальное значение $(2K+1)/(2\pi)$ в точке $\lambda = 0$ и имеет нули в точках $\lambda = \pm k2\pi/(2K+1)$, $k = 1, \dots, K$. Его относительные минимумы и максимумы расположены между этими нулями. Чем меньше K , тем меньше по величине

лики окон $w(\lambda | 0)$ и $w^*(\lambda | 0)$ и тем дальше отстоят нули от начала координат.

С. Оценка Бартлетта¹⁾. Ошибку, связанную с отбрасыванием вклада выборочных ковариаций высших порядков при усечении ряда, можно уменьшить, вводя некоторый демпфирующий (затухающий) множитель. С этой целью, рассматривают оценку

$$\begin{aligned} (34) \quad \hat{f}(v) &= \frac{1}{2\pi} \sum_{r=-K}^K \left(1 - \frac{|r|}{K}\right) \cos vr C_r = \\ &= \frac{1}{2\pi} \sum_{r=-K}^K \frac{1 - \frac{|r|}{K}}{1 - \frac{|r|}{T}} \cos vr c_r. \end{aligned}$$

Коэффициентами w_r здесь являются $w_r = 1 - |r|/K$ для $r = 0, \pm 1, \dots, \pm K$ и $w_r = 0$ для $r = \pm(K+1), \dots, \pm(T-1)$ в отличие от $1 - |r|/T$ и 0 соответственно в случае простого усечения. Коэффициенты w_r^* равны $(1 - |r|/K)/(1 - |r|/T)$ для $r = 0, \pm 1, \dots, \pm K$ и 0 для $r = \pm(K+1), \dots, \pm(T-1)$ в отличие от коэффициентов, равных 1 и 0 соответственно, для простого усечения. Несмещенные оценки $\{C_r\}$ линейно демпфируются до нулевого значения при $r = \pm K$. Соответствующие окна имеют вид

$$(35) \quad w(\lambda | 0) = k_K(\lambda) = \frac{\sin^2(K\lambda/2)}{2\pi K \sin^2(\lambda/2)},$$

$$\begin{aligned} (36) \quad w^*(\lambda | 0) &= \frac{1}{2\pi} \frac{T}{K} \sum_{r=-K}^K \frac{K - |r|}{T - |r|} e^{i\lambda r} = \\ &= \frac{1}{2\pi} \frac{T}{K} \sum_{r=-K}^K \left(1 + \frac{K - T}{T - |r|}\right) e^{i\lambda r} = \\ &= \frac{T}{K} \left\{ h_{2K+1}(\lambda) - \frac{T - K}{2\pi} \sum_{r=-K}^K \frac{\cos \lambda r}{T - |r|} \right\}, \end{aligned}$$

причем последнее выражение нельзя упростить. (См. упр. 7.) Некоторые свойства окна $w(\lambda | 0)$ описаны с помощью эквивалентного ему окна $w^*(\lambda | 0)$ в примере D.

Бартлетт (1950) заметил, что вариабельность оценок спектральной плотности можно уменьшить, если разбить наблюдаемый временной ряд $y_1, \dots, y_T$ на m равных отрезков длины K ($T = mK$) и вычислить выборочную спектральную плотность на каждом из них.

¹⁾ Некоторые авторы не различают пример С, приведенный Бартлеттом (1950), и пример D, являющийся модификацией, более простой для исследования.

Усредняя полученные результаты, приходим к следующей оценке:

$$\begin{aligned}
 (37) \quad & \frac{1}{2\pi} \frac{1}{m} \sum_{g=1}^m \left\{ \frac{1}{K} \sum_{t=1}^K y_{(g-1)K+t}^2 + \right. \\
 & \left. + 2 \sum_{r=1}^{K-1} \cos vr \frac{1}{K} \sum_{t=1}^{K-r} y_{(g-1)K+t} y_{(g-1)K+t+r} \right\} = \\
 & = \frac{1}{2\pi} \frac{1}{T} \sum_{t=1}^T y_t^2 + \frac{2}{2\pi} \sum_{r=1}^{K-1} \left(1 - \frac{r}{K}\right) \times \\
 & \times \cos vr \sum_{g=1}^m \sum_{t=1}^{K-r} y_{(g-1)K+t} y_{(g-1)K+t+r} / [m(K-r)].
 \end{aligned}$$

Каждая двойная сумма в (37) является несмещенной оценкой соответствующего $\sigma(r)$. При этом не используется несколько произведений $y_t y_{t+r}$. Если все такие несмещенные оценки заменить на соответствующие C_r , то в результате получится (34).

Д. Модифицированная оценка Бартлетта. Если в предыдущую оценку ввести еще один сглаживающий множитель $1 - |r|/T$, то в результате получим оценку

$$\begin{aligned}
 (38) \quad \hat{f}(v) &= \frac{1}{2\pi} \sum_{r=-K}^K \left(1 - \frac{|r|}{K}\right) \left(1 - \frac{|r|}{T}\right) \cos vr C_r = \\
 &= \frac{1}{2\pi} \sum_{r=-K}^K \left(1 - \frac{|r|}{K}\right) \cos vr c_r.
 \end{aligned}$$

Коэффициенты w_r равны в этом случае $(1 - |r|/K)(1 - |r|/T)$ для $r = 0, \pm 1, \dots, \pm K$ и нулю для $r = \pm(K+1), \dots, \pm(T-1)$, тогда как в обычной оценке Бартлетта они равны соответственно $1 - |r|/K$ и 0. Коэффициенты w_r равны $1 - |r|/K$ для $r = 0, \pm 1, \dots, \pm K$ и 0 для $r = \pm(K+1), \dots, \pm(T-1)$. В обычной оценке Бартлетта они были равны соответственно величине $(1 - |r|/K)/(1 - |r|/T)$ и 0. Окна здесь таковы:

$$\begin{aligned}
 (39) \quad w(\lambda|0) &= \frac{1}{2\pi} \sum_{r=-K}^K \left[1 - |r| \left(\frac{1}{K} + \frac{1}{T} \right) + \frac{r^2}{TK} \right] \cos \lambda r = \\
 &= \frac{\cos^2(K\lambda/2)}{2\pi T \sin^2(\lambda/2)} + \frac{\sin^2(K\lambda/2)}{2\pi K \sin^2(\lambda/2)} - \frac{\sin K\lambda \cos(\lambda/2)}{4\pi KT \sin^3(\lambda/2)},
 \end{aligned}$$

$$(40) \quad w^*(\lambda|0) = k_K(\lambda) = \frac{\sin^2(K\lambda/2)}{2\pi K \sin^2(\lambda/2)}.$$

Окно $w^*(\lambda | 0)$ неотрицательно, что обеспечивает неотрицательность оценки $\hat{f}(v)$. Максимум $w^*(\lambda | 0)$ расположен в точке $\lambda = 0$ и равен $K/(2\pi)$. Минимум $w^*(\lambda | 0)$, равный нулю, достигается в точках $\lambda = \pm k2\pi/K$, $k = 1, \dots, (K-1)/2$ или $K/2$. Такая модификация оценки Бартлетта рассматривалась у Гренандера и Розенблатта (1957, стр. 146), а также у Хеннана (1960), которые включали ее в общую асимптотическую теорию.

Е. Оценка Даниэля. Помимо оценивания значения $f(\lambda)$ в одной точке $\lambda = v$, можно попытаться оценивать $f(\lambda)$ на целом отрезке. (См. обсуждение этого вопроса Бартлеттом (1946).) Если таким отрезком является $[v-b, v+b]$, $0 < b < \pi$, то интеграл

$\int_{v-b}^{v+b} f(\lambda) d\lambda / (2b)$ иначе можно записать в виде взвешенного интеграла $\int_{-\pi}^{\pi} g(\lambda - v) f(\lambda) d\lambda$, где

$$(41) \quad g(\lambda) = \begin{cases} \frac{1}{2b}, & -b \leq \lambda \leq b, \\ 0, & b < |\lambda|. \end{cases}$$

(Если v не принадлежит отрезку $[-\pi + b, \pi - b]$, то следует использовать периодичность с периодом 2π функций $f(\lambda)$ и $I(\lambda)$.) Для указанного оценивания можно использовать

$$\begin{aligned} (42) \quad \int_{-\pi}^{\pi} g(\lambda - v) I(\lambda) d\lambda &= \frac{1}{2b} \int_{v-b}^{v+b} I(\lambda) d\lambda = \\ &= \frac{1}{4\pi b} \left[c_0 \int_{v-b}^{v+b} d\lambda + 2 \sum_{r=1}^{T-1} c_r \int_{v-b}^{v+b} \cos \lambda r d\lambda \right] = \\ &= \frac{1}{2\pi} \left[c_0 + 2 \sum_{r=1}^{T-1} \frac{\sin(v+b)r - \sin(v-b)r}{2br} c_r \right] = \\ &= \frac{1}{2\pi} \left[c_0 + 2 \sum_{r=1}^{T-1} \frac{2 \cos vr \sin br}{2br} c_r \right] = \\ &= \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} \frac{\sin br}{br} \cos vr c_r = \\ &= \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} \frac{\sin br}{br} \left(1 - \frac{|r|}{T} \right) \cos vr C_r \end{aligned}$$

где $(\sin br)/br = 1$ для $r = 0$. Входящие сюда коэффициенты являются коэффициентами Фурье функции $g(\lambda)$. Именно,

$$w_0^* = \int_{-\pi}^{\pi} g(\lambda) d\lambda = 1 \text{ и}$$

$$(43) \quad w_r^* = \int_{-\pi}^{\pi} g(\lambda) \cos \lambda r d\lambda = \frac{1}{2b} \int_{-b}^b \cos \lambda r d\lambda = \frac{\sin br}{br}, \quad r \neq 0.$$

[Отметим, что $g(\lambda)$ не идентична $w^*(\lambda|0)$.]

Получающееся в итоге окно $w^*(\lambda|0)$ является линейной комбинацией функций $1, \cos \lambda, \dots, \cos \lambda(T-1)$, наилучшим образом аппроксимирующей $g(\lambda)$ в смысле среднеквадратичной ошибки

$$(44) \quad \int_{-\pi}^{\pi} |g(\lambda) - w^*(\lambda|0)|^2 d\lambda.$$

Аппроксимация функции $g(\lambda)$ функцией $w^*(\lambda|0)$ не может быть одинаково хорошей для всех значений λ . Действительно, она не будет очень хорошей в точках разрыва $g(\lambda)$, т. е. в точках $\lambda = \pm b$. Осцилляция аппроксимирующего гармонического ряда вблизи точек разрыва известна под названием явления Гиббса [Ланцош (1956, гл. 4), и Хемминг (1962, гл. 22)]. Ее можно уменьшить, умножая w_r^* на соответствующие веса.

Ряды (42) могут быть усечены, например, до $K = \lfloor \pi/b \rfloor$; при этом $w_r^* = 0$ для $r = \pm(K+1), \dots, \pm(T-1)$. Такая оценка часто называется прямоугольной. Окно $w(\lambda|0)$ отличается от $g(\lambda)$ наличием множителей $1 - |r|/T$ и использованием лишь конечного числа коэффициентов.

Ф. Общие оценки Блэкмена — Тьюки. Блэкмен и Тьюки (1959) предложили две оценки (названные ими соответственно окнами Хеннинга и Хемминга), являющиеся частными случаями оценки

$$(45) \quad \hat{f}(v) = \frac{1}{2\pi} \sum_{r=-K}^K \left(1 - 2a + 2a \cos \frac{\pi r}{K}\right) \left(1 - \frac{|r|}{T}\right) \cos vr C_r = \\ = \frac{1}{2\pi} \sum_{r=-K}^K \left(1 - 2a + 2a \cos \frac{\pi r}{K}\right) \cos vr c_r,$$

$0 \leq a \leq 1/4$. Коэффициенты $w_r^* = 1 - 2a + 2a \cos(\pi r/K)$, $r = 0, \pm 1, \dots, \pm K$, убывают от значения 1 в точке $r = 0$ до значения $1 - 4a$ в точках $r = \pm K$ и равны нулю для $r = \pm(K+1), \dots, \pm(T-1)$. Оценку (45) можно представить иначе в виде

$$(46) \quad \hat{f}(v) = \frac{1-2a}{2\pi} \sum_{r=-K}^K \cos vr c_r + \frac{2a}{2\pi} \sum_{r=-K}^K \cos \frac{\pi r}{K} \cos vr c_r =$$

$$\begin{aligned}
&= \frac{1-2a}{2\pi} \sum_{r=-K}^K \cos vr c_r + \frac{a}{2\pi} \sum_{r=-K}^K \cos\left(v - \frac{\pi}{K}\right) r c_r + \\
&\quad + \frac{a}{2\pi} \sum_{r=-K}^K \cos\left(v + \frac{\pi}{K}\right) r c_r = \\
&= \hat{a} \hat{f}^K\left(v - \frac{\pi}{K}\right) + (1-2a) \hat{f}^K(v) + a \hat{f}^K\left(v + \frac{\pi}{K}\right),
\end{aligned}$$

где $\hat{f}^K(v)$ — усеченная выборочная спектральная плотность (см. пример В), $\hat{f}^K(\pi + \lambda) = \hat{f}^K(\pi - \lambda)$ и $\hat{f}^K(-\pi - \lambda) = \hat{f}^K(-\pi + \lambda)$. Таким образом, для вычисления $\hat{f}(v)$ следует подсчитать $\hat{f}^K(v)$ и взять затем скользящее среднее (46). Спектральные окна, соответствующие оценке (45), выражаются через спектральные окна, соответствующие усеченной выборочной спектральной плотности и имеют вид

$$\begin{aligned}
(47) \quad w(\lambda|0) &= \frac{a}{2\pi} \sum_{r=-K}^K \left(1 - \frac{|r|}{T}\right) \cos\left(\lambda - \frac{\pi}{K}\right) r + \\
&\quad + \frac{1-2a}{2\pi} \sum_{r=-K}^K \left(1 - \frac{|r|}{T}\right) \cos \lambda r + \\
&\quad + \frac{a}{2\pi} \sum_{r=-K}^K \left(1 - \frac{|r|}{T}\right) \cos\left(\lambda + \frac{\pi}{K}\right) r = \\
&= a \left\{ \frac{K+1}{T} k_{K+1}\left(\lambda - \frac{\pi}{K}\right) + \right. \\
&\quad + \left(1 - \frac{K+1}{T}\right) h_{2K+1}\left(\lambda - \frac{\pi}{K}\right) \Big\} + \\
&\quad + (1-2a) \left\{ \frac{K+1}{T} k_{K+1}(\lambda) + \right. \\
&\quad + \left(1 - \frac{K+1}{T}\right) h_{2K+1}(\lambda) \Big\} + \\
&\quad + a \left\{ \frac{K+1}{T} k_{K+1}\left(\lambda + \frac{\pi}{K}\right) + \right. \\
&\quad + \left(1 - \frac{K+1}{T}\right) h_{2K+1}\left(\lambda + \frac{\pi}{K}\right) \Big\},
\end{aligned}$$

$$\begin{aligned}
 (48) \quad w^*(\lambda | 0) &= \frac{a}{2\pi} \sum_{r=-K}^K \cos\left(\lambda - \frac{\pi}{K}r\right) r + \\
 &+ \frac{1-2a}{2\pi} \sum_{r=-K}^K \cos \lambda r + \frac{a}{2\pi} \sum_{r=-K}^K \cos\left(\lambda + \frac{\pi}{K}r\right) r = \\
 &= ah_{2K+1}\left(\lambda - \frac{\pi}{K}\right) + (1-2a)h_{2K+1}(\lambda) + ah_{2K+1}\left(\lambda + \frac{\pi}{K}\right).
 \end{aligned}$$

Максимум $w^*(\lambda | 0)$ расположен в точке $\lambda = 0$ и равен $[2K+1](1-2a)-2a/(2\pi)$.

Г. Окно Хеннинга. Особый интерес представляет частный случай оценки Блэкмена — Тьюки, отвечающий значению $a = 1/4$, т. е. оценка

$$\begin{aligned}
 (49) \quad \hat{f}(v) &= \frac{1}{2\pi} \sum_{r=-K}^K \left(\frac{1}{2} + \frac{1}{2} \cos \frac{\pi}{K}r\right) \cos vrc, = \\
 &= \frac{1}{4} \hat{f}^K\left(v - \frac{\pi}{K}\right) + \frac{1}{2} \hat{f}^K(v) + \frac{1}{4} \hat{f}^K\left(v + \frac{\pi}{K}\right) = \\
 &= \frac{1}{2} \left[\frac{1}{2} \hat{f}^K\left(v - \frac{\pi}{K}\right) + \frac{1}{2} \hat{f}^K(v) \right] + \\
 &+ \frac{1}{2} \left[\frac{1}{2} \hat{f}^K(v) + \frac{1}{2} \hat{f}^K\left(v + \frac{\pi}{K}\right) \right].
 \end{aligned}$$

Как видно из (49), она получается простым усреднением двух средних усеченных выборочных спектральных плотностей. Веса $w_r \doteq (1 + \cos \pi r/K)/2 = \cos^2 \pi r/(2K)$, $r = 0, \pm 1, \dots, \pm K$, убывают от единицы в точке $r = 0$ до нуля в точках $r = \pm K$. Нули функции $w^*(\lambda | 0)$ расположены в точках $\lambda = \pm k\pi/K$, $k = 2, 3, \dots, K$. Максимум $w^*(\lambda | 0)$ достигается при $\lambda = 0$ и равен $K/(2\pi)$. Окно, соответствующее модифицированной оценке Бартлетта (с тем же значением K), имеет точно такой же максимум. В то же время его нули располагаются в точках $\lambda = \pm k\pi/K$ для $k = 2, 4, \dots, K-1$ или K . Иначе говоря, все нули модифицированного окна Бартлетта являются одновременно и нулями окна Хеннинга, но последнее имеет еще и дополнительные нули, расположенные вблизи относительных максимумов модифицированного окна Бартлетта (за исключением максимума в нуле). Этот факт заставляет предполагать, что боковые лепестки окна Хеннинга будут меньшими по сравнению с боковыми лепестками модифицированного окна Бартлетта.

Н. Окно Хемминга. С целью уменьшения размеров первых лепестков Блэкмен и Тьюки (1959) предложили в оценке (45) брать

$a = 0.23$. При этом значении a (45) принимает вид

$$(50) \quad \hat{f}(v) = \frac{1}{2\pi} \sum_{r=-K}^K \left(0.54 + 0.46 \cos \frac{\pi}{K} r \right) \cos vr c_r = \\ = 0.23 \hat{f}^K \left(v - \frac{\pi}{K} \right) + 0.54 \hat{f}^K(v) + 0.23 \hat{f}^K \left(v + \frac{\pi}{K} \right).$$

1. *Оценка Парзена.* Модифицированную оценку Бартлетта можно рассматривать как частный случай оценок с $w_r^* = 1 - (|r|/K)^q$, соответствующий значению $q = 1$. Парзен (1957b) предложил использовать и другие значения q . Если $q = 2$, то

$$(51) \quad \hat{f}(v) = \frac{1}{2\pi} \sum_{r=-K}^K \left(1 - \frac{r^2}{K^2} \right) \left(1 - \frac{|r|}{T} \right) \cos vr C_r = \\ = \frac{1}{2\pi} \sum_{r=-K}^K \left(1 - \frac{r^2}{K^2} \right) \cos vrc_r,$$

а соответствующее ему окно $w^*(\lambda | 0)$

$$(52) \quad w^*(\lambda | 0) = \frac{1}{2\pi} \sum_{r=-K}^K \left(1 - \frac{r^2}{K^2} \right) \cos \lambda r = \\ = \frac{1}{2\pi} \left\{ \frac{\sin \lambda K \cos (\lambda/2)}{2K^2 \sin^3 (\lambda/2)} - \frac{\cos \lambda K}{K \sin^2 (\lambda/2)} \right\}.$$

Максимум этого окна расположен в точке $\lambda = 0$ и равен $(2\pi)^{-1} [4K/3 - 1/(3K)]$.

2. *Оценка Парзена.* Для четного K Парзенем предложена следующая оценка:

$$(53) \quad \hat{f}(v) = \frac{1}{2\pi} \sum_{r=-K/2}^{K/2} \left(1 - 6 \frac{r^2}{K^2} + 6 \frac{|r|^3}{K^3} \right) \left(1 - \frac{|r|}{T} \right) \cos vr C_r + \\ + \frac{2}{2\pi} \sum_{r=-K}^{-K/2-1} \left(1 - \frac{|r|}{K} \right)^3 \left(1 - \frac{|r|}{T} \right) \cos vr C_r + \\ + \frac{2}{2\pi} \sum_{r=K/2+1}^K \left(1 - \frac{r}{K} \right)^3 \left(1 - \frac{r}{T} \right) \cos vrc_r = \\ = \frac{1}{2\pi} \sum_{r=-K/2}^{K/2} \left(1 - 6 \frac{r^2}{K^2} + 6 \frac{|r|^3}{K^3} \right) \cos vrc_r + \\ + \frac{2}{2\pi} \sum_{r=-K}^{-K/2-1} \left(1 - \frac{|r|}{K} \right)^3 \cos vrc_r +$$

$$\begin{aligned}
& + \frac{2}{2\pi} \sum_{r=K/2+1}^K \left(1 - \frac{r}{K}\right)^3 \cos \nu r c_r = \\
& = \frac{2}{2\pi} \sum_{r=-K}^K \left(1 - \frac{|r|}{K}\right)^3 \cos \nu r c_r - \\
& - \frac{1}{2\pi} \sum_{r=-K/2}^{K/2} \left(1 - 2 \frac{|r|}{K}\right)^3 \cos \nu r c_r.
\end{aligned}$$

Функция, равная $1 - 6y^2 + 6|y|^3$ для $|y| \leq 1/2$, $2(1 - |y|)^3$ для $1/2 \leq |y| \leq 1$ и 0 для $1 \leq |y|$, пропорциональна плотности распределения среднего четырех независимых случайных величин, равномерно распределенных на отрезке от -1 до 1 . Функция, равная $1 - |y|$ для $|y| \leq 1$ и 0 для $1 \leq |y|$ (связанная с модифицированной оценкой Бартлетта), пропорциональна плотности распределения среднего двух таких случайных величин. (См. упр. 9.) Но характеристической функцией указанного равномерного распределения является $(\sin t)/t$. Поэтому окно для выборочной спектральной плотности в случае оценки Парзена (53) соответствует характеристической функции $[(\sin t)/t]^4$, а в случае модифицированной оценки Бартлетта — функции $[(\sin t)/t]^3$.

Полагая $K = 2M$ и используя результат упр. 12, найдем окно Парзена из соотношения

$$\begin{aligned}
(54) \quad 2\pi\omega^*(\lambda|0) &= 2 \sum_{r=-2M}^{2M} \left(1 - \frac{|r|}{2M}\right)^3 \cos \lambda r - \\
&- \sum_{r=-M}^M \left(1 - \frac{|r|}{M}\right)^3 \cos \lambda r = \\
&= 1 + \frac{1}{2M^3} \sum_{r=1}^{2M} (2M-r)^3 \cos \lambda r - \\
&- \frac{2}{M^3} \sum_{r=1}^M (M-r)^3 \cos \lambda r = \\
&= 1 + \frac{1}{2M^3} \left\{ -\frac{8M^3}{2} + \frac{2 \sin^2 \lambda M + 12M^3}{4 \sin^2 (\lambda/2)} - \right. \\
&- \frac{3 \sin^2 \lambda M}{4 \sin^4 (\lambda/2)} \left. \right\} - \frac{2}{M^3} \left\{ -\frac{M^3}{2} + \right. \\
&+ \frac{2 \sin^2 (M\lambda/2) + 3M^3}{4 \sin^2 (\lambda/2)} - \frac{3 \sin^2 (M\lambda/2)}{4 \sin^4 (\lambda/2)} \left. \right\} =
\end{aligned}$$

$$= \frac{1}{M^3} \left\{ \frac{\sin^3 M\lambda - 4 \sin^2 (M\lambda/2)}{4 \sin^2 (\lambda/2)} + \right. \\ \left. + \frac{12 \sin^2 (M\lambda/2) - 3 \sin^2 M\lambda}{8 \sin^2 (\lambda/2)} \right\}.$$

Замечая, что

$$(55) \quad \sin^3 M\lambda = (2 \sin (M\lambda/2) \cos (M\lambda/2))^2 = \\ = 4 \sin^2 (M\lambda/2) (1 - \sin^2 (M\lambda/2)),$$

получаем из (54):

$$(56) \quad w^*(\lambda | 0) = \frac{1}{2\pi M^3} \left\{ \frac{3}{2} \left(\frac{\sin (M\lambda/2)}{\sin (\lambda/2)} \right)^4 - \frac{\sin^4 (M\lambda/2)}{\sin^2 (\lambda/2)} \right\} = \\ = \frac{2\pi}{M} \left\{ \frac{3}{2} k_M^2(\lambda) - \frac{\sin^2 (M\lambda/2)}{M^2} h_M^2(\lambda) \right\} = \\ = \frac{8}{2\pi K^2} \left\{ \frac{3}{2} \frac{\sin^4 (K\lambda/4)}{\sin^2 (\lambda/2)} - \frac{\sin^4 (K\lambda/4)}{\sin^2 (\lambda/2)} \right\} = \\ = \frac{2\pi}{K} \left\{ 3k_{K/2}^2(\lambda) - 8 \frac{\sin^2 (K\lambda/4)}{K^2} h_{K/2}^2(\lambda) \right\}.$$

Для больших K и малых λ первый член в (56) существенно больше остальных. Окно $w^*(\lambda | 0)$ неотрицательно; поэтому неотрицательна и оценка $\hat{f}(v)$.

К. Усреднение по дискретным значениям λ . В ряде случаев $I(\lambda)$ вычисляется для $\lambda = 2\pi j/T$, $j = 0, 1, \dots, [T/2]$. В частности, такой подсчет производится в случае, когда T велико и используется быстрое преобразование Фурье (разд. 4.3.5). Следующая оценка тесно связана с оценкой Даниэля (пример Е), но основывается на значениях $I(2\pi j/T)$, $j = 0, 1, \dots, [T/2]$:

$$(57) \quad \hat{f}\left(\frac{2\pi k}{T}\right) = \frac{1}{2n+1} \sum_{j=k-n}^{k+n} I\left(\frac{2\pi j}{T}\right) = \\ = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} c_r \frac{1}{2n+1} \sum_{j=k-n}^{k+n} e^{i2\pi jr/T} = \\ = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} c_r e^{i2\pi kr/T} \frac{1}{2n+1} \sum_{h=-n}^n e^{i2\pi hr/T} = \\ = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} \frac{2\pi}{2n+1} h_{2n+1}\left(\frac{2\pi r}{T}\right) \cos \frac{2\pi kr}{T} c_r,$$

Для значений $v = 2\pi k/T$, $k = 0, 1, \dots, [T/2]$, эта оценка имеет веса

$$(58) \quad \omega_r^* = \frac{2\pi}{2n+1} h_{2n+1} \left(\frac{2\pi r}{T} \right) = \\ = \frac{1}{2n+1} \frac{\sin \pi (2n+1) r/T}{\sin \pi r/T}, \quad r = 0, \pm 1, \dots, \pm (T-1),$$

совпадающие с весами Даниэля (пример Е), в которых величины $br = \pi (2n+1) r/T$ в знаменателе (43) заменены величинами $(2n+1) \sin \pi r/T$.

9.3. АСИМПТОТИЧЕСКИЕ СРЕДНИЕ И КОВАРИАЦИИ ОЦЕНОК СПЕКТРАЛЬНОЙ ПЛОТНОСТИ

9.3.1. Оценивание взвешенных средних от спектральной плотности

В § 9.2 были исследованы математические ожидания оценок спектральных плотностей. Из соотношений (16) и (19) § 9.2 путем соответствующих подстановок могут быть найдены также и вторые моменты этих оценок, т. е. их дисперсии и ковариации. Однако вторые моменты оказывается весьма трудно интерпретировать. Поэтому вместо этих моментов мы рассмотрим предельные значения (при $T \rightarrow \infty$) соответствующим образом нормированных оценок. Рассмотрим последовательность оценок

$$(1) \quad \hat{f}_T(v) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} \omega_r^* \cos vr \, c_{rT},$$

где $\{\omega_r^*\}$ — некоторая фиксированная последовательность ($\omega_r^* = \omega_{-r}^*$), удовлетворяющая условию

$$(2) \quad \sum_{r=-\infty}^{\infty} |\omega_r^*| < \infty.$$

Мы пишем здесь c_{rT} вместо c_r , указывая тем самым на то, что используется ровно T наблюдений. Пусть

$$(3) \quad w^*(\lambda | v) = \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} \omega_r^* \cos \lambda r \cos vr.$$

Тогда (в силу леммы 8.3.1) $w^*(\lambda | v) = w(\lambda | v)$, где

$$(4) \quad w(\lambda | v) = \frac{1}{2\pi} \lim_{T \rightarrow \infty} \sum_{r=-(T-1)}^{T-1} \omega_r^* \left(1 - \frac{|r|}{T} \right) \cos \lambda r \cos vr.$$

Отсюда в свою очередь вытекает, что

$$(5) \quad \lim_{T \rightarrow \infty} \hat{\mathcal{E}}_T(v) = \frac{1}{2\pi} \lim_{T \rightarrow \infty} \sum_{r=-(T-1)}^{T-1} w_r^* \left(1 - \frac{|r|}{T}\right) \cos vr \sigma(r) = \\ = \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} w_r^* \cos vr \sigma(r) = \int_{-\pi}^{\pi} w(\lambda | v) f(\lambda) d\lambda.$$

Таким образом, указанная последовательность (1) является асимптотически несмещенной оценкой взвешенного среднего функции $f(\lambda)$, представляемого правой частью соотношения (5). Поскольку условие (2) влечет за собой непрерывность функции $w(\lambda | v)$, то указанное взвешенное среднее не тождественно самой $f(\lambda)$.

Дисперсия оценки $\hat{f}_T(v)$ имеет порядок $1/T$. Положим

$$(6) \quad \hat{f}_{TK}(v) = \frac{1}{2\pi} \sum_{r=-K}^K w_r^* \cos vrc_r.$$

Применение теоремы 8.3.3 приводит к соотношению

$$(7) \quad \lim_{T \rightarrow \infty} T \text{Cov} [\hat{f}_{TK}(v), \hat{f}_{TK}(v')] = \\ = \frac{1}{(2\pi)^2} \sum_{g,h=-K}^K w_g^* w_h^* \cos vg \cos v'h \times \\ \times \left[4\pi \int_{-\pi}^{\pi} \cos \lambda g \cos \lambda h f^2(\lambda) d\lambda + \sum_{r=-\infty}^{\infty} \kappa(g, -r, h-r) \right] = \\ = 4\pi \int_{-\pi}^{\pi} w_K^*(\lambda | v) w_K^*(\lambda | v') f^2(\lambda) d\lambda + \\ + \frac{1}{(2\pi)^2} \sum_{g,h=-K}^K w_g^* w_h^* \cos vg \cos v'h \sum_{r=-\infty}^{\infty} \kappa(g, -r, h-r),$$

в котором

$$(8) \quad w_K^*(\lambda | v) = \frac{1}{2\pi} \sum_{g=-K}^K w_g^* \cos vg \cos \lambda g.$$

Если выполнено условие $\int_{-\pi}^{\pi} f^2(\lambda) d\lambda < \infty$, то из (2) следует, что при $K \rightarrow \infty$ первое слагаемое в правой части (7) имеет

предел $4\pi \int_{-\pi}^{\pi} w^*(\lambda | v) w^*(\lambda | v') f^2(\lambda) d\lambda$. Если к тому же величина

$\left| \sum_{r=-\infty}^{\infty} \kappa(g, -r, h-r) \right|$ равномерно ограничена по g и h , то при

$K \rightarrow \infty$ существует предел и у второго слагаемого в правой части. Используя аналог следствия 7.7.1 (см. упр. 15), получаем, что предел при $T \rightarrow \infty$ левой части (7) совпадает с $\lim_{T \rightarrow \infty} T \text{Cov} [\hat{f}_T(v), \hat{f}_T(v')]$.

ТЕОРЕМА 9.3.1. При выполнении условия (2) предел математического ожидания $\mathbb{E} \hat{f}_T(v)$ оценки $\hat{f}_T(v)$, определяемой соотношением (1), равен $\int_{-\pi}^{\pi} w(\lambda|v) f(\lambda) d\lambda$, где $w(\lambda|v) = w^*(\lambda|v)$ — функция, указанная в (3). Если при этом спектральная плотность $f(\lambda)$ непрерывна на отрезке $[-\pi, \pi]$, а величина $\left| \sum_{r=-\infty}^{\infty} \kappa(g, -r, h-r) \right|$ равномерно ограничена, то

$$(9) \quad \lim_{T \rightarrow \infty} T \text{Cov} [\hat{f}_T(v), \hat{f}_T(v')] = 4\pi \int_{-\pi}^{\pi} w(\lambda|v) w(\lambda|v') f^2(\lambda) d\lambda + \\ + \frac{1}{(2\pi)^2} \sum_{g,h=-\infty}^{\infty} w_g^* w_h^* \cos vg \cos v'h \sum_{r=-\infty}^{\infty} \kappa(g, -r, h-r).$$

Если

$$(10) \quad y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$$

линейный процесс, то

$$(11) \quad \sum_{r=-\infty}^{\infty} \kappa(g, -r, h-r) = \kappa_4 \frac{\sigma(g) \sigma(h)}{\sigma^4}$$

и второе слагаемое в правой части (9) принимает более простой вид.

СЛЕДСТВИЕ 9.3.1. Если последовательность $\{y_t\}$ порождается с помощью соотношения (10), в котором $\mathbb{E} v_t = 0$, $\mathbb{E} v_t^2 = \sigma^2$, $\mathbb{E} v_t v_s = 0$, $t \neq s$, $\mathbb{E} v_t v_s v_q = 0$, $t \neq s$, $t \neq r$, $t \neq q$, $\mathbb{E} v_t^4 = 3\sigma^4 + \kappa_4 < \infty$, $\mathbb{E} v_t^2 v_s^2 = \sigma^4$, $t \neq s$, $\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty$, то при выполнении условия (2)

$$(12) \quad \lim_{T \rightarrow \infty} T \text{Cov} [\hat{f}_T(v), \hat{f}_T(v')] = \\ = 4\pi \int_{-\pi}^{\pi} w(\lambda|v) w(\lambda|v') f^2(\lambda) d\lambda + \frac{\kappa_4}{\sigma^4} \left[\int_{-\pi}^{\pi} w(\lambda|v) f(\lambda) d\lambda \right] \times \\ \times \left[\int_{-\pi}^{\pi} w(\lambda|v') f(\lambda) d\lambda \right] = 4\pi \int_{-\pi}^{\pi} w(\lambda|v) w(\lambda|v') f^2(\lambda) d\lambda + \\ + \frac{\kappa_4}{\sigma^4} \lim_{T \rightarrow \infty} \mathbb{E} \hat{f}_T(v) \lim_{T \rightarrow \infty} \mathbb{E} \hat{f}_T(v').$$

В разд. 8.4.2 (теорема 8.4.2) было показано, что для линейного процесса, получаемого с помощью независимых и одинаково распределенных случайных величин, всякий конечный набор выборочных ковариаций имеет асимптотически нормальное распределение. Поэтому при любом фиксированном K предельное распределение величины $\sqrt{T}[\hat{f}_{TK}(v) - \mathcal{E}\hat{f}_{TK}(v)]$ будет нормальным ($T \rightarrow \infty$). Применение следствия 7.7.1 показывает, что оценка $\hat{f}_T(v)$ также асимптотически нормальна.

ТЕОРЕМА 9.3.2. Если процесс y_t порождается соотношением (10), в котором $\{v_t\}$ — последовательность независимых одинаково распределенных случайных величин с $\mathcal{E}v_t = 0$, $\mathcal{E}v_t^2 = \sigma^2$, $\mathcal{E}v_t^4 = 3\sigma^4 + \kappa_4 < \infty$ и $\sum_{t=-\infty}^{\infty} |\gamma_t| < \infty$, то при выполнении условия (2) вектор

$$\sqrt{T} \left[\hat{f}_T(v_1) - \int_{-\pi}^{\pi} w(\lambda | v_1) f(\lambda) d\lambda \right], \dots, \sqrt{T} \left[\hat{f}_T(v_n) - \int_{-\pi}^{\pi} w(\lambda | v_n) f(\lambda) d\lambda \right]$$
 имеет в пределе (при $T \rightarrow \infty$) нормальное распределение с нулевым средним и ковариациями, задаваемыми по формуле (12).

Пусть оценка $\hat{f}_T^*(v)$ определяется аналогично оценке (1), но в правой части (1) c_{rT} заменяются на $\dot{c}_{rT}$, где $\dot{c}_{rT}$ обозначает r -ю выборочную ковариацию, использующую отклонения от выборочного среднего (и деленную на T). Тогда $\mathcal{E}\hat{f}_T^*(v)$ имеет тот же предел (5) и к $\hat{f}_T^*(v)$ применимы теорема 9.3.1, следствие 9.3.1 и теорема 9.3.2. Подобным же образом c_{rT} можно заменить на C_{rT} , $\dot{C}_{rT}$ или $\tilde{c}_{rT}$ и получить те же самые предельные результаты.

9.3.2. Асимптотическое смещение

Если представляется желательным состоятельно оценить само значение плотности $f(\lambda)$ в точке $\lambda = v$, а не усредненное значение $\bar{f}(\lambda)$, то для этого следует выбрать такую последовательность оценок, что соответствующая последовательность математических ожиданий будет сходиться к $f(v)$ для широкого класса плотностей. Будем брать коэффициенты $w_r = (T - |r|) w_r^*/T$ зависящими от T и возьмем оценку спектральной плотности в виде

$$\begin{aligned} (13) \quad \hat{f}_T(v) &= \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_{rT} \cos vr C_{rT} = \\ &= \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_{rT}^* \cos vr c_{rT} = \int_{-\pi}^{\pi} w_T^*(\lambda | v) I_T(\lambda) d\lambda, \end{aligned}$$

где $w_{rT} = (T - |r|) w_{rT}^* / T$ и

$$(14) \quad w_T^*(\lambda | \nu) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_{rT}^* \cos \lambda r \cos \nu r.$$

Здесь мы пишем $I_T(\lambda)$ вместо $I(\lambda)$, чтобы указать на количество T используемых наблюдений. Последовательность $\{w_{rT}^*\}$ при каждом T является числовой, $r = 0, 1, \dots$ ($w_{rT}^* = 0$ для $T \leq r$ и $w_{-rT}^* = w_{rT}^*$). Сами же $w_T^*(\lambda | \nu)$ образуют последовательность функций. Математическое ожидание оценки $\hat{f}_T(\nu)$ равно

$$(15) \quad \mathbb{E} \hat{f}_T(\nu) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_{rT} \cos \nu r \sigma(r) = \int_{-\pi}^{\pi} w_T(\lambda | \nu) f(\lambda) d\lambda,$$

где

$$(16) \quad w_T(\lambda | \nu) = \frac{1}{2\pi} \sum_{r=-(T-1)}^{T-1} w_{rT} \cos \lambda r \cos \nu r.$$

Мы хотим, чтобы

$$(17) \quad \lim_{T \rightarrow \infty} \mathbb{E} \hat{f}_T(\nu) = f(\nu)$$

для всех спектральных плотностей из некоторого интересующего нас класса.

Большинство оценок, описанных в разд. 9.2.3, определяется коэффициентами вида

$$(18) \quad w_{rT}^* = \begin{cases} k(r/K_T), & r = 0, \pm 1, \dots, \pm K_T, \\ 0, & r = \pm (K_T + 1), \dots, \pm (T-1), \end{cases}$$

где обозначение K_T для K указывает на зависимость от T . Функция $k(x)$ удовлетворяет при этом условиям

$$(19) \quad k(x) = k(-x),$$

$$(20) \quad k(0) = 1,$$

так что $w_{0T}^* = w_{0T} = 1$ и поэтому $\int_{-\pi}^{\pi} w_T^*(\lambda | \nu) d\lambda = \int_{-\pi}^{\pi} w_T(\lambda | \nu) d\lambda = 1$. Функция $k(x)$ должна быть непрерывной в нуле. Асимптотическое смещение оценки зависит от степени гладкости функции $k(x)$ в нуле. Теоретическое исследование асимптотических смещения, дисперсии и среднеквадратичной ошибки было проведено Парзенем (1957b), построений которого мы и будем придерживаться в дальнейшем.

Предположим, что существуют такие числа $q > 0$ и $k > 0$, что

$$(21) \quad \lim_{x \rightarrow 0} \frac{1 - k(x)}{|x|^q} = k$$

(q является наибольшим показателем, при котором еще $k < \infty$). Если число q целое и четное, то k равно взятому с противоположным знаком значению q -й производной функции $k(x)$ в точке $x = 0$, деленному на $q!$. Если q целое и нечетное, то k равно взятому с противоположным знаком значению q -й правой производной функции $k(x)$ в точке $x = 0$, деленному на $q!$. Из условия $k > 0$ вытекает, что в некоторой окрестности нуля функция $k(x)$ убывает с удалением x от нуля. Для малых $|x|$ функция $k(x)$ хорошо аппроксимируется функцией $1 - k|x|^q$. Если $k(x) = \int_{-\infty}^{\infty} K(v) e^{ivx} dv$ и $q = 2$,

то $k = \frac{1}{2} \int_{-\infty}^{\infty} v^2 K(v) dv$. [При соответствующих условиях $K(v) = \lim \omega_T^*(v/K_T | 0)/K_T$ при $K_T \rightarrow \infty$; см. упр. 17.] Константа k в некотором смысле указывает ширину окон соответствующего семейства.

Асимптотическое смещение оценки спектральной плотности зависит также от степени гладкости самой спектральной плотности. Предположим, что для некоторого $p > 0$

$$(22) \quad \sum_{r=-\infty}^{\infty} |r|^p |\sigma(r)| < \infty.$$

Пусть

$$(23) \quad f^{(s)}(\lambda) = \frac{(-i)^s}{2\pi} \sum_{r=-\infty}^{\infty} s r e^{-i\lambda r} \sigma(r)$$

для целых $s \leq p$. Тогда $f^{(s)}(\lambda)$ является s -й производной $f(\lambda)$. Из условия (22) вытекает, что все производные порядка не выше p ограничены и непрерывны. Пусть $\{K_T\}$ — последовательность целых чисел, зависящих от T таким образом, что $K_T \rightarrow \infty$ и $K_T^m/T \rightarrow 0$, $m = \min(p, q)$ при $T \rightarrow \infty$. Тогда (при $\omega_{rT} = 0$ для $|r| > K_T$)

$$(24) \quad K_T^m |\widehat{g f_T}(v) - f(v)| = \\ = \frac{K_T^m}{2\pi} \sum_{r=-K_T}^{K_T} \left[k\left(\frac{r}{K_T}\right) \left(1 - \frac{|r|}{T}\right) - 1 \right] \cos vr \sigma(r) - \\ - \frac{2K_T^m}{2\pi} \sum_{r=K_T+1}^{\infty} \cos vr \sigma(r) =$$

$$\begin{aligned}
&= \frac{K_T^m}{2\pi} \sum_{r=-K_T}^{K_T} \left[k\left(\frac{r}{K_T}\right) - 1 \right] \cos vr \sigma(r) - \\
&\quad - \frac{K_T^m}{\pi} \sum_{r=1}^{K_T} k\left(\frac{r}{K_T}\right) \frac{r}{T} \cos vr \sigma(r) - \\
&\quad - \frac{K_T^m}{\pi} \sum_{r=K_T+1}^{\infty} \cos vr \sigma(r).
\end{aligned}$$

Первый член в правой части (24) можно записать в виде

$$\begin{aligned}
(25) \quad \frac{1}{2\pi} \sum_{r=-K_T^*}^{K_T^*} \frac{k\left(\frac{r}{K_T}\right) - 1}{\left|\frac{r}{K_T}\right|^m} |r|^m \cos vr \sigma(r) + \\
+ \frac{1}{\pi} \sum_{r=K_T^*+1}^{K_T} \frac{k\left(\frac{r}{K_T}\right) - 1}{\left(\frac{r}{K_T}\right)^m} r^m \cos vr \sigma(r)
\end{aligned}$$

для любого целого $K_T^* < K_T$. В силу условия (21) для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что при $|r/K_T| < \delta$

$$(26) \quad \left| \frac{k\left(\frac{r}{K_T}\right) - 1}{\left|\frac{r}{K_T}\right|^q} + k \right| < \varepsilon.$$

В частности, (26) будет выполнено для всех r , участвующих в формуле (25), если положить $K_T^* = [\delta K_T]$, где δ — указанное выше число. Поэтому при $m = q \leq p$ первый член в (25) с точностью до

$\varepsilon' = \varepsilon \sum_{r=-\infty}^{\infty} |r|^q |\sigma(r)| / (2\pi)$ равен

$$(27) \quad - \frac{k}{2\pi} \sum_{r=-K_T^*}^{K_T^*} |r|^q \cos vr \sigma(r).$$

Последнее выражение стремится к значению $-kf^{(q)}(v)$, где

$$(28) \quad f^{(q)}(v) = \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} |r|^q \cos vr \sigma(r).$$

Если $m = p < q$, то первый член в (25) с точностью до ε' (при $\delta \ll 1$) равен

$$(29) \quad -\frac{k}{2\pi} \sum_{r=-K_T^*}^{K_T^*} \left| \frac{r}{K_T} \right|^{q-p} |r|^p \cos vr \sigma(r).$$

Последнее выражение не превосходит по абсолютной величине

$$(30) \quad \delta^{q-p} \frac{k}{2\pi} \sum_{r=-\infty}^{\infty} |r|^p |\sigma(r)|,$$

которое может быть сделано сколь угодно малым соответствующим выбором δ . Далее, если $|k(x)| \leq M$, то

$$(31) \quad \frac{|1 - k(x)|}{|x|^m} \leq \frac{M+1}{\delta^m}, \quad |x| \geq \delta.$$

Поэтому второе слагаемое в (25) по абсолютной величине не превосходит значения

$$(32) \quad \frac{M+1}{\pi \delta^m} \sum_{r=[\delta K_T]+1}^{\infty} r^m |\sigma(r)|,$$

стремящегося к нулю. Что касается второго члена правой части (24), то его абсолютная величина не превосходит значения

$$(33) \quad M \frac{K_T^m}{\pi T} \sum_{r=1}^{K_T} r |\sigma(r)|,$$

стремящегося к нулю при $p \geq 1$ и $K_T^m/T \rightarrow 0$. Если $p \leq 1$, то (33) не превосходит

$$(34) \quad \frac{M}{\pi} \frac{K_T^{m+1-p}}{T} \sum_{r=1}^{K_T} r^p |\sigma(r)|,$$

стремящегося к 0 при $K_T^{m+1-p}/T \rightarrow 0$. Наконец, третий член правой части (24) мажорируется по абсолютной величине выражением

$$(35) \quad \frac{K_T^m}{\pi} \sum_{r=K_T+1}^{\infty} |\sigma(r)| \leq \frac{1}{\pi} \sum_{r=K_T+1}^{\infty} |r|^p |\sigma(r)|,$$

стремящимся к 0 при условии (22).

ТЕОРЕМА 9.3.3. Пусть

$$(36) \quad \hat{f}_T(v) = \frac{1}{2\pi} \sum_{r=-K_T}^{K_T} k\left(\frac{r}{K_T}\right) \cos vr c_{r,T},$$

где $k(x) = k(-x)$, $k(0) = 1$, $|k(x)| \leq M$ для некоторого M и всех $|x| \leq 1$, и, кроме того, $\lim_{x \rightarrow 0} [1 - k(x)]/|x|^q = k$ для некоторых $q > 0$ и $k > 0$. Пусть $\{K_T\}$ — некоторая последовательность целых чисел, такая, что $K_T \rightarrow \infty$ при $T \rightarrow \infty$. Тогда если для $p \geq q$ выполнено условие

$$(37) \quad \sum_{r=-\infty}^{\infty} |r|^p |\sigma(r)| < \infty$$

и при этом $p \geq 1$ и $K_T^q/T \rightarrow 0$ или $p \leq 1$ и $K_T^{q+1-p}/T \rightarrow 0$, то

$$(38) \quad \lim_{T \rightarrow \infty} K_T^q [\hat{g}_T^f(v) - f(v)] = -\frac{k}{2\pi} \sum_{r=-\infty}^{\infty} |r|^q \cos vr \sigma(r).$$

Если условие (37) выполняется для $p < q$ и при этом $p \geq 1$ и $K_T^p/T \rightarrow 0$ или $p \leq 1$ и $K_T/T \rightarrow 0$, то

$$(39) \quad \lim_{T \rightarrow \infty} K_T^p [\hat{g}_T^f(v) - f(v)] = 0.$$

Условия, указанные в теореме, можно свести в следующую таблицу.

Таблица 9.3.1
УСЛОВИЯ ТЕОРЕМЫ 9.3.3

	$p \leq 1$	$p > 1$
$q \leq p$	$\frac{K_T^{q+1-p}}{T} \rightarrow 0$	$\frac{K_T^q}{T} \rightarrow 0$
$q > p$	$\frac{K_T}{T} \rightarrow 0$	$\frac{K_T^p}{T} \rightarrow 0$

Теорема 9.3.3 показывает, что если степень гладкости q весовой функции превосходит степень гладкости p спектральной плотности, то смещение является величиной порядка $o(K_T^{-p})$. Если же $q \leq p$, то смещение приближенно равно $-kf^{[q]}(v)/K_T^q$, т. е. уменьшается пропорционально величине K_T^{-q} . И в том, и в другом случае оценка $\hat{f}_T(v)$ (при $p > 0$ и $q > 0$) является асимптотически несмещенной. Теорема устанавливает порядок смещения. Если рассматривается класс процессов с вполне определенным значением p , то величину смещения можно (в смысле теоремы 9.3.3) сделать малой, выбирая ядро $k(x)$ с $q > p$. Если $q \leq p$, то смещение уменьшают выбором ядра $k(x)$ с малым значением k . Заметим, что при четном q правая часть (38) равна $(-1)^{1+q/2} kf^{[q]}(v)$.

Соответствующие характеристики для рассмотренных выше примеров приведены в табл. 9.3.2.

Таблица 9.3.2
ХАРАКТЕРИСТИКИ ОКОН

	$k(x)$ для $ x \leq 1$	q	k
В. Усеченная оценка	1	∞	
D. Модифицированная оценка Бартлетта	$1 - x $	1	1
Е. Усеченная оценка Даниэля ($b = \pi/K_T$)	$\sin \pi x / (\pi x)$	2	$\pi^2/6$
F. Оценка Блэкмена — Тьюки	$1 - 2a + 2a \cos \pi x$	2	$\pi^2 a$
G. Окно Хеннинга	$(1 + \cos \pi x)/2$	2	$\pi^2/4$
H. Окно Хемминга	$0.54 + 0.46 \cos \pi x$	2	$0.23\pi^2$
I. Оценка Парзена	$1 - x^2$	2	1
J. Оценка Парзена	$1 - 6x^2 + 6 x ^3, x \leq 1/2,$ $2(1 - x)^3, 1/2 \leq x \leq 1$	2	6

Значение $f(\lambda)$ в точке $\lambda = v$ можно оценить и тогда, когда среднее $\mathbb{E}y_t = \mu$ неизвестно. Для этого можно использовать оценку $\hat{f}_T^*(v)$, которая устроена аналогично $\hat{f}_T(v)$, только вместо c_{rT} используются ковариации c_{rT} , основанные на отклонениях от выборочного среднего. Вместо c_{rT} можно использовать также и $\tilde{c}_{rT}$. В конце § 9.4 показывается, что если $K_T^{m+1}/T \rightarrow 0$ при $T \rightarrow \infty$, то смещения оценок $\hat{f}_T^*(v)$ и $\hat{f}_T(v)$ асимптотически эквивалентны.

9.3.3. Асимптотические дисперсии и ковариации

Перейдем теперь к изучению асимптотического поведения дисперсий и ковариаций оценок (13), в которых $\omega_{rT}^* = k(r/K_T)$, $r = 0, \pm 1, \dots, \pm K_T$ и $\omega_{rT}^* = 0$, $r = \pm(K_T + 1), \dots, \pm(T - 1)$, $k(x) = k(-x)$ и $K_T \rightarrow \infty$, $K_T/T \rightarrow 0$ при $T \rightarrow \infty$. Будем предполагать, что $k(x)$ непрерывна на $[-1, 1]$ и $\sum_{r=-\infty}^{\infty} |\sigma(r)| < \infty$.

Из разд. 8.2.2 следует, что

$$\begin{aligned}
 (40) \quad T \text{Cov}(c_{gT}, c_{hT}) &= \\
 &= \frac{1}{T} \left[\mathbb{E} \sum_{i,s} (y_i - \mu)(y_{i+h} - \mu)(y_s - \mu)(y_{s+g} - \mu) - \right. \\
 &\quad \left. - (T - |g|)(T - |h|) \sigma(g) \sigma(h) \right] =
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{T} \sum_{t,s} [\sigma(t-s) \sigma(t+h-s-g) + \sigma(t-s-g) \times \\
&\quad \times \sigma(t-s+h) + \kappa(h, s-t, s-t+g)] = \\
&= \sum_{r=-(T-1)}^{T-1} \varphi_T(r; g, h) [\sigma(r) \sigma(r+h-g) + \\
&\quad + \sigma(r-g) \sigma(r+h) + \kappa(h, -r, g-r)], \\
&\quad g, h = 0, \pm 1, \dots, \pm(T-1),
\end{aligned}$$

где суммирование ведется по тем s и t , для которых $(1, 1-g) \leq s \leq \min(T-g, T)$ и $\max(1, 1-h) \leq t \leq \min(T-h, T)$, а $T\varphi_T(r; g, h)$ есть число таких пар (s, t) , которые удовлетворяют обоим указанным неравенствам и для которых $t-s = r, r = -(T-1), \dots, (T-1)$. Точные выражения для $T\varphi_T(r; g, h)$ даны в упр. 19. [Для некоторых троек $(r; g, h)$ $\varphi_T(r; g, h) = 0$.] Отметим, что

$$(41) \quad 0 \leq \varphi_T(r; g, h) \leq 1$$

и

$$(42) \quad \lim_{T \rightarrow \infty} \varphi_T(r; g, h) = 1$$

для каждого набора r, g и h . В действительности справедливо даже следующее неравенство:

$$(43) \quad \varphi_T(r; g, h) \geq 1 - \frac{|r| + |g| + |h|}{T}.$$

Используя (40), получаем далее

$$\begin{aligned}
(44) \quad &\frac{T}{K_T} \text{Cov}[\hat{f}_T(\lambda), \hat{f}_T(\nu)] = \\
&= \frac{T}{(2\pi)^2 K_T} \sum_{g,h=-K_T}^{K_T} k\left(\frac{g}{K_T}\right) k\left(\frac{h}{K_T}\right) \times \\
&\quad \times \cos \lambda g \cos \nu h \text{Cov}(c_{gT}, c_{hT}) = \\
&= \frac{1}{(2\pi)^2 K_T} \sum_{g,h=-K_T}^{K_T} k\left(\frac{g}{K_T}\right) k\left(\frac{h}{K_T}\right) e^{i(\lambda g + \nu h)} \times \\
&\quad \times \sum_{r=-(T-1)}^{T-1} \varphi_T(r; g, h) [\sigma(r) \sigma(r+h-g) + \\
&\quad + \sigma(r-g) \sigma(r+h) + \kappa(h, -r, g-r)].
\end{aligned}$$

Рассмотрим входящую в это выражение сумму

$$\begin{aligned}
 (45) \quad & \frac{1}{(2\pi)^2 K_T} \sum_{g, h=-K_T}^{K_T} \sum_{r=-(T-1)}^{T-1} \varphi_T(r; g, h) k\left(\frac{g}{K_T}\right) k\left(\frac{h}{K_T}\right) \times \\
 & \times e^{i(\lambda g + \nu h)} \sigma(r-g) \sigma(r+h) = \\
 & = \frac{1}{(2\pi)^2 K_T} \sum_{u, v=-(K_T+T-1)}^{K_T+T-1} \sum_{r=\max[u-K_T, v-K_T, -(T-1)]}^{\min[u+K_T, v+K_T, T-1]} \varphi_T(r; r-u, v-r) \times \\
 & \times k\left(\frac{r-u}{K_T}\right) k\left(\frac{v-r}{K_T}\right) e^{i(\nu v - \lambda u) + i(\lambda - \nu)r} \sigma(u) \sigma(v).
 \end{aligned}$$

Здесь суммы по r равны нулю, если указанный нижний предел оказывается больше верхнего.

Заметим, что разность между (45) и

$$\begin{aligned}
 (46) \quad & \frac{1}{(2\pi)^2 K_T} \sum_{u, v=-m}^m \sum_{r=\max[u-K_T, v-K_T, -(T-1)]}^{\min[u+K_T, v+K_T, T-1]} \varphi_T(r; r-u, v-r) \times \\
 & \times k\left(\frac{r-u}{K_T}\right) k\left(\frac{v-r}{K_T}\right) e^{i(\nu v - \lambda u) + i(\lambda - \nu)r} \sigma(u) \sigma(v)
 \end{aligned}$$

по абсолютной величине не превосходит

$$(47) \quad \left(2 + \frac{1}{K_T}\right) \frac{4}{(2\pi)^2} \sum_{u=-\infty}^{\infty} \sum_{v=m+1}^{\infty} \sup_{-1 \leq x \leq 1} k^2(x) |\sigma(u)| |\sigma(v)|.$$

Поскольку же $\sum_{r=-\infty}^{\infty} |\sigma(r)| < \infty$, то последняя может быть сделана сколь угодно малой, если m ($\leq K_T$) выбрать достаточно большим и если функция $|k(x)|$ ограничена. Если функция $k(x)$ непрерывна на отрезке $[-1, 1]$, то для $|u| \leq m$, $|v| \leq m$ и достаточно больших значений K_T

$$(48) \quad \left| k\left(\frac{r-u}{K_T}\right) k\left(\frac{v-r}{K_T}\right) - k^2\left(\frac{r}{K_T}\right) \right| < \varepsilon$$

для всех таких r , что $-K_T \leq r-u \leq K_T$, $-K_T \leq v-r \leq K_T$ и $-K_T \leq r \leq K_T$. Кроме того, для $|u| \leq m$, $|v| \leq m$, $|r| \leq m + K_T$

$$(49) \quad \varphi_T(r; r-u, v-r) \geq 1 - \frac{5m + 3K_T}{T}.$$

Это наводит на мысль использовать вместо (46) сумму

$$\begin{aligned}
 (50) \quad & \frac{1}{(2\pi)^2} \sum_{u,v=-m}^m \sum_{r=-K_T}^{K_T} \frac{1}{K_T} k^2 \left(\frac{r}{K_T} \right) e^{i(vv-\lambda u)+i(\lambda-v)r} \sigma(u) \sigma(v) = \\
 & = \left[\sum_{r=-K_T}^{K_T} \frac{1}{K_T} k^2 \left(\frac{r}{K_T} \right) e^{i(\lambda-v)r} \right] \left[\frac{1}{2\pi} \sum_{u=-m}^m e^{-i\lambda u} \sigma(u) \right] \times \\
 & \times \left[\frac{1}{2\pi} \sum_{v=-m}^m e^{ivv} \sigma(v) \right].
 \end{aligned}$$

Действительно, разность между (46) и (50) становится при больших T сколь угодно малой. В то же время, если $\lambda - v = 0$ или $\pm 2\pi$, то при $T \rightarrow \infty$ входящая в (50) сумма по r имеет предел

$$(51) \quad \lim_{K_T \rightarrow \infty} \sum_{r=-K_T}^{K_T} \frac{1}{K_T} k^2 \left(\frac{r}{K_T} \right) = \int_{-1}^1 k^2(x) dx,$$

так что для достаточно больших m предел (50) становится сколь угодно близок к величине

$$(52) \quad f^2(v) \int_{-1}^1 k^2(x) dx.$$

Если же $\lambda - v \neq 0, \pm 2\pi$, то $0 < |\lambda - v| < 2\pi$ и в силу результата упр. 23 (аналога теоремы Римана — Лебега)

$$(53) \quad \lim_{K_T \rightarrow \infty} \sum_{r=-K_T}^{K_T} \frac{1}{K_T} k^2 \left(\frac{r}{K_T} \right) e^{i(\lambda-v)r} = 0.$$

При этом будет равен нулю и предел всего выражения (50).

Рассмотрим теперь следующую составляющую суммы, стоящей в правой части (44):

$$\begin{aligned}
 (54) \quad & \frac{1}{(2\pi)^2 K_T} \sum_{g,h=-K_T}^{K_T} \sum_{r=-(T-1)}^{T-1} \varphi_T(r; g, h) \times \\
 & \times k \left(\frac{g}{K_T} \right) k \left(\frac{h}{K_T} \right) e^{i(\lambda g + v h)} \sigma(r) \sigma(r + h - g) = \\
 & = \frac{1}{(2\pi)^2 K_T} \sum_{u=-(T-1)}^{T-1} \sum_{v=u-2K_T}^{u+2K_T} \sum_{s=\max(u,v)-K_T}^{\min(u,v)+K_T} \varphi_T(u; u-s, v-s) \times \\
 & \times k \left(\frac{u-s}{K_T} \right) k \left(\frac{v-s}{K_T} \right) e^{i(\lambda u + v v) - i(\lambda + v)s} \sigma(u) \sigma(v).
 \end{aligned}$$

Последнюю можно аппроксимировать суммой

$$(55) \quad \frac{1}{(2\pi)^2 K_T} \sum_{u,v=-m}^m \sum_{s=\max(u,v)-K_T}^{\min(u,v)+K_T} \varphi_T(u; u-s, v-s) \times \\ \times k\left(\frac{u-s}{K_T}\right) k\left(\frac{v-s}{K_T}\right) e^{i(\lambda u + \nu v) - i(\lambda + \nu)s} \sigma(u) \sigma(v),$$

а ту в свою очередь суммой

$$(56) \quad \frac{1}{(2\pi)^2} \sum_{u,v=-m}^m \sum_{s=-K_T}^{K_T} \frac{1}{K_T} k^2\left(\frac{s}{K_T}\right) e^{i(\lambda u + \nu v) - i(\lambda + \nu)s} \sigma(u) \sigma(v),$$

поскольку $k[(u-s)/K_T] k[(v-s)/K_T]$ при больших K_T приближенно равно $k^2(-s/K_T) = k^2(s/K_T)$. Если теперь $\lambda + \nu = 0$ или $\pm 2\pi$, то (56) имеет в качестве предела

$$(57) \quad f^2(\nu) \int_{-1}^1 k^2(x) dx.$$

Если же $\nu + \lambda \neq 0, \pm 2\pi$, то предел (56) равен 0.

Наконец, входящая в (44) сумма

$$(58) \quad \frac{1}{(2\pi)^2 K_T} \sum_{g,h=-K_T}^{K_T} \sum_{r=-(T-1)}^{T-1} \varphi_T(r; g, h) k\left(\frac{g}{K_T}\right) k\left(\frac{h}{K_T}\right) \times \\ \times e^{i(\lambda g + \nu h)} \kappa(h, -r, g-r)$$

не превосходит по абсолютной величине

$$(59) \quad \frac{1}{(2\pi)^2 K_T} \sup_{-1 \leq x \leq 1} k^2(x) \sum_{r,s,t=-\infty}^{\infty} |\kappa(r, s, t)|$$

и, следовательно, ее предел при $T \rightarrow \infty$ равен 0, если только

$$\sum_{r,s,t=-\infty}^{\infty} |\kappa(r, s, t)| < \infty.$$

ТЕОРЕМА 9.3.4. Пусть оценка $\hat{f}_T(\nu)$ определена соотношением (36), причем $k(x) = k(-x)$ и функция $k(x)$ непрерывна на отрезке $[-1, 1]$. Предположим, что выполнены условия

$$(60) \quad \sum_{r=-\infty}^{\infty} |\sigma(r)| < \infty,$$

$$(61) \quad \sum_{r,s,t=-\infty}^{\infty} |\kappa(r, s, t)| < \infty.$$

Пусть последовательность целых чисел $\{K_T\}$ такова, что $K_T \rightarrow \infty$ и $K_T/T \rightarrow 0$ при $T \rightarrow \infty$. Тогда

$$(62) \quad \lim_{T \rightarrow \infty} \frac{T}{K_T} \text{Var } \hat{f}_T(0) = 2f^2(0) \int_{-1}^1 k^2(x) dx,$$

$$(63) \quad \lim_{T \rightarrow \infty} \frac{T}{K_T} \text{Var } \hat{f}_T(\pm \pi) = 2f^2(\pi) \int_{-1}^1 k^2(x) dx,$$

$$(64) \quad \lim_{T \rightarrow \infty} \frac{T}{K_T} \text{Var } \hat{f}_T(v) = f^2(v) \int_{-1}^1 k^2(x) dx, \quad v \neq 0, \pm \pi,$$

$$(65) \quad \lim_{T \rightarrow \infty} \frac{T}{K_T} \text{Cov} [\hat{f}_T(\lambda), \hat{f}_T(v)] = 0, \quad v \neq \pm \lambda.$$

Теорема 9.3.4 показывает, что дисперсия $\hat{f}_T(v)$ имеет порядок K_T/T . Поскольку $K_T \rightarrow \infty$, отношение K_T/T будет больше отношения K/T для произвольного K и для всех достаточно больших значений T . Поэтому дисперсия состоятельной оценки вида (36) будет асимптотически большей, чем дисперсия оценок того типа, который рассматривался в разд. 9.3.1.

Если фиксировать целочисленную последовательность $\{K_T\}$, то различные оценки можно сравнивать по величине интеграла $\int_{-1}^1 k^2(x) dx$. Примеры такого сравнения даны в табл. 9.3.3. Как видно из этой таблицы и из табл. 9.3.2, для ядер с $q = 2$ характеристика

Таблица 9.3.3
ЗНАЧЕНИЯ ИНТЕГРАЛОВ ДЛЯ НЕКОТОРЫХ ЯДЕР

$k(x)$ для $ x \leq 1$		$\int_{-1}^1 k^2(x) dx$
В. Усеченная оценка	1	2
Д. Модифицированная оценка Бартлетта	$1 - x $	2/3
Е. Усеченная оценка Даниэля ($b = \pi/K_T$)	$\sin \pi x / (\pi x)$	0.90282336
Ф. Оценка Блэкмена — Тюнки	$1 - 2a + 2a \cos \pi x$	$2(1 - 4a + 6a^2)$
Г. Окно Хеннинга	$(1 + \cos \pi x)/2$	$3/4 = 0.75$
Н. Окно Хемминга	$0.54 + 0.46 \cos \pi x$	0.7948
И. Оценка Парзена	$1 - x^2$	$16/15 = 1.0667$
Ж. Оценка Парзена	$1 - 6x^2 + 6 x ^3, x \leq 1/2,$ $2(1 - x)^3, 1/2 \leq x \leq 1$	$151/280 = 0.5393$

смещения k и величина указанного интеграла имеют тенденцию изменяться противоположным образом.

Отметим, что, как следует из соотношения (44) § 8.3, условие (61) будет выполнено, если, например, $y_t = \mu + \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, где $\sum_{s=-\infty}^{\infty} |\gamma_s| < \infty$ и первые четыре момента последовательности $\{v_t\}$ соответствуют требованиям стационарности и независимости v_t , $t = 0, \pm 1, \dots$.

Если среднее μ неизвестно, то в определении $\hat{f}_T(v)$ следует заменить c_{rT} на $\bar{c}_{rT}$ или на $\tilde{c}_{rT}$. В конце § 9.4 показано, что теорема 9.3.4 приложима и к этим оценкам.

9.3.4. Асимптотическая среднеквадратичная ошибка

Как видно из теорем 9.3.3 и 9.3.4., если T велико, то при соответствующих условиях среднеквадратичная ошибка оценки спектральной плотности в точке v

$$(66) \quad \mathbb{E} [\hat{f}_T(v) - f(v)]^2 = \text{Var } \hat{f}_T(v) + [\mathbb{E} \hat{f}_T(v) - f(v)]^2$$

приблизительно равна ($v \neq 0, \pm \pi$)

$$(67) \quad \frac{K_T}{T} f^2(v) \int_{-1}^1 k^2(x) dx + \left(\frac{1}{K_T^q} \right)^2 k^2(f^{[q]}(v))^2,$$

где

$$(68) \quad f^{[q]}(v) = \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} |r|^q \cos vr \sigma(r).$$

Следует обратить внимание на то, что K_T находится в числителе выражения асимптотической дисперсии для $\hat{f}_T(v)$ и в знаменателе выражения асимптотического смещения. Поэтому, чем больше значение K_T (по отношению к T), тем большей оказывается дисперсия и тем меньшим будет смещение. Можно сказать, что дисперсия и смещение ведут себя противоположным образом. Для того чтобы оба слагаемых в (67) имели одинаковый порядок, необходимо, чтобы величина K_T^{2q+1} имела порядок T . (Если эти слагаемые имеют различные порядки, то одно из них, более высокого порядка, будет доминировать и, кроме того, будет иметь порядок больший, нежели порядок суммы при совпадающих порядках слагаемых.) Для этого можно взять, например, в качестве K_T величину $[\gamma T^{1/(2q+1)}]$.

Тогда

$$\begin{aligned} (69) \quad \lim_{T \rightarrow \infty} T^{2q/(2q+1)} \mathbb{E} [\hat{f}_T(v) - f(v)]^2 = \\ = \gamma f^2(v) \int_{-1}^1 k^2(x) dx + \frac{1}{\gamma^{2q}} \{f^{(q)}(v)\}^2 k^2. \end{aligned}$$

При этом предполагается, что функция $k(x) = k(-x)$ непрерывна на $[-1, 1]$, выполнено условие (21) для $q > 0$ и $k > 0$, выполнено условие (37) для $p = q$, $\sum_{r=-\infty}^{\infty} |\sigma(r)| < \infty$ и $\sum_{r,s,t=-\infty}^{\infty} |\kappa(r, s, t)| < \infty$. Если q — четное целое число, то $(-1)^{q/2} f^{(q)}(v) = f^{(q)}(v)$ есть q -я производная $f(\lambda)$ в точке $\lambda = v$ [см. (23)]. Из справедливости условия (37) для $p = q$ следует, что эта q -я производная существует и непрерывна для всех v .

Порядок среднеквадратичной ошибки равен $T^{-2q/(2q+1)}$. Большим значениям q соответствует меньшее значение этого порядка. В связи с этим оценки Блэкмена — Тьюки и Парзена более предпочтительны, чем модифицированная оценка Бартлетта.

Правая часть (69) зависит от выбранной константы γ , от двух характеристик ядра, $\int_{-1}^1 k^2(x) dx$ и k^2 , а также от $f(v)$ и $f^{(q)}(v)$. Использование асимптотической теории дает возможность сравнивать оценки, ядра которых имеют один и тот же характеристический показатель q . Именно можно сказать, что оценка с ядром $k(x)$ является в асимптотическом плане не хуже оценки с ядром $k^*(x)$ и соответствующей характеристикой k^* , если

$$(70) \quad k^2 \leq k^{*2}, \quad \int_{-1}^1 k^2(x) dx \leq \int_{-1}^1 k^{*2}(x) dx.$$

При этом оценка с ядром $k(x)$ будет асимптотически лучше оценки с ядром $k^*(x)$, если хотя бы одно из неравенств в (70) будет строгим. Оценка с ядром $k^*(x)$ асимптотически допустима, если не существует таких $k(x)$, для которых бы выполнялось (70) и при этом хотя бы одно из неравенств в (70) было строгим. В пяти примерах оценок с $q = 2$, рассмотренных выше, ни одна из оценок не является асимптотически лучше какой-либо другой из них. См. табл. 9.3.4.

Теория, рассмотренная в этом параграфе, является асимптотической. Поэтому естественно ожидать, что для хорошей ее применимости значения T должны быть достаточно большими.

Для малых и умеренных значений T окна $w_T(\lambda | v)$ и $w_T^*(\lambda | v)$ приписывают положительные веса как значениям, близким к v , из-за наличия основного пика, так и значениям, удаленным от v , из-за наличия боковых лепестков. Область тех значений λ , кото-

Таблица 9.3.4

ХАРАКТЕРИСТИКИ ЯДЕР, ВХОДЯЩИХ В СРЕДНЕКВАДРАТИЧНЫЕ ОШИБКИ ОЦЕНОК

	k	k^3	$\int_{-1}^1 k^2(x) dx$
Е. Усеченная оценка Даниэля ($b = \pi/K_T$)	$\pi^2/6$	$\pi^4/36 = 2.7058$	0.90282336
Ф. Оценка Блэкмена — Тьюки	$\pi^2 a$	$\pi^4 a^2$	$2(1-4a+6a^2)$
Г. Окно Хеннинга	$\pi^2/4$	$\pi^4/16 = 6.0881$	$3/4 = 0.75$
Н. Окно Хемминга	$0.23\pi^2$	$0.0529\pi^4 = 5.1529$	0.7948
И. Оценка Парзена	1	1	$16/15 = 1.0667$
Ж. Оценка Парзена	6	36	$151/280 = 0.5393$

рым приписывается значительный вес, называется шириной полосы частот спектрального окна. По этому поводу имеется целый ряд различных определений в которые, однако, мы не будем углубляться. См. Дженкинс (1961) и Парзен (1961b).

9.4. АСИМПТОТИЧЕСКАЯ НОРМАЛЬНОСТЬ ОЦЕНОК СПЕКТРАЛЬНОЙ ПЛОТНОСТИ

При определенных условиях оценки спектральной плотности, изучавшиеся в § 9.3, оказываются асимптотически нормальными. Мы покажем, что величина $(T/K_T)^{1/2} [\hat{f}_T(\nu) - \mathcal{E}\hat{f}_T(\nu)]$ имеет в пределе нормальное распределение. Положим

$$(1) \quad y_t = \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s},$$

где $\sum_{s=-\infty}^{\infty} |\gamma_s| < \infty$, а $\{v_t\}$ — последовательность независимых и одинаково распределенных случайных величин, у которых $\mathcal{E}v_t = 0$, $\mathcal{E}v_t^2 = \sigma^2$ и $\mathcal{E}v_t^4 = 3\sigma^4 + \kappa_4 < \infty$. Найдем предельное распределение разности $U_T - \mathcal{E}U_T$, где

$$(2) \quad U_T = \left(\frac{T}{K_T}\right)^{1/2} \sum_{g=1}^{K_T} k\left(\frac{g}{K_T}\right) \cos \nu g \frac{1}{T} \sum_{t=1}^{T-g} y_t y_{t+g}.$$

Разность между $(T/K_T)^{1/2} [\hat{f}_T(\nu) - \mathcal{E}\hat{f}_T(\nu)]$ и $(U_T - \mathcal{E}U_T)/\pi$ равна

$$(3) \quad \frac{k(0)}{2\pi} \frac{1}{\sqrt{TK_T}} \sum_{t=1}^T [y_t^2 - \sigma(0)]$$

и имеет дисперсию, стремящуюся к 0 при $T \rightarrow \infty$. Пусть

$$(4) \quad y_{t,n} = \sum_{s=-n}^n \gamma_s v_{t-s},$$

$$(5) \quad u_{t,n} = \sum_{|s|>n} \gamma_s v_{t-s},$$

$$(6) \quad U_{Tn} = \left(\frac{T}{K_T}\right)^{1/2} \sum_{g=1}^{K_T} k\left(\frac{g}{K_T}\right) \cos \nu g \frac{1}{T} \sum_{t=1}^{T-g} y_{t,n} y_{t+g,n}^*.$$

Тогда

$$(7) \quad U_T - U_{Tn} = \frac{1}{\sqrt{TK_T}} \sum_{g=1}^{K_T} k\left(\frac{g}{K_T}\right) \cos \nu g \times \\ \times \sum_{t=1}^{T-g} (y_t y_{t+g} - y_{t,n} y_{t+g,n}) = S_1 + S_2 + S_3,$$

где

$$(8) \quad S_1 = \frac{1}{\sqrt{TK_T}} \sum_{g=1}^{K_T} k\left(\frac{g}{K_T}\right) \cos \nu g \sum_{t=1}^{T-g} u_{t,n} y_{t+g,n},$$

$$(9) \quad S_2 = \frac{1}{\sqrt{TK_T}} \sum_{g=1}^{K_T} k\left(\frac{g}{K_T}\right) \cos \nu g \sum_{t=1}^{T-g} y_{t,n} u_{t+g,n},$$

$$(10) \quad S_3 = \frac{1}{\sqrt{TK_T}} \sum_{g=1}^{K_T} k\left(\frac{g}{K_T}\right) \cos \nu g \sum_{t=1}^{T-g} u_{t,n} u_{t+g,n}.$$

Положим $\gamma'_s = \gamma_s$ для $|s| \leq n$, $\gamma'_s = 0$ для $|s| > n$ и $\gamma_s^* = 0$ для $|s| \leq n$, $\gamma_s^* = \gamma_s$ для $|s| > n$. Тогда дисперсию величины S_3 можно представить в виде

$$(11) \quad \frac{1}{TK_T} \sum_{g,h=1}^{K_T} k\left(\frac{g}{K_T}\right) k\left(\frac{h}{K_T}\right) \cos \nu g \cos \nu h \times \\ \times \sum_{s=1}^{T-g} \sum_{t=1}^{T-h} \text{Cov}(u_{s,n} u_{s+g,n}, u_{t,n} u_{t+h,n}) = \\ = \frac{1}{TK_T} \sum_{g,h=1}^{K_T} k\left(\frac{g}{K_T}\right) k\left(\frac{h}{K_T}\right) \cos \nu g \cos \nu h \times \\ \times \sum_{s=1}^{T-g} \sum_{t=1}^{T-h} \sum_{p,q,r,m=-\infty}^{\infty} \gamma_p^* \gamma_q^* \gamma_r^* \gamma_m^* \text{Cov}(v_{s-p} v_{s+g-q}, v_{t-r} v_{t+h-m}) =$$

$$\begin{aligned}
&= \frac{1}{TK_T} \sum_{g,h=1}^{K_T} k\left(\frac{g}{K_T}\right) k\left(\frac{h}{K_T}\right) \cos vg \cos vh \times \\
&\quad \times \sum_{s=1}^{T-g} \sum_{t=1}^{T-h} \left[\sigma^4 \sum_{p,q=-\infty}^{\infty} \gamma_p^* \gamma_{p-s+t}^* \gamma_q^* \gamma_{q-s+t+h-g}^* + \right. \\
&\quad + \sigma^4 \sum_{p,q=-\infty}^{\infty} \gamma_p^* \gamma_{p-s+t+h}^* \gamma_q^* \gamma_{q-s+t-g}^* + \\
&\quad \left. + \kappa_4 \sum_{p=-\infty}^{\infty} \gamma_p^* \gamma_{p+g}^* \gamma_{p-s+t}^* \gamma_{p-s+t+h}^* \right].
\end{aligned}$$

Из этого выражения видно, что дисперсия слагаемого S_3 ограничена сверху величиной

$$\begin{aligned}
(12) \quad K_T^{-1} \sup_{-1 \leq x \leq 1} k^2(x) \cdot \sum_{g,h=1}^{K_T} \left[\sigma^4 \sum_{p,q,r=-\infty}^{\infty} |\gamma_p^*| \cdot |\gamma_{p+r}^*| \cdot |\gamma_q^*| \cdot |\gamma_{q+r+h-g}^*| + \right. \\
+ \sigma^4 \sum_{p,q,r=-\infty}^{\infty} |\gamma_p^*| \cdot |\gamma_{p+r+h}^*| \cdot |\gamma_q^*| \cdot |\gamma_{q+r-g}^*| + \\
\left. + \kappa_4 \sum_{p,r=-\infty}^{\infty} |\gamma_p^*| \cdot |\gamma_{p+g}^*| \cdot |\gamma_{p+r}^*| \cdot |\gamma_{p+r+h}^*| \right] \leq \\
\leq \sup_{-1 \leq x \leq 1} k^2(x) (2\sigma^4 + \kappa_4) \sum_{p,q,r,m=-\infty}^{\infty} |\gamma_p^*| \cdot |\gamma_q^*| \cdot |\gamma_r^*| \cdot |\gamma_m^*| = \\
= \sup_{-1 \leq x \leq 1} k^2(x) (2\sigma^4 + \kappa_4) \left(\sum_{|s| > n} |\gamma_s| \right)^4,
\end{aligned}$$

которая не зависит от T и стремится к нулю при $n \rightarrow \infty$. Подобным же образом,

$$\begin{aligned}
(13) \quad \text{Var } S_1 = \frac{1}{TK_T} \sum_{g,h=1}^{K_T} k\left(\frac{g}{K_T}\right) k\left(\frac{h}{K_T}\right) \cos vg \cos vh \times \\
\times \sum_{s=1}^{T-g} \sum_{t=1}^{T-h} \sum_{p,q,r,m=-\infty}^{\infty} \gamma_p^* \gamma_q^* \gamma_r^* \gamma_m^* \text{Cov}(v_{s-p} v_{s+g-q}, v_{t-r} v_{t+h-m}) \leq \\
\leq \sup_{-1 \leq x \leq 1} k^2(x) (2\sigma^4 + \kappa_4) \left(\sum_{|s| \leq n} |\gamma_s| \right)^2 \left(\sum_{|s| > n} |\gamma_s| \right)^2.
\end{aligned}$$

Этой величиной ограничена сверху и дисперсия слагаемого S_2 . Используя полученные соотношения, неравенство

$$\begin{aligned}
(14) \quad \mathbb{E}[(U_T - \mathbb{E}U_T) - (U_{Tn} - \mathbb{E}U_{Tn})]^2 = \text{Var}(S_1 + S_2 + S_3) \leq \\
\leq 3(\text{Var } S_1 + \text{Var } S_2 + \text{Var } S_3)
\end{aligned}$$

и вследствие 7.7.1, получаем, что предельное распределение разности $U_T - \mathbb{E}U_T$ является пределом по n предельных распределений разностей $U_{Tn} - \mathbb{E}U_{Tn}$.

Что касается U_{Tn} , то эта величина является действительной частью суммы

$$\begin{aligned}
 (15) \quad & \frac{1}{\sqrt{TK_T}} \sum_{g=1}^{K_T} k\left(\frac{g}{K_T}\right) e^{ivg} \sum_{t=1}^{T-g} y_{t,n} y_{t+g,n} = \\
 & = \frac{1}{\sqrt{TK_T}} \sum_{g=1}^{K_T} k\left(\frac{g}{K_T}\right) e^{ivg} \sum_{t=1}^{T-g} \sum_{r,s=-n}^n \gamma_s \gamma_r v_{t-s} v_{t+g-r} = \\
 & = \frac{1}{\sqrt{TK_T}} \sum_{g=1}^{K_T} k\left(\frac{g}{K_T}\right) e^{ivg} \sum_{r,s=-n}^n \sum_{q=1-s}^{T-g-s} \gamma_s \gamma_r v_q v_{q+g+s-r} = \\
 & = \frac{1}{\sqrt{TK_T}} \sum_{r,s=-n}^n \sum_{h=s-r+1}^{K_T+s-r} \sum_{q=1-s}^{T-h-r} k\left(\frac{h+r-s}{K_T}\right) e^{ivh} e^{ivr} \gamma_r e^{-ivs} \gamma_s v_q v_{q+h} = \\
 & = \frac{1}{\sqrt{TK_T}} \sum_{r,s=-n}^n \gamma_r e^{ivr} \gamma_s e^{-ivs} \times \\
 & \times \sum_{h=s-r+1}^{K_T+s-r} k\left(\frac{h+r-s}{K_T}\right) e^{ivh} \sum_{q=1-s}^{T-h-r} v_q v_{q+h}.
 \end{aligned}$$

В то же время, математическое ожидание квадрата разности между действительными частями (15) и

$$\begin{aligned}
 (16) \quad & \frac{1}{\sqrt{TK_T}} \sum_{r,s=-n}^n \gamma_r e^{ivr} \gamma_s e^{-ivs} \times \\
 & \times \sum_{h=1}^{K_T-2n} k\left(\frac{h+r-s}{K_T}\right) e^{ivh} \sum_{q=1}^{T-h} v_q v_{q+h}
 \end{aligned}$$

стремится к 0 при $T \rightarrow \infty$. Действительно, для фиксированных r и s разность между суммируемыми величинами в (15) и (16) состоит из всех тех членов сумм по h и по q , которые входят в одно выражение и не входят в другое. Число таких членов не превосходит $AK_T n + + BTn + Cn^2$, где A , B и C — вполне определенные константы (см. упр. 27 и 28), сами они некоррелированы (см. упр. 29), а математические ожидания квадратов их действительных частей не превосходят

$$(17) \quad \left(\max_{-n \leq r \leq n} |\gamma_r| \right)^4 \sup_{-1 \leq x \leq 1} k^2(x) \sigma^4 / (TK_T).$$

Отсюда следует, что математическое ожидание квадрата соответствующей разности для каждого r и s стремится к нулю при $T \rightarrow \infty$.

Если функция $k(x)$ непрерывна на $[-1, 1]$, то величина $k[(h+r-s)/K_T]$ для достаточно больших K_T и $|r| \leq n$, $|s| \leq n$, $|h| \leq K_T$, $|h+r-s| \leq K_T$ будет сколь угодно близка к значению $k(h/K_T)$. Поэтому разность действительных частей (16) и

$$\begin{aligned} (18) \quad \frac{1}{\sqrt{TK_T}} f_{(n)}(v) \sum_{h=1}^{K_T-2n} k\left(\frac{h}{K_T}\right) e^{ivh} \sum_{q=1}^{T-h} v_q v_{q+h} = \\ = \frac{f_{(n)}(v)}{\sqrt{TK_T}} \sum_{q=1}^{T-1} \sum_{h=1}^{\min(K_T-2n, T-q)} k\left(\frac{h}{K_T}\right) e^{ivh} v_q v_{q+h}, \end{aligned}$$

где $f_{(n)}(v) = \left| \sum_{r=-n}^n \gamma_r e^{iv_r} \right|^2$, также имеет сколь угодно малое среднеквадратичное отклонение. В свою очередь при $T \rightarrow \infty$ стремится к нулю и среднеквадратичное отклонение разности действительных частей (18) и

$$\begin{aligned} (19) \quad \frac{f_{(n)}(v)}{\sqrt{T}} \sum_{q=1}^T \frac{1}{\sqrt{K_T}} \sum_{h=1}^{K_T} k\left(\frac{h}{K_T}\right) \cos vh v_q v_{q+h} = \\ = f_{(n)}(v) \frac{1}{\sqrt{T}} \sum_{q=1}^T W_{qT}, \end{aligned}$$

где

$$(20) \quad W_{qT} = \frac{1}{\sqrt{K_T}} \sum_{h=1}^{K_T} k\left(\frac{h}{K_T}\right) \cos vh v_q v_{q+h}, \quad q = 1, \dots, T.$$

Процесс $\{W_{qT}\}$ является стационарным процессом с конечной зависимостью. Более того,

$$(21) \quad \mathbb{E} W_{qT} = 0,$$

$$\begin{aligned} (22) \quad \mathbb{E} W_{qT}^2 &= \frac{1}{K_T} \sum_{g,h=1}^{K_T} k\left(\frac{g}{K_T}\right) k\left(\frac{h}{K_T}\right) \cos vg \cos vh \mathbb{E} v_q^2 v_{q+g} v_{q+h} = \\ &= \frac{\sigma^4}{K_T} \sum_{h=1}^{K_T} k^2\left(\frac{h}{K_T}\right) \cos^2 vh = \\ &= \frac{\sigma^4}{2} \sum_{h=1}^{K_T} k^2\left(\frac{h}{K_T}\right) (1 + \cos 2vh) / K_T, \end{aligned}$$

$$(23) \quad \mathfrak{W}_{qT} \mathfrak{W}_{q+r,T} = \frac{1}{K_T} \sum_{g,h=1}^{K_T} k\left(\frac{g}{K_T}\right) k\left(\frac{h}{K_T}\right) \times \\ \times \cos v g \cos v h \mathfrak{E} v_q v_{q+g} v_{q+r} v_{q+h+r}, \quad r \neq 0.$$

В силу этого дисперсия выражения (19) равна умноженной на $f_{(n)}^2(v)$ величине (22). Последняя же при $v \neq 0, \pm \pi$ имеет предел

$$(24) \quad \lim_{K_T \rightarrow \infty} \frac{\sigma^4}{2} \sum_{h=1}^{K_T} k^2\left(\frac{h}{K_T}\right) \frac{1}{K_T} = \frac{\sigma^4}{2} \int_0^1 k^2(x) dx,$$

а при $v = 0$ или $v = \pm \pi$ — предел, равный удвоенному значению (24).

Пусть $\{N_T\}$ — еще одна последовательность целых чисел, причем $K_T/N_T \rightarrow 0$ и $N_T/T \rightarrow 0$ при $T \rightarrow \infty$. (Такой, например, является последовательность $N_T = [\sqrt{T K_T}]$.) Пусть M_T — наибольшее целое в ряду T/N_T . Положим

$$(25) \quad Z_{jT} = \frac{1}{\sqrt{N_T}} [\mathfrak{W}_{(j-1)N_T+1,T} + \dots + \mathfrak{W}_{jN_T-K_T,T}], \\ j = 1, \dots, M_T.$$

Тогда $Z_{1T}, \dots, Z_{M_T T}$ — независимые одинаково распределенные случайные величины с $\mathfrak{E} Z_{jT} = 0$ и $\mathfrak{E} Z_{jT}^2$, равным умноженному на $1 - K_T/N_T$ выражению (22). Более того, четвертый момент

$$(26) \quad \mathfrak{E} Z_{1T}^4 = \frac{1}{N_T^2} \sum_{t,s,r,q=1}^{N_T-K_T} \mathfrak{E} W_{tT} W_{sT} W_{rT} W_{qT}$$

оказывается равномерно ограниченным. Действительно, хотя W_{iT} квадратично зависит от переменных v_r , тем не менее все четвертые моменты величин W_{iT} ограничены, так как они включают в себя моменты случайных величин v_r лишь до четвертого порядка. Это связано с тем, что в каждом W_{iT} произведения переменных v_r имеют различные индексы r и что случайные величины v_r независимы. Поскольку $\mathfrak{E} W_{iT} W_{sT} W_{rT} W_{qT} = 0$, если наименьший из индексов отличен от трех остальных, то в (26) имеет смысл рассматривать только слагаемые вида $\mathfrak{E} W_{iT}^4$, $\mathfrak{E} W_{iT}^3 W_{sT}$, $s > t$, и $\mathfrak{E} W_{iT}^2 W_{sT} W_{qT}$, $s > t$, $q > t$. При этом количество слагаемых вида $\mathfrak{E} W_{iT}^4$ равно $N_T - K_T$; количество слагаемых вида $\mathfrak{E} W_{iT}^3 W_{sT}$ ($s > t$) — самое большее $2(N_T - K_T)(N_T - K_T - 1)$; количество слагаемых вида $\mathfrak{E} W_{iT}^2 W_{sT} W_{qT}$ ($s > t$) — не больше $3(N_T - K_T)(N_T - K_T - 1)$. Что касается слагаемых вида $\mathfrak{E} W_{iT}^2 W_{sT} W_{qT}$, $s > t$, $q > t$, $s \neq q$, то они равны нулю, если $|s - q| > K_T$. Поэтому число ненулевых слагаемых такого вида не превосходит $6(N_T - K_T)(N_T - K_T - 1) K_T$. Возникающий здесь бесконечно возрастающий множитель

K_T компенсируется тем, что $|\mathfrak{E} W_{iT}^2 W_{sT} W_{qT}| \leq 4\sigma^8 \sup_{-1 \leq x \leq 1} k^4(x)/K_T$.

Слагаемые остальных типов равномерно ограничены по T (см. упр. 31). Таким образом, четвертый момент $\mathfrak{E} Z_{iT}^4$ оказывается равномерно ограниченным по T . Применяя центральную предельную теорему Ляпунова, получаем, что нормированная сумма

$$(27) \quad \frac{1}{\sqrt{M_T}} \sum_{j=1}^{M_T} Z_{jT}$$

имеет в пределе нормальное распределение с нулевым средним и дисперсией, равной (24) при $v \neq 0, \pm \pi$ и в два раза превышающей (24) при $v = 0, \pm \pi$. (Для этого достаточно в теореме 7.7.3 взять в качестве w_i^T отношение $Z_{iT}/[M_T \mathfrak{E} Z_{iT}^2]^{1/2}$ и положить $\delta = 2$.) Поскольку же среднеквадратичное отклонение разности

$$(28) \quad \frac{1}{\sqrt{T}} \sum_{i=1}^T W_{iT} - \frac{1}{\sqrt{M_T}} \sum_{j=1}^{M_T} Z_{jT}$$

стремится к нулю при $T \rightarrow \infty$, то такое же предельное распределение будет иметь и $\sum_{i=1}^T W_{iT}/\sqrt{T}$.

ТЕОРЕМА 9.4.1. Пусть

$$(29) \quad \hat{f}_T(v) = \frac{1}{2\pi} \sum_{r=-K_T}^{K_T} k\left(\frac{r}{K_T}\right) \cos vr \, c_{rT},$$

где $k(x) = k(-x)$, функция $k(x)$ непрерывна на $[-1, 1]$ и $\{K_T\}$ — целочисленная последовательность, такая, что $K_T \rightarrow \infty$ и $K_T/T \rightarrow 0$ при $T \rightarrow \infty$. Пусть

$$(30) \quad y_t = \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s},$$

где $\sum_{s=-\infty}^{\infty} |\gamma_s| < \infty$, а $\{v_t\}$ — последовательность независимых и одинаково распределенных случайных величин с $\mathfrak{E} v_t = 0$, $\mathfrak{E} v_t^2 = \sigma^2$ и $\mathfrak{E} v_t^4 < \infty$. Тогда $\sqrt{T/K_T} [\hat{f}_T(v) - \mathfrak{E} \hat{f}_T(v)]$ имеет в пределе нормальное распределение с дисперсией, указанной в теореме 9.3.4.

Асимптотически нормальным является и совместное распределение оценок $\hat{f}_T(v_1), \dots, \hat{f}_T(v_n)$ для любого фиксированного числа значений v .

Условия теоремы 9.4.1 можно заменить другими, в которых будет отсутствовать требование $\mathfrak{E} v_t^4 < \infty$. Так, например, Чон и

Хеннан (1968) доказали асимптотическую нормальность в предположении, что $\sum_{i=1}^T (v_i^2 - \sigma^2) / \sqrt{TK_T}$ сходится по вероятности к нулю и $\sum_{s=-\infty}^{\infty} |s| \gamma_s^2 < \infty$. Другие условия нормальности указаны Розенблаттом (1959).

Рассмотрим теперь нормированное отклонение

$$(31) \quad \sqrt{\frac{T}{K_T}} [\hat{f}_T(v) - f(v)] = \sqrt{\frac{T}{K_T}} [\hat{f}_T(v) - \mathcal{E}\hat{f}_T(v)] + \sqrt{\frac{T}{K_T}} [\mathcal{E}\hat{f}_T(v) - f(v)].$$

Пусть p и q — величины, определенные в теореме 9.3.3. Если $p < q$ и $K_T/T \rightarrow 0$ при $p \leq 1$ или $K_T^p/T \rightarrow 0$ при $p \geq 1$, то второе слагаемое в правой части (31) стремится к нулю, если только предел $\lim_{T \rightarrow \infty} T/K_T^{2p+1}$ конечен. Если же $q \leq p$ и $K_T^{q+1-p}/T \rightarrow 0$ при $p \leq 1$ или $K_T^q/T \rightarrow 0$ при $p \geq 1$, то оно будет стремиться к нулю, если $T/K_T^{2q+1} \rightarrow 0$. Как бы то ни было, в обоих этих случаях предельное распределение левой части (31) совпадает с предельным распределением первого члена правой части (31).

Следствие 9.4.1. Если выполнены условия теорем 9.4.1 и 9.3.3 и если предел $\lim_{T \rightarrow \infty} T/K_T^{2p+1}$ конечен при $p < q$ или $T/K_T^{2q+1} \rightarrow 0$ при

$q \leq p$ и $T \rightarrow \infty$, то $\sqrt{T/K_T} [\hat{f}_T(v) - f(v)]$ имеет в пределе нормальное распределение с нулевым средним и дисперсией, указанной в теореме 9.3.4.

Отметим, что для того, чтобы смещение было пренебрежимо малым по сравнению со случайной частью (31), последовательность K_T должна возрастать быстрее, нежели в том случае, когда дисперсия и квадрат смещения имеют один и тот же порядок. Как и ранее, при указанных условиях оценки $\hat{f}_T(v_1), \dots, \hat{f}_T(v_n)$ имеют асимптотически нормальное совместное распределение для любого фиксированного числа значений v .

Если обозначить $\tau^2 = \int_{-1}^1 k(x) dx$ ($v \neq 0, \pm \pi$), то, перефразируя следствие 9.4.1, получим, что для $f(v) > 0$

$$(32) \quad \sqrt{\frac{T}{K_T}} \frac{\hat{f}_T(v) - f(v)}{\tau f(v)} = \sqrt{\frac{T}{K_T}} \frac{1}{\tau} \left[\frac{\hat{f}_T(v)}{f(v)} - 1 \right]$$

имеет стандартное нормальное распределение с нулевым средним и единичной дисперсией. Пусть число $t(\varepsilon)$ таково, что вероятность попадания стандартной нормальной случайной величины на отрезок $[-t(\varepsilon), t(\varepsilon)]$ равна $1 - \varepsilon$. Тогда событие, состоящее в том, что (32) попадает на этот отрезок, можно записать в виде

$$(33) \quad \frac{\hat{f}_T(v)}{1 + t(\varepsilon) \sqrt{K_T/T}} \leq f(v) \leq \frac{\hat{f}_T(v)}{1 - t(\varepsilon) \sqrt{K_T/T}}, \quad v \neq 0, \pm \pi.$$

Последнее соотношение определяет для $f(v)$ доверительный интервал с уровнем доверия, приблизительно равным $1 - \varepsilon$ для больших T (в условиях следствия 9.4.1).

Мы можем воспользоваться также следующим утверждением. Если $Y_T = g(X_T)$, где $p \lim_{T \rightarrow \infty} X_T = \mu$, $\beta_T(X_T - \mu)$ имеет при $T \rightarrow \infty$ предельное нормальное распределение с нулевым средним и дисперсией σ^2 ($\beta_T \rightarrow \infty$), а $g(x)$ имеет производную $g'(\mu)$ в точке $x = \mu$, то $\beta_T[Y_T - g(\mu)]$ имеет в пределе нормальное распределение с нулевым средним и дисперсией $\sigma^2 [g'(\mu)]^2$. Если $X_T = \hat{f}_T(v)$ и $g(x) = \log x$, то $\mu = f(v)$, $g(\mu) = \log f(v)$ и $g'(\mu) = 1/f(v)$ при $f(v) > 0$.

ТЕОРЕМА 9.4.2. В условиях следствия 9.4.1 величина

$$(34) \quad \sqrt{\frac{T}{K_T}} [\log \hat{f}_T(v) - \log f(v)] = \sqrt{\frac{T}{K_T}} \log \frac{\hat{f}_T(v)}{f(v)}$$

при $f(v) > 0$ имеет предельное нормальное распределение с нулевым средним и дисперсией, равной $\tau^2 = \int_{-1}^1 k^2(x) dx$ при $v \neq 0, \pm \pi$ и равной $2\tau^2$ при $v = 0, \pm \pi$.

Сформулированный результат приводит к доверительным интервалам

$$(35) \quad \log \hat{f}_T(v) - t(\varepsilon) \sqrt{\frac{K_T}{T}} \leq \log f(v) \leq \log \hat{f}_T(v) + t(\varepsilon) \sqrt{\frac{K_T}{T}},$$

или

$$(36) \quad \hat{f}_T(v) e^{-t(\varepsilon) \sqrt{K_T/T}} \leq f(v) \leq \hat{f}_T(v) e^{t(\varepsilon) \sqrt{K_T/T}}.$$

Из теоремы 9.4.2 можно заключить, что информативным графическим представлением оценки спектральной плотности является представление в виде зависимости от v ее логарифма $\log \hat{f}_T(v)$, по-

скольку в этом случае (если $\nu \neq 0, \pm \pi$) асимптотическое стандартное отклонение является константой и притом известной, именно $\tau \sqrt{K_T/T}$.

Блэкмен и Тьюки (1959, стр. 22) предложили аппроксимировать распределение оценки $\hat{f}_T(\nu)$, являющейся квадратичной формой от наблюдений, распределением случайной величины $\beta f(\nu) \chi_\alpha^2$, где χ_α^2 — случайная величина, имеющая χ^2 -распределение с α степенями свободы, а α и β выбираются так, чтобы первые два момента величины $\beta f(\nu) \chi_\alpha^2$ были равны соответствующим приближенным моментам для $\hat{f}_T(\nu)$. Таким образом, при $\nu \neq 0, \pm \pi$ возникают условия

$$(37) \quad \beta\alpha = 1,$$

$$(38) \quad 2\beta^2\alpha = \frac{K_T}{T} \int_{-1}^1 k^2(x) dx.$$

Тогда

$$(39) \quad \alpha = \frac{1}{\beta} = 2T \left[K_T \int_{-1}^1 k^2(x) dx \right]^{-1},$$

и для $f(\nu) > 0$ в качестве случайной величины, имеющей χ^2 -распределение с числом степеней свободы α , указанным в (39), следует взять

$$(40) \quad 2T\hat{f}_T(\nu) \left[K_T f(\nu) \int_{-1}^1 k^2(x) dx \right]^{-1}.$$

Если среднее значение μ неизвестно, то оценки для $f(\lambda)$ можно построить с использованием $\hat{c}_{rT}^*$ или $\tilde{c}_{rT}$. Положим

$$(41) \quad \hat{f}_T^*(\nu) = \frac{1}{2\pi} \sum_{r=-K_T}^{K_T} k\left(\frac{r}{K_T}\right) \cos \nu r \hat{c}_{rT}^*.$$

Пусть $m = \min(p, q)$. Тогда

$$\begin{aligned} (42) \quad K_T^m [\hat{\mathcal{G}}\hat{f}_T^*(\nu) - f(\nu)] &= \\ &= K_T^m [\hat{\mathcal{G}}\hat{f}_T(\nu) - f(\nu)] + \\ &+ \frac{K_T^m}{2\pi} \sum_{r=-K_T}^{K_T} k\left(\frac{r}{K_T}\right) \cos \nu r (\hat{c}_{rT}^* - c_{rT}) = \\ &= K_T^m [\hat{\mathcal{G}}\hat{f}_T(\nu) - f(\nu)] + \\ &+ \frac{1}{2\pi} \frac{K_T^m}{T} \sum_{r=-K_T}^{K_T} k\left(\frac{r}{K_T}\right) \cos \nu r T \mathcal{G}(c_{rT}^* - c_{rT}). \end{aligned}$$

Второе слагаемое в правой части (42) не превосходит по модулю

$$(43) \quad \frac{1}{2\pi} \frac{K_T^m}{T} \sup_{-1 \leq x \leq 1} |k(x)| \sum_{r=-K_T}^{K_T} |T \mathfrak{E}(c_{rT}^* - c_{rT})|.$$

Из теоремы 8.3.2 следует, что если $\sum_{r=-\infty}^{\infty} |\sigma(r)| < \infty$, то $|T(\mathfrak{E}_{rT}^* - \mathfrak{E}_{rT})| \leq 3 \sum_{s=-\infty}^{\infty} |\sigma(s)|$ для всех T и r . Поэтому (43) не превосходит

$$(44) \quad \frac{3}{2\pi} \sup_{-1 \leq x \leq 1} |k(x)| \sum_{r=-\infty}^{\infty} |\sigma(r)| \frac{K_T^m}{T} (2K_T + 1).$$

Последняя же величина стремится к 0, если только $K_T^{m+1}/T \rightarrow 0$ при $T \rightarrow \infty$.

ТЕОРЕМА 9.4.3. Пусть $\hat{f}_T^*(v)$ определяется соотношением (41), причем $k(x) = k(-x)$, $k(0) = 1$, $|k(x)| \leq M$ для некоторого M и всех $|x| \leq 1$, и $\lim_{x \rightarrow 0} [1 - k(x)]/|x|^q = k$ для некоторых $q > 0$ и $k > 0$. Пусть $\{K_T\}$ — целочисленная последовательность, такая, что $K_T \rightarrow \infty$ при $T \rightarrow \infty$. Тогда, если

$$(45) \quad \sum_{r=-\infty}^{\infty} |r|^p |\sigma(r)| < \infty$$

для $p \geq q$ и $p \geq 1$ и если $K_T^{q+1}/T \rightarrow 0$ при $T \rightarrow \infty$, то

$$(46) \quad \lim_{T \rightarrow \infty} K_T^q [\hat{\mathfrak{E}}_{f_T}^*(v) - f(v)] = -\frac{k}{2\pi} \sum_{r=-\infty}^{\infty} |r|^q \cos vr \sigma(r).$$

Если же (45) выполняется для $1 \leq p < q$ и $K_T^{p+1}/T \rightarrow 0$ при $T \rightarrow \infty$, то

$$(47) \quad \lim_{T \rightarrow \infty} K_T^p [\hat{\mathfrak{E}}_{f_T}^*(v) - f(v)] = 0.$$

Ковариации оценок $\hat{f}_T^*(v)$ удовлетворяют соотношению

$$(48) \quad \frac{T}{K_T} \text{Cov} [\hat{f}_T^*(\lambda), \hat{f}_T^*(v)] = \\ = \frac{1}{(2\pi)^2 K_T} \sum_{g, h=-K_T}^{K_T} k\left(\frac{g}{K_T}\right) k\left(\frac{h}{K_T}\right) \cos \lambda g \cos v h T \text{Cov}(c_{gT}^*, c_{hT}^*).$$

Отсюда для $2K_T < T$ получаем, что

$$(49) \quad \left| \frac{T}{K_T} \text{Cov} [\hat{f}_T^*(\lambda), \hat{f}_T^*(\nu)] - \frac{T}{K_T} \text{Cov} [\hat{f}_T(\lambda), \hat{f}_T(\nu)] \right| \leq \\ \leq \frac{1}{(2\pi)^2 K_T} \sup_{-1 \leq x \leq 1} k^2(x) \sum_{g, h = -K_T}^{K_T} |T \text{Cov}(c_{gT}^*, c_{hT}^*) - T \text{Cov}(c_{gT}, c_{hT})|.$$

Возьмем для удобства $\mu = 0$. Поскольку

$$(50) \quad c_{gT}^* = c_{gT} - \frac{1}{T^2} \sum_{i=1}^T y_i \left(\sum_{s=g+1}^T y_s + \sum_{s=1}^{T-g} y_s - \frac{T-g}{T} \sum_{s=1}^T y_s \right) = \\ = c_{gT} - \frac{1}{T^2} \sum_{i=1}^T \sum_{s=g+1}^{T-g} y_i y_s - \frac{g}{T^3} \sum_{i,s=1}^T y_i y_s, \quad 0 \leq 2g \leq T-1,$$

рассмотрим для $0 \leq h \leq T-1$, $0 \leq 2g \leq T-1$

$$(51) \quad T \text{Cov} \left(c_{hT}, \frac{1}{T^2} \sum_{i=1}^T \sum_{s=g+1}^{T-g} y_i y_s + \frac{g}{T^3} \sum_{i,s=1}^T y_i y_s \right) = \\ = \frac{1}{T^2} \sum_{r=1}^{T-h} \sum_{t=1}^T \left\{ \sum_{s=g+1}^{T-g} \mathcal{E} [y_r y_{r+h} - \sigma(h)] [y_t y_s - \sigma(t-s)] + \right. \\ \left. + \frac{g}{T} \sum_{s=1}^T \mathcal{E} [y_r y_{r+h} - \sigma(h)] [y_t y_s - \sigma(t-s)] \right\} = \\ = \frac{1}{T^2} \sum_{r=1}^{T-h} \sum_{t=1}^T \left\{ \sum_{s=g+1}^{T-g} [\sigma(r-t) \sigma(r+h-s) + \right. \\ \left. + \sigma(r-s) \sigma(r+h-t) + \kappa(h, t-r, s-r)] + \right. \\ \left. + \frac{g}{T} \sum_{s=1}^T [\sigma(r-t) \sigma(r+h-s) + \sigma(r-s) \sigma(r+h-t) + \right. \\ \left. + \kappa(h, t-r, s-r)] \right\}.$$

Это выражение по абсолютной величине не превосходит

$$(52) \quad \frac{2}{T} \left\{ 2 \left[\sum_{t=-\infty}^{\infty} |\sigma(t)| \right]^2 + \sum_{r,s,t=-\infty}^{\infty} |\kappa(r, s, t)| \right\}.$$

В теореме 9.3.4 предполагается конечность входящих в (52) сумм. Рассмотрим еще

$$(53) \quad T \text{Cov} \left[\frac{1}{T^2} \sum_{i=1}^T \left(\sum_{s=g+1}^{T-g} y_i y_s + \frac{g}{T} \sum_{s=1}^T y_i y_s \right), \frac{1}{T^2} \sum_{r=1}^T \left(\sum_{q=h+1}^{T-h} y_r y_q + \right. \right. \\ \left. \left. + \frac{h}{T} \sum_{q=1}^T y_r y_q \right) \right] =$$

$$\begin{aligned}
&= \frac{1}{T^3} \sum_{t,r=1}^T \left\{ \sum_{s=g+1}^{T-g} \sum_{q=h+1}^{T-h} \mathfrak{E} [y_t y_s - \sigma(t-s)] [y_r y_q - \sigma(r-q)] + \right. \\
&+ \frac{h}{T} \sum_{s=g+1}^{T-g} \sum_{q=1}^T \mathfrak{E} [y_t y_s - \sigma(t-s)] [y_r y_q - \sigma(r-q)] + \\
&+ \frac{g}{T} \sum_{s=1}^T \sum_{q=h+1}^{T-h} \mathfrak{E} [y_t y_s - \sigma(t-s)] [y_r y_q - \sigma(r-q)] + \\
&+ \left. \frac{gh}{T^2} \sum_{s,q=1}^T \mathfrak{E} [y_t y_s - \sigma(t-s)] [y_r y_q - \sigma(r-q)] \right\} = \\
&= \frac{1}{T^3} \sum_{t,r=1}^T \left\{ \sum_{s=g+1}^{T-g} \sum_{q=h+1}^{T-h} [\sigma(t-r) \sigma(s-q) + \right. \\
&+ \sigma(t-q) \sigma(s-r) + \kappa(s-t, r-t, q-t)] + \\
&+ \frac{h}{T} \sum_{s=g+1}^{T-g} \sum_{q=1}^T [\sigma(t-r) \sigma(s-q) + \\
&+ \sigma(t-q) \sigma(s-r) + \kappa(s-t, r-t, q-t)] + \\
&+ \frac{g}{T} \sum_{s=1}^T \sum_{q=h+1}^{T-h} [\sigma(t-r) \sigma(s-q) + \\
&+ \sigma(t-q) \sigma(s-r) + \kappa(s-t, r-t, q-t)] + \\
&+ \left. \frac{gh}{T^2} \sum_{s,q=1}^T [\sigma(t-r) \sigma(s-q) + \sigma(t-q) \sigma(s-r) + \right. \\
&+ \left. \kappa(s-t, r-t, q-t)] \right\}.
\end{aligned}$$

По абсолютной величине (53) не превосходит удвоенного значения (52). Таким образом, для $2K_T < T$ правая часть (49) не превосходит

$$(54) \quad 8 \frac{(2K_T + 1)^2}{(2\pi)^3 K_T T} \sup_{-1 \leq x \leq 1} k^2(x) \left\{ 2 \left[\sum_{t=-\infty}^{\infty} |\sigma(t)|^2 + \sum_{r,s,t=-\infty}^{\infty} |\kappa(r, s, t)| \right] \right\}.$$

ТЕОРЕМА 9.4.4. Теорема 9.3.4 справедлива для оценки $\hat{f}_T^*(v)$, определяемой соотношением (41).

Рассмотрим далее

$$\begin{aligned}
(55) \quad &\sqrt{\frac{T}{K_T}} [\hat{f}_T(v) - \hat{f}_T^*(v)] = \\
&= \sqrt{\frac{T}{K_T}} \frac{1}{2\pi} \sum_{r=-K_T}^{K_T} k\left(\frac{r}{K_T}\right) \cos vr \frac{1}{T^2} \sum_{t=1}^T y_t \left\{ \sum_{s=|r|+1}^{T-|r|} y_s + \frac{|r|}{T} \sum_{s=1}^T y_s \right\}
\end{aligned}$$

для $T > 2K_T$. Математическое ожидание квадрата (55) равно

$$\begin{aligned}
 (56) \quad & \frac{1}{(2\pi)^2} \frac{T}{K_T} \sum_{r, r'=-K_T}^{K_T} k\left(\frac{r}{K_T}\right) k\left(\frac{r'}{K_T}\right) \cos vr \cos vr' \times \\
 & \times \mathbb{E} \frac{1}{T^2} \sum_{t, t'=1}^T \left\{ \sum_{s=|r|+1}^{T-|r|} \sum_{s'=|r'|+1}^{T-|r'|} y_t y_s y_{t'} y_{s'} + \right. \\
 & + \frac{|r'|}{T} \sum_{s=|r|+1}^{T-|r|} \sum_{s'=1}^T y_t y_s y_{t'} y_{s'} + \frac{|r|}{T} \sum_{s=1}^T \sum_{s'=|r'|+1}^{T-r'} y_t y_s y_{t'} y_{s'} + \\
 & \left. + \frac{|rr'|}{T^2} \sum_{s, s'=1}^T y_t y_s y_{t'} y_{s'} \right\} \leq \\
 & \leq 4 \frac{(2K_T+1)^2}{(2\pi)^2 K_T T} \sup_{-1 \leq x \leq 1} k^2(x) \left\{ 3 \left[\sum_{t=-\infty}^{\infty} |\sigma(t)| \right]^2 + \right. \\
 & \left. + \sum_{r, s, t=-\infty}^{\infty} |\kappa(r, s, t)| \right\}.
 \end{aligned}$$

ТЕОРЕМА 9.4.5. В предположениях теоремы 9.4.1 вектор $\sqrt{T/K_T} [\hat{f}_T^*(v_1) - \hat{\mathfrak{F}}_T^*(v_1)], \dots, \sqrt{T/K_T} [\hat{f}_T^*(v_n) - \hat{\mathfrak{F}}_T^*(v_n)]$ имеет предельное нормальное распределение с нулевым средним и с дисперсиями и ковариациями, указанными в теореме 9.3.4.

В разд. 7.5.1 было показано, что спектральная плотность процесса $\{z_t\}$, определяемого соотношением

$$(57) \quad z_t = \sum_{r=0}^q c_r y_{t-r}, \quad t = \dots, -1, 0, 1, \dots,$$

равна

$$(58) \quad h(\lambda) = \left| \sum_{r=0}^q c_r e^{i\lambda r} \right|^2 f(\lambda).$$

Для некоторых целей оказывается полезным с помощью применения соотношения (57) к наблюдаемому ряду $y_1, \dots, y_T$ получать новый ряд $z_{q+1}, \dots, z_T$ и оценивать спектральную плотность $h(v)$ этого нового ряда, а затем в качестве оценки $f(v)$ использовать $\hat{h}(v) / \left| \sum_{r=0}^q c_r e^{iv r} \right|^2$. Такая конструкция скользящего среднего называется *предбелаванием*. Цель ее состоит в том, чтобы получить процесс с достаточно гладкой спектральной плотностью $h(\lambda)$, а это можно сделать, если иметь некоторую предварительную информацию о пиках и впадинах $f(\lambda)$.

Для оценки спектральной функции $G(v)$ [равной $F(v) - F(-v)$ в точках непрерывности $F(v)$] можно использовать

$$(59) \quad \hat{G}(v) = 2 \int_0^v I(\lambda) d\lambda = \frac{1}{\pi} \sum_{r=-(T-1)}^{T-1} c_r \frac{\sin vr}{r},$$

где $(\sin vr)/r = v$ при $r = 0$ (см. упр. 26).

Нормированную спектральную плотность можно оценивать посредством

$$(60) \quad \hat{f}_T(v) = \frac{\hat{f}_T(v)}{c_0} = \frac{1}{2\pi} \sum_{h=-K_T}^{K_T} k\left(\frac{h}{K_T}\right) \cos vhr_h$$

или посредством $\hat{f}_T^*(v) = \hat{f}_T(v)/c_0^*$. При этом предельное распределение

$$(61) \quad \sqrt{\frac{T}{K_T}} [\hat{f}_T(v) - \bar{f}(v)] = \\ = \sqrt{\frac{T}{K_T}} \left\{ \frac{f(v)/\sigma(0) + [\hat{f}_T(v) - f(v)]/\sigma(0)}{1 + [c_0 - \sigma(0)]/\sigma(0)} - \frac{f(v)}{\sigma(0)} \right\}$$

совпадает с предельным распределением

$$(62) \quad \sqrt{\frac{T}{K_T}} \left\{ \frac{\hat{f}_T(v) - f(v)}{\sigma(0)} - \frac{f(v)}{\sigma(0)} \frac{c_0 - \sigma(0)}{\sigma(0)} \right\}.$$

Если к тому же $\sqrt{T/K_T} [c_0 - \sigma(0)] = \sqrt{T} [c_0 - \sigma(0)]/\sqrt{K_T}$ сходится по вероятности к нулю, то предельное распределение (62) совпадает с предельным распределением $\sqrt{T/K_T} [\hat{f}_T(v) - f(v)]/\sigma(0)$.

9.5. ПРИМЕРЫ

Как было указано в § 8.5, выборочные спектральные плотности данных, приведенных в приложении, ведут себя весьма нерегулярным образом. С их помощью нельзя удовлетворительно оценить соответствующие теоретические спектральные плотности. В то же время для каждого из рассматривавшихся там случаев можно получить оценку для $f(\lambda)$ (или $\bar{f}(\lambda)$) согласно разд. 9.2.3. Вольд производил, например, оценивание величины $2\pi\bar{f}(\lambda)$ для рядов, полученных случайным моделированием, с помощью оценки Бартлетта (пример С разд. 9.2.3) при $K = 20$. Соответствующие результаты приводятся в табличном виде и графически (рис. А.2.4, А.2.5 и А.2.6). Значения теоретических нормированных спектральных плотностей (умноженных на 2π) представлены на рис. А.2.7, А.2.8 и А.2.9. Там же пред-

ставлены нормированные плотности, соответствующие выравнивающему процессу авторегрессии второго порядка. Отметим, что чем больше $\sqrt{\beta_2} = \gamma$, тем выше пик спектральной плотности $f(\lambda)$ процесса, т. е. тем более ярко выражен колебательный характер процесса.

Необходимо отметить, что при $\gamma = 0.25$ подобранная спектральная плотность очень близка к истинной. В двух других случаях максимум подобранной плотности наблюдается на той же частоте, что и максимум истинной плотности процесса, причем при $\gamma = 0.7$ согласуются и значения этих максимумов. В то же время, при $\gamma = 0.9$ подобранная плотность имеет много локальных максимумов, отсутствующих у истинной плотности. Эти особенности отражают тот факт, что чем ближе рассматриваемый процесс к последовательности некоррелированных случайных величин, имеющей плоскую спектральную плотность, тем лучше он может быть оценен.

В приложении графики спектральных плотностей приводятся в логарифмическом масштабе. Это сделано по той причине, что асимптотическое стандартное отклонение не зависит в этом случае от значения самой плотности¹⁾. Однако это приводит и к тому, что на графике становятся резко выраженными малые изменения $\hat{f}_T(\lambda)$ в области малых значений $\hat{f}_T(\lambda)$. Заметим, что окно Бартлетта (с $q = 1$) не дает оценок такой гладкости, как другие окна (с $q = 2$). Этот вопрос обсуждается далее в 9.6.

Для индекса Бевериджа цен на пшеницу с выделенным трендом оценка нормированной спектральной плотности производилась при помощи окна Блэкмена — Тьюки с $a = 0.25$ (окно Хеннинга, пример G разд. 9.2.3) для $K = 10, 20$ и 30 . Соответствующие результаты затабулированы в табл. А.1.4 и представлены графически на рис. А.1.4. Следует отметить, что чем меньше K , тем более гладкой является оценивающая функция. При $K = 10$ она имеет один локальный максимум; при $K = 20$ — два заметных локальных максимума и один плохо выраженный; при $K = 30$ оценивающая функция имеет шесть локальных максимумов. Абсолютные по величине локальные максимумы наблюдаются при $K = 20$ и $K = 30$ на частоте 0.065 (соответствующей периоду в 15.4 года), а вторые по величине локальные максимумы — на частоте порядка 0.185 (период 5.4 года). Остальные локальные максимумы при $K = 30$ наблюдаются на частотах 0.285, 0.365, 0.430 и 0.490 (соответствующими периодами являются 3.5, 2.74, 2.33 и 2.04 года).

Оценки спектральной плотности для ряда, представляющего солнечную активность за период с 1749 по 1924 г., представлены графически на рис. А.3.2. При этом использовалось окно Парзена (пример J разд. 9.2.3) с $K = 20, 40$ и 60 . Следует отметить наличие

¹⁾ См. теорему 9.4.2 и сделанное вслед за ней замечание. — *Прим. перев.*

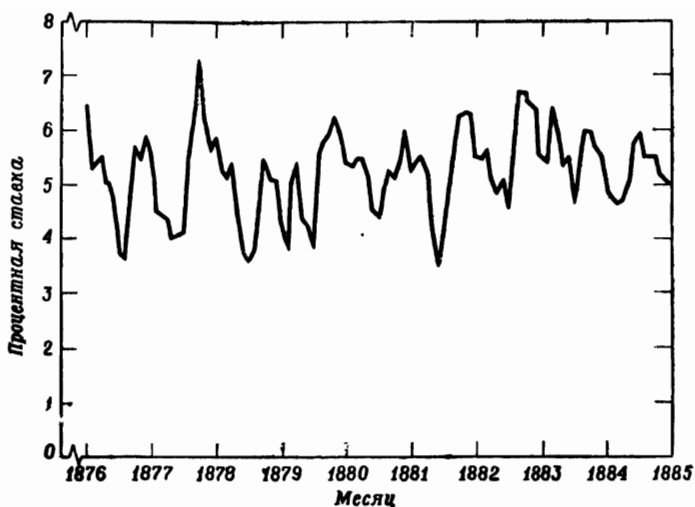


Рис. 9.3.

Процентные ставки коммерческих ценных бумаг на бирже Нью-Йорка за период 1876—1885 гг.

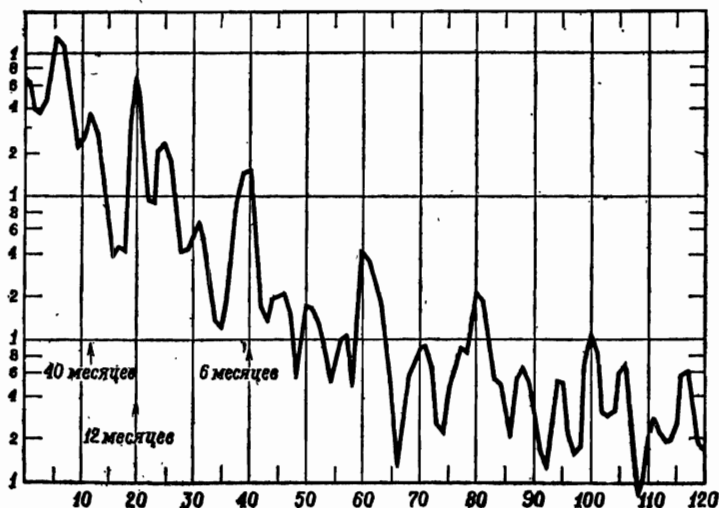


Рис. 9.4.

Оценка спектральной плотности ряда процентных ставок коммерческих ценных бумаг на бирже Нью-Йорка за период 1876—1914 гг

заметного пика на частоте, чуть меньшей 0.09 (период около 11 лет), и на частоте, приблизительно вдвое большей. Первый указывает на наличие цикла с периодом около 11 лет, а второй — на то, что этот цикл не является простой тригонометрической функцией.

В качестве еще одного примера рассмотрим процентные ставки коммерческих ценных бумаг на бирже Нью-Йорка, изучавшиеся Гренджером (1964) на основании ежемесячных данных, имеющихся у Маколея (1938) за период с 1876 по 1914 г. Один из отрезков этого ряда представлен графически на рис. 9.3. Подобранный спектральная плотность изображена на рис. 9.4 в логарифмическом масштабе (120 на горизонтальной шкале соответствует $\lambda = \pi$). При этом использовалась оценка Блэкмена — Тьюки с $a = 0.25$ (окно Хеннинга, пример G разд. 9.2.3) и $K = 120$. Следует отметить наличие пиков на частотах $j/12$, $j = 1, 2, 3, 4, 5$, соответствующих сезонным изменениям. Кроме того, имеется пик, соответствующий периоду 40 месяцев, который можно интерпретировать как цикл деловой активности.

9.6. ОБСУЖДЕНИЕ

Пусть стационарный в широком смысле процесс имеет спектральную плотность. Тогда эта плотность, являясь преобразованием Фурье ковариационной последовательности, определяет ковариационные свойства процесса. Спектральную плотность можно связать с дисперсиями (или энергией) синусоидальных составляющих процесса, имеющих случайные амплитуды и фазы. Оценка спектральной плотности по наблюдаемому временному ряду является в сущности непараметрическим методом, поскольку здесь по конечному множеству наблюдений оценивается функция на $[0, \pi]$, которая не определяется конечным числом параметров. Такая процедура используется обычно в тех случаях, когда исследователь не хочет связывать себя параметрической моделью определенного вида с определенным числом параметров для спектральной плотности или ковариационной последовательности, как это было сделано в гл. 5. Спектральный анализ является более гибким инструментом, чем параметрические выводы. Однако это достигается за счет уменьшения точности оценок. О нем можно говорить скорее как о методе исследования или о методе аналитического представления данных.

Для того чтобы оценивание спектра было информативным, наблюдаемый ряд должен иметь достаточную длину. Необходимая длина ряда зависит от характера спектральной плотности процесса. Спектральную плотность некоррелированной последовательности можно оценить по относительно короткому ряду наблюдений, тогда как для оценки плотности, имеющей много пиков, потребуется значительно большее число наблюдений. Фактически большое число

наблюдений T необходимо и для успешного применения асимптотической теории.

Выбор константы K (или K_T) зависит как от характера оцениваемой спектральной плотности, так и от размера выборки. Чем более гладкой является спектральная плотность, тем меньшим может быть выбрано значение K (для обеспечения меньшей дисперсии). Однако, поскольку спектральная плотность является предметом оценивания, «оптимальное» значение K заранее неизвестно. Поэтому следует испытывать несколько различных значений K . Как видно из приведенных примеров, выбор слишком малого значения K может привести к тому, что какое-то количество пиков спектральной плотности окажется незамеченным. Наоборот, выбор слишком большого K приводит к крайне нерегулярному поведению оценки.

Мы уже замечали, что если $\hat{f}_T(\lambda) \geq 0$, $-\pi \leq \lambda \leq \pi$, то $\hat{f}_T(\lambda)$ можно представить в виде

$$(1) \quad \hat{f}_T(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{j=0}^{K_T} \hat{\alpha}_j e^{i\lambda j} \right|^2,$$

где $\hat{\alpha}_0 = 1$, а $\hat{\alpha}_1, \dots, \hat{\alpha}_{K_T}$ — действительные числа. Из результатов § 7.5 вытекает тогда, что $\hat{f}_T(\lambda)$ можно представить в виде полинома степени K_T от $\cos \lambda$. Поэтому $\hat{f}_T(\lambda)$ может иметь на $(0, \pi)$ не более $K_T - 1$ локальных максимумов и минимумов. Если $\hat{f}_T(\lambda)$ принимает и отрицательные значения, то указанный вывод уже не будет справедлив. В качестве примера такой $\hat{f}_T(\lambda)$ можно привести оценку Бартлетта спектральной плотности, которой соответствует окно, принимающее и отрицательные значения.

В связи с тем, что соответствующая некоррелированным переменным плоская спектральная плотность оценивается наиболее просто, часто оказываются полезными преобразования, приводящие к упрощению спектральной плотности. Говоря более точно, поскольку всякое линейное преобразование временного ряда соответствует умножению спектральной плотности на некоторую передаточную функцию, то при соответствующей информации о характере оцениваемой спектральной плотности преобразование можно выбрать так, чтобы сделать спектральную плотность более плоской. Затем можно оценить новую спектральную плотность и, разделив полученную оценку на соответствующую передаточную функцию, получить оценку исходной спектральной плотности. Подобная процедура обсуждалась в конце § 9.4.

Следует отметить, что с вычислительной точки зрения вместо вычисления ковариаций гораздо более выгодно применять быстрое преобразование Фурье (описанное в разд. 4.3.5). Если интересуются формой спектральной плотности, то можно оценить нормиро-

ванную спектральную плотность $\bar{f}(\lambda) = f(\lambda)/\sigma(0)$. При этом можно использовать ту же асимптотическую теорию, только $f(\lambda)$ следует всюду заменить на $\bar{f}(\lambda)$.

Большинство теоретических и выборочных спектральных плотностей, а также оценок спектральных плотностей представлено в логарифмическом масштабе. Это связано с тем, что асимптотическая дисперсия логарифмов выборочных спектральных плотностей и оценок спектральных плотностей не зависит от значений самих плотностей. Однако представление этих плотностей в обычном масштабе имеет свои преимущества. Значение плотности для каждой частоты соответствует в этом случае дисперсии амплитуд вблизи этой частоты. Далее, колебания спектральной плотности в той области частот, где ее значения малы, не играют особой роли. При использовании же логарифмического масштаба эти колебания становятся преувеличенными, так что становятся заметными столь незначительные локальные максимумы, которые не были видны простым глазом в арифметическом масштабе. Если оценка спектральной плотности $\hat{f}_T(\nu)$ является арифметическим средним значений выборочной спектральной плотности $I_T(\nu)$, то сравнение $\hat{f}_T(\nu)$ и $I_T(\nu)$ также проще производить в обычном масштабе.

В ряде примеров некоторые локальные максимумы соответствуют частотам, кратным той частоте, на которой наблюдается абсолютный максимум. Поскольку в ряд Фурье, аппроксимирующий периодическую функцию или последовательность, не являющуюся тригонометрической, входит несколько слагаемых, то упомянутые вторичные пики могут просто указывать на несинусоидальный характер основной периодической компоненты.

Во введении к этой главе уже отмечалось, что если рассматриваемая здесь оценка спектральной плотности неотрицательна, то ее можно трактовать как спектральную плотность некоторого процесса скользящего среднего. Другой метод оценивания спектральной плотности состоит в оценке коэффициентов процесса авторегрессии, аппроксимирующего процесс, из которого получена выборка. В качестве оценки спектральной плотности исходного процесса используют спектральную плотность выбранного процесса авторегрессии. Такие оценки и оценки типа скользящего среднего ведут себя асимптотически одинаково, однако эквивалентность обоих методов теоретически пока не доказана. Приведенный пример, связанный с индексом Бевериджа цен на пшеницу с выделенным трендом, говорит о том, что следует, по-видимому, предпочесть небольшие запаздывания, но использование слишком малых запаздываний может привести к неправильным выводам. Так, максимум спектральной плотности выбранного для этого случая процесса авторегрессии второго порядка не согласуется ни с одним из максимумов спектральных плотностей подобранных процессов более вы-

соких порядков. Однако преимуществом указанного метода является то, что получаемые коэффициенты дают возможность прогнозировать значения функции и существует разумный способ определения нужного числа запаздываний (§ 5.6).

ЛИТЕРАТУРА

§ 9.2. Бартлетт (1946), (1950), Блэкмен и Тьюки (1959), Гренандер и Розенблатт (1957), Ланцош (1956), Парзен (1957b), (1961b), Хемминг (1962), Хеннан (1960).

§ 9.3. Дженкинс (1961), Парзен (1957b), (1961b).

§ 9.4. Блэкмен и Тьюки (1959), Розенблатт (1959), Чон и Хеннан (1968).

§ 9.5. Гренджер (1964), Маколей (1938).

УПРАЖНЕНИЯ

1. (Разд. 9.2.1) Выведите второе выражение для $\text{Cov}(G, W)$ в (16) непосредственно из первого. Выведите вторую форму записи (19) непосредственно из первой. (Указание. Взять вторую форму записи в (16), выразить $\sin \frac{1}{2} x T / \sin \frac{1}{2} x$ для различных x и проинтегрировать по λ и λ' .)

2. (Разд. 9.2.3) Покажите, что

$$(i) \quad \sum_{r=1}^K \cos \lambda r = \frac{\cos \frac{\lambda}{2} (K+1) \sin \frac{\lambda}{2} K}{\sin \frac{\lambda}{2}},$$

$$(ii) \quad \sum_{r=1}^K \sin \lambda r = \frac{\sin \frac{\lambda}{2} (K+1) \sin \frac{\lambda}{2} K}{\sin \frac{\lambda}{2}},$$

$$(iii) \quad \sum_{r=-K}^K \cos \lambda r = \frac{\sin \frac{\lambda}{2} (2K+1)}{\sin \frac{\lambda}{2}} = 2\pi h_{2K+1}(\lambda),$$

$$(iv) \quad \sum_{r=-K}^K |r| \cos \lambda r = K \frac{\sin \frac{\lambda}{2} (2K+1)}{\sin \frac{\lambda}{2}} - \frac{\sin^2 \frac{\lambda}{2} K}{\sin^2 \frac{\lambda}{2}} =$$

$$= K \frac{\sin \frac{\lambda}{2} (2K+1)}{\sin \frac{\lambda}{2}} + \frac{\cos \lambda K - 1}{2 \sin^2 \frac{\lambda}{2}}.$$

$$\begin{aligned}
 (v) \quad \sum_{r=-K}^K r^2 \cos \lambda r &= K^2 \frac{\sin \frac{\lambda}{2} (2K+1)}{\sin \frac{\lambda}{2}} + \\
 &+ \left(K - \frac{1}{2} \right) \frac{\cos \lambda K}{\sin^2 \frac{\lambda}{2}} - \frac{\sin \frac{\lambda}{2} (2K-1)}{2 \sin^3 \frac{\lambda}{2}}.
 \end{aligned}$$

[Указание. $\sum \cos \lambda r + i \sum \sin \lambda r = \sum e^{i\lambda r}$; $\sum r \cos \lambda r = (d/d\lambda) \sum \sin \lambda r$; $\sum r^2 \cos \lambda r = - (d^2/d\lambda^2) \sum \cos \lambda r$.]

3. (Разд. 9.2.3) Для

$$\hat{f}(v) = \frac{1}{2\pi} \sum_{r=-K}^K \cos vr C_r$$

найдите $\{w_r\}$, $\{w_r^*\}$ и $w(\lambda | 0)$.

4. (Разд. 9.2.3) Убедитесь в том, что для $w^*(\lambda | v)$ из примера А

$$\int_{-\pi}^{\pi} w^*(\lambda | v) I(\lambda) d\lambda = I(v).$$

5. (Разд. 9.2.3) Покажите, что $k_T(\lambda)$ имеет минимумы в точках $\lambda = \pm j2\pi/T$, $j = 1, \dots, [T/2]$, максимум в точке $\lambda = 0$ и локальные максимумы при значениях λ , для которых $\operatorname{tg} \lambda T/2 = T \operatorname{tg} \lambda/2$.

6. (Разд. 9.2.3) Покажите, что $h_{2T-1}(\lambda)$ имеет нули в точках $\lambda = \pm k2\pi/(2T-1)$, $k = 1, \dots, T-1$, максимум при $\lambda = 0$ и локальные максимумы в точках, являющихся корнями уравнения $\operatorname{tg} \lambda (2T-1)/2 = (2T-1) \operatorname{tg} \lambda/2$.

7. (Разд. 9.2.3) Покажите, что

$$\begin{aligned}
 \sum_{r=-K}^K \frac{e^{i\lambda r}}{T-|r|} &= \frac{1}{T} + \sum_{r=1}^K \frac{e^{i\lambda r} + e^{-i\lambda r}}{T-r} = \\
 &= \frac{1}{T} + \left\{ e^{i\lambda T} \sum_{s=T-K}^{T-1} \frac{e^{-i\lambda s}}{s} + e^{-i\lambda T} \sum_{s=T-K}^{T-1} \frac{e^{i\lambda s}}{s} \right\} = \\
 &= \frac{1}{T} + \left\{ 2 \cos \lambda T \sum_{s=T-K}^{T-1} \frac{1}{s} - i e^{i\lambda T} \int_0^{\lambda} \frac{e^{-iu(T-K)} - e^{-iuT}}{1 - e^{-iu}} du + \right. \\
 &\quad \left. + i e^{-i\lambda T} \int_0^{\lambda} \frac{e^{iu(T-K)} - e^{iuT}}{1 - e^{iu}} du \right\}.
 \end{aligned}$$

8. (Разд. 9.2.3) Проверьте соотношение (39).

9. (Разд. 9.2.3) Пусть случайные величины X_1, X_2, X_3, X_4 независимы и одинаково распределены, причем каждая имеет плотность

$$f(x) = \begin{cases} \frac{1}{2}, & -1 \leq x \leq 1, \\ 0, & |x| > 1. \end{cases}$$

Покажите, что плотности распределения сумм $X_1 + X_2$, $X_1 + X_2 + X_3$ и $X_1 + X_2 + X_3 + X_4$ равны соответственно

$$f_2(x) = \begin{cases} \frac{2 - |x|}{4}, & -2 \leq x \leq 2, \\ 0, & |x| \geq 2, \end{cases}$$

$$f_3(x) = \begin{cases} \frac{3 - x^2}{8}, & -1 \leq x \leq 1, \\ \frac{(3 - |x|)^2}{16}, & 1 \leq |x| \leq 3, \\ 0, & |x| \geq 3, \end{cases}$$

$$f_4(x) = \begin{cases} \frac{32 - 12x^2 + 3|x|^3}{96}, & -2 \leq x \leq 2, \\ \frac{(4 - |x|)^3}{96}, & 2 \leq |x| \leq 4, \\ 0, & |x| \geq 4. \end{cases}$$

Проверьте, что характеристическая функция для $f(x)$ равна $(\sin t)/t$ и поэтому характеристическая функция для $f_j(x)$ равна $[(\sin t)/t]^j$.

10. (Разд. 9.2.3) Пусть случайные величины X_1, X_2, X_3, X_4 независимы и принимают только целочисленные значения, каждая с вероятностями

$$\Pr\{X = r\} = \begin{cases} \frac{1}{2K+1}, & r = 0, \pm 1, \dots, \pm K, \\ 0, & |r| > K. \end{cases}$$

Покажите, что соответствующие вероятности для сумм $X_1 + X_2$, $X_1 + X_2 + X_3$ и $X_1 + X_2 + X_3 + X_4$ равны соответственно

$$\begin{aligned} p_2(r) &= \begin{cases} \frac{2K+1-|r|}{(2K+1)^2}, & r = 0, \pm 1, \dots, \pm 2K, \\ 0, & |r| > 2K, \end{cases} \\ p_3(r) &= \begin{cases} \frac{(2K+1)^2 - K(K+1) - r^2}{(2K+1)^3}, & r = 0, \pm 1, \dots, \pm K, \\ \frac{\left[3\left(K + \frac{1}{2}\right) - |r|\right]^2 - \frac{1}{4}}{2(2K+1)^3}, & r = \pm K, \dots, \pm 3K, \\ 0, & |r| > 3K, \end{cases} \\ p_4(r) &= \begin{cases} \frac{4(2K+1)^3 + 2(2K+1) - 3|r| - 6(2K+1)r^2 + 3|r|^3}{6(2K+1)^4}, & r = 0, \pm 1, \dots, \pm 2K, \\ \frac{[2(2K+1) - |r|]^3 - [2(2K+1) - |r|]}{6(2K+1)^4}, & r = \pm 2K, \dots, \pm 4K, \\ 0, & |r| > 4K. \end{cases} \end{aligned}$$

Проверьте, что характеристическая функция исходного распределения равна $\sin(t/2)(2K+1)/[(2K+1)\sin(t/2)]$ и поэтому характеристическая функция распределения $p_j(r)$ равна $\{\sin(t/2)(2K+1)/[(2K+1)\sin(t/2)]\}$.

11. (Разд. 9.2.3) Проверьте следующую формулу «суммирования по частям»:

$$\sum_{r=1}^L a_r (b_{r+1} - b_r) = a_L b_{L+1} - a_0 b_1 - \sum_{r=1}^L b_r (a_r - a_{r-1}).$$

12. (Разд. 9.2.3) Убедитесь, что

$$\sum_{r=1}^L (L-r)^3 \cos \lambda r = -\frac{L^3}{2} + \frac{2 \sin^2 \frac{\lambda}{2} L + 3L^2}{4 \sin^2 \frac{\lambda}{2}} - \frac{3 \sin^2 \frac{\lambda}{2} L}{4 \sin^4 \frac{\lambda}{2}}.$$

[Указание. Использовать формулу суммирования по частям из упр. 11 поочередно с (1) $a_r = (L-r)^3$, $b_r = \sin(\lambda/2)(2r-1)/(2 \sin \lambda/2)$; (2) $a_r = 3r^2 - 3(2L+1)r + 3L^2 + 3L + 1$, $b_r = -\cos \lambda(r-1)/(4 \sin^2 \lambda/2)$; (3) $a_r = 6(L-r+1)$, $b_r = -\sin(\lambda/2)(2r-3)/(8 \sin^3 \lambda/2)$ и (4) $a_r = 6$, $b_r = \cos \lambda(r-2)/(16 \sin^4 \lambda/2)$.]

13. (Разд. 9.2.3) Используйте формулу суммирования по частям для решения упр. 2.

14. (Разд. 9.2.3) Начертите графики $\{\omega_r^*\}$ и $\{\omega_r\}$ для $r = 0, 1, \dots, T-1$ в примерах А, В, С, D, Е с $K = [\pi/b]$, G, H, I и J. Возьмите $T = 20$ и $K = 6$ и используйте везде один и тот же масштаб.

15. (Разд. 9.3.1) Пусть

$$S_T = Z_{kT} + X_{kT}, \quad T = 1, 2, \dots, k = 1, 2, \dots,$$

— последовательность случайных величин, такая, что

$$\mathbb{E} X_{kT}^2 \leq M_k,$$

$$\lim_{k \rightarrow \infty} M_k = 0,$$

$$\mathbb{E} Z_{kT}^2 = L_{kT} \rightarrow L_k \text{ при } T \rightarrow \infty,$$

$$\lim_{k \rightarrow \infty} L_k = L.$$

Покажите, что

$$\lim_{T \rightarrow \infty} \mathbb{E} S_T^2 = L.$$

16. (Разд. 9.3.1) Используя вторые формы записи (16) и (19) § 9.2, докажете соотношение (9) из теоремы 9.3.1.

17. (Разд. 9.3.2) Покажите, что

$$\lim_{KT \rightarrow \infty} \frac{1}{K_T} \omega_T^* \left(\frac{v}{K_T} \middle| 0 \right) = \frac{1}{2\pi} \int_{-1}^1 k(x) e^{ivx} dx,$$

если $k(x)$ интегрируема по Риману на $[-1, 1]$.

19. (Разд. 9.3.3) Покажите, что

$$T\Phi_T(r; g, h) = \begin{cases} T - [|g| + |h| + (g-h)]/2 + r, \\ r = -\{T - [|g| + |h| + (g-h)]/2\}, \dots, (g-h)/2 - [|g| - |h|]/2, \\ T - \max(|g|, |h|), \\ r = \frac{(g-h)}{2} - \frac{||g| - |h||}{2}, \dots, \frac{(g-h)}{2} + \frac{||g| - |h||}{2}, \\ T - [|g| + |h| - (g-h)]/2 - r, \\ r = \frac{(g-h)}{2} + \frac{||g| - |h||}{2}, \dots, T - \frac{|g| + |h| - (g-h)}{2}, \end{cases}$$

и что $T\Phi_T(r; g, h)$ равно нулю для всех других значений r .

20. (Разд. 9.3.3) Докажите (43).

21. (Разд. 9.3.3) Пусть для $-1 \leq a < b \leq 1$

$$g(x) = \begin{cases} C > 0, & a < x \leq b, \\ 0, & \text{в противном случае.} \end{cases}$$

Покажите, что

$$\lim_{K \rightarrow \infty} \sum_{r=-K}^K \frac{1}{K} g\left(\frac{r}{K}\right) e^{idr} = 0$$

для $d \neq k2\pi$, $k = 0, \pm 1, \dots$ [Указание. Суммируя e^{idr} , показать, что

$$\left| \sum_{r=-K}^K \frac{1}{K} g\left(\frac{r}{K}\right) e^{idr} \right| \leq \frac{C}{K} \frac{1}{|\sin d/2|}.$$

22. (Разд. 9.3.3) Пусть $h(x)$ — ступенчатая функция на отрезке $[-1, 1]$, т. е.

$$h(x) = C_j, \quad a_{j-1} < x \leq a_j, \quad j = 1, \dots, n,$$

где $-1 = a_0 < a_1 < \dots < a_n = 1$ и $h(-1) = C_1$. Покажите, что

$$\lim_{K \rightarrow \infty} \sum_{r=-K}^K \frac{1}{K} h\left(\frac{r}{K}\right) e^{idr} = 0$$

для $d \neq k2\pi$, $k = 0, \pm 1, \dots$ [Указание. Применить упр. 21.]

23. (Разд. 9.3.3) Пусть функция $f(x)$ непрерывна на $[-1, 1]$. Покажите, что

$$\lim_{K \rightarrow \infty} \sum_{r=-K}^K \frac{1}{K} f\left(\frac{r}{K}\right) e^{idr} = 0$$

для $d \neq k2\pi$, $k = 0, \pm 1, \dots$ [Указание. Аппроксимировать $f(x)$ ступенчатой функцией и применить упр. 22.]

24. (Разд. 9.3.3) Проверьте таблицу 9.3.3 для окон, отличных от усеченного ядра Даниэля. (См. упр. 25.)

25. (Разд. 9.3.3) Покажите, что

$$\int_{-1}^1 \left(\frac{\sin \pi z}{\pi z} \right)^2 dz = \frac{2}{\pi} \int_0^{2\pi} \frac{\sin y}{y} dy = 4 \sum_{\nu=0}^{\infty} (-1)^{\nu} \frac{(2\pi)^{2\nu}}{(2\nu+1)(2\nu+1)!}.$$

[Указание. Преобразовать $\sin^2 \pi z$ в $(1 - \cos y)/2$ и проинтегрировать второе слагаемое по частям.]

26. (Разд. 9.3.3) Пусть

$$\hat{G}(\nu) = 2 \int_0^{\nu} I(\lambda) d\lambda = \frac{1}{\pi} \left\{ \nu c_0 + 2 \sum_{r=1}^{T-1} \frac{\sin \nu r}{r} c_r \right\}.$$

Покажите, что при надлежащих условиях

$$\lim_{T \rightarrow \infty} T \mathbb{E} [\hat{G}(\nu) - G(\nu)] [\hat{G}(\nu') - G(\nu')] = 8\pi \int_0^{\min(\nu, \nu')} f^2(\lambda) d\lambda.$$

[Указание. Доказательство можно провести по образцу доказательства в упр. 16.]

27. (§ 9.4) Покажите, что если $|r| \leq n$ и $|s| \leq n$, то число слагаемых в разности

$$\sum_{h=s-r+1}^{K_T+s-r} \sum_{q=1-s}^{T-h-r} b_{h,q,r,s} - \sum_{h=1}^{K_T-2n} \sum_{q=1}^{T-h} b_{h,q,r,s}$$

меньше, чем $8Tn + 4K_Tn + 32n^2$. [Указание. Показать, что для значений h между $(1+2n)$ и K_T-2n число слагаемых в указанной разности не превосходит $4n$, а для q между $1+n$ и $T-n$ число слагаемых в разности не превосходит $8n$.]

28. (§ 9.4) Покажите, что более точной оценкой для упр. 27 является $8Tn-4n$.

29. (§ 9.4) Покажите, что случайные величины $v_q v_{q+h}$, $q = 1, \dots, T$, $h = 1, \dots, K_T$, не коррелированы, если $\mathbb{E} v_q v_r v_s v_t = 0$ для не совпадающих попарно индексов q, r, s, t . Предполагается, что $\mathbb{E} v_q = 0$ и $\mathbb{E} v_q v_r = 0$ для $q \neq r$.

30. (§ 9.4) Покажите, что

$$\mathbb{E} \frac{1}{TK_T} \sum_{g,h=1}^{K_T-2n} a_g a_h e^{i\lambda(g-h)} \sum_{q=1}^{T-h} \sum_{s=1}^{T-g} v_q v_{q+h} v_s v_{s+g} = \frac{1}{TK_T} \sum_{h=1}^{K_T-2n} a_h^2 (T-h) \sigma^4,$$

где $\mathbb{E} v_t = 0$, $\mathbb{E} v_t^2 = \sigma^2$ и $\mathbb{E} v_q v_r v_s v_t = 0$ для попарно не совпадающих индексов q, r, s, t .

31. (§ 9.4) Покажите, что

$$\begin{aligned} \mathbb{E} W_{iT}^4 &= \frac{\kappa_4 (\kappa_4 + 3\sigma^4)}{K_T^2} \sum_{h=1}^{K_T} k^4 \left(\frac{h}{K_T} \right) \cos^4 \pi h + \\ &+ 3\sigma^4 (\kappa_4 + 3\sigma^4) \left[\frac{1}{K_T} \sum_{h=1}^{K_T} k^2 \left(\frac{h}{K_T} \right) \cos^2 \pi h \right]^2 \leq \\ &\leq (\kappa_4 + 3\sigma^4) \sup k^4(x) \left\{ \frac{|\kappa_4|}{K_T} + 3\sigma^4 \right\}, \end{aligned}$$

$$\begin{aligned}
\mathbb{E} W_{iT}^2 W_{rT}^2 &= \sigma^8 \left[\frac{1}{K_T} \sum_{h=1}^{K_T} k^2 \left(\frac{h}{K_T} \right) \cos^2 v h \right]^2 < \\
&\leq \sigma^8 \sup k^4(x), \quad r > t + K_T, \\
\mathbb{E} W_{iT}^2 W_{t+r,T}^2 &= \frac{\sigma^8}{K_T^2} \left[\sum_{h=1}^{K_T} k^2 \left(\frac{h}{K_T} \right) \cos^2 v h \right]^2 + \\
&+ \frac{\sigma^4 (\kappa_4 + 2\sigma^4)}{K_T^2} k^2 \left(\frac{r}{K_T} \right) \cos^2 v r \sum_{j=1}^{K_T} k^2 \left(\frac{j}{K_T} \right) \cos^2 v j + \\
&+ \frac{\sigma^4 \kappa_4}{K_T^2} \sum_{h=r+1}^{K_T} k^2 \left(\frac{h}{K_T} \right) k^2 \left(\frac{h-r}{K_T} \right) \cos^2 v h \cos^2 v (h-r) + \\
&+ \frac{2\sigma^8}{K_T^2} \left[\sum_{h=r+1}^{K_T} k \left(\frac{h}{K_T} \right) k \left(\frac{h-r}{K_T} \right) \cos v h \cos v (h-r) \right]^2 + \\
&+ \frac{2\sigma^2 v_3^2}{K_T^2} k \left(\frac{r}{K_T} \right) \cos v r \sum_{k=r+1}^{K_T} k \left(\frac{h}{K_T} \right) k^2 \left(\frac{h-r}{K_T} \right) \cos v h \cos^2 v (h-r) < \\
&\leq \sup k^4(x) \left\{ 3\sigma^8 + \frac{1}{K_T} [\sigma^4 |\kappa_4 + 2\sigma^4| + \sigma^4 |\kappa_4| + 2\sigma^2 v_3^2] \right\}, \\
&1 \leq r \leq K_T
\end{aligned}$$

$$|\mathbb{E} W_{iT}^2 W_{rT} W_{sT}| \leq \frac{4}{K_T} \sigma^8 \sup k^4(x), \quad r > t, s > t, r \neq s,$$

$$|\mathbb{E} W_{iT}^3 W_{rT}| \leq \frac{6}{K_T} v_3^2 \sigma^2 \sup k^4(x), \quad r > t,$$

$$\mathbb{E} W_{iT} W_{sT} W_{rT} W_{qT} = 0 \quad \text{в противном случае.}$$

32. (§ 9.4) Пусть $q \leq p$ и $\sqrt{T/K_T}/K_T^q \rightarrow \beta$ при $T \rightarrow \infty$. Найдите предельное распределение $\sqrt{T/K_T} [\hat{f}_T(v) - f(v)]$ и сформулируйте условия, при которых оно имеет место. Имеет ли этот результат применение и если да, то какое?

33. (§ 9.4) Сравните доверительные интервалы (33) и (36).

34. (§ 9.4) Найдите доверительный интервал для $f(v)$ с уровнем доверия $1 - \varepsilon$, пользуясь статистикой (40). Сравните этот интервал с (33).

35. (§ 9.4) Докажите, что сходимость к пределу в соотношении (28) равномерная.

ЛИНЕЙНЫЕ ТРЕНДЫ И СТАЦИОНАРНЫЕ СЛУЧАЙНЫЕ СОСТАВЛЯЮЩИЕ

10.1. ВВЕДЕНИЕ

В гл. 2, 3 и 4 мы рассматривали задачи статистических выводов для моделей, в которых функции тренда предполагались линейными относительно коэффициентов, а в качестве случайной составляющей рассматривалась последовательность независимых случайных величин с одинаковой дисперсией. В гл. 5 предположение о независимости было заменено предположением о том, что случайная составляющая удовлетворяет некоторому стохастическому разностному уравнению. Теперь же мы займемся оценкой коэффициентов при произвольной ковариационной функции ошибок (разд. 10.2.1 и 10.2.2) и проведем асимптотическое исследование в ситуации, когда случайная составляющая образует стационарный случайный процесс (разд. 10.2.3 и 10.2.4). Поскольку оценки наименьших квадратов можно получать вне зависимости от того, известна или неизвестна случайная составляющая, представляет особый интерес выяснение связи между оценками наименьших квадратов и наилучшими линейными несмещенными оценками.

Т. Андерсоном (1948) было указано на то, что если независимые переменные образуют характеристические векторы ковариационной матрицы ошибок, то критерии типа рассмотренных в § 6.3 подобны и оптимальны, а оценки максимального правдоподобия коэффициентов регрессии в предположении нормальности совпадают с оценками наименьших квадратов (§ 2.4). Ватсон (1952), (1955), а также Ватсон и Хеннан (1956) изучали эффективность оценок наименьших квадратов. Полученное Ватсоном неравенство показывает, что оценки наименьших квадратов эффективны тогда и только тогда, когда независимые переменные образуют совокупность соответствующих характеристических векторов. Магнесс и Макгир (1962) вновь нашли эти условия и доказали их необходимость и

достаточность. Более полное изучение этих условий и обзор литературы можно найти у Ватсона (1967) и Зискинда (1967).

Гренандер (1954), Розенблатт (1956), а также Гренандер совместно с Розенблаттом (1957, гл. 7) исследовали соответствующие задачи в асимптотическом плане. В разд. 10.2.3 их результаты развиваются в форме, аналогичной случаю конечного T . В разд. 10.2.4 показано, что и в общей ситуации с коэффициентами можно обращаться как с нормально распределенными, если только значение T достаточно велико.

В гл. 8 и 9 мы занимались оцениванием ковариаций и спектральной плотности в случае, когда либо математическое ожидание процесса было известно, либо оно оценивалось как константа. В § 10.3 мы рассмотрим оценивание ковариаций и спектральной плотности для случая, когда математическое ожидание процесса является линейной функцией регрессии и коэффициенты оцениваются по методу наименьших квадратов. Это исследование проводится на основе результатов Хеннана (1958).

В гл. 6 для проверки независимости были развиты достаточно общие критерии. Однако построение точных распределений и таблиц процентных точек опиралось там на тот факт, что математическое ожидание было известно или предполагалось постоянным. В § 10.4 мы проведем это исследование для случая, когда используются остатки от оценки регрессии.

10.2. ЭФФЕКТИВНОЕ ОЦЕНИВАНИЕ ФУНКЦИЙ ТРЕНДА

10.2.1. Эффективное оценивание по методу наименьших квадратов

Рассмотрим функцию тренда, линейную относительно неизвестных коэффициентов. Как было показано в § 2.4, в выражения для наилучших линейных несмещенных оценок этих коэффициентов (марковских оценок) входит ковариационная матрица (определенная с точностью до масштабного множителя). Последняя же часто бывает неизвестной. В то же время оценки наименьших квадратов всегда вычислимы и являются несмещенными. Однако они уже могут не иметь минимальные дисперсии (среди несмещенных линейных оценок). В настоящем разделе сравниваются оценки наименьших квадратов и марковские оценки.

Пусть вектор $y = (y_1, \dots, y_T)'$ имеет вектор средних

$$(1) \quad \xi y = Z\beta$$

и ковариационную матрицу

$$(2) \quad \xi(y - Z\beta)(y - Z\beta)' = \Sigma,$$

где $\mathbf{Z} = (z_{it})'$ есть известная матрица размера $T \times p$ ранга p ($\leq T$), β — есть p -компонентный вектор коэффициентов, подлежащих оценке по наблюдению вектора y , а Σ положительно определенная матрица. Тогда марковская оценка вектора β (при известной Σ) равна

$$(3) \quad \mathbf{b} = (\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^{-1}\mathbf{Z}'\Sigma^{-1}y,$$

а оценка наименьших квадратов —

$$(4) \quad \mathbf{b}^* = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'y.$$

Соответствующие ковариационные матрицы равны

$$(5) \quad \mathbb{E}(\mathbf{b} - \beta)(\mathbf{b} - \beta)' = (\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^{-1},$$

$$(6) \quad \mathbb{E}(\mathbf{b}^* - \beta)(\mathbf{b}^* - \beta)' = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\Sigma\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}.$$

Марковской оценкой и оценкой наименьших квадратов произвольной линейной комбинации $\gamma'\beta$ будут соответственно $\gamma'\mathbf{b}$ и $\gamma'\mathbf{b}^*$. Дисперсии этих оценок равны

$$(7) \quad \mathbb{E}(\gamma'\mathbf{b} - \gamma'\beta)^2 = \gamma'(\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^{-1}\gamma,$$

$$(8) \quad \mathbb{E}(\gamma'\mathbf{b}^* - \gamma'\beta)^2 = \gamma'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\Sigma\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\gamma.$$

Обе оценки являются несмещенными. При этом $\text{Var}(\gamma'\mathbf{b}) \leq \leq \text{Var}(\gamma'\mathbf{b}^*)$, т. е.,

$$(9) \quad \gamma'(\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^{-1}\gamma \leq \gamma'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\Sigma\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\gamma$$

для всех γ .

В § 2.4 было показано, что марковская оценка совпадает с оценкой наименьших квадратов в том случае, когда $\mathbf{Z} = \mathbf{V}^*\mathbf{C}$, причем p столбцов матрицы $\mathbf{V}^*$ ($T \times p$) являются характеристическими векторами матрицы Σ и матрица $\mathbf{C}$ не вырождена. Докажем, что эти условия являются и необходимыми.

ТЕОРЕМА 10.2.1. *Оценка наименьших квадратов (4) совпадает с марковской оценкой (3) тогда и только тогда, когда $\mathbf{Z} = \mathbf{V}^*\mathbf{C}$, причем p столбцов матрицы $\mathbf{V}^*$ являются линейно независимыми характеристическими векторами матрицы Σ , а матрица $\mathbf{C}$ не вырождена.*

Доказательство. Исследуемые оценки совпадают тогда и только тогда, когда соотношение $\mathbf{b}^* = \mathbf{b}$ является тождеством относительно y , т. е. тогда и только тогда, когда

$$(10) \quad \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} = \Sigma^{-1}\mathbf{Z}(\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^{-1}.$$

Умножая обе части (10) на Σ слева и на $Z'Z$ справа, получаем

$$(11) \quad \Sigma Z = Z(Z'\Sigma^{-1}Z)^{-1}Z'Z.$$

В то же время существуют невырожденная матрица P размера $p \times p$ и (невырожденная) диагональная матрица D , такие, что

$$(12) \quad P'Z'ZP = I,$$

$$(13) \quad P'Z'\Sigma^{-1}ZP = D^{-1}.$$

(См. упр. 30 гл. 6.) Поэтому, умножая обе части (11) справа на P , получаем

$$(14) \quad \Sigma(ZP) = (ZP)D.$$

Таким образом, столбцы матрицы ZP являются характеристическими векторами матрицы Σ , а диагональные элементы матрицы D — соответствующими им характеристическими корнями. Если положить $ZP = V^*$ и $C = P^{-1}$, то получится утверждение теоремы. ■

Поскольку несмещенные линейные оценки с минимальной дисперсией единственны, то любая несмещенная линейная оценка с минимальной дисперсией будет марковской. Оценки наименьших квадратов имеют минимальную дисперсию тогда и только тогда, когда они удовлетворяют условию теоремы 10.2.1. Условие совпадения оценок наименьших квадратов с марковскими для всех компонент вектора β (т. е. для всех векторов γ) можно сформулировать и иначе. Это совпадение будет в том и только в том случае, когда существует p линейно независимых линейных комбинаций столбцов матрицы Z , являющихся характеристическими векторами матрицы Σ . Еще одна возможная формулировка состоит в том, что

$$(15) \quad \Sigma Z = ZG,$$

где G — невырожденная матрица. Действительно, (15) переходит в (14), если положить $G = PDP^{-1}$.

Посмотрим, как этот критерий можно сформулировать с использованием заданного множества характеристических векторов матрицы Σ . Соответствующий результат выведем с помощью ковариационных матриц оценок. Пусть характеристические корни матрицы Σ равны $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_T$. Пусть матрица V такова, что ее столбцы являются характеристическими векторами матрицы Σ и нормированы таким образом, что $V'V = I$, т. е. $\Sigma V = V\Lambda$, где Λ — диагональная матрица, на диагонали которой стоят характеристические корни матрицы Σ . Если эти корни различны, то соответствующие характеристические векторы определяются однозначно. Далее, любую матрицу Z размера $T \times p$ можно представить в виде $Z = VA$, где V — определенная выше матрица, а A — некоторая матрица размера $T \times p$. Поскольку $V'\Sigma V = \Lambda$ и $V'\Sigma^{-1}V = \Lambda^{-1}$, то

для такой матрицы Z справедливы соотношения

$$(16) \quad Z'Z = A'V'VA = A'A = \sum_{r=1}^T a_r a_r',$$

$$(17) \quad Z'\Sigma Z = A'V'\Sigma VA = A'\Lambda A = \sum_{r=1}^T \lambda_r a_r a_r',$$

$$(18) \quad Z'\Sigma^{-1}Z = A'V'\Sigma^{-1}VA = A'\Lambda^{-1}A = \sum_{r=1}^T \frac{1}{\lambda_r} a_r a_r',$$

в которых $A = (a_1, \dots, a_T)'$.

В том случае, когда имеются кратные характеристические корни, характеристические векторы уже не будут однозначно определенными. Предположим, что различающиеся между собой значения характеристических корней есть $\nu_1 > \nu_2 > \dots > \nu_H$ и

$$(19) \quad \Lambda = \begin{pmatrix} \nu_1 I & 0 & \dots & 0 \\ 0 & \nu_2 I & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \nu_H I \end{pmatrix},$$

где единичные матрицы имеют порядки $m_1, m_2, \dots, m_H$ соответственно. Разобьем на соответствующие блоки также матрицы $V = (V^{(1)}V^{(2)} \dots$

$\dots V^{(H)})$ и $A' = (A^{(1)'}A^{(2)'} \dots A^{(H)'})$. Тогда $Z = \sum_{h=1}^H V^{(h)}A^{(h)}$,

$$(20) \quad Z'Z = \sum_{h=1}^H A^{(h)'}A^{(h)},$$

$$(21) \quad Z'\Sigma Z = \sum_{h=1}^H \nu_h A^{(h)'}A^{(h)},$$

$$(22) \quad Z'\Sigma^{-1}Z = \sum_{h=1}^H \frac{1}{\nu_h} A^{(h)'}A^{(h)}.$$

Матрицы $V^{(h)}$, $h = 1, \dots, H$, являются неоднозначно определенными в том смысле, что $V^{(h)}$ можно заменить на $V^{(h)}Q^{(h)}$, где $Q^{(h)}$ — произвольная ортогональная матрица порядка m_h . При этом матрицы $A^{(h)}$ заменятся на $Q^{(h)'}A^{(h)}$, $h = 1, \dots, H$. Однако соотношения (20), (21) и (22) не изменятся (поскольку $A^{(h)} = V^{(h)'}Z$ для заданной $V^{(h)}$). Сформулируем теперь условие совпадения марковских оценок и оценок наименьших квадратов с использованием рангов матриц $A^{(h)}$, $h = 1, \dots, H$.

ТЕОРЕМА 10.2.2. Матрица ковариаций оценки наименьших квадратов (6) совпадает с матрицей ковариаций марковской оценки (5) тогда и только тогда, когда сумма рангов матриц $\mathbf{A}^{(h)}$, $h = 1, \dots, H$, $\mathbf{Z} = \sum_{h=1}^H \mathbf{V}^{(h)} \mathbf{A}^{(h)}$, равна p .

ДОКАЗАТЕЛЬСТВО. Прежде всего найдется такая невырожденная матрица $\mathbf{P}$, для которой $\mathbf{P}'\mathbf{Z}'\mathbf{Z}\mathbf{P} = \mathbf{I}$ и $\mathbf{P}'\mathbf{Z}'\Sigma\mathbf{Z}\mathbf{P} = \mathbf{D}$, где $\mathbf{D}$ — диагональная матрица с элементами $d_{11} \geq \dots \geq d_{pp} > 0$. Тогда

$$(23) \quad \mathbf{I} = \mathbf{P}'\mathbf{Z}'\mathbf{Z}\mathbf{P} = \sum_{h=1}^H \mathbf{C}^{(h)},$$

$$(24) \quad \mathbf{D} = \mathbf{P}'\mathbf{Z}'\Sigma\mathbf{Z}\mathbf{P} = \sum_{h=1}^H v_h \mathbf{C}^{(h)},$$

$$(25) \quad \mathbf{P}'\mathbf{Z}'\Sigma^{-1}\mathbf{Z}\mathbf{P} = \sum_{h=1}^H \frac{1}{v_h} \mathbf{C}^{(h)},$$

где $\mathbf{C}^{(h)} = \mathbf{P}'\mathbf{A}^{(h)'}\mathbf{A}^{(h)}\mathbf{P}$ — положительно полуопределенная матрица, $h = 1, \dots, H$. Указанные в формулировке ковариационные матрицы (5) и (6) совпадают тогда и только тогда, когда (25) равна $\mathbf{D}^{-1}$. Диагональные элементы матриц (23), (24) и (25) равны

$$(26) \quad 1 = \sum_{h=1}^H c_{ii}^{(h)} \quad i = 1, \dots, p,$$

$$(27) \quad d_{ii} = \sum_{h=1}^H v_h c_{ii}^{(h)}, \quad i = 1, \dots, p,$$

$$(28) \quad \sum_{h=1}^H \frac{1}{v_h} c_{ii}^{(h)}, \quad i = 1, \dots, p.$$

Поскольку матрица $\mathbf{C}^{(h)}$ положительно полуопределена, то $c_{ii}^{(h)} \geq 0$. Из (26) вытекает тогда, что для каждого значения i числа $c_{ii}^{(1)}, \dots, c_{ii}^{(H)}$ образуют набор вероятностей. При этом правая часть (27) есть не что иное как $\mathbb{E}X_i$, где X_i — дискретная случайная величина с распределением вероятностей $\text{Pr}\{X_i = v_h\} = c_{ii}^{(h)}$, $h = 1, \dots, H$, а (28) равно $\mathbb{E}X_i^{-1}$, $i = 1, \dots, p$. Если (25) есть $\mathbf{D}^{-1}$, то (28) равно $1/d_{ii}$. Но так как $\mathbb{E}X_i^{-1} = 1/\mathbb{E}X_i$ для $X_i > 0$ тогда и только тогда, когда X_i с вероятностью 1 есть просто положительная постоянная, то $c_{ii}^{(h)} = 1$ для одного значения h и $c_{ii}^{(h)} = 0$ для всех остальных значений h . (См. лемму 10.2.1 ниже.) Таким образом, во всей совокуп-

ности положительно полуопределенных матриц $C^{(1)}, \dots, C^{(H)}$ имеется всего p диагональных элементов, равных 1. Все остальные диагональные элементы этих матриц равны нулю. Учитывая, что значения $v_h, h = 1, \dots, H$, и $d_{ii}, i = 1, \dots, p$, упорядочены одинаковым образом, получаем отсюда, что те элементы $c_{ii}^{(h)}$, которые равны 1 при некотором h , соответствуют последовательным значениям индекса i . В силу положительной полуопределенности матрицы $C^{(h)}$ из $c_{ii}^{(h)} = 0$ вытекает, что и $c_{ij}^{(h)} = c_{ji}^{(h)} = 0, j = 1, \dots, p$. Таким образом, $C^{(h)}$ имеет не более одного диагонального блока, отличного от нулевого. Но тогда из (23) следует, что остальные недиагональные элементы равны нулю и ранг $C^{(h)}$ равен числу единиц на ее диагонали. Поэтому сумма рангов положительно полуопределенных матриц $C^{(h)}, h = 1, \dots, H$, равна p , а ранг $C^{(h)}$ совпадает с рангом матрицы $A^{(h)}, h = 1, \dots, H$.

Обратно, предположим, что сумма рангов матриц $C^{(h)}, h = 1, \dots, H$, равна p . Из (23) и (24) находим, что $\sum_{g \neq h} (v_g - v_h) C^{(g)} = D - v_h I$. Ранг левой части последнего соотношения не превосходит суммы рангов $H - 1$ матриц $C^{(g)} (g \neq h)$, равной p минус ранг матрицы $C^{(h)}$. Поэтому число тех d_{ii} , которые равны v_h , не меньше ранга $C^{(h)}$. Поскольку последнее утверждение справедливо для всех h , это число в точности равно рангу $C^{(h)}$. Пусть ненулевыми среди матриц $C^{(h)}, h = 1, \dots, H$, являются матрицы $L_1, \dots, L_G$, имеющие ранги $r_1, \dots, r_G$ соответственно, и пусть отвечающие им v_h равны $u_1, \dots, u_G$. Тогда (27) имеет вид $d_{ii} = \sum_{g=1}^G u_g l_{ii}^{(g)}$. Элемент d_{ii} равен для $i = 1, \dots, r_1$ значению u_1 , наибольшему среди $u_g, g = 1, \dots, G$. Поэтому для этих значений i имеем $l_{ii}^{(1)} = 1$ и $l_{ii}^{(g)} = 0, g \neq 1$. Из приведенных ранее соображений ясно, что верхний левый угол матрицы L_1 есть единичная матрица, а верхний левый угол L_g является нулевой матрицей, $g \neq 1$. Поскольку ранг L_1 равен порядку матрицы I , то остальные элементы L_1 равны нулю. Подобные же соображения, примененные к следующей группе уравнений, показывают, что L_2 имеет единичную матрицу в следующем диагональном блоке и нули на остальных местах и т. д. В результате получаем, что каждая матрица L_g имеет диагональный блок, представляющий собой единичную матрицу, причем порядок этой матрицы равен рангу L_g . ■

ЛЕММА 10.2.1. Если $X > 0$, то $\mathcal{E}X\mathcal{E}X^{-1} \geq 1$. При этом $\mathcal{E}X\mathcal{E}X^{-1} = 1$ тогда и только тогда, когда X с вероятностью 1 является положительной постоянной.

Доказательство. Утверждение леммы следует немедленно из соотношения

$$(29) \quad 0 \leq \mathbb{E} \left(\frac{1}{\sqrt{X}} - \frac{\sqrt{X}}{\mathbb{E}X} \right)^2 = \mathbb{E} \frac{1}{X} - 2 \frac{1}{\mathbb{E}X} + \mathbb{E} \frac{X}{(\mathbb{E}X)^2} = \\ = \mathbb{E} \frac{1}{X} - \frac{1}{\mathbb{E}X}. \blacksquare$$

Условие, состоящее в том, что сумма рангов матриц $A^{(1)}, \dots, A^{(H)}$ равна p , эквивалентно тому, что столбцы матрицы

$$(30) \quad Z = \sum_{h=1}^H V^{(h)} A^{(h)} = \sum_{h=1}^H (V^{(h)} Q^{(h)}) (Q^{(h)})' A^{(h)}$$

образуют совокупность p характеристических векторов матрицы Σ , умноженной на некоторую невырожденную матрицу. Действительно, можно выбрать такую матрицу $Q^{(h)}$, чтобы число строк матрицы $Q^{(h)'} A^{(h)}$, не состоящих целиком из нулевых элементов, было в точности равно рангу $A^{(h)}$.

Предположим теперь, что матрица Z не удовлетворяет условию теоремы 10.2.1 (или теоремы 10.2.2). Иными словами, число линейно независимых линейных комбинаций столбцов Z , являющихся характеристическими векторами матрицы Σ , меньше p . Пусть максимальное число таких линейных комбинаций равно p_1 , а сами эти линейные комбинации заданы в виде ZG_1 , где G_1 — матрица размера $p \times p_1$. Пусть G_2 — матрица размера $p \times (p - p_1)$ и ранга $(p - p_1)$, такая, что $(ZG_2)' (ZG_1) = G_2' Z' ZG_1 = 0$, и пусть $G = (G_1 G_2)$. Тогда ZG состоит из p_1 столбцов, являющихся характеристическими векторами G , и $p - p_1$ столбцов, ортогональных первым p_1 столбцам. При этом не существует такой линейной комбинации последних $p - p_1$ столбцов, которая была бы характеристическим вектором матрицы Σ . В этих условиях $ZG_1 = V_1^*$ и $\Sigma V_1^* = V_1^* \Lambda_1^*$, где Λ_1^* — диагональная матрица. Пользуясь тем фактом, что $G_2' Z' ZG_1 = 0$, находим, что

$$(31) \quad (ZG_2)' \Sigma (ZG_1) = G_2' Z' \Sigma V_1^* = G_2' Z' V_1^* \Lambda_1^* = \\ = G_2' Z' ZG_1 \Lambda_1^* = 0,$$

$$(32) \quad (ZG_2)' \Sigma^{-1} (ZG_1) = G_2' Z \Sigma^{-1} V_1^* = G_2' Z' V_1^* (\Lambda_1^*)^{-1} = \\ = G_2' Z' ZG_1 (\Lambda_1^*)^{-1} = 0.$$

Пусть $ZG = Z(G_1 G_2) = (X_1 X_2) = X$ и

$$(33) \quad G^{-1} \beta = \gamma = \begin{pmatrix} \gamma^{(1)} \\ \gamma^{(2)} \end{pmatrix}.$$

Тогда $Z\beta = X\gamma = X_1\gamma^{(1)} + X_2\gamma^{(2)}$, $Zb = Xc$ и $Zb^* = Xc^*$, где $c = G^{-1}b$ есть марковская оценка вектора γ ,

$$(34) \quad c = \left[\begin{pmatrix} X_1' \\ X_2' \end{pmatrix} \Sigma^{-1} (X_1 X_2) \right]^{-1} \begin{pmatrix} X_1' \\ X_2' \end{pmatrix} \Sigma^{-1} y,$$

а $c^* = G^{-1}b^*$ есть оценка наименьших квадратов этого вектора,

$$(35) \quad c^* = \left[\begin{pmatrix} X_1' \\ X_2' \end{pmatrix} (X_1 X_2) \right]^{-1} \begin{pmatrix} X_1' \\ X_2' \end{pmatrix} y = \begin{pmatrix} V_1' y \\ (X_2' X_2)^{-1} X_2' y \end{pmatrix}.$$

Оценка наименьших квадратов для $\gamma^{(1)}$ совпадает, таким образом, с марковской оценкой для $\gamma^{(1)}$ и ортогональна как оценке наименьших квадратов, так и марковской оценке для $\gamma^{(2)}$.

Для более подробного ознакомления с этими задачами см. Ватсон (1967) и Зискинд (1967) [которые развили идеи Андерсона (1948)].

10.2.2. Мера эффективности линейных оценок

Перейдем теперь к рассмотрению вопроса о том, какую погрешность мы допускаем, используя оценку наименьших квадратов вместо марковской в том случае, когда эти оценки отличаются. Ковариационная матрица оценки наименьших квадратов b^* по крайней мере «не меньше», чем ковариационная матрица марковской оценки b , в том смысле, что разность этих двух матриц положительно полуопределена. Точнее говоря, для любого вектора γ

$$(36) \quad \gamma' [(Z'Z)^{-1} Z' \Sigma Z (Z'Z)^{-1} - (Z' \Sigma^{-1} Z)^{-1}] \gamma \geq 0.$$

Последнее соотношение вытекает из (9). Чем «больше» ковариационная матрица b^* по сравнению с ковариационной матрицей b , тем менее эффективно оценивание с помощью метода наименьших квадратов. В качестве меры эффективности возьмем отношение определителей этих ковариационных матриц

$$(37) \quad \text{Eff}(b^*) = \frac{|\Sigma(b - \beta)(b - \beta)'|}{|\Sigma(b^* - \beta)(b^* - \beta)'|} = \frac{|(Z' \Sigma^{-1} Z)^{-1}|}{|(Z'Z)^{-1} Z' \Sigma Z (Z'Z)^{-1}|} = \\ = \frac{|Z'Z|^2}{|Z' \Sigma Z| \cdot |Z' \Sigma^{-1} Z|}.$$

Геометрическую интерпретацию ковариационных матриц для b и b^* можно дать с помощью соответствующих эллипсоидов рассея-

ния, которые имеют вид

$$(38) \quad (x - \beta)' [\mathcal{E}(b - \beta)(b - \beta)']^{-1} (x - \beta) = \\ = (x - \beta)' Z' \Sigma^{-1} Z (x - \beta) = p + 2,$$

$$(39) \quad (x - \beta)' [\mathcal{E}(b^* - \beta)(b^* - \beta)']^{-1} (x - \beta) = \\ = (x - \beta)' Z' Z (Z' \Sigma Z)^{-1} Z' Z (x - \beta) = p + 2.$$

Равномерные распределения на областях, заключенных в пределах эллипсоидов (38) и (39), имеют те же самые векторы средних и те же ковариационные матрицы, что и оценки b и b^* соответственно. Неравенство (36) означает геометрически, что эллипсоид (38) расположен целиком внутри эллипсоида (39). Числитель первого и второго отношений в (37) пропорционален квадрату объема эллипсоида (38), а знаменатель пропорционален квадрату объема эллипсоида (39). (См. Т. Андерсон (1958).)

ТЕОРЕМА 10.2.3. *Eff(b^*) ≤ 1 , причем Eff(b^*) = 1 тогда и только тогда, когда оценка наименьших квадратов совпадает с марковской.*

Для облегчения изучения эффективности оценок наименьших квадратов нам понадобятся некоторые алгебраические результаты.

ЛЕММА 10.2.2. *Если $Z = WQ$ и матрица Q не вырождена, то эффективность оценок наименьших квадратов коэффициентов регрессии по Z совпадает с эффективностью оценок наименьших квадратов коэффициентов регрессии по W .*

ДОКАЗАТЕЛЬСТВО. Эффективность оценок наименьших квадратов коэффициентов регрессии по Z равна (37). Подстановка $Z = WQ$ приводит к соотношению

$$(40) \quad \frac{|Q'W'WQ|^2}{|Q'W'\Sigma WQ| \cdot |Q'W'\Sigma^{-1}WQ|} = \frac{|W'W|^2}{|W'\Sigma W| \cdot |W'\Sigma^{-1}W|},$$

которое и является эффективностью оценок наименьших квадратов коэффициентов регрессии по W . ■

ЛЕММА 10.2.3. *Если $Z = KW$ и $K'\Sigma K = \Psi$, где K — ортогональная матрица, то эффективность оценок наименьших квадратов коэффициентов регрессии по Z , имеющей ковариационную матрицу Σ , совпадает с эффективностью оценок наименьших квадратов коэффициентов регрессии по W , имеющей ковариационную матрицу Ψ .*

ДОКАЗАТЕЛЬСТВО. Эффективность оценок наименьших квадратов коэффициентов регрессии по Z , имеющей ковариационную матрицу Σ , равна (37). Подставляя в (37) $Z = KW$ и $K'\Sigma K = \Psi$ (т. е. $\Sigma =$

$= \mathbf{K}\Psi\mathbf{K}'$), получаем

$$(41) \quad \frac{|\mathbf{W}'\mathbf{K}'\mathbf{K}\mathbf{W}|^2}{|\mathbf{W}'\mathbf{K}'\mathbf{K}\Psi\mathbf{K}'\mathbf{K}\mathbf{W}| \cdot |\mathbf{W}'\mathbf{K}'\mathbf{K}\Psi^{-1}\mathbf{K}'\mathbf{K}\mathbf{W}|} = \frac{|\mathbf{W}'\mathbf{W}|^2}{|\mathbf{W}'\Psi\mathbf{W}| \cdot |\mathbf{W}'\Psi^{-1}\mathbf{W}|},$$

а это и есть эффективность оценок наименьших квадратов коэффициентов регрессии по $\mathbf{W}$, имеющей ковариационную матрицу Ψ .

Предположим, что существует p_1 ($0 < p_1 < p$) линейно независимых линейных комбинаций столбцов матрицы $\mathbf{Z}$, являющихся характеристическими векторами матрицы Σ , и пусть эти линейные комбинации заданы в виде $\mathbf{Z}\mathbf{G}_1 = \mathbf{V}_1^*$. Здесь p_1 является максимальным числом таких линейно независимых комбинаций. Пусть матрица $\mathbf{G}_2$ размера $p \times (p - p_1)$ выбрана так, что $\mathbf{Z}\mathbf{G}_2$ ортогональна $\mathbf{Z}\mathbf{G}_1$. При этом справедливы соотношения (31) и (32). Положим $(\mathbf{Z}\mathbf{G}_1\mathbf{Z}\mathbf{G}_2) = \mathbf{Z}\mathbf{G} = \mathbf{W} = (\mathbf{W}_1\mathbf{W}_2)$. Тогда эффективность оценки наименьших квадратов равна

$$(42) \quad \frac{\begin{vmatrix} \mathbf{W}_1'\mathbf{W}_1 & 0 \\ 0 & \mathbf{W}_2'\mathbf{W}_2 \end{vmatrix}^2}{\begin{vmatrix} \mathbf{W}_1'\Sigma\mathbf{W}_1 & 0 \\ 0 & \mathbf{W}_2'\Sigma\mathbf{W}_2 \end{vmatrix} \cdot \begin{vmatrix} \mathbf{W}_1'\Sigma^{-1}\mathbf{W}_1 & 0 \\ 0 & \mathbf{W}_2'\Sigma^{-1}\mathbf{W}_2 \end{vmatrix}} =$$

$$= \frac{|\mathbf{W}_1'\mathbf{W}_1|^2}{|\mathbf{W}_1'\Sigma\mathbf{W}_1| \cdot |\mathbf{W}_1'\Sigma^{-1}\mathbf{W}_1|} \cdot \frac{|\mathbf{W}_2'\mathbf{W}_2|^2}{|\mathbf{W}_2'\Sigma\mathbf{W}_2| \cdot |\mathbf{W}_2'\Sigma^{-1}\mathbf{W}_2|} =$$

$$= \frac{|\mathbf{W}_2'\mathbf{W}_2|^2}{|\mathbf{W}_2'\Sigma\mathbf{W}_2| \cdot |\mathbf{W}_2'\Sigma^{-1}\mathbf{W}_2|},$$

ввиду того, что $\mathbf{W}_1 = \mathbf{V}_1^*$, где $\mathbf{V}_1^*$ состоит из характеристических векторов Σ [т. е. $\mathbf{W}_1'\Sigma\mathbf{W}_1 = \Lambda_1^*$, $\mathbf{W}_1'\Sigma^{-1}\mathbf{W}_1 = (\Lambda_1^*)^{-1}$]. Иными словами, эффективность оценки наименьших квадратов оказывается равной эффективности оценки наименьших квадратов коэффициентов регрессии по части матрицы $\mathbf{Z}$, ортогональной той части $\mathbf{Z}$, коэффициенты которой удастся эффективно оценить. Геометрическая интерпретация этого факта состоит в том, что имеется такая p_1 -мерная гиперплоскость, проходящая через β , которая имеет одно и то же пересечение с эллипсоидом (38) и с эллипсоидом (39). Отношение объемов зависит при этом от длин главных осей, ортогональных к указанному p_1 -мерному подпространству.

Лемма 10.2.4. Эффективность оценок наименьших квадратов коэффициентов регрессии по $\mathbf{Z}$ равна

$$(43) \quad \frac{1}{|\mathbf{W}'\Lambda\mathbf{W}| \cdot |\mathbf{W}'\Lambda^{-1}\mathbf{W}|},$$

где $\mathbf{W} = \mathbf{V}'\mathbf{Z}\mathbf{P}$, а $\mathbf{P}$ — невырожденная матрица, такая, что

$$(44) \quad \mathbf{W}'\mathbf{W} = \mathbf{P}'\mathbf{Z}'\mathbf{Z}\mathbf{P} = \mathbf{I}.$$

Доказательство. Утверждение леммы вытекает из результатов лемм 10.2.2 и 10.2.3. ■

Если имеется некоторая информация о $\mathbf{Z}$ или о Σ , то могут представлять интерес нижние границы эффективности оценок наименьших квадратов при этих условиях. Пусть, например, матрица Σ известна, но этот факт не используется при оценивании. Насколько при этом могут оказаться неэффективными оценки наименьших квадратов? Эта задача сводится к отысканию минимума (43) по всем матрицам $\mathbf{W}$ размера $T \times p$, таким, что $\mathbf{W}'\mathbf{W} = \mathbf{I}$. [Минимизацию (43) можно выполнять на компактном множестве матриц $\mathbf{W}$, так что минимум существует и достигается.]

Если $p = 1$, то (43) принимает вид

$$(45) \quad \text{Eff}(\mathbf{b}^*) = \frac{1}{\mathbf{w}'\Lambda\mathbf{w} \cdot \mathbf{w}'\Lambda^{-1}\mathbf{w}} = \\ = \frac{1}{\sum_{t=1}^T \lambda_t w_t^2 \sum_{t=1}^T w_t^2 / \lambda_t} = \frac{1}{\sum_{j=1}^H v_j P_j \cdot \sum_{j=1}^H P_j / v_j},$$

где различные значения чисел λ_t обозначены $v_1 > v_2 > \dots > v_H$, а $P_j = \sum_t w_t^2$, причем суммирование производится по тем значениям t , для которых $\lambda_t = v_j$. Знаменатель (45) имеет вид $\mathbb{E}X \cdot \mathbb{E}X^{-1}$, где X — случайная величина с распределением $\text{Pr}\{X = v_j\} = P_j$, $j = 1, \dots, H$. Поэтому к (45) можно применить следующее неравенство Канторовича (1948).

Лемма 10.2.5. (Неравенство Канторовича.) Если случайная величина X такова, что $0 < m \leq X \leq M$, то

$$(46) \quad \mathbb{E}X \cdot \mathbb{E}X^{-1} \leq \frac{(m+M)^2}{4mM} = \frac{1}{4} \left(\frac{m}{M} + \frac{M}{m} + 2 \right).$$

Доказательство. Если $0 < m \leq x \leq M$, то

$$(47) \quad 0 \leq (M-x)(x-m) = (M+m-x)x - Mm,$$

откуда

$$(48) \quad \frac{1}{x} \leq \frac{M+m-x}{Mm}.$$

Поэтому умноженное на $\mathbb{E}X$ математическое ожидание случайной величины $1/X$ удовлетворяет неравенству

$$(49) \quad \mathbb{E}X \cdot \mathbb{E}X^{-1} \leq \frac{1}{Mm} (M+m - \mathbb{E}X) \mathbb{E}X.$$

Отсюда и из соотношения

$$(50) \quad 0 \leq [\mathbb{E}X - (M + m)/2]^2 = (\mathbb{E}X)^2 - (M + m) \mathbb{E}X + \frac{(M + m)^2}{4}$$

вытекает искомое неравенство (46). ■

ТЕОРЕМА 10.2.4. Если $p = 1$, а λ_1 и λ_T — соответственно наибольший и наименьший характеристические корни матрицы Σ , то

$$(51) \quad \text{Eff}(\mathbf{b}^*) \geq \frac{4\lambda_1\lambda_T}{(\lambda_1 + \lambda_T)^2}.$$

Верхняя граница в лемме 10.2.5 достигается для такой случайной величины, у которой $\text{Pr}\{X = m\} = \text{Pr}\{X = M\} = 1/2$. Соответственно, нижняя граница в (51) достигается для $P_1 = P_N$ и $P_j = 0$, $j = 2, \dots, N - 1$, т. е. в том случае, когда z является средним двух характеристических векторов, соответствующих наибольшему и наименьшему характеристическим корням.

Нижняя граница для $\text{Eff}(\mathbf{b}^*)$ может быть записана в виде $4(\lambda_T/\lambda_1)/[1 + (\lambda_T/\lambda_1)]^2$. Иными словами, она является функцией отношения наименьшего характеристического корня матрицы Σ к наибольшему характеристическому корню этой матрицы. В табл. 10.1 приведены значения минимума $\text{Eff}(\mathbf{b}^*)$ для различных значений отношения λ_T/λ_1 . Следует отметить, что для $\lambda_T/\lambda_1 \geq 1/2$ потеря эффективности не очень велика.

Таблица 10.1

НИЖНЯЯ ГРАНИЦА ЭФФЕКТИВНОСТИ СКАЛЯРНОЙ ОЦЕНКИ НАИМЕНЬШИХ
КВАДРАТОВ

λ_T/λ_1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$\min \text{Eff}(\mathbf{b}^*)$	0.331	0.556	0.710	0.816	0.889	0.938	0.969	0.988	0.997	1.000

Если $p \geq 2$, то здесь уже не удастся найти такой удовлетворительной нижней границы, которая была бы достижимой. Например, имеет место оценка

$$(52) \quad \text{Eff}(\mathbf{b}^*) \geq \left[\frac{4\lambda_1\lambda_T}{(\lambda_1 + \lambda_T)^2} \right]^p,$$

однако она не достижима. Ватсон (1967) получил другую оценку снизу:

$$(53) \quad \text{Eff}(\mathbf{b}^*) \geq \frac{4\lambda_1\lambda_2 \dots \lambda_p\lambda_{T-p+1} \dots \lambda_T}{(\lambda_1\lambda_2 \dots \lambda_p + \lambda_{T-p+1} \dots \lambda_T)^2}.$$

При этом он привел два примера, в одном из которых лучшей оказывается оценка (52), а в другом — оценка (53). [См. Ватсон (1955).]

ТЕОРЕМА 10.2.5. *Оценки наименьших квадратов являются эффективными для всех Z тогда и только тогда, когда матрица Σ отличается от единичной матрицы I лишь скалярным множителем.*

ДОКАЗАТЕЛЬСТВО. Единственными матрицами, для которых все векторы характеристические, являются матрицы, указанные в формулировке теоремы. ■

Пусть в качестве примера $A_0 = I$, $A_j = (B^j + B'^j)$, где

$$(54) \quad B = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

для $j = 1, \dots, [T/2]$. Если $T = 2K$, то

$$(55) \quad B^K = B'^K = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix},$$

где каждая из подматриц имеет порядок K . Пусть, подобно тому как это было в гл. 6,

$$(56) \quad \Sigma = \sum_{j=0}^{[T/2]} \gamma_j A_j.$$

Характеристическими корнями будут здесь, во-первых $\sum_j \gamma_j$, соответствующий характеристическому вектору $\sqrt{1/T} \mathbf{e} = \sqrt{1/T} (1, \dots, 1)'$, затем, если T четное, $\sum_j (-1)^j \gamma_j$, соответствующий характеристическому вектору $\sqrt{1/T} (-1, 1, \dots, 1)'$, а также $\sum_j \gamma_j \cos 2\pi jk/T$, соответствующие $\sqrt{2/T} (\cos 2\pi k/T, \cos 4\pi k/T, \dots, 1)'$ и $\sqrt{2/T} \times (\sin 2\pi k/T, \sin 4\pi k/T, \dots, 0)'$, $k = 1, 2, \dots, [(T-1)/2]$. При этом, для каждого $k = 1, 2, \dots, [(T-1)/2]$ корни имеют кратность два. Соответствующая каждому двойному корню пара характеристических векторов, указанных выше, может быть заменена любой парой линейных комбинаций этих векторов, образующей ортогональное преобразование размерности два. Оценка наименьших квадратов будет здесь эффективной тогда и только тогда, когда p столбцов матрицы Z являются независимыми линейными комбинациями p из этих характеристических векторов.

10.2.3. Асимптотическая эффективность оценок наименьших квадратов

Пусть $\{u_t\}$ — случайный процесс с $\mathbb{E}u_t = 0$ и $\mathbb{E}u_t u_s = \sigma(t-s)$. Положим

$$(57) \quad y_t = \sum_{j=1}^p \beta_j z_{jt} + u_t, \quad t = 1, 2, \dots$$

Если последовательные значения $y_1, \dots, y_T$ для любого T образуют вектор y , удовлетворяющий соотношениям (1) и (2) с $\beta = (\beta_1, \dots, \beta_p)'$, $Z = (z_{it})'$ и $\Sigma = [\sigma(t-s)]$, то марковская оценка и оценка наименьших квадратов вектора β , а также ковариационные матрицы этих оценок выражаются соотношениями (3), (4), (5) и (6) соответственно. Нас интересует сейчас вопрос о том, при каких условиях обе ковариационные матрицы (5) и (6) будут асимптотически эквивалентны, в том смысле что

$$(58) \quad \lim_{T \rightarrow \infty} \text{нормированной } \{\mathbb{E}(\mathbf{b}^* - \beta)(\mathbf{b}^* - \beta)'\} = \\ = \lim_{T \rightarrow \infty} \text{нормированной } \{\mathbb{E}(\mathbf{b} - \beta)(\mathbf{b} - \beta)'\}.$$

Нормировка предполагается здесь одинаковой для обеих матриц и такой, что пределы в (58) нетривиальны. Мы ставим этот вопрос и для одной заданной ковариационной последовательности $\{\sigma(h)\}$, и для класса всех ковариационных функций с непрерывными положительными спектральными плотностями. Такие задачи впервые были рассмотрены Гренандером (1954), Розенблаттом (1956), Гренандером и Розенблаттом (1957).

Для того чтобы получить ответ на подобные вопросы, необходимо сделать определенные предположения о последовательности независимых переменных.

Пусть

$$(59) \quad a_{ij}^T(h) = \begin{cases} \sum_{t=1}^{T-h} z_{i,t+h} z_{jt}, & h = 0, 1, \dots, \\ \sum_{t=1-h}^T z_{i,t+h} z_{jt}, & h = 0, -1, \dots \end{cases}$$

Тогда $a_{ij}^T(h) = a_{ji}^T(-h)$. Необходимые предположения таковы.

Условие 10.2.1. $a_{ii}^T(0) \rightarrow \infty$ при $T \rightarrow \infty$, $i = 1, \dots, p$.

Условие 10.2.2.

$$(60) \quad \lim_{T \rightarrow \infty} \frac{z_{i,T+1}^2}{a_{ii}^T(0)} = 0, \quad i = 1, \dots, p.$$

УСЛОВИЕ 10.2.3. Предел при $T \rightarrow \infty$ отношения

$$(61) \quad \frac{a_{ij}^T(h)}{\sqrt{a_{ii}^T(0) a_{jj}^T(0)}} = r_{ij}^T(h)$$

существует для всех i, j, h , принимающих значения $i, j = 1, \dots, p$ и $h = 0, \pm 1, \dots$.

Пусть

$$(62) \quad \lim_{T \rightarrow \infty} r_{ij}^T(h) = \rho_{ij}(h), \quad i, j = 1, \dots, p, h = 0, \pm 1, \dots,$$

и $R(h) = [\rho_{ij}(h)]$.

УСЛОВИЕ 10.2.4. Матрица $R(0)$ не вырождена.

Первое предположение означает, что вклад каждой последовательности z_{it} неограниченно возрастает с ростом T . Благодаря второму предположению исключается возможность заметного вклада последних членов z_{it}^2 в сумму квадратов при больших T . Третье предположение приводит к тому, что корреляции между независимыми переменными при всех достаточно больших T мало отличаются от некоторых фиксированных значений. Наконец, четвертое условие вводится из соображений удобства вычислений.

Последовательность $\{R(h)\}$ является положительно определенной последовательностью эрмитовых матриц, т. е.

$$(63) \quad \sum_{h,k} c' R(h-k) \bar{c}_x x_k \geq 0,$$

где суммирование проводится по любому конечному множеству индексов, $\{x_h\}$ — произвольная последовательность действительных чисел, а c — произвольный комплексный вектор размерности p . Отсюда следует, что существует эрмитова матрица $M(\lambda)$ с положительно полуопределенными приращениями, такая, что

$$(64) \quad R(h) = \int_{-\pi}^{\pi} e^{i\lambda h} dM(\lambda).$$

(См. § 7.3.)

Пусть

$$(65) \quad D_T = \begin{bmatrix} \sqrt{a_{11}^T(0)} & 0 & \dots & 0 \\ 0 & \sqrt{a_{22}^T(0)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{a_{pp}^T(0)} \end{bmatrix}.$$

Тогда

$$(66) \quad R(h) = \lim_{T \rightarrow \infty} D_T^{-1} A_T(h) D_T^{-1},$$

где $A_T(h) = [a_{ij}^T(h)]$.

Для того чтобы найти пределы величин

$$(67) \quad D_T^{-1} Z' \Sigma Z D_T^{-1} = D_T^{-1} \sum_{h=-(T-1)}^{T-1} \sigma(h) A_T(h) D_T^{-1}$$

и

$$(68) \quad D_T^{-1} Z' \Sigma^{-1} Z D_T^{-1},$$

сделаем следующее предположение о последовательности ковариаций.

Условие 10.2.5. *Имеет место представление*

$$(69) \quad \sigma(h) = \int_{-\pi}^{\pi} e^{i\lambda h} f(\lambda) d\lambda$$

с функцией $f(\lambda)$, непрерывной на $[-\pi, \pi]$.

Указанные пределы найдем, аппроксимируя спектральную плотность $f(\lambda)$ тригонометрическими полиномами $f_L(\lambda) = f_L(-\lambda)$ и $f_U(\lambda) = f_U(-\lambda)$, такими, что

$$(70) \quad f_L(\lambda) \leq f(\lambda) \leq f_U(\lambda), \quad -\pi \leq \lambda \leq \pi.$$

Поскольку $0 \leq f(\lambda) \leq f_U(\lambda)$, то $f_U(\lambda)$ является спектральной плотностью. Ей соответствует ковариационная последовательность

$$\sigma_U(h) = \int_{-\pi}^{\pi} e^{i\lambda h} f_U(\lambda) d\lambda. \text{ Если } 0 \leq f_L(\lambda), \text{ то и } f_L(\lambda) \text{ будет спектральной плотностью, порождающей ковариационную последовательность}$$

$$\sigma_L(h) = \int_{-\pi}^{\pi} e^{i\lambda h} f_L(\lambda) d\lambda. \text{ Однако в ряде доказательств неотрицательность } f_L(\lambda) \text{ не является необходимой.}$$

Лемма 10.2.6. Обозначим $\Sigma_L = [\sigma_L(t-s)]$, $\Sigma_U = [\sigma_U(t-s)]$. Если выполнено соотношение (70), то для любого вектора x

$$(71) \quad x' \Sigma_L x \leq x' \Sigma x \leq x' \Sigma_U x.$$

Если, кроме того, $f_L(\lambda) > 0$, то для любого вектора x

$$(72) \quad x' \Sigma_U^{-1} x \leq x' \Sigma^{-1} x \leq x' \Sigma_L^{-1} x.$$

Доказательство. Прежде всего

$$(73) \quad x' \Sigma x = \sum_{t,s=1}^T \sigma(t-s) x_t x_s =$$

$$\begin{aligned}
&= \int_{-\pi}^{\pi} \sum_{t,s=1}^T e^{i\lambda(t-s)} x_t x_s f(\lambda) d\lambda = \\
&= \int_{-\pi}^{\pi} \left| \sum_{t=1}^T x_t e^{i\lambda t} \right|^2 f(\lambda) d\lambda.
\end{aligned}$$

Аналогичные выражения имеют место для $\mathbf{x}' \Sigma_L \mathbf{x}$ и $\mathbf{x}' \Sigma_U \mathbf{x}$. (Если $f_L(\lambda)$ при некотором значении λ отрицательно, то $\sigma_L(u)$ уже не будет ковариационной последовательностью. Однако и в этом случае соотношения (71) и (74) выполняются.) Из (70) вытекает, что

$$\begin{aligned}
(74) \quad \int_{-\pi}^{\pi} \left| \sum_{t=1}^T x_t e^{i\lambda t} \right|^2 f_L(\lambda) d\lambda &\leq \int_{-\pi}^{\pi} \left| \sum_{t=1}^T x_t e^{i\lambda t} \right|^2 f(\lambda) d\lambda \leq \\
&\leq \int_{-\pi}^{\pi} \left| \sum_{t=1}^T x_t e^{i\lambda t} \right|^2 f_U(\lambda) d\lambda.
\end{aligned}$$

Последнее равносильно (71). Что касается (72), то справедливость неравенства $\mathbf{x}' \Sigma \mathbf{x} \leq \mathbf{x}' \Sigma_U \mathbf{x}$ для всех $\mathbf{x}$ эквивалентна тому, что корни уравнения $|\Sigma - \theta \Sigma_U| = 0$ положительны и не превосходят 1. Они же будут являться и корнями уравнения $|\Sigma_U^{-1} - \theta \Sigma^{-1}| = 0$, а отсюда следует справедливость первого неравенства в (72) для всех $\mathbf{x}$. Аналогично доказывается и второе неравенство. ■

ТЕОРЕМА 10.2.6. При условиях 10.2.1, 10.2.2, 10.2.3 и 10.2.5

$$(75) \quad \lim_{T \rightarrow \infty} \mathbf{D}_T^{-1} \mathbf{Z}' \Sigma \mathbf{Z} \mathbf{D}_T^{-1} = 2\pi \int_{-\pi}^{\pi} f(\lambda) d\mathbf{M}(\lambda).$$

Доказательство. Идея состоит в том, чтобы доказать, что левая часть (75) равна $\sum_{h=-\infty}^{\infty} \sigma(h) \mathbf{R}(h)$. Последняя же сумма может быть представлена в виде $\sum_{h=-\infty}^{\infty} \int_{-\pi}^{\pi} \sigma(h) e^{i\lambda h} d\mathbf{M}(\lambda)$, а это и есть правая часть (75). Однако имеется некоторое затруднение, связанное с тем, что число слагаемых в сумме (67) неограниченно возрастает, и с тем, что мы исходим из непрерывности $f(\lambda)$, а не из условия $\sum_{h=-\infty}^{\infty} |\sigma(h)| < \infty$. [Предположение о непрерывности $f(\lambda)$ было сделано по той причине, что оно используется и при доказательстве теоремы 10.2.7.]

По теореме 7.5.2 для любого $\varepsilon > 0$ найдутся такие два тригонометрических полинома, $f_L(\lambda) = \sum_{k=-K}^K a_k e^{i\lambda k}$ и $f_U(\lambda) = \sum_{k=-K}^K b_k e^{i\lambda k}$

с $a_k = a_{-k}$ и $b_k = b_{-k}$, $k = 1, \dots, K$, что выполняется (70) и

$$(76) \quad f_U(\lambda) - f_L(\lambda) \leq \varepsilon, \quad -\pi \leq \lambda \leq \pi.$$

Далее,

$$(77) \quad \begin{aligned} \mathbf{x}' \Sigma_U \mathbf{x} &= \int_{-\pi}^{\pi} \left| \sum_{t=1}^T x_t e^{i\lambda t} \right|^2 f_U(\lambda) d\lambda = \\ &= \int_{-\pi}^{\pi} \sum_{k=-K}^K \sum_{t,s=1}^T b_k x_t x_s e^{i\lambda(k+t-s)} d\lambda = \\ &= 2\pi \sum_{k=-K}^K b_k \sum_{t \in S_k} x_{t+k} x_t, \end{aligned}$$

где $S_k = \{1, \dots, T-k\}$, $k \geq 0$, и $S_k = \{1-k, \dots, T\}$, $k \leq 0$. Для произвольного вектора γ размерности p положим $x_t = (z_{1t}, \dots, z_{pt})' \mathbf{D}_T^{-1} \gamma$. Тогда

$$(78) \quad \gamma' \mathbf{D}_T^{-1} \mathbf{Z}' \Sigma \mathbf{Z} \mathbf{D}_T^{-1} \gamma \leq 2\pi \sum_{k=-K}^K b_k \gamma' \mathbf{D}_T^{-1} \mathbf{A}_T(k) \mathbf{D}_T^{-1} \gamma.$$

Поэтому

$$(79) \quad \begin{aligned} \lim_{T \rightarrow \infty} \gamma' \mathbf{D}_T^{-1} \mathbf{Z}' \Sigma \mathbf{Z} \mathbf{D}_T^{-1} \gamma &\leq 2\pi \sum_{k=-K}^K b_k \gamma' \mathbf{R}(k) \gamma = \\ &= 2\pi \gamma' \sum_{k=-K}^K b_k \int_{-\pi}^{\pi} e^{i\lambda k} d\mathbf{M}(\lambda) \gamma = \\ &= 2\pi \gamma' \int_{-\pi}^{\pi} f_U(\lambda) d\mathbf{M}(\lambda) \gamma. \end{aligned}$$

Подобным же образом получаем, что

$$(80) \quad 2\pi \gamma' \int_{-\pi}^{\pi} f_L(\lambda) d\mathbf{M}(\lambda) \gamma \leq \lim_{T \rightarrow \infty} \gamma' \mathbf{D}_T^{-1} \mathbf{Z}' \Sigma \mathbf{Z} \mathbf{D}_T^{-1} \gamma.$$

Поскольку ε произвольно, из (79) и (80) вытекает

$$(81) \quad \lim_{T \rightarrow \infty} \gamma' \mathbf{D}_T^{-1} \mathbf{Z}' \Sigma \mathbf{Z} \mathbf{D}_T^{-1} \gamma = 2\pi \gamma' \int_{-\pi}^{\pi} f(\lambda) d\mathbf{M}(\lambda) \gamma.$$

Так как последнее соотношение выполняется для любого вектора γ , отсюда и следует утверждение теоремы. ■

Условие 10.2.6. Функция $f(\lambda)$, определяемая в (69), удовлетворяет соотношению $f(\lambda) > 0$, $-\pi \leq \lambda \leq \pi$.

ТЕОРЕМА 10.2.7. При условиях 10.2.1, 10.2.2, 10.2.3, 10.2.5 и 10.2.6

$$(82) \quad \lim_{T \rightarrow \infty} \mathbf{D}_T^{-1} \mathbf{Z}' \Sigma^{-1} \mathbf{Z} \mathbf{D}_T^{-1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{f(\lambda)} d\mathbf{M}(\lambda).$$

Доказательство. В силу следствия 7.5.2 для любого $\varepsilon > 0$ найдется такая пара спектральных плотностей, а именно, $f_L(\lambda) = [2\pi |\sum_{k=0}^K a_k e^{i\lambda k}|^2]^{-1}$ и $f_U(\lambda) = [2\pi |\sum_{k=0}^K b_k e^{i\lambda k}|^2]^{-1}$, соответствующих некоторой паре процессов авторегрессии, что при этом выполняются соотношения (70) и

$$(83) \quad \frac{1}{f_L(\lambda)} - \frac{1}{f_U(\lambda)} \leq \varepsilon, \quad -\pi \leq \lambda \leq \pi.$$

Функция $f_U(\lambda)$ является спектральной плотностью процесса $\{w_t\}$, удовлетворяющего стохастическому разностному уравнению

$$(84) \quad \sum_{k=0}^K b_k w_{t-k} = v_t, \quad t = \dots, -1, 0, 1, \dots,$$

в котором $\{v_t\}$ — последовательность некоррелированных случайных величин с нулевыми средними и единичными дисперсиями. Предположим, что $T > 2K$. Элементы σ_{TU}^{ts} матрицы Σ_U^{-1} равны

$$(85) \quad \begin{aligned} \sigma_{TU}^{ts} &= \sum_{k=0}^{\min(t,s)-1} b_k b_{k+|t-s|} = \\ &= \sigma_{TU}^{T+1-t, T+1-s}, \quad s, t = 1, \dots, K. \end{aligned}$$

Для $K < \max(t, s)$ и $\min(t, s) \leq T - K$ они равны

$$(86) \quad \sigma_{TU}^{ts} = \begin{cases} \sum_{k=0}^{K-|t-s|} b_k b_{k+|t-s|}, & 0 \leq |t-s| \leq K, \\ 0, & K < |t-s|. \end{cases}$$

(См. § 6.2.) Пусть

$$(87) \quad \mathbf{B} = \begin{bmatrix} b_{11} & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ b_{21} & b_{22} & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ b_{31} & b_{32} & b_{33} & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ b_{K1} & b_{K2} & b_{K3} & \dots & b_{KK} & 0 & 0 & \dots & 0 & 0 \\ b_K & b_{K-1} & b_{K-2} & \dots & b_1 & b_0 & 0 & \dots & 0 & 0 \\ 0 & b_K & b_{K-1} & \dots & b_2 & b_1 & b_0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & b_0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & b_1 & b_0 \end{bmatrix},$$

где $b_{11}, b_{21}, b_{22}, \dots, b_{KK}$ выбраны так, что

$$(88) \quad \Sigma_U^{-1} = \mathbf{B}'\mathbf{B}.$$

[Условием (88) треугольная матрица $\mathbf{B}$ определяется однозначно. Последние $T - K$ строк правой части (87) непосредственно восстанавливаются по имеющимся представлениям (85) и (86).] При этом

$$(89) \quad \begin{aligned} \mathbf{x}'\Sigma_U^{-1}\mathbf{x} &= \mathbf{x}'\mathbf{B}'\mathbf{B}\mathbf{x} = \\ &= \sum_{i=1}^K \left(\sum_{j=1}^i b_{ij}x_j \right)^2 + \sum_{t=K+1}^T \left(\sum_{j=0}^K b_j x_{t-j} \right)^2 = \\ &= \sum_{j,k=0}^K \sum_{t=K+1}^T b_j b_k x_{t-j} x_{t-k} + \sum_{i=1}^K \left(\sum_{j=1}^i b_{ij} x_j \right)^2. \end{aligned}$$

Для произвольного вектора γ размерности p положим $x_t = (z_{1t}, \dots, z_{pt}) \mathbf{D}_T^{-1} \gamma$. Тогда

$$(90) \quad \begin{aligned} \gamma' \mathbf{D}_T^{-1} \mathbf{Z}' \Sigma^{-1} \mathbf{Z} \mathbf{D}_T^{-1} \gamma &\geq \gamma' \mathbf{D}_T^{-1} \mathbf{Z}' \Sigma_U^{-1} \mathbf{Z} \mathbf{D}_T^{-1} \gamma = \mathbf{x}'\mathbf{B}'\mathbf{B}\mathbf{x} \geq \\ &\geq \sum_{j,k=0}^K \sum_{t=K+1}^T b_j b_k x_{t-j} x_{t-k} \end{aligned}$$

и

$$(91) \quad \begin{aligned} \lim_{T \rightarrow \infty} \gamma' \mathbf{D}_T^{-1} \mathbf{Z}' \Sigma^{-1} \mathbf{Z} \mathbf{D}_T^{-1} \gamma &\geq \\ &\geq \lim_{T \rightarrow \infty} \sum_{j,k=0}^K b_j b_k \sum_{t=K+1}^T \gamma' \mathbf{D}_T^{-1} \mathbf{z}_{t-j} \mathbf{z}_{t-k}' \mathbf{D}_T^{-1} \gamma = \\ &= \lim_{T \rightarrow \infty} \sum_{j,k=0}^K b_j b_k \gamma' \mathbf{D}_T^{-1} \mathbf{A}_T(k-j) \mathbf{D}_T^{-1} \gamma = \\ &= \sum_{j,k=0}^K b_j b_k \gamma' \mathbf{R}(k-j) \gamma = \\ &= \gamma' \sum_{i,k=0}^K b_i b_k \int_{-\pi}^{\pi} e^{i\lambda(k-i)} d\mathbf{M}(\lambda) \gamma = \\ &= \gamma' \int_{-\pi}^{\pi} \left| \sum_{j=0}^K b_j e^{i\lambda j} \right|^2 d\mathbf{M}(\lambda) \gamma = \\ &= \frac{1}{2\pi} \gamma' \int_{-\pi}^{\pi} \frac{1}{f_U(\lambda)} d\mathbf{M}(\lambda) \gamma. \end{aligned}$$

Аналогично получаем

$$(92) \quad \overline{\lim}_{T \rightarrow \infty} \gamma' D_T^{-1} Z' \Sigma^{-1} Z D_T^{-1} \gamma \leq \frac{1}{2\pi} \gamma' \int_{-\pi}^{\pi} \frac{1}{f_L(\lambda)} dM(\lambda) \gamma.$$

Поскольку ε выбирается произвольно, то из (91) и (92) вытекает, что

$$(93) \quad \lim_{T \rightarrow \infty} \gamma' D_T^{-1} Z' \Sigma^{-1} Z D_T^{-1} \gamma = \frac{1}{2\pi} \gamma' \int_{-\pi}^{\pi} \frac{1}{f(\lambda)} dM(\lambda) \gamma.$$

Это соотношение выполняется для каждого вектора γ , что и доказывает теорему.

Следствие 10.2.1. При выполнении условий 10.2.1—10.2.5

$$(94) \quad \lim_{T \rightarrow \infty} D_T \mathcal{E}(b^* - \beta) (b^* - \beta)' D_T = 2\pi R^{-1}(0) \int_{-\pi}^{\pi} f(\lambda) dM(\lambda) R^{-1}(0);$$

при выполнении условий 10.2.1—10.2.6

$$(95) \quad \lim_{T \rightarrow \infty} D_T \mathcal{E}(b - \beta) (b - \beta)' D_T = \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{f(\lambda)} dM(\lambda) \right]^{-1}.$$

Здесь использован тот факт, что $\lim_{T \rightarrow \infty} D_T^{-1} Z' Z D_T^{-1} = R(0)$.

Пусть

$$(96) \quad S(u) = \{\lambda \mid 2\pi f(\lambda) \leq u\}, \quad m \leq u \leq M,$$

$$(97) \quad T(u) = \int_{S(u)} dM(\lambda),$$

где $m = 2\pi \min f(\lambda)$ и $M = 2\pi \max f(\lambda)$, $-\pi \leq \lambda \leq \pi$. Поскольку $f(\lambda) = f(-\lambda)$, то множество $S(u)$ симметрично относительно нуля, так что матрицы $T(u)$ и $T(u_2) - T(u_1)$, $m \leq u_1 < u_2 \leq M$, действительны и положительно полуопределены. При этом

$$(98) \quad \lim_{T \rightarrow \infty} D_T^{-1} Z' Z D_T^{-1} = \int_m^M dT(u),$$

$$(99) \quad \lim_{T \rightarrow \infty} D_T^{-1} Z' \Sigma Z D_T^{-1} = \int_m^M u dT(u),$$

$$(100) \quad \lim_{T \rightarrow \infty} D_T^{-1} Z' \Sigma^{-1} Z D_T^{-1} = \int_m^M \frac{1}{u} dT(u).$$

Существует такая невырожденная матрица P , что $P'R(0)P = I$, $P' \cdot \lim_{T \rightarrow \infty} D_T^{-1} Z' \Sigma Z D_T^{-1} P = D$, где D — диагональная матрица

с элементами $d_{11} \geq \dots \geq d_{pp}$. Положим $\mathbf{L}(u) = \mathbf{P}'\mathbf{T}(u)\mathbf{P}$, $m \leq u \leq M$. Тогда матричная функция $\mathbf{L}(u)$ и ее приращения положительно полуопределены. Кроме того,

$$(101) \quad \mathbf{I} = \mathbf{P}' \lim_{T \rightarrow \infty} \mathbf{D}_T^{-1} \mathbf{Z}' \mathbf{Z} \mathbf{D}_T^{-1} \mathbf{P} = \int_m^M d\mathbf{L}(u),$$

$$(102) \quad \mathbf{D} = \mathbf{P}' \lim_{T \rightarrow \infty} \mathbf{D}_T^{-1} \mathbf{Z}' \Sigma \mathbf{Z} \mathbf{D}_T^{-1} \mathbf{P} = \int_m^M u d\mathbf{L}(u),$$

$$(103) \quad \mathbf{P}' \lim_{T \rightarrow \infty} \mathbf{D}_T^{-1} \mathbf{Z}' \Sigma^{-1} \mathbf{Z} \mathbf{D}_T^{-1} \mathbf{P} = \int_m^M \frac{1}{u} d\mathbf{L}(u).$$

Две интересующие нас предельные нормированные ковариационные матрицы будут совпадать тогда и только тогда, когда (103) равно $\mathbf{D}^{-1}$. Диагональные элементы матриц (101), (102) и (103) равны

$$(104) \quad 1 = \int_m^M dl_{ii}(u), \quad i = 1, \dots, p,$$

$$(105) \quad d_{ii} = \int_m^M u dl_{ii}(u), \quad i = 1, \dots, p,$$

$$(106) \quad \int_m^M \frac{1}{u} dl_{ii}(u), \quad i = 1, \dots, p.$$

В силу того, что приращения функций $l_{ii}(u)$ неотрицательны, $l_{ii}(u)$ для каждого i является функцией распределения вероятностей. Выражение (105) можно рассматривать как математическое ожидание случайной величины, имеющей такое распределение, а (106) — как математическое ожидание обратной ей по значениям случайной величины. Если (103) есть $\mathbf{D}^{-1}$, то (106) равно $1/d_{ii}$, и из леммы 10.2.1 вытекает, что для каждого i функция $l_{ii}(u)$ имеет только одну точку роста, причем в этой точке она имеет скачок величины 1.

Обратно, пусть каждая из функций $l_{ii}(u)$ имеет только одну точку роста. В соответствии с (104), в этой точке будет скачок величины 1. Пусть скачки диагональных элементов матричной функции $\mathbf{L}(u)$ соответствуют значениям u , среди которых различными являются $u_1 > \dots > u_g$. Поскольку приращения $\mathbf{L}(u)$ положительно полуопределены, возрастание недиагональных элементов может происходить только путем скачков при тех же значениях $u = u_1, \dots, u_g$. Пусть $\mathbf{L}_j$ обозначает приращение $\mathbf{L}(u)$ в точке $u = u_j$, $j =$

$= 1, \dots, G$. Тогда (101) принимает вид

$$(107) \quad \sum_{j=1}^G L_j = I.$$

Матрицы L_j соответствуют отличным от нуля матрицам $C^{(h)}$ из доказательства теоремы 10.2.2. При этом j -й диагональный блок матрицы L_j является единичной матрицей, порядок которой равен числу диагональных элементов $L(u)$, возрастающих в точке u_j . Все другие блоки L_j являются нулевыми матрицами. Ранг матрицы L_j совпадает с порядком единичного блока. Соотношения (101) и (102) принимают вид

$$(108) \quad \sum_{j=1}^G u_j L_j = D,$$

$$(109) \quad \sum_{j=1}^G \frac{1}{u_j} L_j = D^{-1}.$$

ТЕОРЕМА 10.2.8. Для асимптотической эффективности оценок наименьших квадратов при условиях 10.2.1—10.2.6 необходимо и достаточно, чтобы каждый диагональный элемент матричной функции $L(u)$ возрастал только в одной точке.

Отметим, что число диагональных элементов матрицы L_j , равных 1, совпадает с рангом L_j .

СЛЕДСТВИЕ 10.2.2. Для асимптотической эффективности оценок наименьших квадратов необходимо и достаточно, чтобы матричная функция $L(u)$ имела G , $G \leq r$, точек роста, где r — сумма рангов приращений функции $L(u)$.

Множества $S_j = \{\lambda \mid 2\pi f(\lambda) = u_j, -\pi \leq \lambda \leq \pi\}$, $j = 1, \dots, G$, являются множествами значений λ , на которых возрастает $M(\lambda)$. Поэтому

$$(110) \quad P' \int_{S_j} dM(\lambda) P = L_j, \quad j = 1, \dots, G.$$

ТЕОРЕМА 10.2.9. Для асимптотической эффективности оценок наименьших квадратов при условиях 10.2.1—10.2.6 необходимо и достаточно, чтобы на множестве тех λ , при которых возрастает $M(\lambda)$, спектральная плотность $f(\lambda)$ принимала всего G , $G \leq r$, значений, где r равно сумме рангов интегралов $\int dM(\lambda)$ по тем G множествам значений λ , на которых $f(\lambda)$ принимает эти G значений.

Равные единице диагональные элементы матрицы L_j должны располагаться рядом, а диагональные блоки, порождаемые ими,

должны составлять единичные подматрицы. Учитывая это, соотношение (108) можно переписать в виде

$$(111) \quad D = \begin{pmatrix} u_1 I & 0 & \dots & 0 \\ 0 & u_2 I & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & u_G I \end{pmatrix},$$

где матрица $u_j I$ имеет порядок, равный рангу L_j . В силу того, что единичные матрицы в $L_1, \dots, L_G$ не перекрываются,

$$(112) \quad L_i L_j = 0, \quad i \neq j.$$

Множества S_j , $j = 1, \dots, G$, значений λ , при которых возрастает $M(\lambda)$, называются *элементами спектра* $M(\lambda)$. Объединение этих множеств называют *спектром* $M(\lambda)$. Поскольку множество S_j симметрично, интеграл (110) содержит только действительную часть $M(\lambda)$. Фактически соотношение (110) имеет вид

$$(113) \quad 2P' \int_{S_j \cap (0, \pi]} d\mathcal{R}M(\lambda) P = L_j, \quad 0 \notin S_j,$$

$$(114) \quad P' \mathcal{R} [M(0+) - M(0-)] P + 2P' \int_{S_j \cap (0, \pi]} d\mathcal{R}M(\lambda) P = L_j, \quad 0 \in S_j.$$

Отметим, что $\int dM(\lambda) = R(0)$, где интегрирование производится по объединению всех G указанных множеств значений λ , и что последовательность G матриц $\int_{S_j} dM(\lambda)$ является ортогональной в метрике $R(0)$ в силу (112).

ТЕОРЕМА 10.2.10. При условиях 10.2.1—10.2.4 оценки наименьших квадратов асимптотически эффективны для всех стационарных процессов с непрерывными положительными спектральными плотностями тогда и только тогда, когда $M(\lambda)$ возрастает не более, чем при p значениях λ , $0 \leq \lambda \leq \pi$, где p — сумма рангов приращений функции $M(\lambda)$.

Доказательство. Теорема 10.2.9 может оставаться справедливой для всех указанных спектральных плотностей только в том случае, если матрица $\int dM(\lambda)$ положительно полуопределена не более чем в p точках. ■

Следует отметить, что если действительная часть $M(\lambda)$ не возрастает на некотором множестве значений λ , то на этом множестве не возрастает также и мнимая часть $M(\lambda)$. Поэтому указанный критерий можно формулировать как с использованием самой $M(\lambda)$, так и с использованием лишь ее действительной части.

Если спектральная плотность в рассматриваемой задаче неизвестна, то марковские оценки получить не удастся и вместо них можно вычислять оценки наименьших квадратов. Единственная возможность гарантировать асимптотическую эффективность последних состоит в использовании последовательности независимых переменных, удовлетворяющей условиям теоремы 10.2.10. Рассмотрим теперь некоторые примеры. Предположим, что

$$(115) \quad z_t = \alpha_0 + \sum_{j=1}^H (\alpha_j \cos \nu_j t + \beta_j \sin \nu_j t) + \alpha_{H+1} (-1)^t.$$

Тогда

$$(116) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^{T-h} z_t + h z_t' =$$

$$= \alpha_0 \alpha_0' + \frac{1}{2} \sum_{j=1}^H [\cos \nu_j h (\alpha_j \alpha_j' + \beta_j \beta_j') +$$

$$+ \sin \nu_j h (\beta_j \alpha_j' - \alpha_j \beta_j')] + (-1)^h \alpha_{H+1} \alpha_{H+1}'.$$

При этом $R(h) = \sum_{j=0}^{H+1} \cos \nu_j h M_j + \sum_{j=1}^H \sin \nu_j h M_j^*$, где $M_0 = \Gamma \alpha_0 \alpha_0' \Gamma$, $M_j = (1/2) \Gamma (\alpha_j \alpha_j' + \beta_j \beta_j') \Gamma$, $M_j^* = (1/2) \Gamma (\beta_j \alpha_j' - \alpha_j \beta_j') \Gamma$, $M_{H+1} = \Gamma \alpha_{H+1} \alpha_{H+1}' \Gamma$.

Здесь Γ — диагональная матрица, диагональные элементы которой являются величинами, обратными квадратным корням из диагональных элементов матрицы $\frac{1}{2} \sum_{j=0}^{H+1} \alpha_j \alpha_j' + \left(\frac{1}{2}\right) \sum_{j=1}^H \beta_j \beta_j'$, $\nu_0 = 1$ и $\nu_{H+1} = \pi$.

В качестве другого примера рассмотрим независимые переменные $z_{jt} = t^{j-1}$, $j = 1, \dots, p$. Тогда

$$(117) \quad a_{jk}^T(h) = \sum_{t=1}^{T-h} (t+h)^{j-1} t^{k-1} =$$

$$= \sum_{t=1}^{T-h} [t^{j+k-2} + (j-1) h t^{j+k-3} +$$

$$+ \dots + h^{j-1} t^{k-1}],$$

$$h = 0, 1, \dots, T-1.$$

и $a_{jk}^T(0) = \sum_{t=1}^T t^{j+k-2}$ приближенно равно $T^{j+k-1}/(j+k-1)$.
(См. упр. 14.) Таким образом, при каждом $h = 0, \pm 1, \dots$,

$$(118) \quad \rho_{jk}(h) = \lim_{T \rightarrow \infty} \frac{a_{jk}^T(h)/T^{j+k-1}}{(a_{jj}^T(0)/T^{2j-1})^{1/2} (a_{kk}^T(0)/T^{2k-1})^{1/2}} = \\ = \frac{\sqrt{2j-1} \sqrt{2k-1}}{j+k-1}, \quad j, k = 1, \dots, p.$$

Последнее выражение не зависит от h и поэтому $\mathbf{M}(\lambda)$ имеет единственный скачок в точке $\lambda = 0$, равный матрице

$$(119) \quad \mathbf{M}_0 = \left(\frac{\sqrt{2j-1} \sqrt{2k-1}}{j+k-1} \right)$$

ранга p .

Если z_{jt} является полиномом степени $j-1$ с положительным старшим коэффициентом, то получается та же самая последовательность $\{\mathbf{R}(h)\}$ и тот же результат относительно $\mathbf{M}(\lambda)$. Это связано с тем, что при указанном условии старший член доминирует над остальными. Если z_{jt} являются полиномами степени d_j с положительными старшими коэффициентами, $d_1 < d_2 < \dots < d_p$, то в этом случае

$$(120) \quad \rho_{jk}(h) = \frac{\sqrt{2d_j+1} \sqrt{2d_k+1}}{d_j+d_k+1}, \quad j, k = 1, \dots, p, \\ h = 0, \pm 1, \dots,$$

$\mathbf{M}(\lambda)$ имеет скачок при $\lambda = 0$, и (j, k) -й элемент $\mathbf{M}_0$ совпадает с $\rho_{jk}(0)$.

Объединим теперь оба примера. Именно пусть переменные z_{jt} разбиты по индексу j на $H+2$ множества, каждое из которых соответствует некоторой частоте λ_k , $k = 0, 1, \dots, H+1$. Будем предполагать, что $0 = \lambda_0 < \lambda_1 < \dots < \lambda_{H+1} = \pi$ и что частоте λ_k соответствует m_k переменных z_{jt} . Пусть в множестве с индексом $k = 0$ переменные z_{jt} имеют вид $P_j(t)$, $j = 1, \dots, m_0$, где $P_j(t)$ — полином степени d_j с положительным старшим коэффициентом. Пусть в множестве с индексом k , $k = 1, \dots, H$, переменные z_{jt} имеют вид $P_{ki}(t) \cos \lambda_k t$ и $Q_{ki}(t) \sin \lambda_k t$, где $i = 1, \dots, (m_k/2)$, m_k четное и при этом $P_{ki}(t)$ и $Q_{ki}(t)$ — полиномы степени d_{ki} с положительными старшими коэффициентами. Наконец, в последнем из указанных множеств переменные z_{jt} пусть имеют вид $(-1)^t Q_i(t)$, $i = 1, \dots, m_{H+1}$, где $Q_i(t)$ — полином степени $d_{H+1,i}$.

При таких условиях $\rho_{gj}(h) = 0$, если g и j находятся в различных множествах. Если j и k находятся в первом множестве, то $\rho_{jk}(h)$ задается соотношением (120). Для g и j , находящихся во множестве

с индексом k , $k = 1, \dots, H$ значения $\rho_{gj}(h)$ составляют матрицу

$$(121) \quad \begin{pmatrix} \cos \lambda_k h \mathbf{R}_k & -\sin \lambda_k h \mathbf{R}_k \\ \sin \lambda_k h \mathbf{R}_k & \cos \lambda_k h \mathbf{R}_k \end{pmatrix}$$

размера $m_k \times m_k$. Здесь (i, l) -й элемент матрицы $\mathbf{R}_k$ равен $\sqrt{2d_{ki} + 1} \sqrt{2d_{kl} + 1} / (d_{ki} + d_{kl} + 1)$, члены с $\cos \lambda_k t$ находятся в первом, а члены с $\sin \lambda_k t$ — во втором подмножестве множества с индексом k . Наконец, если g и j находятся в последнем множестве, то $\rho_{gj}(h) = (-1)^h \sqrt{2d_{H+1,g} + 1} \sqrt{2d_{H+1,j} + 1} / (d_{H+1,g} + d_{H+1,j} + 1)$. Спектральная функция возрастает только скачками. Скачок в точке $\lambda = 0$ имеет элементы (120) для $j, k = 1, \dots, m_0$ и нули на остальных позициях. Скачки при $\lambda = \pm \lambda_k$ имеют в качестве $(k+1)$ -й диагональной подматрицы матрицу (121), в которой $\cos \lambda_k h$ заменен на $(1/2)$, $\sin \lambda_k h$ — на $\mp (i/2)$, а остальные элементы — на нули. Скачки при $\lambda = \pm \pi$ имеют элементы $(1/2) \sqrt{2d_{H+1,g} + 1} \times \sqrt{2d_{H+1,j} + 1} / (d_{H+1,g} + d_{H+1,j} + 1)$ в последней диагональной подматрице и нули вне ее.

По-видимому, простейшим является случай $p = 1$, $z_t = 1$, $t = 1, 2, \dots$. В этом случае $a^T(h) = T - h$, $h = 0, 1, \dots, T$,

$$(122) \quad \rho(h) = \lim_{T \rightarrow \infty} \frac{a^T(h)}{T} = 1, \quad h = 0, \pm 1, \dots,$$

и каждый элемент $\mathbf{M}(\lambda)$ имеет единичный скачок при $\lambda = 0$. Оценкой наименьших квадратов для β , $\mathcal{E}y_t = \beta$, является $b^* = \sum_{t=1}^T y_t / T = \bar{y}$, и эта оценка асимптотически эффективна.

Предельная ковариационная последовательность (122) и соответствующая ей спектральная функция $\mathbf{M}(\lambda)$ возникают и в другом случае, не равносильном предыдущему. Именно, пусть $\mathcal{E}y_t = \beta \omega_t$, $\omega_t = 1, 2, \dots$. Пусть ω_t принимает значения a и b , причем $a \neq \pm b$ и

$$(123) \quad \frac{m}{n} a^2 + \frac{n-m}{n} b^2 = 1,$$

где m и n — целые числа, $0 < m < n$. Составим последовательность $\{\omega_t\}$ таким образом, чтобы предельные относительные частоты появлений в ней a и b соответственно равнялись m/n и $(n-m)/n$. С этой целью положим $\omega_t = a$ для $t = 1, \dots, m$ и $\omega_t = b$ для $t = m+1, \dots, n$. Затем возьмем $\omega_t = a$ для $t = n+1, \dots, n+2m$ и $\omega_t = b$ для $t = n+2m+1, \dots, n+2n = 3n$ и т. д. Последовательность $\{\omega_t\}$ будет состоять из аналогичных конечных подпоследовательностей длины $n, 2n, 3n, \dots$, причем, в j -й подпоследовательности первые jm элементов равны a , а последние $j(n-m)$ элементов равны b . При $T \rightarrow \infty$ предельная относительная частота появлений a равна

m/n , поскольку приращение числа появлений a в пределах $(j+1)$ -й подпоследовательности равно лишь $(j+1)m$, тогда как общее число элементов в первых j подпоследовательностях равно $j(j+1)n/2$.

Поэтому пределом среднего $\sum_{i=1}^T \omega_i^2/T$ служит (123). Более того, для этой последовательности выполняется и (122). Это вытекает из того факта, что при $j > (h/m)$, $j > h/(n-m)$ и ω_i из j -й последовательности точно $2h$ произведений $\omega_i \omega_{i+h}$ равны ab , точно $jm-h$ произведений $\omega_i \omega_{i+h}$ равны a^2 и точно $j(n-m)-h$ таких произведений равны b^2 . Таким образом, здесь $\rho(h) = 1$ и каждый элемент $\mathbf{M}(\lambda)$ имеет единственный скачок при $\lambda = 0$. Полученный результат показывает, что спектральная функция (или, что равносильно, последовательность ковариаций), не определяет однозначно последовательности независимых переменных. Она не определяет даже предельных частот этих переменных.

Более того, при оценивании β одну из указанных последовательностей нельзя заменять другой. Дело в том, что эти последовательности существенно различаются, так как

$$(124) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=1}^T z_i \omega_i = \frac{m}{n} a + \frac{n-m}{n} b = c$$

отличен от 1. Если $\mathcal{Y}_t = \beta \omega_t$ и если при этом для оценки β использовать $\sum_{i=1}^T z_i y_i/T = \bar{y}$ ($z_i \equiv 1$), то хотя такая оценка и будет иметь

дисперсию, асимптотически эквивалентную дисперсии $\sum_{i=1}^T \omega_i y_i/T$, однако она будет асимптотически смещенной, поскольку $\lim_{T \rightarrow \infty} \bar{y} =$

$$= \lim_{T \rightarrow \infty} \sum_{i=1}^T \beta \omega_i/T = c\beta \neq \beta.$$

Положение здесь существенно отличается от конечномерного случая. Предположим, что $\mathbf{w}'\mathbf{w} = 1$, $\mathbf{w}'\Sigma\mathbf{w} = \lambda^*$ и $\mathbf{w}'\Sigma^{-1}\mathbf{w} = 1/\lambda^*$, где λ^* — простой характеристический корень матрицы Σ , соответствующий характеристическому вектору $\mathbf{v}$, т. е. $\Sigma\mathbf{v} = \lambda^*\mathbf{v}$ и $\Sigma^{-1}\mathbf{v} =$

$$= (1/\lambda^*)\mathbf{v}. \text{ Поскольку } \Sigma = \sum_{i=1}^T \lambda_i \mathbf{v}_i \mathbf{v}_i', \text{ то } \mathbf{w}'\Sigma\mathbf{w} = \sum_{i=1}^T \lambda_i (\mathbf{v}_i'\mathbf{w})^2 \text{ и } \mathbf{w}'\Sigma^{-1}\mathbf{w} = \sum_{i=1}^T (1/\lambda_i) (\mathbf{v}_i'\mathbf{w})^2, \text{ так, что } \mathbf{v}_i'\mathbf{w} = 1 \text{ для } \lambda_i = \lambda^* \text{ и}$$

$\mathbf{v}_i'\mathbf{w} = 0$ для $\lambda_i \neq \lambda^*$. Поэтому в случае $\mathcal{Y} = \beta\mathbf{w}$ оценка $\mathbf{v}'\mathbf{y}$ будет несмещенной эффективной оценкой параметра β . Действительно, $\mathcal{E}\mathbf{v}'\mathbf{y} = \beta\mathbf{v}'\mathbf{w} = \beta$ и $\text{Var}(\mathbf{v}'\mathbf{y}) = \mathbf{v}'\Sigma\mathbf{v} = \lambda^*$. Аналогом вектора $\mathbf{v}$ будет последовательность $\{1\}$, аналогом $\mathbf{v}_t$ — последовательность

$\{\cos \lambda t, \sin \lambda t\}$, и аналогом представления $\Sigma = \sum_{i=1}^T \lambda_i \mathbf{v}_i \mathbf{v}_i'$ — представление $\sigma(t-s) = \int_{-\pi}^{\pi} e^{i\lambda(t-s)} f(\lambda) d\lambda$. Соотношениям $(\mathbf{v}_i' \mathbf{w})^2 = 1$ для $\lambda_i = \lambda^*$ и $(\mathbf{v}_i' \mathbf{w})^2 = 0$ для $\lambda_i \neq \lambda^*$ соответствует соотношение

$$(125) \quad \lim_{T \rightarrow \infty} \int_{-\pi}^{\pi} \frac{1}{T} \left| \sum_{i=1}^T w_i e^{i\lambda_i t} \right|^2 f(\lambda) d\lambda = 2\pi f(0),$$

справедливое для любой непрерывной функции $f(\lambda)$.

Асимптотически функции $\cos vt$ и $\sin vt$ можно рассматривать в некотором смысле как характеристические векторы всех ковариационных матриц стационарных процессов. Дело в том, что при $T \rightarrow \infty$

$$\begin{aligned} (126) \quad \sum_{s=1}^T \sigma(t-s) (\cos vs + i \sin vs) &= \sum_{s=1}^T \sigma(t-s) e^{ivs} = \\ &= e^{ivt} \sum_{s=1}^T \sigma(s-t) e^{iv(s-t)} = \\ &= e^{ivt} \sum_{h=-(t-1)}^{T-t} \sigma(h) e^{ivh} \rightarrow \\ &\rightarrow e^{ivt} \sum_{h=-(t-1)}^{\infty} \sigma(h) e^{ivh}. \end{aligned}$$

Для достаточно больших t правая часть (126) сколь угодно близка к $2\pi f(v) e^{ivt} = 2\pi f(v) (\cos vt + i \sin vt)$.

Если последовательность $\{\mathbf{R}(h)\}$ такова, что функция $\mathbf{M}(\lambda)$ абсолютно непрерывна, то в этом случае оценки наименьших квадратов не могут быть асимптотически эффективными, если спектральная плотность $f(\lambda)$ принимает на спектре $\mathbf{M}(\lambda)$ более p значений.

Так, например, если $\sum_{h=-\infty}^{\infty} |\rho_{ii}(h)| < \infty$, $i = 1, \dots, p$, то $\mathbf{R}(h) =$

$$= \int_{-\pi}^{\pi} e^{i\lambda h} \mathbf{m}(\lambda) d\lambda, \text{ где } m_{jk}(\lambda) = \sum_{s=-\infty}^{\infty} \rho_{jk}(s) e^{i\lambda s} / (2\pi). \text{ Здесь } \mathbf{M}(\lambda)$$

не имеет скачков. Грубо говоря, это будет в том случае, когда $\{\mathbf{z}_t\}$ является реализацией стационарного (векторного) процесса.

10.2.4 Асимптотическая нормальность оценок

При условиях 10.2.1—10.2.4 в § 2.6 было показано, что оценки наименьших квадратов имеют асимптотически нормальное распределение, если независимые ошибки удовлетворяют условию типа Линдеберга. Мы докажем соответствующую теорему, когда случайные составляющие имеют вид $u_t = \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$ и независимые случайные величины v_t удовлетворяют условию типа Линдеберга. Подобная, но более слабая теорема была доказана Хеннаном (1961).

ТЕОРЕМА 10.2.11. Пусть $y_t = \beta' z_t + u_t$, $t = 1, 2, \dots$, где $u_t = \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s}$, v_t — независимые случайные величины со средними 0, дисперсиями σ^2 и функциями распределения $F_t(v)$, $t = \dots, -1, 0, 1, \dots$ и $\sum_{s=-\infty}^{\infty} |\gamma_s| < \infty$. Предположим, что выполнены условия 10.2.1—10.2.4 и, кроме того, при $c \rightarrow \infty$

$$(127) \quad \sup_{t=1,2,\dots} \int_{|v|>c} v^2 dF_t(v) \rightarrow 0.$$

Тогда $D_T(b^* - \beta)$ имеет предельное нормальное распределение со средним 0 и ковариационной матрицей (94).

Доказательство. Поскольку

$$(128) \quad D_T(b^* - \beta) = [D_T^{-1} A_T(0) D_T^{-1}]^{-1} D_T^{-1} \sum_{t=1}^T z_t u_t,$$

предельное распределение $D_T(b^* - \beta)$ совпадает с предельным распределением умноженного на $R^{-1}(0)$ вектора $c_T = D_T^{-1} \sum_{t=1}^T z_t u_t$.

Положим

$$(129) \quad c_{T,k} = D_T^{-1} \sum_{t=1}^T z_t \sum_{s=-k}^k \gamma_s v_{t-s}.$$

Тогда

$$(130) \quad \begin{aligned} & \mathcal{E}(c_T - c_{T,k})'(c_T - c_{T,k}) = \\ &= \text{tr } D_T^{-1} \sum_{t,t'=1}^T z_t z_{t'}' \sum_{|s|,|s'|>k} \gamma_s \gamma_{s'} \mathcal{E} v_{t-s} v_{t'-s'} D_T^{-1} = \\ &= \sigma^2 \text{tr } D_T^{-1} \sum_{|s|,|s'|>k} \gamma_s \gamma_{s'} \sum_t z_t z_{t-s+s'}' D_T^{-1}, \end{aligned}$$

где суммирование по t производится по тем значениям t , для которых t и $t-s+s'$ заключены в пределах от 1 до T . (Возможность

внесения математического ожидания под знак суммы по s и s' доказывается таким же способом, как и в разд. 8.4.3.) Отсюда получаем оценку сверху

$$(131) \quad \mathbb{E} (c_T - c_{T,k})' (c_T - c_{T,k}) = \\ = \sigma^2 \sum_{|s|, |s'| > k} \gamma_s \gamma_{s'} \sum_{i=1}^p r_{ii}^T (s - s') \leq \rho \sigma^2 \left(\sum_{|s| > k} |\gamma_s| \right)^2,$$

которая стремится к нулю при $k \rightarrow \infty$.

Для произвольного вектора α рассмотрим

$$(132) \quad \alpha' c_{T,k} = \alpha' D_T^{-1} \sum_{t=1}^T z_t \sum_{s=-k}^k \gamma_s v_{t-s} = \\ = \sum_{r=k+1}^{T-k} \alpha' D_T^{-1} \sum_{s=-k}^k \gamma_s z_{r+s} v_r + \sum_{r=1-k}^k \alpha' D_T^{-1} \sum_{s=1-r}^k \gamma_s z_{r+s} v_r + \\ + \sum_{r=T-k+1}^{T+k} \alpha' D_T^{-1} \sum_{s=-k}^{T-r} \gamma_s z_{r+s} v_r.$$

Сумма квадратов коэффициентов при $v_{1-k}, \dots, v_k, v_{T-k+1}, \dots, v_{T+k}$ в силу условия 10.2.2 сходится к 0 при $T \rightarrow \infty$. Поэтому предельное распределение $\alpha' c_{T,k}$ совпадает с предельным распределением умноженной на $\sigma \left[\sum_{s,s'=-k}^k \gamma_s \gamma_{s'} \alpha' R (s-s') \alpha \right]^{1/2}$ суммы $\sum_{r=k+1}^{T-k} w_r^T$, где

$$(133) \quad w_r^T = \frac{\alpha' D_T^{-1} \sum_{s=-k}^k \gamma_s z_{r+s}}{\sigma \left(\sum_{r=k+1}^{T-k} \left(\alpha' D_T^{-1} \sum_{s=-k}^k \gamma_s z_{r+s} \right)^2 \right)^{1/2}} v_r.$$

Сумма квадратов коэффициентов при $v_{k+1}, \dots, v_{T-k}$ в (132) равна

$$(134) \quad \sum_{s,s'=-k}^k \gamma_s \gamma_{s'} \alpha' D_T^{-1} \sum_{r=k+1}^{T-k} z_{r+s} z'_{r+s'} D_T^{-1} \alpha,$$

что и дает в пределе $\sum_{s,s'=-k}^k \gamma_s \gamma_{s'} \alpha' R (s-s') \alpha$. Из центральной предель-

ной теоремы Линдберга (теорема 7.7.2) вытекает, что $\sum_{r=k+1}^{T-k} w_r^T$ имеет предельное нормальное распределение со средним 0 и дисперсией 1 (в соответствии с доказательством теоремы 2.6.1). Применение следствия 7.7.1 приводит к утверждению теоремы. ■

Теорема 10.2.11 показывает, что при достаточно большой длине временного ряда с оценками наименьших квадратов можно обращаться так, как если бы они были нормально распределенными.

Предельная ковариационная матрица равна

$$(135) \quad \lim_{T \rightarrow \infty} \mathbf{D}_T \mathbf{E}(\mathbf{b}^* - \boldsymbol{\beta})(\mathbf{b}^* - \boldsymbol{\beta})' \mathbf{D}_T = \mathbf{R}^{-1}(0) \sum_{h=-\infty}^{\infty} \sigma(h) \mathbf{R}(h) \mathbf{R}^{-1}(0),$$

если $\sum_{h=-\infty}^{\infty} |\sigma(h)| < \infty$. Ее можно оценить состоятельно с помощью оценок, аналогичных оценкам $2\pi f(\lambda) = \sum_{h=-\infty}^{\infty} \sigma(h) \cos \lambda h$, именно

$$(136) \quad \mathbf{D}_T \mathbf{A}_T^{-1}(0) \sum_{h=-K_T}^{K_T} k\left(\frac{h}{K_T}\right) c_{h,T}^* \mathbf{A}_T(h) \mathbf{A}_T^{-1}(0) \mathbf{D}_T,$$

где $c_{h,T}^*$ — оценка $\sigma(h)$, основанная на остатках от $\mathbf{b}^{*'} \mathbf{z}_t$, $t = 1, \dots, T$, а $k(x)$ — ядро, типа изучавшегося в разд. 9.3.2. На основании этого можно строить критерии для проверки гипотез о векторе $\boldsymbol{\beta}$ и доверительные интервалы для компонент этого вектора. [Многомерный случай рассматривается у Хеннана (1963)].

10.3. ОЦЕНИВАНИЕ КОВАРИАЦИЙ И СПЕКТРАЛЬНОЙ ПЛОТНОСТИ ПО ОСТАТКАМ ОТ ТРЕНДОВ

10.3.1. Оценивание ковариационных матриц

Рассмотрим снова модель $\mathbf{E} \mathbf{y} = \mathbf{Z} \boldsymbol{\beta}$, $\mathbf{E}(\mathbf{y} - \mathbf{Z} \boldsymbol{\beta})(\mathbf{y} - \mathbf{Z} \boldsymbol{\beta})' = \boldsymbol{\Sigma}$ и займемся вопросами оценивания ковариационной матрицы $\boldsymbol{\Sigma}$. В разд. 10.3.2 мы изучим асимптотическую теорию для случая, когда ковариации соответствуют некоторому стационарному случайному процессу. Асимптотическая теория для выборочной спектральной плотности и для оценок спектральной плотности, основанных на остатках от тренда, подобранного по методу наименьших квадратов, рассматривается в разд. 10.3.3 и 10.3.4 соответственно. Эта теория обобщает проведенное в § 8.3, 8.4 и 9.4 исследование указанных статистик, построенных по остаткам от выборочных средних.

Оценкой наименьших квадратов вектора $\boldsymbol{\beta}$ является

$$(1) \quad \mathbf{b}^* = (\mathbf{Z}' \mathbf{Z})^{-1} \mathbf{Z}' \mathbf{y}.$$

Остатки от подобранной регрессии образуют при этом T -компонентный вектор

$$(2) \quad \mathbf{y} - \mathbf{Z} \mathbf{b}^* = [\mathbf{I} - \mathbf{Z}(\mathbf{Z}' \mathbf{Z})^{-1} \mathbf{Z}'] \mathbf{y},$$

имеющий математическое ожидание $\mathbf{0}$ (ввиду того что $\mathbf{E} \mathbf{b}^* = \boldsymbol{\beta}$). Компоненты этого вектора могут быть использованы для оценива-

ния самой матрицы Σ или функций от Σ . Если о структуре матрицы Σ предварительно ничего не известно, то для ее оценивания можно использовать матрицу $(y - Zb^*)(y - Zb^*)'$. Математическое ожидание последней равно

$$(3) \quad \mathbb{E}(y - Zb^*)(y - Zb^*)' = [I - Z(Z'Z)^{-1}Z'] \times \\ \times \Sigma [I - Z(Z'Z)^{-1}Z']$$

и отлично от Σ . Ранг этой оценки равен $T - p$, в то время как сама Σ имеет ранг T . Если только Σ не обладает специфической структурой, то число подлежащих оценке параметров в Σ будет равно $T(T + 1)/2$. В то же время, в $y - Zb^*$ имеется лишь $T - p$ линейно независимых компонент. (Вектор $y - Zb^*$ ортогонален вектору Zb^* .)

Смысл соотношения (3) можно лучше представить себе в случае, когда оценки наименьших квадратов оказываются эффективными. Именно в этом случае матрица Z представима в виде $Z = V^*C$, где C — невырожденная матрица, p столбцов которой является p характеристическими векторами Σ , соответствующими некоторым p ее характеристическим корням $\lambda_{s_1}, \dots, \lambda_{s_p}$. (Мы предполагаем, что при наличии кратных характеристических корней характеристические векторы определяются таким образом, что среди них содержатся все столбцы матрицы V^* .) Пусть $S = \{s_1, \dots, s_p\}$ — некоторое подмножество целых чисел от 1 до T . Тогда

$$(4) \quad I - Z(Z'Z)^{-1}Z' = I - V^*V'^* = \sum_{s \notin S} v_s v_s',$$

$$(5) \quad [I - Z(Z'Z)^{-1}Z'] \Sigma [I - Z(Z'Z)^{-1}Z'] = \\ = \sum_{s \notin S} v_s v_s' \sum_{t=1}^T \lambda_t v_t v_t' \sum_{r \notin S} v_r v_r' = \sum_{t \notin S} \lambda_t v_t v_t'.$$

Полагая $y = Vx$ ($x = V'y$), получаем

$$(6) \quad y - Zb^* = \sum_{s \notin S} v_s v_s' Vx = \sum_{s \notin S} v_s x_s.$$

Ковариационная матрица вектора x равна диагональной матрице

$$(7) \quad \mathbb{E}(x - \mathbb{E}x)(x - \mathbb{E}x)' = V^* \Sigma V = \Lambda,$$

элементы которой являются характеристическими корнями матрицы Σ . Если y имеет нормальное распределение, то вектор x также будет нормальным, а его компоненты независимыми, причем дисперсия t -й компоненты равна λ_t , $t = 1, \dots, T$. Матрица ковариаций оценки $b^* = C^{-1}V'^*y = C^{-1}V'^*Vx$ равна

$$(8) \quad \mathbb{E}(b^* - \beta)(b^* - \beta)' = C^{-1} \Lambda^* (C')^{-1},$$

где Λ^* — диагональная матрица с элементами $\lambda_{s_1}, \dots, \lambda_{s_p}$. Если характеристические векторы $\mathbf{v}_t, t \notin S$, известны, то для оценки $\lambda_t, t \notin S$, можно использовать значения $x_t^2 = (\mathbf{v}_t' \mathbf{y})^2$. (Если известно, что оценки наименьших квадратов являются эффективными, т. е. если $\mathbf{Z}$ имеет вид $\mathbf{Z} = \mathbf{V}^* \mathbf{C}$, то определено p -мерное линейное пространство, порожаемое $\mathbf{V}^*$.) Эти оценки пригодны для оценивания ковариационной матрицы для $\mathbf{b}^*$, поскольку $\lambda_t, t \in S$, — известные функции от $\lambda_t, t \notin S$, и матрица $\mathbf{C}$ известна. Отметим, что $\mathbf{b}^*$ и $\mathbf{y} - \mathbf{Zb}^*$ не коррелированы, так как

$$(9) \quad \mathbf{Z}(\mathbf{y} - \mathbf{Zb}^*)(\mathbf{b}^* - \beta)' = [\mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}']\mathbf{Z}\mathbf{Z}'(\mathbf{Z}'\mathbf{Z})^{-1} = \\ = (\mathbf{I} - \mathbf{V}^*\mathbf{V}^*)\mathbf{V}^*\mathbf{A}^*\mathbf{C}(\mathbf{Z}'\mathbf{Z})^{-1} = 0.$$

Изучим подробно случай, когда характеристические векторы $\mathbf{v}_1, \dots, \mathbf{v}_T$ известны. Так будет, в частности, когда

$$(10) \quad \Sigma^{-1} = \gamma_0 \mathbf{A}_0 + \dots + \gamma_q \mathbf{A}_q.$$

Здесь $\mathbf{v}_1, \dots, \mathbf{v}_T$ — характеристические векторы матриц $\mathbf{A}_0, \dots, \mathbf{A}_q$, а $\gamma_0, \dots, \gamma_q$ — неизвестные параметры. Примером может служить циклическая модель с $\mathbf{A}_0 = \mathbf{I}$ и

$$(11) \quad \mathbf{A}_j = \frac{1}{2}(\mathbf{B}^j + \mathbf{B}'^j), \quad j = 1, \dots, q,$$

где $\mathbf{B}$ — матрица, имеющая единицы непосредственно над главной диагональю, единицу в левом нижнем углу и нули на всех остальных позициях. При этом берется $q \leq [T/2]$. (Характеристические корни матрицы Σ^{-1} равны $\sum_{j=0}^q \gamma_j \cos 2\pi jt/T, t = 1, \dots, T$; компонентами вектора $\mathbf{v}_t$ являются $\cos 2\pi ts/T$ или $\sin 2\pi ts/T$.) Для удобства предположим, что $\mathbf{Z} = \mathbf{V}^*$ (т. е. $\mathbf{C} = \mathbf{I}$).

Допустим, что p характеристических корней матрицы Σ , соответствующие p столбцам матрицы $\mathbf{V}^*$, функционально не зависят от остальных характеристических корней. В циклической модели это будет при $q = [T/2]$. Тогда $\lambda_T^{-1} = \sum_{j=0}^q \gamma_j$ и

$$(12) \quad \lambda_s^{-1} = \lambda_{T-s}^{-1} = \sum_{j=0}^q \gamma_j \cos \frac{2\pi js}{T}, \quad s = 1, \dots, [(T-1)/2].$$

Если T четное и $q = T/2$, то $\lambda_q^{-1} = \sum_{j=0}^q (-1)^j \gamma_j$. Характеристиче-

ские корни, соответствующие столбцам матрицы V^* , функционально не зависят от остальных корней, если возможными столбцами V^* являются $\sqrt{1/T} (1, \dots, 1)'$, $\sqrt{1/T} (-1, 1, \dots, 1)'$ при четном T и пары линейных комбинаций $\sin 2\pi ts/T$ и $\cos 2\pi ts/T$ для некоторых значений t . Дисперсии оценок компонент вектора β (именно, v_s для $s \in S$) равны λ_s .

10.3.2. Асимптотическое смещение выборочных ковариаций

Рассмотрим теперь ситуацию, когда $y_t = \beta' z_t + u_t$, где $\{u_t\}$ — стационарный в широком смысле случайный процесс с нулевым средним, ковариационной последовательностью $\{\sigma(h)\}$ и спектральной плотностью $f(\lambda)$. Выборочная ковариация порядка h , полученная по выборке длины T , равна

$$(13) \quad C_h^* = \frac{1}{T-h} \sum_{t=1}^{T-h} (y_t - b^{*'} z_t) (y_{t+h} - b^{*'} z_{t+h}) = C_{-h}^*,$$

$$h = 0, 1, \dots, T-1,$$

где

$$(14) \quad b^* = (Z'Z)^{-1} Z'y,$$

$Z = (z_1, \dots, z_T)'$, $y = (y_1, \dots, y_T)'$ и матрица Z имеет ранг p . Ее можно записать иначе:

$$(15) \quad C_h^* = \frac{1}{T-h} \sum_{t=1}^{T-h} y_t y_{t+h} - \frac{1}{T-h} \sum_{t=1}^{T-h} b^{*'} z_t y_{t+h} -$$

$$- \frac{1}{T-h} \sum_{t=1}^{T-h} b^{*'} z_{t+h} y_t + \frac{1}{T-h} \sum_{t=1}^{T-h} b^{*'} z_t z_{t+h}' b^*.$$

ТЕОРЕМА 10.3.1. При условиях 10.2.1–10.2.5

$$(16) \quad \lim_{T \rightarrow \infty} T [E C_h^* - \sigma(h)] = -4\pi \operatorname{tr} R^{-1}(0) \int_{-\pi}^{\pi} \cos \lambda h f(\lambda) dM(\lambda) +$$

$$+ 2\pi \operatorname{tr} R^{-1}(0) \int_{-\pi}^{\pi} f(\lambda) dM(\lambda) R^{-1}(0) \int_{-\pi}^{\pi} e^{i\nu h} dM'(\nu).$$

Доказательство. Из (15) [в силу того, что имеют место равенства

$$y_t - b^{*'} z_t = u_t - (b^* - \beta)' z_t \text{ и } b^* - \beta = (Z'Z)^{-1} \sum_{s=1}^T z_s u_s]$$

получаем, что

$$(17) \quad T [\mathfrak{E} C_h^* - \sigma(h)] = - \frac{T}{T-h} \sum_{t=1}^{T-h} \sum_{s=1}^T z'_t (Z'Z)^{-1} z_s \mathfrak{E} u_s u_{t+h} - \\ - \frac{T}{T-h} \sum_{t=1}^{T-h} \sum_{s=1}^T z'_{t+h} (Z'Z)^{-1} z_s \mathfrak{E} u_s u_t + \\ + \frac{T}{T-h} \sum_{t=1}^{T-h} z'_{t+h} \mathfrak{E} (b^* - \beta) (b^* - \beta)' z_t.$$

Третий член в правой части (17) имеет предел

$$(18) \quad \lim_{T \rightarrow \infty} \text{tr} \mathfrak{E} (b^* - \beta) (b^* - \beta)' \sum_{t=1}^{T-h} z_t z'_{t+h} = \\ = \lim_{T \rightarrow \infty} \text{tr} \mathfrak{E} [D_T (b^* - \beta) (b^* - \beta)' D_T] D_T^{-1} \sum_{t=1}^{T-h} z_t z'_{t+h} D_T^{-1} = \\ = 2\pi \text{tr} R^{-1}(0) \int_{-\pi}^{\pi} f(\lambda) dM(\lambda) R^{-1}(0) \int_{-\pi}^{\pi} e^{i\lambda h} dM(\lambda)$$

в соответствии с (62), (64) и (94) § 10.2. Первый член в правой части (17) равен

$$(19) \quad - \frac{T}{T-h} \text{tr} \sum_{t=1}^{T-h} \sum_{s=1}^T (Z'Z)^{-1} z_s z'_t \sigma(t+h-s) = \\ = - \frac{T}{T-h} \text{tr} \sum_{r=-(T-h-1)}^{T-1} \sum_{s \in S_r} (Z'Z)^{-1} z_s z'_{s+r-h} \sigma(r),$$

где $S_r = \{h-r+1, \dots, T\}$ для $r \leq 0$, $S_r = \{h-r+1, \dots, T-r\}$ для $0 \leq r \leq h$ и $S_r = \{1, \dots, T-r\}$ для $h \leq r$. Второго члена в правой части (17) равен

$$(20) \quad - \frac{T}{T-h} \text{tr} \sum_{t=1}^{T-h} \sum_{s=1}^T (Z'Z)^{-1} z_s z'_{t+h} \sigma(t-s) = \\ = - \frac{T}{T-h} \text{tr} \sum_{r=-(T-h-1)}^{T-1} \sum_{s \in S_r} (Z'Z)^{-1} z_s z'_{s+r} \sigma(r-h).$$

Если $f(\lambda)$ является тригонометрическим полиномом [т. е. $\sigma(r) = 0$ при $|r| > H$ для некоторого H], то пределом (19) будет

$$(21) \quad - \sum_{r=-\infty}^{\infty} \sigma(r) \lim_{T \rightarrow \infty} \text{tr} A_T^{-1}(0) A_T(h-r) = \\ = - \sum_{r=-\infty}^{\infty} \sigma(r) \int_{-\pi}^{\pi} \text{tr} R^{-1}(0) e^{i\lambda(h-r)} dM(\lambda) = \\ = - 2\pi \text{tr} R^{-1}(0) \int_{-\pi}^{\pi} e^{i\lambda h} f(\lambda) dM(\lambda).$$

Подобным же образом

$$\begin{aligned}
 (22) \quad & - \sum_{r=-\infty}^{\infty} \sigma(r-h) \lim_{T \rightarrow \infty} \text{tr} \mathbf{A}_T^{-1}(0) \mathbf{A}_T(-r) = \\
 & = - \sum_{r=-\infty}^{\infty} \sigma(r-h) \int_{-\pi}^{\pi} \text{tr} \mathbf{R}^{-1}(0) e^{-i\lambda r} d\mathbf{M}(\lambda) = \\
 & = -2\pi \text{tr} \mathbf{R}^{-1}(0) \int_{-\pi}^{\pi} e^{-i\lambda h} f(\lambda) d\mathbf{M}(\lambda).
 \end{aligned}$$

Если $f(\lambda)$ непрерывна, то найдутся такие тригонометрические полиномы $f_L(\lambda)$ и $f_U(\lambda)$, что $f_L(\lambda) \leq f(\lambda) \leq f_U(\lambda)$ и $f_U(\lambda) - f_L(\lambda) \leq \varepsilon$. Указанные выше пределы сохраняются для последовательностей $\sigma_L(h) = \int_{-\pi}^{\pi} e^{i\lambda h} f_L(\lambda) d\lambda$ и $\sigma_U(h) = \int_{-\pi}^{\pi} e^{i\lambda h} f_U(\lambda) d\lambda$. Пусть

$(\mathbf{Z}'\mathbf{Z})^{-1} = \sum_{j=1}^p \alpha_j \alpha_j'$ и $x_{tj} = a\alpha_j' z_t + b\alpha_j' z_{t-h}$, $t = 1, \dots, T$, для произвольных a и b , причем $z_0 = \dots = z_{-h+1} = 0$. Тогда $x_j' \Sigma_L x_j \leq x_j' \Sigma x_j \leq x_j' \Sigma_U x_j$, как и в разд. 10.2.3. Соответствующая аргументация пригодна и для

$$\begin{aligned}
 (23) \quad & \sum_{j=1}^p \sum_{t,s=1}^T x_{tj} \sigma(t-s) x_{sj} = \\
 & = \text{tr} (\mathbf{Z}'\mathbf{Z})^{-1} \sum_{t,s=1}^T [a^2 z_t z_s' + ab(z_t z_{s-h}' + z_{t-h} z_s') + b^2 z_{t-h} z_{s-h}'] \sigma(t-s).
 \end{aligned}$$

Поскольку предел существует для любых значений a и b , то это означает, что существуют пределы и для отдельно взятых членов соотношения (23), стоящих при a^2 , ab и b^2 . Отсюда следует, что при непрерывной спектральной плотности $f(\lambda)$ пределом сумм (19) и (20) служат суммы (21) и (22). ■

Предположим теперь, что оценки наименьших квадратов асимптотически эффективны. Тогда матричная функция $\mathbf{M}(\lambda)$ имеет скачки $(\mathbf{M}_j - i\mathbf{M}_j^*)/2$ при $\lambda = \nu_j$ и $(\mathbf{M}_j + i\mathbf{M}_j^*)/2$ при $\lambda = -\nu_j$, $j = 1, \dots, H$, если при $\lambda = 0, \pm \pi$ скачков нет. (Матрицы $\mathbf{M}_j$ и $\mathbf{M}_j^*$ действительны.) Первое слагаемое в правой части (16) равно

$$(24) \quad -4\pi \text{tr} \mathbf{R}^{-1}(0) \sum_{j=1}^H \cos \nu_j h f(\nu_j) \mathbf{M}_j.$$

а второе

$$(25) \quad 2\pi \operatorname{tr} \mathbf{R}^{-1}(0) \sum_{j=1}^H f(v_j) \mathbf{M}_j \mathbf{R}^{-1}(0) \sum_{k=1}^H (\cos v_k h \mathbf{M}_k + \sin v_k h \mathbf{M}_k^*) = \\ = 2\pi \operatorname{tr} \mathbf{R}^{-1}(0) \sum_{j=1}^H \cos v_j h f(v_j) \mathbf{M}_j,$$

поскольку

$$(26) \quad \operatorname{tr} \mathbf{R}^{-1}(0) \mathbf{M}_j \mathbf{R}^{-1}(0) \mathbf{M}_i^* = 0$$

в силу того, что $\mathbf{M}_j = \mathbf{M}_j'$ и $\mathbf{M}_i^* = -\mathbf{M}_i'^*$. (Если $v_j = \pi$, то $\mathbf{M}_j^* = 0$.)

ТЕОРЕМА 10.3.2. Пусть выполнены условия 10.2.1—10.2.5. Если при этом $\mathbf{M}(\lambda)$ имеет скачки $(\mathbf{M}_j - i\mathbf{M}_j^*)/2$ при $\lambda = v_j > 0$, $j = 1, \dots, H$, то

$$(27) \quad \lim_{T \rightarrow \infty} T [\mathbb{E} C_h^* - \sigma(h)] = -2\pi \operatorname{tr} \mathbf{R}^{-1}(0) \sum_{j=1}^H \cos v_j h f(v_j) \mathbf{M}_j.$$

Если $\mathbf{M}(\lambda)$ имеет, кроме того, скачок $\mathbf{M}_0$ при $\lambda = 0$, то

$$(28) \quad \lim_{T \rightarrow \infty} T [\mathbb{E} C_h^* - \sigma(h)] = -2\pi \operatorname{tr} \mathbf{R}^{-1}(0) \sum_{j=0}^H \cos v_j h f(v_j) \mathbf{M}_j.$$

Теорема 10.3.2 является обобщением теоремы 8.3.2. Значение теорем 10.3.1 и 10.3.2 состоит в том, что смещение оценки C_h^* имеет порядок $1/T$. Следует отметить, что асимптотическое смещение можно оценить состоятельно.

Покажем теперь, что при соответствующих условиях предельное распределение величины $\sqrt{T} [C_h^* - \sigma(h)]$ совпадает с предельным распределением величины $\sqrt{T} [C_h - \sigma(h)]$ и с предельным распределением, полученным в разд. 8.4.2.

ТЕОРЕМА 10.3.3. Если выполнены условия 10.2.1—10.2.5, то

$$(29) \quad \operatorname{plim}_{T \rightarrow \infty} \sqrt{T} (C_h^* - C_h) = 0,$$

$$\text{где } C_h = \sum_{t=1}^{T-h} (y_t - \beta' z_t) (y_{t+h} - \beta' z_{t+h}) / (T-h).$$

Доказательство. Прежде всего

$$(30) \quad \sqrt{T} (C_h^* - C_h) = -\frac{\sqrt{T}}{T-h} \sum_{t=1}^{T-h} (\mathbf{b}^* - \beta)' z_t u_{t+h} - \\ - \frac{\sqrt{T}}{T-h} \sum_{t=1}^{T-h} (\mathbf{b}^* - \beta)' z_{t+h} u_t + \frac{\sqrt{T}}{T-h} \sum_{t=1}^{T-h} (\mathbf{b}^* - \beta)' z_t z_{t+h}' (\mathbf{b}^* - \beta),$$

где $u_t = y_t - \beta' z_t$. Последний член в (30) равен

$$(31) \quad \frac{\sqrt{T}}{T-h} (b^* - \beta)' A_T(h) (b^* - \beta) = \\ = \frac{\sqrt{T}}{T-h} (b^* - \beta)' D_T [D_T^{-1} A_T(h) D_T^{-1}] D_T (b^* - \beta).$$

Поскольку $\lim_{T \rightarrow \infty} D_T^{-1} A_T(h) D_T^{-1} = R(h)$ и $\mathbb{E} D_T (b^* - \beta) (b^* - \beta)' D_T$ имеет предел при $T \rightarrow \infty$, то (31) сходится по вероятности к 0. Квадрат второго члена в (30) равен

$$(32) \quad \frac{T}{(T-h)^2} (b^* - \beta)' D_T \left(D_T^{-1} \sum_{t,s=1}^{T-h} z_{t+h} u_t u_s z'_{s+h} D_T^{-1} \right) D_T (b^* - \beta).$$

Для произвольного вектора α

$$(33) \quad \mathbb{E} \left(\alpha' D_T^{-1} \sum_{t=1}^{T-h} z_{t+h} u_t \right)^2 = \mathbb{E} \alpha' D_T^{-1} \sum_{t,s=1}^{T-h} z_{t+h} u_t u_s z'_{s+h} D_T^{-1} \alpha = \\ = \alpha' D_T^{-1} \sum_{t,s=1}^{T-h} z_{t+h} z'_{s+h} \sigma(t-s) D_T^{-1} \alpha = \\ = \alpha' D_T^{-1} \sum_{r=-(T-h-1)}^{T-h-1} \sum_{s \in S_r} z_{s+h+r} z'_{s+h} \sigma(r) D_T^{-1} \alpha$$

имеет в условиях теоремы предел $\alpha' \sum_{r=-\infty}^{\infty} \sigma(r) R(r) \alpha$. Поэтому и второй член в правой части (30) сходится по вероятности к 0. Что касается первого члена, то он оценивается аналогичным образом. ■

ТЕОРЕМА 10.3.4. Если выполнены условия 10.2.1—10.2.5 и условия теоремы 8.4.2, то вектор $\sqrt{T} [C_0^* - \sigma(0)], \dots, \sqrt{T} [C_n^* - \sigma(n)]$ имеет предельное нормальное распределение с нулевыми средними и ковариациями, приведенными в теореме 8.4.2.

При надлежащих предположениях относительно моментов четвертого порядка можно показать, что пределы ковариаций $\sqrt{T} [C_h^* - \sigma(h)]$ совпадают с пределами ковариаций $\sqrt{T} [C_h - \sigma(h)]$ и без обязательного выполнения всех условий, необходимых для асимптотической нормальности.

10.3.3. Асимптотическое смещение выборочной спектральной плотности

Рассмотрим выборочную спектральную плотность, строящуюся по остаткам от подобранной по методу наименьших квадратов регрессии, т. е.

$$(34) \quad I^*(\lambda) = \frac{1}{2\pi T} \left| \sum_{t=1}^T (y_t - \mathbf{b}^{*'} \mathbf{z}_t) e^{i\lambda t} \right|^2.$$

ТЕОРЕМА 10.3.5. Пусть выполнены условия 10.2.1—10.2.5 и, кроме того, предел

$$(35) \quad \lim_{T \rightarrow \infty} \frac{1}{\sqrt{T}} \mathbf{D}_T^{-1} \sum_{t=1}^T \mathbf{z}_t e^{i\lambda t} = \mathbf{g}(\lambda)$$

существует для всех λ и $\sum_{h=-\infty}^{\infty} |\sigma(h)| < \infty$. Тогда

$$(36) \quad \lim_{T \rightarrow \infty} \mathbb{E} I^*(\lambda) = f(\lambda) - 2f(\lambda) \mathbf{g}'(\lambda) \mathbf{R}^{-1}(0) \overline{\mathbf{g}(\lambda)} + \\ + \mathbf{g}'(\lambda) \mathbf{R}^{-1}(0) \int_{-\pi}^{\pi} f(\nu) d\mathbf{M}(\nu) \mathbf{R}^{-1}(0) \overline{\mathbf{g}(\lambda)}.$$

ДОКАЗАТЕЛЬСТВО. Преобразуем (34) следующим образом:

$$(37) \quad I^*(\lambda) = \frac{1}{2\pi} \sum_{h=-(T-1)}^{T-1} \left(1 - \frac{|h|}{T} \right) C_h^* \cos \lambda h = \\ = \frac{1}{2\pi T} \sum_{s,t=1}^T \hat{u}_s \hat{u}_t e^{i\lambda(t-s)} = \frac{1}{2\pi T} \left| \sum_{t=1}^T \hat{u}_t e^{i\lambda t} \right|^2,$$

где вектор $\hat{\mathbf{u}} = (\hat{u}_1, \dots, \hat{u}_T)'$ имеет вид

$$(38) \quad \hat{\mathbf{u}} = [\mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'] \mathbf{u}.$$

Положим $\boldsymbol{\theta} = (e^{i\lambda}, e^{i\lambda 2}, \dots, e^{i\lambda T})'$. Тогда

$$(39) \quad 2\pi I^*(\lambda) = \frac{1}{T} |\hat{\mathbf{u}}' \boldsymbol{\theta}|^2 = \frac{1}{T} \boldsymbol{\theta}' \hat{\mathbf{u}} \hat{\mathbf{u}}' \bar{\boldsymbol{\theta}} = \\ = \frac{1}{T} \boldsymbol{\theta}' [\mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'] \mathbf{u} \mathbf{u}' [\mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'] \bar{\boldsymbol{\theta}} = \\ = \frac{1}{T} \boldsymbol{\theta}' \mathbf{u} \mathbf{u}' \bar{\boldsymbol{\theta}} - \frac{1}{T} [\boldsymbol{\theta}' \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}' \mathbf{u} \mathbf{u}' \bar{\boldsymbol{\theta}} + \\ + \boldsymbol{\theta}' \mathbf{u} \mathbf{u}' \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}' \bar{\boldsymbol{\theta}}] + \\ + \frac{1}{T} \boldsymbol{\theta}' \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}' \mathbf{u} \mathbf{u}' \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}' \bar{\boldsymbol{\theta}}.$$

Математическое ожидание первого члена правой части (39) равно

$$(40) \quad 2\pi f(\lambda) = \sum_{r=-(T-1)}^{T-1} \left(1 - \frac{|r|}{T}\right) \sigma(r) \cos \lambda r$$

и стремится при $T \rightarrow \infty$ к $2\pi f(\lambda)$. Математическое ожидание последнего члена равно

$$\begin{aligned} (41) \quad \frac{1}{T} \theta' Z (Z'Z)^{-1} Z' \Sigma Z (Z'Z)^{-1} Z' \bar{\theta} &= \\ &= \frac{1}{T} \theta' Z \mathbf{g} (b^* - \beta) (b^* - \beta)' Z' \bar{\theta} = \\ &= \frac{1}{T} \theta' Z D_T^{-1} [D_T \mathbf{g} (b^* - \beta) (b^* - \beta)' D_T] D_T^{-1} Z' \bar{\theta} = \\ &= \text{tr} [D_T \mathbf{g} (b^* - \beta) (b^* - \beta)' D_T] \frac{1}{T} D_T^{-1} Z' \bar{\theta} \theta' Z D_T^{-1}. \end{aligned}$$

Предел матрицы, стоящей в квадратных скобках, указан в следствии 10.2.1. Другая матрица в (41) равна $\overline{g_T(\lambda)} g'(\lambda)$, где

$$(42) \quad g_T(\lambda) = \frac{1}{\sqrt{T}} D_T^{-1} \sum_{t=1}^T z_t e^{i\lambda t},$$

сходится к $g(\lambda)$. Таким образом, предел математического ожидания последнего члена правой части (39) равен

$$(43) \quad 2\pi g'(\lambda) R^{-1}(0) \int_{-\pi}^{\pi} f(v) dM(v) R^{-1}(0) \overline{g(\lambda)}.$$

Математическое ожидание среднего члена равно

$$\begin{aligned} (44) \quad -\frac{1}{T} [\theta' Z (Z'Z)^{-1} Z' \Sigma \bar{\theta} + \theta' \Sigma Z (Z'Z)^{-1} Z' \bar{\theta}] &= \\ &= -\frac{1}{T} \sum_{t,s,r=1}^T z'_t [A_T(0)]^{-1} z_s \sigma(s-r) [e^{i\lambda(t-r)} + e^{i\lambda(r-t)}] = \\ &= -\sum_{h=-(T-1)}^{T-1} \sigma(h) \sum_{r \in S_h} \sum_{t=1}^T z'_t D_T^{-1} [D_T^{-1} A_T(0) D_T^{-1}]^{-1} D_T^{-1} z_{r+h} \times \\ &\quad \times \frac{1}{T} [e^{i\lambda(t-r)} + e^{i\lambda(r-t)}], \end{aligned}$$

где $S_h = \{1, \dots, T-h\}$, $h \geq 0$, и $S_h = \{1-h, \dots, T\}$, $h \leq 0$. Оно сходится к

$$\begin{aligned} (45) \quad -2 \sum_{h=-\infty}^{\infty} \sigma(h) \cos \lambda h g'(\lambda) R^{-1}(0) \overline{g(\lambda)} &= \\ &= -4\pi f(\lambda) g'(\lambda) R^{-1}(0) \overline{g(\lambda)}. \end{aligned}$$

Этим завершаем доказательство. ■

Отметим, что

$$(46) \quad \int_{-\pi}^{\pi} g_T(\lambda) \overline{g_T(\lambda)} d\lambda = \frac{1}{T} \int_{-\pi}^{\pi} D_T^{-1} \sum_{t,s=1}^T z_t z_s' D_T^{-1} e^{i\lambda(t-s)} d\lambda = \\ = \frac{2\pi}{T} D_T^{-1} A_T(0) D_T^{-1}.$$

Интеграл квадрата абсолютной величины каждой компоненты вектора $g_T(\lambda)$ равен $2\pi/T$ и стремится к 0 при $T \rightarrow \infty$. Поэтому $g(\lambda) = 0$ для почти всех λ . Если $\{z_t\}$ ведет себя, грубо говоря, как реализация стационарного случайного процесса, то $g(\lambda) = 0$ для всех λ . При этом каждая компонента $g(\lambda)$ соответствует $A(\lambda) + iB(\lambda)$, а каждый диагональный член матрицы $g(\lambda) \overline{g'(\lambda)}$ соответствует $R^2(\lambda)$. В этом случае $I^*(\lambda)$ оказывается асимптотически несмещенной оценкой.

Следствие 10.3.1. Если $g(\lambda) = 0$, то $\lim_{T \rightarrow \infty} \mathbb{E} I^*(\lambda) = f(\lambda)$.

Рассмотрим теперь такую последовательность независимых переменных $\{z_t\}$, что оценки наименьших квадратов оказываются при этом асимптотически эффективными для каждой непрерывной положительной спектральной плотности. Из теоремы 10.2.10 следует тогда, что

$$(47) \quad \lim_{T \rightarrow \infty} T \int_{-\pi}^{\pi} g_T(\lambda) \overline{g_T'(\lambda)} f(\lambda) d\lambda = \sum_{j=1}^G f(\nu_j) M_j$$

для каждой непрерывной положительной спектральной плотности $f(\lambda)$. Существует матрица P (не обязательно единственная), такая, что $P'R(0)P = I$ и матрица $P' \sum_{j=1}^G f(\nu_j) M_j P$ диагональна для каждого набора значений $f(\nu_j)$, $j = 1, \dots, G$. При этом матрицы $P'M_j P = L_j$ будут диагональными, причем их диагональные элементы равны 1 или 0 и $\sum_{j=1}^G L_j = I$. Пусть $g_T^*(\lambda) = P'g_T(\lambda)$. (См. упр. 32 гл. 6.) Тогда

$$(48) \quad \lim_{T \rightarrow \infty} T \int_{-\pi}^{\pi} g_T^*(\lambda) \overline{g_T^{*'}(\lambda)} f(\lambda) d\lambda = \sum_{j=1}^G f(\nu_j) L_j,$$

где

$$(49) \quad L_j = \begin{bmatrix} 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

$\mathbf{g}$ и $\mathbf{I}$ есть j -я единичная диагональная подматрица. Далее, естественно предположить, что

$$(50) \quad \lim_{T \rightarrow \infty} \mathbf{g}_T^*(\nu_j) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{g}_{(j)}^* \\ 0 \\ \vdots \\ 0 \end{bmatrix}; \quad j = 1, \dots, G,$$

$$(51) \quad \lim_{T \rightarrow \infty} \mathbf{g}_T^*(\lambda) = 0, \quad \lambda \neq \pm \nu_1, \dots, \pm \nu_G.$$

При этих условиях $\lim_{T \rightarrow \infty} \mathcal{E} I^*(\lambda) = f(\lambda)$, $\lambda \neq \pm \nu_1, \dots, \pm \nu_G$.

Тогда

$$(52) \quad \mathbf{g}'(\nu_j) \mathbf{R}^{-1}(0) \sum_{k=1}^G f(\nu_k) \mathbf{M}_k \mathbf{R}^{-1}(0) \overline{\mathbf{g}(\nu_j)} = \\ = f(\nu_j) \mathbf{g}_{(j)} \overline{\mathbf{g}_{(j)}} = f(\nu_j) \mathbf{g}'(\nu_j) \mathbf{R}^{-1}(0) \overline{\mathbf{g}(\nu_j)}.$$

Итак, получена

ТЕОРЕМА 10.3.6. Если выполнены условия 10.2.1 — 10.2.5, (35) существует для каждого λ , $\sum_{h=-\infty}^{\infty} |\sigma(h)| < \infty$, кроме того, для $G \leq \rho$

$$(53) \quad \mathbf{R}(h) = \sum_{j=1}^G (\cos \nu_j h \mathbf{M}_j + \sin \nu_j h \mathbf{M}_j^*),$$

$\mathbf{g}(\lambda) = 0$, $\lambda \neq \pm \nu_1, \dots, \pm \nu_G$, и если $\mathbf{M}_k \mathbf{R}^{-1}(0) \mathbf{g}(\nu_j) = 0$, $k \neq j$, $k, j = 1, \dots, G$, то

$$(54) \quad \lim_{T \rightarrow \infty} \mathcal{E} I^*(\lambda) = f(\lambda), \quad \lambda \neq \pm \nu_1, \dots, \pm \nu_G,$$

$$(55) \quad \lim_{T \rightarrow \infty} \mathcal{E} I^*(\pm \nu_j) = f(\nu_j) [1 - \mathbf{g}'(\nu_j) \mathbf{R}^{-1}(0) \overline{\mathbf{g}(\nu_j)}],$$

$$j = 1, \dots, G.$$

Выборочная спектральная плотность $I^*(\lambda)$ может оказаться асимптотически смещенной для тех значений λ , при которых $\mathbf{M}(\lambda)$ имеет скачки, т. е. для значений, соответствующих частотам периодичности последовательности $\{z_t\}$. Если множитель $1 - \mathbf{g}'(\nu_j) \times \times \mathbf{R}^{-1}(0) \overline{\mathbf{g}(\nu_j)}$ в (55) известен и отличен от нуля, то оценку значения спектральной плотности $f(\nu_j)$ можно получить делением $I^*(\nu_j)$ на этот множитель.

Рассмотрим несколько примеров. Пусть $z_{1t} \equiv 1$ (т. е. $v_1 = 0$),

$$(56) \quad z_{2k-2,t} = \cos v_k t, \quad k = 2, \dots, G-1,$$

$$(57) \quad z_{2k-1,t} = \sin v_k t, \quad k = 2, \dots, G-1,$$

и $z_{2G-2,t} = (-1)^t$, $t = 1, 2, \dots$. Тогда

$$(58) \quad R(h) = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \cos v_1 h & -\sin v_1 h & 0 & \dots & 0 \\ 0 & \sin v_1 h & \cos v_1 h & 0 & \dots & 0 \\ 0 & 0 & 0 & \cos v_2 h & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & (-1)^h \end{bmatrix}.$$

Используя результат упр. 13, можно получить, что $g(\lambda) = 0$ для $\lambda \neq \pm v_1, \dots, \pm v_G$. Первой компонентой вектора $g(0)$ является 1; $g(v_k)$ имеют $(2k-2)$ -е компоненты, равные $1/\sqrt{2}$, и $(2k-1)$ -е компоненты, равные $i/\sqrt{2}$, $k = 2, \dots, G-1$. Последняя компонента вектора $g(\pi)$ равна 1. Остальные компоненты $g(v_j)$ равны нулю, $j = 1, \dots, G$. Тогда $\lim_{T \rightarrow \infty} \mathcal{I}^*(v_j) = 0$, $j = 1, \dots, G$.

Если в z_t опустить $\cos v_k t$ или $\sin v_k t$, то $\lim_{T \rightarrow \infty} \mathcal{I}^*(v_k) = f(v_k)/2$.

В качестве другого примера рассмотрим $z_t = t$. В этом случае

$$(59) \quad g_T(\lambda) = \frac{\sum_{t=1}^T t e^{i\lambda t}}{\sqrt{T} \sqrt{T(T+1)(2T+1)/6}} =$$

$$= -\frac{1 - (T+1)e^{i\lambda T} + Te^{i\lambda(T+1)}}{4T \sin^2(\lambda/2) \sqrt{T(T+1)(2T+1)/6}}, \quad \lambda \neq 0,$$

$$(60) \quad g_T(0) = \frac{T+1}{2\sqrt{T(T+1)(2T+1)/6}}.$$

При этом $g(\lambda) = 0$, $\lambda \neq 0$, и $g(0) = \sqrt{3}/2$. Отсюда $\lim_{T \rightarrow \infty} \mathcal{I}^*(\lambda) = f(\lambda)$, $\lambda \neq 0$, и $\lim_{T \rightarrow \infty} \mathcal{I}^*(0) = f(0)/4$.

10.3.4. Асимптотическое смещение оценок спектральной плотности

Рассмотрим теперь

$$(61) \quad \hat{f}_T^*(v) = \frac{1}{2\pi} \sum_{r=-K_T}^{K_T} k\left(\frac{r}{K_T}\right) \cos v r c_{rT},$$

где

$$(62) \quad \dot{c}_{rT} = \frac{1}{T} \sum_{i=1}^{T-r} (y_i - \mathbf{b}^{*'} \mathbf{z}_i) (y_{i+r} - \mathbf{b}^{*'} \mathbf{z}_{i+r}) = \dot{c}_{-rT},$$

$$r = 0, 1, \dots, T-1,$$

$k(x)$ — симметричная функция и $\{K_T\}$ — возрастающая последовательность целых чисел типа исследовавшихся в гл. 9. Тогда

$$(63) \quad \hat{f}_T^*(v) - f(v) = \hat{f}_T(v) - f(v) + \hat{f}_T^*(v) - \hat{f}_T(v),$$

где

$$(64) \quad \hat{f}_T(v) = \frac{1}{2\pi} \sum_{r=-K_T}^{K_T} k\left(\frac{r}{K_T}\right) \cos vr c_{rT},$$

$$(65) \quad c_{rT} = \frac{1}{T} \sum_{i=1}^{T-r} (y_i - \beta' \mathbf{z}_i) (y_{i+r} - \beta' \mathbf{z}_{i+r}) =$$

$$= \frac{1}{T} \sum_{i=1}^{T-r} u_i u_{i+r} = c_{-rT}, \quad r = 0, 1, \dots, T-1.$$

Слагаемое $\hat{f}_T(v) - f(v)$ в (63) было изучено в гл. 9. Рассмотрим второе слагаемое. Имеем

$$(66) \quad \hat{f}_T^*(v) - \hat{f}_T(v) = \frac{1}{2\pi} \sum_{r=-K_T}^{K_T} k\left(\frac{r}{K_T}\right) \cos vr (\dot{c}_{rT} - c_{rT}),$$

где для $r \geq 0$

$$(67) \quad \dot{c}_{rT} - c_{rT} = -\frac{1}{T} (\mathbf{b}^* - \beta)' \sum_{i=1}^{T-r} \mathbf{z}_i u_{i+r} - \frac{1}{T} (\mathbf{b}^* - \beta)' \sum_{i=1}^{T-r} \mathbf{z}_{i+r} u_i +$$

$$+ \frac{1}{T} (\mathbf{b}^* - \beta)' \sum_{i=1}^{T-r} \mathbf{z}_i \mathbf{z}_{i+r}' (\mathbf{b}^* - \beta).$$

Умноженное на T последнее слагаемое в (67) оценивается по абсолютной величине следующим образом:

$$(68) \quad \left| \sum_{i=1}^{T-r} (\mathbf{b}^* - \beta)' \mathbf{z}_i \mathbf{z}_{i+r}' (\mathbf{b}^* - \beta) \right| \leq$$

$$< \left(\sum_{i=1}^{T-r} [(\mathbf{b}^* - \beta)' \mathbf{z}_i]^2 \sum_{i=1}^{T-r} [(\mathbf{b}^* - \beta)' \mathbf{z}_{i+r}]^2 \right)^{1/2} <$$

$$< \sum_{i=1}^T [(\mathbf{b}^* - \beta)' \mathbf{z}_i]^2 =$$

$$= (\mathbf{b}^* - \beta)' \mathbf{A}_T(0) (\mathbf{b}^* - \beta) =$$

$$= \sum_{i,s=1}^T \mathbf{z}_i' \mathbf{A}_T^{-1}(0) \mathbf{z}_s u_i u_s.$$

Здесь мы использовали неравенство Коши—Шварца. Пусть $A_T^{-1}(0) = P'P$ и $z_s^* = Pz_s$, $s = 1, \dots, T$. Тогда математическое ожидание правой части (68) равно

$$(69) \quad \mathbb{E} \sum_{t,s=1}^T z_t' A_T^{-1}(0) z_s u_t u_s = \sum_{t,s=1}^T z_t^* z_s^* \sigma(t-s) = \\ = \sum_{q=-(T-1)}^{T-1} \sum_{t \in S_q} z_t^* z_{t+q}^* \sigma(q),$$

где $S_q = \{1, \dots, T - q\}$ для $q \geq 0$ и $S_q = \{1 - q, \dots, T\}$ для $q \leq 0$. Для любого $h \geq 0$

$$(70) \quad \left| \sum_{t=1}^{T-h} z_t^* z_{t+h}^* \right| \leq \sum_{t=1}^p \left| \sum_{t=1}^{T-h} z_{it}^* z_{i,t+h}^* \right| \leq \\ \leq \sum_{i=1}^p \sqrt{\sum_{t=1}^{T-h} z_{it}^2} \sqrt{\sum_{t=1}^{T-h} z_{i,t+h}^2} \leq \\ \leq \sum_{i=1}^p \sum_{t=1}^T z_{it}^2 = p.$$

Умноженное на T математическое ожидание абсолютной величины последнего члена в (67) ограничено поэтому значением $p \sum_{q=-\infty}^{\infty} |\sigma(q)|$, которое по предположению конечно.

Умноженная на T абсолютная величина первого члена в правой части (67) имеет математическое ожидание

$$(71) \quad \mathbb{E} \left| (b^* - \beta)' \sum_{t=1}^{T-r} z_t u_{t+r} \right| = \mathbb{E} \left| \sum_{s=1}^T \sum_{t=1}^{T-r} z_s' A_T^{-1}(0) z_t u_s u_{t+r} \right| = \\ = \mathbb{E} \left| \sum_{s=1}^T \sum_{t=1}^{T-r} z_s^* z_t^* u_s u_{t+r} \right| = \\ = \mathbb{E} \left| \left(\sum_{s=1}^T u_s z_s^* \right)' \left(\sum_{t=1}^{T-r} u_{t+r} z_t^* \right) \right| \leq \\ \leq \left[\mathbb{E} \left(\sum_{s=1}^T u_s z_s^* \right)' \left(\sum_{s'=1}^T u_{s'} z_{s'}^* \right) \right]^{1/2} \times \\ \times \left[\mathbb{E} \left(\sum_{t=1}^{T-r} u_{t+r} z_t^* \right)' \left(\sum_{t'=1}^{T-r} u_{t'+r} z_{t'}^* \right) \right]^{1/2},$$

поскольку $\mathbb{E} |x'y| \leq \mathbb{E} \sum_{i=1}^p |x_i| |y_i| \leq \sqrt{\mathbb{E} \sum_{i=1}^p x_i^2} \sqrt{\mathbb{E} \sum_{i=1}^p y_i^2}$, а квадрат (71) не превосходит

$$\begin{aligned} (72) \quad & \mathbb{E} \sum_{s,s'=1}^T z_s^* z_{s'}^* u_s u_{s'} \mathbb{E} \sum_{t,t'=1}^{T-r} z_t^* z_{t'}^* u_{t+r} u_{t'+r} = \\ & = \sum_{s,s'=1}^T z_s^* A_T^{-1}(0) z_{s'} \sigma(s'-s) \sum_{t,t'=1}^{T-r} z_t^* A_T^{-1}(0) z_{t'} \sigma(t'-t) \leq \\ & \leq \left[p \sum_{q=-\infty}^{\infty} |\sigma(q)| \right]^2. \end{aligned}$$

Таким образом, математические ожидания умноженных на T абсолютных величин первого и второго членов правой части (67) также ограничены значением $p \sum_{q=-\infty}^{\infty} |\sigma(q)|$.

ТЕОРЕМА 10.3.7.

$$(73) \quad |\mathbb{E}(c_{rT}^* - c_{rT})| \leq \mathbb{E}|c_{rT}^* - c_{rT}| \leq \frac{3p}{T} \sum_{q=-\infty}^{\infty} |\sigma(q)|.$$

Из (66) имеем

$$(74) \quad \mathbb{E} |\hat{f}_T^*(v) - \hat{f}_T(v)| \leq \frac{1}{2\pi} \sum_{r=-K_T}^{K_T} \left| k\left(\frac{r}{K_T}\right) \right| \mathbb{E} |c_{rT}^* - c_{rT}|.$$

ТЕОРЕМА 10.3.8.

$$\begin{aligned} (75) \quad & |\mathbb{E} \hat{f}_T^*(v) - \mathbb{E} \hat{f}_T(v)| \leq \mathbb{E} |\hat{f}_T^*(v) - \hat{f}_T(v)| \leq \\ & \leq \frac{3p}{2\pi T} (2K_T + 1) \sup_{-1 \leq x \leq 1} |k(x)| \sum_{q=-\infty}^{\infty} |\sigma(q)|. \end{aligned}$$

ТЕОРЕМА 10.3.9.

$$(76) \quad \lim_{T \rightarrow \infty} \frac{T}{K_T} |\mathbb{E} \hat{f}_T^*(v) - \mathbb{E} \hat{f}_T(v)| \leq \frac{3p}{\pi} \sup_{-1 \leq x \leq 1} |k(x)| \sum_{q=-\infty}^{\infty} |\sigma(q)|.$$

ТЕОРЕМА 10.3.10. Предположим, что $k(x) = k(-x)$, $k(0) = 1$, $|k(x)| \leq M$ для некоторого M и для всех $|x| \leq 1$. Предположим, что для некоторых $q > 0$ и $k > 0$ $\lim_{x \rightarrow 0} [1 - k(x)]/|x|^q = k$. Пусть $\{K_T\}$ — последовательность целых чисел, такая, что $K_T \rightarrow \infty$ при $T \rightarrow \infty$. Пусть, кроме того,

$$(77) \quad \sum_{r=-\infty}^{\infty} |r|^p |\sigma(r)| < \infty$$

для некоторого $p \geq q$. Тогда, если $K_T^{q+1}/T \rightarrow 0$ при $T \rightarrow \infty$, то

$$(78) \quad \lim_{T \rightarrow \infty} \frac{\hat{f}_T^*(v) - \hat{f}_T(v)}{\hat{f}_T(v) - f(v)} = 0.$$

Доказательство. Справедливость теоремы вытекает из теорем 10.3.9 и 9.3.3. ■

Следствием последней теоремы является то, что если K_T растет достаточно медленно по отношению к T , а именно если $K_T^{q+1}/T \rightarrow 0$ при $T \rightarrow \infty$, то добавочное смещение оценки из-за наличия неизвестной регрессии оказывается несущественным по сравнению со смещением, возникающим при известном математическом ожидании. Если $q = 2$ и $p \geq 2$, то указанное условие будет иметь вид $K_T^3/T \rightarrow 0$; если $K_T = [\gamma T^\alpha]$, то при этом получаем условие $\alpha < 1/3$.

Мы хотим показать, что асимптотическое распределение разности $\hat{f}_T^*(v) - f(v)$ совпадает с асимптотическим распределением $\hat{f}_T(v) - f(v)$. Теорема 10.3.8 и обобщенное неравенство Чебышёва приводят к следующему результату.

ТЕОРЕМА 10.3.11. Пусть функция $g(T)$ такова, что $g(T) \rightarrow 0$ при $T \rightarrow \infty$. Если $\sum_{r=-\infty}^{\infty} |\sigma(r)| < \infty$ и функция $k(x)$ ограничена на $[-1, 1]$, то

$$(79) \quad \text{plim}_{T \rightarrow \infty} g(T) \frac{T}{K_T} [\hat{f}_T^*(v) - \hat{f}_T(v)] = 0.$$

СЛЕДСТВИЕ 10.3.2. Если $\sum_{r=-\infty}^{\infty} |\sigma(r)| < \infty$ и функция $k(x)$ ограничена на $[-1, 1]$, то

$$(80) \quad \text{plim}_{T \rightarrow \infty} \sqrt{\frac{T}{K_T}} [\hat{f}_T^*(v) - \hat{f}_T(v)] = 0.$$

Из следствия 9.4.1 и следствия 10.3.2 выводим следующий результат.

ТЕОРЕМА 10.3.12. Пусть $\hat{f}_T^*(v)$ определяется соотношениями (61) и (62), в которых функция $k(x)$ непрерывна на $[-1, 1]$ и $K_T/T \rightarrow 0$ при $T \rightarrow \infty$. Пусть $y_t = \beta' z_t + u_t$, где

$$(81) \quad u_t = \sum_{s=-\infty}^{\infty} \gamma_s v_{t-s},$$

$\sum_{s=-\infty}^{\infty} |\gamma_s| < \infty$, а $\{v_t\}$ — последовательность независимых и одинаково распределенных случайных величин с $\mathbb{E} v_t = 0$, $\mathbb{E} v_t^2 = \sigma^2$ и $\mathbb{E} v_t^4 < \infty$.

$< \infty$. Если выполнены условия теоремы 9.4.3 и если $\lim_{T \rightarrow \infty} T/K_T^{2p+1}$ конечен при $p < q$ или $T/K_T^{2q+1} \rightarrow 0$ при $q \leq p$, то $\sqrt{T/K_T} [\hat{f}_T^*(v) - f(v)]$ имеет в пределе нормальное распределение с нулевым средним и дисперсией, указанной в теореме 9.3.4.

Асимптотическое распределение оценки $f(v)$ не изменяется при использовании остатков от регрессии, подобранной по методу наименьших квадратов. Теорема 10.3.11 показывает, что для получения в пределе нетривиальной случайной величины разность $\hat{f}_T^*(v) - \hat{f}_T(v)$ следовало бы домножать на T/K_T , в то время как разность $\hat{f}_T(v) - f(v)$ для той же цели следует домножать на $\sqrt{T/K_T}$. Таким образом, разность $\hat{f}_T^*(v) - \hat{f}_T(v)$ обычно имеет больший порядок малости по сравнению с $\hat{f}_T(v) - f(v)$.

Большая часть результатов § 10.3 содержится в работе Хеннана (1958).

10.4. ПРОВЕРКА НЕЗАВИСИМОСТИ

10.4.1. Случай, когда оценки наименьших квадратов являются эффективными

В гл. 6 мы рассматривали задачу проверки нулевой гипотезы о том, что ковариационная матрица совокупности T наблюдений пропорциональна единичной матрице, против частных альтернатив наличия сериальной корреляции при условии, что математические ожидания наблюдаемых величин или все равнялись нулю, или все были равны некоторой отличной от нуля постоянной величине. В настоящем параграфе мы займемся этой задачей в предположении, что последовательность математических ожиданий наблюдаемых величин образует линейную функцию регрессии. Подобным же образом в этой ситуации могут быть исследованы задачи проверки порядка зависимости и оценки порядка зависимости.

Будем предполагать, что плотность распределения вектора $y = (y_1, \dots, y_T)'$ равна

$$(1) \quad K \exp [-(\gamma_0 Q_0 + \dots + \gamma_q Q_q)/2],$$

где

$$(2) \quad Q_j = (y - Z\beta)' A_j (y - Z\beta), \quad j = 0, 1, \dots, q,$$

Z — матрица размера $T \times p$ ранга p ($p < T - q$), а β — вектор размерности p . Тогда вектор y нормально распределен с вектором

средних

$$(3) \quad \mathbb{E}y = Z\beta$$

и ковариационной матрицей

$$(4) \quad \Sigma = (\gamma_0 A_0 + \dots + \gamma_q A_q)^{-1}.$$

Константа K равна

$$(5) \quad K = (2\pi)^{-T/2} |\gamma_0 A_0 + \dots + \gamma_q A_q|^{1/2}.$$

Наиболее интересна проверка нулевой гипотезы $\gamma_1 = 0$ для случая, когда $q = 1$ и $A_0 = I$, т. е. проверка гипотезы независимости против альтернативы, состоящей в том, что ковариационная матрица имеет вид $\Sigma = (\gamma_0 I + \gamma_1 A_1)^{-1}$. При этом предполагается, что A_1 и Z — известные матрицы, а γ_0 , γ_1 и β неизвестны. Обычно альтернативной гипотезой (характеризующейся матрицей A_1) является наличие сериальной корреляции.

Прежде всего рассмотрим случай $Z = V^*C$, где матрица C невырождена, а V^* состоит из p характеристических векторов v_s , $s = s_1, \dots, s_p$, матрицы Σ (причем $V^*V^* = I$). Функция регрессии равна здесь $Z\beta = V^*C\beta = V^*\alpha$, где $\alpha = C\beta$ — некоторое преобразование параметров.

Как было показано в разд. 6.6.2, в этом случае статистики

$$(6) \quad a_h = v_{s_h}' y, \quad h = 1, \dots, p,$$

$$(7) \quad Q_j^* = \left(y - \sum_{h=1}^p a_h v_{s_h} \right)' A_j \left(y - \sum_{h=1}^p a_h v_{s_h} \right) = \\ = y' A_j y - \sum_{h=1}^p \lambda_{js_h} a_h^2, \quad j = 0, 1, \dots, q,$$

(где λ_{js_h} — характеристический корень матрицы A_j , соответствующий характеристическому вектору v_{s_h}) образуют достаточное множество статистик для параметров β_h , $h = 1, \dots, p$, и γ_j , $j = 0, 1, \dots, q$. Здесь a_h — оценка наименьших квадратов и марковская оценка для α_h , $h = 1, \dots, p$. Если $\gamma_{i+1} = \dots = \gamma_q = 0$, то наилучший критерий для проверки нулевой гипотезы $\gamma_i = 0$ основывается на статистике Q_i^* и область принятия гипотезы зависит от условного распределения Q_i^* при заданных значениях $Q_0^*, \dots, Q_{i-1}^*$. В частности, критерий для проверки гипотезы $\gamma_1 = 0$ при $\gamma_2 = \dots = \gamma_q = 0$ основывается на статистике

$$(8) \quad r_1^* = \frac{Q_1^*}{Q_0^*}.$$

При $\gamma_1 = 0$ распределение r_1^* не зависит ни от β (теорема 6.7.1), ни от γ_0 (теорема 6.7.2).

ТЕОРЕМА 10.4.1. Если вектор y имеет плотность (1) с $q = 1$, а Q_0 и Q_1 определены соотношением (2), причем матрица $\gamma_0 A_0 + \gamma_1 A_1$ положительно определена, и если матрица Z состоит из r линейно независимых линейных комбинаций r характеристических векторов матриц A_0 и A_1 , то равномерно наиболее мощный подобный критерий для проверки нулевой гипотезы $\gamma_1 = 0$ против альтернатив $\gamma_1 < 0$ с уровнем значимости ε имеет критическую область $r_1^* > c_1^*$, где $r_1^* = Q_1^*/Q_0^*$,

$$(9) \quad Q_j^* = (y - Zb^*)' A_j (y - Zb^*), \quad j = 0, 1,$$

$b^* = (Z'Z)^{-1}Z'y$, а c_1^* определяется так, чтобы при нулевой гипотезе $\Pr \{r_1^* > c_1^*\} = \varepsilon$. Равномерно наиболее мощный подобный критерий против альтернатив $\gamma_1 > 0$ имеет критическую область $r_1^* < c_1^{**}$, где c_1^{**} определяется так, чтобы $\Pr \{r_1^* < c_1^{**}\} = \varepsilon$ при $\gamma_1 = 0$. Равномерно наиболее мощный несмещенный критерий против альтернатив $\gamma_1 \neq 0$ имеет критическую область, являющуюся объединением областей $r_1^* < c_{L1}^*$ и $r_1^* > c_{U1}^*$, где c_{L1}^* и c_{U1}^* определяются так, чтобы $\Pr \{c_{L1}^* \leq r_1^* \leq c_{U1}^*\} = 1 - \varepsilon$

$$(10) \quad \int_{c_{L1}^*}^{c_{U1}^*} r_1^* f(r_1^*) dr_1^* = (1 - \varepsilon) \int_{-\infty}^{\infty} r_1^* f(r_1^*) dr_1^*,$$

где $f(r_1^*)$ — плотность распределения r_1^* при $\gamma_1 = 0$.

Наиболее интересны случаи, когда $A_0 = I$. Пусть матрица характеристических векторов A_1 равна V , причем $V'V = I$. Если положить $y = Vx$ ($x = V'y$), то вектор x будет иметь распределение $N \left[\sum_{h=1}^p \alpha_h v_{sh}, (\gamma_0 I + \gamma_1 \Lambda)^{-1} \right]$, где Λ — диагональная матрица с диагональными элементами $\lambda_1 \geq \dots \geq \lambda_T$, являющимися характеристическими корнями матрицы A_1 . (Все характеристические корни матрицы $A_0 = I$ равны единице.) Соответствующие квадратичные формы равны

$$(11) \quad Q_0^* = x'x - \sum_{s \in S} x_s^2 = \sum_{s \notin S} x_s^2,$$

$$(12) \quad Q_1^* = x' \Lambda x - \sum_{s \in S} \lambda_s x_s^2 = \sum_{s \notin S} \lambda_s x_s^2,$$

где $S = \{s_1, \dots, s_p\}$. Тогда

$$(13) \quad r_1^* = \frac{\sum_{s \in S} \lambda_s x_s^2}{\sum_{s \in S} x_s^2}.$$

Если $\gamma_1 = 0$, то распределение величины r_1^* не зависит от γ_0 . При $\gamma_1 = 0$ распределение величины r_1^* совпадает с распределением отношения (13), в котором x_s — независимые нормальные величины с нулевыми средними и единичными дисперсиями.

Распределения общего типа, рассмотренные в § 6.7, соответствуют рассматриваемым здесь с заменой $\lambda_1, \dots, \lambda_T$ или $\lambda_1, \dots, \lambda_{T-1}$ (когда включается математическое ожидание) на $\lambda_s, s \in S$. В частности, если $T - p$ корней являются двойными, скажем если это корни $v_1 > \dots > v_H$, то $\text{Pr}\{r_1^* \geq R\}$ указывается в теореме 6.7.4. Если двойными являются все корни ($v_1 > \dots > v_H$) за исключением одного простого корня $v_{H+1} (< v_H)$, то в этом случае $\text{Pr}\{r_1^* \geq R\}$ дается теоремой 6.7.6.

Если $\gamma_1 = 0$, то

$$(14) \quad \mathfrak{E} r_1^{*h} = \frac{\mathfrak{E} Q_1^{*h}}{\mathfrak{E} Q_0^{*h}}.$$

Если $\gamma_0 = 1$, то знаменатель в (14) равен

$$(15) \quad \mathfrak{E} Q_0^{*h} = 2^h \frac{\Gamma\left[\frac{T-p}{2} + h\right]}{\Gamma\left[\frac{T-p}{2}\right]}, \quad -(T-p)/2 < h.$$

Числитель в (14) можно найти, воспользовавшись семинвариантами квадратичной формы Q_1^* . Действительно, k -й семинвариант представляет собой умноженную на $k!$ величину

$$(16) \quad v_k^* = \frac{2^{k-1}}{k} \sum_{i \in S} \lambda_i^k = \frac{2^{k-1}}{k} \sum_{i=1}^T \lambda_i^k - \frac{2^{k-1}}{k} \sum_{h=1}^p \lambda_{s_h}^k.$$

Тогда несколько первых моментов выражаются соотношениями (126) и (127) § 6.7, в которых v_k заменяются на v_k^* . Если $\lambda_1, \dots, \lambda_T$ и $\lambda_{s_1}, \dots, \lambda_{s_p}$ симметричны относительно нуля, то $v_{2l-1}^* = 0$, $l = 1, 2, \dots$, $\mathfrak{E} Q_1^{*2} = 2v_0^*$ и $\mathfrak{E} Q_1^{*4} = 24v_4^* + 12v_2^{*2}$. Могут быть также использованы и приближенные распределения, описанные в § 6.8.

Особый интерес представляет циклический случай с $T = ph$ и независимыми переменными $\cos 2\pi jt/T$ и $\sin 2\pi jt/T$. Если при этом

p четное, то в качестве независимых переменных берут

$$(17) \quad 1, \cos \frac{2\pi}{p} t, \sin \frac{2\pi}{p} t, \cos \frac{4\pi}{p} t, \sin \frac{4\pi}{p} t, \dots, \\ \cos \frac{2\pi \left(\frac{p}{2} - 1\right)}{p} t, \sin \frac{2\pi \left(\frac{p}{2} - 1\right)}{p} t, (-1)^t, \\ t = 1, \dots, T.$$

Соответствующие характеристические корни равны

$$(18) \quad 1, \cos \frac{2\pi}{p}, \cos \frac{2\pi}{p}, \cos \frac{4\pi}{p}, \cos \frac{4\pi}{p}, \dots, \\ \cos \frac{2\pi \left(\frac{p}{2} - 1\right)}{p}, \cos \frac{2\pi \left(\frac{p}{2} - 1\right)}{p}, -1.$$

Тогда $\dot{v}_{2l-1} = 0$, $l = 1, 2, \dots$, и

$$(19) \quad \dot{v}_{2l} = \frac{2^{2l-1}}{2l} \sum_{t=1}^T \cos^{2l} \frac{2\pi}{T} t - \frac{2^{2l-1}}{2l} \sum_{j=1}^p \cos^{2l} \frac{2\pi}{p} j = \\ = \frac{T}{4l} \sum'_{j_T} \binom{2l}{j_T} - \frac{p}{4l} \sum'_{j_p} \binom{2l}{j_p},$$

где символы Σ' указывают на то, что входящие в правую часть (19) суммы берутся соответственно по всем значениям j_T , $j_T = 0, 1, \dots, 2l$, для которых $2(j_T - l)$ делится на T , и по всем значениям j_p , $j_p = 0, 1, \dots, 2l$, для которых $2(j_p - l)$ делится на T . (См. разд. 6.7.7.) При этом

$$(20) \quad \dot{v}_{2l} = \frac{T-p}{4l} \binom{2l}{l} + \frac{T}{2l} \sum_{g=1}^{[2l/T]} \binom{2l}{l - gT/2} - \frac{p}{2l} \sum_{g=1}^{[2l/p]} \binom{2l}{l - gp/2},$$

для $T = ph$. Если $l < p/2$, то (20) переходит в

$$(21) \quad \dot{v}_{2l} = \frac{T-p}{4l} \binom{2l}{l}, \quad l < p/2, \quad p = 2, 4, \dots$$

(Отметим, что $T \geq 2p$.) В частности,

$$(22) \quad \dot{v}_2 = (T-p)/2, \quad p = 4, 6, \dots,$$

$$(23) \quad \dot{v}_4 = 3(T-p)/4, \quad p = 6, 8, \dots$$

Можно также показать, что $\dot{v}_2 = T/2 - 2$ для $p = 2$, $\dot{v}_4 = 3T/4 - 4$ для $p = 2$ и 4 , $T \geq 6$. Тогда $\mathfrak{E}Q_1^2 = T - 4$ для $p = 2$ и $T - p$ для $p = 4, 6, \dots$, $\mathfrak{E}Q_1^4 = 3(T^2 - 2T - 16)$ для $p = 2, 4$, $T \geq 6$.

Таблица 10.2

ВЕРХНИЕ 100%-ПРОЦЕНТНЫЕ ТОЧКИ РАСПРЕДЕЛЕНИЯ ЦИКЛИЧЕСКОГО СЕРИАЛЬНОГО
КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ, СТРОЯЩЕГОСЯ ПО ОСТАТКАМ ОТ ОЦЕНКИ
ТРИГОНОМЕТРИЧЕСКОГО ТРЕНДА *

Номера жорней	$P = 2$			$P = 2, 4$			$P = 2, 3, 6$			$P = 2, 12/5, 3, 4, 6,$ 12		
	T	5%	1%	T	5%	1%	T	5%	1%	T	5%	1%
4	6	0.450	0.490	8	0.636	0.693						
6	8	0.484	0.607				12	0.592	0.744			
8	10	0.453	0.601	12	0.515	0.661						
10	12	0.426	0.572									
12	14	0.402	0.544	16	0.439	0.582	18	0.442	0.592	24	0.441	0.592
14	16	0.382	0.519									
16	18	0.364	0.496	20	0.388	0.523						
18	20	0.348	0.476				24	0.369	0.504			
20	22	0.334	0.458	24	0.351	0.478						
22	24	0.321	0.442									
24	26	0.310	0.427	28	0.323	0.441	30	0.323	0.445	36	0.323	0.445
26	28	0.300	0.414									
28	30	0.290	0.402	32	0.300	0.414						
30	32	0.282	0.390				36	0.291	0.403			
32	34	0.274	0.380	36	0.282	0.391						
34	36	0.266	0.370									
36	38	0.260	0.361	40	0.267	0.371	42	0.267	0.371	48	0.267	0.371
38	40	0.254	0.353									
40	42	0.248	0.345	44	0.254	0.354						
42	44	0.242	0.338				48	0.248	0.346			
44	46	0.237	0.331	48	0.243	0.338						
46	48	0.233	0.324									
48	50	0.228	0.318	52	0.233	0.325	54	0.233	0.325	60	0.233	0.325
50	52	0.224	0.313									
52	54	0.220	0.307	56	0.224	0.313						
54	56	0.216	0.302				60	0.220	0.308			
56	58	0.212	0.297	60	0.216	0.302						
58	60	0.209	0.292									

* Значения P указывают на периоды включенных в оценку тригонометрических составляющих.

и $\mathcal{E}Q_1^{*4} = 3[T^2 - 2(p-3)T + p(p-6)]$ для $p = 6, 8, \dots$.
Поэтому $\mathcal{E}r_{2l-1}^* = 0$, $\mathcal{E}r_1^{*2} = (T-4)/[T(T-2)]$ для $p = 2$,

$$(24) \quad \mathcal{E}r_1^{*2} = (T-p+2)^{-1}, \quad p = 4, 6, \dots,$$

$$\begin{aligned} \mathcal{E}r_1^{*4} &= 3(T^2 - 2T - 16)/[(T-2)T(T+2)(T+4)] \text{ для } p=2 \text{ и } T \geq 6, \\ \mathcal{E}r_1^{*4} &= 3(T^2 - 2T - 16)/[(T-4)(T-2)T(T+2)] \text{ для } p=4, \text{ и} \\ (25) \quad \mathcal{E}r_1^{*4} &= 3[(T-p+2)(T-p+4)]^{-1}, \quad p=6, 8, \dots \end{aligned}$$

Аппроксимируем плотность r_1^* плотностью

$$(26) \quad \frac{\Gamma(2m)}{\Gamma^2(m)} \frac{(b^2 - x^2)^{m-1}}{(2b)^{2m-1}}, \quad -b \leq x \leq b,$$

второй момент которой равен $b^2/(2m+1)$. Возьмем $b = \cos 2\pi/T$ и приравняем $b^2/(2m+1)$ второму моменту $\mathcal{E}r_1^{*2}$ коэффициента r_1^* . Тогда для $p=4, 6, \dots$ получим $m = [b^2(T-p+2) - 1]/2$. Четвертый момент плотности (26) равен $3b^4/[(2m+1)(2m+3)]$. При указанной подгонке получаем для него выражение

$$(27) \quad 3[(T-p+2)(T-p+2) + 2/b^2]^{-1}.$$

Сравнивая это с (25) для $p=6, 8, \dots$, видим, что при значениях b , близких к 1, разность между (25) и (27) невелика. В работе Р. Андерсона и Т. Андерсона (1950) приведена таблица значений R , при которых $\text{Pr}\{r_1^* \geq R\} = 0.05$ и 0.01 . Эта таблица включена в настоящую книгу как табл. 10.2¹⁾. Для выравнивания полугодовых данных используются константа и функция $(-1)^t$, имеющая период $P=2$. Для поквартальных данных используются константа, функция $(-1)^t$ с периодом $P=2$, а также $\cos(\pi t/2)$ и $\sin(\pi t/2)$ с периодом $P=4$. Если рассматриваются данные за каждые два месяца, то используются константа, функция $(-1)^t$ с периодом $P=2$, функции $\cos 2\pi t/3$ и $\sin 2\pi t/3$ с периодом $P=3$, $\cos \pi t/3$ и $\sin \pi t/3$ с периодом $P=6$. Выравнивание ежемесячных данных проводится при помощи константы, функции $(-1)^t$ с периодом $P=2$ и $\cos 2\pi jt/12$ и $\sin 2\pi jt/12$ с периодами $P=12/j$, $j=1, 2, 3, 4, 5$.

Если посмотреть на таблицу, то можно отметить близость процентных точек для всех случаев, соответствующих одной и той же степени свободы, т. е. одному и тому же числу остающихся корней. Так, распределение величины $r_1^*/(\cos 2\pi/T)$ для $P=2$ весьма близко к распределению $r_1^*/(\cos \pi/T)$, основанному на сумме квадратов последовательных разностей. (См. табл. 6.3.)

10.4.2. Общий случай

Рассмотрим теперь указанную задачу в ситуации, когда факт совпадения столбцов матрицы $\mathbf{Z}$ с характеристическими векторами матрицы $\mathbf{A}_1$ не учитывается. Пусть остатки от подобранной по методу

¹⁾ У Р. Андерсона и Т. Андерсона (1950) значения для $T=6$ и $P=2$ указаны неверно.

наименьших квадратов регрессии равны

$$(28) \quad \hat{u} = y - Zb^* = y - Z(Z'Z)^{-1}Z'y = [I - Z(Z'Z)^{-1}Z']y.$$

Тогда сериальный коэффициент корреляции равен

$$(29) \quad r_1^* = \frac{\hat{u}'A_1\hat{u}}{\hat{u}'\hat{u}} = \\ = \frac{y'[I - Z(Z'Z)^{-1}Z']A_1[I - Z(Z'Z)^{-1}Z']y}{y'[I - Z(Z'Z)^{-1}Z'][I - Z(Z'Z)^{-1}Z']y} = \\ = \frac{y'B'A_1By}{y'B'By},$$

где

$$(30) \quad B = I - Z(Z'Z)^{-1}Z'.$$

При этом матрица B размера $T \times T$ имеет ранг $T - p$ как матрица, ортогональная Z . Матрица B идемпотентна, поскольку

$$(31) \quad B^2 = [I - Z(Z'Z)^{-1}Z'] [I - Z(Z'Z)^{-1}Z'] = \\ = I - 2Z(Z'Z)^{-1}Z' + Z(Z'Z)^{-1}Z'Z(Z'Z)^{-1}Z' = \\ = I - 2Z(Z'Z)^{-1}Z' + Z(Z'Z)^{-1}Z' = \\ = I - Z(Z'Z)^{-1}Z' = B.$$

Характеристические корни идемпотентной матрицы равны 1 и 0. Матрица B имеет характеристический корень 1 кратности $T - p$, равной ее рангу, и характеристический корень 0 кратности p . Поскольку $B'B = B^2$, то $B'B$ имеет те же самые характеристические корни. Таким образом, существует ортогональная матрица Q , которая приводит $B'B$ к диагональному виду, т. е.

$$(32) \quad Q'B'BQ = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix},$$

где матрица I в (32) имеет порядок $T - p$. Если $BQ = (CG)$, где C имеет размер $T \times (T - p)$, то (32) равно

$$(33) \quad Q'B'BQ = \begin{pmatrix} C' \\ G' \end{pmatrix} (C \ G) = \begin{pmatrix} C'C & C'G \\ G'C & G'G \end{pmatrix}.$$

Поскольку $G'G = 0$, то $G = 0$ и $BQ = (C0)$. Пусть $y = Qu$ и вектор u представлен в виде $u = (u^{*'} u^{**'})'$, где u^* состоит из $T - p$

компонент. Тогда

$$(34) \quad r_1^* = \frac{u' Q' B' A_1 B Q u}{u' Q' B' B Q u} = \\ = \frac{(u'' u^{**}) \begin{pmatrix} C' \\ 0 \end{pmatrix} A_1 (C \ 0) \begin{pmatrix} u^* \\ u^{**} \end{pmatrix}}{(u'' u^{**}) \begin{pmatrix} C' \\ 0 \end{pmatrix} (C \ 0) \begin{pmatrix} u^* \\ u^{**} \end{pmatrix}} = \frac{u'' C' A_1 C u^*}{u'' u^*}.$$

При нулевой гипотезе $\gamma_1 = 0$ ковариационная матрица вектора u равна $\sigma^2 I$ и такова же ковариационная матрица вектора u . Поскольку $u = Qu$ и $Bu = BQu = (C0)u = Cu^*$, то $C' Bu = C' Cu^* = u^*$. Поэтому $\xi u^* = C' B \xi u = C' B Z \beta = 0$.

При нулевой гипотезе вектор u^* имеет распределение $N(0, \sigma^2 I)$. Пусть характеристическими корнями матрицы $C' A_1 C$ являются $v_1 \geq \dots \geq v_{T-p}$. Поскольку в матрице $Q' B' A_1 B Q$ в левом верхнем углу расположена матрица $C' A_1 C$, а остальные элементы — нули, указанные корни являются также и характеристическими корнями матрицы $B' A_1 B$. Более того, они являются характеристическими корнями матрицы $B B' A_1 = B^2 A_1 = B A_1$.

ТЕОРЕМА 10.4.2. При $\gamma_1 = 0$ распределение величины r_1^* совпадает с распределением отношения

$$(35) \quad \frac{\sum_{i=1}^{T-p} v_i x_i^2}{\sum_{i=1}^{T-p} x_i^2},$$

в котором $v_1 \geq \dots \geq v_{T-p}$ представляют собой $T-p$ характеристических корней матрицы $[I - Z(Z'Z)^{-1}Z']A_1$, составляющих все отличные от нуля характеристические корни этой матрицы, а случайные величины $x_1, \dots, x_{T-p}$ независимы и нормально распределены с нулевыми средними и единичными дисперсиями.

Задача, таким образом, состоит в том, чтобы найти характеристические корни матрицы $[I - Z(Z'Z)^{-1}Z']A_1 = A_1 - Z(Z'Z)^{-1}Z'A_1$. При этом можно использовать теорию § 6.7 и 6.8. Семиинвариант k -го порядка числителя (35) равен $k!$, умноженному на

$$(36) \quad \frac{2^{k-1}}{k} \sum_{i=1}^{T-p} v_i^k = \frac{2^{k-1}}{k} \text{tr} \{ [I - Z(Z'Z)^{-1}Z'] A_1 \}^k = \\ = \frac{2^{k-1}}{k} \text{tr} \{ A_1 - Z(Z'Z)^{-1}Z'A_1 \}^k.$$

При разложении правой части (36) следует учитывать возможность того, что матрицы A_1 и $Z(Z'Z)^{-1}Z'$ могут оказаться неперестановоч-

ными. См. упр. 21 и 22. Дурбин и Ватсон (1950) не обратили внимания на эту возможную потерю коммутативности.

Моменты величины r_1^* могут быть вычислены через ее семиинварианты. По первым четырем моментам можно подобрать бета-распределение, аппроксимирующее истинное распределение величины r_1^* .

Во многих случаях вычисление характеристических корней матрицы $\mathbf{BA}_1$ является довольно утомительной процедурой. В связи с этим Дурбин и Ватсон (1950), (1951) получили неравенства для этих корней и как следствие неравенства для точного распределения r_1^* . Мы покажем сейчас, что $T - p$ характеристических корней матрицы $\mathbf{BA}_1$, не относящихся к тем p векторам, для которых $\mathbf{Bx} = 0$, будут соответственно не больше чем $T - p$ наибольших характеристических корней матрицы $\mathbf{A}_1$ и не меньше чем $T - p$ наименьших характеристических корней этой матрицы. Мы получим эти неравенства с помощью следующих двух лемм. Пусть $\text{ch}_i(\mathbf{A})$ обозначает i -й по величине характеристический корень матрицы $\mathbf{A}$, т. е. $\text{ch}_1(\mathbf{A}) \geq \dots \geq \text{ch}_T(\mathbf{A})$.

Лемма 10.4.1. Для любой симметрической матрицы $\mathbf{A}$ и произвольных векторов $\alpha_1, \dots, \alpha_{t-1}$

$$(37) \quad \text{ch}_t(\mathbf{A}) \leq \max_{\substack{\mathbf{x}'\alpha_j=0 \\ j=1, \dots, t-1}} \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{x}}.$$

Доказательство. Если матрица $\mathbf{A}$ диагональна, то $\mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{i=1}^T \lambda_i x_i^2$, где $\lambda_1 \geq \dots \geq \lambda_T$ — характеристические корни $\mathbf{A}$. Поэтому

$$\begin{aligned} (38) \quad \max_{\substack{\mathbf{x}'\alpha_j=0 \\ j=1, \dots, t-1}} \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{x}} &\geq \max_{\substack{\mathbf{x}'\alpha_j=0 \\ j=1, \dots, t-1 \\ x_{t+1}=\dots=x_T=0}} \left(\sum_{i=1}^T \lambda_i x_i^2 \right) / \left(\sum_{i=1}^T x_i^2 \right) = \\ &= \max_{\substack{\mathbf{x}'\alpha_j=0 \\ j=1, \dots, t-1 \\ x_{t+1}=\dots=x_T=0}} \left(\sum_{i=1}^t \lambda_i x_i^2 \right) / \left(\sum_{i=1}^t x_i^2 \right) \geq \\ &\geq \lambda_t \left(\sum_{i=1}^t x_i^2 \right) / \left(\sum_{i=1}^t x_i^2 \right) = \text{ch}_t(\mathbf{A}). \end{aligned}$$

Поскольку всякую симметрическую матрицу $\mathbf{A}$ можно привести к диагональному виду некоторым ортогональным преобразованием, причем и характеристические корни, и $\mathbf{x}'\mathbf{x}$ остаются без изменений, то отсюда следует, что утверждение леммы сохраняет силу для любой симметрической матрицы. (См. Курант и Гильберт (1937).)

Аналогичные аргументы приводят к следующей лемме.

ЛЕММА 10.4.2. Для всякой симметрической матрицы A и произвольных векторов $\alpha_{i+1}, \dots, \alpha_T$

$$(39) \quad \text{ch}_i(A) \geq \min_{\substack{x' \alpha_j = 0 \\ j=i+1, \dots, T}} \frac{x'Ax}{x'x}.$$

ТЕОРЕМА 10.4.3. Если $\lambda_1 \geq \dots \geq \lambda_T$ — характеристические векторы симметрической матрицы A_1 , $B = I - Z(Z'Z)^{-1}Z'$, где матрица Z размера $T \times p$ имеет ранг $p < T$, и $v_1 \geq \dots \geq v_{T-p}$ — характеристические корни матрицы $B'A_1B$, отличные от нуля, то

$$(40) \quad \lambda_{i+p} \leq v_i \leq \lambda_i, \quad i = 1, \dots, T-p.$$

ДОКАЗАТЕЛЬСТВО. Те p корней матрицы $B'A_1B$, которые опущены в формулировке теоремы, соответствуют p векторам, аннулируемым матрицей B . Положим $H = Z(Z'Z)^{-1}Z'$. Тогда H является положительно полуопределенной матрицей ранга p . Поскольку $B + H = I$, то выполнение соотношения $Hx = 0$ влечет за собой и выполнение соотношения $x = Bx$. Пусть $v_1, \dots, v_T$ — совокупность (линейно независимых) характеристических векторов матрицы A_1 . Тогда

$$\begin{aligned} (41) \quad \text{ch}_i(A_1) &= \max_{\substack{x'v_j=0 \\ j=1, \dots, i-1}} \frac{x'A_1x}{x'x} \geq \max_{\substack{x'v_j=0 \\ j=1, \dots, i-1 \\ x'H=0}} \frac{x'A_1x}{x'x} = \\ &= \max_{\substack{x'v_j=0 \\ j=1, \dots, i-1 \\ x'H=0}} \frac{x'B'A_1Bx}{x'x} \geq v_i. \end{aligned}$$

Аналогично

$$\begin{aligned} (42) \quad \text{ch}_{i+p}(A_1) &= \min_{\substack{x'v_j=0 \\ j=i+p+1, \dots, T}} \frac{x'A_1x}{x'x} \leq \\ &\leq \min_{\substack{x'v_j=0 \\ j=i+p+1, \dots, T \\ x'H=0}} \frac{x'A_1x}{x'x} = \\ &= \min_{\substack{x'v_j=0 \\ j=i+p+1, \dots, T \\ x'H=0}} \frac{x'B'A_1Bx}{x'x} \leq v_i. \end{aligned}$$

СЛЕДСТВИЕ 10.4.1. В условиях теоремы 10.4.3

$$(43) \quad \sum_{i=1}^{T-p} \lambda_{i+p} x_i^2 \leq \sum_{i=1}^{T-p} v_i x_i^2 \leq \sum_{i=1}^{T-p} \lambda_i x_i^2.$$

ТЕОРЕМА 10.4.4. В условиях теоремы 10.4.3

$$(44) \quad \Pr \left\{ \frac{\sum_{t=1}^{T-p} \lambda_{t+p} w_t^2}{\sum_{t=1}^{T-p} w_t^2} \geq R \right\} \leq \Pr \left\{ \frac{\sum_{t=1}^{T-p} v_t w_t^2}{\sum_{t=1}^{T-p} w_t^2} \geq R \right\} \leq \\ \leq \Pr \left\{ \frac{\sum_{t=1}^{T-p} \lambda_t w_t^2}{\sum_{t=1}^{T-p} w_t^2} \geq R \right\}.$$

Пусть случайные величины $w_1, \dots, w_{T-p}$ независимы и нормально распределены с нулевыми средними и дисперсиями 1. Если при этом независимые переменные соответствуют наибольшим характеристическим корням матрицы A_1 , то левая часть (44) равна 1 минус функция распределения соответствующей сериальной корреляции. Если же независимые переменные соответствуют наименьшим по величине характеристическим корням матрицы A_1 , то тогда той же величине равна правая часть соотношения (44). Было бы хорошо, если бы мы знали, что некоторые из независимых переменных являются характеристическими векторами матрицы A_1 .

ТЕОРЕМА 10.4.5. Предположим, что p^* линейно независимых линейных комбинаций столбцов матрицы Z являются p^* характеристическими векторами матрицы A_1 . Пусть $\lambda'_1 \geq \dots \geq \lambda'_{T-p^*}$ — характеристические корни A_1 , соответствующие остальным ее характеристическим векторам. Тогда если вектор y имеет распределение $N(Z\beta, \sigma^2 I)$, то

$$(45) \quad \Pr \left\{ \frac{\sum_{t=1}^{T-p} \lambda'_{t+p-p^*} w_t^2}{\sum_{t=1}^{T-p} w_t^2} \geq R \right\} \leq \Pr \{r_1^* \geq R\} \leq \\ \leq \Pr \left\{ \frac{\sum_{t=1}^{T-p} \lambda'_t w_t^2}{\sum_{t=1}^{T-p} w_t^2} \geq R \right\},$$

где r_1^* определяется соотношением (29), а $w_1, \dots, w_{T-p}$ — независимые нормально распределенные случайные величины с нулевыми средними и единичными дисперсиями.

Для данной матрицы A_1 и определенных p^* характеристических векторов в Z можно затабулировать те значения R , для которых прав-

вая часть (45) равна ε . Обозначим эти R через R_U . Тогда если наблюдаемое значение r_1^* превосходит R_U , то это превышение следует считать значимым для уровня ε . Пусть R_L — такое R , при котором левая часть (45) равна ε . Тогда, если наблюдаемое r_1^* окажется меньше R_L , это различие оказывается незначимым для уровня ε . Если же наблюдаемое значение r_1^* заключено между R_L и R_U , то определенного заключения вынести нельзя. Эта процедура является критерием проверки независимости против альтернативы о положительной зависимости ($\gamma_1 = 0$ против $\gamma_1 < 0$).

Дурбин и Ватсон (1951) затабулировали величины $d_L = 2(1 - R_U)$ и $d_U = 2(1 - R_L)$ в случае использования суммы квадратов последовательных разностей для уровней $\varepsilon = 0.05, 0.025, 0.01$ и $p = 2, 3, 4, 5, 6$. При этом предполагалось, что одним из столбцов матрицы Z является $\varepsilon = (1, \dots, 1)'$. Если для $p = 2$ статистики, стоящие в левой и правой частях неравенства (45), переписать соответственно как деленные на $\sum_{i=1}^{T-2} w_i^2$ суммы $\sum_{i=2}^{T-2} \lambda'_i w_i^2 + \lambda'_1 w_1^2$ и $\sum_{i=2}^{T-2} \lambda'_i w_i^2 + \lambda'_{T-1} w_1^2$ (где символы $w_1, \dots, w_{T-2}$ используются для обозначения $T-2$ независимых стандартных нормальных величин), то можно непосредственно заметить, что разность между этими статистиками равна

$$(46) \quad \frac{w_1^2 (\lambda'_1 - \lambda'_{T-1})}{\sum_{i=1}^{T-2} w_i^2} = 2 \cos \frac{\pi}{T} \frac{w_1^2}{\sum_{i=1}^{T-2} w_i^2},$$

что приближенно соответствует величине $2 \cos(\pi/T)/(T-2)$. Используя таблицы Дурбина и Ватсона, можно показать, что разность между R_U и R_L для $p = 2$ близка к этой величине. Упрощенное эмпирическое правило состоит здесь в том, что R_U приближенно равно R из табл. 6.3 (для данного T) плюс $\sum_{j=1}^{p-1} \cos(\pi_j/T)/(T-p)$, а R_L приближенно равно R минус последняя сумма. В табл. 10.3 приведены процентные точки для $p = 2$.

Особого интереса заслуживает случай полиномиальной регрессии. Предположим, что A_1 соответствует использованию суммы квадратов последовательных разностей (см. (41) из § 6.5) и $\tau = (1, 2, \dots, T)'$. Тогда линейный тренд имеет вид $\beta_1 \varepsilon + \beta_2 \tau$. Вектор τ близок к характеристическому вектору матрицы A_1 , соответствующему характеристическому корню 1, именно

$$(47) \quad A_1 \tau = \tau + \frac{1}{2} (1, 0, \dots, 0, -1)'$$

(См. упр. 24.) Это наводит на мысль о том, что процентные точки остатков от регрессии $\beta_1 \varepsilon + \beta_2 \tau$ должны быть близки к R_L для $p =$

Таблица 10.3

НИЖНЯЯ И ВЕРХНЯЯ ГРАНИЦЫ ДЛЯ ВЕРХНИХ 100%-ПРОЦЕНТНЫХ ТОЧЕК РАСПРЕДЕЛЕНИЙ СЕРИАЛЬНОГО КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ, СТРОЯЩЕГОСЯ ПО ПОСЛЕДОВАТЕЛЬНОМУ РАЗНОСТЯМ ОСТАТКОВ ОТ ТРЕНДА, СОСТОЯЩЕГО ИЗ КОНСТАНТЫ И НЕЗАВИСИМОЙ ПЕРЕМЕННОЙ

T	Уровни значимости ε					
	5%		2.5%		1%	
	R_L	R_U	R_L	R_U	R_L	R_U
15	0.32	0.46	0.385	0.525	0.465	0.595
16	0.315	0.45	0.38	0.51	0.455	0.58
17	0.31	0.435	0.375	0.495	0.45	0.465
18	0.305	0.42	0.37	0.485	0.44	0.55
19	0.30	0.41	0.36	0.47	0.435	0.535
20	0.295	0.40	0.36	0.46	0.425	0.525
21	0.29	0.39	0.35	0.45	0.42	0.515
22	0.285	0.38	0.345	0.44	0.415	0.50
23	0.28	0.37	0.34	0.43	0.405	0.49
24	0.275	0.365	0.335	0.42	0.40	0.48
25	0.275	0.355	0.33	0.41	0.395	0.475

= 2. (Другой подход состоит в том, что вектор, t -я компонента которого равна $(T + 1 - 2t)/(T - 1)$, близок к вектору с t -й компонентой $\cos \pi (t - 1/2)/T$.) Возьмем матрицу A_1 из разд. 6.5.4. Тогда

$$(48) \quad A_1 e = e - \frac{1}{2} (1, 0, \dots, 0, 1)'$$

Таким образом, вектор e мало отличается от характеристического вектора матрицы A_1 , соответствующего характеристическому корню 1. Это заставляет предположить, что распределение отношения $y'B'A_1 y / y'B'y$ близко к распределению отношения $\sum_{i=2}^T \lambda_i w_i^2 / \sum_{i=2}^T w_i^2$. В самом деле, $T - 1$ наименьших характеристических корней матрицы A_1 и $T - 1$ характеристических корней матрицы $B'A_1 B$ равны соответственно $1/2$ и $-1/2$ для $T = 2$; $(0, -\sqrt{2}/2 = -0.7071)$ и $(0, -2/3 = -0.6667)$ для $T = 3$; $(\cos 2\pi/5 = 0.3090, \cos 3\pi/5 = -0.3090, \cos 4\pi/5 = -0.8090)$ и $[(1/\sqrt{5} - 1)/4 = 0.3090, -0.2500, -(1/\sqrt{5} + 1)/4 = -0.8090]$ для $T = 4$.

Для проверки независимости в условиях, когда оценки наименьших квадратов не являются эффективными, Дурбин (1970) предложил другую процедуру. Предположим, что $Z = (V_1' Z_2)'$, где p_1 столбцов матрицы V_1' являются p_1 характеристическими векторами матрицы A_1 , $V_1' V_1 = I$, $V_1' Z_2 = 0$ и что никакая линейная комбина-

ция $p_2 = p - p_1$ столбцов матрицы V_1^* не является линейной комбинацией остальных $T - p_1$ характеристических векторов матрицы A_1 . Предположим еще, что матрица V_2^* , состоящая из p_2 характеристических векторов матрицы A_1 , «близка» к матрице Z_2 , $V_2^{**} V_2^* = I$ и $V_1^{**} V_2^* = 0$. Пусть

$$(49) \quad \begin{pmatrix} b^{(1)} \\ b^{(2)} \\ c \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ 0 & Z_2' Z_2 & Z_2' V_2^* \\ 0 & V_2^{**} Z_2 & I \end{pmatrix}^{-1} \begin{pmatrix} V_1^{**} y \\ Z_2' y \\ V_2^{**} y \end{pmatrix} =$$

$$= \begin{pmatrix} V_1^{**} y \\ (C \ D) \begin{pmatrix} Z_2' y \\ V_2^{**} y \end{pmatrix} \\ (D' \ E) \begin{pmatrix} Z_2' y \\ V_2^{**} y \end{pmatrix} \end{pmatrix},$$

где

$$(50) \quad \begin{pmatrix} C & D \\ D' & E \end{pmatrix} = \begin{pmatrix} Z_2' Z_1 & Z_2' V_2^* \\ V_2^{**} Z_2 & I \end{pmatrix}^{-1}.$$

Пусть P и Q — матрицы размера $p_2 \times p_2$, удовлетворяющие соотношениям $PP' = C$ и $QQ' = E$. Пусть

$$(51) \quad v = y - V_1^* b^{(1)} - Z_2 b^{(2)} - V_2^* c,$$

$$(52) \quad w = (Z_2 - V_2^* V_2^{**} Z_2) P Q^{-1} c.$$

Поскольку

$$(53) \quad \mathfrak{E} \begin{pmatrix} b^{(1)} \\ b^{(2)} \\ c \end{pmatrix} = \begin{pmatrix} \beta^{(1)} \\ \beta^{(2)} \\ 0 \end{pmatrix},$$

то $\mathfrak{E}v = \mathfrak{E}w = 0$, и поскольку вектор регрессии c не коррелирован с остатками, то $\mathfrak{E}vw' = 0$. Ковариационная матрица вектора v равна (в предположении, что вектор y имеет ковариационную матрицу I)

$$(54) \quad \mathfrak{E}vv' = I - (V_1^* Z V_2^*) \begin{pmatrix} I & 0 & 0 \\ 0 & Z_2' Z_2 & Z_2' V_2^* \\ 0 & V_2^{**} Z_2 & I \end{pmatrix}^{-1} \begin{pmatrix} V_1^{**} \\ Z_2' \\ V_2^{**} \end{pmatrix} =$$

$$= I - V_1^* V_1^{**} - (Z_2 V_2^*) \begin{pmatrix} C & D \\ D' & E \end{pmatrix} \begin{pmatrix} Z_2' \\ V_2^{**} \end{pmatrix}.$$

Ковариационная матрица вектора w равна

$$(55) \quad \mathcal{E}ww' = (Z_2 - V_2^*V_2^{**'}Z_2)PQ^{-1}QQ'(Q')^{-1}P'(Z_2' - Z_2'V_2^*V_2^{**'}) = \\ = (Z_2 - V_2^*V_2^{**'}Z_2)C(Z_2' - Z_2'V_2^*V_2^{**'}).$$

Поскольку $Z_2'V_2^* = -C^{-1}D$ и $E - D'C^{-1}D = I$, то

$$(56) \quad \mathcal{E}(v + w)(v + w)' = \mathcal{E}vv' + \mathcal{E}ww' = I - V_1^*V_1^{**'} - V_2^*V_2^{**'},$$

что совпадает с ковариационной матрицей остатков от регрессии на $V^* = (V_1^* V_2^*)$. Распределение отношения

$$(57) \quad \frac{(v + w)' A_1 (v + w)}{(v + w)' (v + w)}$$

при нулевой гипотезе совпадает с распределением отношения

$$(58) \quad \frac{x'(I - V^*V^{**'}) A_1 (I - V^*V^{**'}) x}{x'(I - V^*V^{**'}) (I - V^*V^{**'}) x} = \frac{x'A_1x - x'V^*\Lambda^*V^{**'}x}{x'x - x'V^*V^{**'}x},$$

в котором x имеет распределение $N(0, I)$. Последнее же отношение является сериальной корреляцией, основывающейся на остатках от V^* . При этом Λ^* — диагональная матрица, диагональные элементы которой равны характеристическим корням, соответствующим столбцам матрицы V^* .

Можно использовать также и сериальный коэффициент корреляции, основывающийся на последовательных разностях. Константу можно включить в регрессию, полагая $p_1 = 1$ и $V_1^* = \sqrt{1/T} \varepsilon = \sqrt{1/T} (1, \dots, 1)'$. Если другие независимые переменные, составляющие Z_2 , изменяются медленно, то их совокупность может оказаться близкой к матрице V_2^* , состоящей из p_2 векторов-столбцов с компонентами $\cos \pi s(t - 1/2)/T$, $t = 1, \dots, T$, $s = 1, \dots, p_2$. При этом процентной точкой (58) будет R_L .

ЛИТЕРАТУРА

§ 10.1. Т. Андерсон (1948), Ватсон (1952), (1955), (1967), Ватсон и Хеннан (1956), Гренандер (1954), Гренандер и Розенблатт (1957), Зискинд (1967), Магнесс и Макгир (1962), Розенблатт (1956), Хеннан (1958).

§ 10.2. Т. Андерсон (1948), (1958), Ватсон (1955), (1967), Гренандер (1954), Гренандер и Розенблатт (1957), Зискинд (1967), Канторович (1948), Розенблатт (1956), Хеннан (1961), (1963).

§ 10.3. Хеннан (1958).

§ 10.4. Р. Андерсон и Т. Андерсон (1950), Дурбин (1970), Дурбин и Ватсон (1950), (1951), Курант и Гильберт (1937).

УПРАЖНЕНИЯ

1. (Разд. 10.2.1) Покажите, что ковариационная матрица векторов $\mathbf{b}$ и $\mathbf{b}^*$ равна

$$\begin{aligned} & \begin{pmatrix} \mathbf{z}(\mathbf{b} - \beta)(\mathbf{b} - \beta)' & \mathbf{z}(\mathbf{b} - \beta)(\mathbf{b}^* - \beta)' \\ \mathbf{z}(\mathbf{b}^* - \beta)(\mathbf{b} - \beta)' & \mathbf{z}(\mathbf{b}^* - \beta)(\mathbf{b}^* - \beta)' \end{pmatrix} = \\ &= \begin{pmatrix} (\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^{-1} & (\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^{-1} \\ (\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^{-1} & (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\Sigma\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} \end{pmatrix} = \\ &= \begin{pmatrix} (\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^{-1} & 0 \\ 0 & (\mathbf{Z}'\mathbf{Z})^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{Z}'\Sigma^{-1}\mathbf{Z} & \mathbf{Z}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{Z} & \mathbf{Z}'\Sigma\mathbf{Z} \end{pmatrix} \begin{pmatrix} (\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^{-1} & 0 \\ 0 & (\mathbf{Z}'\mathbf{Z})^{-1} \end{pmatrix}. \end{aligned}$$

2. (Разд. 10.2.1) Докажите, что для невырожденной матрицы $\mathbf{G}$ соотношение (15) равносильно соотношению $\mathbf{Z} = \mathbf{V}^*\mathbf{C}$, в котором $\Sigma\mathbf{V}^* = \mathbf{V}^*\mathbf{A}$, матрица $\mathbf{A}$ диагональна, а $\mathbf{C}$ не вырождена.

3. (Разд. 10.2.1) Докажите, что если ранг матрицы $\mathbf{A}$ размера $T \times p$ равен r , то существует ортогональная матрица $\mathbf{Q}$ размера $T \times T$, такая, что последние $T - r$ строк матрицы $\mathbf{Q}'\mathbf{A}$ состоят целиком из нулей.

4. (Разд. 10.2.1) Составьте алгоритм для определения максимального числа линейно независимых линейных комбинаций столбцов матрицы $\mathbf{Z}$, являющихся характеристическими векторами матрицы Σ .

5. (Разд. 10.2.2) Покажите, что оценка $\mathbf{b}^*$ является эффективной тогда и только тогда, когда ранг ковариационной матрицы вектора $(\mathbf{b}', \mathbf{b}^{*'})$ равен p .

6. (Разд. 10.2.3) Покажите, что из условий $\sum_{i=1}^T z_i^2 \rightarrow \infty$ и $z_{T+1}^2 / \sum_{i=1}^T z_i^2 \rightarrow 0$ при $T \rightarrow \infty$ еще не вытекает существование предела $\lim_{T \rightarrow \infty} \sum_{i=1}^T z_i z_{i+h} / \sum_{i=1}^T z_i^2$. (Указание. Построить пример, в котором этот предел не существует, полагая $z_i^2 = \pm 1$ в такой последовательности, чтобы $\sum_{i=1}^{T-1} z_i z_{i+1} / \sum_{i=1}^{T-1} z_i^2$ достигало 0 и 1/2 для произвольно больших значений T .)

7. (Разд. 10.2.3) Покажите, что если $\sum_{h=-\infty}^{\infty} |\rho_{ii}(h)| < \infty$, $i = 1, \dots, p$, то существует положительно полуопределенная матричная функция $\mathbf{m}(\lambda)$, такая, что

$$\mathbf{R}(h) = \int_{-\pi}^{\pi} e^{i\lambda h} \mathbf{m}(\lambda) d\lambda.$$

8. (Разд. 10.2.3) Покажите, что если матрица $\mathbf{A}$ положительно определена, то из выполнения неравенства $\mathbf{u}'\mathbf{A}\mathbf{u} \leq \mathbf{u}'\mathbf{u}$ для всех $\mathbf{u}$ вытекает выполнение для всех $\mathbf{u}$ и неравенства $\mathbf{u}'\mathbf{u} \leq \mathbf{u}'\mathbf{A}^{-1}\mathbf{u}$ и обратно.

9. (Разд. 10.2.3) Пусть матрица $\mathbf{A}$ симметрична, а матрица $\mathbf{B}$ положительно определена. Покажите, что выполнение неравенства $\mathbf{x}'\mathbf{A}\mathbf{x} \leq \mathbf{x}'\mathbf{B}\mathbf{x}$ для всех $\mathbf{x}$ равносильно тому, что все корни уравнения $|\mathbf{A} - \theta\mathbf{B}| = 0$ не превосходят 1. (Указание. Использовать упр. 30 гл. 6.)

10. (Разд. 10. 2. 3) Докажите, что

$$2\pi \int_{-\pi}^{\pi} f(\lambda) dM(\lambda) = \int_m^M u dT(u).$$

11. (Разд. 10. 2. 3) Пусть L_j — положительно полуопределенная матрица размера $p \times p$ ранга n_j , $j = 1, \dots, G$. Покажите, что если $\sum_{j=1}^G n_j = p$ и

$$\sum_{j=1}^G L_j = I,$$

то

$$L_i L_j = 0, \quad i \neq j.$$

(Указание. Существует ортогональная матрица Q_1 , такая, что матрица $Q_1' L_1 Q_1 = D_1$ диагональна. Доказать, что ранг D_1 равен n_1 , матрица $Q_1' \sum_{j=2}^G L_j Q_1 = D_1^*$ диагональна и имеет ранг $p - n_1$,

$$D_1 = \begin{pmatrix} D_{11} & 0 \\ 0 & 0 \end{pmatrix}, \quad D_1^* = \begin{pmatrix} 0 & 0 \\ 0 & D_{12} \end{pmatrix},$$

$D_{11} = I$, $D_{12} = I$, и применить индукцию.)

12. (Разд. 10. 2. 3) Покажите, что из равенства $\int_S d\mathcal{R}M(\lambda) = 0$ вытекает равенство $\int_S d\mathcal{J}M(\lambda) = 0$. (Указание. Использовать тот факт, что матрица $\int_S dM(\lambda)$ эрмитова и положительно полуопределена.)

13. (Разд. 10.2.3) Покажите, что для $h = 0, 1, \dots$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^{T-h} \cos vt \cos \lambda(t+h) = \begin{cases} \frac{1}{2} \cos \lambda h, & 0 < v = \lambda < \pi, \\ 0, & 0 \leq v \neq \lambda \leq \pi, \end{cases}$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^{T-h} \cos vt \sin \lambda(t+h) = \begin{cases} \frac{1}{2} \sin \lambda h, & 0 < v = \lambda < \pi, \\ 0, & 0 \leq v \neq \lambda \leq \pi, \end{cases}$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^{T-h} \sin vt \sin \lambda(t+h) = \begin{cases} \frac{1}{2} \cos \lambda h, & 0 < v = \lambda < \pi, \\ 0, & 0 \leq v \neq \lambda \leq \pi. \end{cases}$$

14. (Разд. 10.2.3) Покажите, что

$$\lim_{T \rightarrow \infty} \frac{1}{T^{n+1}} \sum_{t=1}^T t^n = \frac{1}{n+1}, \quad n = 0, 1, \dots,$$

(Указание.

$$\int_0^T t^n dt < \sum_{t=1}^T t^n < \int_1^{T+1} t^n dt.)$$

15. (Разд. 10.2.3) Покажите, что для $n = 0, 1, \dots$

$$\lim_{T \rightarrow \infty} \frac{1}{T^{n+1}} \sum_{t=1}^T t^n \cos vt = 0,$$

$$\lim_{T \rightarrow \infty} \frac{1}{T^{n+1}} \sum_{t=1}^T t^n \sin vt = 0,$$

$$0 < v < 2\pi.$$

Указание. Этот результат для $n = 0$ вытекает из леммы 4.4.1, а для $n = 1$ из упр. 2 гл. 9. Далее следует провести индукцию с использованием упр. 11 гл. 9 и соотношений

$$2 \sin \frac{1}{2} v \cos vt = \sin v \left(t + \frac{1}{2} \right) - \sin v \left(t - \frac{1}{2} \right),$$

$$2 \sin \frac{1}{2} v \sin vt = \cos v \left(t - \frac{1}{2} \right) - \cos v \left(t + \frac{1}{2} \right).$$

16. (Разд. 10.2.3) Покажите, что для $n = 0, 1, \dots$ и для $h = 0, 1, \dots$

$$\lim_{T \rightarrow \infty} \frac{2}{T^{n+1}} \sum_{t=1}^{T-h} t^n \cos vt \cos \lambda (t+h) = \begin{cases} \frac{1}{n+1} \cos \lambda h, & 0 < v = \lambda < \pi, \\ 0, & 0 \leq v \neq \lambda \leq \pi, \end{cases}$$

$$\lim_{T \rightarrow \infty} \frac{2}{T^{n+1}} \sum_{t=1}^{T-h} t^n \cos vt \sin \lambda (t+h) = \begin{cases} \frac{1}{n+1} \sin \lambda h; & 0 < v = \lambda < \pi, \\ 0, & 0 \leq v \neq \lambda \leq \pi, \end{cases}$$

$$\lim_{T \rightarrow \infty} \frac{2}{T^{n+1}} \sum_{t=1}^{T-h} t^n \sin vt \sin \lambda (t+h) = \begin{cases} \frac{1}{n+1} \cos \lambda h, & 0 < v = \lambda < \pi, \\ 0, & 0 \leq v \neq \lambda \leq \pi. \end{cases}$$

(Указание. Использовать упр. 2 из гл. 4 и упр. 14 и 15.)

17. (Разд. 10.3.2) Докажите (26).

18. (Разд. 10.4.1) Покажите, что если функция регрессии имеет вид $\mu_1 + \mu_2 (-1)^t$ и r_1^* определяется циклическим образом, то $r_1^* = 0$, $T = 4$,

$$\Pr \{r_1^* \geq R\} = \begin{cases} \frac{1}{2} - R, & -\frac{1}{2} \leq R \leq \frac{1}{2}, \quad T = 6, \\ \frac{1}{2} (1 - \sqrt{2} R)^2, & 0 \leq R \leq \frac{\sqrt{2}}{2}, \quad T = 8. \end{cases}$$

19. (Разд. 10.4.1) Покажите, что если функция регрессии имеет вид $\mu_1 + \mu_2 (-1)^t + \mu_3 \cos \pi t/2 + \mu_4 \sin \pi t/2$ и r_1^* определяется циклическим образом, то

$$\Pr \{r_1^* \geq R\} = \frac{1}{2} (1 - \sqrt{2} R), \quad -\frac{\sqrt{2}}{2} \leq R \leq \frac{\sqrt{2}}{2}, \quad T = 8.$$

20. (Разд. 10.4.2) Покажите, что характеристические корни симметрической матрицы B размера $T \times T$ равны 1 (кратности m) и 0 (кратности $T - m$), если $B^2 = B$ имеет ранг m . (Указание. Если характеристический вектор w матрицы B соответствует корню μ , то в матрице B^2 этот характеристический вектор соответствует корню μ^2 .)

21. (Разд. 10.4.2) Покажите, что для любых двух квадратных матриц A и B одинакового размера

$$\text{tr}(A+B)^2 = \text{tr}(A^2 + 2AB + B^2),$$

$$\text{tr}(A+B)^3 = \text{tr}(A^3 + 3A^2B + 3AB^2 + B^3),$$

$$\text{tr}(A+B)^4 = \text{tr}(A^4 + 4A^3B + 4A^2B^2 + 2ABAB + 4AB^3 + B^4).$$

22. (Разд. 10.4.2) Покажите, что (36) равно

$$\text{tr} A_1 - \text{tr}(Z'Z)^{-1} Z'A_1Z, \quad k=1,$$

$$\text{tr} A_1^2 - 2 \text{tr}(Z'Z)^{-1} Z'A_1^2Z + \text{tr}[(Z'Z)^{-1} Z'A_1Z]^2, \quad k=2,$$

$$\begin{aligned} & \frac{4}{3} \{ \text{tr} A_1^3 - 3 \text{tr}(Z'Z)^{-1} Z'A_1^3Z + 3 \text{tr}(Z'Z)^{-1} Z'A_1^2Z (Z'Z)^{-1} Z'A_1Z - \\ & \quad - \text{tr}[(Z'Z)^{-1} Z'A_1Z]^3 \}, \quad k=3. \end{aligned}$$

23. (Разд. 10.4.2) Покажите, что если $z_t = t - (T+1)/2$, $t = 1, \dots, T$ и A_1 — матрица, соответствующая сумме квадратов последовательных разностей, то

$$\text{tr} A_1 - \text{tr}(Z'Z)^{-1} Z'A_1Z = 1 - \frac{(T-2)(T+3)}{T(T+1)} = \frac{6}{T(T+1)}.$$

24. (Разд. 10.4.2) Пусть $\tau^{(j)} = (1, 2^j, 3^j, \dots, T^j)'$, $\tau^{(0)} = \varepsilon = (1, 1, \dots, 1)'$ и A_1 — матрица, соответствующая сумме квадратов последовательных разностей. Покажите, что

$$\begin{aligned} A_1 \tau^{(j)} &= \tau^{(j)} + \binom{j}{2} \tau^{(j-2)} + \dots + \binom{j}{2[j/2]} \tau^{(j-2[j/2])} + \\ &+ \frac{1}{2} (1, 0, \dots, 0, T^j - (T+1)^j)'. \end{aligned}$$

25. (Разд. 10.4.2) Начертите графики функций $(T+1-2t)/(T-1)$ и $\cos \pi(t-1/2)/T$ для значений $t = 1, \dots, T$.

26. (Разд. 10.4.2) Пусть A_1 — матрица из разд. 6.5.4, а $B = I - (1/T) \times \mathbf{1}\mathbf{1}'$. Покажите, что характеристические корни матрицы $B'A_1B$ равны

$$-\frac{1}{2}, 0, \quad T=2,$$

$$-\frac{2}{3}, 0, 0, \quad T=3,$$

$$\frac{\sqrt{5}-1}{4}, -\frac{1}{4}, -\frac{\sqrt{5}+1}{4}, 0, \quad T=4.$$

СТАТИСТИЧЕСКИЕ ДАННЫЕ

А.1. ИНДЕКС БЕВЕРИДЖА ЦЕН НА ПШЕНИЦУ

В § 4.5 был кратко описан ежегодный индекс цен, построенный Бевериджем (1921), по которому с 1500 по 1869 г. продавалась пшеница на европейских рынках. В табл. А.1.1 приведены как значения самого индекса, так и значения, полученные после выделения тренда в соответствии с § 4.5. Полученный в результате выделения тренда индекс представлен графически на рис. А.1.1. [Выделение тренда из указанного ряда рассматривалось также Гренджером и Хагсом (1969).]

Первые 60 корреляций затабулированы в табл. А.1.2 и представлены графически на рис. А.1.2. При этом $r_k^* = C_k^*/C_0^*$.

Бевериджем была также приближенно построена соответствующая спектрограмма $R^2(v)$ для частот вида k/N , где N — целые числа (как правило, делящиеся на 4) в интервале от 276 до 356. Беверидж (1922) вычислял значения $R^2(v)$ последовательно для частот, вблизи которых наблюдались локальные максимумы. Результаты вычислений приведены в табл. А.1.3. (Мы не исправляли ошибок вычислений, поскольку среди них как будто бы нет таких, которые могли бы привести к ошибочным выводам.) Значения выборочной спектральной плотности пропорциональны значениям спектрограммы. Последние представлены в логарифмическом масштабе на рис. А.1.3. Значения частот на рис. А.1.3 взяты с шагом, приблизительно равным 0.002.

Значения коэффициентов корреляции из табл. А.1.2 использовались для вычисления оценки $2\pi\hat{f}_T^*(\lambda)$ с окном Блэкмена — Тьюки при $a = 0,25$ (окно Хеннинга). Значения K полагались равными 10, 20 и 30, а значения λ — равными $2\pi j/200$, $j = 1, \dots, 100$. Полученные оценки табулированы в табл. А.1.4 и представлены в логарифмическом масштабе на рис. А.1.4. В табл. 5.5 были приведены оценки коэффициентов процессов авторегрессии, которыми можно приблизить ряд Бевериджа после выделения из него тренда. Умноженные на 2π оценки нормированной спектральной плотности по данным табл. 5.5 для $p = 2, 6, 8$ представлены в логарифмическом масштабе на рис. А.1.5. При этом в качестве оценок использовались спектральные плотности соответствующих процессов авторегрессии.

Таблица А.1.1

ИНДЕКС БЕВЕРИДЖА ЦЕН НА ПШЕНИЦУ С ВЫДЕЛЕННЫМ ТРЕНДОМ

Реаль- ный индекс		Индекс с выде- ленным трендом	Реаль- ный индекс		Индекс с выде- ленным трендом	Реаль- ный индекс		Индекс с выде- ленным трендом
1500	17	106	1534	16	76	1568	34	77
1501	19	118	1535	22	102	1569	36	80
1502	20	124	1536	22	100	1570	43	93
1503	15	94	1537	16	73	1571	55	112
1504	13	82	1538	19	86	1572	64	131
1505	14	88	1539	17	74	1573	79	158
1506	14	87	1540	17	74	1574	59	113
1507	14	88	1541	19	76	1575	47	89
1508	14	88	1542	20	80	1576	48	87
1509	11	68	1543	24	96	1577	49	87
1510	16	98	1544	28	112	1578	45	79
1511	19	115	1545	36	144	1579	53	90
1512	23	135	1546	20	80	1580	55	90
1513	18	104	1547	14	54	1581	55	87
1514	17	96	1548	18	69	1582	54	83
1515	20	110	1549	27	100	1583	56	85
1516	20	107	1550	29	103	1584	52	76
1517	18	97	1551	36	129	1585	76	110
1518	14	75	1552	29	100	1586	113	161
1519	16	86	1553	27	90	1587	68	97
1520	21	111	1554	30	100	1588	59	84
1521	24	125	1555	38	123	1589	74	106
1522	15	78	1556	50	156	1590	78	111
1523	16	86	1557	24	71	1591	69	97
1524	20	102	1558	25	71	1592	78	108
1525	14	71	1559	30	81	1593	73	100
1526	16	81	1560	31	84	1594	88	119
1527	25.5	129	1561	37	97	1595	98	131
1528	25.8	130	1562	41	105	1596	109	143
1529	26	129	1563	36	90	1597	106	138
1530	26	125	1564	32	78	1598	87	112
1531	29	139	1565	47	112	1599	77	99
1532	20	97	1566	42	100	1600	77	97
1533	18	90	1567	37	86	1601	63	80

Таблица А.1.1 (продолжение)

<i>Реаль- ный индекс</i>			<i>Реаль- ный индекс</i>			<i>Реаль- ный индекс</i>		
<i>Индекс с выде- ленным трендом</i>			<i>Индекс с выде- ленным трендом</i>			<i>Индекс с выде- ленным трендом</i>		
1602	70	90	1637	114	105	1672	72	84
1603	70	90	1638	103	97	1673	89	106
1604	63	80	1639	98	93	1674	114	134
1605	61	77	1640	103	99	1675	102	122
1606	66	81	1641	101	99	1676	85	102
1607	78	98	1642	110	107	1677	88	107
1608	93	115	1643	109	106	1678	97	115
1609	97	94	1644	98	96	1679	94	113
1610	77	93	1645	84	82	1680	88	104
1611	83	100	1646	90	88	1681	79	92
1612	81	99	1647	120	116	1682	74	84
1613	82	100	1648	124	122	1683	79	86
1614	78	94	1649	136	134	1684	95	101
1615	75	88	1650	120	119	1685	70	74
1616	80	92	1651	135	136	1686	72	75
1617	87	100	1652	100	102	1687	63	66
1618	72	82	1653	70	72	1688	60	62
1619	65	73	1654	60	63	1689	74	76
1620	74	81	1655	72	76	1690	75	79
1621	91	99	1656	70	75	1691	91	97
1622	115	124	1657	71	77	1692	126	134
1623	99	106	1658	94	103	1693	161	169
1624	99	106	1659	95	104	1694	109	111
1625	115	121	1660	110	120	1695	108	109
1626	101	105	1661	154	167	1696	110	111
1627	90	84	1662	116	126	1697	130	128
1628	95	97	1663	99	108	1698	166	163
1629	108	109	1664	82	91	1699	143	137
1630	147	148	1665	76	85	1700	103	99
1631	112	114	1666	64	73	1701	89	85
1632	108	108	1667	63	74	1702	76	72
1633	99	97	1668	68	80	1703	93	88
1634	96	92	1669	64	74	1704	82	77
1635	102	97	1670	67	78	1705	71	66
1636	105	98	1671	71	83	1706	69	64

Таблица А.1.1 (продолжение)

Реаль- ный индекс			Индекс с выде- ленным трендом			Реаль- ный индекс			Индекс с выде- ленным трендом		
1707	75	69	1742	94	90	1777	135	94			
1708	134	125	1743	85	81	1778	125	87			
1709	183	175	1744	89	84	1779	116	79			
1710	113	108	1745	109	102	1780	132	87			
1711	108	103	1746	110	102	1781	133	88			
1712	121	115	1747	109	100	1782	144	94			
1713	139	134	1748	120	109	1783	145	94			
1714	109	108	1749	116	104	1784	146	92			
1715	90	90	1750	101	90	1785	138	85			
1716	88	89	1751	113	99	1786	139	84			
1717	88	89	1752	109	95	1787	154	93			
1718	93	94	1753	105	90	1788	181	108			
1719	106	107	1754	94	80	1789	185	108			
1720	89	89	1755	102	85	1790	151	86			
1721	79	79	1756	141	117	1791	139	78			
1722	91	91	1757	135	112	1792	157	87			
1723	96	94	1758	118	95	1793	155	85			
1724	111	110	1759	115	91	1794	191	103			
1725	112	111	1760	111	88	1795	248	130			
1726	104	103	1761	127	100	1796	185	95			
1727	94	94	1762	124	97	1797	168	84			
1728	98	101	1763	113	88	1798	176	87			
1729	88	90	1764	122	95	1799	243	120			
1730	94	96	1765	130	101	1800	289	139			
1731	81	80	1766	137	106	1801	251	117			
1732	77	76	1767	148	113	1802	232	105			
1733	84	84	1768	142	108	1803	207	94			
1734	92	91	1769	143	108	1804	276	125			
1735	96	94	1770	176	131	1805	250	114			
1736	102	101	1771	184	136	1806	216	98			
1737	95	93	1772	164	119	1807	205	93			
1738	98	91	1773	146	106	1808	206	94			
1739	125	122	1774	147	105	1809	208	94			
1740	162	159	1775	124	88	1810	226	104			
1741	113	110	1776	119	84	1811	302	140			

Таблица А.1.1 (продолжение)

<i>Реаль- Индекс ный с выде- индекс ленным трендом</i>			<i>Реаль- Индекс ный с выде- индекс ленным трендом</i>			<i>Реаль- Индекс ный с выде- индекс ленным трендом</i>		
1812	261	121	1832	151	82	1851	180	86
1813	207	96	1833	144	80	1852	223	165
1814	209	96	1834	138	78	1853	294	138
1815	280	130	1835	145	82	1854	300	141
1816	381	178	1836	156	88	1855	297	138
1817	266	126	1837	184	102	1856	232	107
1818	197	94	1838	216	117	1857	179	82
1819	177	86	1839	204	107	1858	180	81
1820	170	84	1840	186	95	1859	215	97
1821	152	76	1841	197	101	1860	258	116
1822	156	77	1842	183	92	1861	236	107
1823	141	71	1843	175	88	1862	202	92
1824	142	71	1844	183	92	1863	174	79
1825	137	69	1845	230	115	1864	179	81
1826	161	82	1846	278	139	1865	210	94
1827	189	93	1847	179	90	1866	268	119
1828	226	114	1848	161	80	1867	267	118
1829	194	103	1849	150	74	1868	208	93
1830	217	110	1850	159	78	1869	224	102
1831	199	105						

Примечание: В графе «Реальный индекс» представлены значения среднего индекса, полученные усреднением отдельных индексов для соответствующих рынков. Начало отсчета определяется следующим условием: средняя цена за период с 1700 по 1745 г. равна 100.

Значения, указанные в графе «Индекс с выделенным трендом», получены для каждого года как процентное отношение реального индекса для этого года к среднему значению реальных индексов за период в 31 год, для которого данный год является серединой.

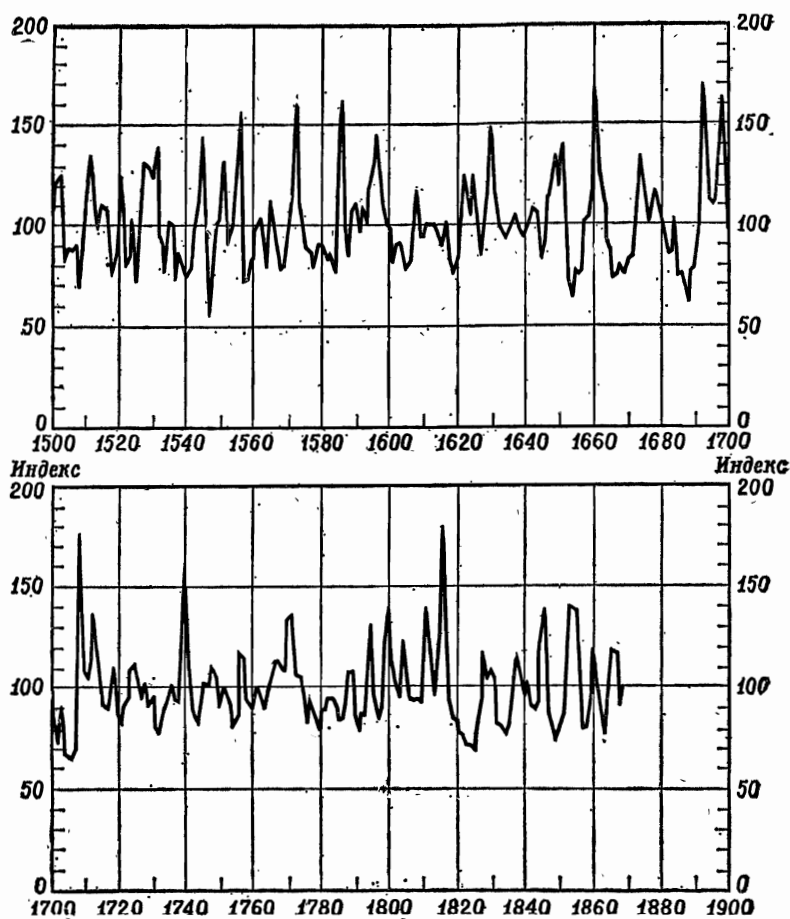


Рис. А.1.1.

Индекс Бевериджа цен на пшеницу с выделенным трендом.

Таблица А.1.2

КОРРЕЛЯЦИОННАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ИНДЕКСА БЕВЕРИДЖА ЦЕН НА ПШЕНИЦУ С ВЫДЕЛЕННЫМ ТРЕНДОМ

Порядок корреляции							
k	r_k^*	k	r_k^*	k	r_k^*	k	r_k^*
1	0.562	16	0.158	31	0.060	46	-0.036
2	0.103	17	0.109	32	-0.008	47	-0.013
3	-0.075	18	0.002	33	-0.039	48	0.042
4	-0.092	19	-0.075	34	0.007	49	0.062
5	-0.082	20	-0.062	35	0.056	50	0.065
6	-0.136	21	-0.021	36	0.010	51	0.050
7	-0.211	22	-0.062	37	-0.004	52	0.009
8	-0.261	23	-0.088	38	-0.015	53	-0.027
9	-0.192	24	-0.084	39	-0.047	54	-0.053
10	-0.070	25	-0.076	40	-0.047	55	-0.073
11	-0.003	26	-0.091	41	0.008	56	-0.106
12	-0.015	27	-0.052	42	0.034	57	-0.084
13	-0.012	28	-0.032	43	0.065	58	-0.019
14	0.047	29	-0.012	44	0.099	59	0.003
15	0.101	30	0.059	45	0.009	60	0.010

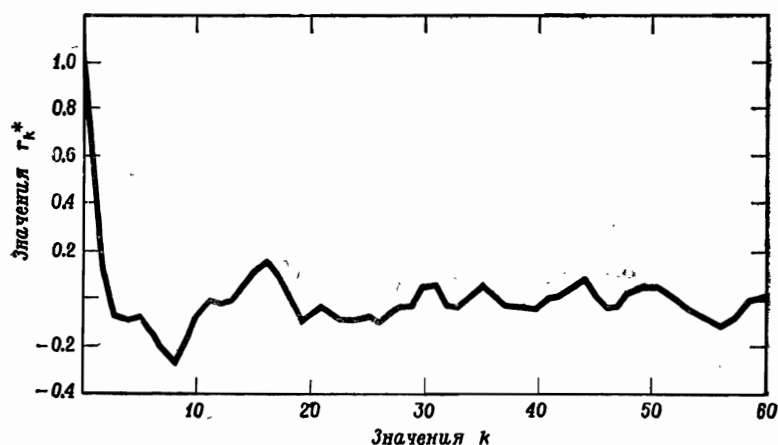


Рис. А.1.2.

Корреляции индекса Бевериджа цен на пшеницу с выделенным трендом.

Таблица А.1.3

СПЕКТРОГРАММА ИНДЕКСА БЕВЕРИДЖА ЦЕН НА ПШЕНИЦУ С ВЫДЕЛЕННЫМ ТРЕНДОМ

<i>Период</i>	<i>k</i>	<i>N</i>	$A\left(\frac{2\pi k}{N}\right)$	$B\left(\frac{2\pi k}{N}\right)$	$\frac{NR^2\left(\frac{2\pi k}{N}\right)}{300}$	<i>Частота</i>
2.000	150	300	0.11		0.01	0.5000
2.049	164	336	-0.40	-0.09	0.19	0.4881
2.054	148	304	0.48	-0.72	0.77	0.4868
2.061	165	340	0.38	-0.57	0.54	0.4853
2.069	145	300	0.25	0.63	0.46	0.4833
2.074	162	336	-0.61	0.51	0.71	0.4821
2.080	150	312	0.92	-0.50	1.14	0.4808
2.087	138	288	-0.52	-0.11	0.27	0.4792
2.095	147	308	-0.91	0.90	1.69	0.4773
2.105	152	320	0.90	0.07	0.86	0.4750
2.112	136	288	0.90	0.80	1.38	0.4722
2.133	150	320	0.89	0.15	0.84	0.4688
2.154	143	308	0.48	0.23	0.29	0.4643
2.182	132	288	1.32	-0.59	1.99	0.4583
2.200	140	308	-0.13	-0.60	0.39	0.4546
2.222	144	320	-0.32	-0.62	0.52	0.4500
2.261	138	312	0.50	-0.22	0.31	0.4423
2.286	140	320	-0.38	-0.85	0.93	0.4375
2.316	133	308	1.39	-1.05	3.11	0.4318
2.333	132	308	-0.10	-0.25	0.08	0.4286
2.353	136	320	0.90	0.07	0.86	0.4250
2.364	132	312	-0.12	-0.63	0.43	0.4231
2.370	135	320	0.05	-0.28	0.08	0.4219
2.375	128	304	0.29	-0.43	0.27	0.4210
2.381	126	300	-0.19	-1.22	1.53	0.4200
2.385	130	310	-1.00	-0.89	1.86	0.4194
2.391	138	330	-1.30	-0.54	2.18	0.4182
2.395	129	309	-0.72	0.60	0.90	0.4175
2.400	130	312	0.34	0.68	0.60	0.4167
2.412	136	328	-0.08	-0.65	0.47	0.4146
2.417	144	348	0.63	0.57	0.69	0.4138
2.435	138	336	0.44	0.01	0.22	0.4107
2.452	124	304	-1.40	-0.51	2.23	0.4079
2.462	130	320	-0.25	1.49	2.44	0.4062
2.476	126	312	-0.38	0.35	0.27	0.4038

Таблица А.1.3 (продолжение)

Период	k	N	$A\left(\frac{2\pi k}{N}\right)$	$B\left(\frac{2\pi k}{N}\right)$	$\frac{NR^2\left(\frac{2\pi k}{N}\right)}{300}$	Частота
2.483	116	288	-0.07	0.74	0.53	0.4028
2.500	128	320	-0.24	1.19	1.56	0.4000
2.512	129	324	0.86	0.39	0.97	0.3982
2.516	124	312	0.45	0.24	0.26	0.3974
2.529	119	301	-0.19	-0.31	0.13	0.3954
2.545	132	336	-1.39	-0.81	2.89	0.3929
2.555	126	322	0.38	0.50	0.42	0.3913
2.571	119	306	1.25	0.55	1.91	0.3889
2.588	119	308	0.30	0.43	0.28	0.3864
2.600	120	312	1.02	-0.39	1.25	0.3846
2.615	117	306	-0.75	-0.24	0.63	0.3824
2.625	112	294	-0.45	1.36	2.01	0.3810
2.643	112	296	0.95	-0.62	1.27	0.3784
2.667	117	312	-0.92	1.20	2.38	0.3750
2.687	112	301	1.23	-0.02	1.52	0.3721
2.692	117	315	-0.04	0.23	0.06	0.3714
2.706	119	322	-0.27	1.33	1.97	0.3696
2.714	112	304	0.83	1.17	2.10	0.3684
2.727	110	300	0.86	1.46	2.87	0.3667
2.733	105	287	2.05	1.19	6.16	0.3658
2.735	102	279	2.44	1.23	7.82	0.3656
2.737	114	312	2.23	1.00	6.22	0.3654
2.741	108	296	2.43	0.25	5.86	0.3649
2.750	112	308	0.90	-0.84	1.55	0.3636
2.762	126	348	-0.57	-0.04	0.37	0.3621
2.769	117	324	1.49	0.23	2.28	0.3611
2.778	117	325	1.20	-0.92	2.48	0.3600
2.800	120	336	-1.01	-0.19	1.18	0.3571
2.818	110	310	0.55	1.07	1.49	0.3548
2.833	114	323	0.78	-0.10	0.67	0.3529
2.846	104	296	0.41	0.42	0.34	0.3514
2.857	112	320	0.96	0.21	1.03	0.3500
2.875	112	322	0.35	0.14	0.15	0.3478
2.888	108	312	1.51	0.26	2.43	0.3462
2.895	114	330	-0.69	-1.57	3.21	0.3454
2.909	110	320	0.70	-1.11	1.84	0.3438

Таблица А.1.3. (продолжение)

Период	k	N	$A\left(\frac{2\pi k}{N}\right)$	$B\left(\frac{2\pi k}{N}\right)$	$\frac{NR^2\left(\frac{2\pi k}{N}\right)}{300}$	Частота
2.933	105	308	-0.04	0.39	0.16	0.3409
2.947	114	336	-0.93	-1.19	2.57	0.3393
2.960	100	296	-0.00	-1.15	1.30	0.3378
3.000	100	300	-0.29	-0.39	0.23	0.3333
3.040	100	304	0.09	0.75	0.58	0.3290
3.077	104	320	0.05	1.18	1.50	0.3250
3.111	108	336	0.91	-0.44	1.15	0.3214
3.143	98	308	2.01	0.23	4.20	0.3182
3.167	96	304	0.46	-1.05	1.33	0.3158
3.200	100	320	0.43	0.95	1.16	0.3125
3.217	92	296	1.25	0.00	1.55	0.3108
3.250	96	312	-1.22	-0.47	1.80	0.3077
3.273	99	324	-0.55	1.18	1.82	0.3056
3.286	98	322	-0.11	0.99	1.07	0.3044
3.304	92	304	0.13	0.75	0.59	0.3026
3.333	96	320	0.90	1.58	3.54	0.3000
3.364	88	296	1.76	0.98	4.00	0.2973
3.375	96	324	0.55	0.92	1.24	0.2963
3.385	91	308	0.35	1.03	1.21	0.2954
3.400	95	323	1.12	2.37	7.41	0.2941
3.407	81	276	2.98	2.81	14.90	0.2935
3.412	102	348	1.27	-3.98	15.53	0.2931
3.417	96	328	3.08	-2.24	15.84	0.2927
3.429	84	288	3.11	-1.40	11.16	0.2917
3.444	90	310	0.09	-0.99	1.03	0.2903
3.455	88	304	0.55	0.29	0.39	0.2895
3.462	91	315	1.57	1.02	4.87	0.2889
3.500	88	308	1.20	-0.94	2.38	0.2857
3.524	84	296	1.41	-1.18	3.31	0.2838
3.538	91	322	0.50	-1.45	2.53	0.2826
3.556	90	320	0.02	-0.43	0.20	0.2812
3.571	91	325	0.80	-0.69	1.21	0.2800
3.600	90	324	-1.03	0.82	1.88	0.2778
3.619	84	304	1.18	1.23	2.94	0.2763
3.636	88	320	1.14	0.13	1.39	0.2750
3.643	84	306	-0.16	0.27	0.10	0.2745

Таблица А.1.3 (продолжение)

Период	k	N	$A\left(\frac{2\pi k}{N}\right)$	$B\left(\frac{2\pi k}{N}\right)$	$\frac{NR^2\left(\frac{2\pi k}{N}\right)}{300}$	Частота
3.667	84	308	-2.14	-1.07	5.87	0.2727
3.679	84	309	0.34	-1.90	3.83	0.2718
3.692	78	288	1.28	-0.22	1.63	0.2708
3.700	80	296	0.90	-0.59	1.18	0.2703
3.714	84	312	1.15	1.78	4.65	0.2692
3.727	77	287	-0.45	-1.65	2.72	0.2683
3.750	84	315	0.64	-0.06	0.44	0.2667
3.778	81	306	-1.17	-0.68	1.86	0.2647
3.800	80	304	1.60	0.80	3.24	0.2632
3.833	84	322	-1.12	-1.63	4.17	0.2609
3.857	84	324	1.63	0.45	3.08	0.2593
3.888	72	280	-0.15	0.66	0.43	0.2571
3.895	76	296	-0.66	1.00	1.42	0.2568
3.923	78	306	0.64	-1.61	3.06	0.2549
3.962	78	309	-0.67	1.74	3.59	0.2524
4.000	75	300	1.47	-1.13	3.64	0.2500
4.077	78	318	0.57	-0.26	0.41	0.2453
4.111	72	296	1.13	-1.70	4.13	0.2432
4.143	70	290	-0.50	0.23	0.30	0.2414
4.167	78	325	1.21	0.32	1.70	0.2400
4.173	77	322	0.66	-1.46	2.77	0.2391
4.200	70	294	-0.99	-0.41	1.02	0.2381
4.250	76	323	0.50	-2.73	8.32	0.2353
4.286	70	300	-0.65	0.79	1.04	0.2333
4.333	72	312	-1.50	-1.30	4.10	0.2308
4.353	68	296	-2.85	-0.24	8.05	0.2297
4.364	66	288	-2.98	0.75	9.07	0.2292
4.375	72	315	-2.47	0.87	7.19	0.2286
4.385	78	342	-0.50	2.55	7.72	0.2281
4.400	70	308	-1.38	3.27	12.89	0.2273
4.412	68	300	0.08	3.62	13.11	0.2267
4.417	72	318	0.87	3.85	16.48	0.2264
4.429	70	310	1.80	2.41	9.32	0.2258
4.444	72	320	2.15	0.83	5.66	0.2250
4.471	68	304	0.91	0.79	1.48	0.2237
4.500	68	306	1.87	0.72	4.09	0.2222
4.571	70	320	-0.21	0.04	0.22	0.2188

Таблица А.1.3 (продолжение)

Период	k	N	$A\left(\frac{2\pi k}{N}\right)$	$B\left(\frac{2\pi k}{N}\right)$	$\frac{NR^2\left(\frac{2\pi k}{N}\right)}{300}$	Частота
4.600	70	322	-0.08	1.24	1.65	0.2174
4.667	72	336	0.19	0.93	1.00	0.2143
4.750	64	304	-0.12	2.28	5.28	0.2105
4.800	60	288	2.44	1.08	6.84	0.2083
4.857	63	306	-1.06	-1.30	2.89	0.2059
4.888	64	312	-1.80	2.11	8.00	0.2051
4.933	60	296	1.57	1.58	4.91	0.2027
5.000	60	300	1.85	1.00	4.30	0.2000
5.067	60	304	-0.05	3.98	16.09	0.1974
5.091	66	336	-0.73	5.55	35.05	0.1964
5.100	60	306	5.71	2.98	42.34	0.1961
5.111	63	322	5.70	0.29	34.91	0.1956
5.125	64	328	3.97	2.90	26.38	0.1951
5.143	63	324	2.46	2.46	13.09	0.1944
5.200	60	312	0.02	0.30	0.10	0.1923
5.250	56	294	1.74	1.92	6.56	0.1905
5.333	60	320	0.71	-4.46	21.72	0.1875
5.400	60	324	1.04	3.71	16.06	0.1852
5.415	60	325	4.27	1.90	23.66	0.1846
5.429	56	304	4.72	-0.28	22.61	0.1842
5.455	55	300	1.37	-3.73	15.76	0.1833
5.500	56	308	-1.04	1.49	3.39	0.1818
5.555	54	300	2.40	-0.68	6.23	0.1800
5.600	60	336	0.46	1.21	1.88	0.1786
5.667	54	306	5.31	-1.97	32.72	0.1765
5.692	52	296	2.05	-3.91	19.18	0.1757
5.714	56	320	0.35	-2.13	4.97	0.1750
5.750	56	322	1.39	-0.33	2.18	0.1739
5.800	50	290	3.55	-2.75	19.47	0.1724
5.846	52	304	0.00	-2.29	5.35	0.1710
5.933	60	356	4.37	0.91	23.63	0.1685
6.000	50	300	-3.50	-0.12	12.29	0.1667
6.111	54	330	-0.79	-1.90	4.66	0.1636
6.143	49	301	0.74	-2.96	9.32	0.1628
6.167	48	296	-0.22	-2.94	8.56	0.1622
6.200	50	310	-2.02	-3.38	16.02	0.1613
6.250	52	325	-3.23	-0.11	11.30	0.1600

Таблица А.1.3 (продолжение)

Период	k	N	$A\left(\frac{2\pi k}{N}\right)$	$B\left(\frac{2\pi k}{N}\right)$	$\frac{NR^2\left(\frac{2\pi k}{N}\right)}{300}$	Частота
6.286	49	308	-1.72	-0.59	3.41	0.1591
6.333	48	304	-1.52	1.29	4.02	0.1579
6.400	50	320	0.80	2.74	8.71	0.1562
6.500	48	312	0.69	-0.73	0.94	0.1538
6.571	49	322	1.49	-0.77	3.02	0.1522
6.667	48	320	0.25	0.21	0.11	0.1500
6.727	44	296	0.08	-0.13	0.02	0.1486
6.750	48	324	-0.20	-1.66	3.01	0.1482
6.800	45	306	0.23	-0.65	0.48	0.1471
6.909	44	304	0.58	2.56	7.00	0.1447
6.933	45	312	1.68	2.01	7.15	0.1442
7.000	44	308	3.10	-2.17	14.74	0.1429
7.143	42	300	1.83	-1.86	6.79	0.1400
7.200	45	324	0.54	-3.93	16.96	0.1389
7.333	42	308	1.52	-2.81	10.46	0.1364
7.400	40	296	-2.33	-2.72	12.65	0.1351
7.417	48	356	1.50	-4.01	21.72	0.1348
7.429	42	312	-3.80	-1.49	17.28	0.1346
7.500	42	315	0.17	1.50	2.40	0.1333
7.600	40	304	-2.33	-1.37	7.43	0.1316
7.667	42	322	-1.46	-2.61	9.57	0.1304
7.750	40	310	1.38	-0.39	2.13	0.1290
7.857	42	330	-0.50	0.28	0.36	0.1273
8.000	39	312	-3.96	1.34	18.63	0.1250
8.091	44	356	4.32	-0.98	23.23	0.1236
8.200	35	287	1.62	-0.64	2.90	0.1220
8.222	36	296	0.19	-0.56	0.34	0.1216
8.333	39	325	0.21	0.91	0.95	0.1200
8.500	38	323	0.17	3.19	10.41	0.1176
8.667	36	312	2.51	-1.01	7.59	0.1154
8.800	35	308	2.97	0.83	9.77	0.1136
9.000	34	306	-1.51	-0.57	2.65	0.1111
9.200	35	322	-0.16	-1.56	2.65	0.1087
9.333	36	336	-0.74	0.64	1.08	0.1071
9.500	32	304	1.08	1.07	2.26	0.1053
9.667	30	290	5.03	0.37	24.55	0.1034
9.750	32	312	4.46	-3.56	33.89	0.1026

Таблица А.1.3 (продолжение)

Период	k	N	$A\left(\frac{2\pi k}{N}\right)$	$B\left(\frac{2\pi k}{N}\right)$	$\frac{NR^2\left(\frac{2\pi k}{N}\right)}{300}$	Частота
9.818	33	324	1.21	-4.94	27.90	0.1018
10.000	32	320	-1.19	-0.83	2.25	0.1000
10.200	30	306	0.86	-0.22	0.80	0.0980
10.250	32	328	-0.69	1.10	1.84	0.0976
10.400	30	312	1.88	-1.65	6.52	0.0962
10.500	28	294	2.46	-1.82	9.19	0.0952
10.750	28	301	1.47	-3.13	11.98	0.0930
10.800	30	324	1.00	-4.75	25.48	0.0926
11.000	28	308	-3.85	-4.26	33.84	0.0909
11.200	30	336	-2.48	0.55	7.24	0.0893
11.500	28	322	-1.32	-0.66	2.34	0.0870
11.667	24	280	0.46	1.42	2.07	0.0857
12.000	26	312	-2.47	-4.04	23.30	0.0833
12.143	28	340	-0.22	-4.37	21.66	0.0824
12.333	24	296	-2.44	2.74	11.43	0.0811
12.500	26	325	-1.22	2.63	9.13	0.0800
12.667	24	304	2.28	5.19	32.58	0.0789
12.800	25	320	5.70	3.26	46.01	0.0781
12.875	24	309	6.46	0.77	43.58	0.0777
13.000	24	312	4.26	-4.32	38.23	0.0769
13.333	24	320	0.40	0.37	0.32	0.0750
13.500	24	324	2.56	-2.09	11.79	0.0741
13.667	24	328	3.49	-1.34	15.28	0.0732
14.000	22	308	1.15	-1.00	2.38	0.0714
14.500	20	290	-3.78	-0.18	13.82	0.0690
14.667	21	308	-1.50	4.23	20.69	0.0682
15.000	20	300	6.32	-2.66	46.83	0.0667
15.200	20	304	1.19	-8.52	75.04	0.0658
15.250	20	305	-0.28	-8.65	76.17	0.0656
15.286	21	321	-2.35	-7.15	60.62	0.0654
15.333	21	322	-3.89	-6.55	62.29	0.0652
15.500	20	310	-6.92	-2.02	59.11	0.0645
16.000	20	320	-1.46	4.52	24.02	0.0625
16.667	18	300	5.21	-0.39	27.33	0.0600
17.000	18	306	2.56	-6.35	47.84	0.0588
17.333	18	312	-3.04	-6.65	54.55	0.0577
17.500	16	280	-6.18	-4.45	54.12	0.0571

Таблица А.1.3 (продолжение)

Период	k	N	$A\left(\frac{2\pi k}{N}\right)$	$B\left(\frac{2\pi k}{N}\right)$	$\frac{NR^2\left(\frac{2\pi k}{N}\right)}{300}$	Частота
18.000	17	306	-4.40	1.25	21.29	0.0556
18.500	16	296	-1.46	2.25	7.10	0.0540
19.000	16	304	1.00	-0.23	1.07	0.0526
19.750	16	316	-4.73	-1.59	26.25	0.0506
20.000	16	320	-5.71	1.69	37.88	0.0500
21.000	14	294	0.78	2.61	7.28	0.0476
22.000	14	308	1.87	1.58	6.18	0.0454
23.000	14	322	-2.45	-1.43	8.61	0.0435
24.000	12	288	0.45	5.19	26.10	0.0417
24.667	12	296	4.31	1.99	22.21	0.0405
25.000	13	325	3.86	-0.19	14.94	0.0400
26.000	12	312	1.23	-1.34	3.43	0.0385
27.000	12	324	0.50	-0.33	0.38	0.0370
28.000	11	308	-0.49	0.68	0.72	0.0357
29.000	10	290	1.08	-2.12	5.46	0.0345
30.000	10	300	-1.53	-2.34	7.81	0.0333
31.000	10	310	-1.98	0.13	4.06	0.0322
32.000	10	320	-0.37	0.51	0.42	0.0312
33.000	10	330	0.96	-0.78	1.68	0.0303
34.000	9	306	-3.00	-2.15	13.90	0.0294
35.000	8	280	-4.64	1.79	23.11	0.0286
36.000	8	288	-1.65	4.85	23.29	0.0278
37.000	8	296	2.08	3.92	19.47	0.0270
38.000	8	304	2.99	0.56	9.37	0.0263
40.000	8	320	-1.44	-0.63	2.63	0.0250
41.000	8	328	-1.93	0.93	5.01	0.0244
42.000	7	294	0.93	3.02	9.75	0.0238
44.000	7	308	3.00	-0.14	9.27	0.0227
45.000	7	315	1.69	-1.99	7.14	0.0222
46.000	7	322	0.16	-2.27	5.58	0.0217
48.000	6	288	-0.76	-0.09	0.56	0.0208
50.000	6	300	1.83	2.19	8.14	0.0200
52.000	6	312	4.77	-0.57	24.03	0.0192
53.000	6	318	4.22	-2.60	26.08	0.0189
54.000	6	324	2.84	-4.01	26.09	0.0185
55.000	6	330	3.54	-3.30	25.82	0.0182

Таблица А.1.3 (продолжение)

Период	k	N	$A\left(\frac{2\pi k}{N}\right)$	$B\left(\frac{2\pi k}{N}\right)$	$\frac{NR^2\left(\frac{2\pi k}{N}\right)}{300}$	Частота
56.000	6	336	3.31	-2.36	18.47	0.0178
58.000	5	290	3.89	1.49	16.82	0.0172
60.000	5	300	-3.08	-0.93	10.32	0.0167
62.000	5	310	-1.62	0.39	2.88	0.0161
64.000	5	320	-0.78	0.13	0.66	0.0156
66.000	5	330	-0.56	-0.56	0.69	0.0152
68.000	5	340	2.90	-1.88	13.58	0.0147
70.000	4	280	-0.69	-0.16	0.47	0.0143
74.000	4	296	-1.20	0.82	2.07	0.0135
76.000	4	304	-0.66	1.17	1.83	0.0132
78.000	4	312	0.58	1.26	2.00	0.0128
80.000	4	320	0.77	0.82	1.34	0.0125
84.000	4	336	0.26	0.69	0.62	0.0119

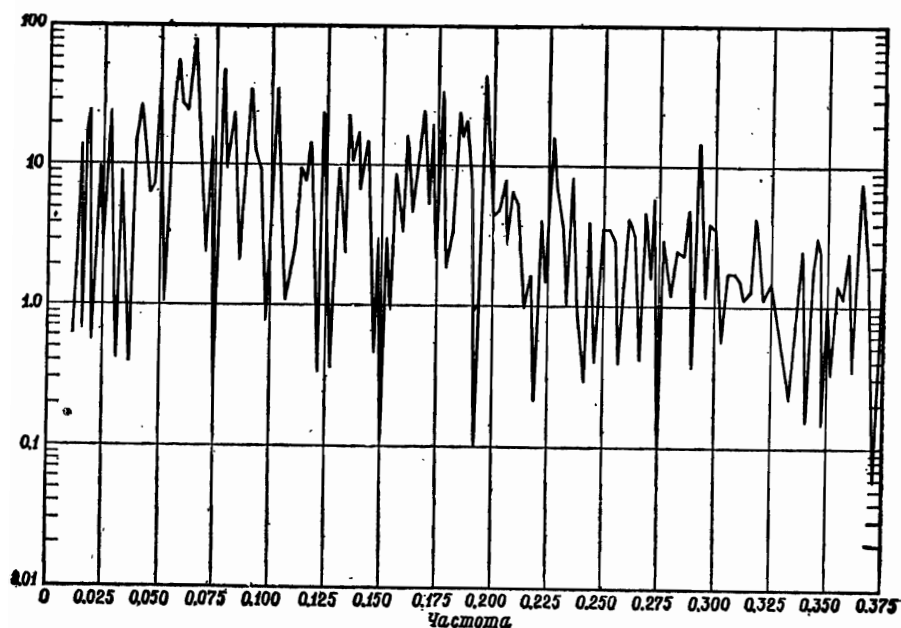


Рис. А.1.3.

Спектрограмма индекса Бевериджа цен на пшеницу с выделенным трендом.

Таблица А.1.4

НОРМИРОВАННАЯ ОЦЕНКА БЛЭКМЕНА — ТЬЮКИ ДЛЯ СПЕКТРАЛЬНОЙ ПЛОТНОСТИ

Частота	K			Частота	K		
	10	20	30		10	20	30
0.005	0.2751	0.1293	0.0971	0.205	0.1910	0.2156	0.2116
0.010	0.2785	0.1426	0.1152	0.210	0.1836	0.1951	0.1775
0.015	0.2840	0.1648	0.1412	0.215	0.1758	0.1742	0.1495
0.020	0.2914	0.1954	0.1718	0.220	0.1677	0.1546	0.1297
0.025	0.3003	0.2336	0.2059	0.225	0.1593	0.1371	0.1177
0.030	0.3104	0.2781	0.2455	0.230	0.1508	0.1225	0.1115
0.035	0.3211	0.3265	0.2942	0.235	0.1422	0.1110	0.1080
0.040	0.3320	0.3760	0.3544	0.240	0.1337	0.1023	0.1046
0.045	0.3427	0.4232	0.4245	0.245	0.1254	0.0962	0.0997
0.050	0.3527	0.4646	0.4975	0.250	0.1174	0.0919	0.0931
0.055	0.3615	0.4968	0.5625	0.255	0.1097	0.0889	0.0859
0.060	0.3688	0.5175	0.6072	0.260	0.1026	0.0866	0.0800
0.065	0.3743	0.5252	0.6229	0.265	0.0959	0.0848	0.0767
0.070	0.3778	0.5199	0.6066	0.270	0.0898	0.0830	0.0765
0.075	0.3793	0.5027	0.5627	0.275	0.0843	0.0812	0.0788
0.080	0.3785	0.4761	0.5008	0.280	0.0793	0.0792	0.0818
0.085	0.3758	0.4431	0.4330	0.285	0.0748	0.0769	0.0836
0.090	0.3710	0.4071	0.3705	0.290	0.0708	0.0743	0.0828
0.095	0.3646	0.3710	0.3209	0.295	0.0673	0.0712	0.0789
0.100	0.3566	0.3376	0.2873	0.300	0.0641	0.0677	0.0724
0.105	0.3475	0.3085	0.2685	0.305	0.0612	0.0639	0.0647
0.110	0.3375	0.2845	0.2608	0.310	0.0587	0.0597	0.0572
0.115	0.3270	0.2657	0.2593	0.315	0.0563	0.0555	0.0509
0.120	0.3163	0.2515	0.2589	0.320	0.0541	0.0513	0.0464
0.125	0.3056	0.2414	0.2559	0.325	0.0520	0.0476	0.0433
0.130	0.2951	0.2346	0.2483	0.330	0.0500	0.0443	0.0413
0.135	0.2852	0.2305	0.2364	0.335	0.0480	0.0417	0.0398
0.140	0.2759	0.2289	0.2222	0.340	0.0462	0.0398	0.0385
0.145	0.2672	0.2298	0.2097	0.345	0.0444	0.0385	0.0375
0.150	0.2593	0.2330	0.2030	0.350	0.0426	0.0377	0.0369
0.155	0.2520	0.2384	0.2056	0.355	0.0410	0.0371	0.0367
0.160	0.2453	0.2454	0.2187	0.360	0.0394	0.0365	0.0369
0.165	0.2392	0.2533	0.2406	0.365	0.0379	0.0359	0.0372
0.170	0.2333	0.2608	0.2671	0.370	0.0365	0.0350	0.0370
0.175	0.2277	0.2666	0.2919	0.375	0.0352	0.0340	0.0361
0.180	0.2222	0.2692	0.3089	0.380	0.0340	0.0329	0.0344
0.185	0.2165	0.2677	0.3133	0.385	0.0330	0.0317	0.0321
0.190	0.2107	0.2613	0.3032	0.390	0.0320	0.0307	0.0296
0.195	0.2045	0.2500	0.2800	0.395	0.0311	0.0298	0.0275
0.200	0.1980	0.2344	0.2476	0.400	0.0303	0.0293	0.0261

Таблица А.1.4 (продолжение)

К				К			
Частота	10	20	30	Частота	10	20	30
0.405	0.0296	0.0290	0.0258	0.455	0.0232	0.0229	0.0210
0.410	0.0289	0.0290	0.0265	0.460	0.0226	0.0214	0.0186
0.415	0.0283	0.0290	0.0280	0.465	0.0220	0.0201	0.0172
0.420	0.0277	0.0291	0.0299	0.470	0.0214	0.0190	0.0167
0.425	0.0270	0.0290	0.0316	0.475	0.0209	0.0182	0.0169
0.430	0.0264	0.0287	0.0324	0.480	0.0204	0.0177	0.0174
0.435	0.0258	0.0281	0.0320	0.485	0.0201	0.0174	0.0177
0.440	0.0252	0.0271	0.0303	0.490	0.0198	0.0173	0.0178
0.445	0.0245	0.0258	0.0275	0.495	0.0196	0.0172	0.0177
0.450	0.0239	0.0244	0.0242	0.500	0.0196	0.0172	0.0177

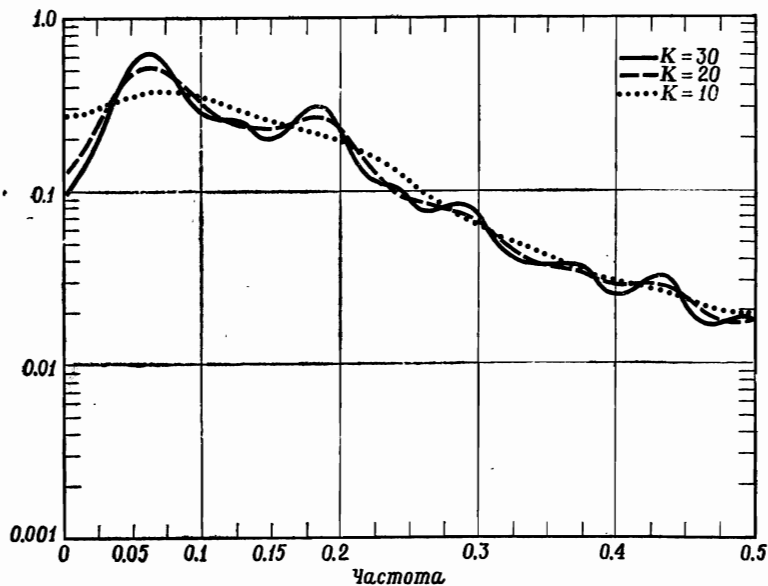


Рис. А.1.4.

Нормированная оценка Блэкмена — Тьюки ($a = 0.25$) спектральной плотности индекса Бевериджа цен на пшеницу с выделенным трендом.

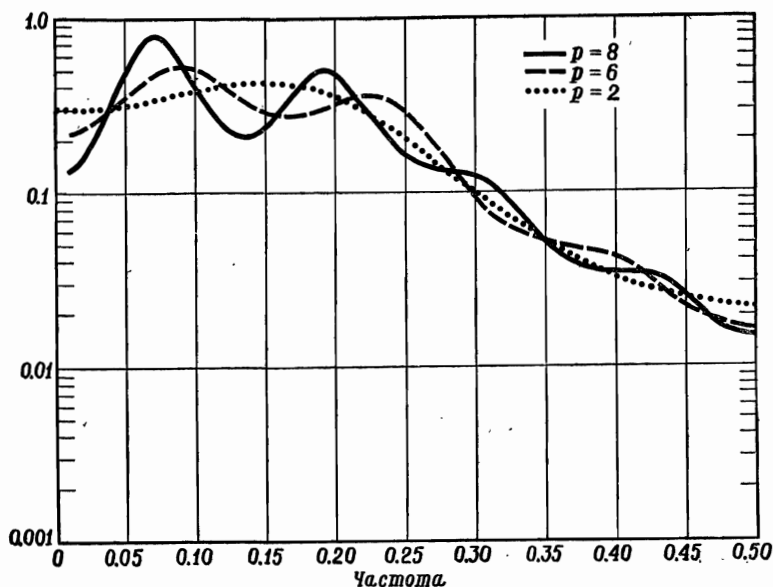


Рис. А.1.5.

Спектральные плотности подобранных процессов авторегрессии для индекса Бевериджа цен на пшеницу с выделенным трендом.

А.2. ТРИ ПРОЦЕССА АВТОРЕГРЕССИИ ВТОРОГО ПОРЯДКА, ПОЛУЧЕННЫЕ С ПОМОЩЬЮ СЛУЧАЙНЫХ ЧИСЕЛ

Используя случайные числа, Вольд (1965) получил реализации процессов авторегрессии вида

$$(1) \quad y_t + \beta_1 y_{t-1} + \beta_2 y_{t-2} = u_t,$$

где $\beta_1 = -\gamma$ и $\beta_2 = \gamma^2$. Корни характеристического уравнения для (1) равны $\gamma e^{\pm i2\pi/6}$. Случайные отклонения u_t представляли собой независимую выборку из нормального распределения с нулевым средним и дисперсией $(1 - \gamma^6)/(1 + \gamma^2)$. Значение последней выбрано так с тем, чтобы дисперсия y_t была равна 1. Каждая реализация начиналась со значений $y_{-1} = u_{-1}$ и $y_0 = u_0 + \gamma u_{-1}$ и имела длину $T = 200$. Вольд получил по 100 подобных реализаций для трех значений параметра γ : $\gamma = 0.25$, $\gamma = 0.7$, $\gamma = 0.9$. В табл. А.2.1 приведены вторые реализации для $\gamma = 0.25$, $\gamma = 0.7$ и первая реализация для $\gamma = 0.9$. На рис. А.2.1 — А.2.3 эти реализации представлены в графическом виде. У Вольда имеются графики и для других реализаций.

Таблица А.2.1
НАБЛЮДАЕМЫЙ РЯД

А. $\beta_1 = -0.25, \beta_2 = 0.0625$. (Выборка №2 Вольда)

t	y_t	t	y_t	t	y_t	t	y_t
1	1.2268	51	-0.1108	101	0.4048	151	-1.0168
2	0.3973	52	0.8899	102	-1.0996	152	0.7853
3	0.1334	53	0.9237	103	-0.5004	153	-1.4676
4	-0.3104	54	1.2217	104	-1.1657	154	-0.9854
5	-0.4614	55	-0.1803	105	0.3554	155	-1.8534
6	0.8008	56	0.8025	106	-0.5300	156	-1.7285
7	1.7893	57	-1.1287	107	0.2337	157	-1.1020
8	-0.5069	58	-0.0972	108	-1.1382	158	1.2012
9	1.6061	59	2.6301	109	-1.4920	159	1.6259
10	2.4069	60	0.5941	110	-0.7520	160	1.1087
11	-0.2187	61	0.6991	111	0.2857	161	1.8593
12	1.1178	62	0.4784	112	-0.7418	162	1.3407
13	0.3816	63	-3.0912	113	0.4471	163	2.0435
14	-0.4851	64	-0.5962	114	1.3734	164	0.5902
15	-1.3512	65	-1.1888	115	-1.1441	165	-0.9144
16	-0.8527	66	-2.0397	116	-1.5980	166	-1.2349
17	0.1690	67	-0.5262	117	0.3228	167	-1.8843
18	-0.0702	68	-0.7187	118	-0.2360	168	-0.2991
19	-0.8201	69	1.2073	119	-1.2372	169	0.2676
20	-0.4147	70	0.9935	120	-0.5845	170	-0.4726
21	-0.9631	71	-1.2825	121	2.1724	171	1.5744
22	-0.5125	72	-0.6821	122	1.8074	172	0.2603
23	-0.8878	73	-1.3399	123	0.1891	173	-1.5125
24	-1.1251	74	-1.2830	124	0.2040	174	-0.9015
25	0.1423	75	-0.3670	125	-0.2196	175	-1.3728
26	0.0267	76	-1.9386	126	-1.9188	176	1.0273
27	0.2646	77	-0.8190	127	0.8898	177	-0.2447
28	-1.2587	78	-2.0741	128	-1.0344	178	0.2098
29	-0.4948	79	0.3179	129	-0.1844	179	-1.1214
30	0.9793	80	-2.4611	130	-0.6813	180	-0.3899
31	-0.2555	81	-0.4927	131	-0.1136	181	-0.7416
32	-0.8822	82	-0.5582	132	-1.8186	182	-0.1831
33	-0.4405	83	-1.3925	133	0.5493	183	0.7315
34	0.1329	84	-1.0148	134	1.5896	184	0.6945
35	0.3030	85	-0.6616	135	0.1114	185	-0.9982
36	2.6584	86	-1.2155	136	-0.1702	186	0.2781
37	0.0972	87	2.0345	137	0.8204	187	-0.0199
38	-0.3308	88	2.3948	138	-2.2823	188	-0.3416
39	0.6895	89	0.5303	139	0.2210	189	0.2993
40	0.2827	90	-0.0755	140	-0.5385	190	0.9690
41	0.9795	91	-0.3050	141	0.0591	191	-0.9832
42	0.1590	92	0.2302	142	0.0450	192	1.1158
43	-1.7317	93	-0.0099	143	-0.1498	193	-0.7260
44	0.2491	94	-0.4606	144	-2.5719	194	0.7717
45	1.9132	95	0.6743	145	0.9076	195	0.5904
46	0.7171	96	0.7211	146	0.6395	196	1.2021
47	1.1631	97	-0.7078	147	0.9806	197	-0.2385
48	0.7018	98	0.0573	148	0.9813	198	-1.3084
49	0.6421	99	-0.0437	149	-0.0761	199	-0.2723
50	-1.4017	100	0.1556	150	-0.6912	200	-0.6236

$$y_{-1} = -0.0729, y_0 = 0.5054$$

Таблица А.2.1 (продолжение)

В. $\beta_1 = -0.7, \beta_2 = 0.49$. (Выборка №1 Вальда)

t	y_t	t	y_t	t	y_t	t	y_t
1	1.2134	51	-0.7713	101	0.5270	151	-1.4259
2	0.7306	52	0.7760	102	-0.6797	152	0.3695
3	0.0047	53	1.4720	103	-0.8929	153	-0.4131
4	-0.6077	54	1.4803	104	-1.1720	154	-0.9219
5	-0.7256	55	-0.0246	105	0.1055	155	-1.7906
6	0.5013	56	-0.0096	106	0.0993	156	-1.8542
7	1.9442	57	-1.0581	107	0.3260	157	-1.0438
8	0.3980	58	-0.5495	108	-0.7960	158	1.2637
9	0.7894	59	2.1837	109	-1.6632	159	2.3931
10	1.9233	60	1.7427	110	-1.1313	160	1.6726
11	0.3883	61	0.7170	111	0.3249	161	1.3340
12	0.3789	62	-0.0817	112	0.0993	162	0.8640
13	0.1451	63	-2.9210	113	0.4263	163	1.3986
14	-0.4891	64	-1.8409	114	1.2138	164	0.6851
15	-1.3703	65	-0.8355	115	-0.5171	165	-0.9469
16	-1.1521	66	-1.0947	116	-1.9294	166	-1.7676
17	0.1012	67	-0.4288	117	-0.5809	167	-2.0686
18	0.5038	68	-0.3307	118	0.2083	168	-0.5067
19	-0.3252	69	1.0528	119	-0.4882	169	0.8371
20	-0.6444	70	1.4121	120	-0.6738	170	0.3914
21	-1.0141	71	-0.6820	121	1.5455	171	1.2198
22	-0.6302	72	-1.4068	122	2.3861	172	0.5329
23	-0.5947	73	-1.6419	123	0.8122	173	-1.3982
24	-0.8493	74	-1.2459	124	-0.3867	174	-1.6421
25	-0.0111	75	-0.1707	125	-0.8740	175	-1.4496
26	0.3456	76	-1.0378	126	-1.8908	176	0.8325
27	0.4591	77	-0.9262	127	0.1803	177	0.8271
28	-0.8977	78	-1.7189	128	-0.0395	178	0.4369
29	-0.9831	79	-0.1265	129	-0.0130	179	-1.0428
30	0.5643	80	-1.3646	130	-0.5450	180	-1.0205
31	0.4553	81	-0.7803	131	-0.3392	181	-0.7700
32	-0.5585	82	-0.3447	132	-1.4244	182	-0.0564
33	-0.8012	83	-0.8773	133	-0.0401	183	0.9177
34	-0.1381	84	-1.0018	134	1.7319	184	1.0668
35	0.4880	85	-0.6640	135	1.0323	185	-0.5963
36	2.4647	86	-0.8573	136	-0.2043	186	-0.4871
37	1.0511	87	1.5475	137	0.0413	187	-0.1692
38	-0.6219	88	2.9393	138	-1.8527	188	-0.1330
39	-0.3329	89	1.3459	139	-0.6485	189	0.2940
40	0.1428	90	-0.5445	140	-0.1303	190	0.9634
41	1.0182	91	-1.2413	141	0.3912	191	-0.4270
42	0.5887	92	-0.3627	142	0.3350	192	0.3573
43	-1.4436	93	0.2857	143	-0.0820	193	-0.3867
44	-0.7501	94	0.0257	144	-2.2300	194	0.3658
45	1.5649	95	0.5038	145	-0.2982	195	0.7249
46	1.6647	96	0.7555	146	1.0837	196	1.2030
47	1.2739	97	-0.3891	147	1.6008	197	0.0886
48	0.4376	98	-0.4210	148	1.2053	198	-1.4585
49	0.1100	99	-0.1851	149	-0.1471	199	-1.0327
50	-1.3420	100	0.2117	150	-1.1781	200	-0.5138

$$y_{-1} = -0.0729, y_0 = 0.4403$$

Таблица А.2.1 (продолжение)

С. $\beta_1 = -0.9, \beta_2 = 0.81$. (Выборка №2. Вольда)

t	y_t	t	y_t	t	y_t	t	y_t
1	-0.9606	51	-0.0564	101	-0.8298	151	-0.7219
2	-1.0378	52	0.9502	102	-0.9909	152	-0.9665
3	-0.5484	53	0.2213	103	-0.2148	153	-0.8216
4	-1.0541	54	-0.8695	104	-0.0358	154	0.8067
5	-0.4834	55	-1.4278	105	0.9816	155	1.3913
6	0.6311	56	-1.0505	106	1.5830	156	1.2016
7	1.4114	57	0.4253	107	0.6708	157	-0.5651
8	0.3825	58	0.9019	108	-0.7696	158	-2.4274
9	-1.8239	59	0.7936	109	-1.6945	159	-1.6577
10	-1.4407	60	0.2804	110	0.0077	160	0.9373
11	-0.3132	61	-0.7642	111	1.3638	161	3.0506
12	-0.0874	62	-0.9103	112	1.0386	162	1.8579
13	-0.3918	63	-0.1189	113	-0.3156	163	-1.1132
14	-0.5888	64	1.4175	114	-1.0937	164	-2.9611
15	-0.5612	65	1.7326	115	-1.1822	165	-2.0225
16	0.0961	66	0.3792	116	-0.4685	166	0.5278
17	0.4954	67	-1.2491	117	0.7767	167	1.6486
18	0.1312	68	-0.4060	118	1.5031	168	1.8253
19	-0.5580	69	0.9696	119	1.1802	169	0.4094
20	-0.4685	70	2.1885	120	-0.8624	170	-0.7162
21	-0.1463	71	1.3596	121	-1.4084	171	-0.4403
22	0.5760	72	-0.8464	122	0.6630	172	0.4893
23	-0.0495	73	-1.2709	123	2.0564	173	0.9551
24	-1.0003	74	-0.5210	124	2.0627	174	0.4074
25	-1.8508	75	1.1009	125	0.7909	175	-0.5918
26	-1.0105	76	1.4884	126	-0.9423	176	-1.8504
27	0.7121	77	1.0229	127	-2.0920	177	-1.1401
28	0.4883	78	-0.8339	128	-0.5903	178	0.2450
29	0.2780	79	-2.0272	129	0.8511	179	1.3324
30	-0.2835	80	-1.4314	130	1.4750	180	1.3814
31	-0.5369	81	1.1195	131	0.7712	181	0.8292
32	-0.4541	82	1.5164	132	-1.5542	182	-0.3209
33	-0.2950	83	0.2551	133	-1.9841	183	-0.9560
34	-1.2701	84	-0.9049	134	-0.1077	184	0.4385
35	-1.0686	85	-0.8063	135	0.5974	185	0.8297
36	0.3006	86	0.0818	136	1.4342	186	0.9752
37	1.7958	87	1.8423	137	0.3418	187	-0.5203
38	1.3815	88	1.4011	138	-1.6969	188	-1.4231
39	-1.0157	89	-0.7910	139	-1.7782	189	-0.8991
40	-2.0178	90	-2.1526	140	-0.5346	190	0.3139
41	-0.6857	91	-2.0495	141	0.0004	191	0.7310
42	0.7797	92	-0.5901	142	0.2964	192	0.5303
43	0.8469	93	2.3577	143	0.5667	193	0.0094
44	0.0207	94	1.7757	144	0.4114	194	-0.3464
45	0.0693	95	-0.3490	145	-0.1130	195	-0.7523
46	0.8074	96	-1.9200	146	0.0279	196	-1.2773
47	1.2451	97	-1.3727	147	0.1295	197	-1.1861
48	0.0664	98	0.5940	148	0.2099	198	-0.1403
49	-0.8526	99	1.6419	149	-0.6045	199	1.6277
50	-0.1221	100	0.5217	150	-0.4056	200	1.0874

$$y_{-1} = -0.0630, y_0 = -0.5200$$

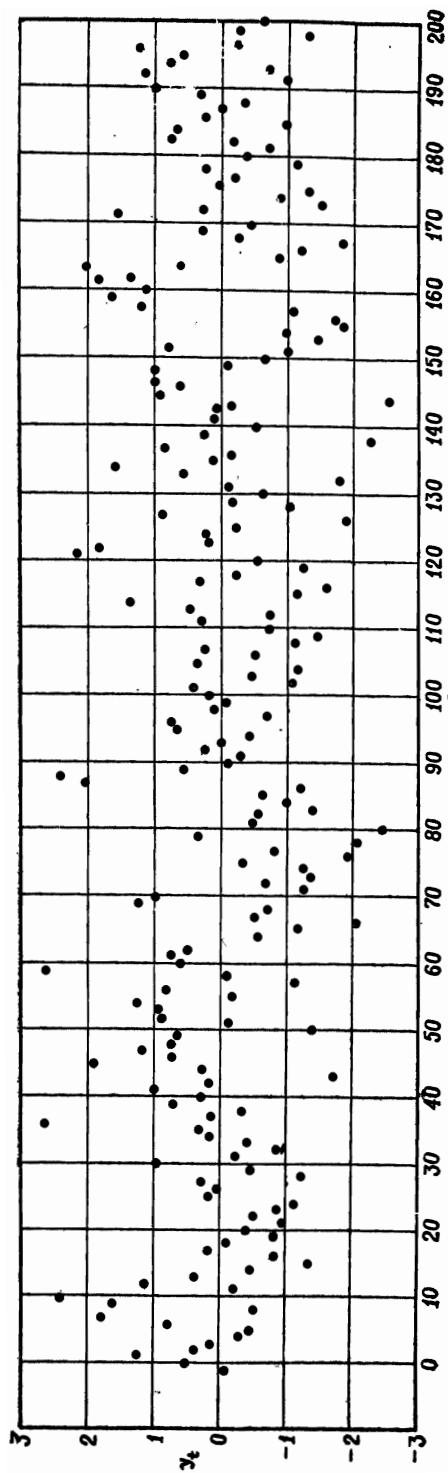


Рис. А.2.1

Результат моделирования процесса авторегрессии с $\gamma = 0.25$.

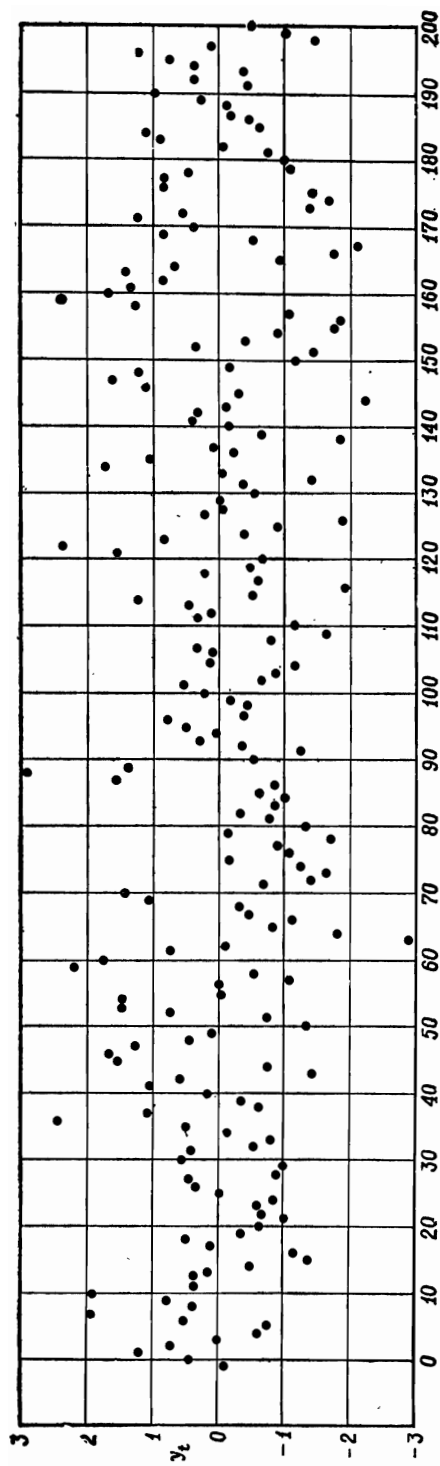


Рис. А.2.2.

Результат моделирования процесса авторегрессии с $\gamma = 0.7$.

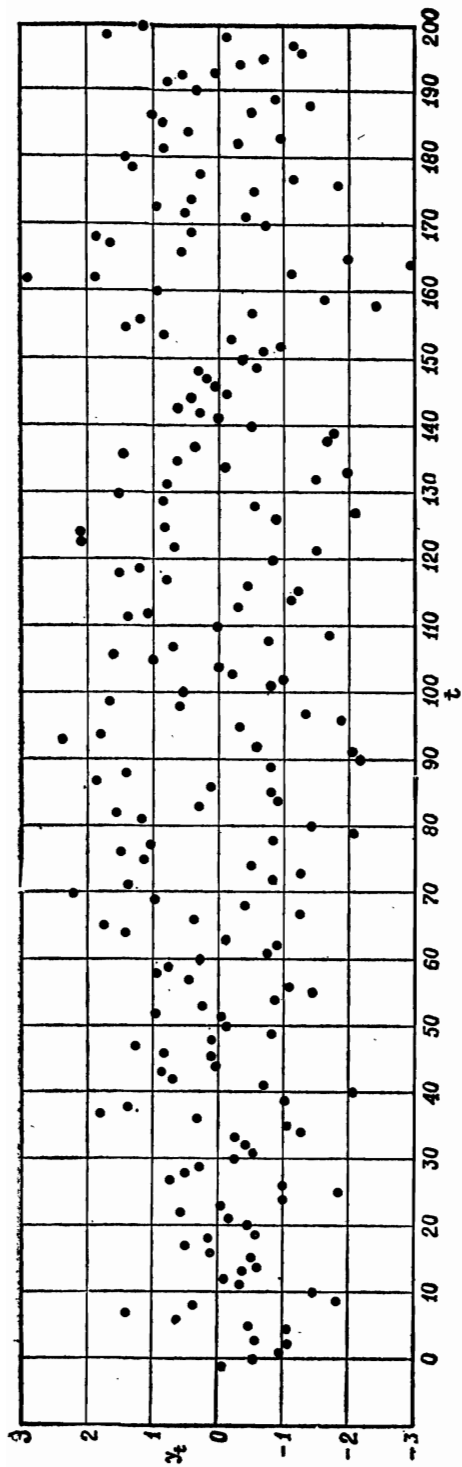


Рис. А.2.3.

Результат моделирования процесса авторегрессии с $\gamma = 0.9$.

Таблица А.2.2¹⁾
ДИСПЕРСИИ И КОРРЕЛЯЦИИ

А. $\beta_1 = -0.25$, $\beta_2 = 0.0625$. (Выборка №2 Вальда)

Корреляции

h	C_h/C_0	h	C_h/C_0	h	C_h/C_0	h	C_h/C_0
1	0.2473	26	0.0967	51	-0.0142	76	0.0515
2	0.1120	27	-0.1038	52	0.1814	77	-0.0103
3	0.0492	28	0.0269	53	0.0548	78	0.0435
4	-0.0868	29	0.0736	54	-0.0191	79	-0.1610
5	-0.0511	30	-0.0659	55	0.0126	80	-0.0683
6	0.0715	31	-0.0833	56	-0.1752	81	0.0749
7	-0.0224	32	-0.0679	57	-0.0746	82	0.0174
8	0.0167	33	-0.0796	58	0.0725	83	-0.0086
9	-0.0250	34	0.1311	59	0.0299	84	-0.1150
10	0.0291	35	0.1565	60	0.1072	85	-0.1234
11	0.0224	36	0.0443	61	0.0078	86	-0.0354
12	0.0906	37	-0.0264	62	0.0766	87	0.0468
13	0.0936	38	-0.0260	63	0.0151	88	0.1223
14	-0.0458	39	-0.0195	64	-0.1189	89	0.1668
15	-0.0656	40	0.0092	65	0.0401	90	-0.0106
16	-0.0813	41	0.0579	66	-0.0452	91	0.0304
17	-0.0677	42	0.1178	67	-0.1240	92	-0.0031
18	0.0124	43	0.0315	68	-0.1144	93	0.0038
19	-0.0920	44	-0.0223	69	0.0108	94	0.0671
20	-0.0527	45	-0.0206	70	-0.0334	95	-0.0456
21	-0.1819	46	0.0391	71	-0.0599	96	-0.1538
22	-0.1307	47	0.2125	72	-0.0062	97	-0.1372
23	0.0587	48	0.0303	73	0.0478	98	-0.1337
24	0.0278	49	0.1453	74	-0.0907	99	-0.0197
25	0.0023	50	0.0041	75	0.1844	100	0.0565

$$C_0 = 1.1248$$

¹⁾ Пояснения к табл. А.2.2 — А.2.4 и к рис. А.2.4 — А.2.9 см. на стр. 714.—
Прим. перев.

Таблица А.2.2 (продолжение)

В. $\beta_1 = -0.7$, $\beta_2 = 0.49$. (Выборка №2 Вальда)

Корреляции

h	C_h/C_0	h	C_h/C_0	h	C_h/C_0	h	C_h/C_0
1	0.5113	26	-0.0110	51	-0.0376	76	0.1672
2	-0.0473	27	-0.0877	52	0.1676	77	0.0497
3	-0.3001	28	0.0399	53	0.1763	78	-0.0920
4	-0.2549	29	0.1242	54	0.0466	79	-0.2366
5	-0.0187	30	-0.0005	55	-0.1064	80	-0.1108
6	0.1613	31	-0.1583	56	-0.2419	81	0.1287
7	0.0997	32	-0.2075	57	-0.1292	82	0.1821
8	-0.0174	33	-0.0826	58	0.0752	83	0.0397
9	-0.1026	34	0.1879	59	0.1397	84	-0.1851
10	-0.0562	35	0.2784	60	0.1258	85	-0.2533
11	0.0579	36	0.1175	61	0.0377	86	-0.1009
12	0.1689	37	-0.0864	62	0.0106	87	0.1201
13	0.1357	38	-0.1646	63	-0.0243	88	0.2470
14	-0.0435	39	-0.1078	64	-0.0454	89	0.1869
15	-0.1557	40	0.0305	65	0.0227	90	-0.0286
16	-0.1329	41	0.1551	66	-0.0369	91	-0.1088
17	-0.0103	42	0.1619	67	-0.1326	92	-0.0552
18	0.0901	43	0.0052	68	-0.1089	93	0.0818
19	0.0174	44	-0.1359	69	0.0122	94	0.1601
20	-0.1108	45	-0.1114	70	0.0313	95	0.0282
21	-0.2329	46	0.0724	71	-0.0160	96	-0.1766
22	-0.1449	47	0.2417	72	-0.0476	97	-0.2563
23	0.0673	48	0.1541	73	-0.0523	98	-0.1860
24	0.1277	49	0.0337	74	-0.0234	99	0.0093
25	0.0629	50	-0.1000	75	0.1654	100	0.1701

$$C_0 = 1.0837$$

Таблица А.2.2 (продолжение)

С. $\beta_1 = -0.9$, $\beta_2 = 0.81$. (Выборка №1 Вольда)

Корреляции

h	C_h/C_0	h	C_h/C_0	h	C_h/C_0	h	C_h/C_0
1	0.5011	26	-0.1094	51	-0.1857	76	-0.0906
2	-0.3858	27	-0.4210	52	-0.2829	77	-0.1983
3	-0.7726	28	-0.3308	53	-0.1160	78	-0.0742
4	-0.3711	29	0.0565	54	0.1330	79	0.1669
5	0.3371	30	0.3689	55	0.2395	80	0.2586
6	0.6311	31	0.3321	56	0.1117	81	0.0823
7	0.2843	32	-0.0101	57	-0.0918	82	-0.2097
8	-0.2853	33	-0.3317	58	-0.1697	83	-0.3125
9	-0.5178	34	-0.3582	59	-0.0819	84	-0.1023
10	-0.2291	35	-0.0762	60	0.0897	85	0.2119
11	0.2813	36	0.2481	61	0.1699	86	0.3365
12	0.5056	37	0.3487	62	0.1023	87	0.1426
13	0.2361	38	0.1280	63	-0.0524	88	-0.2054
14	-0.2348	39	-0.2110	64	-0.1794	89	-0.3622
15	-0.4581	40	-0.3080	65	-0.1506	90	-0.1475
16	-0.2402	41	-0.0980	66	-0.0135	91	0.1975
17	0.1880	42	0.1736	67	0.1098	92	0.2815
18	0.4203	43	0.2451	68	0.1333	93	0.0362
19	0.2518	44	0.0404	69	0.0326	94	-0.2656
20	-0.1416	45	-0.2251	70	-0.1287	95	-0.2595
21	-0.3953	46	-0.2614	71	-0.2059	96	0.0260
22	-0.2578	47	-0.0504	72	-0.1163	97	0.2556
23	0.1495	48	0.2085	73	0.0803	98	0.2187
24	0.4303	49	0.2903	74	0.1935	99	-0.0230
25	0.3123	50	0.0979	75	0.1078	100	-0.2242

$$C_0 = 1.2224$$

Таблица А.2.3

НОРМИРОВАННАЯ ВЫБОРОЧНАЯ СПЕКТРАЛЬНАЯ ПЛОТНОСТЬ

А. $\beta_1 = -0.25$, $\beta_2 = 0.0625$. (Выборка №2 Вольда)

k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$	k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$	k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$	k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$
0	1.5087						
1	1.5029	26	0.5243	51	0.1236	76	0.2538
2	0.5606	27	0.9970	52	1.7227	77	3.8915
3	1.0695	28	0.4683	53	0.1063	78	0.2601
4	3.8996	29	2.2405	54	0.4195	79	0.4844
5	3.3980	30	1.5876	55	0.5291	80	1.9956
6	0.6895	31	0.7753	56	0.3043	81	0.0868
7	0.8402	32	1.1055	57	0.1007	82	0.3921
8	1.7680	33	1.0494	58	1.5100	83	0.1838
9	1.9260	34	6.0570	59	0.3572	84	0.0284
10	0.1815	35	0.5693	60	0.0249	85	0.7412
11	0.7431	36	0.4151	61	2.2647	86	0.1808
12	0.6993	37	0.0128	62	1.2785	87	0.3800
13	3.0737	38	0.1682	63	0.8900	88	0.1344
14	0.9221	39	0.2226	64	1.0273	89	0.0748
15	0.6206	40	0.3264	65	0.5327	90	1.6157
16	5.8050	41	1.6880	66	0.3617	91	0.2269
17	1.2231	42	1.0313	67	0.0274	92	0.3397
18	1.4275	43	0.3604	68	0.8100	93	1.2194
19	1.0515	44	0.3232	69	3.5494	94	0.7040
20	1.2780	45	0.3270	70	1.1472	95	0.5200
21	1.1196	46	0.9212	71	0.3027	96	0.0886
22	1.1128	47	0.9579	72	0.7541	97	0.7174
23	1.7322	48	0.7229	73	0.2701	98	0.5979
24	3.0603	49	0.2063	74	0.2194	99	1.6979
25	0.4891	50	0.1732	75	0.1076	100	0.4220

Таблица А.2.3 (продолжение)

В. $\beta_1 = -0.7$, $\beta_2 = 0.49$. (Выборка №2 Вольда)

k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$	k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$	k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$	k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$
0	1.1189						
1	0.9664	26	1.0541	51	0.0877	76	0.0825
2	0.4030	27	1.8778	52	1.1702	77	1.0673
3	0.7823	28	1.0098	53	0.0617	78	0.0662
4	2.6466	29	5.0956	54	0.2595	79	0.1388
5	2.3570	30	3.5765	55	0.3281	80	0.5218
6	0.5149	31	1.7633	56	0.1525	81	0.0210
7	0.5969	32	2.4765	57	0.0426	82	0.1066
8	1.4281	33	2.4593	58	0.7922	83	0.0414
9	1.5064	34	15.0876	59	0.1625	84	0.0074
10	0.1523	35	1.2011	60	0.0080	85	0.1818
11	0.6217	36	0.8276	61	1.0016	86	0.0500
12	0.6459	37	0.0150	62	0.5806	87	0.1026
13	2.8789	38	0.3897	63	0.3768	88	0.0285
14	0.8066	39	0.4910	64	0.3807	89	0.0209
15	0.5380	40	0.5550	65	0.2053	90	0.4074
16	6.0875	41	2.7688	66	0.1410	91	0.0567
17	1.4132	42	1.4481	67	0.0067	92	0.0891
18	1.5177	43	0.5413	68	0.2936	93	0.2865
19	1.2673	44	0.4815	69	1.2444	94	0.1767
20	1.6300	45	0.3843	70	0.3671	95	0.1313
21	1.5202	46	1.0904	71	0.0891	96	0.0178
22	1.7333	47	0.9157	72	0.2535	97	0.1789
23	2.5668	48	0.7222	73	0.0851	98	0.1440
24	4.8883	49	0.2032	74	0.0706	99	0.3831
25	0.7624	50	0.1366	75	0.0374	100	0.1068

Таблица А.2.3 (продолжение)

С. $\beta_1 = -0.9, \beta_2 = 0.81$. (Выборка №1 Вольда)

k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$	k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$	k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$	k	$\frac{2\pi}{C_0} I\left(\frac{2\pi k}{200}\right)$
0	0.1835						
1	0.5806	26	3.6521	51	0.0902	76	0.0538
2	0.5042	27	2.6660	52	0.4664	77	0.0186
3	0.5937	28	0.9321	53	0.0246	78	0.0037
4	0.2787	29	0.6410	54	0.2046	79	0.1264
5	0.0850	30	0.1841	55	0.0311	80	0.0095
6	0.0430	31	4.7018	56	0.0621	81	0.0104
7	0.4200	32	15.7755	57	0.1469	82	0.0102
8	0.2520	33	24.4907	58	0.1169	83	0.0259
9	0.0175	34	6.2684	59	0.0315	84	0.0380
10	0.7423	35	8.8254	60	0.1073	85	0.0065
11	0.1452	36	0.1490	61	0.0048	86	0.0848
12	0.1305	37	7.4089	62	0.0045	87	0.0058
13	0.2365	38	0.0970	63	0.0708	88	0.0471
14	0.0259	39	2.6194	64	0.2145	89	0.0210
15	0.0354	40	1.3477	65	0.0586	90	0.0692
16	1.2747	41	0.2229	66	0.0730	91	0.0013
17	1.0938	42	1.8378	67	0.0349	92	0.0106
18	0.1598	43	0.9971	68	0.1006	93	0.0134
19	0.0107	44	0.2457	69	0.1786	94	0.0117
20	0.0607	45	0.1095	70	0.1164	95	0.0339
21	2.6171	46	0.5497	71	0.0238	96	0.0153
22	0.3760	47	0.5808	72	0.0073	97	0.0466
23	0.2992	48	1.0833	73	0.2072	98	0.0297
24	0.7754	49	0.1481	74	0.0119	99	0.0239
25	0.0146	50	0.3898	75	0.0025	100	0.0039

Таблица А.2.4

НОРМИРОВАННЫЕ ОЦЕНКИ БАРТЛЕТТА ДЛЯ СПЕКТРАЛЬНОЙ ПЛОТНОСТИ

А. $\beta_1 = -0.25$, $\beta_2 = 0.0625$. (Выборка №2 Вольда)

k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$	k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$	k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$	k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$
0	1.5349						
1	1.5345	26	1.2662	51	0.8224	76	0.7793
2	1.5334	27	1.2455	52	0.8153	77	0.7755
3	1.5316	28	1.2244	53	0.8092	78	0.7713
4	1.5291	29	1.2030	54	0.8041	79	0.7665
5	1.5258	30	1.1813	55	0.7999	80	0.7613
6	1.5217	31	1.1594	56	0.7965	81	0.7558
7	1.5169	32	1.1374	57	0.7940	82	0.7498
8	1.5112	33	1.1155	58	0.7921	83	0.7436
9	1.5048	34	1.0936	59	0.7908	84	0.7372
10	1.4975	35	1.0720	60	0.7900	85	0.7306
11	1.4894	36	1.0507	61	0.7897	86	0.7240
12	1.4805	37	1.0298	62	0.7898	87	0.7173
13	1.4707	38	1.0094	63	0.7900	88	0.7107
14	1.4600	39	0.9896	64	0.7905	89	0.7043
15	1.4484	40	0.9705	65	0.7909	90	0.6982
16	1.4360	41	0.9521	66	0.7914	91	0.6923
17	1.4226	42	0.9346	67	0.7917	92	0.6869
18	1.4084	43	0.9180	68	0.7919	93	0.6819
19	1.3933	44	0.9023	69	0.7918	94	0.6775
20	1.3774	45	0.8877	70	0.7913	95	0.6736
21	1.3607	46	0.8741	71	0.7905	96	0.6704
22	1.3432	47	0.8616	72	0.7892	97	0.6679
23	1.3249	48	0.8501	73	0.7875	98	0.6660
24	1.3060	49	0.8398	74	0.7853	99	0.6649
25	1.2864	50	0.8306	75	0.7826	100	0.6646

Таблица А.2.4 (продолжение)

В. $\beta_1 = -0.7$, $\beta_2 = 0.49$ (Выборка №2 Вольда)

k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$	k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$	k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$	k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$
0	1.2499						
1	1.2506	26	2.2623	51	0.5270	76	0.2781
2	1.2522	27	2.4148	52	0.4735	77	0.2712
3	1.2530	28	2.5912	53	0.4326	78	0.2621
4	1.2515	29	2.7682	54	0.4062	79	0.2497
5	1.2466	30	2.9190	55	0.3936	80	0.2341
6	1.2389	31	3.0173	56	0.3920	81	0.2160
7	1.2314	32	3.0422	57	0.3974	82	0.1969
8	1.2289	33	2.9823	58	0.4058	83	0.1788
9	1.2378	34	2.8375	59	0.4139	84	0.1633
10	1.2648	35	2.6194	60	0.4197	85	0.1519
11	1.3152	36	2.3489	61	0.4223	86	0.1452
12	1.3912	37	2.0529	62	0.4218	87	0.1430
13	1.4909	38	1.7591	63	0.4184	88	0.1447
14	1.6077	39	1.4920	64	0.4128	89	0.1491
15	1.7315	40	1.2691	65	0.4052	90	0.1551
16	1.8497	41	1.0987	66	0.3956	91	0.1618
17	1.9500	42	0.9798	67	0.3839	92	0.1687
18	2.0228	43	0.9037	68	0.3702	93	0.1755
19	2.0639	44	0.8572	69	0.3551	94	0.1821
20	2.0756	45	0.8254	70	0.3393	95	0.1886
21	2.0673	46	0.7952	71	0.3241	96	0.1949
22	2.0538	47	0.7579	72	0.3105	97	0.2004
23	2.0526	48	0.7096	73	0.2994	98	0.2049
24	2.0803	49	0.6515	74	0.2908	99	0.2078
25	2.1489	50	0.5885	75	0.2841	100	0.2087

Таблица А.2.4 (продолжение)

С. $\beta_1 = -0.9$, $\beta_2 = 0.81$. (Выборка №1 Вольда)

k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$	k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$	k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$	k	$\frac{2\pi}{C_0} f_T \left(\frac{2\pi k}{200} \right)$
0	0.4574						
1	0.4812	26	1.9682	51	0.2360	76	0.0302
2	0.5128	27	1.7071	52	0.2329	77	0.0398
3	0.4687	28	1.4244	53	0.1873	78	0.0529
4	0.3259	29	1.0239	54	0.1303	79	0.0636
5	0.2040	30	1.7788	55	0.1071	80	0.0535
6	0.2295	31	6.3443	56	0.0990	81	0.0329
7	0.3226	32	13.1094	57	0.1029	82	0.0301
8	0.3489	33	15.6088	58	0.1136	83	0.0375
9	0.3579	34	11.5562	59	0.1064	84	0.0389
10	0.3921	35	6.2179	60	0.0837	85	0.0427
11	0.3584	36	3.9872	61	0.0633	86	0.0480
12	0.2611	37	3.5342	62	0.0655	87	0.0435
13	0.1995	38	2.7475	63	0.1049	88	0.0384
14	0.2327	39	1.9772	64	0.1382	89	0.0439
15	0.4434	40	1.7151	65	0.1105	90	0.0475
16	0.7550	41	1.5282	66	0.0615	91	0.0372
17	0.8461	42	1.2119	67	0.0679	92	0.0228
18	0.7107	43	0.8837	68	0.1192	93	0.0174
19	0.7688	44	0.5833	69	0.1422	94	0.0212
20	1.0923	45	0.4081	70	0.1074	95	0.0289
21	1.2004	46	0.4382	71	0.0662	96	0.0373
22	0.8373	47	0.5520	72	0.0697	97	0.0414
23	0.5144	48	0.5691	73	0.0858	98	0.0371
24	0.9066	49	0.4415	74	0.0659	99	0.0281
25	1.7041	50	0.2890	75	0.0353	100	0.0239

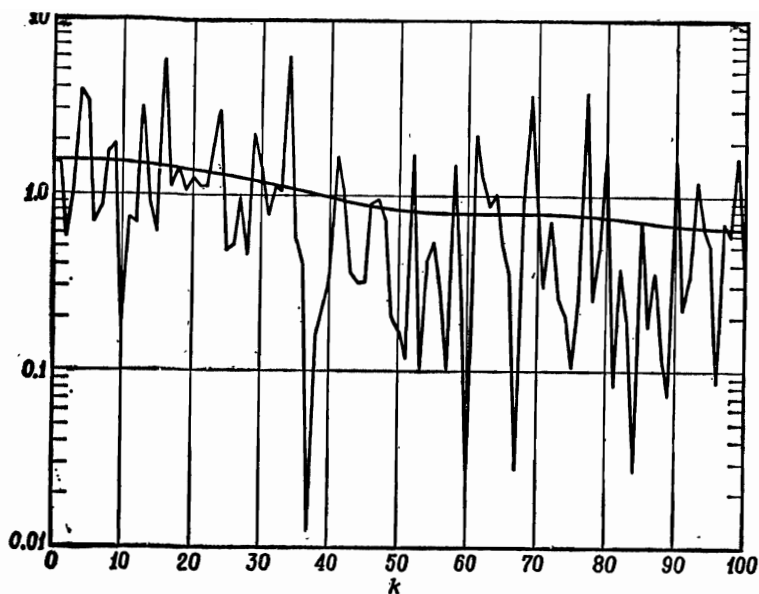


Рис. А.2.4.

Нормированная выборочная спектральная плотность и нормированная оценка Бартлетта для ряда с $\gamma = 0.25$.

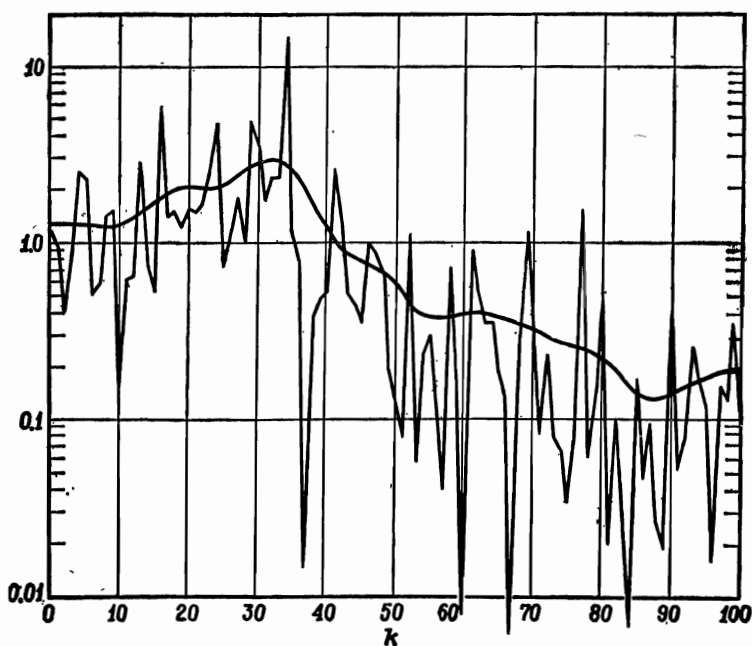


Рис. А.2.5.

Нормированная выборочная спектральная плотность и нормированная оценка Бартлетта для ряда $\gamma = 0.7$.

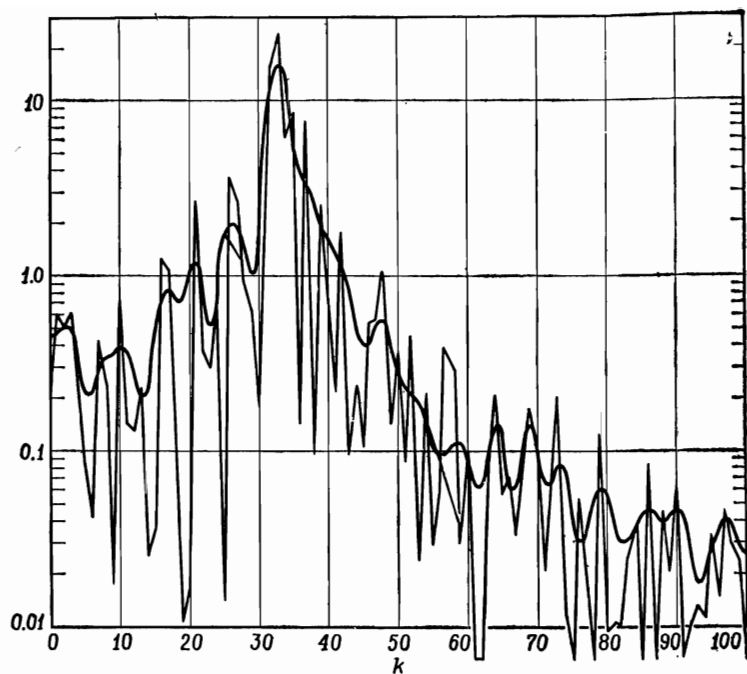


Рис. А.2.6.

Нормированная выборочная спектральная плотность и нормированная оценка Бартлетта для ряда с $\gamma = 0.9$.

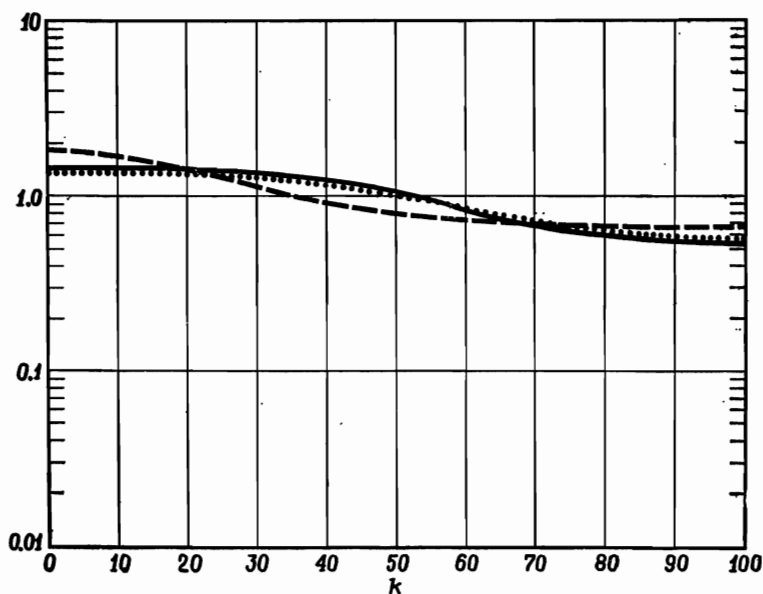


Рис. А.2.7.

Нормированные спектральные плотности исходного процесса (сплошная линия) и двух подобранных процессов для $\gamma = 0.25$.

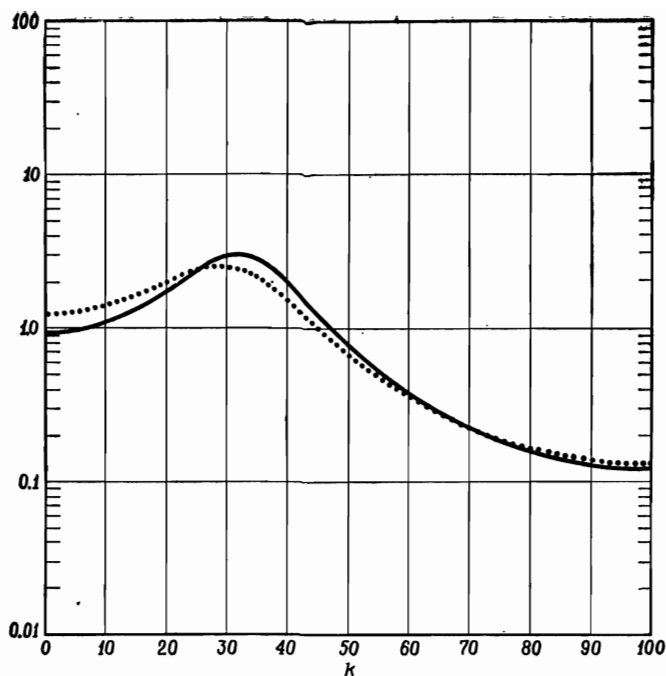


Рис. А.2.8.

Нормированные спектральные плотности исходного (сплошная линия) и подобранного процессов для $\gamma = 0.7$.

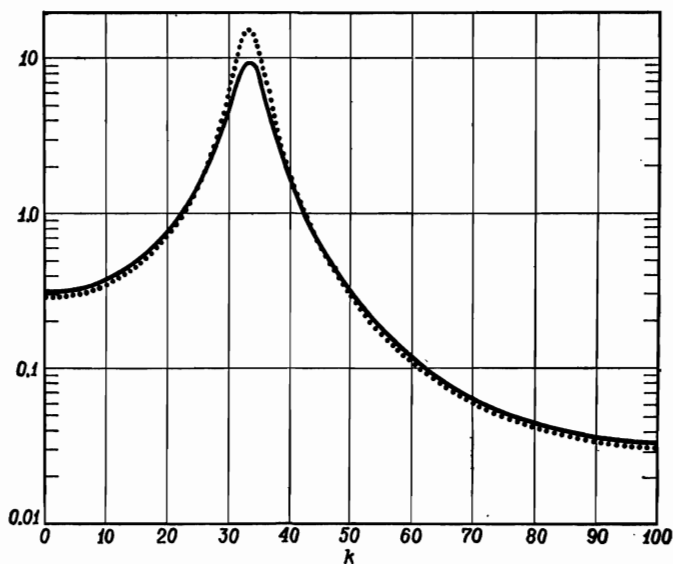


Рис. А.2.9.

Нормированные спектральные плотности исходного (сплошная линия) и подобранного процессов для $\gamma = 0.9$.

Для каждой из указанных трех реализаций в табл. А.2.2 приведены значения выборочной дисперсии и последовательность выборочных корреляций. Умноженные на 2π и деленные на выборочную дисперсию S_0 значения выборочных спектральных плотностей для этих трех случаев приведены в табл. А.2.3. Соответствующие графики построены на рис. А.2.4—А.2.6. В табл. А.2.4 указаны значения оценок Бартлетта (пример С из разд. 9.2.3) нормированных спектральных плотностей при $K = 20$. Графики этих оценок также представлены на рис. А.2.4—А.2.6. На рис. А.2.7—А.2.9 изображены графики (умноженных на 2π) нормированных спектральных плотностей исходных и подобранных (т. е. оцененных по реализациям) процессов авторегрессии второго порядка. (Для $\gamma = 0.25$ были исследованы две реализации.) Вертикальный масштаб на рис. А.2.4—А.2.9 логарифмический, а по горизонтальной оси откладываются значения $j = 200 \lambda / (2\pi)$.

А.3. ЧИСЛА СОЛНЕЧНОЙ АКТИВНОСТИ

О числах солнечной активности говорилось в § 5.9. Числа солнечной активности по Вольфу за период с 1749 по 1924 г. приведены в табл. А.3.1. На рис. 5.1 представлен график данных Вальдмейера. Соответствующая корреляционная последовательность изображена графически в § 5.9.

На основании данных о солнечной активности в каждом полугодии с 1749 по 1901 г., полученных суммированием соответствующих ежемесячных чисел солнечной активности, Шустер (1906) построил спектрограмму для частот вида k/N . Полученные им результаты приведены в табл. А.3.2 и изображены графически на рис. А.3.1 в логарифмическом масштабе. На рис. А.3.2 представлены графики оценок спектральной плотности, вычисленных Шерф (1964) по ежегодным данным (1749—1924) с использованием окон Парзена (пример J разд. 9.2.3) со значениями $K = 20, 40$ и 60 .

Таблица А.3.1

ЧИСЛА СОЛНЕЧНОЙ АКТИВНОСТИ ПО ВОЛЬФУ, 1749—1924 гг.

Год	Число Вольфа	Год	Число Вольфа	Год	Число Вольфа	Год	Число Вольфа
1749	80.9	1793	46.9	1837	138.3	1881	54.3
1750	83.4	1794	41.0	1838	103.2	1882	59.7
1751	47.7	1795	21.3	1839	85.8	1883	63.7
1752	47.8	1796	16.0	1840	63.2	1884	63.5
1753	30.7	1797	6.4	1841	36.8	1885	52.2
1754	12.2	1798	4.1	1842	24.2	1886	25.4
1755	9.6	1799	6.8	1843	10.7	1887	13.1
1756	10.2	1800	14.5	1844	15.0	1888	6.8
1757	32.4	1801	34.0	1845	40.1	1889	6.3
1758	47.6	1802	45.0	1846	61.5	1890	7.1
1759	54.0	1803	43.1	1847	98.5	1891	35.6
1760	62.9	1804	47.5	1848	124.3	1892	73.0
1761	85.9	1805	42.2	1849	95.9	1893	84.9
1762	61.2	1806	28.1	1850	66.5	1894	78.0
1763	45.1	1807	10.1	1851	64.5	1895	64.0
1764	36.4	1808	8.1	1852	54.2	1896	41.8
1765	20.9	1809	2.5	1853	39.0	1897	26.2
1766	11.4	1810	0.0	1854	20.6	1898	26.7
1767	37.8	1811	1.4	1855	6.7	1899	12.1
1768	69.8	1812	5.0	1856	4.3	1900	9.5
1769	106.1	1813	12.2	1857	22.8	1901	2.7
1770	100.8	1814	13.9	1858	54.8	1902	5.0
1771	81.6	1815	35.4	1859	93.8	1903	24.4
1772	66.5	1816	45.8	1860	95.7	1904	42.0
1773	34.8	1817	41.1	1861	77.2	1905	63.5
1774	30.6	1818	30.4	1862	89.1	1906	53.8
1775	7.0	1819	23.9	1863	44.0	1907	62.0
1776	19.8	1820	15.7	1864	47.0	1908	48.5
1777	92.5	1821	6.6	1865	30.5	1909	43.9
1778	154.4	1822	4.0	1866	16.3	1910	18.6
1779	125.9	1823	1.8	1867	7.3	1911	5.7
1780	84.8	1824	8.5	1868	37.3	1912	3.6
1781	68.1	1825	16.6	1869	73.9	1913	1.4
1782	38.5	1826	36.3	1870	139.1	1914	9.6
1783	22.8	1827	49.7	1871	111.2	1915	47.4
1784	10.2	1828	62.5	1872	101.7	1916	57.1
1785	24.1	1829	67.0	1873	66.3	1917	103.9
1786	82.9	1830	71.0	1874	44.7	1918	80.6
1787	132.0	1831	47.8	1875	17.1	1919	63.6
1788	130.9	1832	27.5	1876	11.3	1920	37.6
1789	118.1	1833	8.5	1877	12.3	1921	26.1
1790	89.9	1834	13.2	1878	3.4	1922	14.2
1791	66.6	1835	56.9	1879	6.0	1923	5.8
1792	60.0	1836	121.5	1880	32.3	1924	16.7

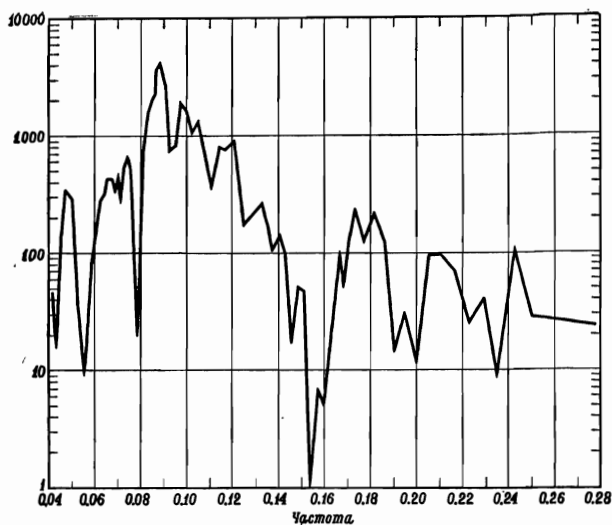


Рис. А.3.1.

Спектрограмма Шустера чисел солнечной активности.

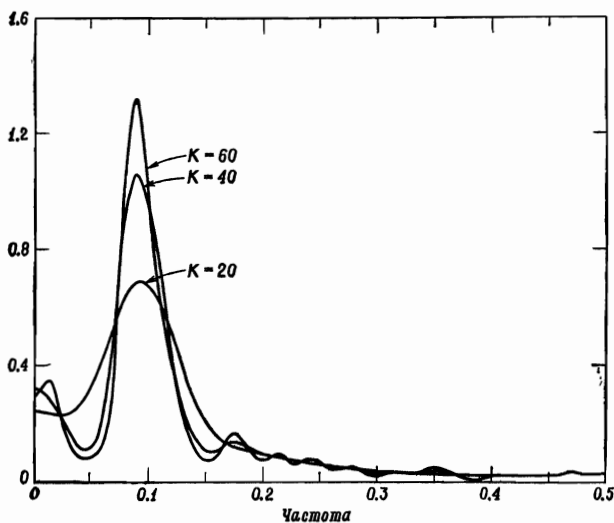


Рис. А.3.2.

Оценка спектральной плотности для чисел солнечной активности за период с 1749 по 1924 г.

Таблица А.3.2

СПЕКТРОГРАММА ШУСТЕРА ДЛЯ ЧИСЕЛ СОЛНЕЧНОЙ АКТИВНОСТИ

<i>Частота</i>	<i>Период в годах</i>	<i>Спектрограмма</i>	<i>Частота</i>	<i>Период в годах</i>	<i>Спектрограмма</i>
0.0417	24	48	0.1111	9	364
0.0435	23	17	0.1143	8.75	812
0.0455	22	112	0.1176	8.5	770
0.0476	21	349	0.1212	8.25	933
0.0500	20	298	0.1250	8	177
0.0526	19	35	0.1333	7.5	264
0.0556	18	9	0.1356	7.375	173
0.0588	17	83	0.1379	7.25	106
0.0625	16	278	0.1403	7.125	149
0.0645	15.5	340	0.1429	7	99
0.0658	15.25	436	0.1455	6.875	18
0.0667	15	434	0.1481	6.75	52
0.0680	14.75	432	0.1509	6.625	49
0.0690	14.5	342	0.1538	6.5	1
0.0704	14.25	474	0.1568	6.375	7
0.0714	14	278	0.1600	6.25	5
0.0727	13.75	550	0.1633	6.125	24
0.0741	13.5	696	0.1667	6	102
0.0755	13.25	552	0.1686	5.93	54
0.0769	13	198	0.1702	5.875	122
0.0784	12.75	21	0.1739	5.75	236
0.0800	12.5	105	0.1778	5.625	128
0.0816	12.25	675	0.1818	5.5	224
0.0833	12	1464	0.1860	5.375	129
0.0843	11.86	1951	0.1905	5.25	15
0.0851	11.75	2338	0.1951	5.125	31
0.0870	11.5	3700	0.2000	5	12
0.0889	11.25	4230	0.2051	4.875	97
0.0909	11	2724	0.2105	4.75	99
0.0930	10.75	742	0.2162	4.625	70
0.0952	10.5	853	0.2222	4.5	25
0.0976	10.25	2026	0.2286	4.375	41
0.1000	10	1677	0.2353	4.25	9
0.1026	9.75	1050	0.2424	4.125	109
0.1053	9.5	1313	0.2500	4	29
0.1081	9.25	603	0.2778	3.75	24

ЛИТЕРАТУРА

- А.1. Беверидж (1921), (1922), Гренджер и Харс (1969).
 А.2. Вольф (1965).
 А.3. Шерф (1964), Шустер (1906).

Приложение В

РЕШЕНИЯ ИЗБРАННЫХ УПРАЖНЕНИЙ

Глава 2, упражнение 7. Поскольку $\mathfrak{L}(y - Zb) = 0$, то

$$\begin{aligned}\mathfrak{L}(y - Zb)(b - \beta)' &= \mathfrak{L}(y - Zb)b' = \\ &= \mathfrak{L}(I - ZA^{-1}Z')yy'ZA^{-1} = \\ &= (I - ZA^{-1}Z')(\sigma^2I + Z\beta\beta'Z')ZA^{-1} = 0.\end{aligned}$$

Глава 2, упражнение 8. Уравнение $AB = I$ разбивается на два уравнения

$$A_{11}B_{12} + A_{12}B_{22} = 0,$$

$$A_{21}B_{12} + A_{22}B_{22} = I.$$

Из первого находим, что $B_{12} = -A_{11}^{-1}A_{12}B_{22}$ и отсюда получаем (а). Подстановка выражения для B_{12} во второе уравнение дает

$$(A_{22} - A_{21}A_{11}^{-1}A_{12})B_{22} = I.$$

Таким образом, матрица $A_{22} - A_{21}A_{11}^{-1}A_{12}$ невырождена и справедливо (б).

Глава 2, упражнение 10. (а) Пусть B — матрица коэффициентов регрессии $z_i^{(2)}$ на $z_i^{(1)}$. Тогда

$$\sum_{i=1}^T z_i^{(1)} z_i^{(2)'} = \sum_{i=1}^T z_i^{(1)} z_i^{(1)'} B', \quad B' = A_{11}^{-1} A_{12}, \quad B = A_{11}^{-1} A_{21},$$

$$\tilde{z}_i^{(2)} = z_i^{(2)} - Bz_i^{(1)} = z_i^{(2)} - A_{21}A_{11}^{-1}z_i^{(1)}.$$

Пусть $b^{*(1)}$ — вектор коэффициентов регрессии y_t на $z_t^{(1)}$. Тогда

$$b^{*(1)} = A_{11}^{-1} \sum_{i=1}^T z_i^{(1)} y_i,$$

$$\begin{aligned}
 \tilde{y}_t &= y_t - \mathbf{b}^{(1)'} \mathbf{z}_t^{(1)} = y_t - \sum_{\tau=1}^T y_{\tau} \mathbf{z}_{\tau}^{(1)'} \mathbf{A}_{11}^{-1} \mathbf{z}_t^{(1)}, \\
 \mathcal{Z} \tilde{y}_t &= \beta' \mathbf{z}_t - \sum_{\tau=1}^T \beta' \mathbf{z}_{\tau} \mathbf{z}_{\tau}^{(1)'} \mathbf{A}_{11}^{-1} \mathbf{z}_t^{(1)} = \\
 &= \beta' \mathbf{z}_t - \sum_{\tau=1}^T (\beta^{(1)'} \beta^{(2)'}) \begin{pmatrix} \mathbf{z}_{\tau}^{(1)} \\ \mathbf{z}_{\tau}^{(2)} \end{pmatrix} \mathbf{z}_{\tau}^{(1)'} \mathbf{A}_{11}^{-1} \mathbf{z}_t^{(1)} = \\
 &= \beta^{(1)'} \mathbf{z}_t^{(1)} + \beta^{(2)'} \mathbf{z}_t^{(2)} - \beta^{(1)'} \sum_{\tau=1}^T \mathbf{z}_{\tau}^{(1)} \mathbf{z}_{\tau}^{(1)'} \mathbf{A}_{11}^{-1} \mathbf{z}_t^{(1)} - \\
 &\quad - \beta^{(2)'} \sum_{\tau=1}^T \mathbf{z}_{\tau}^{(2)} \mathbf{z}_{\tau}^{(1)'} \mathbf{A}_{11}^{-1} \mathbf{z}_t^{(1)} = \\
 &= \beta^{(2)'} (\mathbf{z}_t^{(2)} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{z}_t^{(1)}) = \beta^{(2)'} \tilde{\mathbf{z}}_t^{(2)}.
 \end{aligned}$$

(b) Нормальные уравнения для $\tilde{\mathbf{b}}^{(2)}$ имеют вид

$$(i) \quad \sum_{t=1}^T \tilde{\mathbf{z}}_t^{(2)} \tilde{y}_t = \sum_{t=1}^T \tilde{\mathbf{z}}_t^{(2)} \tilde{\mathbf{z}}_t^{(2)'} \tilde{\mathbf{b}}^{(2)}.$$

Подстановка в (i) выражения для $\tilde{\mathbf{z}}_t^{(2)}$ дает

$$\begin{aligned}
 \sum_{t=1}^T (\mathbf{z}_t^{(2)} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{z}_t^{(1)}) \left(y_t - \sum_{\tau=1}^T y_{\tau} \mathbf{z}_{\tau}^{(1)'} \mathbf{A}_{11}^{-1} \mathbf{z}_t^{(1)} \right) &= \\
 &= \sum_{t=1}^T (\mathbf{z}_t^{(2)} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{z}_t^{(1)}) (\mathbf{z}_t^{(2)'} - \mathbf{z}_t^{(1)'} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}) \tilde{\mathbf{b}}^{(2)}.
 \end{aligned}$$

Последнее соотношение можно переписать в виде

$$(ii) \quad \sum_{t=1}^T \mathbf{z}_t^{(2)} y_t - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \sum_{t=1}^T \mathbf{z}_t^{(1)} y_t = (\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}) \tilde{\mathbf{b}}^{(2)}.$$

В то же время нормальные уравнения для $\mathbf{b}^{(1)}$ и $\mathbf{b}^{(2)}$ имеют вид

$$\begin{aligned}
 \mathbf{A}_{11} \mathbf{b}^{(1)} + \mathbf{A}_{12} \mathbf{b}^{(2)} &= \sum_{t=1}^T \mathbf{z}_t^{(1)} y_t, \\
 \mathbf{A}_{21} \mathbf{b}^{(1)} + \mathbf{A}_{22} \mathbf{b}^{(2)} &= \sum_{t=1}^T \mathbf{z}_t^{(2)} y_t.
 \end{aligned}$$

Первое из них дает

$$\mathbf{b}^{(1)} = -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{b}^{(2)} + \mathbf{A}_{11}^{-1} \sum_{t=1}^T \mathbf{z}_t^{(1)} y_t.$$

Подстановка последнего выражения во второе уравнение приводит его к виду

$$-\mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{b}^{(2)} + \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \sum_{t=1}^T \mathbf{z}_t^{(1)} y_t + \mathbf{A}_{22} \mathbf{b}^{(2)} = \sum_{t=1}^T \mathbf{z}_t^{(2)} y_t,$$

совпадающему с (ii).

Глава 3, упражнение 5. Выражение, содержащееся в указании, равно

$$\sum_{t=1}^T \left[\sum_{k=0}^{p+1} \binom{p+1}{k} t^k - t^{p+1} \right] = \sum_{k=0}^p \binom{p+1}{k} \sum_{t=1}^T t^k,$$

а первая сумма в указании равна $\sum_{t=2}^{T+1} t^{p+1}$.

Глава 3, упражнение 7. В соответствии с упр. 5,

$$(k+1) \psi_k(T) = (T+1)^{k+1} - 1 - \sum_{r=0}^{k-1} \binom{k+1}{r} \psi_r(T),$$

где $\psi_0(T) = T$. Из этого соотношения поочередно находим $\psi_1(T)$, ..., $\psi_8(T)$.

Глава 3, упражнение 30. Условия ортогональности $\varphi^*(s)$ приводят к $r+1$ однородным уравнениям

$$c_0 \sum_{s=-m}^m s^{2j} + c_2 \sum_{s=-m}^m s^{2j+2} + \dots + c_{2r} \sum_{s=-m}^m s^{2j+2r} = 0, \quad j = 0, 1, \dots, r.$$

Матрица коэффициентов имеет поэтому вид

$$\sum_{s=-m}^m (1, s^2, \dots, s^{2r})' (1, s^2, \dots, s^{2r}),$$

и из соотношения

$$\sum_{s=-m}^m \left(\sum_{i=0}^r x_i s^{2i} \right)^2 = 0$$

вытекает, что $x_0 = 0$ и $\sum_{i=1}^r x_i s^{2i} = 0$, $s = 1, \dots, m$. Если $m \geq r$, то существует единственное решение $x_0 = x_1 = \dots = 0$ и матрица коэффициентов положительно определена. Таким образом, она будет невырожденной и $\varphi^*(s) = c_1 s + c_3 s^3 + \dots + c_{2r-1} s^{2r-1} + s^{2r+1}$. Из предыдущего вытекает также, что $\varphi_{i,2m+1}^*(s)$ при четном i содержит только четные степени s .

Глава 3, упражнение 38. В силу (2), (3) и (7) для любых действительных чисел d_1 и d_2 и линейных операторов $\mathcal{O}_1$ и $\mathcal{O}_2$

$$\begin{aligned} (c_1 \mathcal{O}_1 + c_2 \mathcal{O}_2)(d_1 u_t + d_2 v_t) &= c_1 \mathcal{O}_1(d_1 u_t + d_2 v_t) + c_2 \mathcal{O}_2(d_1 u_t + d_2 v_t) = \\ &= c_1(d_1 \mathcal{O}_1 u_t + d_2 \mathcal{O}_1 v_t) + c_2(d_1 \mathcal{O}_2 u_t + d_2 \mathcal{O}_2 v_t) = \\ &= d_1(c_1 \mathcal{O}_1 u_t + c_2 \mathcal{O}_2 u_t) + d_2(c_1 \mathcal{O}_1 v_t + c_2 \mathcal{O}_2 v_t) = \\ &= d_1(c_1 \mathcal{O}_1 + c_2 \mathcal{O}_2) u_t + d_2(c_1 \mathcal{O}_1 + c_2 \mathcal{O}_2) v_t. \end{aligned}$$

Глава 3, упражнение 52. В силу леммы 3.4.2 $\Delta^{q+1} f(t) = 0$, $t = 1, \dots, T - q - 1$, поскольку $\Delta^{q+1} f(t)$ для каждого t определяется значениями $f(t)$ в $q+2$ точках. Поскольку тот же оператор аннулирует тренд на всем рассматриваемом интервале, то тренд является полиномиальным степени не выше q в силу задачи 51.

Глава 3, упражнение 57. Имеем

$$\sum_{i,j,k,l=1}^n a_{ij} b_{kl} \xi_{iij} \mu_{kl} = \kappa_4 \sum_{i=1}^n a_{ii} b_{ii} + \sigma^4 \sum_{i=1}^n a_{ii} \sum_{i=1}^n b_{ii} + 2\sigma^4 \sum_{i,j=1}^n a_{ij} b_{ij}.$$

Далее следует использовать соотношение (39). См. также упр. 7 гл. 8.

Глава 3, упражнение 63. В силу (39) $\xi Q_1 / \xi Q_2 = \text{tr} A / \text{tr} B$. Дисперсия равна

$$\text{Var} \sum_{s,t=1}^T \left(a_{st} - \frac{\text{tr} A}{\text{tr} B} b_{st} \right) u_{st} u_t,$$

и искомый результат получается непосредственным применением леммы 3.4.4.

Глава 3, упражнение 64. Случайная величина

$$\begin{aligned} & \sqrt{T} \left[\frac{u' A_T u}{u' B_T u} - \frac{\alpha}{\beta} \right] = \\ & = \sqrt{T} \left[\frac{\alpha \sigma^2 + (u' A_T u - \alpha \sigma^2)}{\beta \sigma^2 + (u' B_T u - \beta \sigma^2)} - \frac{\alpha}{\beta} \right] = \sqrt{T} \left[\frac{\frac{\alpha}{\beta} + \frac{u' A_T u - \alpha \sigma^2}{\beta \sigma^2}}{1 + \frac{u' B_T u - \beta \sigma^2}{\beta \sigma^2}} - \frac{\alpha}{\beta} \right] \end{aligned}$$

имеет предельное распределение, совпадающее с предельным распределением случайной величины

$$\sqrt{T} \left\{ \left[\frac{\alpha}{\beta} + \frac{u' A_T u - \alpha \sigma^2}{\beta \sigma^2} \right] \left[1 - \frac{u' B_T u - \beta \sigma^2}{\beta \sigma^2} \right] - \frac{\alpha}{\beta} \right\},$$

которое, в свою очередь, совпадает с предельным распределением величины

$$\sqrt{T} \left\{ \frac{u' A_T u - \alpha \sigma^2}{\beta \sigma^2} - \frac{\alpha}{\beta} \frac{u' B_T u - \beta \sigma^2}{\beta \sigma^2} \right\} = \frac{\sqrt{T} [u' A_T u - (\alpha/\beta) u' B_T u]}{\beta \sigma^2}.$$

Отсюда вытекает, что рассматриваемая в задаче статистика имеет в пределе нормальное распределение с нулевым средним и дисперсией

$$\frac{1}{\beta^2 \sigma^4} \lim_{T \rightarrow \infty} T \text{Var} \left(u' A_T u - \frac{\alpha}{\beta} u' B_T u \right).$$

Результат, содержащийся в указании, получается непосредственным применением упр. 63 для каждого фиксированного T . Если $a_{11}^{(T)} / b_{11}^{(T)} = \dots = a_{TT}^{(T)} / b_{TT}^{(T)}$, то $a_{ii}^{(T)} = k b_{ii}^{(T)}$, $i = 1, \dots, T$, для некоторого k , и $\alpha = k\beta$. Тогда

$$a_{ii}^{(T)} - \frac{\alpha}{\beta} b_{ii}^{(T)} = \left(k - \frac{k\beta}{\beta} \right) b_{ii}^{(T)} = 0, \quad i = 1, \dots, T,$$

для каждого T . Относительно более общего результата см. упр. 34 гл. 8.

Глава 4, упражнение 6. Элементы искомой матрицы P можно найти из равенства $N = MP$, используя соотношение $e^{i\lambda} = \cos \lambda + i \sin \lambda$. Более прямой метод состоит в использовании соотношений $\cos(2\pi - \lambda) = \cos \lambda$ и $\sin(2\pi - \lambda) = -\sin \lambda$ для получения результатов типа (9) — (13) в несколько

инном варианте. Например, что касается соотношений (9), то

$$\sum_{t=1}^T \cos \frac{2\pi j}{T} t \cos \frac{2\pi (T-k)}{T} t = \sum_{t=1}^T \cos \frac{2\pi j}{T} t \cos \frac{2\pi k}{T} t, \quad 0 \leq j, \quad k \leq [T/2].$$

Далее матрицу P находят непосредственным перемножением, $P = M'N$, и использованием полученных вариантов соотношений (9) — (13). Если T четное, то $p_{1T} = 1$, $p_{2t,t} = p_{2t,T-t} = 1/\sqrt{2}$, $p_{2t+1,t} = i/\sqrt{2}$, $p_{2t+1,T-t} = -i/\sqrt{2}$, $t = 1, \dots, T/2-1$, $p_{T,T/2} = 1$, а все остальные элементы P равны нулю. Если же T нечетное, то $p_{1T} = 1$, $p_{2t,t} = p_{2t,T-t} = 1/\sqrt{2}$, $p_{2t+1,t} = i/\sqrt{2}$, $p_{2t+1,T-t} = -i/\sqrt{2}$, $t = 1, \dots, (T-1)/2$, а все остальные элементы равны нулю.

Глава 4, упражнение 18. Используя соотношения (10) — (13) и (33), можно записать:

$$\begin{aligned} a_0 &= \frac{1}{n} \sum_{t=1}^n \bar{y}_t = \bar{y}, \\ a(kh) &= \frac{2}{n} \sum_{t=1}^n \bar{y}_t \cos \frac{2\pi k}{n} t, \\ b(kh) &= \frac{2}{n} \sum_{t=1}^n \bar{y}_t \sin \frac{2\pi k}{n} t, \quad k = 1, \dots, n/2 - 1, \\ a_{T/2} &= \frac{1}{n} \sum_{t=1}^n \bar{y}_t (-1)^t. \end{aligned}$$

Для сравнения можно взять соотношения (17) и (18), соответствующие $n = 12$, $h = 3$. Положим

$$\begin{aligned} \mathbf{x}^* &= \sqrt{\frac{n}{2}} \left(\sqrt{2} a_0, a(1h), b(1h), a(2h), \dots, b\left(\frac{n-2}{2}h\right), \sqrt{2} a_{T/2} \right)', \\ \mathbf{y}^* &= (\bar{y}_1, \dots, \bar{y}_n)'. \end{aligned}$$

В силу соотношений (26) и (27) § 4.2 $\mathbf{y}^* = M^* \mathbf{x}^*$, где матрица M^* ортогональна, и

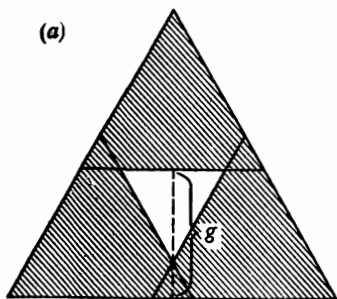
$$h \sum_{t=1}^n (\bar{y}_t - \bar{y})^2 = h [\mathbf{y}^{*'} \mathbf{y}^* - n \bar{y}^2] = h [\mathbf{x}^{*'} \mathbf{x}^* - n a_0^2].$$

Последнее же является числителем в (30). Наконец, заметим, что в силу (15) и отмеченного выше

$$\begin{aligned} (T-n)s^2 &= \sum_{t=1}^T y_t^2 - T \bar{y}^2 - h \sum_{t=1}^n (\bar{y}_t - \bar{y})^2 \\ &= \sum_{t=1}^n \sum_{j=0}^{h-1} y_{t+nj}^2 - h \sum_{t=1}^n \bar{y}_t^2 = \\ &= \sum_{t=1}^n \sum_{j=0}^{h-1} (y_{t+nj} - \bar{y}_t)^2. \end{aligned}$$

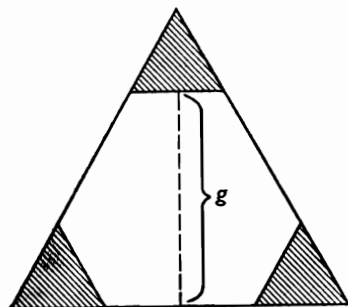
Глава 4, упражнение 27.

(a)



$$\max x_i \geq g$$

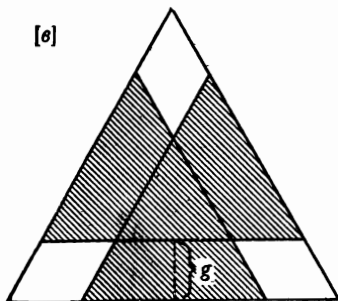
$$1/3 \leq g \leq 1/2$$



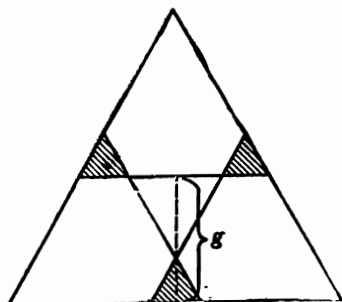
$$\max x_i \geq g$$

$$1/3 \leq g \leq 1$$

[e]



Второе по величине значение $x_i \geq g$

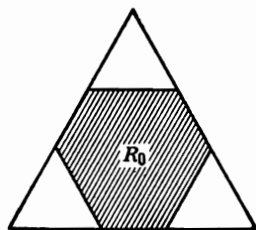
$$0 \leq g \leq 1/3$$


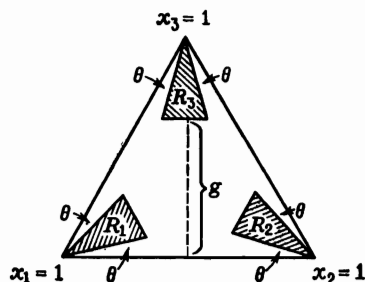
Второе по величине значение $x_i \geq g$

$$1/3 \leq g \leq 1/2$$

Глава 4, упражнение 28.

$$(i) R_0 : x_i \leq g, \quad i = 1, 2, 3$$



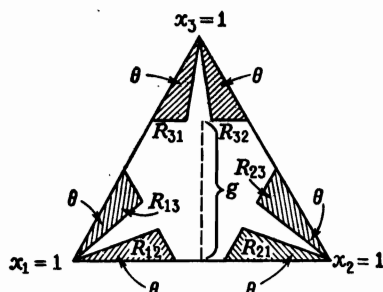


$$(ii) R_i: x_i > g, x_j \leq g^* x_k,$$

i, j, k различные

$$g^* \sin \theta = \sin (\pi/3 - \theta)$$

$$\theta = \arctg \frac{\sqrt{3}}{2g^* + 1}, g^* \geq 1$$



$$(iii) R_{ij}: x_i > g, x_j > g^* x_k,$$

i, j, k различные

$$\theta = \arctg \frac{\sqrt{3}}{2g^* + 1}, g^* \geq 1$$

Процедура, предложенная Уиттлом, состоит в том, что принимается решение $\rho_1 = \rho_2 = \rho_3 = 0$, если $x_i \leq g$, $i = 1, 2, 3$. Это есть область R_0 . Решение $\rho_1 > 0$ принимается, если $x_1 > g$. Для $g \geq 1/2$ это есть область $R_1 \cup R_{12} \cup R_{13}$. В этом случае решение $\rho_2 = \rho_3 = 0$ принимается, если $x_2 \leq \bar{g}(x_2 + x_3)$ и $x_3 \leq \bar{g}(x_2 + x_3)$ или, что равносильно, если $x_2 \leq g^* x_3$ и $x_3 \leq g^* x_2$. Это есть область R_1 . Решение приводится только для $g \geq 1/2$ из соображений большей наглядности. Для $1/3 < g < 1/2$ решение $\rho_1 > 0$ принимается, если $x_1 > g$, $x_1 > x_2$ и $x_1 > x_3$.

Глава 5, упражнение 6. Левая часть доказываемого соотношения равна

$$P_q(t+1)a^{t+1} - P_q(t)a^{t+1} = [P_q(t+1) - P_q(t)]a^{t+1} = P_{q-1}(t)a^{t+1}$$

в силу представления (13) § 3.4.

Глава 5, упражнение 7. Доказательство проводится по индукции. Для $m = 1$ $(\mathcal{P} - a)a^t = 0$. Предположим, что результат верен для $m = k$. Тогда

$$(\mathcal{P} - a)^{k+1}P_k(t)a^t = (\mathcal{P} - a)^k(\mathcal{P} - a)P_k(t)a^t = (\mathcal{P} - a)^kP_{k-1}(t)a^{t+1} = 0$$

в силу упр. 6 и предположения индукции.

Глава 5, упражнение 9. Полином степени p от $\mathcal{P}$ можно записать в виде

$$\sum_{r=0}^p \beta_r \mathcal{P}^{p-r} = P_{p-m}(\mathcal{P})(\mathcal{P} - a)^m,$$

где $P_{p-m}(t)$ — полином степени $p-m$. Тогда

$$\begin{aligned}\sum_{r=0}^p \beta_r w_{t-r} &= \sum_{r=0}^p \beta_r \mathcal{P}^{p-r} w_{t-p} = \\ &= P_{p-m}(\mathcal{P}) (\mathcal{P}-a)^m P_{m-1}(t-p) a^{t-p} = 0,\end{aligned}$$

в соответствии с результатом упр. 7.

Глава 5, упражнение 10. Используем метод решения упр. 9. Тогда, применяя дважды результат упр. 7, получаем

$$\begin{aligned}\sum_{r=0}^p \beta_r w_{t-r} &= P_{p-2m}(\mathcal{P}) (\mathcal{P}-\alpha e^{i\theta})^m (\mathcal{P}-\alpha e^{-i\theta})^m \times \\ &\times [P_{m-1}(t) \alpha^t e^{it\theta} + \bar{P}_{m-1}(t) \alpha^t e^{-it\theta}] = 0.\end{aligned}$$

Отметим, что решением является любое выражение вида $P_{m-1}(t) \alpha^t e^{it\theta} + Q_{m-1}(t) \alpha^t e^{-it\theta}$, где $Q_{m-1}(t)$ — полином степени $m-1$. Однако действительным такое решение будет только, когда $Q_{m-1}(t) = \bar{P}_{m-1}(t)$.

Глава 5, упражнение 17.

$$\Lambda^\tau = \begin{bmatrix} \lambda^\tau & \binom{\tau}{1} \lambda^{\tau-1} & \binom{\tau}{2} \lambda^{\tau-2} & \binom{\tau}{3} \lambda^{\tau-3} & \dots & \binom{\tau}{n-2} \lambda^{\tau-n+2} & \binom{\tau}{n-1} \lambda^{\tau-n+1} \\ 0 & \lambda^\tau & \binom{\tau}{1} \lambda^{\tau-1} & \binom{\tau}{2} \lambda^{\tau-2} & \dots & \binom{\tau}{n-3} \lambda^{\tau-n+3} & \binom{\tau}{n-2} \lambda^{\tau-n+2} \\ 0 & 0 & \lambda^\tau & \binom{\tau}{1} \lambda^{\tau-1} & \dots & \binom{\tau}{n-4} \lambda^{\tau-n+4} & \binom{\tau}{n-3} \lambda^{\tau-n+3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda^\tau & \binom{\tau}{1} \lambda^{\tau-1} \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda^\tau \end{bmatrix},$$

где n — порядок матрицы и $\binom{\tau}{j} = 0$ для $j > \tau$. В справедливости этого результата можно убедиться по индукции.

Глава 5, упражнение 18. Характеристическое уравнение $|\Lambda - vI| = 0$ имеет здесь вид $(\lambda - v)^n = 0$, где n — порядок матрицы Λ . Уравнение $\Lambda x = \lambda x$, $x = (x_1, \dots, x_n)'$, имеет решение $x = (x, 0, \dots, 0)'$, которое является единственным (если не считать различий в первой компоненте для $x \neq 0$).

Глава 5, упражнение 19. В покомпонентной записи уравнение $-\tilde{B}v = x_i v$, $v = (v_1, \dots, v_p)'$, имеет вид $-\sum_{r=1}^p \beta_r v_r = x_i v_i$, $v_{r-1} = x_i v_r$, $r = 2, \dots, p$. Если $v_i = x_i^{p-1}$, то единственным решением будет $v_r = x_i^{p-r}$, $r = 1, \dots, p$. Иначе говоря, каждому отдельному корню соответствует единственный характеристический вектор. Если $-\tilde{B} = S\Lambda S^{-1}$ и при этом матрица Λ диагональна, а матрица S не вырождена, то $-\tilde{B}S = S\Lambda$ и матрица $-\tilde{B}$ имеет p различных характеристических векторов. Следовательно, она имеет p различных характеристических корней.

Глава 5, упражнение 20. Сумма $\sum_{j=1}^p c_{1j} x_j^T c_{1j}$ в (22) является первым ди-

агональным элементом матрицы $C\Lambda^T C^{-1} = (-\tilde{B})^T$. Для решения задачи необходимо показать, что эта величина совпадает с δ_T из (28) § 5.2. Обозначим первую строку матрицы $(-\tilde{B})^{T+1}$ через $(\alpha_{T1}, \alpha_{T2}, \dots, \alpha_{T,p-1}, \alpha_{Tp})$. Тогда первой строкой матрицы $(-\tilde{B})^{T+2} = (-\tilde{B})^{T+1}(-\tilde{B})$ будет $(\alpha_{T2} - \alpha_{T1}\beta_1, \alpha_{T3} - \alpha_{T1}\beta_2, \dots, \alpha_{Tp} - \alpha_{T1}\beta_{p-1}, -\alpha_{T1}\beta_p)$. Таким образом, соответствующие величины удовлетворяют тем же самым рекуррентным соотношениям, что и величины δ_T . [См. (14) или (22) из § 5.2.] Кроме того, они удовлетворяют тем же самым начальным условиям.

Глава 5, упражнение 27. Следуя указанию, умножим обе части (31) слева на C^{-1} и справа на $(C')^{-1}$. В силу (10) $C^{-1}BC = -\Lambda$, так что $A^* - \Lambda A^* \Lambda = \Sigma^*$. Если $A^* = (a_{ij}^*)$, $\Sigma^* = (\sigma_{ij}^*)$ и диагональные элементы матрицы Λ равны $\lambda_1, \dots, \lambda_p$, то $a_{ij}^* (1 - \lambda_i \lambda_j) = \sigma_{ij}^*$, $i, j = 1, \dots, p$. Поскольку характеристические корни матрицы $-B$ лежат в единичном круге, то A^* определяется однозначно и $A = \Sigma C A^* C'$ определяется из (31) единственным образом.

Глава 5, упражнение 40.

(а) График совокупного дохода, остающегося после уплаты налогов (в сопоставимых ценах) приведен на стр. 727.

(б) Средние y_t, y_{t-1} и y_{t-2} , $t = 1, \dots, 21$, равны 59.0980, 57.2485 и 55.8937 соответственно. Нормальные уравнения относительно $\hat{\gamma}_2, \hat{\beta}_1$ и $\hat{\beta}_2$ представляются в виде

$$\begin{pmatrix} 770 & 681 & 639 \\ 681 & 1287 & 1050 \\ 639 & 1050 & 1175 \end{pmatrix} \begin{pmatrix} \hat{\gamma}_2 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = - \begin{pmatrix} 821 \\ 1316 \\ 946 \end{pmatrix}.$$

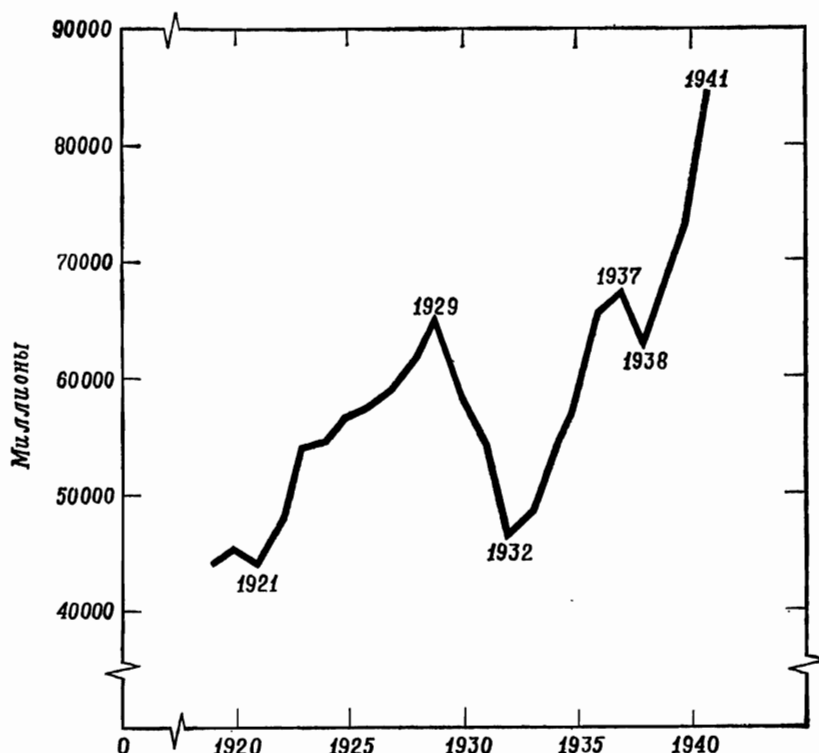
Сумма квадратов значений y_t , взятых относительно среднего, равна 1807.38. Соответствующие оценки равны $\hat{\beta}_1 = -1.223$, $\hat{\beta}_2 = 0.5103$, $\hat{\gamma}_1 = -13.13$, $\hat{\gamma}_2 = -0.4095$, $\hat{\sigma} = 4.045$. (Оценка параметра σ , деленная на «число степеней свободы» $21 - 4 = 17$, равна 4.496.) Оценкой ковариационной матрицы для $\hat{\gamma}_2, \hat{\beta}_1$ и $\hat{\beta}_2$ является умноженная на $\hat{\sigma}^2$ (или s^2) и на 10^{-4} матрица

$$\begin{pmatrix} 25.79 & -8.11 & -6.78 \\ -8.11 & 31.25 & -23.52 \\ -6.78 & -23.52 & 33.22 \end{pmatrix}.$$

Оценки стандартных отклонений оцениваемых коэффициентов равны при использовании s соответственно 0.228, 0.251 и 0.259.

(с) $\hat{\mu} = 44.7$, $\hat{\delta} = 1.425$.

(д) Характеристическое уравнение, соответствующее данному стохастическому разностному уравнению, имеет корни $0.6115 \pm 0.3693 i$. Аргумент равен $31^\circ 8'$ и соответствует частоте $31^\circ 8' / 360^\circ = 0.08648$, или периоду 11.57 лет, а модуль равен 0.7144.



Обсуждение. Модель авторегрессии второго порядка для дохода можно получить из следующей экономической модели:

$$Y_t = C_t + I_t + G_t,$$

$$C_t = \alpha_0 + \alpha Y_{t-1} + \alpha^* t,$$

$$I_t = \beta_0 + \beta [C_t + G_t - (C_{t-1} + G_{t-1})] + \beta^* t,$$

$$G_t = \gamma_0 + \gamma^* t,$$

где Y_t , C_t , I_t и G_t — соответственно доход, потребление, капиталовложения и правительственные расходы за t -й год. При этом $\beta_1 = -\alpha(1 + \beta)$ и $\beta_2 = \alpha\beta$. Оценка для α равна 0.713. Она определяет часть дохода, предназначенную для потребления. Оценка для β равна 0.716. Она представляет компоненту капиталовложений, выявляемую при изменении неинвестиционных расходов. Малое значение последнего коэффициента можно объяснить недостаточной уверенностью деловых кругов, связанной со спадом деловой активности на протяжении большей части периода наблюдений.

Глава 6, упражнение 13. Подстановка (16) в (20) дает (i), что доказывает утверждение указания. Из (i) и (ii) вытекает

$$(iii) \quad \int_0^{\infty} (Q_i - \mu) h_i(Q_i | Q_0, \dots, Q_{i-1}; \gamma_i^{(1)}) dQ_i = 0.$$

В силу (16) данное условное распределение имеет положительную массу в интервалах $(-\infty, c)$ и (c, ∞) . В противном случае $\varepsilon_i = 0$ или 1. Тогда приходим к противоречию: из (i) получаем, что $\mu < c$, а из (iii) — что $\mu > c$.

Глава 6, упражнение 27. (а) Вектор $y = (y_1, y_2, y_3, y_4)'$ имеет нормальное распределение с вектором средних 0 и ковариационной матрицей Σ , где

$$\Sigma^{-1} = \begin{pmatrix} \gamma_0 + \gamma_1 & 0 & 0 & 0 \\ 0 & \gamma_0 + \gamma_1 & 0 & 0 \\ 0 & 0 & \gamma_0 - \gamma_1 & 0 \\ 0 & 0 & 0 & \gamma_0 - \gamma_1 \end{pmatrix}.$$

Тогда $Y_1 = (\gamma_0 + \gamma_1)(y_1^2 + y_2^2)$ и $Y_2 = (\gamma_0 - \gamma_1)(y_3^2 + y_4^2)$ — независимые случайные величины, каждая из которых имеет распределение χ^2 с 2 степенями свободы, так что плотность совместного распределения величин Y_1 и Y_2 равна

$$\frac{1}{4} e^{-(Y_1 + Y_2)/2}, \quad Y_1, Y_2 \geq 0.$$

Якобиан преобразования

$$Y_1 = (\gamma_0 + \gamma_1)(Q_0 + Q_1)/2$$

$$Y_2 = (\gamma_0 - \gamma_1)(Q_0 - Q_1)/2$$

равен $(\gamma_0^2 - \gamma_1^2)/2$ и плотность совместного распределения величин Q_0 и Q_1 равна (ii).

(b) Якобиан преобразования $Q_0 = Q_0$, $r = Q_1/Q_0$ равен Q_0 . Из (а) вытекает тогда, что плотностью совместного распределения величин Q_0 и r является (iii).

(с) Маргинальная плотность величины Q_0 равна интегралу от (iii) по множеству всех возможных значений r , т. е. равна

$$\begin{aligned} \frac{\gamma_0^2 - \gamma_1^2}{8} Q_0 e^{-\gamma_0 Q_0/2} \int_{-1}^1 e^{-\gamma_1 Q_0 r/2} dr = \\ = \frac{\gamma_0^2 - \gamma_1^2}{8} Q_0 e^{-\gamma_0 Q_0/2} \frac{2}{\gamma_1 Q_0} (e^{\gamma_1 Q_0/2} - e^{-\gamma_1 Q_0/2}), \end{aligned}$$

что совпадает с (iv). Ее можно также записать в виде

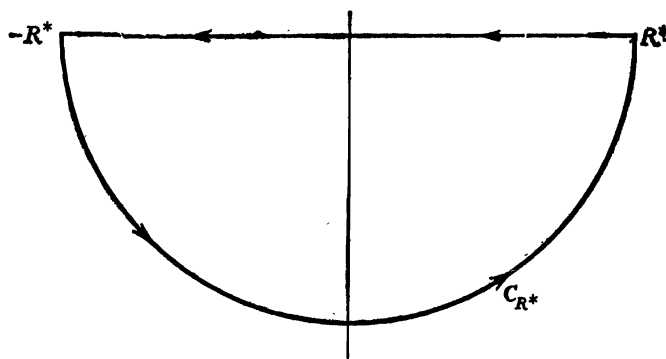
$$\frac{\gamma_0^2 - \gamma_1^2}{2\gamma_1} e^{-\gamma_0 Q_0/2} \operatorname{sh} \gamma_1 Q_0/2, \quad 0 \leq Q_0.$$

(d) Условная плотность величины r при заданном Q_0 равна отношению (iii) и (iv).

(е) Маргинальная плотность величины r является интегралом от (iii) по множеству всех значений Q_0 , именно

$$\begin{aligned} \frac{\gamma_0^2 - \gamma_1^2}{8(\gamma_0 + \gamma_1 r)^2} \int_0^\infty [(\gamma_0 + \gamma_1 r) Q_0]^{4/2-1} e^{-(\gamma_0 + \gamma_1 r) Q_0} [(\gamma_0 + \gamma_1 r) dQ_0] = \\ = \frac{\gamma_0^2 - \gamma_1^2}{8(\gamma_0 + \gamma_1 r)^2} 2^{4/2} \Gamma(4/2) = \frac{\gamma_0^2 - \gamma_1^2}{2(\gamma_0 + \gamma_1 r)^2}. \end{aligned}$$

Этот интеграл содержит χ^2 -плотность с 4 степенями свободы. Отметим, что $\gamma_0 + \gamma_1 r > 0$, поскольку $|\gamma_1| < \gamma_0$ и $|r| \leq 1$.



Глава 6, упражнение 37. Характеристическая функция случайной величины u , приведенная в (69), имеет только простые полюсы. Предположим, что $v_{m+1} \leq R \leq v_m$ и $H > 2$. Плотность величины u равна тогда

$$f(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iut} \left(\prod_{j=1}^H [1 - 2it(v_j - R)] \right)^{-1} dt.$$

Ее можно найти интегрированием вдоль замкнутого контура, содержащего все полюсы, расположенные в нижней полуплоскости. Например, можно взять следующий контур (см. рис.).

Радиус R^* выбирается здесь достаточно большим, с тем чтобы внутри контура оказались все полюсы, расположенные в нижней половине комплексной плоскости. Тогда по теореме о вычетах выражение

$$\begin{aligned} & \frac{1}{2\pi} \int_{R^*}^{-R^*} e^{-iut} \left(\prod_{j=1}^H [1 - 2it(v_j - R)] \right)^{-1} dt + \\ & + \frac{1}{2\pi} \int_{C_{R^*}} e^{-iuz} \left(\prod_{j=1}^H [1 - 2iz(v_j - R)] \right)^{-1} dz \end{aligned}$$

равно умноженной на $2\pi i$ сумме вычетов подинтегрального выражения относительно полюсов, расположенных внутри указанного контура. Поскольку $v_{m+1} \leq R \leq v_m$, то эти полюсы расположены в точках $z = (-i/2)/(v_j - R)$, $j = 1, \dots, m$. (Если $R = v_m$, то имеется $m - 1$ полюсов, так как при этом не будет полюса при $j = m$.) Имеем

$$[1 - 2iz(v_j - R)]^{-1} = \frac{\frac{i}{2(v_j - R)}}{z + \frac{i}{2(v_j - R)}}.$$

Поэтому полюс в точке $-z = -(i/2)/(v_j - R)$ равен

$$\frac{i}{4\pi} e^{-u/[2(v_j - R)]} \frac{(v_j - R)^{H-2}}{\prod_{\substack{k=1 \\ k \neq j}}^H (v_j - v_k)}$$

и умноженная на $2\pi i$ сумма вычетов равна

$$(i) \quad -\frac{1}{2} \sum_{j=1}^m e^{-u/2(v_j-R)} \frac{(v_j-R)^{H-2}}{\prod_{\substack{k=1 \\ k \neq j}}^H (v_j-v_k)}, \quad v_{m+1} \leq R \leq v_m, \\ m=1, \dots, H-1.$$

Тогда $-f(u)$ для $u \geq 0$ равно (i) минус предел интеграла вдоль C_{R^*} при $R^* \rightarrow \infty$. Последний интеграл стремится к нулю, поскольку знаменатель подынтегрального выражения имеет порядок $O(R^{*H})$, числитель по абсолютной величине не превосходит 1 и длина дуги C_{R^*} равна πR^* . Наконец, $\text{Pr}\{r \geq R\} = \int_0^\infty f(u) du =$

$$= \frac{1}{2} \sum_{j=1}^m \frac{(v_j-R)^{H-2}}{\prod_{\substack{k=1 \\ k \neq j}}^H (v_j-v_k)} \cdot 2(v_j-R), \quad v_{m+1} \leq R \leq v_m, \\ m=1, \dots, H-1,$$

а это и есть (52). Чтобы получить $f(u)$ для $u < 0$, необходимо провести интегрирование вдоль замкнутого контура, охватывающего полюсы в верхней полуплоскости.

Глава 6, упражнение 69. Производная левой части равна

$$2 \int_0^\pi \frac{dx}{a+b \cos x}.$$

Следуя указанию, получаем

$$\begin{aligned} 2 \int_0^\pi \frac{dx}{(a+b) \cos^2(x/2) + (a-b) \sin^2(x/2)} &= \\ = 2 \int_0^\pi \frac{\sec^2(x/2) dx}{(a+b) \left[1 + \frac{a-b}{a+b} \tan^2(x/2) \right]} &= \frac{4}{\sqrt{a^2-b^2}} \int_0^\infty \frac{dy}{1+y^2} = \\ &= \frac{2\pi}{\sqrt{a^2-b^2}}. \end{aligned}$$

Производная правой части имеет то же самое выражение. Мы показали в сущности, что функция

$$(i) \quad C(a, b) = \int_0^\pi \log(a+b \cos x) dx - \pi \log \frac{a + \sqrt{a^2-b^2}}{2}, \quad a > b \geq 0,$$

не зависит от первого аргумента. Положим $a = kb > 0$, $k > 1$. Тогда $C(kb, b) = C^*(b)$ и $C^*(b) = \int_0^\pi \log(k + \cos x) dx - \pi \log(k + \sqrt{k^2-1})/2$, так что $C(a, b) \equiv C^*$. При $b \rightarrow 0$ левая часть (i) остается, таким образом, постоянной, тогда как правая часть стремится к нулю. Следовательно, $C(a, b) \equiv 0$.

Глава 7, упражнение 28. Существование искомых процессов скользящего среднего и авторегрессии обеспечивается соответственно следствиями 7.5.1 и 7.5.2. В каждом из этих случаев

$$|\sigma(h) - \sigma_m(h)| = \left| \int_{-\pi}^{\pi} e^{i\lambda h} [f(\lambda) - f_m(\lambda)] d\lambda \right| < \varepsilon \int_{-\pi}^{\pi} d\lambda.$$

Глава 7, упражнение 29. Процессы $\{y_t\}$ и $\{w_t\}$ независимы. Поэтому в силу результата упр. 9 спектральная плотность процесса $\{z_t\}$ равна

$$\frac{\sigma^2}{2\pi |e^{i\lambda} + \beta|^2} + \frac{\tau^2}{2\pi}.$$

Спектральная плотность процесса $\{z_t^*\}$ выводится из (51) и требований равенства двух плотностей, которые можно записать в виде $K\alpha = \beta\tau^2$ и $K(1 + \alpha^2) = \sigma^2 + \tau^2 + \beta^2\tau^2$. Тогда $K = \beta\tau^2/\alpha$ и α является корнем уравнения $\beta\tau^2\alpha^2 - (\sigma^2 + \tau^2 + \beta^2\tau^2)\alpha + \beta\tau^2 = 0$. Оба корня этого уравнения действительны, поскольку

$$(\sigma^2 + \tau^2 + \beta^2\tau^2)^2 - 4(\beta\tau^2)^2 = (\sigma^2 + \tau^2 - \beta^2\tau^2)^2 + 4\sigma^2\beta^2\tau^2 \geq 0.$$

Глава 7, упражнение 33. В силу следствия 7.5.3

$$\begin{aligned} f(\lambda) &= \frac{\sigma^2}{2\pi} \left| 1 - \frac{1}{2m+1} \sum_{s=-m}^m e^{i\lambda s} \right|^2 = \frac{\sigma^2}{2\pi} \left| e^{i\lambda m} - \frac{1}{2m+1} \sum_{s=0}^{2m} e^{i\lambda s} \right|^2 = \\ &= \frac{\sigma^2}{2\pi} \left| e^{i\lambda m} - \frac{1}{2m+1} \cdot \frac{\sin \lambda(m+1/2)}{\sin(\lambda/2)} e^{i\lambda m} \right|^2 = \\ &= \frac{\sigma^2}{2\pi} \left| 1 - \frac{\sin \lambda(m+1/2)}{(2m+1) \sin \lambda/2} \right|^2. \end{aligned}$$

См. (71).

Глава 8, упражнение 2. Из определения (34) и из того факта, что произведение $(y_t - \mu)(y_{t+h} - \mu)(y_{t+r} - \mu)(y_{t+s} - \mu)$ можно записать 4! различными способами (перестановкой скобок), вытекает, что $v(h, r, s)$ имеет 24 симметрии. В силу (36) те же самые симметрии имеет и $\kappa(h, r, s)$. Не указанные в (а) и (б) остальные 15 форм получаются применением результатов (а) к каждой из форм, указанных в (б):

$$\begin{aligned} \kappa(h, r, s) &= \kappa(-h, s-h, r-h) = \kappa(r-h, -h, s-h) = \\ &= \kappa(r-h, s-h, -h) = \kappa(s-h, -h, r-h) = \\ &= \kappa(s-h, r-h, -h); \\ \kappa(h, r, s) &= \kappa(-r, s-r, h-r) = \kappa(h-r, -r, s-r) = \\ &= \kappa(h-r, s-r, -r) = \kappa(s-r, -r, h-r) = \\ &= \kappa(s-r, h-r, -r); \\ \kappa(h, r, s) &= \kappa(-s, r-s, h-s) = \kappa(h-s, -s, r-s) = \\ &= \kappa(h-s, r-s, -s) = \kappa(r-s, -s, h-s) = \\ &= \kappa(r-s, h-s, -s). \end{aligned}$$

Глава 8, упражнение 12. Имеем

$$\begin{aligned}\mathcal{E}(y_t - \mu)^2 (y_s - \mu) &= \sum_{r,p,q=-\infty}^{\infty} \gamma_r \gamma_p \gamma_q \mathcal{E} v_{t-r} v_{t-p} v_{s-q} = \\ &= \kappa_3 \sum_{r=-\infty}^{\infty} \gamma_r^2 \gamma_{r+s-t}.\end{aligned}$$

Поскольку $\text{Cov}[(y_t - \mu)^2, y_s] = \mathcal{E}(y_t - \mu)^2 (y_s - \mu)$, то

$$\begin{aligned}\text{Cov}(\bar{y}, C_0) &= \frac{1}{T^2} \sum_{t,s=1}^T \mathcal{E}(y_t - \mu)^2 (y_s - \mu) = \\ &= \frac{\kappa_3}{T^2} \sum_{t,s=1}^T \sum_{r=-\infty}^{\infty} \gamma_r^2 \gamma_{r+s-t}\end{aligned}$$

и отсюда следует (i). Тогда (i) по абсолютной величине не превосходит

$$\frac{\kappa_3}{T} \max_r |\gamma_r| \sum_{r=-\infty}^{\infty} |\gamma_r| \sum_{h=-(T-1)}^{T-1} \left(1 - \frac{|h|}{T}\right) |\gamma_{r+h}|.$$

В силу леммы 8.3.1 вторая сумма ограничена сверху своим пределом $\sum_{h=-\infty}^{\infty} |\gamma_h|$.

Отсюда и следует (ii).

Глава 8, упражнение 17. При сделанных предположениях $\text{Cov}(y_t y_s, y_t y_s) = \sigma^4$, $t \neq s$, $\text{Cov}(y_t^2, y_s^2) = 2\sigma^4 + \kappa_4$ и $\text{Cov}(y_t y_s, y_t y_{s'}) = 0$ в остальных случаях. Поэтому [см. (76)] для $\lambda \neq \pm \lambda'$

$$\begin{aligned}(2\pi T)^2 \text{Cov}[I(\lambda), I(\lambda')] &= \sum_{t,s,t',s'=1}^T e^{i\lambda(t-s) + i\lambda'(t'-s')} \text{Cov}(y_t y_s, y_{t'} y_{s'}) = \\ &= T\kappa_4 + \sigma^4 \sum_{t,s=1}^T [e^{i(\lambda+\lambda')(t-s)} + e^{i(\lambda-\lambda')(t-s)}] = \\ &= T\kappa_4 + \sigma^4 \left[\frac{\sin^2 \frac{1}{2}(\lambda + \lambda')T}{\sin^2 \frac{1}{2}(\lambda + \lambda')} + \frac{\sin^2 \frac{1}{2}(\lambda - \lambda')T}{\sin^2 \frac{1}{2}(\lambda - \lambda')} \right].\end{aligned}$$

Заметим, что $I(\lambda) = I(-\lambda)$. Выражения для дисперсии представляют собой частные случаи формулы для ковариации.

Глава 8, упражнение 26. В силу (45)

$$\begin{aligned}\sum_{h=-\infty}^{\infty} \sigma^2(h) &= \sigma^4 \sum_{h,s,s'=-\infty}^{\infty} \gamma_s \gamma_{s+h} \gamma_{s'} \gamma_{s'+h} \leq \\ &\leq \sigma^4 \max_t |\gamma_t| \sum_{h,s,s'=-\infty}^{\infty} |\gamma_s| \cdot |\gamma_{s+h}| \cdot |\gamma_{s'}| = \\ &= \sigma^4 \max_t |\gamma_t| \left(\sum_{s=-\infty}^{\infty} |\gamma_s| \right)^2.\end{aligned}$$

Глава 9, упражнение 21. Левая часть неравенства, стоящего в указании, равна

$$\left| \sum_{r=0}^{2K} \frac{1}{K} g\left(\frac{r-K}{K}\right) e^{idr} \right|,$$

а $g((r-K)/K) = C$ для $A = aK + K < r \leq bK + K = B$ и равна нулю в противном случае. Пусть $[x]$ обозначает целую часть числа x , а $\{x\}$ — наименьшее целое число, большее чем x . Тогда выписанное выше выражение равно

$$\frac{C}{K} \left| \sum_{r=\{A\}}^{[B]} e^{idr} \right| = \frac{C}{K} \left| \frac{1 - e^{id([B] - \{A\} + 1)}}{1 - e^{id}} \right| \leq \frac{C}{K |\sin d/2|},$$

поскольку $|1 - e^{id}| = 2 |\sin d/2|$.

Глава 9, упражнение 22. Пусть

$$h_j(x) = \begin{cases} C_j, & a_{j-1} < x \leq a_j, \\ 0, & \text{в противном случае, } j = 1, \dots, n, \end{cases}$$

так что $h(x) = \sum_{j=1}^n h_j(x)$. Результат упр. 21 сохраняется и для $C < 0$ и для функции $g(x)$, такой, что $g(x) = C$, $a \leq x \leq b$, и $g(x) = 0$ вне отрезка $[a, b]$. Поэтому условие $h(-1) = C_1$ не ограничительно. Желаемый предел равен тогда

$$\sum_{j=1}^n \lim_{K \rightarrow \infty} \sum_{r=-K}^K \frac{1}{K} h_j\left(\frac{r}{K}\right) e^{idr},$$

но при каждом j соответствующее слагаемое стремится к нулю.

Глава 9, упражнение 23. Поскольку $f(x)$ равномерно непрерывна на $[-1, 1]$, то для любого $\varepsilon > 0$ существует такое $\delta > 0$, что $|f(x) - f(y)| < \varepsilon$ при $|x - y| < \delta$. Зафиксируем $\varepsilon > 0$ и выберем такое δ и такое разбиение $-1 = a_0 < a_1 < \dots < a_n = 1$, что

$$\max_{j=1, \dots, n} |a_j - a_{j-1}| < \delta.$$

Выберем a'_j так, чтобы $a_{j-1} < a'_j \leq a_j$, $j = 1, \dots, n$. Определим $h(x) = f(a'_j)$, $a_{j-1} < x \leq a'_j$, $j = 1, \dots, n$, и положим $h(-1) = f(a'_1)$. Тогда $|f(x) - h(x)| < \varepsilon$ для $-1 \leq x \leq 1$ и

$$\begin{aligned} \left| \sum_{r=-K}^K \frac{1}{K} f\left(\frac{r}{K}\right) e^{idr} - \sum_{r=-K}^K \frac{1}{K} h\left(\frac{r}{K}\right) e^{idr} \right| &\leq \\ &\leq \frac{1}{K} \sum_{r=-K}^K \left| f\left(\frac{r}{K}\right) - h\left(\frac{r}{K}\right) \right| < \left(2 + \frac{1}{K}\right) \varepsilon. \end{aligned}$$

В силу упр. 22

$$\lim_{K \rightarrow \infty} \sum_{r=-K}^K \frac{1}{K} h\left(\frac{r}{K}\right) e^{idr} = 0$$

для $d \neq k2\pi$, $k = 0, \pm 1, \dots$, и искомый результат вытекает из произвольности выбора ε .

Глава 10, упражнение 10. Для любого p -компонентного комплексного вектора $\mathbf{c} \neq 0$ справедливо равенство $\mathbf{c}' \mathbf{R}(h) \bar{\mathbf{c}} = \int_{-\pi}^{\pi} e^{i\lambda h} \mathbf{c}' d\mathbf{M}(\lambda) \bar{\mathbf{c}}$ в силу (64). Поскольку матрица $\mathbf{R}(0)$ положительно определена, а приращения матрицы $\mathbf{M}(\lambda)$ положительно полуопределены, то функция $\mathbf{c}' \mathbf{M}(\lambda) \bar{\mathbf{c}} / \mathbf{c}' \mathbf{R}(0) \bar{\mathbf{c}}$ является функцией распределения. Без потери общности можно предположить, что $\mathbf{c}' \mathbf{R}(0) \bar{\mathbf{c}} = 1$. Тогда $2\pi \int_{-\pi}^{\pi} f(\lambda) \mathbf{c}' d\mathbf{M}(\lambda) \bar{\mathbf{c}}$ является математическим ожиданием случайной величины $U = 2\pi f(X)$, где X имеет функцию распределения $\mathbf{c}' \mathbf{M}(\lambda) \bar{\mathbf{c}}$. Это математическое ожидание можно также выразить в виде $\int_{-\pi}^{\pi} u \mathbf{c}' d\mathbf{T}(u) \bar{\mathbf{c}}$, где $\mathbf{c}' \mathbf{T}(u) \bar{\mathbf{c}}$ — функция распределения величины U . [См. (96) и (97).] Искомый результат вытекает теперь из произвольности выбора вектора $\mathbf{c}$.

СПИСОК ЛИТЕРАТУРЫ

КНИГИ

- Андерсон О. (Anderson O.)
(1929) Die Korrelationsrechnung in der Konjunkturforschung, Schroeder, Bonn.
- Андерсон Р. (Anderson R. L.)
(1941) Serial Correlation in the Analysis of Time Series, unpublished thesis, Iowa State College, Ames, Iowa.
- Андерсон Р. и Хаузмэн (Anderson R. L., Houseman E. E.)
(1942) Tables of Orthogonal Polynomial Values Extended to $N=104$ (Research Bulletin 297), Agricultural Experiment Station, Iowa State College, Ames, Iowa.
- Андерсон Т. (Anderson T. W.)
(1958) An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, Inc., New York. (Русский перевод: Андерсон Т. У., Введение в многомерный статистический анализ, Физматгиз, М., 1963.)
- Бартлетт (Bartlett M. S.)
(1966) An Introduction to Stochastic Processes with Special Reference to Methods and Applications (Second Edition), Cambridge Univ. Press, Cambridge.
- Блэкмен и Тьюки (Blackman R. B., Tukey J. W.)
(1959) The Measurement of Power Spectra from the Point of View of Communication Engineering, Dover Publications, Inc., New York (printed originally 1958, *Bell System Tech. J.*, 37, 185—282, 485—569).
- Бокс и Дженкинс (Box George E. P., Jenkins Gwilym M.)
(1970) Time Series Analysis Forecasting and Control, Holden-Day, Inc., San Francisco. (Русский перевод: Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление, вып. 1, «Мир», М., 1974. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление, вып. 2, «Мир», М., 1974.)
- Бюро переписей США (United States Bureau of the Census)
(1955) Statistical Abstract of The United States, U. S. Government Printing Office, Washington, D. C.
- Бьюйс-Баллот (Buijs Ballot C. H. D.)
(1847) Les Changement Périodiques de Température, Utrecht.
- Вальд (Wald A.)
(1936) Berechnung und Ausschaltung von Saisonschwankungen, Springer-Verlag, Vienna.
- Вальдмейер (Waldmeier M.)
(1961) The Sunspot Activity in the Years 1610—1960, Schulthess & Co., Zürich.

- Ватсон (Watson G. S.)
(1952) *Serial Correlation in Regression Analysis*, Mimeo Series No. 49, Institute of Statistics, Univ. of North Carolina, Chapel Hill, N. C.
- Винер (Wiener Norbert)
(1949) *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*, Technology Press of the Massachusetts Institute of Technology, Cambridge, Mass.
- Вольд (Wold Herman)
(1954) *A Study in the Analysis of Stationary Time Series (Second Edition)*, Almqvist and Wiksell Book Co., Uppsala.
- (1965) *Bibliography on Time Series and Stochastic Processes*, Oliver and Boyd Ltd., Edinburgh.
- Грейбилл (Graybill Franklin A.)
(1961) *An Introduction to Linear Statistical Models*, vol. I, McGraw-Hill Book Co., Inc., New York.
- Гренандер и Розенблатт (Grenander Ulf, Rosenblatt Murray)
(1957) *Statistical Analysis of Stationary Time Series*, John Wiley & Sons, Inc., New York.
- Гренджер и Хатанака (Granger C. W. J., Hatanaka M.)
(1964) *Spectral Analysis of Economic Time Series*, Princeton Univ. Press, Princeton, N. J. (Русский перевод: Гренджер К., Хатанака М., *Спектральный анализ временных рядов в экономике*, «Статистика», М., 1972.)
- Департамент сельского хозяйства США (United States Department of Agriculture)
(1939) *Agricultural Statistics*, United States Government Printing Office, Washington, D. C.
- Дрейпер и Смит (Draper N. R., Smith H.)
(1966) *Applied Regression Analysis*, John Wiley & Sons, Inc., New York. (Русский перевод: Дрейпер Н. и Смит Г., *Прикладной регрессионный анализ*, «Статистика», М., 1973.)
- Дуб (Doob J. L.)
(1953) *Stochastic Processes*, John Wiley & Sons, Inc., New York. (Русский перевод: Дуб Дж., *Вероятностные процессы*, ИЛ, М., 1956.)
- Дэвис (Davis Harold)
(1941) *The Analysis of Economic Time Series (The Cowles Commission for Research in Economics, Monograph No. 6)*, Principia Press, Inc., Bloomington, Ind.
- Жордан (Jordan Charles)
(1939) *Calculus of Finite Differences*, Budapest (reprinted in 1947 by Chelsea Publishing Co., New York.)
- Зигмунд (Zygmund A.)
(1959) *Trigonometric Series, Vol. I*, Cambridge Univ. Press, Cambridge. (Русский перевод: Зигмунд А., *Тригонометрические ряды*, т. I, «Мир», М., 1965.)
- Кемпторн (Kempthorne Oscar)
(1952) *The Design and Analysis of Experiments*, John Wiley & Sons, Inc., New York.
- Кендалл (Kendall Maurice G.)
(1946a) *The Advanced Theory of Statistics, Vol. II*, Charles Griffin and Co., Ltd., London.
- (1946b) *Contributions to the Study of Oscillatory Time-Series*, Nat. Inst. Econ. Soc. Res. Occasional Papers IX, Cambridge Univ. Press, Cambridge.
- Кендалл и Стьюарт (Kendall Maurice G., Stuart Alan)
(1961) *The Advanced Theory of Statistics, Vol. 2*, Charles Griffin and Co., Ltd., London, and Hafner Publishing Co., Inc., New York. (Русский перевод: Кендалл М., Стьюарт А., *Статистические выводы и связи*, «Наука», М., 1973.)

- (1966) *The Advanced Theory of Statistics*, Vol. 3, Charles Griffin and Co., Ltd., London, and Hafner Publishing Co., Inc., New York.
- Колмогоров А. Н.
(1933) *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlin; reprinted in 1946 by Chelsea Publishing Co., New York. (Русский перевод: Колмогоров А. Н., Основные понятия теории вероятностей, ОНТИ, М., 1936, «Наука», М., 1974.)
- Крамер (Cramér Harald)
(1946) *Mathematical Methods of Statistics*, Princeton Univ. Press, Princeton, N. J. (Русский перевод: Крамер Г., Математические методы статистики, 1-е изд., ИЛ, М., 1948, 2-е изд., «Мир», М., 1975.)
- Кузнец (Kuznets S.)
(1954) *National Income and Its Composition: 1919—1938*, National Bureau of Economic Research, New York.
- Курант и Гильберт (Courant R., Hilbert D.)
(1937) *Methoden der Mathematischen Physik*, Springer-Verlag, Berlin. (Русский перевод: Курант Р., Гильберт Д., Методы математической физики, т. I, Гостехиздат, 1951.)
- Ланцош (Lanczos Cornelius)
(1956) *Applied Analysis*, Prentice-Hall, Inc., Englewood Cliffs, N. J. (Русский перевод: Ланцош К., Практические методы прикладного анализа, Физматгиз, М., 1961.)
- Лаплас (Laplace P. S.)
(1829) *Traité de Mécanique Céleste, Tome Second (Second Édition)*, Bachelier Successeur de Mme. Ve. Courcier, Paris.
- Леман (Lehmann E. L.)
(1959) *Testing Statistical Hypotheses*, John Wiley & Sons, Inc., New York. (Русский перевод: Леман Э., Проверка статистических гипотез, «Наука», М., 1964.)
- Лозэ (Loève Michel)
(1963) *Probability Theory (Third Edition)*, D. Van Nostrand Co., Inc., New York. (Русский перевод: Лозэ М., Теория вероятностей, ИЛ, М., 1962.)
- Лукач (Lukacs Eugene)
(1960) *Characteristic Functions*, Charles Griffin and Co., Ltd, London, and Hafner Publishing Co., Inc., New York.
- Маколей (Macaulay F. R.)
(1938) *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields and Stock Prices in the United States Since 1856*, National Bureau of Economic Research, New York.
- Миллер (Miller Kenneth S.)
(1960) *An Introduction to the Calculus of Finite Differences and Difference Equations*, Henry Holt & Co., New York.
- Парзен (Parzen Emanuel)
(1962) *Stochastic Processes*, Holden-Day, Inc., San Francisco.
- Плэккетт (Plackett R. L.)
(1960) *Principles of Regression Analysis*, Clarendon Press, Oxford.
- Рао С. (Rao C. Radhakrishna)
(1952) *Advanced Statistical Methods in Biometric Research*, John Wiley & Sons, Inc., New York.
- Розенблатт (Rosenblatt Murray)
(1962) *Random Processes*, Oxford Univ. Press, New York.
- Тернбулл и Эйкен (Turnbull H. W., Aitken A. C.)
(1952) *An Introduction to the Theory of Canonical Matrices (Third Edition)*, Blackie and Son Ltd., London (reprinted in 1961 by Dover Publications, Inc., New York).
- Тинтнер (Tintner Gerhard)
(1940) *The Variate Difference Method*, Principia Press, Inc., Bloomington, Ind.

- (1952) *Econometrics*, John Wiley & Sons, Inc., New York.
 Уилкс (Wilks Samuel S.)
 (1962) *Mathematical Statistics*, John Wiley & Sons, Inc., New York. (Русский перевод: Уилкс С., Математическая статистика, «Наука», М., 1967.)
 Уильямс (Williams E. J.)
 (1959) *Regression Analysis*, John Wiley & Sons, Inc., New York.
 Уиттекер и Ватсон (Whittaker E. T., Watson G. N.)
 (1943) *A Course of Modern Analysis*, Cambridge Univ. Press, Cambridge, and The Macmillan Co., New York. (Русский перевод: Уиттекер Э. Т. и Ватсон Дж. Н., Курс современного анализа, т. I, II, «Наука», М., 1963.)
 Уиттекер и Робинсон (Whittaker E. T., Robinson G.)
 (1926) *The Calculus of Observations, A Treatise on Numerical Mathematics* (Second Edition), Blackie & Son Ltd., London. (Русский перевод: Уиттекер Э. и Робинсон Г., Математическая обработка результатов измерений, ОНТИ, 1933.)
 Уиттл (Whittle Peter)
 (1951) *Hypothesis Testing in Time Series Analysis*, Almqvist and Wiksell Book Co., Uppsala.
 Феллер (Feller William)
 (1968) *An Introduction to Probability Theory and Its Applications*, Vol. I (Third Edition), John Wiley & Sons, Inc., New York. (Русский перевод: Феллер В., Введение в теорию вероятностей и ее приложения, т. I, Мир, «М», 1964.)
 Фишер и Йейтс (Fisher R. A., Yates Frank)
 (1963) *Statistical Tables for Biological, Agricultural and Medical Research* (Sixth Edition), Oliver and Boyd Ltd., Edinburgh.
 Халмош (Halmos Paul R.)
 (1958) *Finite-Dimensional Vector Spaces*, D. Van Nostrand Co., Inc., Princeton, N. J. (Русский перевод: Халмош П., Конечномерные векторные пространства, Физматгиз, М., 1963.)
 Хемминг (Hamming Richard W.)
 (1962) *Numerical Methods for Scientists and Engineers*, McGraw-Hill Book Co., Inc., New York. (Русский перевод: Хемминг Р., Численные методы для научных работников и инженеров, «Наука», М., 1972.)
 Хеннан (Hannan E. J.)
 (1960) *Time Series Analysis*, Methuen and Co., Ltd., London, and John Wiley & Sons, Inc., New York. (Русский перевод: Хеннан Э., Анализ временных рядов, «Наука», М., 1964.)
 Хобсон (Hobson E. W.)
 (1907) *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, Cambridge Univ. Press, Cambridge.
 Шеффе (Scheffé Henry)
 (1959) *The Analysis of Variance*, John Wiley & Sons, Inc., New York. (Русский перевод: Шеффе Г., Дисперсионный анализ, Физматгиз, М., 1963.)

СТАТЬИ

- Акутович (Akutowicz Edwin J.)
 (1957) On an explicit formula in least squares prediction, *Math. Scand.*, 5, 261—266.
 Андерсон Р. (Anderson R. L.)
 (1942) Distribution of the serial correlation coefficient, *Ann. Math. Statist.*, 13, 1—13.
 Андерсон Р. и Андерсон Т. (Anderson R. L., Anderson T. W.)
 (1950) Distribution of the circular serial correlation coefficient for residuals from a fitted Fourier series, *Ann. Math. Statist.*, 21, 59—81.

- Андерсон Т. (Anderson T. W.)
(1948) On the theory of testing serial correlation, *Skand. Aktuarietidskr.*, **31**, 88—116.
(1959) On asymptotic distributions of estimates of parameters of stochastic difference equations, *Ann. Math. Statist.*, **30**, 676—687.
(1963) Determination of the order of dependence in normally distributed time series, Proc. Symp. Time Series Anal. Brown Univ. (Murray Rosenblatt, ed.), John Wiley & Sons, Inc., New York, 425—446.
- Андерсон Т. и Рубин (Anderson T. W., Rubin Herman)
(1950) The asymptotic properties of estimates of the parameters of a single equation in a complete system of stochastic equations, *Ann. Math. Statist.*, **21**, 570—582.
- Андерсон Т. и Уолкер А. (Anderson T. W., Walker A. M.)
(1964) On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process, *Ann. Math. Statist.*, **35**, 1296—1303.
- Бартлетт (Bartlett M. S.)
(1935) Some aspects of the time-correlation problem in regard to tests of significance, *J. Roy. Statist. Soc.*, **98**, 536—543.
(1946) On the theoretical specification and sampling properties of autocorrelated time-series, *J. Roy. Statist. Soc. Supp.*, **8**, 27—41, 85—97. (Corrigenda (1948), **10**, 200.)
(1950) Periodogram analysis and continuous spectra, *Biometrika*, **37**, 1—16.
- Бартлетт и Диананда (Bartlett M. S., Diananda P. H.)
(1950) Extensions of Quenouille's test for autoregressive schemes, *J. Roy. Statist. Soc. Ser. B*, **12**, 108—115.
- Бевеидж (Beveridge W. H.)
(1921) Weather and harvest cycles, *Econ. J.*, **31**, 429—452.
(1922) Wheat prices and rainfall in Western Europe, *J. Roy. Statist. Soc.*, **85**, 412—459.
- Бирнбаум (Birnbbaum Allan)
(1959) On the analysis of factorial experiments without replication, *Technometrics*, **1**, 343—357.
(1961) A multi-decision procedure related to the analysis of single degrees of freedom, *Ann. Inst. Statist. Math. Tokyo*, **12**, 227—236.
- Ватсон (Watson G. S.)
(1955) Serial correlation in regression analysis, I, *Biometrika*, **42**, 327—341.
(1956) On the joint distribution of the circular serial correlation coefficients, *Biometrika*, **43**, 161—168.
(1967) Linear least squares regression, *Ann. Math. Statist.*, **38**, 1679—1699.
- Ватсон и Дурбин (Watson G. S., Durbin J.)
(1951) Exact tests of serial correlation using noncircular statistics, *Ann. Math. Statist.*, **22**, 446—451.
- Ватсон и Хеннан (Watson G. S., Hannan E. J.)
(1956) Serial correlation in regression analysis, II, *Biometrika*, **43**, 436—449.
- Гейссер (Geisser Seymour)
(1956) The modified mean square successive difference and related statistics, *Ann. Math. Statist.*, **27**, 819—824.
- Гренандер (Grenander Ulf)
(1954) On the estimation of the regression coefficients in the case of an autocorrelated disturbance, *Ann. Math. Statist.*, **25**, 252—272.
- Гренджер и Харс (Granger C. W. J., Hughes A. O.)
(1969) A new look at some old data: the Beveridge wheat price series, unpublished.
- Даниэлс (Daniels H. E.)
(1956) The approximate distribution of serial correlation coefficients, *Biometrika*, **43**, 169—185.
- Дженкинс (Jenkins G. M.)
(1956) Tests of hypotheses in the linear autoregressive model, II. Null distributions for higher order schemes: non-null distributions, *Biometrika*, **43**, 186—199.

- (1961) General considerations in the analysis of spectra, *Technometrics*, **3**, 133—166.
Диананда (Diananda P. H.)
- (1953) Some probability limit theorems with statistical applications, *Proc. Cambridge Philos. Soc.*, **49**, 239—246.
- Диксон (Dixon Wilfrid J.)
- (1944) Further contributions to the problem of serial correlation, *Ann. Math. Statist.*, **15**, 119—144.
- Дуб (Doob J. L.)
- (1944) The elementary Gaussian processes, *Ann. Math. Statist.*, **15**, 229—282.
- Дурбин (Durbin J.)
- (1959) Efficient estimation of parameters in moving-average models, *Biometrika*, **46**, 306—316.
- (1960a) Estimation of parameters in time-series regression models, *J. Roy. Statist. Soc. Ser. B.*, **22**, 139—153.
- (1960b) The fitting of time-series models, *Rev. Inst. Internat. Statist.*, **28**, 233—244.
- (1963) Trend elimination for the purpose of estimating seasonal and periodic components of time series, *Proc. Symp. Time Series Anal. Brown Univ.* (Murray Rosenblatt, ed.), John Wiley & Sons, Inc., New York, 3—16.
- (1970) An alternative to the bounds test for testing serial correlation in least-squares regression, *Econometrica*, **38**, 422—429.
- Дурбин и Ватсон (Durbin J., Watson G. S.)
- (1950) Testing for serial correlation in least squares regression, I, *Biometrika*, **37**, 409—428.
- (1951) Testing for serial correlation in least squares regression, II, *Biometrika*, **38**, 159—178.
- Зискинд (Zyskind George)
- (1967) On canonical forms, non-negative covariance matrices, and best and simple least squares linear estimators in linear models, *Ann. Math. Statist.*, **38**, 1092—1109.
- Ирвин (Irwin J. O.)
- (1955) A unified derivation of some well-known frequency distributions of interest in biometry and statistics, *J. Roy. Statist. Ser. A*, **118**, 389—404.
- Камат (Kamat A. R.)
- (1955) Modified mean square successive difference with an exact distribution, *Sankhya*, **15**, 295—302.
- Канторович Л. В.
- (1948) Функциональный анализ и прикладная математика, *УМН*, **3** (6), 89—185.
- Карлин и Траукс (Karlin Samuel, Traux Donald)
- (1960) Slippage problems, *Ann. Math. Statist.*, **31**, 296—324.
- Кейв-Браун-Кейв (Cave-Browne-Cave F. E.)
- (1904) On the influence of the time factor on the correlation between the barometric heights at stations more than 1000 miles apart, *Proc. Roy. Soc. London*, **74**, 403—413.
- Кенуй (Quenouille M. H.)
- (1947) A large-sample test for the goodness of fit of autoregressive schemes, *J. Roy. Statist. Soc.*, **110**, 123—129.
- (1949a) The joint distribution of serial correlation coefficients, *Ann. Math. Statist.*, **20**, 561—571.
- (1949b) Approximate tests of correlation in time-series, *J. Roy. Statist. Soc. Ser. B*, **11**, 68—84.
- (1953) Modifications to the variate-difference method, *Biometrika*, **40**, 383—408.
- Колмогоров А. Н.
- (1941a) Стационарные последовательности в гильбертовом пространстве, *Бюллетень МГУ*, **2**, 6, 1—40.

- (1941b) Интерполирование и экстраполирование стационарных случайных последовательностей, *Изв. АН СССР, сер. мат.*, 5, № 5, 3—14.
- Коуден (Cowden D. J.)
- (1962) Weights for fitting polynomial secular trends, University of North Carolina School of Business Administration Technical Paper No. 4, Chapel Hill, N. C.
- Крэдок (Craddock J. M.)
- (1967) An experiment in the analysis and prediction of time series, *The Statistician*, 17, 257—268.
- Кудо (Kudo Akiō)
- (1960) The symmetric multiple decision problems, *Mem. Fac. Sci. Kyushu Univ. Ser. A*, 14, 179—206.
- Кули, Льюис и Уэлш (Cooley James W., Lewis Peter A. W., Welch Peter D.)
- (1967) Application of the Fast Fourier Transform to computation of Fourier integrals, Fourier series, and convolution integrals, *IEEE Transactions on Audio and Electroacoustics*, AU—15, 79—84.
- Кули и Тьюки (Cooley J. W., Tukey J. W.)
- (1965) An algorithm for the machine calculation of complex Fourier series, *Math. Comput.*, 19, 297—301.
- Купменс (Koopmans Tjalling)
- (1942) Serial correlation and quadratic forms in normal variates, *Ann. Math. Statist.*, 13, 14—33.
- Купменс, Рубин и Лейпник (Koopmans T. C., Rubin H., Leipnik R. B.)
- (1950) Measuring the equation systems of dynamic economics, Statistical Inference in Dynamic Economic Models (Cowles Commission Monograph No. 10, T. C. Koopmans, ed.), John Wiley & Sons, Inc., New York, 53—237.
- Лейпник (Leipnik R. B.)
- (1947) Distribution of the serial correlation coefficient in a circularly correlated universe, *Ann. Math. Statist.*, 18, 80—87.
- Леман (Lehmann E. L.)
- (1957) A theory of some multiple decision problems, II, *Ann. Math. Statist.*, 28, 547—572.
- Магнесс и Макгир (Magness T. A., McGuire J. B.)
- (1962) Comparison of least squares and minimum variance estimates of regression parameters, *Ann. Math. Statist.*, 33, 462—470.
- Манн и Вальд (Mann H. B., Wald A.)
- (1943a) On stochastic limit and order relationships, *Ann. Math. Statist.*, 14, 217—226.
- (1943b) On the statistical treatment of linear stochastic difference equations, *Econometrica*, 11, 173—220.
- Марсалья (Marsaglia G.)
- (1954) Iterated limits and the central limit theorem for dependent variables, *Proc. Amer. Math. Soc.*, 5, 987—991.
- Моран (Moran P. A. P.)
- (1947) Some theorems on time series, I, *Biometrika*, 34, 281—291.
- Мэдоу (Madow William G.)
- (1945) Note on the distribution of the serial correlation coefficient, *Ann. Math. Statist.*, 16, 308—310.
- фон Нейман (von Neumann John)
- (1941) Distribution of the ratio of the mean square successive difference to the variance, *Ann. Math. Statist.*, 12, 367—395.
- Олшен (Olshen Richard A.)
- (1967) Asymptotic properties of the periodogram of a discrete stationary process, *J. Appl. Prob.*, 4, 508—528.
- Парзен (Parzen Emanuel)
- (1957a) A central limit theorem for multilinear stochastic processes, *Ann. Math. Statist.*, 28, 252—256.
- (1957b) On consistent estimates of the spectrum of a stationary time series, *Ann. Math. Statist.*, 28, 329—348.

- (1961a) An approach to time series analysis, *Ann. Math. Statist.*, **32**, 951—989.
- (1961b) Mathematical considerations in the estimation of spectra, *Technometrics*, **3**, 167—190.
- Рао М. (Rao M. M.)
- (1960) Estimation by periodogram, *Trabajos Estadist.*, **11**, 123—137.
- Розенблатт (Rosenblatt Murray)
- (1956) Some regression problems in time series analysis, Proc. Third Berkeley Symposium on Mathematical Statistics and Probability (J. Neyman, ed.), Vol. 1, Univ. of California Press, Berkeley and Los Angeles, 165—186.
- (1959) Statistical analysis of stochastic processes with stationary residuals, Probability and Statistics: The Harald Cramér Volume (Ulf Grenander, ed.) Almqvist and Wiksell Book Co., Stockholm, 246—275.
- Рубин (Rubin Herman)
- (1945) On the distribution of the serial correlation coefficient, *Ann. Math. Statist.*, **16**, 211—125.
- Рунге (Runge C.)
- (1903) Über die Zerlegung empirisch gegebener periodischer Funktionen in Sinuswellen, *Z. Math. Phys.*, **48**, 443—456.
- Сарган (Sargan J. D.)
- (1953) An approximate treatment of the properties of the correlogram and periodogram, *J. Roy. Statist. Soc. Ser. B*, **15**, 140—152.
- Стивенс (Stevens W. L.)
- (1939) Solution to a geometrical problem in probability, *Ann. Eugenics*, **9**, 315—320.
- Стьюдент («Student»)
- (1914) The elimination of spurious correlation due to position in time and space, *Biometrika*, **10**, 179—180.
- Тинтнер (Tintner Gerhard)
- (1955) The distribution of the variances of variate differences in the circular case, *Metron*, **17** (3—4), 43—52.
- Уиттл (Whittle Peter)
- (1952) The simultaneous estimation of a time series harmonic components and covariance structure, *Trabajos Estadist.*, **3**, 43—57.
- (1954) A statistical investigation of sunspot observations with special reference to H. Alfvén's sunspot model, *The Astrophys. J.*, **120**, 251—260.
- (1959) Sur la distribution du maximum d'un polynôme trigonométrique à coefficients aléatoires, *Colloq. Internat. Centre Nat. Rech. Sci.*, **87**, 173—183.
- Уолкер А. (Walker A. M.)
- (1952) Some properties of asymptotic power functions of goodness-of-fit tests for linear autoregressive schemes, *J. Roy. Statist. Soc. Ser. B*, **14**, 117—134.
- (1954) The asymptotic distribution of serial correlation coefficients for autoregressive processes with dependent residuals, *Proc. Cambridge Philos. Soc.*, **50**, 60—64.
- (1961) Large-sample estimation of parameters for moving-average models, *Biometrika*, **48**, 343—357.
- (1962) Large-sample estimation of parameters for autoregressive processes with moving-average residuals, *Biometrika*, **49**, 117—131.
- (1965) Some asymptotic results for the periodogram of a stationary time series, *J. Austral. Math. Soc.*, **5**, 107—128.
- (1968) Large-sample properties of least-squares estimators of harmonic components in a time series with stationary residuals, I. Independent residuals, Technical Report, Department of Statistics, Stanford University, Stanford, Calif.
- Уолкер Дж. Т. (Walker G. T.)
- (1914) Correlation in seasonal variation of weather. III. On the criterion for the reality of relationships or periodicities, *Mem. Indian Meteorol. Dept.*, **21** (9), 13—15.
- (1931) On periodicity in series of related terms, *Proc. Roy. Soc. London Ser. A*, **131**, 518—532.

- Фишер (Fisher R. A.)
(1929) Tests of significance in harmonic analysis, *Proc. Roy. Soc. London Ser. A*, 125, 54—59.
(1940) On the similarity of the distributions found for the test of significance in harmonic analysis, and in Stevens's problem in geometrical probability, *Ann. Eugenics*, 10, 14—17.
- Хаавелмо (Haavelmo Trygve)
(1947) Methods of measuring the marginal propensity to consume, *J. Amer. Statist. Assoc.*, 42, 105—122.
- Харт (Hart B. J.)
(1942) Significance levels for the ratio of the mean square successive difference to the variance, *Ann. Math. Statist.*, 13, 445—447.
- Харт и фон Нейман (Hart B. J., von Neumann John)
(1942) Tabulation of the probabilities for the ratio of the mean square successive difference to the variance, *Ann. Math. Statist.*, 13, 207—214.
- Хеннан (Hannan E. J.)
(1958) The estimation of the spectral density after-trend removal, *J. Roy. Statist. Soc. Ser. B*, 20, 323—333.
(1961) A central limit theorem for systems of regression, *Proc. Cambridge Philos. Soc.*, 57, 583—588.
(1963) Regression for time series, *Proc. Symp. Time Series Anal. Brown Univ.* (Murray Rosenblatt, ed.), John Wiley & Sons, Inc., New York, 17—37.
(1964) The estimation of a changing seasonal pattern, *J. Amer. Statist. Assoc.*, 59, 1063—1077.
- Хефдинг и Роббинс (Hoeffding Wassily, Robbins Herbert)
(1948) The central limit theorem for dependent random variables, *Duke Math. J.*, 15, 773—780.
- Хукер (Hooker R. H.)
(1905) On the correlation of successive observations, illustrated by corn prices, *J. Roy. Statist. Soc.*, 68, 696—703.
- Чон и Хеннан (Cheong H. A., Hannan E. J.)
(1968) The asymptotic distribution of spectral estimates, unpublished.
- Шерф (Schaefer M. Casini)
(1964) Estimation of the covariance and autoregressive structure of a stationary time series, Technical Report, Department of Statistics, Stanford University, Stanford, Calif.
- Шеффе (Scheffé Henry)
(1970) Multiple testing versus multiple estimation. Improper confidence sets. Estimation of directions and ratios, *Ann. Math. Statist.*, 41, 1—29.
- Шустер (Schuster A.)
(1898) On the investigation of hidden periodicities with application to a supposed 26-day period of meteorological phenomena, *Terr. Mag. Atmos. Elect.*, 3, 13.
(1906) On the periodicities of sunspots, *Philos. Trans. Roy. Soc. London Ser. A*, 206, 69—100.
- Эйкер (Eicker F.)
(1963) Asymptotic normality and consistency of the least squares estimators for families of linear regressions, *Ann. Math. Statist.*, 34, 447—456.
- Юл (Yule G. Udny)
(1927) On a method for investigating periodicities in disturbed series with special reference to Wolfer's sunspot numbers, *Philos. Trans. Roy. Soc. London Ser. A*, 226, 267—298.
- Яглом А. М.
Введение в теорию стационарных случайных функций, *УМН*, VII, вып. 5, (1952), 3—168.
- Янг (Young L. C.)
(1941) On randomness in ordered sequences, *Ann. Math. Statist.*, 12, 293—300.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Автоковариация, см. Выборочная ковариация; Сериальная корреляция
Автокорреляция 284, см. также Выборочная ковариация; Ковариационная функция; Сериальная корреляция

Авторегрессии процесс, см. Авторегрессии процесс с остатками в виде скользящего среднего; Вероятностная структура процесса авторегрессии; Реализации трех процессов авторегрессии; Статистические выводы для процессов авторегрессии

Авторегрессии процесс с остатками в виде скользящего среднего
вероятностная структура 444—447
параметризация с помощью корреляций 265
представление с помощью оператора запаздывания и в виде скользящего среднего 263, 264

оценивание параметров
использование выборочных корреляций 265—270
использование представления в виде скользящего среднего 270, 271
Амплитуда 15, 110

Барицентрические координаты 146, 186, 398, 723, 724

Бесселева функция 135

Быстрое преобразование Фурье 154—156, 599

Бьюис-Баллота таблица 122, 123, 129

Векторный случайный процесс авторегрессии первого порядка 203—208

ковариационная структура 425, 426
Вероятностная структура процесса авторегрессии, см. также Авторегрессии процесс с остатками в виде скользящего среднего; Статистические выводы для процессов авторегрессии; Стационарный случайный процесс

бесконечный порядок 459—461
векторный процесс первого порядка 203—208
ковариационная структура 199—201, 413

Юла — Уолкера уравнения 200
представление процесса

в виде бесконечного скользящего среднего 192, 193, 195—197, 413, 442, 443

в виде векторного процесса первого порядка 203, 208

с помощью оператора запаздывания 194

присоединенное алгебраическое уравнение (характеристическое уравнение) 195, 277, 278, 413, 473, 724
как характеристическое уравнение векторного процесса 206

с корнем, равным единице 197
прогноз с минимальной среднеквадратичной ошибкой 203

с независимыми переменными 201, 202

с ошибкой 191, 414, 473

- спектральная плотность 442—444, 473
стохастическое разностное уравнение 189, 190, 277, 286, 195—197, 472
флуктуация 16, 17, 201
- Выборочная ковариация**
построенная по остаткам от тренда
асимптотическое смещение 646
определение 643
предельное распределение 647
случай известного среднего
ее асимптотические дисперсия и ковариация 503—507
ее ковариация 486—488
как несмещенная оценка 477
предельное распределение 518—521
случай неизвестного среднего
асимптотическое среднее 502
ее асимптотическая ковариация 508—510
ее ковариация 489, 490
математическое ожидание 484—486, 539
определения 477
- Выборочная корреляция**, ее асимптотическая нормальность 260—263, 529—536, см. также Сериальная корреляция
- Выборочная спектральная плотность**
несостоятельность 514, 517, 525
случай известного среднего
асимптотические ковариация и среднее 511—516
асимптотическое представление 526—529
интерпретация 421
ковариация 492—495, 540
определение 478
предельное распределение 517, 525
среднее 417, 490, 491
случай неизвестного среднего
асимптотические ковариация и среднее 511—513, 516
предельное распределение 525
среднее 491, 492
строющаяся по остаткам от тренда
асимптотическое среднее 648—652
определение 648
- Выборочная спектральная функция распределения**
асимптотическая ковариация 606
как оценка спектральной функции распределения 595
состоятельность 537, 595
- Выборочное среднее** 477
асимптотическое среднее и дисперсия 497—501
асимптотическое распределение 517, 518
дисперсия 480
ковариация выборочного среднего и выборочной дисперсии 540
- Выборочные тригонометрические коэффициенты**, см. также Тригонометрическая регрессия с периодами, являющимися целыми делителями длины ряда; Тригонометрическая регрессия с периодами, не являющимися целыми делителями длины ряда
случай известного среднего
асимптотическая нормальность 521—525
асимптотические ковариации 517
выражения для них 172—174
ковариации 160, 495—497
определение 157, 477, 478
среднее 160, 161
случай неизвестного среднего
асимптотическая нормальность 524
ковариация 177
определение 175, 478
среднее 175—177
- Выбор порядка зависимости в моделях сериальной корреляции**, см. также Критерии порядка зависимости в моделях сериальной корреляции
случай известного среднего 301—308
выбор с помощью односторонних критериев 303—305
выбор с помощью двусторонних несмещенных критериев 306, 307
неймановская структура и подобие областей 303, 304
случай неизвестного среднего 331
- Гауссовская модель** 286
Гиббса явление 558
Гильбертово пространство 449—455, см. также Линейное векторное пространство
критерий Коши сходимости 450
линейное многообразие 453
норма 450
ортогональность 454
полнота 450
порожденное стационарным процессом 449, 450
проекция на подпространство 455
расстояние 450
теорема о проекции 454
- Дурбина — Ватсона статистика** 669, 670, 676

Индекс Бевериджа цен на пшеницу
183, 184

- корреляции 683
- оценка спектральной плотности 596, 693—695
- подобраный процесс авторегрессии 275, 276
- спектрограмма 537, 684—692
- статистические данные 677—682
- Интенсивность 124, см. также Выборочная спектральная плотность; Спектрограмма; Тригонометрическая регрессия с периодами, не являющимися целыми делителями длины ряда

Квадратичные формы

- для вероятностных моделей с зависимостью 288, 289
- системы с двойными корнями 324, 325
- системы форм 308—311, 334—336
- для независимых случайных величин дисперсии 83, 107
- ковариации 87, 88
- для нормальных случайных величин
 - моменты 366, 367
 - примеры 538, 539
 - производящая функция моментов 366, 400
 - семиинварианты 366, 660, 665
 - характеристические функции 337—339, 341
- для случайных величин, образующих стационарный процесс
 - дисперсия 481, 482
 - ковариации 483, 484, 547—549
 - математическое ожидание 481, 546, 547
- канонические формы 339—341
- строящиеся по остаткам 326, 327
- независимость выборочного среднего и остатков 327
- характеристические корни и векторы 308—310

Квадратичных форм отношения, см. также

- Сериальная корреляция
- асимптотическая нормальность для случайных величин, образующих стационарный процесс 542
- независимые случайные величины
- асимптотическая нормальность 107
- распределение 338

Квадратурная спектральная плотность
426**Ковариационная матрица, оценка по методу наименьших квадратов 640—657**

- Ковариационная функция
 - аппроксимация ковариационными функциями процессов скользящего среднего и авторегрессии 473
 - векторного процесса 425
 - взвешенного среднего 62, 63
 - в циклической модели 389—392
 - временного ряда, составленного из r -х разностей исходного ряда 80
 - как преобразование Фурье спектральной функции распределения 416
 - комплекснозначного процесса 425
 - производящая функция ковариаций 252
 - процесса авторегрессии 199—201
 - процесса скользящего среднего 251, 252
 - случайного процесса 408
 - стационарного процесса 199—201
- Корреляционная функция, см. Ковариационная функция; Стационарный случайный процесс
- Коспектральная плотность 426
- Коэффициент корреляции (пирсоновский) 126, 284
- распределение 374
- Критерии порядка зависимости в моделях сериальной корреляции
 - случай известного среднего
 - достаточные статистики 290
 - использование сериального коэффициента корреляции 294, 298
 - неймановская структура критерия с подобной областью 291
 - полнота семейства распределений 291
 - равномерно наиболее мощный несмещенный критерий 294—298
 - равномерно наиболее мощный односторонний критерий 290—294, 298—301
 - случай неизвестного среднего
 - достаточные статистики 326
 - использование сериального коэффициента корреляции 331
 - равномерно наиболее мощный несмещенный критерий 329, 330
 - равномерно наиболее мощный односторонний критерий 329
 - строящийся по остаткам от тренда 657—659
- Критерий Уолкера 137, 138
- Фишера 145—148
- Шустера 135, 136
- Кронекерово произведение матриц 226

Линейная регрессия, см. также Нелинейная регрессия; Полиномиальная регрессия; Тригонометрическая регрессия с периодами, не являющимися делителями длины ряда; Тригонометрическая регрессия с периодами, являющимися делителями длины ряда доверительные области для коэффициентов 22, 23

критерии для коэффициентов 22, 23

линейные оценки в случае ковариационной матрицы общего вида

наилучшие линейные несмещенные оценки 30—32, 609—616

нижняя граница эффективности оценок наименьших квадратов 619, 620

условия эквивалентности марковских оценок и оценок наименьших квадратов 31, 615, 616, 621 эллипсоиды рассеяния 616, 617 эффективность оценок наименьших квадратов 616—621, 673

линейные оценки в случае стационарной случайной составляющей асимптотическая нормальность 638—640

асимптотическая эффективность оценок наименьших квадратов 622—635

предельная ковариационная матрица оценок наименьших квадратов и марковских оценок и ее состоятельная оценка 518, 640

условия асимптотической эффективности оценок наименьших квадратов 631, 632, 636, 637

марковские оценки 30—32, 609—616

оценки наименьших квадратов в случае некоррелированных наблюдений, имеющих одинаковую дисперсию

асимптотическая нормальность 35—38

геометрическое представление 23—25

достаточность в случае нормально распределенных наблюдений 22, 40

ковариации 21

линейные несмещенные оценки с наименьшей дисперсией 22

линейные преобразования независимых переменных 25—30

математическое ожидание 21

нормальность при нормально распределенных наблюдениях 22 ортогональность остатков 23 состоятельность оценки дисперсии 38, 39

теорема Гаусса — Маркова 22, 39, 40

прогнозирование 32—35

доверительные интервалы 41

Линейное векторное пространство 451, 452, см. также Гильбертово пространство

метрика 452

неравенство Коши — Шварца 451

треугольника 451, 452

норма 451

полнота 452

последовательность Коши 452

расстояние 452

скалярное произведение 451

Линейное прогнозирование стационарных процессов, см. также Вероятностная структура процесса авторегрессии

дисперсия 458

как проектирование 454, 455

с минимальной среднеквадратичной ошибкой 453, 455

Линейные преобразования стационарных процессов

влияние на спектральную плотность 447—449

ковариационная функция 434

спектральная функция 435

стационарность 434

Линейный процесс, см. Скользящего среднего процесс

Линейный фильтр 434

Логистическая кривая 94

Марковские оценки, см. Линейная регрессия

Матрица 308—311, см. также Квадратичные формы

диагонализация симметричной матрицы 309, 398

диагонализируемая 204, 205

жорданова каноническая форма 204, 205, 279

неравенства для характеристических корней 666, 667

характеристические корни и векторы 308, 309

полиномов от матриц 310

циркулянт 311

- эрмитова 426, 623
Метод переменных разностей
вычисление разностей 76—78
ковариация ряда, составленного из
конечных разностей 79, 80
критерий для проверки гипотез о сте-
пени тренда 89—94, 107
определение степени тренда 89—94
оценивание дисперсии ошибки 82—
88
Мультипликативная модель 74
- Наилучшие линейные несмещенные
оценки, см. Линейная регрессия
Найквиста частота 422, 423
Неймана — Пирсона фундаментальная
лемма 395, 396
Неймановская структура 52, 53, 99,
100, 291, 303, 304, 395
Нелинейная регрессия 94—97
Неравенство Канторовича 619
Коши — Шварца 451
- Обновление 456
Окно, см. Оценка спектральной плот-
ности (общая теория); Оценки спек-
тральной плотности (примеры)
Оператор вычисления разностей 77,
78, 104, 190
запаздывания 194
линейный 76, 77, 104
сдвига 76, 190
Ортогональные полиномы 28, 29, 45—
47, 98, 102, 103, см. также Полино-
миальная регрессия
Отношение к скользящему среднему 74
Оценка спектральной плотности (об-
щая теория)
взвешенного среднего 564—567
как плотность конечного процесса
скользящего среднего 656, 657
масштаб при построении графиков
596, 600
нормированная 595—600
случай известного среднего
асимптотическая дисперсия и ко-
вариация 573—579
асимптотическая нормальность
оценки и ее логарифма 581—590
асимптотическое смещение 567—
573
выражение с помощью квадратич-
ных форм 544—549
доверительные интервалы при
больших выборках 589
как линейная комбинация выбо-
рочных ковариаций 545—547, 550,
551
окна 551
приближенное распределение 590
состоятельность 567—573
случай неизвестного среднего
асимптотическая дисперсия и ко-
вариация 591—593
асимптотическая нормальность 594
асимптотическое смещение 590, 591
состоятельность 594
получение с помощью подбора вырав-
нивающего процесса авторегрессии
595, 596, 600, 601
строющаяся по остаткам от тренда
652—657
численные примеры 595—598, 693—
695, 702, 703, 711—713, 715
Оценки спектральной плотности (при-
меры)
Бартлетта 555, 556
Бартлетта модифицированная 556,
557
Блэкмена — Тьюки 558—560
выборочная спектральная плотность
551—555
Даниэля 557, 558
Парзена 561—563
прямоугольное 558
усеченная выборочная спектраль-
ная плотность 554, 555
усреднение по дискретным значени-
ям частоты 563, 564
характеристики окон 573, 578, 581
Хемминга 560, 561
Хеннинга 560
- Передаточная функция 435
Передаточная функция мощности 435
Период 15, 109
Периодограмма 124, см. также Выбо-
рочная спектральная плотности; Спе-
ктрограмма
Подмена частот 422, 423
Полиномиальная регрессия
критерии для отдельного коэффици-
ента 48
определение степени регрессии 48—
60
достаточные статистики 52
подобные области 51—55
ортогональные полиномы 28, 29,
45—47, 102, 103
оценки наименьших квадратов 45—47

- асимптотическая эффективность (для стационарной случайной составляющей) 632—636
 прогнозирование 57—60
 Предбеливание 594, 599
 Преобразование Фурье 116—119, см. также Быстрое преобразование Фурье; Тригонометрические функции
 Примеры статистического анализа временных рядов, см. также Индекс Бевериджа цен на пшеницу, Реализации трех процессов авторегрессии, Ряды солнечной активности
 индекс Доу — Джонса средних цен на акции 104
 население США 107
 отношения чисел голосов, поданных за республиканцев и демократов 381
 поступление масла на пяти рынках 122—125, 334
 потребление мясных продуктов в США 57—60, 94, 97, 98
 процентные ставки коммерческих бумаг 597, 598
 совокупный доход 281, 726, 727
- Реализации трех процессов авторегрессии**
 выборочные спектральные плотности 271, 272, 705—707, 711, 712
 дисперсии и корреляции 537, 702—704
 критерии порядка 272
 наблюдаемые ряды 696—701
 оценки спектральных плотностей 595—598, 708—712
 подобранные процессы авторегрессии 536, 537
 спектральные плотности 271, 272
 подобранных процессов 536, 537, 712, 713
- Регулярный процесс** 456
- Ряды солнечной активности**
 коэффициент корреляции 273, 274
 критерий порядка зависимости 274, 275
 оценка спектральной плотности 596, 598, 717
 спектрограмма 716, 717
 статистические данные 272, 273, 714, 715
- Ряды Фурье** 116—118, см. также Тригонометрические функции
 для ступенчатых функций 432
 сходимость 432
- Сглаживание** 60—76, см. также Линейные преобразования стационарных процессов
 дисперсии и ковариации сглаженных значений 62, 63, 66—68
 коэффициенты 65, 66, 80—82, 102, 103, 105
 с использованием оператора вычисления разностей 80—105
 с использованием полиномиальной регрессии 63—66
 устранение сезонных колебаний 72—76
 формулы Спенсера 71
- Сезонные изменения** 72—76
- Семинварианты четвертого порядка стационарного случайного процесса** 481, 538
 процесса скользящего среднего 506
- Сериальная корреляция (сериальный коэффициент корреляции)** 283, 284, 298
 моменты 356, 357, 364, 365
 приближенные распределения 372—381
 производящие функции моментов 322, 400
 распределение при двойных корнях 342—355
 использование характеристических функций 399
 распределение при одном простом корне 348, 349
 случай известного среднего 325—327, 331
 независимость от выборочного среднего 327
 строящаяся по остаткам от тренда 663—670
 неравенства для распределения 668—670
 распределение 665
- Сериальный коэффициент корреляции, см. Сериальная корреляция**
Сериальный коэффициент корреляции в модели, использующей последовательные разности
 приближенное распределение 377—381
 случай известного среднего 316—322
 случай неизвестного среднего
 моменты 366—368, 401
 определение 326
 производящая функция моментов 362
 распределение 355, 356, 400

- строящийся по остаткам от тренда (статистика Дурбина—Ватсона) 669, 670, 676
 Серийный коэффициент корреляции в циклической модели (циклический серийный коэффициент корреляции) определение циклической модели 389—392
 приближенное распределение 375—377
 случай известного среднего моменты 357, 368—371
 определения и канонические формы 311—316
 приближенная производящая функция моментов 375
 производящая функция моментов 362—364, 400, 401
 распределение 352—355, 399
 случай неизвестного среднего моменты 365, 366, 370, 401, 402
 определение 326
 производящая функция моментов 362, 364, 401
 распределение 352—355, 399
 распределение при зависимых наблюдениях 385—388
 совместное распределение 381—383
 таблица процентных точек 354
 основные распределения 383—385
 строящийся по остаткам от тригонометрического тренда моменты 662, 663
 определение 333, 334
 пример 334
 распределение 675
 таблица процентных точек 662
 частный 383—385, 403, 404
 Систематическая составляющая 13
 Скользящего среднего процесс 191
 как представление стационарного процесса общего вида
 для процесса с конечным числом отличных от нуля ковариаций 252, 253
 для регулярного процесса 459
 определяемое спектральной плотностью 435, 436
 ковариационная структура 251, 252
 представление
 в виде процесса авторегрессии бесконечного порядка 253
 с использованием линейных операторов 441, 442
 прогнозирование с минимальной среднеквадратичной ошибкой 442
 семинвариант четвертого порядка 505, 506
 спектральная плотность 435—442, 473
 характеристическое уравнение 252
 Спектральная плотность (стационарного процесса)
 векторного процесса 426
 интерпретация 421
 как преобразование Фурье ковариационной функции 415
 квадратурная 426
 конечного процесса скользящего среднего 438—442
 непрерывность 541
 коспектральная 426
 нормированная 424
 представление Фурье 436
 приближение спектральными плотностями процессов авторегрессии и скользящего среднего 446, 624—627
 простого скользящего среднего 449
 процесса авторегрессии 442—444
 с остатками в виде скользящего среднего 444, 445
 с ошибкой 473
 процесса скользящего среднего 435—437
 регулярного процесса 458
 ряда, получаемого вычислением разностей и сглаживанием 448
 ряда, составленного из разностей 447
 сглаженного ряда 448
 Спектральная функция
 дискретная 419, 420
 определение 416, 418
 подмена частот 422, 423
 примеры 418, 419
 разложение 421, 422
 свертывание спектра 422, 423
 Найквиста частота 423
 Спектральное представление стационарного процесса 431—434
 Спектр возрастающей матричной функции 632
 Спектрограмма 124, см. также Периодограмма; Выборочная спектральная плотность
 Статистические выводы для процессов авторегрессии, см. также Вероятностная структура процесса авторегрессии; Серийная корреляция
 доверительные интервалы при больших выборках 241
 достаточные статистики 211, 287, 288, 393
 критерии для коэффициентов 241—244

- критерии проверки порядка процесса 244—251, 272—276
 использующий частную корреляцию 250
 отношения правдоподобия 244
 χ^2 -критерии для больших выборок 244—249
 оценки максимального правдоподобия в скалярном случае
 асимптотическая нормальность 227, 230, 237, 238
 для стационарного процесса первого порядка 388, 389, 404
 модифицированные уравнения 213, 214
 нормальные уравнения 210
 оценка асимптотической ковариационной матрицы 238—240
 с использованием частных корреляций 214
 состоятельность 223, 224
 численные примеры 271—276, 281, 282
 оценки максимального правдоподобия для векторного процесса первого порядка 215—238
 Стационарный случайный процесс гауссовский 407, 409—411
 детерминированный 457
 ковариационная функция 199—201
 корреляционная функция 199
 определение (в узком смысле) 189, 409
 определение (в широком смысле) 191
 примеры 409—415, 469, 470
 разложение Вольда 456—461
 регулярный 457
 условие регулярности 459
 семинварианты четвертого порядка 481, 538
 спектральное представление 427—434
 с циклическими составляющими 410, 411
 Стирлинга числа второго рода 104
 Стохастические интегралы 427—430, 471, 472
 Стохастическое разностное уравнение, см. Вероятностная структура процесса авторегрессии
 Суммирование по Чезаро 498, 540, 541
 Суммы степеней натуральных чисел 98, 99
 Сходимость в среднем (в среднеквадратичном) 412, 450
- Таблицы процентных точек и распределений
 сериальный коэффициент корреляции, строящийся по последовательным разностям 379, 380, 670
 циклический сериальный коэффициент корреляции 354, 376, 662
 Теорема Вейерштрасса об аппроксимации 445
 Вольда о разложении 456—461
 Гаусса — Маркова 22, 39, 40
 центральная предельная, см. Центральная предельная теорема
 Тетраэдр правильный
 объем 398
 равномерное распределение на нем 343—348
 Тренд 15, 43, 44
 Тригонометрическая регрессия с периодами, не являющимися целыми делителями длины ряда 156, см. также Выборочные тригонометрические коэффициенты; Тригонометрическая регрессия с периодами, являющимися делителями длины ряда
 выборочные коэффициенты Фурье при известном среднем и независимых ошибках 157—165
 их выражение 172—175
 квадратичная форма от них 164, 165
 состоятельность и асимптотическая нормальность 180, 181, 186
 выборочные коэффициенты Фурье при неизвестном среднем и независимых ошибках 175—177
 оценки максимального правдоподобия (наименьших квадратов) параметров при произвольной частоте 179
 состоятельность и асимптотическая нормальность 180, 181
 функция интенсивности
 выборочная 162, 180—183
 теоретическая, когда тренд обладает единственной частотой 162, 163, 177, 178
 Тригонометрическая регрессия с периодами, являющимися целыми делителями длины ряда, см. также Тригонометрическая регрессия с периодами, не являющимися целыми делителями длины ряда
 доверительные множества для параметров 130—132
 инвариантность процедур 133, 134
 критерий Уолкера 137, 138

- Фишера 145—148
Шустера 135, 136
процедуры выбора положительных амплитуд
не более двух амплитуд 149—153
одной при известной дисперсии 140, 141
одной при неизвестной дисперсии 143—148
при известной дисперсии 137—140
среди части амплитуд при неизвестной дисперсии 143
решение о включении тригонометрических составляющих 133—154
Тригонометрические функции 15
в представлении Фурье конечной последовательности 112—115
в представлении Фурье периодической функции 116—118
ортogonalность 28, 110, 111, 116
Тригонометрический полином 445
- Фаза 15, 109
Фейера ядро 491, 499, 551
- Характеристическая функция 118, см. также Квадратичные формы
Харди класс H^2 461
 χ^2 -распределение нецентральное 134, 435
ограниченная полнота 186
- Центральная предельная теорема
для векторного стационарного процесса с конечной зависимостью 466
для линейного стационарного процесса 466—468
для независимых и одинаково распределенных случайных величин 464
для стационарного процесса с конечной зависимостью 464—466, 468, 469
Линдберга 463
Ляпунова 463, 464
Циклический сериальный коэффициент корреляции, см. Сериальный коэффициент корреляции в циклической модели
- Частная корреляция
выборочная сериальная 214, 250, 298, 383—385, 404
стационарного процесса 396
Частота 15, 109
Частотная характеристика 435
- Ширина спектра 533
Шум 13
- Эрмитова матрица 426, 623
- Юла — Уолкера уравнения 200

ОГЛАВЛЕНИЕ

От редактора перевода	5
Из предисловия автора	7
Глава 1. Введение	11
Литература	19
Глава 2. Использование регрессионного анализа	20
2.1. Введение	20
2.2. Общая теория наименьших квадратов	21
2.3. Линейные преобразования независимых переменных; ортогональные независимые переменные	25
2.4. Коррелированные переменные	30
2.5. Прогнозирование	32
2.6. Асимптотическая теория	35
Литература	39
Упражнения	39
Глава 3. Тренды и сглаживание	43
3.1. Введение	43
3.2. Полиномиальные тренды	44
3.3. Сглаживание	60
3.4. Метод переменных разностей	76
3.5. Нелинейные тренды	94
3.6. Обсуждение	97
Литература	98
Упражнения	98
Глава 4. Циклические тренды	108
4.1. Введение	108
4.2. Преобразования и представления	109
4.3. Статистические выводы для случая, когда периоды тренда являются делителями длины ряда	119

4.4. Статистические выводы для случая, когда периоды тренда не являются делителями длины ряда	156
4.5. Обсуждение	183
Литература	184
Упражнения	184
Глава 5. Линейные вероятностные модели с конечным числом параметров	189
5.1. Введение	189
5.2. Процессы авторегрессии	191
5.3. Редукция общего скалярного уравнения к векторному уравнению первого порядка	203
5.4. Оценки максимального правдоподобия в случае нормального распределения	209
5.5. Асимптотическое распределение оценок максимального правдоподобия	215
5.6. Статистические выводы о моделях авторегрессии, основанные на теории больших выборок	238
5.7. Модель скользящего среднего	251
5.8. Процесс авторегрессии с остатками в виде скользящего среднего	263
5.9. Некоторые примеры	271
5.10. Обсуждение	276
Литература	277
Упражнения	277
Глава 6. Сериальная корреляция	283
6.1. Введение	283
6.2. Типы моделей	286
6.3. Равномерно наиболее мощные критерии для проверки заданного порядка зависимости	290
6.4. Выбор порядка зависимости как задача со многими решениями	301
6.5. Модели: системы квадратичных форм	308
6.6. Случай, когда средние значения неизвестны	325
6.7. Распределения сериальных коэффициентов корреляции	334
6.8. Аппроксимация распределений сериальных коэффициентов корреляции	372
6.9. Совместные и условные распределения сериальных коэффициентов корреляции	381
6.10. Распределения для случая зависимых наблюдений	385
6.11. Оценки максимального правдоподобия	388
6.12. Обсуждение	392
Литература	393
Упражнения	393
Глава 7. Стационарные случайные процессы	406
7.1. Введение	406
7.2. Стационарные случайные процессы, определения и примеры	407
7.3. Спектральная плотность и спектральная функция	415
7.4. Спектральное представление стационарного случайного процесса	427

7.5. Линейные операции над стационарными процессами	434
7.6. Гильбертово пространство и теория прогнозирования	449
7.7. Некоторые предельные теоремы	461
Литература	469
Упражнения	469
 Глава 8. Выборочные среднее, ковариации и спектральная плотность . .	475
8.1. Введение	475
8.2. Определения выборочных среднего, ковариаций, спектральной плотности и их моментов	476
8.3. Асимптотические средние значения и ковариации выборочных среднего, ковариаций и спектральной плотности	497
8.4. Асимптотические распределения выборочных среднего, ковариаций и спектральной плотности	517
8.5. Примеры	536
8.6. Обсуждение	537
Литература	538
Упражнения	538
 Глава 9. Оценивание спектральной плотности	543
9.1. Введение	543
9.2. Оценки, основанные на выборочных ковариациях	544
9.3. Асимптотические средние и ковариации оценок спектральной плотности	564
9.4. Асимптотическая нормальность оценок спектральной плотности	581
9.5. Примеры	595
9.6. Обсуждение	598
Литература	601
Упражнения	601
 Глава 10. Линейные тренды и стационарные случайные составляющие . . .	608
10.1. Введение	608
10.2. Эффективное оценивание функций тренда	609
10.3. Оценивание ковариаций и спектральной плотности по остаткам от трендов	640
10.4. Проверка независимости	657
Литература	672
Упражнения	673
 Приложение А. Статистические данные	677
А.1. Индекс Бевериджа цен на пшеницу	677
А.2. Три процесса авторегрессии второго порядка, полученные с помощью случайных чисел	695
А.3. Числа солнечной активности	714
 Приложение В. Решения избранных упражнений	718
 Список литературы	735
 Предметный указатель	744

Т. Андерсон

СТАТИСТИЧЕСКИЙ АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

Редактор *И. А. Маховая*

Художник *Н. В. Зотова*

Художественный редактор *В. И. Шаповалов*

Технический редактор *Т. А. Максимова*

Корректор *В. С. Соколов*

Сдано в набор 17.09.75. Подписано к печати 21.06.76.

Бумага тип. № 2 $60 \times 90^{1/16} = 23,75$ бум. л. 47,50 печ. л.

Уч.-изд. л. 47,90. Изд. № 1/7571. Цена 3 р. 52 к. Зак. № 261

ИЗДАТЕЛЬСТВО «МИР»

Москва, 1-й Рижский пер., 2

Отпечатано в ордена Трудового Красного Знамени
Ленинградской типографии № 2 имени Евгении Соколовой

Союзполиграфпрома при Государственном комитете

Совета Министров СССР по делам издательств,

полиграфии и книжной торговли,

198052, Ленинград, Л-52, Измайловский проспект, 29,

с матриц Головного предприятия республиканского

производственного объединения «Полиграфкнига»

Госкомиздата УССР, г. Киев, ул. Довженко, 3