П. БИЛЛИНГСЛИ

СХОДИМОСТЬ ВЕРОЯТНОСТНЫХ МЕР

Перевод с английского
А. В. ПРОХОРОВА
под редакцией
В. В. САЗОНОВА

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
Москва 1977

517.8
Б 61
УДК 519.21

CONVERGENCE
OF PROBABILITY
MEASURES

PATRICK BILLINGSLEY

John Wiley & Sons, Inc.,
New York•London•Sydney•Toronto

Сходимость вероятностных мер. Биллингсли П. перев. с англ. Главная редакция физико-математической литературы издательства «Наука», М., 1977, 352 стр.

Книга посвящена общей теории слабой сходимости вероятностных мер в метрических пространствах. Развитые в последние пятнадцать лет методы изучения распределений (в частности, сходимости распределений) в функциональных пространствах оказались весьма плодотворными. Книга дает достаточно полное изложение материала. При этом автор на ряде примеров постарался продемонстрировать широкую применимость общих результатов в задачах теории стохастических процессов.

Илл. 6, библ. 112.

Б $\frac{20203-095}{053(02)-77}$ 32-77

ПРЕДИСЛОВИЕ РЕДАКТОРА ПЕРЕВОДА

Слабую сходимость вероятностных мер в ее простейших формах начали изучать давно — так, теорему Муавра — Лапласа можно рассматривать как некоторое утверждение о слабой сходимости вероятностных мер на прямой. И все же, современная теория слабой сходимости вероятностных мер — это молодая теория, насчитывающая лишь несколько десятилетий и связанная в первую очередь с именами А. Н. Колмогорова, Дж. Дуба, М. Донскера, Ю. В. Прохорова, А. В. Скорохода, а также Л. Ле Кама и С. Варадарайна. Автор настоящей книги П. Биллингсли не относится к числу основателей этой теории, однако он прекрасно в ней разобрался, успешно ее применял и изложил с присущим ему мастерством.

Первая глава книги посвящена общим вопросам слабой сходимости вероятностных мер в метрических пространствах (включая фундаментальную теорему Ю. В. Прохорова о равносильности плотности и относительной компактности семейств мер). Автор ведет изложение в той степени общности, которая, с одной стороны, достаточна для большинства приложений, а с другой — позволяет сосредоточиться на главном и не заниматься рассмотрением специальных вопросов, связанных с общими ситуациями. В результате часть теории осталась за пределами книги, однако заинтересованный читатель на основе знакомства с изложенным сможет легко сориентироваться и в соответствующей более специальной литературе, кое-что в этом плане содержится в помещенном в конце книги Добавлении III).

Далее изучаются вероятностные меры и их слабая сходимость в конкретных пространствах C (пространство непрерывных функций на $[0,1]$) и D (пространство функций на $[0,1]$, не имеющих разрывов второго рода).

Здесь, в частности, доказывается функциональная центральная предельная теорема (ее называют также теоремой Донскера или принципом инвариантности).

Заключительная глава посвящена более специальным вопросам — характеризации диффузионных процессов и, главным образом, доказательству функциональных центральных предельных теорем для некоторых последовательностей зависимых случайных величин (стационарных последовательностей с φ-перемешиванием, последовательностей функций от них; мартингалов). Многие результаты здесь принадлежат самому автору.

В книге имеется немало ярких и важных для приложений примеров применения общей теории. Так, автор не раз обращается к рассмотрению предельного поведения эмпирических функций распределения, изучается случайная замена времени и т. д.

Изложение довольно замкнуто в себе (этому, в частности, способствуют помещенные в конце книги добавления). Для понимания книги достаточно по существу знакомства с университетским курсом анализа и основами теории вероятностей. Первые три главы можно рекомендовать для первоначального ознакомления с предметом.

Нет сомненья, что предлагаемая в русском переводе книга П. Биллингсли найдет своего читателя и будет полезна аспирантам и студентам старших курсов математических специальностей, а также научным работникам — специалистам в области теории вероятностей и ее приложений.

В. В. Сазонов

ИЗ ПРЕДИСЛОВИЯ АВТОРА

Теоремы об асимптотических распределениях в теории вероятностей и математической статистике изначально основывались на классической теории слабой сходимости функций распределения в евклидовом пространстве, т. е. сходимости в точках непрерывности предельной функции распределения. Последние десятилетия отмечены созданием и широким применением более объемлющей теории слабой сходимости вероятностных мер в метрических пространствах. Имеется целый ряд асимптотических результатов, которые могут быть сформулированы в рамках классической теории, но требуют для доказательства этой более общей теории, развитие которой не является, таким образом, самоцелью. Эта книга — о методах слабой сходимости в метрических пространствах и их приложениях, показывающих силу и полезность этих методов.

Во введении разъясняются определения и показывается, как теория дает решение задач, возникающих за ее пределами. В главе 1 излагаются фундаментальные общие теоремы, которые затем применяются в главе 2 к пространству непрерывных функций, заданных на единичном интервале, и в главе 3 к пространству функций с разрывами первого рода. Результаты первых трех глав используются в главе 4 для вывода целого ряда предельных теорем для зависимых последовательностей случайных величин.

От читателя книги ожидается знакомство с обычной теорией вероятностей, основанной на теории меры, и с топологией метрических пространств, однако общая (неметрическая) топология не используется, а те немногие результаты, которые требуются из функционального анализа, доказываются в тексте и в добавлениях.

Преодолевая желание поместить примеры и применения после изложения общей теории и вынудить читателя тем самым упорно продолжать чтение до конца, я вместо этого распределил их равномерно по всей книге с тем, чтобы они иллюстрировали теорию по мере ее развития.

Я приношу благодарность Сёрен Иохансон, Самуэлю Карлину, Дэвиду Кендаллу, Роналду Пайку и Флемингу Топсё, которые прочитали большие части книги в рукописи; книга многим обязана их детальным предложениям, и я им очень признателен. Мне приятно также поблагодарить Мэри Вульридж за ее печатание, бодрое, быстрое и безошибочное.

Патрик Биллингсли

Чикаго, март 1968

ВВЕДЕНИЕ

Предельная теорема Муавра — Лапласа утверждает: если

$$F_n(x) = \mathsf{P}\left\{\frac{S_n - np}{\sqrt{npq}} \leqslant x\right\} \tag{1}$$

есть функция распределения нормированного числа «успехов» в n испытаниях Бернулли и если

$$F(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{1}{2}u^2}\, du \tag{2}$$

есть функция распределения стандартного нормального закона, то

$$F_n(x) \to F(x) \tag{3}$$

для всех x ($n \to \infty$, вероятность p «успеха» фиксирована).

Мы говорим о произвольных функциях распределения F_n и F на прямой, что F_n *слабо сходится* к F, и записываем $F_n \Rightarrow F$, если (3) выполняется для всех точек x непрерывности F. Таким образом, теорема Муавра — Лапласа утверждает, что (1) слабо сходится к (2); так как функция (2) всюду непрерывна, то в данном случае оговорка о точках непрерывности не нужна. Если F_n и F определяются равенствами

$$F_n(x) = \begin{cases} 0 & \text{при } x < 1/n, \\ 1 & \text{при } x \geqslant 1/n \end{cases} \tag{4}$$

и

$$F(x) = \begin{cases} 0 & \text{при } x < 0, \\ 1 & \text{при } x \geqslant 0, \end{cases} \tag{5}$$

то снова $F_n \Rightarrow F$, и здесь оговорка о точках непрерывности действительно необходима: (3) не выполняется в точке $x = 0$.

Для лучшего понимания феномена слабой сходимости, который составляет основу большого класса предельных теорем теории вероятностей, рассмотрим вероятностные меры P_n и P, порожденные произвольными функциями распределения F_n и F. Эти вероятностные меры, заданные на классе борелевских подмножеств прямой, однозначно определяются соотношениями

$$P_n(-\infty, x] = F_n(x), \qquad P(-\infty, x] = F(x).$$

Поскольку F непрерывна в x тогда и только тогда, когда множество $\{x\}$, состоящее из одной точки x, имеет P-меру нуль, то $F_n \Rightarrow F$ означает, что

$$P_n(-\infty, x] \to P(-\infty, x], \quad \text{если} \quad P\{x\} = 0. \tag{6}$$

Пусть ∂A обозначает границу множества A на прямой; ∂A состоит из тех точек, которые являются пределами последовательностей точек, принадлежащих A, и точек, не принадлежащих A. Так как граница множества $(-\infty, x]$ состоит всего из одной точки x, то (6) эквивалентно соотношению

$$P_n(A) \to P(A), \quad \text{если} \quad P(\partial A) = 0, \tag{7}$$

где символом A мы обозначили множество $(-\infty, x]$. Существо дела в том, что сходимость $F_n \Rightarrow F$ имеет место тогда и только тогда, когда (7) выполняется для *любого* борелевского множества A (результат, доказываемый в главе 1).

Будем выделять термином *множество P-непрерывности* такие борелевские множества A, для которых $P(\partial A) = 0$, и будем говорить, что P_n *слабо сходится* к P, и записывать $P_n \Rightarrow P$, если $P_n(A) \to P(A)$ для любого множества P-непрерывности (иными словами, если выполняется (7)). Как уже отмечалось выше, соотношение $P_n \Rightarrow P$ справедливо тогда и только тогда, когда для соответствующих функций распределения имеет место соотношение $F_n \Rightarrow F$.

Эта иная формулировка понятия слабой сходимости делает более ясной причину, по которой мы допускаем, что (3) может не выполняться, когда F имеет скачок в точке x. Без этого допущения (4) не будет слабо сходиться к (5), однако подобный пример может показаться искусственным. Если мы сосредоточим внимание на вероятностных мерах P_n и P, мы увидим, что соотношение $P_n(A) \to P(A)$ может не выполняться при $P(\partial A) > 0$ даже в теореме Муавра — Лапласа. Меры P_n и P, порожденные функциями распределения (1) и (2), удовлетворяют соотношениям

$$P_n(A) = \mathsf{P}\left\{\frac{S_n - np}{\sqrt{npq}} \in A\right\} \tag{8}$$

и

$$P(A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-\frac{1}{2}u^2}\, du \tag{9}$$

для борелевских множеств A. Если теперь A состоит из счетного множества точек

$$\frac{k - np}{\sqrt{npq}}, \quad n = 1, 2, \ldots, \quad k = 0, 1, \ldots, n,$$

то $P_n(A) = 1$ для всех n и $P(A) = 0$, так что соотношение $P_n(A) \to P(A)$ невозможно. Поскольку здесь ∂A есть вся действительная прямая, то это согласуется с (7).

Несмотря на то, что понятие слабой сходимости функций распределения «привязано» к действительной прямой (или евклидову пространству, во всяком случае), понятие слабой сходимости вероятностных мер может быть сформулировано для общих метрических пространств, что является реальным поводом для предпочтения этого последнего понятия. Пусть S — произвольное метрическое пространство, и пусть $\mathcal{S}$ — класс борелевских подмножеств S ($\mathcal{S}$ представляет собой σ-поле, порожденное открытыми множествами). Рассмотрим вероятностные меры P_n и P, заданные на $\mathcal{S}$. В точности, как прежде, мы определяем слабую сходимость $P_n \Rightarrow P$ требованием, чтобы (7) выполнялось для всех борелевских множеств A. В главе 1 мы изучим общую теорию этой сходимости и увидим, как она интерпретируется в различных конкретных метрических пространствах. Мы докажем, например, что P_n слабо сходится к P тогда и только тогда, когда

$$\int_S f\,dP_n \to \int_S f\,dP, \quad \text{если} \quad f \in C(S), \tag{10}$$

где $C(S)$ обозначает класс ограниченных непрерывных действительных функций на S. (Для согласования с общематематической традицией мы примем (10) в качестве определения слабой сходимости, так что (7) превратится из определения в необходимое и достаточное условие.)

Глава 2 посвящена слабой сходимости в пространстве $C = C[0,1]$ с равномерной топологией; C есть пространство всех непрерывных действительных функций на замкнутом единичном интервале $[0,1]$, метризованное расстоянием между двумя функциями $x = x(t)$ и $y = y(t)$:

$$\rho(x, y) = \sup_{0 \leqslant t \leqslant 1} |x(t) - y(t)|. \tag{11}$$

Приведенный в главе 2 в качестве приложения пример показывает полезность общей трактовки слабой сходимости — трактовки, выходящей за рамки классического евклидова случая. Пусть $\xi_1, \xi_2, \ldots$ — последовательность независимых одинаково распределенных случайных величин, определенных на некотором вероятностном пространстве $(\Omega, \mathscr{B}, \mathsf{P})$. Если ξ_n имеют среднее 0 и дисперсию σ^2, то в соответствии с центральной предельной теоремой Линдеберга — Леви распределения нормированных сумм

$$\frac{1}{\sigma\sqrt{n}} S_n = \frac{1}{\sigma\sqrt{n}} (\xi_1 + \ldots + \xi_n) \tag{12}$$

слабо сходятся при n, стремящемся к бесконечности, к нормальному распределению (9).

Мы можем сформулировать уточненный вариант центральной предельной теоремы, если докажем слабую сходимость распределений некоторых случайных функций, построенных по частичным суммам S_n. Для любого целого числа n и для любой выборочной точки ω построим на единичном интервале ломаную, которая линейна на каждом из подинтервалов $[(i-1)/n, i/n]$,

$i=1, 2, \ldots, n$, и принимает значение $S_i(\omega)/\sigma\sqrt{n}$ в точке i/n ($S_0(\omega)=0$) (рис. 1). Иными словами, построим функцию $X_n(\omega)$, значение которой в точке t из $[0,1]$ равно

$$X_n(t, \omega)=\frac{1}{\sigma\sqrt{n}}S_{i-1}(\omega)+\frac{t-(i-1)/n}{1/n}\frac{1}{\sigma\sqrt{n}}\xi_i(\omega),$$

$$\text{если } t\in\left[\frac{i-1}{n}, \frac{i}{n}\right]. \qquad (13)$$

При каждом ω функция $X_n(\omega)$ является элементом пространства C. Определим P_n — распределение $X_n(\omega)$ в C, полагая для борелевских подмножеств A пространства C (борелевских множеств относительно метрики (11))

$$P_n(A)=\mathsf{P}\{\omega: X_n(\omega)\in A\}.$$

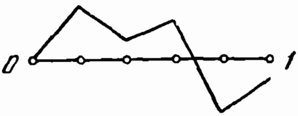

Рис. 1.

(Такое определение возможно вследствие того, что отображение $\omega\to X_n(\omega)$ оказывается измеримым в обычном понимании.) В главе 2 мы доказываем, что

$$P_n\Rightarrow W, \qquad (14)$$

где W — винеровская мера. Мы также доказываем существование в пространстве C винеровской меры, которая описывает распределение вероятностей для траектории частицы в броуновском движении.

Если $A=\{x: x(1)\leqslant a\}$, то, поскольку значение функции $X_n(\omega)$ при $t=1$ равно $X_n(1, \omega)=S_n(\omega)/\sigma\sqrt{n}$,

$$P_n(A)=\mathsf{P}\left\{\omega: \frac{1}{\sigma\sqrt{n}}S_n(\omega)\leqslant a\right\}.$$

Оказывается, что $W(\partial A)=0$, так что из (14) следует

$$\mathsf{P}\left\{\omega: \frac{1}{\sigma\sqrt{n}}S_n(\omega)\leqslant a\right\}\to W\{x: x(1)\leqslant a\}.$$

Оказывается, кроме того, что

$$W\{x: x(1)\leqslant \alpha\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\alpha}e^{-\frac{1}{2}u^2}\,du,$$

так что утверждение (14) на самом деле включает в себя теорему Линдеберга — Леви.

Сумма $\xi_1 + \ldots + \xi_n$ может быть интерпретирована как положение в момент времени n в случайном блуждании. Центральная предельная теорема утверждает, что это положение (надлежащим образом нормированное) при больших n распределено приближенно так же, как положение частицы в броуновском движении в момент времени $t = 1$. Соотношение (14) означает, что вся траектория случайного блуждания в течение первых n шагов при больших n распределена приближенно как траектория частицы в броуновском движении вплоть до момента времени $t = 1$ (рис. 2).

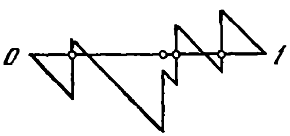

Рис. 2.

Чтобы конкретно убедиться в том, что (14) содержит информацию, выходящую за пределы центральной предельной теоремы, рассмотрим множество

$$A = \{x: \sup_{0 \leqslant t \leqslant 1} x(t) \leqslant \alpha\}.$$

Снова оказывается, что $W(\partial A) = 0$, так что из (14) следует

$$\lim_{n \to \infty} \mathsf{P}\left\{\omega: \frac{1}{\sigma\sqrt{n}} \max_{1 \leqslant k \leqslant n} S_k(\omega) \leqslant \alpha\right\} = \\ = \lim_{n \to \infty} P_n(A) = W\{x: \sup_{0 \leqslant t \leqslant 1} x(t) \leqslant \alpha\}. \quad (15)$$

Если мы оценим самую правую часть равенства (15) (это можно сделать несколькими способами, например, вычислением предела слева для некоторой специально выбранной последовательности $\{\xi_n\}$, для которой это легко сделать), то мы получим предельную теорему для распределения $\max_{k \leqslant n} S_k$ в предположениях теоремы Линдеберга — Леви.

В качестве последнего примера, относящегося к случайной функции $X_n(\omega)$, рассмотрим множество A тех x в C, для которых множество $\{t: x(t) > 0\}$ имеет меру Лебега, не превосходящую α (предполагается, что $0 \leqslant \alpha \leqslant 1$). Как и прежде, $P_n(A) \to W(A)$. Поскольку мера Лебега множества $\{t: X_n(t, \omega) > 0\}$ есть по существу доля тех частичных сумм $S_1, S_2, \ldots, S_n$, которые превосходят 0, то это рассуждение приводит к закону арксинуса (справедливому в предположениях теоремы Линдеберга — Леви). Глава 2 содержит детали всех этих выводов.

Мы можем, таким образом, использовать теорию слабой сходимости в C для получения целого класса предельных теорем для функций от частичных сумм $S_1, S_2, \ldots, S_n$. Тот факт, что винеровская мера W есть слабый предел распределений в C, соответствующих случайным функциям $X_n(\omega)$, может быть также использован для доказательства теорем относительно W, а мера W представляет самостоятельный интерес.

В главе 3 общая теория слабой сходимости прилагается к другому пространству функций на $[0, 1]$ — пространству $D = D\,[0,1]$ функций, не имеющих разрывов второго рода. Это пространство является естественным, например, для анализа поведения эмпирических функций распределения. Пусть $\xi_1, \xi_2, \ldots$ — независимые случайные величины на $(\Omega, \mathscr{B}, \mathsf{P})$ и каждая из них равномерно распределена на $[0,1]$. Пусть $F_n(t, \omega)$ — эмпирическая функция распределения $\xi_1(\omega), \ldots, \xi_n(\omega)$; $F_n(t, \omega)$ при $0 \leqslant t \leqslant 1$ равна доле целых k, $1 \leqslant k \leqslant n$, для которых $\xi_k(\omega) \leqslant t$. Пусть далее $Y_n(\omega)$ является элементом D, значение которого в каждой точке t равно

$$Y_n(t, \omega) = \sqrt{n}\,(F_n(t, \omega) - t).$$

Если D надлежащим образом метризовано, то позволительно говорить о распределениях случайных функций $Y_n(\omega)$ в D, и можно доказать, что эти распределения слабо сходятся при $n \to \infty$ к некоторому пределу. Точно так же, как и для случайного элемента пространства C, определенного равенством (13), мы можем затем вывести предельное распределение для случайной величины

$$\sup_{0 \leqslant t \leqslant 1} \sqrt{n}\,(F_n(t, \omega) - t) = \sup_{0 \leqslant t \leqslant 1} Y_n(t, \omega)$$

и ряда других аналогичных величин, которые появляются в математической статистике.

Глава 4 посвящена слабой сходимости распределений случайных функций, определяемых через различные последовательности зависимых случайных величин. Многие выводы глав 2, 3 и 4, хотя и не требуют понятия функциональных пространств для своей формулировки, вряд ли могли быть получены без применения методов функционального анализа.

Стандартный теоретико-множественный подход к определению вероятности и топология метрических пространств используются в книге с самого начала. Несмотря на то, что повсюду в книге господствует точка зрения функционального анализа (функция есть точка пространства), никакие сведения из функционального анализа не предполагаются известными (кроме изначальной готовности рассматривать функцию как точку пространства); все необходимые результаты из функционального анализа доказываются в тексте или же в Добавлении I (в котором также для облегчения ссылок собраны вместе некоторые результаты из топологии метрических пространств).

Примечание. Основными работами, которые привели к развитию этой теории, являются: Колмогоров (1931), Эрдеш и Кац (1946 и 1947), Дуб (1956) и Донскер (1951 и 1952). Прохоров (1953 и 1956) и Скороход (1956) придали теории современный вид. Ле Кам (1960) и Варадарайн (1958а и 1961а) распространили ее на общие топологические пространства.

Глава 1

СЛАБАЯ СХОДИМОСТЬ В МЕТРИЧЕСКИХ ПРОСТРАНСТВАХ

§ 1. Меры в метрических пространствах

Пусть S — метрическое пространство. Мы будем изучать вероятностные меры, определенные на классе $\mathcal{S}$ борелевских множеств пространства S. Класс $\mathcal{S}$ представляет собой σ-поле, порожденное открытыми множествами (т. е. наименьшее σ-поле, содержащее все открытые множества). Вероятностная мера на $\mathcal{S}$ представляет собой неотрицательную счетно аддитивную функцию множеств P такую, что $P(S) = 1$.

Если вероятностные меры P_n и P удовлетворяют соотношению $\int\limits_S f\,dP_n \to \int\limits_S f\,dP$ для каждой действительной ограниченной непрерывной функции f на S, мы будем говорить, что P_n *слабо сходится* к P, и будем писать $P_n \Rightarrow P$. В этой главе мы ставим целью детальное изучение этого понятия; мы начнем с изучения некоторых свойств вероятностных мер на $(S, \mathcal{S})$.

Хотя нам придется иногда предполагать сепарабельность или полноту, большинство теорем этой главы справедливы для произвольного метрического пространства S. Пространства, появляющиеся в наших приложениях, обычно сепарабельны и полны; поскольку они редко обладают дополнительными свойствами регулярности, такими как локальная компактность, мы никогда не будем налагать дальнейших ограничений.

Меры и интегралы. Теорема 1.1. *Каждая вероятностная мера на $(S, \mathcal{S})$ регулярна: если $A \in \mathcal{S}$ и $\varepsilon > 0$, то существует замкнутое множество F и открытое множество G такие, что $F \subset A \subset G$ и $P(G - F) < \varepsilon$.*

Доказательство. Обозначим метрику на S через $\rho(x, y)$ и расстояние от x до A через $\rho(x, A)$ *). Если A замкнуто, то можно принять $F = A$ и $G = \{x: \rho(x, A) < \delta\}$ при некотором δ, поскольку последние множества убывают при $\delta \downarrow 0$ и в пересечении дают A. Поэтому нам достаточно показать, что класс $\mathcal{G}$ борелевских множеств с объявленным свойством является σ-полем **). Пусть даны множества A_n из $\mathcal{G}$; выберем замкнутые множества F_n и открытые множества G_n такие, что $F_n \subset A_n \subset G_n$ и $P(G_n - F_n) < \varepsilon/2^{n+1}$. Если $G = \bigcup_n G_n$ и если $F = \bigcup_{n \leqslant n_0} F_n$, где n_0 выбрано так, что $P\left(\bigcup_n F_n - F\right) < \varepsilon/2$, тогда $F \subset \bigcup_n A_n \subset G$ и $P(G - F) < \varepsilon$. Таким образом, $\mathcal{G}$ является классом множеств, замкнутым относительно операции образования счетных объединений; так как класс $\mathcal{G}$, очевидно, замкнут относительно взятия дополнений, то теорема полностью доказана.

Из теоремы 1.1 вытекает, что мера P однозначно определена значениями $P(F)$ для замкнутых множеств F. Теорема 1.3 показывает, что мера P однозначно определена значениями $\int f\,dP$ ***) для действительных ограниченных непрерывных функций f, заданных на S. Обозначим через $C(S)$ класс таких функций f. На стр. 304 ****) показано, что каждая f из $C(S)$ $\mathcal{S}$-измерима. Доказательство теоремы 1.3 использует следующую теорему, которая показывает, как аппроксимировать индикатор (или характеристическую функцию) I_F замкнутого множества F элементами $C(S)$.

---

*) По поводу терминологии и некоторых теорем, касающихся метрических пространств, см. Добавление I.

**) Мы определили класс $\mathcal{S}$ борелевских множеств как σ-поле, порожденное открытыми множествами; это то же самое, что и σ-поле, порожденное замкнутыми множествами. Это σ-поле наиболее пригодно для целей данной теории. По поводу других (по большей части неподходящих) σ-полей см. задачу 6.

***) В случае, если интеграл берется по всему пространству, мы не указываем области интегрирования.

****) Эта страница принадлежит Добавлению II, где рассматривается бо́льшая часть вопросов об измеримости.

Теорема 1.2. *Если F замкнуто, а ε положительно, то существует функция f в $C(S)$ такая, что $f(x)=1$ при $x\in F$, $f(x)=0$ при $\rho(x,F)\geqslant\varepsilon$ и $0\leqslant f(x)\leqslant 1$ для всех x. Функция f может быть выбрана равномерно непрерывной.*

Доказательство. Определим непрерывную функцию φ действительного переменного, полагая

$$\varphi(t)=\begin{cases} 1 & \text{при } t\leqslant 0,\\ 1-t & \text{при } 0\leqslant t\leqslant 1,\\ 0 & \text{при } 1\leqslant t.\end{cases} \tag{1.1}$$

Если

$$f(x)=\varphi\left(\frac{1}{\varepsilon}\rho(x,\ F)\right), \tag{1.2}$$

то функция f обладает требуемыми свойствами — она даже равномерно непрерывна. Рис. 3 изображает эту функцию f для множества $F=[a,b]$ на прямой.

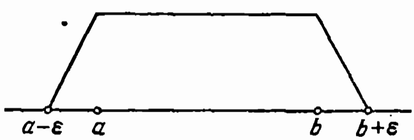

Рис. 3.

Теорема 1.3. *Вероятностные меры P и Q на $(S,\mathscr{S})$ совпадают, если*

$$\int f\,dP=\int f\,dQ \tag{1.3}$$

для любой функции f из $C(S)$.

Доказательство. Предположим, что множество F замкнуто. Будем исходить из (1.1) и определим для любого положительного целого числа u

$$\varphi_u(t)=\varphi(ut) \tag{1.4}$$

и

$$f_u(x)=\varphi_u(\rho(x,\ F)). \tag{1.5}$$

Тогда $\{f_u\}$ — невозрастающая последовательность элементов $C(S)$, сходящаяся поточечно к I_F. В силу теоремы об ограниченной сходимости $P(F)=\lim_u\int f_u\,dP$ и

$Q(F) = \lim_u \int f_u \, dQ$, так что если равенство (1.3) справедливо для всех f из $C(S)$, то $P(F) = Q(F)$. Поскольку меры P и Q совпадают на всех замкнутых множествах, из теоремы 1.1 следует, что P и Q тождественны.

Таким образом, значения $\int f \, dP$ для f из $C(S)$ полностью определяют значения $P(A)$ для A из $\mathscr{S}$. Этот факт лежит в основе круга идей, группирующихся вокруг понятия слабой сходимости; хотя мы определили слабую сходимость требованием сходимости интегралов от функций из $C(S)$, в следующем параграфе мы будем ее характеризовать в терминах сходимости значений мер некоторых множеств.

Плотность. Нижеследующее понятие плотности оказывается важным как в теории слабой сходимости, так и в ее применениях. Вероятностная мера P на $(S, \mathscr{S})$ *плотна*, если для каждого положительного ε существует компактное (стр. 296) множество K такое, что $P(K) > 1 - \varepsilon$. Ясно, что мера P плотна тогда и только тогда, когда она обладает σ-компактным носителем*). Согласно теореме 1.1 мера P плотна в том и только том случае, если для любого A из $\mathscr{S}$ $P(A)$ является точной верхней гранью $P(K)$ по компактным подмножествам K множества A.

В σ-компактном пространстве каждая вероятностная мера плотна (это верно, в частности, для k-мерного евклидова пространства). Более полезным является нижеследующий результат, включающий также случай евклидова пространства.

Теорема 1.4. *Если S сепарабельно и полно, то каждая вероятностная мера на $(S, \mathscr{S})$ плотна.*

Доказательство. Поскольку S сепарабельно, то для любого n существует последовательность $A_{n1}, A_{n2}, \ldots$ открытых $1/n$-сфер, покрывающих S. Выберем

---

*) Носителем вероятностной меры называется множество A из $\mathscr{S}$ такое, что $P(A) = 1$; множество σ-компактно, если оно может быть представлено как счетное объединение компактных множеств. Характеризация плотной меры P, как меры, обладающей σ-компактным носителем, не годится в качестве определения, ибо оно не может быть обобщено надлежащим образом на семейства вероятностных мер (см. § 6).

i_n так, чтобы $P\left(\bigcup_{i \leqslant i_n} A_{ni}\right) > 1 - \varepsilon/2^n$. В силу предположения о полноте вполне ограниченное множество $\bigcap_{n \geqslant 1} \bigcup_{i \leqslant i_n} A_{ni}$ обладает компактным замыканием K (см. стр. 296). Поскольку, очевидно, $P(K) > 1 - \varepsilon$, то теорема доказана.

Теорема 1.4 неверна без предположения о полноте; вопрос о том, можно ли избавиться от предположения сепарабельности, равносилен «проблеме меры». Эта тема обсуждается в Добавлении III *).

Примечание. Теорема 1.4 принадлежит Уламу (см. Окстоби и Улам (1939); Ле Кам (1960) ввел термин «плотность» (tight)).

Задачи **)

1. Скажем, что функция f разделяет множества A и B, если $f(x) = 0$ для x из A, $f(x) = 1$ для x из B и $0 \leqslant f(x) \leqslant 1$ для всех x. Если расстояние между A и B положительно, эти множества могут быть разделены равномерно непрерывной функцией f [теорема 1.2]. Если A и B имеют непересекающиеся замыкания, но расстояние между ними равно 0, то они могут быть разделены непрерывной функцией f [$f(x) = \rho(x, A)/\rho(x, A) + \rho(x, B))$], но не могут быть разделены равномерно непрерывной функцией. Не существует непрерывной функции f, разделяющей A и B в случае, если их замыкания пересекаются; не существует функции f, разделяющей A и B, если они пересекаются.

2. Приведите примеры различных топологий, которые приводят к одному и тому же классу борелевских множеств.

3. Если S можно вложить как открытое множество в некоторое полное метрическое пространство, то [Келли (1968, стр. 275)] оно топологически полно. Так как локально компактное пространство S является открытым в своем пополнении, то такое пространство топологически полно. Следовательно, теорема 1.4 применима, если пространство S сепарабельно и локально компактно. Такое пространство

---

*) Хотя теорема 1.4 в том виде, в котором она дана, достаточна для всех приложений, содержащихся в этой книге, естественно поинтересоваться ее возможными обобщениями. Добавление III как раз и посвящено вопросам этого рода.

**) Некоторые задачи включают понятия, которые не требуются для понимания основного текста; задачи, решения которых были бы использованы далее в тексте, не даются. Простое утверждение следует понимать в смысле «покажите, что». Квадратные скобки содержат намеки или указания для решений.

S является σ-компактным [будучи объединением открытых множеств с компактными замыканиями и, следовательно (стр. 295), счетным объединением таких множеств]. Отсюда непосредственно вытекает, что любая вероятностная мера на нем плотна; примером является евклидово пространство.

4. Пусть S — гильбертово пространство со счетным ортонормальным базисом $x_1, x_2, \ldots$ Так как S сепарабельно и полно, то теорема 1.4 применима. Однако любое множество, имеющее внутренние точки, не является компактом [такое множество должно для некоторых x и ε содержать все точки вида $x + \varepsilon x_n$], так что S не является ни локально компактным, ни [теорема Бэра о категории; Келли (1968, стр. 267)] σ-компактным. Если мера P приписывает положительную массу каждому элементу счетного всюду плотного множества, то P не обладает локально компактным носителем в относительной топологии.

5. Приспособьте задачу 4 к общему банахову пространству счетного числа измерений [существуют точки $x_1, x_2, \ldots$ с $\sup_n \|x_n\| < \infty$ и $\inf_{m \neq n} \|x_m - x_n\| > 0$; см. Банах (1932, стр. 83)]; пространство $C[0, 1]$, важное в теории вероятностей, является примером такого пространства; это объясняет тот факт, что теория, основанная на локальной компактности, мало полезна в данном круге вопросов. (См. также задачу 5 в § 3.)

6. Мы определили $\mathscr{S}$ как σ-поле, порожденное открытыми множествами, что можно обозначить записью $\mathscr{S} = \sigma$ {открытые множества}. Подобно этому определим $\mathscr{S}_1 = \sigma$ {замкнутые G_δ-множества} (множество является G_δ-множеством, если оно представляется счетным пересечением открытых множеств), определим $\mathscr{S}_2 = \sigma\{C(S)\}$ (наименьшее σ-поле, по отношению к которому любая функция из $C(S)$ измерима) и определим $\mathscr{S}_3 = \sigma${открытые сферы}, $\mathscr{S}_4 = \sigma${компактные множества} и $\mathscr{S}_5 = \sigma${компактные G_δ-множества}. В метрическом пространстве каждое замкнутое множество есть множество типа G_δ. Используйте последний факт и теорему 1.2 для доказательства того, что

$$\mathscr{S} = \mathscr{S}_1 = \mathscr{S}_2 \supset \mathscr{S}_3 \supset \mathscr{S}_4 = \mathscr{S}_5.$$

Покажите, что $\mathscr{S} = \mathscr{S}_3$, если S сепарабельно. Покажите, что $\mathscr{S} = \mathscr{S}_5$, если S σ-компактно (это так, если S сепарабельно и локально компактно). Возможно, что $\mathscr{S}_2 \neq \mathscr{S}_3$ (даже при том, что S локально компактно): возьмите S дискретным, но не счетным. Возможно, что $\mathscr{S}_3 \neq \mathscr{S}_4$ (даже если S сепарабельно и полно): возьмите в качестве S гильбертово пространство из задачи 4. (Иное положение в общем топологическом пространстве, где следует рассматривать два класса множеств: борелевские множества — их считают иногда элементами $\mathscr{S}$, а иногда — $\mathscr{S}_4$, и бэровские множества — их считают иногда элементами $\mathscr{S}_2$, а иногда — $\mathscr{S}_5$; терминология неоднозначна.)

7. В связи с понятием плотности интересно отметить следующий факт: допустим, что функция множеств P определена на $(S, \mathscr{S})$, и допустим лишь, что она конечно аддитивна. Если для любого A из $\mathscr{S}$ имеем $P(A) = \sup P(K)$, где K пробегает компактные подмножества A, то P непременно счетно аддитивна.

§ 2. Свойства слабой сходимости

Мы определили слабую сходимость $P_n \Rightarrow P$ как сходимость интегралов $\int f\,dP_n \to \int f\,dP$ для любой функции f из класса $C(S)$ действительных ограниченных непрерывных функций на S. Заметим, что так как интегралы $\int f\,dP$ полностью определяют P (теорема 1.3), последовательность $\{P_n\}$ не может слабо сходиться одновременно к двум различным пределам. Заметим также, что слабая сходимость зависит только от топологии S, но не от вида метрики, ее порождающей: две метрики, порождающие одну и ту же топологию, приводят к одинаковым классам $\mathcal{S}$ и $C(S)$ и, следовательно, к одному и тому же понятию слабой сходимости*).

Теорема об условиях слабой сходимости. В следующей теореме предлагаются полезные условия, эквивалентные слабой сходимости; любое из этих условий может служить определением. Множество A из $\mathcal{S}$, граница которого ∂A удовлетворяет соотношению $P(\partial A) = 0$, называется *множеством непрерывности P* или *P-непрерывным множеством* (∂A замкнуто и поэтому принадлежит $\mathcal{S}$).

Теорема 2.1. *Пусть P_n и P — вероятностные меры, определенные на $(S, \mathcal{S})$. Следующие пять условий эквивалентны:*

(i) $P_n \Rightarrow P$.

(ii) $\lim_n \int f\,dP_n = \int f\,dP$ *для всех ограниченных равномерно непрерывных действительных функций f.*

(iii) $\limsup_n P_n(F) \leqslant P(F)$ *для всех замкнутых F.*

(iv) $\liminf_n P_n(G) \geqslant P(G)$ *для всех открытых G.*

(v) $\lim_n P_n(A) = P(A)$ *для всех P-непрерывных множеств A.*

---

*) Введем в пространстве $Z(S)$ всех вероятностных мер на $(S, \mathcal{S})$ топологию, взяв в качестве базы окрестностей меры P множества: совокупность таких мер Q, что $\left|\int f_i\,dP - \int f_i\,dQ\right| < \varepsilon$ для $i = 1, \ldots, k$, где ε положительно и f_i принадлежат $C(S)$. Тогда слабая сходимость будет равносильна сходимости в этой топологии. Топологическая структура пространства $Z(S)$ (которая сама по себе для нас не важна) рассматривается в Добавлении III.

Следующие два примера поясняют значение этих условий. Пусть P — единичная масса в точке x (значение $P(A)$ равно 1 или 0 в соответствии с тем, лежит ли x в A, или нет *), и пусть P_n — единичная масса в x_n. Если $x_n \to x$, то $\int f\,dP_n = f(x_n) \to f(x) = \int f\,dP$ для всех f из $C(S)$, так что $P_n \Rightarrow P$. Если x_n не сходится к x, то для некоторого положительного числа ε и бесконечно многих n мы имеем $\rho(x_n, x) > \varepsilon$. Если $f(y) = \varphi(\varepsilon^{-1}\rho(x, y))$, где функция φ определена соотношением (1.1), то $f \in C(S)$, $f(x) = 1$ и для бесконечно многих n $f(x_n) = 0$; следовательно, P_n не может сходиться слабо к P. Таким образом, $P_n \Rightarrow P$ в том и только том случае, если $x_n \to x$. Этот пример мы будем часто использовать в дальнейшем. (Многие гипотетические утверждения о слабой сходимости, которые в действительности неверны, могут быть опровергнуты с помощью соответствующей специализации данного примера.) Поскольку A является множеством непрерывности P тогда и только тогда, когда $x \notin \partial A$, легко проверить в этом случае эквивалентность (i) и (v). Если $x_n \to x$, но все x_n отличаются от x, то имеется строгое неравенство в (iii) для $F = \{x\}$ и строгое неравенство в (iv) для дополнительного множества $G = F^c$, кроме того, если x_n все различны и $A = \{x_2, x_4, \ldots\}$, то $P_n(A)$ не сходится ни к $P(A)$, ни к какому-либо другому пределу.

На прямой с обычной метрикой теорема Муавра - Лапласа также иллюстрирует условия теоремы. В качестве более простого, но тоже весьма важного примера рассмотрим меру P_n, соотносящую массу $1/n$ каждой точке i/n, $i = 1, 2, \ldots, n$. Тогда P_n сходится слабо к мере Лебега P, сосредоточенной на единичном интервале (это следует из того, что $\int f\,dP_n$ представляет собой сумму, аппроксимирующую интеграл $\int f\,dP$, который рассматривается как интеграл Римана). Если A состоит из рациональных точек, то $P_n(A) = 1$ не сходится к $P(A) = 0$; если G — открытое множество, содержащее рациональные точки и имеющее лебегову меру, близкую

*) Предполагается, что каждое подмножество, о котором идет речь, принадлежит $\mathcal{S}$.

к 0, тогда в (iv) имеет место строгое неравенство. Мы докажем теорему 2.1, устанавливая импликации в следующей диаграмме:

$$\begin{tikzcd}
(\mathrm{i}) \arrow[r] & (\mathrm{ii}) \arrow[r] & (\mathrm{iii}) \arrow[r, leftrightarrow] \arrow[d, leftrightarrow] \arrow[ll, bend right] & (\mathrm{iv}) \\
 & & (\mathrm{v}) &
\end{tikzcd}$$

Разумеется, импликация (i) $\to$ (ii) тривиальна.

Доказательство (ii) $\to$ (iii). Допустим, что (ii) справедливо, множество F замкнуто и пусть $\delta > 0$. При достаточно малом ε множество $G = \{x: \rho(x, F) < \varepsilon\}$ удовлетворяет неравенству $P(G) < P(F) + \delta$, поскольку множества такого вида убывают к F при $\varepsilon \downarrow 0$. Если функция $f(x)$ определена формулой (1.2), то она равномерно непрерывна на S, $f(x) = 1$ на F, $f(x) = 0$ на дополнении G^c к G и $0 \leqslant f(x) \leqslant 1$ для все x. Поскольку (ii) выполняется, мы имеем $\lim_n \int f\,dP_n = \int f\,dP$, что вместе с соотношениями

$$P_n(F) = \int_F f\,dP_n \leqslant \int f\,dP_n$$

и

$$\int f\,dP = \int_G f\,dP \leqslant P(G) < P(F) + \delta$$

дает

$$\limsup_n P_n(F) \leqslant \lim_n \int f\,dP_n = \int f\,dP < P(F) + \delta.$$

Так как δ было произвольным, отсюда следует (iii).

Доказательство (iii) $\to$ (i). Допустим, что (iii) справедливо и что $f \in C(S)$. Мы сначала покажем, что

$$\limsup_n \int f\,dP_n \leqslant \int f\,dP. \tag{2.1}$$

Преобразуя f линейно (с положительным коэффициентом при члене с первой степенью), мы можем свести нашу проблему к случаю, когда $0 < f(x) < 1$ для всех x. Пусть при целом k (временно фиксированном) F_i

будет замкнутым множеством: $F_i = \{x: i/k \leqslant f(x)\}$, $i = 0, 1, \ldots, k$. Так как $0 < f(x) < 1$, то мы имеем

$$\sum_{i=1}^{k} \frac{i-1}{k} P\left\{x: \frac{i-1}{k} \leqslant f(x) < \frac{i}{k}\right\} \leqslant$$
$$\leqslant \int f\,dP < \sum_{i=1}^{k} \frac{i}{k} P\left\{x: \frac{i-1}{k} \leqslant f(x) < \frac{i}{k}\right\}.$$

Сумма в правой части равна

$$\sum_{i=1}^{k} \frac{i}{k} [P(F_{i-1}) - P(F_i)] = \frac{1}{k} + \frac{1}{k} \sum_{i=1}^{k} P(F_i).$$

Это и аналогичное преобразование суммы в левой части дают неравенство

$$\frac{1}{k} \sum_{i=1}^{k} P(F_i) \leqslant \int f\,dP < \frac{1}{k} + \frac{1}{k} \sum_{i=1}^{k} P(F_i). \tag{2.2}$$

Если (iii) справедливо, то $\limsup_n P_n(F_i) \leqslant P(F_i)$ при каждом i, и поэтому (применяем правое неравенство в (2.2) к P_n, а левое неравенство — к P)

$$\limsup_n \int f\,dP_n \leqslant \frac{1}{k} + \int f\,dP.$$

Полагая $k \to \infty$, мы получаем (2.1).

Применяя (2.1) к $-f$, получим $\liminf_n \int f\,dP_n \geqslant$ $\geqslant \int f\,dP$, что вместе с (2.1) доказывает слабую сходимость.

Эквивалентность (iii) и (iv) легко доказывается путем перехода к дополнениям.

Доказательство (iii) $\to$ (v). Обозначим через A° внутренность множества A, а через A^- — его замыкание. Если (iii) выполняется, то справедливо также (iv) и, следовательно, для любого A:

$$P(A^-) \geqslant \limsup_n P_n(A^-) \geqslant \limsup_n P_n(A) \geqslant$$
$$\geqslant \liminf_n P_n(A) \geqslant \liminf_n P_n(A^\circ) \geqslant P(A^\circ). \tag{2.3}$$

Если $P(\partial A) = 0$, тогда крайние члены этого неравенства равны $P(A)$, и мы получаем $\lim_n P_n(A) = P(A)$.

Доказательство (v)$\to$(iii). $\partial\{x: \rho(x, F) \leqslant \delta\}$ содержится *) в $\{x: \rho(x, F) = \delta\}$, так что эти границы при различных δ не пересекаются, и поэтому не более чем счетное множество их имеет положительную P-меру. Следовательно, для некоторой последовательности положительных δ_k, сходящейся к 0, множества $F_k = \{x: \rho(x, F) \leqslant \delta_k\}$ являются множествами непрерывности P. Если (v) выполняется, то $\limsup_n P_n(F) \leqslant \lim_n P_n(F_k) = P(F_k)$ для каждого k; если F замкнуто, то $F_k \downarrow F$, так что (iii) выполняется. Этим завершается доказательство теоремы 2.1.

Другие критерии. Иногда удобно доказывать слабую сходимость, проверяя сходимость $P_n(A) \to P(A)$ для некоторого специального класса множеств A.

Теорема 2.2. *Пусть $\mathcal{U}$ — подкласс класса $\mathcal{S}$ такой, что* (i) *$\mathcal{U}$ замкнут относительно взятия конечных пересечений и* (ii) *каждое открытое множество в S является конечным или счетным объединением элементов $\mathcal{U}$. Если $P_n(A) \to P(A)$ для каждого A из $\mathcal{U}$, то $P_n \Rightarrow P$.*

Доказательство. Если $A_1, \ldots, A_m$ принадлежат $\mathcal{U}$, то их пересечение также принадлежит $\mathcal{U}$; следовательно, по известной формуле включения-исключения для вероятности объединения событий

$$P_n\left(\bigcup_{i=1}^{m} A_i\right) = \sum_i P_n(A_i) - \sum_{ij} P_n(A_iA_j) + \sum_{ijk} P_n(A_iA_jA_k) - \ldots \to \sum_i P(A_i) - \sum_{ij} P(A_iA_j) + \sum_{ijk} P(A_iA_jA_k) - \ldots = P\left(\bigcup_{i=1}^{m} A_i\right).$$

Если G — открытое множество, то $G = \bigcup_i A_i$ для некоторой последовательности $\{A_i\}$ элементов $\mathcal{U}$. Задавшись числом ε, выберем m так, чтобы $P\left(\bigcup_{i \leqslant m} A_i\right) > P(G) - \varepsilon$.

*) Включение может быть строгим — например, в дискретном пространстве.

Из только что доказанного равенства

$$P(G)-\varepsilon<P\Big(\bigcup_{i\leqslant m}A_i\Big)=\lim_n P_n\Big(\bigcup_{i\leqslant m}A_i\Big)\leqslant\liminf_n P_n(G).$$

Так как ε было произвольным, условие (iv) предыдущей теоремы выполнено.

Пусть $S(x, \varepsilon)$ обозначает (открытую) ε-сферу с центром в точке x.

Следствие 1. *Пусть $\mathscr{U}$ — класс множеств такой, что* (i) *$\mathscr{U}$ замкнут относительно взятия конечных пересечений и* (ii) *для любого элемента x из S и любого положительного ε существует множество A в $\mathscr{U}$ такое, что $x\in A^\circ\subset A\subset S(x, \varepsilon)$. Если S сепарабельно и если $P_n(A)\to P(A)$ для любого A из $\mathscr{U}$, то $P_n\Rightarrow P$.*

Доказательство. Из условия (ii) *) следует, что для любой точки x открытого множества G имеет место включение $x\in A^\circ\subset A\subset G$ для некоторого A из $\mathscr{U}$. Поскольку S сепарабельно, в $\mathscr{U}$ существует (см. стр. 295) конечная или бесконечная последовательность $\{A_i\}$ такая, что $G\subset\bigcup_i A_i^\circ$ и $A_i\subset G$, откуда $G=\bigcup_i A_i$. Таким образом, $\mathscr{U}$ удовлетворяет предположениям теоремы 2.2.

Следствие 2. *Предположим, что для каждого конечного пересечения A открытых сфер мы имеем $P_n(A)\to P(A)$ при условии, что A — множество непрерывности P. Если S сепарабельно, то $P_n\Rightarrow P$.*

Доказательство. Границы $\partial S(x, \varepsilon)$, принадлежащие множествам $\{y: \rho(x, y)=\varepsilon\}$, не пересекаются (при фиксированном x), и поэтому имеют P-меру 0 при всех ε, за исключением конечного или счетного их числа. Из соотношения $\partial(A\cap B)\subset(\partial A)\cup(\partial B)$ вытекает, что класс $\mathscr{U}$ множеств непрерывности P, которые являются конечными пересечениями сфер, удовлетворяет предположениям следствия 1, откуда и следует нужный результат.

Условимся называть подкласс $\mathscr{V}$ класса $\mathscr{S}$ *определяющим сходимость классом*, если сходимость $P_n(A)\to P(A)$ для всех множеств P-непрерывности A из $\mathscr{V}$

---

*) Это условие несколько сильнее, чем условие, требующее, чтобы внутренности элементов $\mathscr{U}$ образовывали базу топологии пространства S.

всегда влечет за собой слабую сходимость P_n к P. Следствие 2 можно теперь сформулировать следующим образом: в сепарабельном пространстве конечные пересечения сфер образуют класс, определяющий сходимость.

Условимся называть в дальнейшем класс $\mathscr{V}$ *определяющим классом*, если P и Q тождественны всякий раз, когда их значения на $\mathscr{V}$ совпадают. Класс замкнутых множеств является определяющим классом, и таково же любое поле, которое порождает $\mathscr{S}$. Хотя каждый определяющий сходимость класс, очевидно, является также и определяющим классом, следующий пример показывает, что обратное неверно. Пусть S — полуоткрытый интервал $[0,1)$ с обычной метрикой; пусть $\mathscr{V}$ — класс множеств $[a, b)$ с $0 < a < b < 1$. Тогда $\mathscr{V}$ является определяющим классом, но не определяющим сходимость классом (в этом легко убедиться, взяв в качестве P_n и P распределения, приписывающие единичные массы точкам $1-1/n$ и 0, соответственно). Хотя данный пример искусствен, мы увидим, что применения изобилуют реальными примерами определяющих классов, которые не являются классами, определяющими сходимость.

В заключение этого параграфа дадим еще одно условие слабой сходимости. Напомним, что последовательность $\{x_n\}$ действительных чисел сходится к x тогда и только тогда, когда любая подпоследовательность $\{x_{n'}\}$ содержит другую подпоследовательность $\{x_{n''}\}$, которая сходится к x (удобно обозначать последовательность целых чисел через $\{n'\}$ вместо $\{n_k\}$ и подпоследовательность $\{n'\}$ через $\{n''\}$ вместо $\{n_{k_l}\}$). Из этого факта легко получить аналогичное утверждение для слабой сходимости.

Теорема 2.3. *$P_n \Rightarrow P$ в том и только том случае, когда любая подпоследовательность $\{P_{n'}\}$ содержит другую подпоследовательность $\{P_{n''}\}$ такую, что $P_{n''} \Rightarrow P$.*

Мы будем иногда иметь дело со слабой сходимостью P_t к P, когда t — непрерывный параметр, стремящийся к бесконечности. Разумеется, это означает по определению, что

$$\lim_{t\to\infty} \int f\,dP_t = \int f\,dP \tag{2.4}$$

для любой функции f из $C(S)$. Если функция f фиксирована, соотношение (2.4) выполняется тогда и только тогда, когда

$$\lim_{n\to\infty}\int f\,dP_{t_n}=\int f\,dP \tag{2.5}$$

для любой последовательности $\{t_n\}$, сходящейся к бесконечности. Таким образом, $P_t\Rightarrow P$ при $t\to\infty$ в том и только том случае, когда $P_{t_n}\Rightarrow P$ для любой последовательности $\{t_n\}$, сходящейся к бесконечности, так что при переходе к непрерывно меняющемуся параметру ничего существенно нового не требуется. Мы можем также допустить, что t приближается непрерывным образом к некоторому конечному значению t_0.

Примечание. Теорема 2.1 была известна уже Александрову (1940—1943). Теорема 2.2 принадлежит Колмогорову и Прохорову (1954). Другие изложения теории см. в книгах: Гихман, Скороход (1965), Хеннекен, Тортра (1974) и Партасарати (1967).

Сопряженным пространством к банахову пространству $C(S)$ является пространство $C^*(S)$ конечных знакопеременных *) мер на $\mathcal{S}$. Относительная топология пространства $Z(S)$ вероятностных мер, порожденная слабой * топологией (или $C(S)$-топологией) пространства $C^*(S)$, — это топология, описанная в первой сноске этого параграфа (отсюда термин «слабая» в нашей «слабой сходимости»); см. Данфорд, Шварц (1962, стр. 283 и 414). Варадарайн (1958a и 1961a) исследовал топологическую структуру $Z(S)$; см. также Добавление III.

О расширении теории на общие топологические пространства см. Ле Кам (1960), Варадарайн ([1958a и 1961a) и Каллианпур (1961).

Если метрическое пространство S не сепарабельно, то σ-поле $\mathcal{S}_0$, порожденное сферами, может быть меньше, чем $\mathcal{S}$. Дадли (1966 и 1967) рассматривает теорию слабой сходимости применительно лишь к множествам из $\mathcal{S}_0$ и $\mathcal{S}_0$-измеримым функциям.

Если $P_n\Rightarrow P$, то естественно задаться вопросом, выполняется ли $P_n(A)\to P(A)$ равномерно для данного класса множеств P-непрерывности; см. Ранга Рао (1962) и Биллингсли, Топсё (1967).

Задачи

1. Если S счетно и дискретно, то $P_n\Rightarrow P$ тогда и только тогда, когда $P_n\{x\}\to P\{x\}$ для любого одноточечного множества $\{x\}$.

2. Пусть P_n и P заданы плотностями p_n и p относительно меры λ на $(S, \mathcal{S})$. Если $p_n(x)\to p(x)$ всюду, за исключением x из множества, имеющего λ-меру 0, то $P_n\Rightarrow P$ [см. теорему Шеффе на стр. 306]. Постройте пример, показывающий, что сходимость

---

*) signed; в русской литературе такие меры называют обычно обобщенными мерами; см., например, Халмош (1953). (*Прим. перев.*)

$p_n(x) \to p(x)$ может не иметь места на множестве положительной меры даже в том случае, когда $P_n \Rightarrow P$.

3. Пусть $P_n \Rightarrow P$. Соотношение $\int f\,dP_n \to \int f\,dP$ может не выполняться, если f ограничена, но не непрерывна, или если f непрерывна, но не ограничена (даже если интегралы существуют). Приведите примеры. Если S — компакт, то вторая возможность не осуществляется. Что будет в случае, если S не является компактом, но P обладает компактным носителем?

4. Класс множеств непрерывности P (мера P фиксирована) образует поле.

5. Если $\mathcal{U}$ — определяющий класс, если $P_n(A) \to Q(A)$ для $A \in \mathcal{U}$ и если $P_n \Rightarrow P$, то отсюда еще *не* следует, что $P = Q$. [Определим вероятностные меры на прямой соотношениями $P_n\{n^{-1}\} = P_n\{1 + n^{-1}\} = P\{0\} = P\{1\} = 1/2$ и $Q\{0\} = 1$. Пусть B состоит из точек 0, 1, n^{-1} и $1 + n^{-1}$, $n = 1, 2, \ldots$ Определим $\mathcal{U}$ как поле борелевских множеств A таких, что или (i) $A \cap B$ конечно и $0 \notin A$, или же (ii) $A^c \cap B$ конечно и $0 \notin A^c$. (Этот пример принадлежит О. Бьернсону.)]

6. Определите, что должно понимать под определяющими классами и классами, определяющими сходимость, элементов пространства $C(S)$. Приведите пример определяющего класса, который не является классом, определяющим сходимость. Покажите, что класс, равномерно плотный в $C(S)$, служит классом, определяющим сходимость.

7. Если f ограничена и полунепрерывна сверху (стр. 298), то из $P_n \Rightarrow P$ следует $\limsup_n \int f\,dP_n \leqslant \int f\,dP$.

8. Пусть $\{f_\theta\}$ — семейство действительных функций на S, равностепенно непрерывных в каждой точке x (для каждого x и ε существует δ такое, что при всех θ из $\rho(x, y) < \delta$ следует $|f_\theta(x) - f_\theta(y)| < \varepsilon$). Если семейство $\{f_\theta\}$ равномерно ограничено, а S сепарабельно, то из $P_n \Rightarrow P$ следует, что $\int f_\theta\,dP_n \to \int f_\theta\,dP$ равномерно по θ. [Покажите сначала, что для каждого ε существует счетное разбиение $\Delta_\varepsilon = \{D_{\varepsilon k}\}$ пространства S на множества P-непрерывности такие, что $|f_\theta(x) - f_\theta(y)| < \varepsilon$ при всех θ, если x и y принадлежат одному и тому же $D_{\varepsilon k}$. Затем замените интеграл $\int f_\theta\,dP$ суммой $\sum_k f_\theta(x_k)P(D_{\varepsilon k})$, где x_k принадлежит $D_{\varepsilon k}$ (аналогично для $\int f_\theta\,dP_n$), и примените теорему Шеффе (стр. 306).] (Этот результат принадлежит Ранга Рао (1962); по поводу обобщений см. Биллингсли, Топсё (1967) и Топсё (1967а и 1967б).)

§ 3. Некоторые частные случаи

Евклидово пространство. Обозначим через R^k k-мерное евклидово пространство, которое мы всегда будем рассматривать с обычной метрикой $\rho(x, y) = |x - y| = \left[\sum_{i=1}^{k}(x_i - y_i)^2\right]^{1/2}$. Обозначим через $\mathfrak{R}^k$ класс борелевских множеств в R^k; элементы $\mathfrak{R}^k$ мы будем называть *k-мерными борелевскими множествами*, а в случае $k = 1$ — *линейными борелевскими множествами*. Запись $y \leqslant x$ $[y < x]$ будет означать, что $y_i \leqslant x_i$ $[y_i < x_i]$ для всех $i = 1, 2, \ldots, k$. (Заметим, что условие $y < x$ сильнее, чем условие $y \leqslant x$ и $y \neq x$.) Интервал есть множество вида

$$(a, b] = \{x\colon a < x \leqslant b\}. \tag{3.1}$$

Наконец, обозначим через $e = (1, \ldots, 1)$ вектор, все координаты которого равны 1.

Произвольная вероятностная мера P на $(R^k, \mathfrak{R}^k)$ имеет функцию распределения F, определенную равенством

$$F(x) = P\{y\colon y \leqslant x\}, \qquad x \in R^k. \tag{3.2}$$

Сравним понятие слабой сходимости $P_n \Rightarrow P$ с обычным понятием сходимости соответствующих функций распределения F_n к F.

Функция F непрерывна в x тогда и только тогда, когда для каждого положительного ε существует положительное число δ такое, что из $x - \delta e < y < x + \delta e$ следует $|F(x) - F(y)| < \varepsilon$. По определению F непрерывна сверху *) в x, если для каждого положительного ε существует положительное δ такое, что из $x \leqslant y < x + \delta e$ следует $|F(x) - F(y)| < \varepsilon$. Из определения (3.2) следует, что F не убывает по любой переменной. Следовательно, F непрерывна сверху в x тогда и только тогда, когда $F(x)$ совпадает с $\inf\limits_{\delta > 0} F(x + \delta e) = \inf\limits_{\delta > 0} P\{y\colon y \leqslant x + \delta e\}$. Этот инфинум является как раз P-мерой пересе-

*) Понятия «непрерывность сверху» и «непрерывность снизу» аналогичны, таким образом, непрерывности справа и слева в одномерном случае. *(Прим. перев.)*

чения $\bigcap_{\delta>0}\{y: y \leqslant x+\delta e\}=\{y: y \leqslant x\}$. Тем самым F непрерывна сверху в каждой точке x.

Поскольку F есть функция, неубывающая по каждой переменной и всюду непрерывная сверху, она является непрерывной в x тогда и только тогда, когда она непрерывна снизу в этой точке, иначе говоря, тогда и только тогда, когда для каждого положительного ε существует положительное δ такое, что из $x-\delta e < y \leqslant x$ следует $|F(x)-F(y)| < \varepsilon$. Еще раз используя монотонность, мы видим, что это условие в свою очередь эквивалентно равенству $F(x)=\sup\limits_{\delta>0} F(x-\delta e)$. Поскольку супремум является P-мерой объединения $\bigcup_{\delta>0}\{y: y \leqslant x-\delta e\}=\{y: y<x\}$, мы видим, что F непрерывна в x тогда и только тогда, когда $F(x)=P\{y: y<x\}$. Поскольку множество $\{y: y \leqslant x\}-\{y: y<x\}$ является в точности границей $\{y: y \leqslant x\}$, F непрерывна в x тогда и только тогда, когда $\{y: y \leqslant x\}$ является множеством непрерывности меры P.

Для функций распределения F_n и F введем определение сходимости $F_n \Rightarrow F$, означающее, что $F_n(x) \rightarrow F(x)$ в точках x непрерывности F. Из доказанного выше следует, что если $P_n \Rightarrow P$, то для соответствующих функций распределения имеет место сходимость $F_n \Rightarrow F$. Интервал $(a, b]$ определяется $2k$ гиперплоскостями размерности $k-1$, содержащими все его грани; пусть $\mathscr{U}$ есть класс интервалов, для которых все эти гиперплоскости имеют P-меру 0. Каждая вершина элемента $\mathscr{U}$ есть точка непрерывности F и класс $\mathscr{U}$ замкнут относительно образования конечных пересечений. Поскольку положительную P-меру может иметь только счетное число параллельных гиперплоскостей, из следствия 1 теоремы 2.2 следует, что если $P_n(A) \rightarrow P(A)$ для каждого A из $\mathscr{U}$, то $P_n \Rightarrow P$. Вероятность $P(a, b]$ есть сумма $\sum \pm F(x)$, где x пробегает множество 2^k вершин интервала $(a, b]$ (и аналогично выражается $P_n(a, b]$ через F_n). Поэтому из $F_n \Rightarrow F$ следует, что $P_n(A) \rightarrow P(A)$ для каждого A из $\mathscr{U}$. Следовательно, соотношения $P_n \Rightarrow P$ и $F_n \Rightarrow F$ эквивалентны.

Таким образом, понятие слабой сходимости сводится в R^k к обычному понятию сходимости функций распределения. Другими словами, множества $\{y: y \leqslant x\}$

образуют класс, определяющий сходимость. Доказательство, приведенное выше, также показывает, что прямоугольники $(a, b]$ образуют класс, определяющий сходимость.

Окружность. Подобного же рода результат, очевидно, справедлив, если S — единичная окружность на комплексной плоскости: $P_n \Rightarrow P$ тогда и только тогда, когда $P_n(A) \to P(A)$ для каждой дуги A, концы которой имеют P-меру 0. Последовательность $\{x_1, x_2, \ldots\}$ точек S (комплексные числа, равные по модулю 1) считается равномерно распределенной, если доля точек на каждой дуге пропорциональна ее длине, в том смысле, что

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} I_A(x_j) = P(A), \tag{3.3}$$

где I_A есть индикатор или характеристическая функция A, а P есть мера Лебега на окружности, нормированная так, что $P(S) = 1$. Если P_n обозначает эмпирическое распределение последовательности первых n точек, т. е. меру, сопоставляющую массу $1/n$ каждой из точек $x_1, x_2, \ldots, x_n$, то это условие сводится к сходимости $P_n(A) \to P(A)$ для дуг A; таким образом, последовательность равномерно распределена тогда и только тогда, когда $P_n \Rightarrow P$. Поэтому, если каждая дуга содержит соответствующую долю точек в смысле (3.3), то это же самое справедливо для любого другого борелевского множества, граница которого имеет меру Лебега, равную 0. Мы доказываем в § 7 известную теорему Вейля, согласно которой $\{x_1, x_2, \ldots\}$ равномерно распределена тогда и только тогда, когда условие $n^{-1} \sum_{j=1}^{n} (x_j)^u \to 0$ справедливо для каждого отличного от нуля целого числа u.

Пространство R^∞. Результаты, установленные для R^k, переносятся во всех существенных отношениях на топологическое произведение счетной последовательности прямых, иначе говоря, на пространство R^∞ последовательностей $x = (x_1, x_2, \ldots)$ действительных чисел (см. стр. 299). Топология в R^∞ имеет в качестве базовых окрестностей точки x множества вида

$$N_{k,\varepsilon}(x) = \{y\colon |x_i - y_i| < \varepsilon, \quad i = 1, \ldots, k\}, \tag{3.4}$$

где $\varepsilon > 0$, а $k = 1, 2, \ldots$ В этой топологии R^∞ является полным сепарабельным метрическим пространством. (Мы всегда будем подразумевать под R^∞ топологическое произведение *счетного* числа прямых.)

Обозначим через π_k естественную проекцию из R^∞ на R^k: $\pi_k(x) = (x_1, \ldots, x_k)$. *Конечномерное множество* (или цилиндр) есть по определению множество вида $\pi_k^{-1}H$ при $k \geqslant 1$ и $H \in \mathscr{R}^k$. Поскольку каждая проекция π_k непрерывна и, следовательно, измерима (см. стр. 304), то конечномерные множества принадлежат σ-полю $\mathscr{R}^\infty$ борелевских множеств в R^∞. Обозначим через $\mathscr{F}$ класс конечномерных множеств. Поскольку каждое множество (3.4) принадлежит $\mathscr{F}$ и поскольку R^∞ сепарабельно, то $\mathscr{F}$ порождает $\mathscr{R}^\infty$. Поскольку $\mathscr{F}$ — поле (конечно аддитивное), то, следовательно, $\mathscr{F}$ — определяющий класс.

Для фиксированных k и x множества (3.4) при различных значениях ε имеют непересекающиеся границы (из $\varepsilon > \delta$ следует $N_{k,\varepsilon}^- \subset N_{k,\delta}^0$). Применяя следствие 1 теоремы 2.2 к классу множеств P-непрерывности из $\mathscr{F}$, мы находим тем самым, что $\mathscr{F}$ является даже классом, определяющим сходимость. Таким образом, $P_n \Rightarrow P$ тогда и только тогда, когда сходимость $P_n(A) \to P(A)$ справедлива для всех конечномерных множеств P-непрерывности.

Пространство C. В $C = C[0, 1]$, пространстве непрерывных функций на $[0, 1]$ с равномерной метрикой $\rho(x, y) = \sup_t |x(t) - y(t)|$ (см. стр. 301), положение заметно отличается от того, что мы имели в R^∞. Выберем точки $t_1, \ldots, t_k$ из $[0, 1]$. Пусть $\pi_{t_1 \ldots t_k}$ — отображение, сопоставляющее точке x из C точку $(x(t_1), \ldots, x(t_k))$ в R^k. Конечномерные множества здесь определяются как множества вида $\pi_{t_1 \ldots t_k}^{-1} H$ с $H \in \mathscr{R}^k$. Поскольку отображение $\pi_{t_1 \ldots t_k}$ непрерывно, то эти множества содержатся в классе $\mathscr{C}$ борелевских множеств в C. С другой стороны, замкнутая сфера $\{y: \rho(x, y) < \varepsilon\}$ является пределом конечномерных множеств $\{y: |x(i/n) - y(i/n)| \leqslant \varepsilon, i = 1, \ldots, n\}$; поскольку C сепарабельно, каждое открытое множество есть счетное объединение открытых сфер, а следовательно, и замкнутых сфер, так что конечномерные множества порождают $\mathscr{C}$. Так как они образуют поле, то класс конечномерных множеств является определяющим классом.

Приведем пример, показывающий, что конечномерные множества *не* образуют класс, определяющий сходимость. Пусть P приписывает единичную массу точке 0 (0 — это в данном случае функция, тождественно равная 0), а P_n — точке (функции) x_n, где

$$x_n(t)=\begin{cases} nt & \text{при} \quad 0\leqslant t\leqslant 1/n, \\ 2-nt & \text{при} \quad 1/n\leqslant t\leqslant 2/n, \\ 0 & \text{при} \quad 2/n\leqslant t\leqslant 1. \end{cases} \tag{3.5}$$

Поскольку x_n не сходится к 0 равномерно (т. е. в топологии C), то мера P_n не может слабо сходиться к P. (Например, если $A=S(0, 1/2)$, то $P(\partial A)=0$, и в то же время $P_n(A)=0$ не сходится к $P(A)=1$.) Но соотношение $P_n(A)\to P(A)$ справедливо для конечномерных множеств непрерывности P. В самом деле, если $A=\pi^{-1}_{t_1 \dots t_k} H$ и если $2/n$ меньше наименьшего из не равных нулю t_i, то $P_n(A)=P(A)$.

Конечномерные множества являются, таким образом, определяющим классом в C, но не классом, определяющим сходимость. Трудность, интерес и польза понятия слабой сходимости в C связаны с тем, что это понятие требует выхода за рамки того, что относится к конечномерным множествам.

Произведение пространств. Пусть $S=S'\times S''$ — произведение метрических пространств S' и S''. Если S сепарабельно (для этого нужно, чтобы S' и S'' были сепарабельны), тогда σ-поля $\mathcal{S}$, $\mathcal{S}'$ и $\mathcal{S}''$ борелевских множеств в этих пространствах связаны соотношением $\mathcal{S}=\mathcal{S}'\times\mathcal{S}''$ (см. стр. 308).

Два маргинальных распределения вероятностной меры P на $(S, \mathcal{S})$ определяются как $P'(A')=P(A'\times S'')$, $A'\in\mathcal{S}'$, и $P''(A'')=P(S'\times A'')$, $A''\in\mathcal{S}''$.

Теорема 3.1. *Если S сепарабельно, то для сходимости $P_n\Rightarrow P$ необходимо и достаточно выполнения условия $P_n(A'\times A'')\to P(A'\times A'')$ для любого множества A' непрерывности меры P' и любого множества A'' непрерывности меры P'', где P' и P'' являются маргинальными распределениями меры P.*

Доказательство. Обозначим через ∂, ∂' и ∂'' граничные операторы в S, S' и S'', соответственно. Так как

$$\partial(A' \times A'') \subset ((\partial' A') \times S'') \cup (S' \times (\partial'' A'')), \qquad (3.6)$$

то условие теоремы необходимо.

Для доказательства достаточности применим следствие 1 теоремы 2.2 к классу $\mathcal{U}$ множеств $A' \times A''$, где A' и A'' — множества непрерывности P' и P'', соответственно. Класс $\mathcal{U}$ замкнут относительно образования конечных пересечений и по предположению $P_n(A) \to P(A)$ для A из $\mathcal{U}$.

Задаваясь точкой (x', x'') в S и числом $\varepsilon > 0$, рассмотрим множества

$$A_\delta = \{y' \colon \rho'(x', y') < \delta\} \times \{y'' \colon \rho''(x'', y'') < \delta\}.$$

При различных δ множества $\partial'\{y' \colon \rho'(x', y') < \delta\}$ не пересекаются и множества $\partial''\{y'' \colon \rho''(x'', y'') < \delta\}$ не пересекаются, поэтому при некотором δ, где $0 < \delta < \varepsilon$, A_δ принадлежит $\mathcal{U}$. Если в S ввести метрику

$$\rho((x', x''), (y', y'')) = \max\{\rho'(x', y'), \rho''(x'', y'')\},$$

то A_δ будет сферой с центром (x', x'') и радиусом δ. Следовательно, $\mathcal{U}$ удовлетворяет предположению следствия 1 теоремы 2.2, что и требовалось доказать.

Достаточность условия теоремы 3.1 показывает, что измеримые прямоугольники образуют класс, определяющий сходимость (и даже больше, так как из того, что $P(\partial(A' \times A'')) = 0$ не следует $P'(\partial' A) = P''(\partial'' A) = 0$).

Для данных вероятностных мер P' и P'', заданных на $(S', \mathcal{S}')$ и $(S'', \mathcal{S}'')$, соответственно, их произведение $P' \times P''$ есть вероятностная мера на $\mathcal{S}' \times \mathcal{S}''$ и, следовательно, если S сепарабельно, мера на $\mathcal{S}$. Следующая теорема, в которой P'_n и P' являются вероятностными мерами на $(S, \mathcal{S}')$, а P''_n и P'' — вероятностными мерами на $(S'', \mathcal{S}'')$, представляет собой прямое следствие теоремы 3.1.

Теорема 3.2. *Если S сепарабельно, то $P'_n \times P''_n \Rightarrow P' \times P''$ тогда и только тогда, когда $P'_n \Rightarrow P'$ и $P''_n \Rightarrow P''$.*

Задачи

1. Покажите непосредственно, что $F_n \Rightarrow F$ тогда и только тогда, когда $F_n(x) \to F(x)$ для всех x из некоторого всюду плотного множества, и $F_n \Rightarrow F$ тогда и только тогда, когда $\limsup_n F_n(x) \leqslant F(x)$

и $\liminf_n F_n(x-0) \geqslant F(x-0)$ для всех x (здесь $F(x-0) = \sup_{y<x} F(y)$).

2. Если $k > 1$, то множество точек разрыва функции распределения F, хотя и обладает всюду плотным дополнением, не обязано быть счетным. $(k-1)$-мерная гиперплоскость может содержать самое большее счетное множество разрывов, если она не перпендикулярна ни к одной из осей. [Чтобы облегчить себе понимание задачи, рассмотрите сначала гиперплоскость $\{(x_1, x_2) : x_1 = -x_2\}$ в R^2.]

3. Если $F_n \Rightarrow F$ и если F непрерывна в каждой точке замкнутого множества A, то

$$\sup_{x \in A} |F_n(x) - F(x)| \to 0.$$

4. Расстояние Леви $\lambda(F, G)$ между двумя одномерными функциями распределения определяется как точная нижняя грань всех положительных ε таких, что $F(x-\varepsilon) - \varepsilon \leqslant G(x) \leqslant F(x+\varepsilon) + \varepsilon$ для всех x. Интерпретируйте $\lambda(F, G)$ геометрически с точки зрения взаимного расположения графиков F и G. Покажите, что $F_n \Rightarrow F$ тогда и только тогда, когда $\lambda(F_n, F) \to 0$; докажите, что совокупность одномерных функций распределения представляет собой сепарабельное полное метрическое пространство относительно метрики λ.

5. Задача 5 § 1 адаптирует предшествующую ей задачу к общему банахову пространству счетного числа измерений. В пространстве C возможен простой непосредственный анализ: используйте функции x_n, определенные в (3.5).

6. Равномерное распределение на единичном квадрате и равномерное распределение на его диагонали имеют одинаковые маргинальные распределения. Сопоставьте этот факт с теоремой 3.2.

7. Обобщите факты, касающиеся мер в произведении пространств (приведенные в конце параграфа), показав, что для счетного произведения сепарабельных пространств конечномерные множества (соответствующим образом определенные) образуют класс, определяющий сходимость (пространство R^∞ представляет собой частный случай).

§ 4. Сходимость по распределению

Теория слабой сходимости может быть изложена как теория сходимости по распределению. Это не требует никаких новых идей, и многие результаты приобретают при этом прозрачную и компактную форму.

Случайные элементы. Пусть X отображает вероятностное пространство $(\Omega, \mathscr{B}, \mathrm{P})$ в метрическое пространство S. Если X измеримо (в том смысле, что $X^{-1}\mathscr{S} \subset \mathscr{B}$; см. стр. 304), то мы назовем X *случайным элементом*. Мы будем говорить, что случайный элемент X задан *на* своей области определения Ω (или $(\Omega, \mathscr{B}, \mathrm{P})$) и *в* своей области изменения S, и будем называть X случайным элементом S. Если $S = R^1$, то мы называем X случайной величиной; если $S = R^k$, то X есть случайный вектор; если $S = C$, то X именуется случайной функцией.

Случайные величины и случайные векторы являются привычными объектами изучения, а пример полезной функции (измеримость которой, впрочем, не была доказана) содержится во введении (см. формулу (13)). Целый ряд случайных функций возникает естественным образом в теории вероятностей.

Распределение *) случайного элемента X есть вероятностная мера $P = \mathsf{P}X^{-1}$ на $(S, \mathscr{S})$:

$$P(A) = \mathsf{P}\left(X^{-1}A\right) = \mathsf{P}\{\omega:\ X(\omega) \in A\} = \\ = \mathsf{P}\{X \in A\}, \quad A \in \mathscr{S} \tag{4.1}$$

(мы обычно опускаем в записи аргумент ω). В случае $S = R^k$ мы определим, кроме того, ***функцию распределения*** X:

$$F(x) = P\{y:\ y \leqslant x\} = \mathsf{P}\{X \leqslant x\}, \quad x \in R^k.$$

Заметим, что вероятностная мера P определяется на пространстве произвольной природы, в то время как распределение P всегда задано на метрическом пространстве. Как правило, распределение P содержит всю существенную информацию о случайном элементе X. Если h — измеримая функция на S $(h^{-1}\mathscr{R}^1 \subset \mathscr{S})$, то с помощью формулы замены переменных (см. стр. 305) получаем равенство

$$\int h(X)\,d\mathsf{P} = \int h\,dP, \tag{4.2}$$

понимаемое так, что оба интеграла существуют или не существуют одновременно и имеют одинаковое значение, если существуют. Используя обычное обозначение для математического ожидания, вместо (4.2) получаем

$$\mathsf{E}\{h(X)\} = \int h\,dP. \tag{4.3}$$

Любая вероятностная мера на любом метрическом пространстве является распределением некоторого случайного элемента на некотором вероятностном пространстве:

*) Это понятие не имеет ничего общего с понятием распределения по Шварцу.

пусть дана мера P на $(S, \mathcal{S})$, тогда, если мы положим $(\Omega, \mathcal{B}, \mathsf{P}) = (S, \mathcal{S}, \mathsf{P})$ и

$$X(\omega) = \omega, \qquad \omega \in \Omega = S, \tag{4.4}$$

то X оказывается случайным элементом, определенным на Ω, со значениями в S и имеет в качестве своего распределения меру P. Хотя класс распределений, таким образом, совпадает с классом вероятностных мер на метрических пространствах, мы обычно называем меру на метрическом пространстве распределением только тогда, когда она действительно является распределением некоторого заранее данного случайного элемента.

Сходимость по распределению. Мы говорим, что последовательность $\{X_n\}$ случайных элементов *сходится по распределению* к случайному элементу X, и записываем

$$X_n \xrightarrow{\mathcal{D}} X, \tag{4.5}$$

если распределения P_n элементов X_n слабо сходятся к распределению P элемента X:

$$P_n \Rightarrow P. \tag{4.6}$$

Хотя это определение, разумеется, не имеет смысла в случае, если пространство образов S (область изменения) и топология на нем не являются одними и теми же для всех случайных элементов X, X_1, X_2, ..., исходные вероятностные пространства (области определения) могут быть все различными. Мы обычно не упоминаем об этих исходных пространствах, ибо их структуры участвуют в рассуждениях лишь постольку, поскольку они индуцируют распределения на S. Поэтому мы пишем $\mathsf{P}\{X_n \in A\}$ вместо того, чтобы писать $\mathsf{P}_n\{X_n \in A\}$, и пишем $\mathsf{E}\{f(X_n)\}$ вместо того, чтобы писать $\int f(X_n)\,d\mathsf{P}_n$ или $\mathsf{E}_n\{f(X_n)\}$. Поскольку из формулы замены переменных (4.2) вытекает равенство $\int\limits_S f(x)\,P(dx) = \int\limits_\Omega f(X)\,d\mathsf{P}$ $\left(\text{и аналогичное равенство для } \int f\,dP_n\right)$, мы имеем $X_n \xrightarrow{\mathcal{D}} X$ тогда и только тогда, когда $\mathsf{E}\{f(X_n)\} \to \mathsf{E}\{f(X)\}$ для каждой $f \in C(S)$.

Теорема 2.1 утверждает эквивалентность следующих пяти утверждений. Назовем множество A из $\mathcal{S}$ X-непрерывным множеством, если $\mathsf{P}\{X \in \partial A\} = 0$.

(i) $X_n \xrightarrow{\mathcal{D}} X$.

(ii) $\lim_n \mathsf{E}\{f(X_n)\} = \mathsf{E}\{f(X)\}$ *для всех ограниченных, равномерно непрерывных действительных функций* f.

(iii) $\limsup_n \mathsf{P}\{X_n \in F\} \leqslant \mathsf{P}\{X \in F\}$ *для всех замкнутых* F.

(iv) $\liminf_n \mathsf{P}\{X_n \in G\} \geqslant \mathsf{P}\{X \in G\}$ *для всех открытых* G.

(v) $\lim_n \mathsf{P}\{X_n \in A\} = \mathsf{P}\{X \in A\}$ *для всех X-непрерывных множеств* A.

Каждая теорема о слабой сходимости может быть переформулирована подобным же образом.

Может оказаться полезной нижеследующая смешанная терминология: если X_n суть случайные элементы S, P_n — соответствующие распределения и P — вероятностная мера на $(S, \mathcal{S})$, то мы говорим, что X_n *сходится по распределению* к P, и пишем

$$X_n \xrightarrow{\mathcal{D}} P \tag{4.7}$$

в случае, когда $P_n \Rightarrow P$. Существует очевидный соответствующий вариант теоремы 2.1.

Весьма удобно иметь возможность переходить от одного из трех эквивалентных понятий (4.5), (4.6) и (4.7) к другому, и мы будем широко этим пользоваться. Это, главным образом, вопрос удобства фразеологии. Например, если случайные величины X_n имеют асимптотическое нормальное распределение со средним μ и дисперсией σ^2, мы будем выражать этот факт записью

$$X_n \xrightarrow{\mathcal{D}} N(\mu, \sigma^2). \tag{4.8}$$

Безразлично при этом, будем ли мы интерпретировать $N(\mu, \sigma^2)$ как меру на $(R^1, \mathcal{R}^1)$ с плотностью $(2\pi\sigma^2)^{-1/2}e^{-(u-\mu)^2/2\sigma^2}$ и понимать (4.8) в смысле (4.7), или же интерпретировать $N(\mu, \sigma^2)$ как случайную величину с распределением

$$\mathsf{P}\{N(\mu, \sigma^2) \in A\} = \frac{1}{\sqrt{2\pi}\,\sigma} \int_A e^{-(u-\mu)^2/2\sigma^2}\, du, \quad A \in \mathcal{R}^1, \tag{4.9}$$

и понимать (4.8) в смысле (4.5). (Существует сколько угодно подобных случайных величин $N(\mu, \sigma^2)$ на многих вероятностных пространствах; конструкция (4.4) представляет одну из них.) Мы ограничимся для $N(\mu, \sigma^2)$ этим (двойным) значением и введем запись N для $N(0, 1)$:

$$\mathsf{P}\{N \in A\} = \frac{1}{\sqrt{2\pi}} \int_A e^{-u^2/2}\, du, \quad A \in \mathscr{R}^1. \tag{4.10}$$

Сходимость по вероятности. Многие стандартные понятия и результаты, касающиеся сходимости по распределению обычных случайных величин, обобщаются на случайные элементы. Если для элемента a из S для каждого положительного ε справедливо

$$\mathsf{P}\{\rho(X_n, a) \geqslant \varepsilon\} \to 0, \tag{4.11}$$

то мы говорим, что X_n *сходится по вероятности* к a, и записываем это в виде

$$X_n \xrightarrow{P} a. \tag{4.12}$$

Если рассматривать a как случайный элемент с постоянным значением, то, как легко доказать, $X_n \xrightarrow{P} a$ тогда и только тогда, когда $X_n \xrightarrow{\mathscr{D}} a$. Иными словами, предельное соотношение $X_n \xrightarrow{P} a$ верно тогда и только тогда, когда распределение X_n слабо сходится к вероятностной мере, отвечающей массе, равной 1 в точке a. Случайные элементы X_n в (4.12), как говорилось, могут быть определены на различных вероятностных пространствах; однако область изменения S должна быть общей для них всех.

Если X_n и Y_n имеют общую область определения, есть смысл говорить о расстоянии $\rho(X_n, Y_n)$, т. е. о функции со значением $\rho(X_n(\omega), Y_n(\omega))$ в точке ω. Если S сепарабельно, то $\rho(X_n, Y_n)$ является случайной величиной (см. стр. 309). В следующей теореме мы предположим, что для каждого n X_n и Y_n действительно имеют общую область определения и что S сепарабельно.

Теорема 4.1. *Если* $X_n \xrightarrow{\mathscr{D}} X$ *и* $\rho(X_n, Y_n) \xrightarrow{P} 0$, *то* $Y_n \xrightarrow{\mathscr{D}} Y$.

Доказательство. Если положить $F_\varepsilon = \{x: \rho(x, F) \leqslant \varepsilon\}$, то

$$\mathsf{P}\{Y_n \in F\} \leqslant \mathsf{P}\{\rho(X_n, Y_n) \geqslant \varepsilon\} + \mathsf{P}\{X_n \in F_\varepsilon\}.$$

Поскольку F_ε замкнуто, то из предположений теоремы следует, что

$$\limsup_n \mathsf{P}\{Y_n \in F\} \leqslant \limsup_n \mathsf{P}\{X_n \in F_\varepsilon\} \leqslant \mathsf{P}\{X \in F_\varepsilon\}.$$

Если F замкнуто, то $F_\varepsilon \downarrow F$ при $\varepsilon \downarrow 0$ и результат вытекает из теоремы 2.1 (ее варианта, соответствующего сходимости по распределению).

В следующей теореме*) мы предполагаем, что для каждого n случайные элементы Y_n, X_{1n}, X_{2n}, ... имеют общую область определения и что S сепарабельно.

Теорема 4.2. *Предположим, что для каждого u $X_{un} \xrightarrow{\mathscr{D}} X_u$ при $n \to \infty$ и что $X_u \xrightarrow{\mathscr{D}} X$ при $u \to \infty$. Предположим далее, что для каждого положительного ε*

$$\lim_{u\to\infty} \limsup_{n\to\infty} \mathsf{P}\{\rho(X_{un}, Y_n) \geqslant \varepsilon\} = 0. \tag{4.13}$$

Тогда $Y_n \xrightarrow{\mathscr{D}} X$ при $n \to \infty$.

Доказательство. Определяя F_ε, как и ранее, равенством $F_\varepsilon = \{x: \rho(x, F) \leqslant \varepsilon\}$, мы имеем

$$\mathsf{P}\{Y_n \in F\} \leqslant \mathsf{P}\{X_{un} \in F_\varepsilon\} + \mathsf{P}\{\rho(X_{un}, Y_n) \geqslant \varepsilon\}.$$

В силу предположения $X_{un} \xrightarrow{\mathscr{D}} X_u$ $(n \to \infty)$ имеем

$$\limsup_{n\to\infty} \mathsf{P}\{Y_n \in F\} \leqslant \mathsf{P}\{X_u \in F_\varepsilon\} + \limsup_{n\to\infty} \mathsf{P}\{\rho(X_{un}, Y_n) \geqslant \varepsilon\}.$$

В силу (4.13) и предположения $X_u \xrightarrow{\mathscr{D}} X$ $(u \to \infty)$ имеем

$$\limsup_{n\to\infty} \mathsf{P}\{Y_n \in F\} \leqslant \mathsf{P}\{X \in F_\varepsilon\}.$$

Результат теоремы выводится теперь так же, как в предыдущем случае (теорема 4.1).

---

*) Материал конца этого параграфа не является стержневым для теории; после беглого прочтения к нему можно обращаться в дальнейшем по мере надобности.

Предположим далее, что все X, X_1, X_2, ... имеют общую область определения и что S сепарабельно. Если для всякого положительного ε

$$\mathsf{P}\{\rho(X_n, X) \geqslant \varepsilon\} \to 0,$$

то мы говорим, что X_n сходится по вероятности к X, и пишем

$$X_n \xrightarrow{P} X. \tag{4.14}$$

Ввиду предположений о сепарабельности и (что более важно) об общей области определения для X_n это понятие не обобщает (4.12).

Теорема 4.3. *Если* $X_n \xrightarrow{P} X$, *то* *)

$$\mathsf{P}(\{X_n \in A\} + \{X \in A\}) \to 0 \tag{4.15}$$

для каждого X-непрерывного множества A.

Доказательство. Для каждого положительного ε

$$\mathsf{P}\{X_n \in A,\ X \notin A\} \leqslant \mathsf{P}\{\rho(X_n, X) \geqslant \varepsilon\} + \\ + \mathsf{P}\{\rho(X, A) < \varepsilon,\ X \notin A\}.$$

Из этого неравенства, а также из такого же неравенства, в котором A заменяется на A^c и, наконец, из предположения $X_n \xrightarrow{P} X$ мы получаем

$$\limsup_{n\to\infty} \mathsf{P}(\{X_n \in A\} + \{X \in A\}) \leqslant \\ \leqslant \mathsf{P}\{\rho(X, A) < \varepsilon,\ X \notin A\} + \mathsf{P}\{\rho(X, A^c) < \varepsilon,\ X \in A\}.$$

При $\varepsilon \to 0$ правая часть этого неравенства стремится к $\mathsf{P}\{X \in \partial A\} = 0$.

Поскольку из (4.15) вытекает $\mathsf{P}\{X_n \in A\} \to \mathsf{P}\{X \in A\}$, то из теоремы 4.3 следует, что сходимость $X_n \xrightarrow{P} X$ влечет сходимость $X_n \xrightarrow{\mathscr{D}} X$.

Произведение пространств **). Пусть X' и X'_n есть случайные элементы S', а X'' и X''_n — случайные эле-

---

*) *Некоторые обозначения*: E^c — дополнение E, $E_1 - E_2 = E_1 \cap E_2^c$ — разность E_1 и E_2, и $E_1 + E_2 = (E_1 - E_2) \cup (E_2 - E_1)$ — симметрическая разность E_1 и E_2.

**) Предпоследняя сноска применима и к материалу этой части параграфа.

менты S''. В оставшейся части этого параграфа мы предполагаем, что X' и X'' имеют одинаковую область определения, что X'_n и X''_n имеют одинаковую область определения при каждом n и что S' и S'' сепарабельны, так что (X', X'') и (X'_n, X''_n) суть случайные элементы $S' \times S''$ (см. стр. 309). Мы разыскиваем условия, при которых

$$(X'_n, X''_n) \xrightarrow{\mathscr{D}} (X', X''). \tag{4.16}$$

Случайные элементы X' и X'' по определению независимы, если $\mathsf{P}\{X' \in A',\ X'' \in A''\} = \mathsf{P}\{X' \in A'\} \times \times \mathsf{P}\{X'' \in A''\}$. Если X' и X'' независимы и X'_n и X''_n независимы при каждом n, то согласно теореме 3.2 соотношение (4.16) эквивалентно тому, что

$$X'_n \xrightarrow{\mathscr{D}} X', \qquad X''_n \xrightarrow{\mathscr{D}} X''. \tag{4.17}$$

Хотя из соотношения (4.16) факт, устанавливаемый формулой (4.17), вытекает даже без предположения о независимости, обратное неверно (например, возьмем вектор (X', X'') равномерно распределенным на единичном квадрате, а (X'_n, X''_n) — распределенным равномерно на диагонали этого квадрата). Мы будем заниматься случаями, в которых X' и X'' независимы, однако X'_n и X''_n удовлетворяют некоторому условию, более слабому, чем независимость. Если они удовлетворяют условию (4.16) с независимыми X' и X'', то естественно рассматривать X'_n и X''_n как величины, независимые асимптотически.

Из теоремы 3.1 следует, что соотношение (4.16) выполняется тогда и только тогда, когда

$$\mathsf{P}\{X'_n \in A',\ X''_n \in A''\} \to \mathsf{P}\{X' \in A',\ X'' \in A''\} \tag{4.18}$$

для всех X'-непрерывных множеств A' и для всех X''-непрерывных множеств A''.

Теорема 4.4. *Если* $X'_n \xrightarrow{\mathscr{D}} X'$ *и* $X''_n \xrightarrow{P} a''$, *то* $(X'_n, X''_n) \xrightarrow{\mathscr{D}} (X', a'')$.

Доказательство. Мы должны проверить соотношение (4.18) при X'', тождественно равном a''. Предположим, что A' есть X'-непрерывное множество и что A'' есть X''-непрерывное множество (иными словами, $a'' \notin$

$\notin \partial A''$). Если $a'' \in A''$, то $\mathsf{P}\{X''_n \notin A''\} \to 0$ и (4.18) следует из $X'_n \xrightarrow{\mathscr{D}} X'$ и неравенства

$$\mathsf{P}\{X'_n \in A'\} - \mathsf{P}\{X''_n \notin A''\} \leqslant \\ \leqslant \mathsf{P}\{X'_n \in A',\ X''_n \in A''\} \leqslant \mathsf{P}\{X'_n \in A'\}.$$

Если $a'' \notin A''$, то (4.18) следует из соотношения

$$\mathsf{P}\{X'_n \in A',\ X''_n \in A''\} \leqslant \mathsf{P}\{X''_n \in A''\} \to 0.$$

В следующей теореме мы предположим, что X''_n сходится по вероятности к случайному элементу Y'', который не обязательно является константой; при этом естественно нужно потребовать, чтобы Y'' и все (X'_n, X''_n) имели бы одинаковую область определения $(\Omega, \mathscr{B}, \mathsf{P})$. Пусть $\mathscr{B}_0$ есть (конечно аддитивное) поле, содержащееся в $\mathscr{B}$; обозначим через $\sigma(\mathscr{B}_0)$ σ-поле, порожденное $\mathscr{B}_0$.

Теорема 4.5. *Предположим, что X' и X'' независимы и что X'' имеет то же распределение, что и Y''. Если $X''_n \xrightarrow{P} Y''$, если*

$$\mathsf{P}(\{X'_n \in A'\} \cap E) \to \mathsf{P}\{X' \in A'\}\,\mathsf{P}(E) \tag{4.19}$$

для каждого X'-непрерывного множества A' и каждого множества E из поля $\mathscr{B}_0$ и если каждый элемент X'_n $\sigma(\mathscr{B}_0)$-измерим, то $(X'_n, X''_n) \xrightarrow{\mathscr{D}} (X', X')$.

Заметим, что условие (4.19) влечет сходимость $X'_n \xrightarrow{\mathscr{D}} X'$ (нужно положить $E = \Omega$). Заметим также, что область определения (X', X'') не обязана совпадать с общей областью определения Y'' и (X'_n, X''_n). Мы можем заменить в теореме 5.1 (X', X'') на (Y', Y''), если область определения Y'' служит в то же время и областью определения некоторого случайного элемента Y', который не зависит от Y'' и имеет надлежащее распределение; однако допущение о существовании Y' в общем случае было бы ненужным ограничением.

Доказательство. Фиксируем X'-непрерывное множество A' и X''-непрерывное множество A''. Мы должны доказать (4.18) или, что (в силу независимости X' и X'' и одинаковой распределенности X'' и Y'') то же самое,

$$\mathsf{P}\{X'_n \in A',\ X''_n \in A''\} \to \mathsf{P}\{X' \in A'\}\,\mathsf{P}\{Y'' \in A''\}. \tag{4.20}$$

Поскольку $X''_n \xrightarrow{P} Y''$, то из теоремы 4.3 следует, что (4.20) в свою очередь то же самое, что

$$\mathsf{P}\{X'_n \in A', Y'' \in A''\} \to \mathsf{P}\{X' \in A'\}\,\mathsf{P}\{Y'' \in A''\}. \quad (4.21)$$

Обозначим $E_n = \{X'_n \in A'\}$ и $\alpha = \mathsf{P}\{X' \in A'\}$; и пусть g есть индикатор множества $\{Y'' \in A''\}$. Тогда соотношение (4.21) принимает вид

$$\int_{E_n} g\,d\mathsf{P} \to \alpha \int g\,d\mathsf{P}. \quad (4.22)$$

Так как каждый случайный элемент X'_n $\sigma(\mathscr{B}_0)$-измерим, то каждое множество E_n принадлежит $\sigma(\mathscr{B}_0)$.

Мы докажем, что предельное соотношение (4.22) справедливо для *произвольной* ($\mathscr{B}$-измеримой) интегрируемой функции g. Если g — индикатор множества, принадлежащего $\mathscr{B}_0$, то (4.22) справедливо, так как совпадает с (4.19). Очевидно, что (4.22) выполняется и тогда, когда g является простой $\mathscr{B}_0$-измеримой функцией. Если g интегрируема и $\sigma(\mathscr{B}_0)$-измерима, то для любого положительного ε существует простая $\mathscr{B}_0$-измеримая функция g_ε такая, что $\mathsf{E}\{|g - g_\varepsilon|\} < \varepsilon$; но тогда

$$\limsup_{n\to\infty}\left|\int_{E_n} g\,d\mathsf{P} - \alpha\int g\,d\mathsf{P}\right| \leqslant (1 + |\alpha|)\,\mathsf{E}\{|g - g_\varepsilon|\},$$

и (4.22) выполняется для всех таких функций g.

Допустим, наконец, что функция g $\mathscr{B}$-измерима и интегрируема, но не обязательно $\sigma(\mathscr{B}_0)$-измерима. Воспользовавшись свойствами условных математических ожиданий *), мы по-прежнему имеем, поскольку $E_n \in \sigma(\mathscr{B}_0)$,

$$\int_{E_n} g\,d\mathsf{P} = \int_{E_n} \mathsf{E}\{g \,\|\, \sigma(\mathscr{B}_0)\}\,d\mathsf{P} \to \alpha\int \mathsf{E}\{g \,\|\, \sigma(\mathscr{B}_0)\}\,d\mathsf{P} = \alpha\int g\,d\mathsf{P}.$$

Таким образом, (4.22) выполняется для всех интегрируемых функций g, что и требовалось доказать.

Примечание. Идеи доказательства теоремы 4.5 заимствованы у Реньи (1958).

---

*) См. Дуб (1956) и Биллингсли (1969). Основные результаты этой книги не используют таких понятий, как условные вероятности и условные математические ожидания.

Задачи

1. Для случайных величин X, Y_n и X из $X_n \xrightarrow{\mathscr{D}} X$ и $Y_n \xrightarrow{P} 0$ следует, что $X_n + Y_n \xrightarrow{\mathscr{D}} X$ и $X_n Y_n \xrightarrow{P} 0$.

2. Для случайных векторов X_n, Y_n X и случайных величин Z_n

(а) $Y_n \xrightarrow{\mathscr{D}} X$, если $X_n \xrightarrow{\mathscr{D}} X$ $|X_n - Y_n| \leqslant Z_n |X_n|$ и $Z_n \xrightarrow{P} 0$,

(б) $Y_n \xrightarrow{\mathscr{D}} X$, если $X_n \xrightarrow{\mathscr{D}} X$, $|X_n - Y_n| \leqslant Z_n |Y_n|$ и $Z_n \xrightarrow{P} 0$.

[Сведите (б) к (а), используя тот факт что из $|x-y| \leqslant \varepsilon |y|$ при $\varepsilon < 1/2$ следует $|x-y| \leqslant 2\varepsilon |x|$.]

3. Три симметричные монеты бросают независимо друг от друга. Пусть событие E_{ij} состоит в том, что монеты i и j падают одной и той же стороной, пусть $X' = X'_n = X''_n = Y''$ — индикатор события E_{13}, X'' — индикатор события E_{12}, и пусть $\mathscr{B}_0 = \{E_{12}, E_{23}\}$. Утверждение теоремы 4.5 не выполняется, хотя, за исключением требования, что $\mathscr{B}_0$ должно быть полем, все предположения теоремы удовлетворены.

4. Пусть $X_1, X_2, \ldots$ независимы и имеют одинаковое распределение P на S. Пусть $P_{n,\omega}$ — эмпирическая мера для $X_1(\omega), \ldots$ $\ldots, X_n(\omega)$; именно, $P_{n,\omega}(A)$ есть доля тех k, $1 \leqslant k \leqslant n$, для которых $X_k(\omega) \in A$:

$$P_{n,\omega}(A) = \frac{1}{n} \sum_{k=1}^{n} I_A(X_k(\omega))$$

Покажите, что если S сепарабельно, то $P_{n,\omega} \Rightarrow P$ с вероятностью 1. [Используйте усиленный закон больших чисел для испытаний Бернулли и следствие 1 из теоремы 2.2.] (Этот результат принадлежит Варадарайну (1958b); см. обобщения у Ранго Рао (1962).)

5. Покажите, что случайные величины X_n и X удовлетворяют предельному соотношению $X_n \xrightarrow{\mathscr{D}} X$ тогда и только тогда, когда

$$\mathsf{E}\{F(X_n)\} \to \mathsf{E}\{F(X)\} \tag{4.23}$$

для каждой непрерывной функции распределения F. Если случайная величина Y имеет функцию распределения F и не зависит от X и всех X_n (считайте, что все они имеют общую область определения), то (4.23) равносильно тому, что $\mathsf{P}\{Y \leqslant X_n\} \to \mathsf{P}\{Y \leqslant X\}$.

6. Для вероятностной меры P, заданной на $(S, \mathscr{S})$, пример (4.4) показывает, как построить на вероятностном пространстве $(\Omega, \mathscr{B}, \mathsf{P})$ случайный элемент X с распределением P. Если S сепарабельно и полно, можно взять в качестве P лебегову меру на борелевских множествах единичного интервала Ω. [Пусть $\mathscr{A}_k = \{A_{ku}\}$ — разбиение S на множества непрерывности P с диаметром, меньшим чем $1/k$, и $\mathscr{T}_k = \{I_{ku}\}$ — разбиение Ω на подинтервалы длины $\mathsf{P}(I_{ku}) = P(A_{ku})$; выберите разбиения так, чтобы $\mathscr{A}_{k+1}$ было измельчением $\mathscr{A}_k$, а $\mathscr{T}_{k+1}$ — измельчением $\mathscr{T}_k$. Придайте $X_k(\omega)$ для $\omega \in I_{ku}$ некоторое значение из множества A_{ku} и покажите что последовательность $\{X_k(\omega)\}$ фундаментальна при каждом ω; используя тео-

рему 4.1, покажите далее, что $X(\omega) = \lim_k X_k(\omega)$ имеет распределение P.] (Скороход (1956) получил более сильный результат, а именно, если $P_n \Rightarrow P$, то случайные элементы X_n и X с этими распределениями могут быть построены на единичном интервале так, что $X_n(\omega) \to X(\omega)$ при каждом ω.)

7. Покажите, что $(X'_n, X''_n) \xrightarrow{\mathcal{D}} (X', X'')$, если X' и X'' независимы, $X''_n \xrightarrow{P} Y''$, где Y'' имеет то же распределение, что и X'', и соотношение (4.19) выполняется для каждого множества E из σ-поля, порожденного элементом Y''.

§ 5. Слабая сходимость и отображения

Непрерывные отображения. Если h — измеримое отображение S в другое метрическое пространство S' (с метрикой ρ' и σ-полем $\mathcal{S}'$ борелевских множеств), то каждая вероятностная мера P на $(S, \mathcal{S})$ индуцирует на $(S', \mathcal{S}')$ единственную вероятностную меру Ph^{-1}, определяемую равенством $Ph^{-1}(A) = P(h^{-1}A)$ при $A \in \mathcal{S}'$. Нас интересуют условия, при которых из $P_n \Rightarrow P$ следует $P_n h^{-1} \Rightarrow Ph^{-1}$. Одно такое условие состоит в том, что отображение h непрерывно, так как в этом случае функция $f(h(x))$ ограничена и непрерывна на S всякий раз, когда $f(y)$ ограничена и непрерывна на S', так что из $P_n \Rightarrow P$ следует $\int f(h(x))\, P_n(dx) \to \int f(h(x))\, P(dx)$ или, после преобразования интегралов (см. стр. 305),

$$\int f(y)\, P_n h^{-1}(dy) \to \int f(y)\, Ph^{-1}(dy).$$

Например, естественная проекция π_k из R^∞ в R^k непрерывна, так что из сходимости $P_n \Rightarrow P$ следует $P_n \pi_k^{-1} \Rightarrow P\pi_k^{-1}$ при каждом k. Покажем обратное, т. е. что если при каждом k имеет место $P_n \pi_k^{-1} \Rightarrow P\pi_k^{-1}$, то $P_n \Rightarrow P$. Из непрерывности π_k с легкостью следует, что $\partial \pi_k^{-1} H \subset \pi_k^{-1} \partial H$ при $H \subset R^k$. Используя специальные свойства π_k, мы докажем, что имеет место включение в другом направлении. Если $x \in \pi_k^{-1} \partial H$, так что $\pi_k x \in \partial H$, то найдутся точки $\alpha^{(u)}$ в H и точки $\beta^{(u)}$ в H^c такие, что $\alpha^{(u)} \to \pi_k x$ и $\beta^{(u)} \to \pi_k x$ $(u \to \infty)$. Так как точки

$(\alpha_1^{(u)}, \ldots, \alpha_k^{(u)}, x_{k+1}, \ldots)$ принадлежат $\pi_k^{-1}H$ и сходятся к x, а точки $(\beta_1^{(u)}, \ldots, \beta_k^{(u)}, x_{k+1}, \ldots)$ принадлежат $(\pi_k^{-1}H)^c$ и также сходятся к x, то $x \in \partial(\pi_k^{-1}H)$. Таким образом, $\partial\pi_k^{-1}H = \pi_k^{-1}(\partial H)$. Если $P_n\pi_k^{-1} \Rightarrow P\pi_k^{-1}$, то $P_n(A) \to P(A)$ для множеств $A = \pi_k^{-1}H$ с $H \in \mathscr{R}^k$ и $P(\pi_k^{-1}\partial H) = 0$. Так как $P(\pi_k^{-1}\partial H) = 0$ эквивалентно $P(\partial\pi_k^{-1}H) = 0$, то $P_n(A) \to P(A)$ для всех конечномерных множеств непрерывности P. Поскольку конечномерные множества образуют определяющий сходимость класс, то, следовательно, $P_n \Rightarrow P$.

Мы называем $P\pi_k^{-1}$ *конечномерными распределениями* или *мерами*, соответствующими P. Мы показали, что вероятностные меры на $(R^\infty, \mathscr{R}^\infty)$ слабо сходятся тогда и только тогда, когда слабо сходятся все соответствующие конечномерные распределения.

Конечномерные распределения вероятностной меры P на $(C, \mathscr{C})$ мы определяем, как всевозможные меры вида $P\pi_{t_1 \ldots t_k}^{-1}$, где $\pi_{t_1 \ldots t_k}$ являются проекциями, определенными в § 3. Так как эти проекции непрерывны, слабая сходимость вероятностных мер на $(C, \mathscr{C})$ влечет слабую сходимость соответствующих конечномерных распределений. Однако обратное утверждение может не выполняться (как показал противоречащий пример, класс конечномерных множеств не определяет сходимость). Действительно, если $P\,[P_n]$ приписывает единичную массу функции, тождественно равной нулю [соответствующей функции $x_n(t)$, определенной в (3.5)], то P_n не сходится слабо к P, хотя $P_n\pi_{t_1 \ldots t_k}^{-1} \Rightarrow P\pi_{t_1 \ldots t_k}^{-1}$ для всех множеств $(t_1, \ldots, t_k)$. С другой стороны, так как конечномерные множества образуют определяющий класс, то вероятностная мера на $(C, \mathscr{C})$ однозначно определяется своими конечномерными распределениями.

Основная теорема. Мы видели, что из $P_n \Rightarrow P$ следует $P_nh^{-1} \Rightarrow Ph^{-1}$, если h — непрерывное отображение S в S', однако предположение о непрерывности мы можем ослабить. Предположим лишь, что h измеримо, и обозначим через D_h множество точек разрыва h. Имеем $D_h \in S$ (даже в случае, когда h неизмеримо; см. стр. 309).

Теорема 5.1. *Если $P_n \Rightarrow P$ и $P(D_h) = 0$, то $P_n h^{-1} \Rightarrow Ph^{-1}$.*

Доказательство. Мы покажем, что если F — замкнутое подмножество S', то

$$\limsup_{n\to\infty} P_n h^{-1}(F) \leqslant Ph^{-1}(F).$$

Так как $P_n \Rightarrow P$, мы имеем

$$\limsup_n P_n(h^{-1}F) \leqslant \limsup_n P_n\left((h^{-1}F)^-\right) \leqslant P\left((h^{-1}F)^-\right).$$

Следовательно, достаточно доказать, что $P((h^{-1}F)^-) = P(h^{-1}F)$. Последнее равенство является следствием предположения $P(D_h) = 0$ и того факта, что $(h^{-1}F)^- \subset D_h \cup (h^{-1}F)$.

Укажем два прямых следствия доказанной теоремы. Для случайного элемента X из S $h(X)$ является случайным элементом из S' (мы по-прежнему предполагаем, что h измеримо).

Следствие 1. *Если $X_n \xrightarrow{\mathscr{D}} X$ и $\mathsf{P}\{X \in D_h\} = 0$, то $h(X_n) \xrightarrow{\mathscr{D}} h(X)$.*

Следствие 2. *Если $X_n \xrightarrow{P} a$ и если h непрерывно в a, то $h(X_n) \xrightarrow{P} h(a)$.*

Например, для обычных случайных величин из $(X_n, Y_n) \xrightarrow{\mathscr{D}} (X, Y)$ следует $X_n + Y_n \xrightarrow{\mathscr{D}} X + Y$. Факты такого рода, постоянно используемые в теории вероятностей и математической статистике, в конечном счете все связаны с теоремой 5.1.

Особый интерес приобретает теорема 5.1, когда S' является действительной прямой, т. е. h — обычная действительная измеримая функция.

Теорема 5.2. (i). *Если $P_n \Rightarrow P$, то $P_n h^{-1} \Rightarrow Ph^{-1}$ для любой действительной измеримой функции h такой, что $P(D_h) = 0$.*

(ii) *Если $P_n h^{-1} \Rightarrow Ph^{-1}$ для всех ограниченных непрерывных действительных функций h, то $P_n \Rightarrow P$.*

(iii) *Если $P_n \Rightarrow P$ и h — действительная ограниченная измеримая функция такая, что $P(D_h) = 0$, то $\int h\,dP_n \to \int h\,dP$.*

Доказательство. Часть (i) утверждения теоремы следует из теоремы 5.1. Если $P_n h^{-1} \Rightarrow Ph^{-1}$, то заменой переменной получаем, что $\int f(h(x))\, P_n(dx) \to \int f(h(x))\, P(dx)$ для каждой f из $C(R^1)$. Если h ограничена числом M, то, взяв

$$f(t) = \begin{cases} -M & \text{при} \quad t \leqslant -M, \\ t & \text{при} \quad -M \leqslant t \leqslant M, \\ M & \text{при} \quad M \leqslant t, \end{cases}$$

мы находим, что $\int h\, dP_n \to \int h\, dP$. Таким образом, часть (iii) является следствием (i), а (ii) следует из определения слабой сходимости. (Мы можем усилить (ii), заменив непрерывность h на равномерную непрерывность.)

В приложениях мы обычно интересуемся установлением слабой сходимости в различных метрических пространствах для того, чтобы иметь возможность, используя часть (i) настоящей теоремы, доказывать слабую сходимость мер, индуцируемых на действительной прямой с помощью различных действительных функций h.

Предельный переход под знаком интеграла. В части (iii) теоремы 5.2 требование ограниченности может быть ослаблено. Эту проблему проще рассматривать в терминах случайных величин X_n и X, имеющих распределения $P_n h^{-1}$ и Ph^{-1}.

Теорема 5.3. *Если* $X_n \xrightarrow{\mathscr{D}} X$, *то* $\mathsf{E}\{|X|\} \leqslant \liminf_n \mathsf{E}\{|X_n|\}$.

Доказательство. Положим в части (iii) теоремы 5.2

$$h(x) = \begin{cases} |x| & \text{при} \quad |x| \leqslant \alpha, \\ 0 & \text{при} \quad |x| > \alpha; \end{cases}$$

если $\mathsf{P}\{|X| = \alpha\} = 0$, то

$$\int\limits_{\{|X| \leqslant \alpha\}} |X|\, d\mathsf{P} = \lim_{n\to\infty} \int\limits_{\{|X_n| \leqslant \alpha\}} |X_n|\, d\mathsf{P} \leqslant \liminf_{n\to\infty} \mathsf{E}\{|X_n|\}.$$

Если мы устремим α к бесконечности по значениям, удовлетворяющим условию $\mathsf{P}\{|X| = \alpha\} = 0$, то получим искомый результат.

Величины X_n называются *равномерно интегрируемыми*, если

$$\lim_{\alpha\to\infty} \sup_n \int_{\{|X_n|\geqslant\alpha\}} |X_n|\,d\mathsf{P} = 0.$$

Если X_n равномерно интегрируемы, то

$$\sup_n \mathsf{E}\{|X_n|\} < \infty. \tag{5.1}$$

С другой стороны, если для некоторого положительного ε

$$\sup_n \mathsf{E}\left\{|X_n|^{1+\varepsilon}\right\} < \infty,$$

то X_n равномерно интегрируемы, так как в этом случае

$$\int_{\{|X_n|\geqslant\alpha\}} |X_n|\,d\mathsf{P} \leqslant \frac{1}{\alpha^\varepsilon}\mathsf{E}\left\{|X_n|^{1+\varepsilon}\right\}.$$

Величины X_n равномерно интегрируемы также тогда, когда существует случайная величина Y такая, что $\mathsf{E}\{|Y|\} < \infty$ и

$$\mathsf{P}\{|X_n|\geqslant\alpha\} \leqslant \mathsf{P}\{|Y|\geqslant\alpha\}, \quad n\geqslant 1, \quad \alpha>0, \tag{5.2}$$

поскольку при этом (см. (3) на стр. 305)

$$\int_{\{|X_n|\geqslant\alpha\}} |X_n|\,d\mathsf{P} \leqslant \int_{\{|Y|\geqslant\alpha\}} |Y|\,d\mathsf{P}.$$

Теорема 5.4. *Предположим, что* $X_n \xrightarrow{\mathcal{D}} X$. *Если* X_n *равномерно интегрируемы, то*

$$\mathsf{E}(X_n) \to \mathsf{E}(X); \tag{5.3}$$

если X *и* X_n *неотрицательны и интегрируемы, то из* (5.3) *следует, что* X_n *равномерно интегрируемы.*

Доказательство. Если X_n равномерно интегрируемы, то интегрируемость X следует из (5.1) и теоремы 5.3. Положим в части (iii) теоремы 5.2

$$h_\alpha(x) = \begin{cases} x & \text{при } |x|<\alpha, \\ 0 & \text{при } |x|\geqslant\alpha. \end{cases}$$

Так как $X_n \xrightarrow{\mathcal{D}} X$, то

$$\mathsf{E}\{h_\alpha(X_n)\} \to \mathsf{E}\{h_\alpha(X)\}, \tag{5.4}$$

если $\mathsf{P}\{|X|=\alpha\}=0$. Но

$$\mathsf{E}\{X_n\}-\mathsf{E}\{h_\alpha(X_n)\}=\int\limits_{\{|X_n|\geqslant\alpha\}} X_n\,d\mathsf{P}, \tag{5.5}$$

$$\mathsf{E}\{X\}-\mathsf{E}\{h_\alpha(X)\}=\int\limits_{\{|X|\geqslant\alpha\}} X\,d\mathsf{P}. \tag{5.6}$$

Так как из этих трех соотношений вытекает

$$\limsup_{n\to\infty}|\mathsf{E}\{X_n\}-\mathsf{E}\{X\}|\leqslant$$
$$\leqslant\sup_n\int\limits_{\{|X_n|\geqslant\alpha\}}|X_n|\,d\mathsf{P}+\int\limits_{\{|X|\geqslant\alpha\}}|X|\,d\mathsf{P},$$

то (5.3) действительно следует из условия равномерной интегрируемости.

С другой стороны, если X и X_n интегрируемы и (5.3) выполняется, то из (5.4), (5.5) и (5.6) следует, что

$$\int\limits_{\{|X_n|\geqslant\alpha\}} X_n\,d\mathsf{P}\to\int\limits_{\{|X|\geqslant\alpha\}} X\,d\mathsf{P}, \tag{5.7}$$

если $\mathsf{P}\{|X|=\alpha\}=0$. Так как величина X интегрируема, то мы можем при данном ε выбрать α так, чтобы правая часть (5.7) была меньше ε. Но тогда левая часть меньше ε при всех n, превосходящих некоторое n_0. Так как X_n неотрицательны и по отдельности интегрируемы, то отсюда следует равномерная интегрируемость.

Поскольку из $X_n\xrightarrow{\mathscr{D}}X$ вытекает $X_n^r\xrightarrow{\mathscr{D}}X^r$, то теоремы 5.3 и 5.4 непосредственно обобщаются на моменты более высокого порядка, чем первый.

Ни в одной из этих теорем не требуется, чтобы X и X_n (и Y в (5.2)) имели общую область определения. Если же это так и если $X_n(\omega)\to X(\omega)$ для почти всех ω или если $X_n\xrightarrow{P}X$, то $X_n\xrightarrow{\mathscr{D}}X$. Таким образом, теорема 5.3 содержит в себе лемму Фату, а теорема 5.4 содержит стандартные теоремы о переходе к пределу под знаком интеграла (вариант с (5.2) содержит теорему Лебега о мажорированной сходимости).

Обобщение теоремы 5.1 *). Пусть h_n и h — измеримые отображения S в S'. Можно ли утверждать, что из

*) Этот результат в дальнейшем не используется.

$P_n \Rightarrow P$ следует $P_n h_n^{-1} \Rightarrow P h^{-1}$, когда h_n сходится к h в некотором смысле? Пусть E — множество таких x, для которых предельное соотношение $h_n x_n \to hx$ не выполняется для некоторой последовательности $\{x_n\}$, сходящейся к x. (Если h_n тождественно равно h, то $E = D_h$.) Нетрудно показать, что $x \in E^c$ тогда и только тогда, когда для любого ε существуют k и δ такие, что при $i \geqslant k$ и $\rho(x, y) < \delta$ имеет место $\rho'(hx, h_i y) < \varepsilon$. Предположим, что E принадлежит $\mathcal{S}$ (что, как показано на стр. 309, выполняется автоматически, если S' сепарабельно).

Теорема 5.5. *Если* $P_n \Rightarrow P$ *и* $P(E) = 0$, *то* $P_n h_n^{-1} \Rightarrow P h^{-1}$.

Доказательство. Мы покажем, что $P(h^{-1}G) \leqslant \leqslant \liminf_n P_n(h_n^{-1}G)$ для открытых множеств G из S'. Из вышеприведенной характеризации точек E^c следует, что если $x \in E^c$ и если hx лежит в открытом множестве G, то найдутся k и δ такие, что $h_i y \in G$ при $i \geqslant k$ и $\rho(x, y) < \delta$, так что x является внутренней точкой множества $T_k = \bigcap_{i \geqslant k} h_i^{-1} G$. Таким образом, $h^{-1}G \subset E \cup \cup \bigcup_k \overset{\circ}{T}_k$. Поскольку $P(E) = 0$, мы имеем $P(h^{-1}G) \leqslant \leqslant P\left(\bigcup_k \overset{\circ}{T}_k\right)$. Так как $\overset{\circ}{T}_k \subset \overset{\circ}{T}_{k+1}$, то при любом данном $\varepsilon > 0$ $P(h^{-1}G) < P(\overset{\circ}{T}_k) + \varepsilon$ для достаточно больших значений k. Из $P_n \Rightarrow P$ следует, что $P(\overset{\circ}{T}_k) \leqslant \liminf_n P_n(\overset{\circ}{T}_k)$; так как $\overset{\circ}{T}_k \subset h^{-1}G$ при всех достаточно больших n, то мы получаем $P(\overset{\circ}{T}_k) \leqslant \liminf_n P_n(h_n^{-1}G)$. Таким образом, справедливо неравенство $P(h^{-1}G) \leqslant \liminf_n P_n(h_n^{-1}G) + \varepsilon$, что в силу произвольности ε и завершает доказательство.

Если $h_n = h$, то этот результат сводится к теореме 5.1. Если h всюду непрерывна и h_n сходится к h равномерно на компактных множествах, то E пусто, так что условие теоремы удовлетворено. Более общим образом, $P(E) = 0$, если $P(D_h) = 0$ и если имеет место равномерная сходимость на компактных множествах. С другой стороны, если $D_h = E = 0$, то h_n сходится к h равномерно на компактных множествах.

Примечание. В случае евклидова пространства теорема 5.1 иногда называется теоремой Манна — Вальда (см. Манн, Вальд (1943) и Чернов (1956)); следствие 2 для рациональных функций h представляет собой теорему Слуцкого (см. Слуцкий (1925)). Теорема 5.4, которая обобщает результат Прохорова (1956), принадлежит Г. Рубину; см. Андерсон (1963). По поводу обобщений теорем 5.1 и 5.5 и обратных им теорем см. Топсё (1967а и 1967b).

Задачи

1. Если борелевская функция f имеет производную в точке a (причем это единственное предположение относительно f) и $X_n \xrightarrow{P} a$, то $f(X_n) = f(a) + f'(a)(X_n - a) + (X_n - a) Z_n$, где $Z_n \xrightarrow{P} 0$. Обобщите это на производные более высокого порядка.

2. Предположим, что f имеет производную в точке 0. Если $X_n Y_n \xrightarrow{\mathscr{D}} Y$ и $Y_n \xrightarrow{P} 0$, то

$$X_n (f(Y_n) - f(0)) \xrightarrow{\mathscr{D}} f'(0) Y.$$

3. Предположим, что f_n и f имеют непрерывные производные, причем $f'_n(x) \to f'(x)$ равномерно по x. Если $X_n Y_n \xrightarrow{\mathscr{D}} Y$ и $Y_n \xrightarrow{P} 0$, то

$$X_n (f_n(Y_n) - f_n(0)) \xrightarrow{\mathscr{D}} f'(0) Y.$$

4. Предположим, что $X_n \xrightarrow{\mathscr{D}} X$. Приведите примеры, когда X_n интегрируемы, а X неинтегрируема, и наоборот. Покажите, что из $\mathsf{E}\{X_n\} \to \mathsf{E}\{X\}$, вообще говоря, не вытекает равномерная интегрируемость X_n.

5. Покажите, что X_n равномерно интегрируемы тогда и только тогда, когда $\sup_n \mathsf{E}\{|X_n|\} < \infty$, и для любого ε найдется такое δ, что при всех n из $\mathsf{P}(E) < \delta$ следует $\int_E |X_n| \, d\mathsf{P} < \varepsilon$.

6. Пусть P — мера Лебега на единичном интервале, а P_n приписывает массы $1/n$ некоторым точкам, выбранным по одной в интервалах $((i-1)/n, i/n)$, $i = 1, \ldots, n$. Покажите, что $P_n \Rightarrow P$, и выведите из теоремы 5.2, что ограниченная функция, непрерывная почти всюду, интегрируема по Риману.

§ 6. Теорема Прохорова

Относительная компактность. Пусть Π — семейство вероятностных мер на $(S, \mathscr{S})$. Мы назовем Π *относительно компактным*, если *любая последовательность элементов* Π *содержит слабо сходящуюся подпоследовательность*, т. е. если для любой последовательности $\{P_n\}$ в Π

существуют подпоследовательность $\{P_{n'}\}$ и вероятностная мера Q (определенная на $(S, \mathscr{S})$, но не обязательно являющаяся элементом П) такие, что $P_{n'} \Rightarrow Q$ *). Хотя сходимость $P_{n'} \Rightarrow Q$ не имеет смысла при $Q(S) < 1$, мы еще раз подчеркиваем, что $Q(S) = 1$ (как это обсуждается ниже, мы не допускаем никакого исчезновения массы).

Нам необходимо уметь определять, является ли заданное семейство П относительно компактным или нет. Например, допустим, что нам известно про вероятностные меры P_n и P, определенные на $(C, \mathscr{C})$, что конечномерные распределения P_n слабо сходятся к конечномерным распределениям P. Мы видели в § 5, что мера P_n не обязана при этом сходиться слабо к P. Предположим, однако, что нам также известно, что последовательность $\{P_n\}$ относительно компактна. Тогда любая подпоследовательность $\{P_{n'}\}$ содержит другую подпоследовательность $\{P_{n''}\}$ и слабо сходящуюся к некоторому пределу Q. Конечномерные распределения Q должны быть слабыми пределами конечномерных распределений $\{P_{n''}\}$ и следовательно, должны совпадать с конечномерными распределениями P $\left(P_{n''}\pi^{-1}_{t_1\ldots t_k} \Rightarrow Q\pi^{-1}_{t_1\ldots t_k}\right.$ и $P_n\pi^{-1}_{t_1\ldots t_k} \Rightarrow P\pi^{-1}_{t_1\ldots t_k}$ для любого набора $(t_1, \ldots, t_k)$, так что $\left.Q\pi^{-1}_{t_1\ldots t_k} = P\pi^{-1}_{t_1\ldots t_k}\right)$. Но тогда, так как вероятностная мера на C полностью определяется своими конечномерными распределениями (конечномерные распределения образуют определяющий класс), Q и P сами должны совпадать. Таким образом, любая подпоследовательность $\{P_{n'}\}$ содержит другую подпоследовательность, слабо сходящуюся к P, и из теоремы 2.3 следует, что вся последовательность $\{P_n\}$ слабо сходится к P. Заметим, что если последовательность $\{P_n\}$ действительно слабо сходится к P, то она относительно компактна, так что требование относительной компактности не является чрезмерным.

Допустим, нам известно, что последовательность $\{P_n\}$ относительно компактна и что для любого набора

*) Мы берем это предложение в качестве *определения*; на самом деле это есть понятие *секвенциальной* компактности в пространстве $Z(S)$, введенном в сноске на стр. 21. Дальнейшие топологические сведения см. в Добавлении III.

$(t_1, \ldots, t_k)$ $P_n\pi^{-1}_{t_1 \ldots t_k}$ слабо сходится к некоторой вероятностной мере $\mu_{t_1 \ldots t_k}$ на $(R^k, \mathscr{R}^k)$. Подчеркнем, что мы не предполагаем заранее, что $\mu_{t_1 \ldots t_k}$ являются конечномерными распределениями какой-либо вероятностной меры на $(C, \mathscr{C})$. Как и выше, любая подпоследовательность $\{P_{n'}\}$ содержит другую подпоследовательность $\{P_{n''}\}$, слабо сходящуюся к некоторому пределу. Так как этот предел должен иметь $\mu_{t_1 \ldots t_k}$ своими конечномерными распределениями, он является единственным. Поэтому $\{P_n\}$ слабо сходится к *некоторой* мере P.

Эти идеи дают нам мощный метод для доказательства слабой сходимости в C и других функциональных пространствах. Сначала доказывается, что конечномерные распределения слабо сходятся, а затем, что рассматриваемая последовательность относительно компактна. Для того чтобы использовать этот метод, нам необходим эффективный критерий относительной компактности

Рассмотрим вначале R^1. Пусть Π — семейство вероятностных мер на $(R^1, \mathscr{R}^1)$. Если дана последовательность $\{P_n\}$ из Π, то мы можем применить к последовательности $\{F_n\}$ соответствующих функций распределения классический принцип выбора Хелли (см. стр. 311). Существует подпоследовательность $\{F_{n'}\}$ и некоторая функция F такие, что

$$F_{n'}(x) \to F(x) \tag{6.1}$$

справедливо для всех точек x, где F непрерывна. Функцию F можно взять непрерывной справа, и в этом случае на $(R^1, \mathscr{R}^1)$ существует конечная мера μ такая, что

$$\mu(a, b] = F(b) - F(a). \tag{6.2}$$

Может, однако, случиться, что $\mu(R^1) < 1$. Например, если P_n соответствует единичной массе в точке n, то F тождественно равна 0, вне зависимости от выбора подпоследовательности $\{F_{n'}\}$, так что $\mu(R^1) = 0$. Если P_n есть равномерное распределение на $[-n, n]$, то единственной возможностью является тождество $F(x) \equiv 1/2$ — и снова $\mu(R^1) = 0$. В этих примерах масса «ускользает в бесконечность» в очевидном и интуитивно понятном смысле.

Предположим, однако, что мера μ, определенная с помощью F в (6.2), удовлетворяет условию $\mu(R^1) = 1$. Тогда μ является вероятностной мерой, и поскольку (6.1) справедливо в точках непрерывности F, мы имеем (см. § 3) $P_{n'} \Rightarrow \mu$. Таким образом, П будет относительно компактно, если мы каким-либо образом обеспечим, чтобы каждая из мер μ, с которой мы встретимся, удовлетворяла бы условию $\mu(R^1) = 1$. Далее, $\mu(R^1) = 1$, если для любого положительного ε существуют a и b такие, что $\mu(a, b] \geqslant 1 - \varepsilon$. Предположим, что для каждого положительного ε существуют такие a и b, что

$$P_n(a, b] > 1 - \varepsilon, \quad n = 1, 2, \ldots \tag{6.3}$$

Поскольку неравенство (6.3) сохраняется, если a уменьшается, а b увеличивается, то можно считать a и b точками непрерывности F; в этом случае из (6.3) и (6.1) вытекает неравенство $\mu(a, b] \geqslant 1 - \varepsilon$.

Отсюда следует, что семейство П вероятностных мер на $(R^1, \mathscr{R}^1)$ относительно компактно, если для каждого положительного ε найдутся a и b такие, что $P(a, b] > 1 - \varepsilon$ для всех мер P из П; это условие предотвращает «утечку массы», о которой говорилось выше. С другой стороны, если это условие не выполняется, то существует некоторое положительное ε такое, что независимо от того, каковы числа a и b, $P(a, b] \leqslant 1 - \varepsilon$ для некоторой меры P из П. Выберем в П такую меру P_n, что $P_n(-n, n] \leqslant 1 - \varepsilon$. Если бы некоторая подпоследовательность $\{P_{n'}\}$ слабо сходилась к вероятностной мере Q, то мы получили бы для каждого x

$$Q(-x, x) \leqslant \liminf_{n'} P_{n'}(-x, x) \leqslant$$
$$\leqslant \liminf_{n'} P_{n'}(-n', n'] \leqslant 1 - \varepsilon,$$

что невозможно. Таким образом, рассматриваемое условие одновременно необходимо и достаточно.

Так как интервал $(a, b]$ имеет компактное замыкание и так как каждое компактное множество может быть заключено в такой интервал, то рассматриваемому условию можно придать следующую форму: семейство вероятностных мер на $(R^1, \mathscr{R}^1)$ относительно компактно тогда и только тогда, когда для любого положительного ε найдется компактное множество K такое, что $P(K) > 1-\varepsilon$ для всех P из Π. Теперь это условие имеет смысл в произвольном метрическом пространстве.

Семейство Π вероятностных мер на общем метрическом пространстве S называется *плотным*, если для каждого положительного ε существует компактное множество K такое, что $P(K) > 1-\varepsilon$ для всех P из Π. Если Π состоит только из единственной меры, это определение сводится к определению, сформулированному в § 1. Мы только что убедились в том, что в случае R^1 условие плотности является необходимым и достаточным для относителной компактности. Следующие теоремы, принадлежащие Прохорову, распространяют достаточность на произвольные метрические пространства и необходимость на сепарабельные полные пространства.

Теорема 6.1. *Если* Π *плотно, то оно относительно компактно.*

Теорема 6.2. *Предположим, что* S *сепарабельно и полно. Если* Π *относительно компактно, то оно плотно.*

Мы будем ссылаться на эти две теоремы, как на теорему Прохорова, называя теорему 6.1 и теорему 6.2, соответственно, прямой и обратной половинами теоремы Прохорова.

Прямая теорема. Докажем сначала прямую теорему, которая более полезна для приложений. Мы установим ее последовательно для R^k, для R^∞, для σ-компактного S (т. е. для S, являющегося счетным объединением компактных множеств) и, наконец, для общего случая. Каждый из последних трех случаев изучается путем сведения к предыдущему.

Случай R^k. Здесь доказательство практически не отличается от уже упомянутого доказательства для R^1. Если $\{P_n\}$ — последовательность мер из Π, то из теоремы Хелли (стр. 311) следует, что последовательность $\{F_n\}$ соответствующих функций распределения содержит подпоследовательность $\{F_{n'}\}$ такую, что

$$F_{n'}(x) \to F(x) \tag{6.4}$$

в точках непрерывности F, где F — непрерывная сверху функция. На $(R^k, \mathscr{R}^k)$ существует мера μ такая, что $\mu(a, b]$ вычисляется через смешанные разности значений

F в вершинах k-мерного прямоугольника $(a, b]$ *). Если мы докажем теперь, что $\mu(R^k) = 1$, то получим $P_{n'} \Rightarrow \mu$.

При данном ε выберем компактное множество K и R^k такое, что $P_{n'}(K) > 1 - \varepsilon$ для всех n', что возможно вследствие плотности семейства Π. Затем выберем a и b такими, что $K \subset (a, b]$, и такими, что все 2^k вершин прямоугольника $(a, b]$ являются точками непрерывности F (это возможно в силу того, что только счетное множество параллельных друг другу $(k-1)$-мерных гиперплоскостей может иметь положительную μ-меру). Так как $P_{n'}(a, b]$ вычисляется как смешанная разность значений $F_{n'}$ в вершинах $(a, b]$, то из (6.4) следует, что $P_{n'}(a, b] \to \mu(a, b]$. Из $P_{n'}(a, b] \geqslant P_{n'}(K) > 1 - \varepsilon$ следует теперь, что $\mu(a, b] \geqslant 1 - \varepsilon$. Так как ε произвольно, то $\mu(R^k) = 1$. Таким образом, семейство Π относительно компактно.

В случае R^∞ нам понадобится следующая лемма.

Лемма 1. *Если Π — плотное семейство на $(S, \mathcal{S})$ и если h — непрерывное отображение S в S', то $\{Ph^{-1}: P \in \Pi\}$ есть плотное семейство на $(S', \mathcal{S}')$.*

Доказательство. При данном ε выберем в S компактное множество K такое, что $P(K) > 1 - \varepsilon$ при всех P из Π. Если $K' = hK$, то K' компактно (см. стр. 298) и $h^{-1}K' \supset K$, так что $Ph^{-1}(K') > 1 - \varepsilon$ для всех P из Π.

Случай R^∞. Если Π — плотное семейство на $(R^\infty, \mathcal{R}^\infty)$, то согласно лемме 1 $\{P\pi_k^{-1}: P \in \Pi\}$ является при каждом k плотным семейством на $(R^k, \mathcal{R}^k)$. Согласно прямой части теоремы Прохорова для R^k, только что доказанной с помощью теоремы Хелли, мы можем выбрать из заданной последовательности $\{P_n\}$ элементов Π подпоследовательность $\{P_{n'}\}$ такую, что $P_{n'}\pi_k^{-1}$ слабо сходится к вероятностной мере μ_k на $(R^k, \mathcal{R}^k)$. С помощью диагонального метода (как на стр. 300) мы можем выбрать последовательность $\{P_{n'}\}$ так, чтобы было $P_{n'}\pi_k^{-1} \Rightarrow \Rightarrow \mu_k$ для всех k.

Поскольку меры μ_k, очевидно, удовлетворяют условиям согласованности теоремы Колмогорова (см.

---

*) Полагая $h_i = b_i - a_i$, имеем $\mu(a, b] = F(a_1 + h_1, \ldots, a_k + h_k) - F(a_1, a_2 + h_2, \ldots) - F(a_1 + h_1, a_2, \ldots) + \ldots + (-1)^k F(a_1, \ldots, a_k)$. (*Прим. перев.*)

стр. 312), то существует вероятностная мера Q на $(R^\infty, \mathcal{R}^\infty)$ такая, что $Q\pi_k^{-1} = \mu_k$ для всех k. (В § 3 было показано, что σ-поле $\mathcal{R}^\infty$ борелевских множеств в R^∞ совпадает с σ-полем, порожденным конечномерными множествами, и это есть то σ-поле, которое участвует в теореме Колмогорова.) Но тогда $P_{n'}\pi_k^{-1} \Rightarrow Q\pi_k^{-1}$ для каждого k, так что конечномерные распределения $P_{n'}$ сходятся к конечномерным распределениям Q, откуда, как отмечено в начале § 5, вытекает $P_{n'} \Rightarrow Q$. Итак, плотность влечет относительную компактность в R^∞.

Для σ-компактного и общего случая нам понадобятся еще две леммы. Предположим, что S_0 есть борелевское подмножество S:

$$S_0 \in \mathcal{S}. \tag{6.5}$$

Множество S_0 само является метрическим пространством в относительной топологии. Обозначим через $\mathcal{S}_0$ класс борелевских множеств для S_0. Из (6.5) следует (см. стр. 307), что

$$\mathcal{S}_0 = \{A\colon\ A \subset S_0,\ A \in \mathcal{S}\} \tag{6.6}$$

и, в частности,

$$\mathcal{S}_0 \subset \mathcal{S}. \tag{6.7}$$

Если P — вероятностная мера на $(S, \mathcal{S})$ с $P(S_0) = 1$, то пусть P^r (здесь индекс r обозначает сужение — restriction) есть вероятностная мера на $(S_0, \mathcal{S}_0)$, полученная сужением области определения $\mathcal{S}$ до $\mathcal{S}_0$ (см. (6.7)). Если P — вероятностная мера на $(S_0, \mathcal{S}_0)$, то пусть P^e (e означает продолжение — extension) есть вероятностная мера на $(S, \mathcal{S})$ с $P^e(A) = P(A \cap S_0)$ для $A \in \mathcal{S}$ (см. (6.6)). Отметим, что $P^e(S_0) = 1$.

Если P — вероятностная мера на $(S, \mathcal{S})$ с $P(S_0) = 1$, то

$$(P^r)^e = P; \tag{6.8}$$

если P — вероятностная мера на $(S_0, \mathcal{S}_0)$, то

$$(P^e)^r = P. \tag{6.9}$$

Если в лемме 1 и теореме 5.1 мы возьмем в качестве h тождественное отображение S_0 в S, то получим следующую лемму:

Лемма 2. *Если* Π — *плотное семейство мер на* $(S_0, \mathscr{S}_0)$, *то* $\Pi^e = \{P^e\colon\ P \in \Pi\}$ *есть плотное семейство на* $(S, \mathscr{S})$. *Если* $P_n \Rightarrow P$ *в* $(S_0, \mathscr{S}_0)$, *то* $P_n^e \Rightarrow P^e$ *в* $(S, \mathscr{S})$.

Лемма 3. *Если* $P_n \Rightarrow P$ *в* $(S, \mathscr{S})$ *и* $P_n(S_0) = P(S_0) = 1$, *то* $P_n^r \Rightarrow P^r$ *в* $(S_0, \mathscr{S}_0)$.

Доказательство. Всякое открытое множество в S_0 можно представить в виде $G_0 = G \cap S_0$, где G — открытое множество в S. Поскольку $P_n^r(G_0) = P_n(G)$ и $P^r(G_0) = P(G)$, то из $\liminf_n P_n(G) \geqslant P(G)$ следует, что $\liminf_n P_n^r(G_0) \geqslant P^r(G_0)$.

σ-компактный случай. Если пространство S σ-компактно, то S сепарабельно и, следовательно, может быть гомеоморфно вложено в R^∞ (см. стр. 300). Так как S σ-компактно, то таков же его образ при гомеоморфизме; в частности, этот образ является борелевским подмножеством R^∞. Таким образом, S гомеоморфно борелевскому подмножеству R^∞. Согласно теореме 5.1 слабая сходимость сохраняется при гомеоморфизме и, следовательно, сохраняется и относительная компактность Π. Поскольку компактность множеств сохраняется при гомеоморфизме, то сохраняется и плотность Π. Поэтому мы можем заменить S на его гомеоморфный образ.

Мы можем, таким образом, предположить, что S — борелевское подмножество R^∞. Если Π плотно на $(S, \mathscr{S})$, то по лемме 2, примененной к R^∞ и его подмножеству S, Π^e плотно в $(R^\infty, \mathscr{R}^\infty)$. Поскольку, как уже доказано, теорема 6.1 справедлива в этом объемлющем пространстве, то Π^e относительно компактно, так что для любой последовательности $\{P_n\}$ в Π соответствующая последовательность $\{P_n^e\}$ содержит подпоследовательность $\{P_{n'}^e\}$, слабо сходящуюся в смысле $(R^\infty, \mathscr{R}^\infty)$ к некоторой мере Q. В силу плотности самого семейства Π для каждого ε существует компактное подмножество K в S такое, что $P_{n'}^e(K) = P_{n'}(K) > 1 - \varepsilon$ для всех n', так что $Q(S) \geqslant Q(K) \geqslant \limsup_{n'} P_{n'}^e(K) \geqslant 1 - \varepsilon$. Таким образом, S является носителем Q, так же как и всех $P_{n'}^e$, и, следовательно, по лемме 3 и (6.9) $P_{n'}$ сходится слабо к Q^r в смысле $(S, \mathscr{S})$. Таким образом, из плотности следует относительная компактность, если S σ-компактно *).

Общий случай. Каково бы ни было S, если положить $S_0 = \bigcup_i K_i$, где K_i — такое компактное множество в S, что $P(K_i) > 1 - 1/i$ для всех P из Π, то S_0 имеет меру, равную 1, относительно любого элемента Π и $\Pi^r = \{P^r : P \in \Pi\}$ есть плотное семейство в $(S_0, \mathscr{S}_0)$. В соответствии с только что разобранным случаем Π^r относительно компактно. Поэтому для каждой последовательности $\{P_n\}$ в Π соответствующая последовательность $\{P_n^r\}$ содержит подпоследовательность $\{P_{n'}^r\}$, слабо сходящуюся в смысле $(S_0, \mathscr{S}_0)$ к некоторой мере Q. По лемме 2 и (6.8) $P_{n'}$ слабо сходятся в смысле $(S, \mathscr{S})$ к Q^e. Таким образом, Π относительно компактно, что доказывает теорему 6.1 в полной общности.

Обратная теорема. Перейдем теперь к доказательству теоремы 6.2, которая представляет собой обратное утверждение теоремы Прохорова. Отметим, что эта теорема в случае, когда Π состоит из единственной меры, сводится к теореме 1.4; доказательство является усовершенствованием ранее данного доказательства.

Допустим, что для любых положительных ε и δ существует конечная совокупность $A_1, \ldots, A_n$ δ-шаров таких, что $P\left(\bigcup_{i \leqslant n} A_i\right) > 1 - \varepsilon$ для всех P из Π. Нижеследующее рассуждение показывает, что в этом случае Π плотно. При данном ε выберем для каждого целого числа k конечное множество $1/k$-шаров $A_{k1}, \ldots, A_{kn_k}$ таких, что $P\left(\bigcup_{i \leqslant n_k} A_{ki}\right) > 1 - \varepsilon/2^k$ для всех P из Π. Если K — замыкание вполне ограниченного множества $\bigcap_{k \geqslant 1} \bigcup_{i \leqslant n_k} A_{ki}$, то $P(K) > 1 - \varepsilon$ и, так как S предполагается полным, K есть компакт (см. стр. 296).

Мы докажем теорему 6.2, показав, что если условие, установленное в предыдущем абзаце, не выполняется, то Π не является относительно компактным. Допустим, что

---

*) Сепарабельное пространство S гомеоморфно подмножеству R^∞. Более сильное предположение о σ-компактности S обеспечивается принадлежностью его образа пространству $\mathscr{R}^\infty$; это удобно, так как мы опираемся на леммы 2 и 3.

существуют такие положительные ε и δ, что любая конечная совокупность $A_1, \ldots, A_n$ δ-шаров удовлетворяет условию $P\left(\bigcup_{i \leqslant n} A_i\right) \leqslant 1 - \varepsilon$ для некоторой P из Π. Поскольку S по предположению сепарабельно, то оно представляет собой объединение последовательности A_1, $A_2, \ldots$ открытых шаров радиуса δ. Пусть $B_n = \bigcup_{i \leqslant n} A_i$. Выберем P_n в Π так, что $P_n(B_n) \leqslant 1 - \varepsilon$. Предположим, что некоторая подпоследовательность $\{P_{n'}\}$ слабо сходится к некоторому пределу Q. Так как B_m открыто, то мы имели бы тогда $P(B_m) \leqslant \liminf_{n'} P_{n'}(B_m)$ при каждом фиксированном m. Но тогда, поскольку при всех достаточно больших n' $B_m \subset B_{n'}$, мы имели бы $P(B_m) \leqslant \leqslant \liminf_{n'} P_{n'}(B_{n'}) \leqslant 1 - \varepsilon$; но B_m возрастают к S, следовательно, это невозможно, что и завершает доказательство.

Усиленные варианты теоремы 6.2 см. в Добавлении III.

Пусть X_n — случайные элементы S. Мы называем последовательность $\{X_n\}$ плотной, если плотна последовательность $\{P_n\}$, где P_n есть распределение X_n. Если S — это R^∞ или C, мы отождествляем конечномерные распределения X_n с конечномерными распределениями P_n. То, что изложено в начале этого параграфа, может быть интерпретировано в терминах случайных элементов X_n и X из C следующим образом: если конечномерные распределения X_n сходятся слабо к конечномерным распределениям X и последовательность $\{X_n\}$ плотна, то $X_n \xrightarrow{\mathscr{D}} X$.

Примечание. В своем первоначальном доказательстве теоремы 6.1 Прохоров (1956) предполагал, что S сепарабельно и полно; данное здесь обобщение его доказательства принадлежит Варадарайну (1958а и 1961а). См. также Ле Кам (1960).

Задачи

1. Если семейство Π плотно, то его элементы обладают общим σ-компактным носителем. Обратное неверно (исключая некоторые случаи, например, когда Π состоит из единственной меры).

2. Пусть Π состоит из мер, каждая из которых приписывает единичную массу какой-либо из точек множества A. Выведите непосредственно из определения, что Π относительно компактно тогда и только тогда, когда A^- есть компакт в S.

3. Последовательность вероятностных мер на прямой плотна тогда и только тогда, когда соответствующие функции распределения удовлетворяют соотношениям $\lim_{x\to\infty} F_n(x) = 1$ и $\lim_{x\to-\infty} F_n(x) = 0$ равномерно по n. Класс нормальных распределений является плотным тогда и только тогда, когда средние и дисперсии ограничены.

4. Покажите, что если F — непрерывная справа, неубывающая действительная функция, причем $0 \leqslant F(x) \leqslant 1$, то существуют функции распределения F_n такие, что $F_n(x) \to F(x)$ в любой точке непрерывности x.

5. Если последовательность $\{|X_n|^\delta\}$ равномерно интегрируема (см. стр. 51) при некотором $\delta > 0$, то последовательность $\{X_n\}$ плотна.

6. Вероятностные меры на произведении $S' \times S''$ образуют плотное семейство тогда и только тогда, когда оба семейства маргинальных распределений плотны в S' и S'', соответственно.

7. Допустим, что S сепарабельно и локально компактно. Поскольку S может быть снабжено метрикой, в которой оно полно [задача 3 § 1], то теорема 6.2 применима. Из того факта, что любое компактное множество в S может быть включено в множество внутренних точек более широкого компактного множества, следует, что $P_n \Rightarrow P$ тогда и только тогда, когда $\int f\, dP_n \to \int f\, dP$ для каждой (ограниченной) непрерывной функции f с компактным носителем (результат, который невозможен, например, в C, так как в C непрерывная функция с компактным носителем тождественно равна 0; см. задачу 5 в § 3).

8. В связи с леммой 3 отметим следующее. Пусть Π — плотное семейство на $(S, \mathcal{S})$ и $P(S_0) = 1$ для всех $P \in \Pi$. Отсюда не следует, вообще говоря, что $\Pi^r = \{P^r\colon P \in \Pi\}$ есть плотное семейство на $(S_0, \mathcal{S}_0)$. [Возьмите $S = [0, 1]$ и $S_0 = (0, 1)$ и рассмотрите меры с одноточечными носителями.] Соответствующий пример можно привести даже тогда, когда Π состоит из единственной меры. [См. замечание 2 к теореме 1 в Добавлении III.]

§ 7. Первые приложения

В этом параграфе мы рассматриваем с точки зрения предыдущей теории некоторые хорошо известные вероятностные результаты, которые часто используются в последующих главах.

Гладкие функции. При доказательстве теорем 1.2 и 1.3 и при доказательстве импликации (ii) → (iii) в теореме 2.1 мы использовали функцию (1.1). Эти рассуждения по-прежнему проходят, если вместо (1.1) мы используем любую равномерно непрерывную функцию φ такую, что $\varphi(t) = 1$ при $t \leqslant 0$, $0 \leqslant \varphi(t) \leqslant 1$ при $0 \leqslant t \leqslant 1$ и $\varphi(t) = 0$ при $t \geqslant 1$. Можно построить такие функции со свойствами гладкости более сильными, чем равномерная непрерывность. Определим φ вне $[0,1]$, как и раньше, а при $0 \leqslant t \leqslant 1$ положим

$$\varphi(t) = \alpha^{-1} \int_t^1 e^{-1/s(1-s)}\, ds, \tag{7.1}$$

где

$$\alpha = \int_0^1 e^{-1/s(1-s)}\, ds.$$

Тогда при любом целом v функция φ имеет на всей прямой ограниченную непрерывную производную порядка v.

Пусть P_n и P — вероятностные меры на $(R^1, \mathscr{R}^1)$.

Теорема 7.1. *Если $\int f\, dP_n \to \int f\, dP$ для любой ограниченной непрерывной функции f, обладающей ограниченными непрерывными производными любого порядка, то $P_n \Rightarrow P$.*

Доказательство. Рассмотрим функции распределения F_n и F, соответствующие P_n и P. Если $\varphi_u(t) = \varphi(ut)$, где φ определена в (7.1), то при каждом u

$$\limsup_{n\to\infty} F_n(x) \leqslant \lim_{n\to\infty} \int \varphi_u(y-x)\, P_n(dy) =$$
$$= \int \varphi_u(y-x)\, P(dy) \leqslant F\left(x + \frac{1}{u}\right).$$

Используя $\varphi_u(y - x + 1/u)$ вместо $\varphi_u(y-x)$, мы видим, что $\liminf_n F_n(x) \geqslant F(x - 1/u)$ при каждом u. Следовательно, $F_n(x) \to F(x)$, если функция F непрерывна в x, т. е. $P_n \Rightarrow P$.

Центральная предельная теорема. Теорема 7.1 может быть использована для вывода разных вариантов классической центральной предельной теоремы для схемы серии. Пусть при каждом n

$$\xi_{n1}, \ldots, \xi_{nk_n} \tag{7.2}$$

— независимые случайные величины со средним 0 и конечной дисперсией σ_{nk}^2. (Вероятностное пространство, на котором определяются величины (7.2), может меняться в зависимости от n.) Положим $S_n = \xi_{n1} + \ldots + \xi_{nk_n}$ и допустим, что дисперсия этой суммы $s_n^2 = \sigma_{n1}^2 + \ldots + \sigma_{nk_n}^2$ положительна. Напомним, что N обозначает нормально распределенную случайную величину со средним 0 и дисперсией 1. Докажем сначала теорему Линдеберга:

Теорема 7.2. *Если*

$$\frac{1}{s_n^2} \sum_{k=1}^{k_n} \int_{\{|\xi_{nk}| \geqslant \varepsilon s_n\}} \xi_{nk}^2 \, d\mathsf{P} \to 0 \qquad (7.3)$$

$(n \to \infty)$ *при любом положительном* ε, *то* $S_n/s_n \xrightarrow{\mathscr{D}} N$.

Доказательство. Достаточно показать, что

$$\mathsf{E}\{f(S_n/s_n)\} \to \mathsf{E}\{f(N)\} \qquad (7.4)$$

для любой ограниченной непрерывной функции f, обладающей ограниченными непрерывными производными любого порядка. Зафиксируем такую функцию f и определим

$$g(h) = \sup_x \left| f(x+h) - f(x) - f'(x)h - \frac{1}{2} f''(x) h^2 \right| \qquad (7.5)$$

(функция g измерима по Борелю). Согласно известным формулам для остаточного члена формулы Тейлора существует константа K, зависящая только от f, такая, что

$$g(h) \leqslant K \min\{h^2, |h|^3\}. \qquad (7.6)$$

Мы имеем $g(h) \leqslant Kh^2$ и $g(h) \leqslant K|h|^3$; первое неравенство хорошо действует при h большом, а второе — при h, близком к 0. Поскольку f фиксирована всюду в последующих рассуждениях, то и постоянная K также фиксирована. Из определения (7.5) следует, что

$$\left| [f(x+h_1) - f(x+h_2)] - \left[f'(x)(h_1 - h_2) + \frac{1}{2} f''(x)(h_1^2 - h_2^2) \right] \right| \leqslant$$
$$\leqslant g(h_1) + g(h_2). \qquad (7.7)$$

Если бы ξ_{nk} были все нормально распределены, то $\mathsf{E}\{f(S_n/s_n)\}$ совпало бы с $\mathsf{E}\{f(N)\}$. Если мы последовательно заменим ξ_{nk} нормальными величинами η_{nk} со средним 0 и дисперсиями σ_{nk}^2, то мы получим последовательность

$$\begin{aligned}
&\mathsf{E}\left\{f\left(s_n^{-1}\left(\xi_{n1}+\ldots+\xi_{nk_n}\right)\right)\right\},\\
&\mathsf{E}\left\{f\left(s_n^{-1}\left(\xi_{n1}+\ldots+\xi_{nk_n-1}+\eta_{nk_n}\right)\right)\right\},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\mathsf{E}\left\{f\left(s_n^{-1}\left(\xi_{n1}+\eta_{n2}+\ldots+\eta_{nk_n}\right)\right)\right\},\\
&\mathsf{E}\left\{f\left(s_n^{-1}\left(\eta_{n1}+\ldots+\eta_{nk_n}\right)\right)\right\}.
\end{aligned}$$

Первым членом последовательности является $\mathsf{E}\{f(S_n/s_n)\}$, а последним — $\mathsf{E}\{f(N)\}$. Далее развиваемая идея состоит в том, чтобы показать, что каждый член последовательности близок к следующему (при большом n) настолько, что даже первый и последний члены оказываются близкими.

Поскольку в (7.4) существенно только совместное распределение величин (7.2), а свойства вероятностного пространства, на котором они определены, не играют роли, мы можем, переходя к новому пространству (скажем, $\left(R^{2k_n}, \mathscr{R}^{2k_n}\right)$, ввести случайные величины $\eta_{n1}, \ldots, \eta_{nk_n}$ так, что η_{nk} нормально распределена со средним 0 и дисперсией σ_{nk}^2, причем все $2k_n$ случайные величины

$$\xi_{n1}, \ldots, \xi_{nk_n}, \quad \eta_{n1}, \ldots, \eta_{nk_n} \tag{7.8}$$

независимы. Если

$$\zeta_{nk}=\sum_{1\leqslant i<k}\xi_{ni}+\sum_{k<i\leqslant k_n}\eta_{ni}, \qquad 1\leqslant k\leqslant k_n,$$

то, поскольку $\zeta_{nk_n}+\xi_{nk_n}=S_n$ и поскольку $\zeta_{n1}+\eta_{n1}$ имеет распределение s_nN,

$$\left|\mathsf{E}\left\{f\left(\frac{S_n}{s_n}\right)\right\}-\mathsf{E}\{f(N)\}\right|\leqslant$$
$$\leqslant\sum_{k=1}^{k_n}\left|\mathsf{E}\left\{f\left(\frac{\zeta_{nk}+\xi_{nk}}{s_n}\right)-f\left(\frac{\zeta_{nk}+\eta_{nk}}{s_n}\right)\right\}\right|.$$

Вследствие независимости членов последовательности (7.8) три величины ζ_{nk}, ξ_{nk} и η_{nk} являются независимыми при каждом значении k, и поэтому

$$\mathsf{E}\left\{f'\left(\frac{\zeta_{nk}}{s_n}\right)(\xi_{nk}-\eta_{nk})\right\}=\mathsf{E}\left\{f''\left(\frac{\zeta_{nk}}{s_n}\right)(\xi_{nk}^2-\eta_{nk}^2)\right\}=0.$$

Из (7.7) вытекает, что

$$\left|\mathsf{E}\left\{f\left(\frac{S_n}{s_n}\right)\right\}-\mathsf{E}\{f(N)\}\right|\leqslant\sum_{k=1}^{k_n}\mathsf{E}\left\{g\left(\frac{\xi_{nk}}{s_n}\right)+g\left(\frac{\eta_{nk}}{s_n}\right)\right\}.$$

Доказательство, таким образом, будет завершено, если мы покажем, что

$$\sum_{k=1}^{k_n}\mathsf{E}\left\{g\left(\frac{\xi_{nk}}{s_n}\right)\right\}\to 0 \tag{7.9}$$

и

$$\sum_{k=1}^{k_n}\mathsf{E}\left\{g\left(\frac{\eta_{nk}}{s_n}\right)\right\}\to 0. \tag{7.10}$$

При данном $\varepsilon>0$ разобьем математическое ожидание в (7.9) на интеграл по множеству $\{|\xi_{nk}|\leqslant\varepsilon s_n\}$ и интеграл по дополнительному множеству. Используя (7.6), оценим подинтегральную функцию в первом интеграле величиной $K|\xi_{nk}/s_n|^3$ и во втором интеграле — величиной $K|\xi_{nk}/s_n|^2$:

$$\mathsf{E}\left\{g\left(\frac{\xi_{nk}}{s_n}\right)\right\}\leqslant K\varepsilon\frac{\sigma_{nk}^2}{s_n^2}+K\frac{1}{s_n^2}\int\limits_{\{|\xi_{nk}|\geqslant\varepsilon s_n\}}\xi_{nk}^2\,d\mathsf{P}.$$

Таким образом,

$$\sum_{k=1}^{k_n}\mathsf{E}\left\{g\left(\frac{\xi_{nk}}{s_n}\right)\right\}\leqslant K\varepsilon+K\frac{1}{s_n^2}\sum_{k=1}^{k_n}\int\limits_{\{|\xi_{nk}|\geqslant\varepsilon s_n\}}\xi_{nk}^2\,d\mathsf{P}, \tag{7.11}$$

и (7.9) следует из (7.3).

Поскольку (7.11) справедливо также и с η_{nk} вместо ξ_{nk}, то для доказательства (7.10) нам достаточно показать, что

$$\frac{1}{s_n^2}\sum_{k=1}^{k_n}\int\limits_{\{|\eta_{nk}|\geqslant\varepsilon s_n\}}\eta_{nk}^2\,dP \tag{7.12}$$

стремится к 0 при каждом положительном ε. Но (7.12) не превосходит

$$\frac{1}{s_n^2}\sum_{k=1}^{k_n}\frac{1}{\varepsilon s_n}\mathsf{E}\{|\eta_{nk}|^3\}=\frac{1}{\varepsilon s_n^3}\sum_{k=1}^{k_n}\sigma_{nk}^3\mathsf{E}\{|N|^3\}.$$

Поскольку

$$\frac{\sigma_{nk}^2}{s_n^2}\leqslant\varepsilon^2+\frac{1}{s_n^2}\int\limits_{\{|\xi_{nk}|\geqslant\varepsilon s_n\}}\xi_{nk}^2\,dP,$$

из (7.3) вытекает, что $\max\limits_{k\leqslant k_n}\sigma_{nk}/s_n\to 0$, откуда в свою очередь следует, что $\sum\limits_{k=1}^{k_n}\sigma_{nk}^3/s_n^3\to 0$. Таким образом, (7.12) действительно стремится к 0, что и завершает доказательство.

Если случайные величины ξ_{nk} имеют моменты порядка $2+\delta$, то нормированная сумма в (7.3) не превосходит величины $\varepsilon^{-\delta}s_n^{-2-\delta}\sum\limits_{k=1}^{k_n}\mathsf{E}\{|\xi_{nk}|^{2+\delta}\}$, что доказывает теорему Ляпунова:

Теорема 7.3. *Если для некоторого положительного* δ

$$\frac{1}{s_n^{2+\delta}}\sum_{k=1}^{k_n}\mathsf{E}\{|\xi_{nk}|^{2+\delta}\}\to 0,$$

то $S_n/s_n\xrightarrow{\mathcal{D}}N$.

Наконец, из теоремы 7.2 мы можем вывести теорему Линдеберга — Леви:

Теорема 7.4. *Если* $\xi_1, \xi_2, \ldots$ *независимы и одинаково распределены со средним* 0 *и конечной дисперсией* $\sigma^2 > 0$, *то*

$$\frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{n}\xi_k \xrightarrow{\mathscr{D}} N.$$

Для доказательства этого последнего результата положим $k_n = n$ и $\xi_{ni} = \xi_i$; нормированная сумма в (7.3) не превосходит интеграла по множеству $\{|\xi_1| \geqslant \varepsilon\sigma\sqrt{n}\}$ от величины ξ_1^2/σ^2.

Характеристические функции. Пусть P — вероятностная мера на $(R^k, \mathscr{R}^k)$. Характеристическая функция $p(t)$ меры P определяется равенством

$$p(t) = \int e^{it\cdot x}P(dx), \qquad t \in R^k,$$

где $t \cdot x = \sum_{u=1}^{k} t_u x_u$ обозначает скалярное произведение.

Пусть Q — другая вероятностная мера на $(R^k, \mathscr{R}^k)$, и пусть $q(t)$ — ее характеристическая функция. Мы докажем следующую теорему единственности:

Теорему 7.5. *Если* $p(t) = q(t)$ *для всех* t, *то* $P = Q$.

При каждом t $e^{it\cdot x}$ как функция от x является элементом $C(R^k)$ (или ее действительная и мнимая части являются элементами $C(R^k)$). Таким образом, в случае R^k теорема единственности представляет собой усиление теоремы 1.3, которая утверждает, что P определяется значениями $\int f\,dP$ при $f \in C(S)$. Мы будем доказывать единственность без вывода формулы обращения, а с помощью теоремы Вейерштрасса о приближении функций тригонометрическими многочленами.

Мы знаем из § 3, что прямоугольники $(a, b]$ образуют определяющий класс; так как каждый такой прямоугольник представляется как объединение возрастающей последовательности замкнутых прямоугольников, то достаточно проверить, что P и Q совпадают на множествах

$$A = [a_1, b_1] \times \ldots \times [a_k, b_k].$$

В силу рассуждений, использованных при доказательстве теоремы 1.3, это будет так, если для каждого целого u

$$\int f\,dP = \int f\,dQ \tag{7.13}$$

для всех f, имеющих вид $f(x_1, \ldots, x_k) = f_1(x_1) \ldots$ $\ldots f_k(x_k)$, где $f_j(s) = \varphi_u(\rho(s, [a_j, b_j]))$, ρ обозначает расстояние на прямой и φ_u определяется формулой (1.4).

Зафиксируем u. При данном ε выберем r столь большим, чтобы каждая функция f_j обращалась в нуль вне интервала $[-r, r]$, и в то же самое время столь большим, чтобы для куба $I_r = \{x\colon |x_m| \leqslant r,\ m = 1, \ldots, k\}$ выполнялись неравенства $P(I_r^c) < \varepsilon$ и $Q(I_r^c) < \varepsilon$. Так как $f_j(-r) = f_j(r)$, то по теореме Вейерштрасса*) функция $f_j(s)$ может быть равномерно приближена в $[-r, r]$ конечной тригонометрической суммой $\sum\limits_l \alpha_l e^{il\pi s/r}$ периода $2r$. Перемножив эти суммы для различных значений j, мы видим, что f может быть равномерно приближена в I_r конечной тригонометрической суммой

$$g(x) = \sum_l \gamma_l e^{it^{(l)}\cdot x} \tag{7.14}$$

периода $2r$ по каждой переменной. Выберем сумму (7.14) так, чтобы $|f(x) - g(x)| < \varepsilon$ для x из I_r. Так как f всюду ограничена единицей, то g ограничена величиной $1 + \varepsilon$ в I_r и, следовательно, по периодичности всюду в R^k. Таким образом, $|f - g|$ ограничена величиной ε в I_r и величиной $2 + \varepsilon$ за пределами I_r. Предполагая, что $0 < \varepsilon < 1$, мы имеем

$$\int |f - g|\,dP \leqslant \varepsilon + (2 + \varepsilon) P(I_r^c) < 4\varepsilon.$$

Аналогично, $\int |f - g|\,dQ < 4\varepsilon$, так что

$$\left|\int f\,dP - \int f\,dQ\right| < \left|\int g\,dP - \int g\,dQ\right| + 8\varepsilon.$$

---

*) См., например, Титчмарш (1951).

Равенство $\int g\,dP = \int g\,dQ$ является теперь непосредственным следствием условия $p = q$ и (7.14); так как ε произвольно, то (7.13) действительно выполняется.

Предположим теперь, что P_n и P — вероятностные меры на $(R^k, \mathcal{R}^k)$ с характеристическими функциями p_n и p. Докажем теорему непрерывности:

Теорема 7.6. *Сходимость $P_n \Rightarrow P$ имеет место тогда и только тогда, когда $p_n(t) \to p(t)$ при каждом t.*

Необходимость следует из того факта, что $e^{it\cdot x}$ ограничена и непрерывна по x при каждом t. Мы докажем достаточность с помощью двух промежуточных утверждений.

Если $p_n(t)$ сходятся в каждой точке к некоторому пределу и если последовательность $\{P_n\}$ плотна, то $P_n \Rightarrow P$ для некоторой вероятностной меры P.

Пусть $g(t) = \lim_n p_n(t)$; мы не делаем никаких предположений относительно g. Если $\{P_n\}$ плотна, то по теореме Прохорова она относительно компактна, так что любая подпоследовательность $\{P_{n'}\}$ содержит некоторую другую подпоследовательность $\{P_{n''}\}$, сходящуюся слабо к некоторому пределу, характеристическая функция которого равна $\lim_{n''} p_{n''}(t) = g(t)$. По теореме единственности существует только один такой предел, который, как мы видели (см. теорему 2.3), равен P.

Обратим внимание на то, что это утверждение не выполняется без предположения о плотности [возьмем, например, $k = 1$ и распределение P_n равномерным на $(-n, n)$]. Заметим также, что вышеприведенное доказательство во многом аналогично рассуждениям в начале § 6, в которых использовались плотность и сходимость конечномерных распределений для вероятностных мер на $(C, \mathscr{C})$.

Покажем теперь, что *если предельная функция $g(t) = \lim_n p_n(t)$ непрерывна при $t = 0$, то $\{P_n\}$ плотна* (и, следовательно, слабо сходится к некоторому пределу). Очевидно, что $\{P_n\}$ плотна тогда, когда плотна каждая из k соответствующих последовательностей одномерных маргинальных распределений. Характеристические функции $p_n(s, 0, \ldots, 0)$ маргинальных распределений для первой координаты сходятся поточечно к пределу $g(s, 0, \ldots, 0)$, непрерывному при $s = 0$, и аналогично для других координат. Таким образом, каждая последовательность маргинальных распределений удовлетворяет предположению, и достаточно разобрать случай $k = 1$. В этом случае по теореме Фубини

$$\frac{1}{u}\int_{-u}^{u}(1 - p_n(t))\,dt = \int\left[\frac{1}{u}\int_{-u}^{u}(1 - e^{itx})\,dt\right]P_n(dx) =$$

$$= 2\int\left(1 - \frac{\sin ux}{ux}\right)P_n(dx) \geqslant 2\int_{|x|\geqslant 2/u}\left(1 - \frac{1}{|ux|}\right)P_n(dx) \geqslant$$

$$\geqslant P_n\left\{x\colon |x| \geqslant \frac{2}{u}\right\} \quad (7.15)$$

(в частности, первый интеграл равен действительному числу). Поскольку функция g непрерывна в начале координат, то при любом положительном ε существует такое u, что $u^{-1}\int_{-u}^{u}|1 - g(t)|\,dt < \varepsilon$. По теореме об ограниченной сходимости при n, превосходящих некоторое n_0, мы имеем

$$u^{-1}\int_{-u}^{u}|1 - p_n(t)|\,dt < 2\varepsilon$$

и, следовательно, в силу (7.15) $P_n\{x\colon |x| \geqslant a\} < 2\varepsilon$ с $a = 2/u$. Увеличивая при необходимости a, мы можем гарантировать, что это неравенство выполняется для конечного множества n, не превышающих упомянутое n_0: последовательность $\{P_n\}$ действительно плотна.

Если g — характеристическая функция вероятностной меры, она автоматически непрерывна. Теорема непрерывности доказана.

Поучительно сравнить следующие пары утверждений, среди которых все справедливы, за исключением 6°. Здесь P_n и P — вероятностные меры на $(R^k, \mathscr{R}^k)$ или на $(C, \mathscr{C})$, соответственно; в первом случае p_n и p являются характеристическими функциями мер P_n и P, а g — функция, заданная на R^k; во втором случае, $P_n\pi^{-1}_{t_1 \ldots t_l}$ и $P\pi^{-1}_{t_1 \ldots t_l}$ — конечномерные распределения мер P_n и P, а $\mu_{t_1 \ldots t_l}$ — вероятностные меры на $(R^l, \mathscr{R}^l)$.

Пространство R^k	*Пространство C*
1. Функция $p(t)$ определяет P.	1°. Меры $P\pi^{-1}_{t_1\ldots t_l}$ определяют P
2. Если $p_n(t)$ сходятся к некоторой функции $g(t)$ при каждом t и если последовательность $\{P_n\}$ плотна, то P_n слабо сходятся к некоторой мере P.	2°. Если $P_n\pi^{-1}_{t_1\ldots t_l}$ слабо сходятся к некоторой мере $\mu_{t_1\ldots t_l}$ при каждом $(t_1, \ldots, t_l)$ и если последовательность $\{P_n\}$ плотна, то P_n слабо сходятся к некоторому пределу P.
3. Утверждение 2 неверно без предположения о плотности.	3°. Утверждение 2° неверно без предположения о плотности.
4. Если $p_n(t)$ сходятся к некоторой функции $g(t)$ при каждом t и если функция g непрерывна в 0, то последовательность $\{P_n\}$ — плотная (и, следовательно, в силу 2 — слабо сходящаяся).	
5. Если $p_n(t) \to p(t)$ при каждом t и если $\{P_n\}$ плотна, то $P_n \Rightarrow P$.	5°. Если $P_n\pi^{-1}_{t_1\ldots t_l} \Rightarrow P\pi^{-1}_{t_1\ldots t_l}$ при каждом $(t_1, \ldots, t_l)$ и если $\{P_n\}$ плотна, то $P_n \Rightarrow P$.
6. Если $p_n(t) \to p(t)$ при каждом t, то $P_n \Rightarrow P$.	6°. [Если $P_n\pi^{-1}_{t_1\ldots t_l} \Rightarrow P\pi^{-1}_{t_1\ldots t_l}$ при каждом $(t_1, \ldots, t_l)$, то $P_n \Rightarrow P$.]

Утверждение 1 есть не что иное, как теорема единственности для характеристических функций; 1° утверждает тот факт, что конечномерные множества в C образуют определяющий класс. Доказательства 1 и 1° не имеют между собой ничего общего.

Утверждения 2 и 2° выводятся, соответственно, из 1 и 1° посредством параллельных рассуждений.

Контрпримеры подтверждают утверждения 3 и 3°.

Утверждение 4 не имеет аналога 4°.

Далее, 5 и 5° выводятся, соответственно, из 2 и 2° параллельными рассуждениями; однако предположение о плотности в 5 излишне в силу 4 и того факта, что p должна быть непрерывна. Исключение предположения о плотности приводит от 5 к 6 и от 5° к 6°, при этом 6 верно (в силу 4 и непрерывности p), а 6° неверно (так как нет 4°).

Утверждение 6° не только ложно само по себе (как вывод, примененный ко всем P и ко всем $\{P_n\}$), но, более того, *не существует ни одной меры P, для которой* *оно справедливо* (как вывод, примененный ко всем $\{P_n\}$): определим h_n: $C \to C$ как $h_n(x) = x + x_n$, где x_n заданы равенством (3.5), и при данном P положим $P_n = Ph_n^{-1}$. Для некоторого достаточно малого δ открытое множество $A = \{x\colon x(t) < x(0) + {}^1/_2,\ t \leqslant \delta\}$ удовлетворяет неравенству $P(A) > 0$. Если $A_n = A - x_n$, то множество $\limsup_n A_n$ пусто и, следовательно,

$$\liminf_n P_n(A) \leqslant \limsup_n P_n(A) = \limsup_n P(A_n) = 0.$$

По теореме 2.1 P_n не может слабо сходиться к P, однако соответствующие конечномерные распределения все же сходятся.

Только что приведенное рассуждение показывает также, что 4° не может быть справедливым. В этой связи заметим, что 3° следует из того факта, что 6° ложно; поскольку 6 справедливо, то 3 требует специального контрпримера (предельная функция g должна иметь разрыв в 0).

Если бы 4° было справедливо, то бо́льшая часть главы 2 была бы излишней (а если бы 4° было справедливо для пространства D, то излишней была бы бо́льшая часть главы 3). Добавим теперь замечание по поводу 6°: поскольку конечномерные распределения P и всех P_n полностью определяют P и все P_n, условия сходимости $P_n \Rightarrow P$ в принципе могут быть выражены в терминах одних только конечномерных распределений *). Однако в действительности мы должны иметь дело со всеми множествами $(t_1, \ldots, t_i)$ одновременно при $n \to \infty$ — нельзя фиксировать какой-то набор $(t_1, \ldots, t_i)$ и рассматривать, что происходит при $n \to \infty$. Именно здесь плотность играет существенную роль.

Прием Крамера — Уолда. С помощью нижеследующего простого приема, предложенного Крамером и Уолдом, задачи, включающие случайные векторы в R^k, часто могут быть сведены к задачам, включающим лишь обычные случайные величины в R^1. Предположим, что k-мерные случайные векторы $X_n = (X_{n1}, \ldots, X_{nk})$ и $X = (X_1, \ldots, X_k)$ удовлетворяют условию

$$\sum_{j=1}^{k} t_j X_{nj} \xrightarrow{\mathcal{D}} \sum_{j=1}^{k} t_j X_j$$

*) См. Бартошинский (1961).

для каждой точки $t = (t_1, \ldots, t_k)$ из R^k. Тогда характеристические функции $p_n(s) = \mathsf{E}\left\{\exp\left(is\sum_{j=1}^{k} t_j X_{nj}\right)\right\}$ этих одномерных случайных величин сходятся к $p(s) = \mathsf{E}\left\{\exp\left(is\sum_{j=1}^{k} t_j X_j\right)\right\}$ для каждого действительного s. Полагая $s = 1$, мы видим, что

$$\mathsf{E}\{e^{it\cdot X_n}\} \to \mathsf{E}\{e^{it\cdot X}\}.$$

Поскольку t было произвольным, сходимость $X_n \xrightarrow{\mathscr{D}} X$ следует теперь из теоремы непрерывности для характеристических функций.

Теорема 7.7. *В пространстве R^k X_n сходятся по распределению к X тогда и только тогда, когда каждая линейная комбинация компонент X_n сходится по распределению к соответствующей линейной комбинации компонент X.*

В терминах § 2 полупространства в R^k образуют класс, определяющий сходимость; этот факт, по-видимому, не может быть доказан без привлечения идей из анализа Фурье.

Локальная и интегральная предельные теоремы. Если P_n и P суть вероятностные меры на R^k с плотностями p_n и p относительно меры Лебега и если

$$p_n(x) \to p(x) \tag{7.16}$$

для всех x вне некоторого множества лебеговой меры 0, то по теореме Шеффе (стр. 306)

$$P_n \Rightarrow P. \tag{7.17}$$

Соотношение (7.16), называемое локальной предельной теоремой, влечет соотношение (7.17), называемое интегральной предельной теоремой.

Существует аналогичный результат для случая, когда P имеет плотность, но масса, соответствующая P_n, сосредоточена на некоторой решетке в R^k. Пусть $\delta(n) = (\delta_1(n), \ldots, \delta_k(n))$ — точка в R^k с положительными координатами, и пусть $\alpha(n) = (\alpha_1(n), \ldots, \alpha_k(n))$ — произвольная точка в R^k; обозначим через L_n решетку, образуемую всеми точками R^k вида

$$(u_1\delta_1(n) - \alpha_1(n), \ldots, u_k\delta_k(n) - \alpha_k(n)),$$

где $u_1, \ldots, u_k$ независимо принимают значения $0, \pm 1, \pm 2, \ldots$ Если x есть точка из L_n, то интервал $(x - \delta(n), x]$ (см. (3.1) по поводу обозначений) представляет собой ячейку объема $v_n = \delta_1(n) \cdots \delta_k(n)$ и R^k представляет собой счетное объединение таких непересекающихся ячеек.

Предположим теперь, что P_n есть вероятностная мера в R^k с $P_n(L_n) = 1$, и пусть для $x \in L_n$ $p_n(x)$ обозначает массу (возможно, равную 0) в этой точке. Пусть P имеет плотность p относительно меры Лебега.

Теорема 7.8. *Предположим, что*

$$\max\{\delta_1(n), \ldots, \delta_k(n)\} \to 0. \tag{7.18}$$

Предположим, далее, что если $\{x_n\}$ — любая последовательность и x — любая точка из R^k и если x_n принадлежит L_n и меняется вместе с n таким образом, что

$$x_n \to x, \tag{7.19}$$

то

$$\frac{p_n(x_n)}{v_n} \to p(x). \tag{7.20}$$

Тогда $P_n \Rightarrow P$.

Доказательство. Определим плотность вероятности q_n на R^k, положив $q_n(y) = p_n(x)/v_n$, если y лежит в ячейке $(x - \delta(n), x]$, определяемой точкой x решетки L_n. Так как из (7.19) следует (7.20), то мы имеем

$$q_n(y) \to p(y) \tag{7.21}$$

для каждого y. Пусть X_n имеет плотность q_n; определим Y_n на том же вероятностном пространстве, полагая $Y_n = x$, если $X_n \in (x - \delta(n), x]$ при $x \in L_n$. Нам надо доказать, что $Y_n \xrightarrow{\mathcal{D}} P$. Поскольку $|X_n - Y_n| \leqslant |\delta(n)|$, это будет следовать из (7.18) и теоремы 4.1, если мы докажем, что $X_n \xrightarrow{\mathcal{D}} P$. Однако ввиду (7.21) последнее соотношение есть следствие теоремы Шеффе.

Слабая сходимость на окружности и торе. Пусть S — единичная окружность на комплексной плоскости, и пусть $e_u(x) = x^u$, $u = 0, \pm 1, \pm 2, \ldots$ Коэффициенты Фурье меры P на S определяются для целых u равенством $p(u) = \int e_u(x) P(dx)$. Поскольку мера P определена значениями интеграла $\int f\,dP$ для f из $C(S)$ и поскольку каждая функция f из $C(S)$ может быть по теореме Вейерштрасса равномерно аппроксимирована линейными комбинациями функций e_u, то p определяет P. Рассуждая так же, как в случае прямой, можно получить из этой теоремы единственности теорему непрерывности: $P_n \Rightarrow P$ тогда и только тогда, когда соответствующие коэффициенты Фурье удовлетворяют условию: $p_n(u) \to p(u)$ для любого целого u. Поскольку окружность компактна, то на этот раз не возникает вопроса о плотности последовательности $\{P_n\}$; следовательно, $\{P_n\}$ слабо сходится к некоторому пределу, если для каждого u существует $\lim_n p_n(u)$. Если P представляет собой нормированную меру Лебега на окружности, то $p(u)$ обращается в 0 при $u \neq 0$ (и, разумеется, равняется 1 при $u = 0$). Отсюда немедленно следует критерий Вейля: последовательность $(x_1, x_2, \ldots)$ равномерно распределена на окружности тогда и только тогда, когда $n^{-1}\sum_{j=1}^{n}(x_j)^u \to 0$ для каждого u, отличного от 0. В частности, степени x^j равномерно распределены, если x не является корнем из единицы.

Окружность может быть развернута в интервал $[0, 1)$ с помощью соответствия $e^{2\pi ix} \longleftrightarrow x$. Коэффициенты Фурье вероятностной меры P на $[0,1)$ определяются равенством $p(u) = \int e^{2\pi iux}P(dx)$ для целых u; разумеется, теоремы единственности и непрерывности для окружности могут быть переформулированы для интервала $[0,1)$. Несмотря на то, что здесь слабая сходимость должна рассматриваться в той топологии, которую $[0, 1)$ наследует от окружности (например, в этой топологии $1 - 1/n \to 0$), мы можем полностью игнорировать это различие, если предельная мера приписывает точке 0 меру 0. Таким образом, если P_n, P суть вероятностные меры на R^1 с носителем в $[0,1)$ и если $P\{0\} = 0$, то $P_n \Rightarrow P$ тогда и только тогда, когда

$$\int e^{2\pi iux}P_n(dx) \to \int e^{2\pi iux}P(dx)$$

для каждого целого u. Говорят, что последовательность $(x_1, x_2, \ldots)$ действительных чисел равномерно распределена по модулю 1, если эмпирическое распределение последовательности дробных частей $(\{x_1\}, \{x_2\}, \ldots)$ слабо сходится к равномерному распределению на единичном интервале. Критерий Вейля приобретает следующий вид: $(x_1, x_2, \ldots)$ равномерно распределена по модулю 1 тогда и только тогда, когда $n^{-1}\sum_{j=1}^{n} e^{2\pi i u x_j} \to 0$ для каждого целого u, отличного от нуля. Этот критерий удовлетворяется, если $x_j = j\xi$, где ξ — иррациональное число.

С помощью k-мерного варианта теоремы Вейерштрасса можно получить аналогичные результаты для k-мерного тора — т. е. произведения k окружностей. (Эти результаты могут быть перенесены даже на общую компактную группу; роль тригонометрических функций будут играть при этом характеры.) Рассматривая случай $k = 2$ и развертывая тор в квадрат, мы видим, что если P_n и P суть вероятностные меры в R^2, сосредоточенные на квадрате $\{(x, y)\colon 0 \leqslant x,\ y < 1\}$, и если нижнее и левое ребра квадрата имеют P-меру 0, то $P_n \Rightarrow P$ тогда и только тогда, когда

$$\int e^{2\pi i\,(ux+vy)}\,dP_n \to \int e^{2\pi i\,(ux+vy)}\,dP$$

для любой пары целых чисел u и v.

Мы говорим, что последовательность $((x_1, y_1), (x_2, y_2), \ldots)$ на плоскости равномерно распределена по модулю 1, если эмпирические распределения, полученные приведением всех координат по модулю 1, слабо сходятся к равномерному распределению на квадрате; это справедливо в том и только том случае, когда $n^{-1}\sum_{j=1}^{n} e^{2\pi i\,(ux_j+vy_j)} \to 0$, если только u и v не обращаются в 0 одновременно. Это условие удовлетворяется, когда $(x_j, y_j) = (j\xi, j\eta)$, где ξ, η и 1 линейно независимы (не существуют целые числа u и v, не равные одновременно 0, такие, что $u\xi + v\eta$ является целым числом). Возьмем другой случай и предположим, что ξ и η удовлетворяют ровно одному ограничению: $u\xi + v\eta$ является целым числом тогда и только тогда, когда $u = 2w$ и $v = 3w$

для некоторого целого числа w. Тогда

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n} e^{2\pi i\,(u/\xi+v/\eta)}=\begin{cases}1, & \text{если } u/2=v/3 \text{ целое число},\\ 0 & \text{во всех других случаях.}\end{cases}$$

Рассматривая коэффициенты Фурье, можно понять, что эмпирические распределения слабо сходятся к распределению, равномерному на четырех наклонных линиях, показанных на рис. 4.

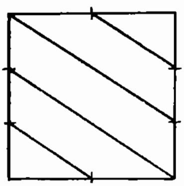

Рис. 4.

Результат Крамера — Уолда имеет здесь очевидную аналогию. Например, последовательность точек плоскости $((x_1, y_1), (x_2, y_2), \ldots)$ равномерно распределена по модулю 1 тогда и только тогда, когда для всех целых u и v, не равных одновременно 0, линейная последовательность $(ux_1+vy_1,\ ux_2+vy_2, \ldots)$ равномерно распределена по модулю 1.

Примечание. Теорема 7.2 доказывается здесь с помощью варианта метода Линдеберга, предложенного Леви; см. Леви (1925, стр. 246—249). С другими подходами к центральной предельной теореме и с характеристическими функциями можно познакомиться по Феллеру (1967b). Для линейных пространств существует теория характеристических функций, тесно связанная со слабой сходимостью; см. Прохоров (1960) и Гросс (1963), а также имеющиеся там ссылки.

См. Крамер, Уолд (1936) по поводу их приема. См. Харди, Райт (1960) о равномерном распределении по модулю 1, в частности, в связи с последним из приведенных нами примеров.

Задачи

1. Замените подинтегральное выражение в (7.1) полиномом степени $2k$, выбранным так, чтобы φ имела ограниченные непрерывные производные вплоть до порядка k ($k=3$ достаточно для доказательства теоремы Линдеберга).

2. Используйте прием Крамера — Уолда для получения k-мерного варианта теоремы Линдеберга — Леви.

3. Говорят, что измеримая функция $h(x)$ на прямой имеет среднее значение, если существует предел

$$M\{h\}=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} h(x)\,dx;$$

$h(x)$ имеет распределеиие, если вероятиостные меры

$$P_T(A) = \frac{1}{2T}\lambda\{x: |x| \leqslant T,\ h(x) \in A\} \qquad A \in \mathscr{R}^1,$$

(здесь λ — мера Лебега) сходятся слабо при $T \to \infty$. Если $h(x)$ — почти периодическая функция, то она ограничена, и при любом t функция e^{itx} является почти периодической (см. Бор (1932)); с помощью первого предложения, используемого при выводе теоремы иепрерывности для характеристических функций, покажите, что h имеет распределение.

4. Рассмотрим на прямой вероятностные меры, обладающие моментами всех порядков. Говорят, что мера P определяется своими момеитами, если из $\int x^u P(dx) = \int x^u Q(dx)$, $u = 1, 2, \ldots$, следует $P = Q$. (См. Феллер (1967b, стр. 282, 283, 587) по поводу примера меры P, не определяемой своими моментами, и по поводу критерия определяемости P своими моментами.) Если $\sup_n \left|\int x^u P_n(dx)\right| < \infty$ при каждом u и если $P_n \Rightarrow P$, то моменты P_n сходятся к момеитам P [теорема 5.4]. Если моменты P_n сходятся к моментам P и если P определяется своими момеитами, то $P_n \Rightarrow P$. [Покажите сначала, что $\{P_n\}$ плотна (задача 5 из § 6) и затем проведите рассуждение, близкое даниому выше доказательству теоремы непрерывности для характеристических функций.] Используйте методику Крамера — Уолда для распространеиия этого результата иа R^k [рассмотрите смешанные моменты всех порядков].

5. Если k меияется вместе с n так, что $(k - np)/\sqrt{npq} \to x$, то

$$\binom{n}{k} p^k q^{n-k} npq \to \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

(см. Феллер (1967a), стр. 186). Используйте теорему 7.8 для вывода теоремы Муавра — Лапласа. Примените ту же самую методику к гипергеометрическому распределению (см. задачу 2 на стр. 71 и задачу 10 на стр. 197 у Феллера (1967a)) и к распределению частот для последовательности полииомиальных испытаиий (рассмотрите частоты появления всех исходов, кроме одного, с тем, чтобы предельное распределение имело плотность).

6. Используйте задачу 8 § 2 для доказательства того, что если вероятностные меры в R^k слабо сходятся, то соответствующие характеристические функции сходятся равиомерно на ограиичениых множествах.

7. Обобщите теорему 7.8: замените меру Лебега на R^k общей мерой на сепарабельном пространстве S и рассмотрите разбиения $\{C_{ni}\}$ пространства S, удовлетворяющие условию $\lim_n \max_i \operatorname{diam} C_{ni} = 0$.

Глава 2

ПРОСТРАНСТВО *C*

§ 8. Слабая сходимость и плотность в *C*

Во введении были разъяснены некоторые полезные следствия, вытекающие из факта слабой сходимости в пространстве $C = C[0,1]$ непрерывных функций на $[0,1]$. Пространство C наделяется при этом топологией равномерной сходимости, задаваемой при помощи расстояния между двумя точками x и y (двумя функциями $x(t)$ и $y(t)$ от $t \in [0,1]$):

$$\rho(x, y) = \sup_t |x(t) - y(t)|.$$

Слабая сходимость. Слабая сходимость распределений в C не вытекает, вообще говоря, из слабой сходимости конечномерных распределений. Однако мы видели в начале § 6, что вторая сходимость влечет первую в предположении относительной компактности. Так как пространство C сепарабельно и полно (см. стр. 301), из теоремы Прохорова (теоремы 6.1 и 6.2) следует, что относительная компактность семейства вероятностных мер на $(C, \mathscr{C})$ эквивалентна плотности этого семейства. Таким образом, мы получаем следующий результат:

Теорема 8.1. *Пусть P_n и P — вероятностные меры на $(C, \mathscr{C})$. Если конечномерные распределения мер P_n слабо сходятся к конечномерным распределениям мер P и, кроме того, семейство $\{P_n\}$ плотно, то $P_n \Rightarrow P$.*

Чтобы использовать эту теорему для доказательства слабой сходимости в C, нам необходимо лишь установить, что означает в этом случае плотность последовательности $\{P_n\}$.

Плотность. Модуль непрерывности элемента $x \in C$ определяется соотношением

$$w_x(\delta) = w(x, \delta) = \sup_{|s-t|<\delta} |x(s) - x(t)|, \quad 0 < \delta \leqslant 1. \qquad (8.1)$$

Пусть $\{P_n\}$ — последовательность вероятностных мер на $(C, \mathscr{C})$.

Теорема 8.2. *Последовательность $\{P_n\}$ плотна тогда и только тогда, когда выполнены следующие два условия*:

(i) *Для каждого положительного η существует такое a, что*

$$P_n\{x\colon |x(0)| > a\} \leqslant \eta, \quad n \geqslant 1. \qquad (8.2)$$

(ii) *Для каждой пары положительных чисел ε и η существуют δ, $0 < \delta < 1$, и целое число n_0 такие, что*

$$P_n\{x\colon w_x(\delta) \geqslant \varepsilon\} \leqslant \eta, \quad n \geqslant n_0. \qquad (8.3)$$

Условие (i) обеспечивает плотность последовательности $\{P_n\pi_0^{-1}\}$. В связи с (8.3) заметим, что функция $w_x(\delta)$ непрерывна по x и поэтому измерима.

Доказательство. Допустим, что последовательность $\{P_n\}$ плотна. Для данных ε и η выберем компактное множество K такое, что $P_n(K) > 1 - \eta$ при всех n. По теореме Арцела — Асколи (стр. 302) имеем $K \subset \{x\colon |x(0)| \leqslant a\}$ для достаточно больших a и $K \subset \{x\colon w_x(\delta) < \varepsilon\}$ для достаточно малых δ, так что (i) и (ii) выполняются (с $n_0 = 1$ в (ii)). Этим доказана необходимость условий (i) и (ii).

Так как любая отдельная вероятностная мера P на $(C, \mathscr{C})$ плотна (теорема 1.4), то из необходимости условия (ii) следует, что для каждой пары ε и η существует такое δ, что $P\{x\colon w_x(\delta) \geqslant \varepsilon\} \leqslant \eta$. Поэтому, если $\{P_n\}$ удовлетворяет условию (ii), то мы можем гарантировать выполнение неравенства (8.3) для конечного множества значений n, не превосходящих n_0 (при необходимости уменьшая δ). Таким образом, мы можем предположить при доказательстве достаточности, что n_0 в (8.3) всегда равно 1.

Допустим, что $\{P_n\}$ удовлетворяет условиям (i) и (ii) с $n_0 = 1$ в (8.3). При данном η выберем a так, что для множества

$$A = \{x\colon |x(0)| \leqslant a\}$$

при всех n $P_n(A) \geqslant 1 - \frac{1}{2}\eta$, и выберем δ_k так, что если

$$A_k = \{x\colon\ w_x(\delta_k) < 1/k\},$$

то $P_n(A_k) \geqslant 1 - \eta/2^{k+1}$ при всех n. Если K — замыкание множества $A \cap \bigcap_{k=1}^{\infty} A_k$, то $P_n(K) \geqslant 1 - \eta$ при всех n и по теореме Арцела — Асколи K компактно. Отсюда вытекает, что последовательность $\{P_n\}$ плотна.

Как показывает доказательство, мы можем требовать, чтобы (8.3) выполнялось для *всех* P_n. При таком изменении теорема остается верной, даже если $\{P_n\}$ заменить произвольным семейством Π. В применениях мы будем часто доказывать (8.3), заменяя величины δ, ε, η на, скажем, $\frac{1}{2}\delta$, 3ε, 9η, что, конечно, не меняет дела.

Теорема 8.2 преобразует условие плотности в C, заменяя компактность ее характеризацией с помощью теоремы Арцела — Асколи. Следующая теорема делает лишь малый шаг вперед по сравнению с этим. Тем не менее она удовлетворяет наши настоящие потребности.

Теорема 8.3. *Последовательность* $\{P_n\}$ *плотна, если выполняются следующие два условия*:

(i) *Для каждого положительного* η *существует* a *такое, что*

$$P_n\{x\colon\ |x(0)| > a\} \leqslant \eta, \quad n \geqslant 1. \tag{8.4}$$

(ii) *Для любых положительных* ε *и* η *существуют* δ, $0 < \delta < 1$, *и целое* n_0 *такие, что*

$$\frac{1}{\delta} P_n\{x\colon \sup_{t \leqslant s \leqslant t+\delta} |x(s) - x(t)| \geqslant \varepsilon\} \leqslant \eta, \quad n \geqslant n_0, \tag{8.5}$$

для всех t.

Разумеется, аргумент t в соотношении (8.5) меняется в пределах $0 \leqslant t \leqslant 1$; если $t > 1 - \delta$, то мы ограничиваем величину s в (8.5) пределами $t \leqslant s \leqslant 1$. Отметим, что (8.5) формально выполнено, если $\delta > 1/\eta$; но мы требуем, чтобы $\delta < 1$.

Доказательство. Фиксируем δ и положим

$$A_t = \{x\colon \sup_{t \leqslant s \leqslant t+\delta} |x(s) - x(t)| \geqslant \varepsilon\}.$$

Далее, и s и t лежат в интервалах вида $[i\delta, (i+1)\delta]$. Если $|s-t|<\delta$, то эти интервалы либо совпадают, либо примыкают друг к другу; отсюда вытекает, что

$$P_n\{x\colon\ w_x(\delta)\geqslant 3\varepsilon\}\leqslant P_n\Big(\bigcup_{i<\delta^{-1}} A_{i\delta}\Big). \tag{8.6}$$

Так как

$$P_n\Big(\bigcup_{i<\delta^{-1}} A_{i\delta}\Big)\leqslant \sum_{i<\delta^{-1}} P_n(A_{i\delta}), \tag{8.7}$$

то из неравенства (8.5) следует, что $P_n\{x\colon\ w_x(\delta)\geqslant 3\varepsilon\}\leqslant$ $\leqslant(1+[1/\delta])\delta\eta<2\eta$ (напомним, что $\delta<1$) *). Поэтому условие (ii) влечет соответствующее условие в теореме 8.2. Так как условие (i) осталось тем же самым, то мы получаем желаемый результат.

Сформулируем теперь следствие, которое содержит существенный пункт доказательства.

Следствие. *Если* $0=t_0<t_1<\ldots<t_r=1$ *и*

$$t_i-t_{i-1}\geqslant\delta,\quad 2\leqslant i\leqslant r-1, \tag{8.8}$$

то

$$P\{x\colon\ w_x(\delta)\geqslant 3\varepsilon\}\leqslant$$
$$\leqslant\sum_{i=1}^{r} P\{x\colon \sup_{t_{i-1}\leqslant s\leqslant t_i}|x(s)-x(t_{i-1})|\geqslant\varepsilon\}. \tag{8.9}$$

Отметим, что выполнение неравенства (8.8) при $i=1$ и $i=r$ не предполагается.

Замена в неравенстве (8.6) левой части правой не приводит к существенному огрублению, но в (8.7) это уже не так, и условие (ii) не является необходимым **). Допустим, однако, что множества $A_{i\delta}$ (δ фиксировано, i меняется) независимы относительно P_n и все имеют одну и ту же вероятность p (что будет по крайней мере приближенно верно для многих мер P_n, встречающихся в этой главе). Тогда

$$P\Big(\bigcup_{i<\delta^{-1}} A_{i\delta}\Big)=1-(1-p)^{(1+[1/\delta])}.$$

---

*) Мы используем символ $[\alpha]$ для обозначения целой части α — наибольшего целого числа, не превосходящего α.

**) См. задачу 1.

Если эта величина ограничена сверху числом η, то (при $\eta \leqslant 1/2$) получаем оценку

$$p/\delta \leqslant (1+[1/\delta])\,p \leqslant -\log(1-\eta) \leqslant 2\eta, \qquad (8.10)$$

которая имеет тот же самый порядок, что и оценка (8.5). Поэтому любой эффективный критерий плотности за пределами теоремы 8.3 должен каким-либо образом учитывать зависимость между $A_{i\delta}$.

Случайные функции. Пусть X отображает пространство $(\Omega, \mathscr{B}, \mathsf{P})$ в C. Пока мы не будем предполагать, что отображение X измеримо $(X^{-1}\mathscr{C} \subset \mathscr{B})$. При каждом ω из Ω $X(\omega)$ есть элемент C — непрерывная заданная на $[0,1]$ функция, значение которой в точке t мы обозначим $X(t, \omega)$. При фиксированном t пусть $X(t)$ обозначает действительную функцию на Ω со значением $X(t, \omega)$ в точке ω; $X(t)$ есть композиция $\pi_t X$. Аналогично, пусть $(X(t_1), \ldots, X(t_k))$ обозначает отображение Ω в R^k со значением $(X(t_1, \omega), \ldots, X(t_k, \omega))$ в ω.

Если $A = \{x \in C\colon x(t) \leqslant \alpha\}$, то $A \in \mathscr{C}$ и $\{\omega\colon X(t, \omega) \leqslant \alpha\} = X^{-1}A$. Отсюда следует, что если X — случайный элемент C, т. е. если $X^{-1}\mathscr{C} \subset \mathscr{B}$, то при каждом t $X(t)$ является случайной величиной $(X(t)^{-1}\mathscr{R}^1 \subset \mathscr{B})$ и, следовательно, каждый набор $(X(t_1), \ldots, X(t_k))$ является случайным вектором. Предположим, с другой стороны, что $X(t)$ является случайной величиной при каждом t. Если B — замкнутая сфера в C с центром x и радиусом δ, то $X^{-1}B = \{\omega\colon X(\omega) \in B\} = \bigcap_r \{\omega\colon x(r) - \delta \leqslant X(r, \omega) \leqslant x(r) + \delta\}$, где пересечение берется по всем рациональным числам, так что $X^{-1}B \in \mathscr{B}$. Так как C сепарабельно, то $X^{-1}\mathscr{C} \subset \mathscr{B}$, т. е. X есть случайный элемент C. Таким образом, X является случайной функцией (случайным элементом C) тогда и только тогда, когда при каждом t $X(t)$ является случайной величиной. (Мы просто еще раз доказали, что конечномерные множества порождают $\mathscr{C}$.)

Допустим теперь, что $\{X_n\}$ — последовательность случайных функций. По определению последовательность считается плотной, если последовательность соответствующих распределений плотна. Согласно теореме 8.2 последовательность $\{X_n\}$ плотна в том и только том случае, когда последовательность $\{X_n(0)\}$ плотна (на прямой) и для любых положительных ε и η существуют δ, $0 < \delta < 1$, и целое n_0 такие, что

$$P\{w(X_n, \delta) \geqslant \varepsilon\} \leqslant \eta, \quad n \geqslant n_0 \tag{8.11}$$

(здесь $w(x, \delta)$ — всего лишь другое обозначение для $w_x(\delta)$). Это условие обеспечивает не слишком сильное колебание случайных функций X_n.

Теорема 8.3 может быть преобразована тем же самым образом: последовательность $\{X_n\}$ будет плотна, если плотна последовательность $\{X_n(0)\}$ и для любых положительных ε и η существуют δ, $0 < \delta < 1$, и целое число n_0 такие, что

$$\frac{1}{\delta} \mathsf{P}\{\sup_{t \leqslant s \leqslant t+\delta} |X_n(s) - X_n(t)| \geqslant \varepsilon\} \leqslant \eta \tag{8.12}$$

при $n \geqslant n_0$ и $0 \leqslant t \leqslant 1$ (где s в случае $1 - \delta < t \leqslant 1$ ограничено пределами $t \leqslant s \leqslant 1$).

Как объяснено во введении, нас в первую очередь интересуют случайные функции, которые строятся следующим способом. Пусть $\xi_1, \xi_2, \ldots$ — случайные величины на некотором вероятностном пространстве $(\Omega, \mathscr{B}, \mathsf{P})$. Мы пока не делаем предположений о таких специальных свойствах, как стационарность и независимость. Определим $S_n = \xi_1 + \ldots + \xi_n$, $S_0 = 0$ и построим X_n, используя частичные суммы $S_0, S_1, \ldots, S_n$. Положим для точек вида i/n из $[0,1]$

$$X_n\left(\frac{i}{n}, \omega\right) = \frac{1}{\sigma\sqrt{n}} S_i(\omega) \tag{8.13}$$

(мы ограничиваемся нормирующими множителями вида $\sigma/\sqrt{n}$, но другие множители также возможны). Для прочих точек t из $[0,1]$ мы определим $X_n(t, \omega)$ с помощью линейной интерполяции: если $t \in [(i-1)/n, i/n]$, то

$$\begin{aligned} X_n(t, \omega) &= \frac{(i/n) - t}{1/n} X_n\left(\frac{i-1}{n}\right) + \frac{t - (i-1)/n}{1/n} X_n\left(\frac{i}{n}\right) = \\ &= \frac{1}{\sigma\sqrt{n}} S_{i-1}(\omega) + n\left(t - \frac{i-1}{n}\right)\frac{1}{\sigma\sqrt{n}} \xi_i(\omega). \end{aligned} \tag{8.14}$$

Так как $i - 1 = [nt]$ при $t \in [(i-1)/n, i/n]$, то можно определить функцию X_n более сжато формулой

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega). \tag{8.15}$$

Ввиду того, что ξ_i и, следовательно, S_i являются случайными величинами, из (8.15) следует, что $X_n(t)$ — тоже случайная величина при каждом t. Поэтому X_n являются случайными функциями.

Когда эти случайные функции образуют плотную последовательность? Так как $X_n(0) = 0$, то очевидно, что последовательность $\{X_n(0)\}$ плотна. Используя определение (8.15), мы можем превратить условие (8.12) в ограничение на флуктуации частичных сумм S_i. Если $t = k/n$ и $t + \delta = j/n$, где k и j целые, то (8.12) сведется к

$$\frac{1}{\delta} \mathsf{P}\left\{\max_{i \leqslant n\delta} \frac{1}{\sigma\sqrt{n}} |S_{k+i} - S_k| \geqslant \varepsilon\right\} \leqslant \eta. \tag{8.16}$$

Несмотря на то, что условия (8.12) и (8.16), вообще говоря, различаются, если t и δ не являются целыми кратными $1/n$, различие это, как мы покажем, несущественно для наших целей. Рассмотрим целые j и k, определенные неравенствами

$$\frac{k}{n} \leqslant t < \frac{k+1}{n}, \quad \frac{j-1}{n} \leqslant t + \frac{\delta}{2} < \frac{j}{n}.$$

Ввиду того, что X_n кусочно-линейна, имеем

$$\sup_{t \leqslant s \leqslant t + \frac{1}{2}\delta} |X_n(s) - X_n(t)| \leqslant 2 \max_{0 \leqslant i \leqslant j-k} \frac{1}{\sigma\sqrt{n}} |S_{k+i} - S_k|.$$

Если $n \geqslant 4/\delta$, то $j - k < n\delta$, так что максимум справа не уменьшится, если ограничение $i \leqslant j - k$ ослабить до $i \leqslant n\delta$. Поэтому, если (8.16) выполняется при всех k и всех $n \geqslant n_0$, то (8.12) выполняется при всех t и всех $n \geqslant \max\{n_0, 4/\delta\}$ в предположении, что ε, η и δ в (8.12) заменены, соответственно, на 2ε, 2η и $\frac{1}{2}\delta$, что сводится лишь к перемене обозначений.

Таким образом, последовательность $\{X_n\}$ плотна, если для любых положительных ε и η существуют δ, $0 < \delta < 1$, и целое n_0 такие, что (8.16) выполняется при всех k и при всех $n \geqslant n_0$. Максимум в (8.16), который распространяется на значения $i \leqslant n\delta$, легче вычисляется, если заменить $n\delta$ целым числом. Если $n\delta$ равно некоторому целому m (что верно, конечно, не всегда), тогда неравенство (8.16) превращается в неравенство

$$\mathsf{P}\left\{\max_{i\leqslant m}|S_{k+i}-S_k|\geqslant\frac{\varepsilon}{\sqrt{\delta}}\,\sigma\sqrt{m}\right\}\leqslant\eta\delta.$$

Положим $\lambda=\varepsilon/\sqrt{\delta}$ (если δ мало, то λ велико); неравенство (8.16) сводится тогда к

$$\mathsf{P}\{\max_{i\leqslant m}|S_{k+i}-S_k|\geqslant\lambda\sigma\sqrt{m}\}\leqslant\frac{\eta\varepsilon^2}{\lambda^2}.$$

Так как $\eta\varepsilon^2$ положительно, то мы получаем следующую теорему.

Теорема 8.4. *Пусть последовательность $\{X_n\}$ определена равенством* (8.15). *Последовательность $\{X_n\}$ плотна, если для любого положительного ε существуют λ, $\lambda>1$, и целое n_0 такие, что если $n\geqslant n_0$, то*

$$\mathsf{P}\{\max_{i\leqslant n}|S_{k+i}-S_k|\geqslant\lambda\sigma\sqrt{n}\}\leqslant\frac{\varepsilon}{\lambda^2}\tag{8.17}$$

при всех k.

Мы требуем, чтобы $\lambda>1$ ((8.17) тривиально, если $\lambda\leqslant\sqrt{\varepsilon}$), что соответствует требованию $\delta<1$ в теореме 8.3. В конкретных случаях доказательство (8.17) приводит к большим значениям λ.

Доказательство. При данных ε и η выберем δ, $0<\delta<1$, и n_0 так, чтобы при $n\geqslant n_0$ (8.16) выполнялось для всех k. Так как (8.16) становится лишь более строгим при уменьшении ε и η, то мы предположим, что $0<\varepsilon$, $\eta<1$. По предположению (с $\eta\varepsilon^2$ вместо ε) существуют λ, $\lambda>1$, и n_1 такие, что

$$\mathsf{P}\{\max_{i\leqslant n}|S_{k+i}-S_k|\geqslant\lambda\sigma\sqrt{n}\}\leqslant\frac{\eta\varepsilon^2}{\lambda^2}\tag{8.18}$$

при $n\geqslant n_1$ и $k\geqslant 1$. Положим $\delta=\varepsilon^2/\lambda^2$; так как $\lambda>1>\varepsilon$, мы имеем $0<\delta<1$. Пусть n_0 — целое число, превосходящее n_1/δ. Если $n\geqslant n_0$, то $[n\delta]\geqslant n_1$ и из (8.18) следует, что

$$\mathsf{P}\{\max_{i\leqslant[n\delta]}|S_{k+i}-S_k|\geqslant\lambda\sigma\sqrt{[n\delta]}\}\leqslant\frac{\eta\varepsilon^2}{\lambda^2}.$$

Так как $\lambda\sqrt{[n\delta]}\leqslant\varepsilon\sqrt{n}$ и $\eta\varepsilon^2/\lambda^2=\eta\delta$, то (8.16) выполняется для всех k, если только $n\geqslant n_0$. Таким образом, теорема 8.4 доказана.

Как нетрудно видеть, достаточно лишь, чтобы (8.17) выполнялось при $k\leqslant n\lambda^2/\varepsilon$; нам, однако, это не потребуется. Если последовательность $\{\xi_n\}$ стационарна, то (8.17) сводится к

$$\mathsf{P}\left\{\max_{i\leqslant n}|S_i|\geqslant\lambda\sigma\sqrt{n}\right\}\leqslant\frac{\varepsilon}{\lambda^2}. \tag{8.19}$$

Мы можем объединить σ с λ и требовать выполнения условия

$$\mathsf{P}\left\{\max_{i\leqslant n}|S_i|\geqslant\lambda\sqrt{n}\right\}\leqslant\frac{\varepsilon}{\lambda^2} \tag{8.20}$$

при $\lambda>\sigma$.

Мы увидим позднее, что в некоторых ситуациях предположение теоремы 8.4 становится не только достаточным, но и необходимым.

Координатные величины. Проекция π_t со значением $\pi_t(x)=x(t)$ при $x\in C$ есть случайная величина, заданная на $(C,\mathscr{C})$. Мы будем часто обозначать эту случайную величину через x_t: при фиксированном t x_t является функцией на C со значением $x(t)$ в точке x. Если задана вероятностная мера P на $(C,\mathscr{C})$, то $\{x_t\colon 0\leqslant t\leqslant 1\}$ есть стохастический процесс. Мы представляем себе t как временной параметр, а x_t обычно называются координатными величинами или функциями. Распределение x_t зависит от меры P, и мы будем говорить о распределении x_t *относительно* P; мы будем часто писать $P\{x_t\in H\}$ вместо $P\{x\colon x_t\in H\}$ и $\int x_t\,dP$ вместо $\int x_tP(dx)$. Наконец, и в том случае, когда t есть сложное выражение, мы будем часто называть $x(t)$ по-прежнему координатной величиной.

Примечание. Общая теория слабой сходимости в C (теорема 8.2, например) создана Прохоровым (1956). Теорема 8.4, являющаяся более удобным вариантом теоремы 8.3, возникла в результате бесед с Ф. Топсё.

Стоун (1963) изучал слабую сходимость в пространстве непрерывных функций на $[0,\infty)$. Ламперти (1962b) изучал пространства функций, удовлетворяющих условию Гёльдера.

Задача

1. Пусть $X(t) = t\xi$, где ξ — такая случайная величиина, что $P\{|\xi| \geqslant a\} \sim a^{-\frac{1}{2}}$ при $a \to \infty$, и пусть при каждом n P_n совпадает с распределением X. Тогда последовательность $\{P_n\}$ плотна, но не удовлетворяет условию (ii) теоремы 8.3.

§ 9. Существование винеровской меры

Винеровская мера. Винеровская мера W есть вероятностная мера на $(C, \mathscr{C})$, обладающая следующими двумя свойствами. Во-первых, при каждом t случайная величина x_t нормально распределена по отношению к W со средним 0 и дисперсией t:

$$W\{x_t \leqslant \alpha\} = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\alpha} e^{-u^2/2t}\, du. \tag{9.1}$$

Если $t = 0$, то подразумевается, что $W\{x_0 = 0\} = 1$. Во-вторых, стохастический процесс $\{x_t\colon 0 \leqslant t \leqslant 1\}$ имеет независимые приращения относительно W *): если

$$0 \leqslant t_0 \leqslant t_1 \leqslant \ldots \leqslant t_k \leqslant 1, \tag{9.2}$$

то случайные величины

$$x_{t_1} - x_{t_0},\quad x_{t_2} - x_{t_1},\ \ldots,\ x_{t_k} - x_{t_{k-1}} \tag{9.3}$$

независимы по W. В этом параграфе мы докажем существование такой меры W.

Если W обладает указанными двумя свойствами и если $s \leqslant t$, то x_t (нормальная случайная величина со средним 0 и дисперсией t) есть сумма независимых величин x_s (нормальных случайных величин со средним 0 и дисперсией s) и $x_t - x_s$, так что величина $x_t - x_s$ должна быть нормальной со средним 0 и дисперсией $t - s$ (ее характеристическая функция равна отношению характеристических функций x_t и x_s). Таким

---

*) Мы часто для краткости вместо «относительно W» будем писать «по W». (*Прим. перев.*)

образом, если (9.3) выполняется, то

$$W\{x_{t_i}-x_{t_{i-1}}\leqslant \alpha_i,\ i=1,\ \dots,\ k\}=$$
$$=\prod_{i=1}^{k}\frac{1}{\sqrt{2\pi(t_i-t_{i-1})}}\int_{-\infty}^{\alpha_i}e^{-u^2/2\,(t_i-t_{i-1})}\,du. \qquad (9.4)$$

В частности, приращения стационарны (распределение x_t-x_s по W зависит только от разности $t-s$), равно как и независимы.

Если представлять себе $x(t)$ как (одну) координату положения движущейся частицы в момент времени t, то сама функция x будет описывать историю изменения выбранной координаты при движении частицы от момента $t=0$ до момента $t=1$. Винеровская мера задает на этих траекториях x распределение, подходящее для описания броуновского движения — движения частиц цветочной пыльцы в воде.

При доказательстве существования W мы сталкиваемся с проблемой доказательства существования на $(C,\mathscr{C})$ вероятностной меры с заданными конечномерными распределениями. При произвольном их выборе существует самое большое одна такая мера, но может вообще не быть ни одной. (Например, не существует меры P, относительно которой распределение x_t приписывает единичную массу точке 0 при $t<1/2$ и точке 1 при $t\geqslant 1/2$.)

Теорема 9.1. *На $(C,\mathscr{C})$ существует вероятностная мера W такая, что выполняется соотношение* (9.1) *и случайные величины* (9.3) *независимы по W всякий раз, когда выполняется* (9.2).

Доказательство. Пусть $\xi_1, \xi_2, \dots$ независимы и нормально распределены (на некотором вероятностном пространстве $(\Omega,\mathscr{B},\mathsf{P})$) со средним 0 и дисперсией 1. Пусть случайная функция X_n определена равенством (8.15) при $\sigma=1$:

$$X_n(t,\ \omega)=\frac{1}{\sqrt{n}}S_{[nt]}(\omega)+(nt-[nt])\frac{1}{\sqrt{n}}\xi_{[nt]+1}(\omega). \qquad (9.5)$$

Пусть P_n — распределение функции X_n на C. Эта мера корректно определена вследствие того, что, как мы показали в предыдущем параграфе, X_n измерима $(X_n^{-1}\mathscr{C}\subset\mathscr{B})$.

Мы сначала покажем, что конечномерные распределения P_n слабо сходятся к конечномерным распределениям, которые должны быть у W. После этого мы докажем, что последовательность $\{P_n\}$ плотна. Ясно, что предел любой сходящейся слабо подпоследовательности $\{P_{n'}\}$ будет удовлетворять требованиям, предъявляемым к W. Идея заключается в том, что P_n аппроксимируют предполагаемую меру W.

Конечномерное распределение $P_n\pi^{-1}_{t_1 \ldots t_k}$ является в точности распределением случайного вектора $(X_n(t_1), \ldots$ $\ldots, X_n(t_k))$. Рассмотрим некоторый момент времени t. Из равенства (9.5) и предположения о нормальности ξ_n следует, что $X_n(t)$ нормально распределена со средним 0 и дисперсией

$$\frac{[nt]}{n} + \frac{(nt-[nt])^2}{n};$$

эта дисперсия отличается от t самое большее на $2/n$. Таким образом, $X_n(t) \xrightarrow{\mathcal{D}} N(0, t)$ (по поводу этого обозначения см. (4.9)).

Очевидно, что тот же самый способ рассуждений применим в случае нескольких временны́х точек; поэтому конечномерные распределения слабо сходятся к конечномерным распределениям, которые должна иметь мера W.

Для доказательства плотности $\{P_n\}$ применим теорему 8.4. Чтобы доказать (8.20), обозначим

$$E_i = \{\max_{j<i} |S_j| < 2\lambda\sqrt{n} \leqslant |S_i|\}. \tag{9.6}$$

Из свойств стационарности и независимости ξ_n вытекают неравенства

$$\begin{aligned} \mathsf{P}\{\max_{i\leqslant n}|S_i| \geqslant 2\lambda\sqrt{n}\} &\leqslant \\ \leqslant \mathsf{P}\{|S_n| \geqslant \lambda\sqrt{n}\} &+ \sum_{i=1}^{n-1} \mathsf{P}(E_i \cap \{|S_n - S_i| \geqslant \lambda\sqrt{n}\}) = \\ = \mathsf{P}\{|S_n| \geqslant \lambda\sqrt{n}\} &+ \sum_{i=1}^{n-1} \mathsf{P}(E_i)\mathsf{P}\{|S_{n-i}| \geqslant \lambda\sqrt{n}\} \leqslant \\ \leqslant P\{|S_n| \geqslant \lambda\sqrt{n}\} &+ \sum_{i=1}^{n-1} \mathsf{P}(E_i)\mathsf{P}\{|S_{n-i}| \geqslant \lambda\sqrt{n-i}\}. \end{aligned} \tag{9.7}$$

Так как величина $S_j/\sqrt{j}$ нормально распределена со средним 0 и дисперсией 1 и, следовательно, имеет конечный третий абсолютный момент a, не зависящий от j $(a=2\sqrt{2/\pi})$, то из (9.7) и неравенства Чебышева следует, что

$$\mathsf{P}\{\max_{i\leqslant n}|S_i|\geqslant 2\lambda\sqrt{n}\}\leqslant\frac{a}{\lambda^3}+\sum_{i=1}^{n-1}\mathsf{P}(E_i)\frac{a}{\lambda^3}.$$

Так как E_i не пересекаются, то мы получаем

$$\mathsf{P}\{\max_{i\leqslant n}|S_i|\geqslant 2\lambda\sqrt{n}\}\leqslant\frac{2a}{\lambda^3},\quad n=1,\ 2,\ \ldots \tag{9.8}$$

По заданному ε выберем λ так, что $2a/\lambda<\varepsilon$ и $\lambda>1$. Тогда (9.8) приводит к неравенству

$$\mathsf{P}\{\max_{i\leqslant n}|S_i|\geqslant 2\lambda\sqrt{n}\}\leqslant\frac{\varepsilon}{\lambda^2},$$

что равносильно (8.20), за исключением несущественного множителя 2 в левой части. Следовательно, по теореме 8.4 последовательность $\{P_n\}$ плотна, что завершает доказательство.

Таким образом, винеровская мера существует. В следующих параграфах мы выведем некоторые факты относительно W. Сейчас же мы только покажем, что элементы C имеют неограниченную вариацию, за исключением множества винеровской меры 0. Это интересно тем, что указывает смысл, в котором винеровские траектории (элементы C, выбранные согласно вероятностной мере W) отличаются высокой нерегулярностью — они непрерывны, но и только.

Элемент x пространства C имеет ограниченную вариацию тогда и только тогда, когда существует число M_x такое, что

$$\sum_{i=1}^{k}|x(t_i)-x(t_{i-1})|\leqslant M_x$$

для всех $t_0,\ t_1,\ \ldots,\ t_k$, где $0\leqslant t_0\leqslant t_1\leqslant\ldots\leqslant t_k\leqslant 1$. Поэтому достаточно показать, что если

$$f_n(x)=\sum_{i=1}^{2^n}\left|x\left(\frac{i}{2^n}\right)-x\left(\frac{i-1}{2^n}\right)\right|,$$

то $f_n(x) \to \infty$ всюду, за исключением множества x, имеющего W-меру 0. Далее, $|N|$ (см. (4.10)) имеет положительное среднее b и дисперсию c $\left(b = \sqrt{2/\pi}\right.$ и $c = 1 - 2/\pi\left.\right)$. Так как величины $|x(i2^{-n}) - x((i-1)2^{-n})|$ по отношению к мере W независимы и имеют среднее $2^{-n/2}b$ и дисперсию $2^{-n}c$, то среднее $f_n(x)$ равно $2^{n/2}b$, а дисперсия равна c. Из неравенства Чебышева следует теперь, что $W\{x\colon f_n(x) > \alpha\} \to 1$ для любого α. Так как $f_{n+1}(x) \geqslant f_n(x)$, то далее получаем, что $f_n(x) \to \infty$ всюду, за исключением множества x винеровской меры 0.

Мы будем также использовать символ W для обозначения случайного элемента со значениями, принадлежащими C, и с винеровской мерой в качестве распределения; так что W есть измеримое отображение некоторого пространства $(\Omega, \mathscr{B}, \mathsf{P})$ в C с тем свойством, что

$$\mathsf{P}\{\omega\colon W(\omega) \in A\} = W(A), \qquad A \in \mathscr{C},$$

где W в правой части не что иное, как обычная винеровская мера. Тот факт, что такой случайный элемент существует, следует из доводов, опирающихся на (4.4). Подобное двоякое использование символа W не должно вызывать недоразумения.

Мы будем обозначать через $W_t(\omega)$ или $W(t, \omega)$ значение случайной функции $W(\omega)$ в точке t. Таким образом, $\{W_t\colon 0 \leqslant t \leqslant 1\}$ — стохастический процесс, траектории которого непрерывны для *всех* ω. Этот процесс мы будем называть ***винеровским процессом*** или ***процессом броуновского движения***. Соотношение $X_n \xrightarrow{\mathscr{D}} W$ можно интерпретировать либо в соответствии с (4.5), либо в соответствии с (4.7) в зависимости от того, как истолковывается W, как случайная функция или как мера — смысл один и тот же для обеих интерпретаций.

Броуновский мост. Случайный элемент X из C называется *гауссовским*, если все его конечномерные распределения нормальны. Распределение гауссовского случайного элемента в C полностью задается средними $\mathsf{E}\{X(t)\}$ и смешанными моментами $\mathsf{E}\{X(s)X(t)\}$, $0 \leqslant s, t \leqslant 1$, так как они определяют конечномерные распределения. Моменты W равны

$$\mathsf{E}\{W_t\} = 0 \tag{9.9}$$

и

$$\mathsf{E}\{W_s W_t\} = s, \quad \text{если} \quad s \leqslant t. \tag{9.10}$$

При изучении поведения эмпирических функций распределения нам потребуется гауссовский элемент W° в C, распределение которого задается соотношениями

$$\mathsf{E}\{W_t^\circ\} = 0 \tag{9.11}$$

и

$$\mathsf{E}\{W_s^\circ W_t^\circ\} = s(1-t), \quad \text{если} \quad s \leqslant t. \tag{9.12}$$

Мы могли бы доказать существование такого элемента W° методами теоремы 9.1, однако проще построить W° с помощью W, полагая

$$W_t^\circ = W_t - tW_1, \qquad 0 \leqslant t \leqslant 1. \tag{9.13}$$

Очевидно, что так определенный элемент W° является гауссовским случайным элементом в C и соотношения (9.1) и (9.12) следуют из (9.9) и (9.10).

Случайный элемент W° называется *броуновским мостом* (или связанным броуновским движением). Отметим, что $W_0^\circ = W_1^\circ = 0$ с вероятностью 1. Из (9.11) и (9.12) вытекает, что

$$\mathsf{E}\{(W_t^\circ - W_s^\circ)^2\} = (t-s)(1-(t-s)), \quad \text{если} \quad s \leqslant t, \tag{9.14}$$

и что

$$\mathsf{E}\{W_{s_2}^\circ - W_{s_1}^\circ)(W_{t_2}^\circ - W_{t_1}^\circ)\} = \\ = -(s_2 - s_1)(t_2 - t_1), \quad \text{если} \quad s_1 \leqslant s_2 \leqslant t_1 \leqslant t_2. \tag{9.15}$$

Мы будем также использовать символ W° для обозначения распределения случайного элемента W° на C. Если преобразование h: $C \to C$ переводит функцию x в функцию, принимающую в точке t значение $x(t) - tx(1)$, то меры W° и W в терминах h связаны соотношением $W^\circ = Wh^{-1}$.

Сепарабельные стохастические процессы *). Свяжем нашу конструкцию винеровской меры с понятием сепарабельного стохастического процесса. По теореме Колмогорова (стр. 316) существует стохастический процесс $\{\xi_t\colon 0 \leqslant t \leqslant 1\}$ на некотором вероятностном пространстве $(\Omega, \mathcal{B}, \mathsf{P})$, для которого распределение любой ко-

*) Эта часть параграфа может быть опущена.

нечной совокупности ξ_t совпадает с соответствующим конечномерным распределением для меры W. В стандартной схеме $(\Omega, \mathscr{B})$ есть произведение прямых $(R^1, \mathscr{R}^1)$, t пробегает отрезок $[0, 1]$, а ξ_t являются координатными функциями.

Множество Ω_0 точек ω, для которых $\xi_t(\omega)$ непрерывна по t, $0 \leqslant t \leqslant 1$, не обязано удовлетворять условию $\mathsf{P}(\Omega_0) = 1$; действительно, часто при использовании стандартной схемы Ω_0 даже не принадлежит $\mathscr{B}$. Возможно, однако, с помощью иной процедуры найти *сепарабельный* процесс с предписанными конечномерными распределениями. Процесс $\{\xi_t\colon 0 \leqslant t \leqslant 1\}$ на $(\Omega, \mathscr{B}, \mathsf{P})$ является сепарабельным *), если $\mathscr{B}$ полно относительно P и если в $\mathscr{B}$ существует множество E меры 0, а в $[0, 1]$ счетное множество T_0 такое, что для любых α, β и для любого открытого интервала I мы имеем

$$\{\omega\colon \alpha \leqslant \xi_t(\omega) \leqslant \beta,\ t \in I \cap T_0\} - \\ - \{\omega\colon \alpha \leqslant \xi_t(\omega) \leqslant \beta,\ t \in I \cap T\} \subset E, \qquad (9.16)$$

где символом T обозначен отрезок $[0, 1]$. Левое множество в этой разности принадлежит $\mathscr{B}$, потому что T_0 счетно; так как $\mathsf{P}(E) = 0$ и $\mathscr{B}$ полно относительно P, то правое множество в разности также должно принадлежать $\mathscr{B}$ и иметь ту же самую вероятность.

Если процесс $\{\xi_t\}$ сепарабелен и имеет конечномерные распределения, соответствующие броуновскому движению, то **) множество Ω_0 точек ω, для которых выборочная траектория ($\xi_t(\omega)$ как функция t) непрерывна, принадлежит $\mathscr{B}$ и удовлетворяет условию $\mathsf{P}(\Omega_0) = 1$. Если мы отобразим Ω_0 в C, переводя ω в соответствующую выборочную траекторию, и если с помощью этого преобразования мы перенесем P (точнее, ограничение P на Ω_0) на $(C, \mathscr{C})$, то мы придем к W, что и является иным способом построения винеровской меры.

Фактически, задача доказательства непрерывности выборочных траекторий сепарабельных процессов совпадает с задачей построения мер на $(C, \mathscr{C})$. Если любой сепарабельный процесс с некоторыми заданными конечномерными распределениями имеет также непрерывные

---

*) См. Дуб (1956, стр. 53).
**) См. Дуб (1956, стр. 353).

с вероятностью 1 выборочные траектории, то, как показывает проведенное построение, на $(C, \mathscr{C})$ существует вероятностная мера с этими конечномерными распределениями. Докажем обратное.

Теорема 9.2. *Если $\{\xi_t\colon 0 \leqslant t \leqslant 1\}$ — сепарабельный стохастический процесс и если существует на $(C, \mathscr{C})$ вероятностная мера с конечномерными распределениями, совпадающими с конечномерными распределениями процесса $\{\xi_t\}$, то выборочные траектории процесса непрерывны с вероятностью* 1.

Доказательство. При доказательстве этого результата, не ограничивая общности, можно предположить, что T_0 содержит все рациональные точки единичного интервала. В этом случае (9.16) выполняется, если I — замкнутый интервал, концы которого имеют рациональные координаты. Мы воспользуемся соотношениями

$$\left(\bigcup_\theta E_\theta\right) + \left(\bigcup_\theta E'_\theta\right) \subset \bigcup_\theta (E_\theta + E'_\theta) \tag{9.17}$$

и

$$\left(\bigcap_\theta E_\theta\right) + \left(\bigcap_\theta E'_\theta\right) \subset \bigcup_\theta (E_\theta + E'_\theta), \tag{9.18}$$

каждое из которых справедливо в любой области изменения индекса θ *).

Обозначим через V систему чисел

$$V\colon\ k;\ r_0, \ldots, r_k;\ \alpha_1, \ldots, \alpha_k, \tag{9.19}$$

где k — произвольное целое число, r_i и α_i — рациональные числа и $0 = r_0 < \ldots < r_k = 1$. Имеется счетное множество таких систем V. Для заданной системы V и положительного числа ε определим

$$\Omega_{T_0}(V, \varepsilon) = \bigcap_{i=1}^{k} \{\omega\colon\ \alpha_i \leqslant \xi_t(\omega) \leqslant \alpha_i + \varepsilon,\ t \in [r_{i-1}, r_i] \cap T_0\} \tag{9.20}$$

и

$$\Omega_{T_0} = \bigcap_\varepsilon \bigcup_V \Omega_{T_0}(V, \varepsilon), \tag{9.21}$$

где пересечение производится по положительным рациональным ε. Обозначим $\Omega_T(V, \varepsilon)$ и Ω_T множества, зада-

*) Напомним, что знаком $+$ обозначается симметрическая разность.

ваемые теми же формулами (9.20) и (9.21), но с заменой T_0 на $T = [0, 1]$.

Так как $\Omega_T(V, \varepsilon) + \Omega_{T_0}(V, \varepsilon) \subset E$, из (9.17) и (9.18) следует, что

$$\Omega_T + \Omega_{T_0} \subset E. \tag{9.22}$$

Далее множество $\Omega_{T_0}(V, \varepsilon)$ представляет собой счетное пересечение множеств из $\mathscr{B}$ и, следовательно, само принадлежит $\mathscr{B}$. Отсюда $\Omega_{T_0} \in \mathscr{B}$. Из (9.22) вытекает, что Ω_T принадлежит $\mathscr{B}$ и $\mathsf{P}(\Omega_T) = \mathsf{P}(\Omega_{T_0})$. Так как Ω_T есть не что иное, как множество ω с непрерывными траекториями, то нам осталось лишь доказать, что

$$\mathsf{P}(\Omega_{T_0}) = 1. \tag{9.23}$$

Пусть $(t_1, t_2, \ldots)$ — занумерованные в каком-либо порядке точки множества T_0. Для системы V в (9.19) и положительного ε пусть $H_{T_0}(V, \varepsilon)$ обозначает множество точек $z = (z_1, z_2, \ldots)$ в R^∞ таких, что для каждого $i = 1, \ldots, k$ неравенство $\alpha_i \leqslant z_u \leqslant \alpha_i + \varepsilon$ выполняется для координаты с любым индексом u, для которого $t_u \in [r_{i-1}, r_i]$. Тогда $H_{T_0}(V, \varepsilon) \in \mathscr{R}^\infty$. Если отображение $\varphi\colon \Omega \to R^\infty$ определить равенством $\varphi(\omega) = (\xi_{t_1}(\omega), \xi_{t_2}(\omega), \ldots)$, то $\varphi^{-1}H_{T_0}(V, \varepsilon) = \Omega_{T_0}(V, \varepsilon)$. Если $H_{T_0} = \bigcap_\varepsilon \bigcup_V H_{T_0}(V, \varepsilon)$, то $H_{T_0} \in \mathscr{R}^\infty$ и

$$\varphi^{-1}H_{T_0} = \Omega_{T_0}. \tag{9.24}$$

Определим теперь отображение $\psi\colon C \to R^\infty$ равенством $\psi(x) = (x(t_1), x(t_2), \ldots)$. Пусть P — вероятностная мера на $(C, \mathscr{C})$, конечномерные распределения которой совпадают с конечномерными распределениями $\{\xi_t\}$. Если H — конечномерное множество в R^∞, то

$$\mathsf{P}(\varphi^{-1}H) = P(\psi^{-1}H). \tag{9.25}$$

Так как конечномерные множества образуют поле, порождающее $\mathscr{R}^\infty$, то (9.25) выполняется также для любого множества H из $\mathscr{R}^\infty$. Поэтому по формуле (9.24) $\mathsf{P}(\Omega_{T_0}) = P(\psi^{-1}H_{T_0})$. Но $\psi^{-1}H_{T_0} = C$, откуда следует (9.23).

Таким образом, доказательство непрерывности сепарабельных процессов и построение мер на C — это по существу одно и то же.

Примечание. О других конструкциях винеровской меры и об истории вопроса можно прочитать в монографии Ито, Маккина (1968). Интересное изложение ранней работы Винера см. в статьях книги «Norbert Wiener, 1894—1964», *Bull. Amer. Math Soc.* 72 (1966), 1, II.

Задачи

1. Пусть r пробегает совокупность рациональных чисел в $[0, 1]$ и ξ_r — случайный процесс с конечномерными распределениями, соответствующими броуновскому движению. Используя аргументы доказательства теоремы 9.1, но не применяя понятие плотности и не упоминая даже о пространстве C, покажите, что процесс $\{\xi_r\}$ с вероятностью 1 равномерно непрерывен по r. Определите для иррациональных t $\xi_t = \lim\limits_{r \to t} \xi_r$ и с помощью рассуждений, предшествующих теореме 9.2, постройте винеровскую меру на C заново.

2. Покажите, что винеровская мера не обладает локально компактным носителем.

3. При $0 \leqslant t < \infty$ определим $V_t = (1+t) \overset{\circ}{W}_{t/(1+t)}$, где W° — броуновский мост. Процесс $\{V_t\colon t \geqslant 0\}$ имеет выборочные траектории, которые непрерывны для всех ω, его конечномерные распределения являются гауссовскими и моменты его равны $\mathsf{E}\{V_t\} = 0$ и $\mathsf{E}\{V_s V_t\} = \min(s, t)$. Таким образом, процесс описывает броуновское движение на интервале $[0, \infty)$.

§ 10. Теорема Донскера

Теорема. Пусть даны случайные величины $\xi_1, \xi_2, \ldots$, определенные на $(\Omega, \mathscr{B}, \mathsf{P})$, пусть $S_n = \xi_1 + \ldots + \xi_n$ — их частичные суммы, и пусть случайный элемент X_n пространства C определен равенством (8.15):

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega). \tag{10.1}$$

Следующая теорема о сходимости по распределению случайного элемента X_n принадлежит Донскеру. Во введении были приведены примеры ее применения.

Теорема 10.1. *Предположим, что случайные величины ξ_n независимы и одинаково распределены со средним* 0 *и конечной ненулевой дисперсией* σ^2:

$$\mathsf{E}\{\xi_n\} = 0, \qquad \mathsf{E}\{\xi_n^2\} = \sigma^2. \tag{10.2}$$

Тогда для случайных функций X_n, определенных формулой (10.1), *выполняется соотношение*

$$X_n \xrightarrow{\mathscr{D}} W. \tag{10.3}$$

Доказательство. Мы сначала покажем, что конечномерные распределения X_n сходятся к конечномерным распределениям W. Рассмотрим сначала один момент времени s; мы должны доказать, что

$$X_n(s) \xrightarrow{\mathscr{D}} W_s. \tag{10.4}$$

Так как по неравенству Чебышева

$$\left| X_n(s) - \frac{1}{\sigma\sqrt{n}} S_{[ns]} \right| \leqslant \frac{1}{\sigma\sqrt{n}} \xi_{[ns]+1} \xrightarrow{P} 0, \tag{10.5}$$

то соотношение (10.4) будет следовать из теоремы 4.1, если доказать, что

$$\frac{1}{\sigma\sqrt{n}} S_{[ns]} \xrightarrow{\mathscr{D}} W_s.$$

Но этот факт является прямым следствием центральной предельной теоремы Леви — Линдеберга (теорема 7.4) и того факта, что $[ns]/n \to s$.

Рассмотрим теперь два момента времени s и t, $s < t$. Мы должны доказать, что

$$(X_n(s),\ X_n(t)) \xrightarrow{\mathscr{D}} (W_s,\ W_t);$$

это будет вытекать из следствия 1 к теореме 5.1, если мы покажем, что

$$(X_n(s),\ X_n(t) - X_n(s)) \xrightarrow{\mathscr{D}} (W_s,\ W_t - W_s).$$

В силу (10.5) и в точности такого же соотношения с s, замененным на t, достаточно доказать, что

$$\left(\frac{1}{\sigma\sqrt{n}} S_{[ns]},\ \frac{1}{\sigma\sqrt{n}} (S_{[nt]} - S_{[ns]}) \right) \xrightarrow{\mathscr{D}} (W_s,\ W_t - W_s). \tag{10.6}$$

Так как в силу независимости ξ_n компоненты вектора, стоящего слева, независимы, то соотношение (10.6) следует из центральной предельной теоремы и теоремы 3.2. Аналогичным образом рассматриваются случаи трех и большего числа моментов времени. Следовательно, конечномерные распределения сходятся надлежащим образом.

Мы докажем свойство плотности с помощью следующей леммы, которая является несколько более общей, чем в данный момент требуется.

Лемма. *Пусть* $\xi_1, \ldots, \xi_m$ — *независимые случайные величины со средним* 0 *и конечными дисперсиями* σ_i^2; *положим* $S_i = \xi_1 + \ldots + \xi_i$ *и* $s_i^2 = \sigma_1^2 + \ldots + \sigma_i^2$. *Тогда*

$$\mathsf{P}\{\max_{i \leqslant m} |S_i| \geqslant \lambda s_m\} \leqslant 2\mathsf{P}\{|S_m| \geqslant (\lambda - \sqrt{2})s_m\}. \quad (10.7)$$

Для доказательства неравенства (10.7) (заметим, что оно тривиально при $\lambda \leqslant \sqrt{2}$) рассмотрим множества

$$E_i = \{\max_{j < i} |S_j| < \lambda s_m \leqslant |S_i|\}.$$

Ясно, что

$$\mathsf{P}\{\max_{i \leqslant m} |S_i| \geqslant \lambda s_m\} \leqslant \mathsf{P}\{|S_m| \geqslant (\lambda - \sqrt{2})s_m\} + \\ + \sum_{i=1}^{m-1} \mathsf{P}(E_i \cap \{|S_m| < (\lambda - \sqrt{2})s_m\}). \quad (10.8)$$

Учитывая, что неравенства $|S_i| > \lambda s_m$ и $|S_m| < (\lambda - \sqrt{2})s_m$ вместе дают $|S_m - S_i| \geqslant \sqrt{2}\, s_m$, получаем из неравенства Чебышева и предположения о независимости ξ_i, что сумма в (10.8) не превосходит величины

$$\sum_{i=1}^{m-1} \mathsf{P}(E_i)\mathsf{P}\{|S_m - S_i| \geqslant \sqrt{2}\, s_m\} \leqslant \\ \leqslant \sum_{i=1}^{m-1} \mathsf{P}(E_i)\frac{1}{2s_m^2} \sum_{k=i+1}^{m} \sigma_k^2 \leqslant \frac{1}{2}\sum_{i=1}^{m-1} \mathsf{P}(E_i) \leqslant \\ \leqslant \frac{1}{2}\mathsf{P}\{\max_{i \leqslant m} |S_i| \geqslant \lambda s_m\}. \quad (10.9)$$

Теперь неравенства (10.8) и (10.9) вместе приводят к (10.7).

Применяя лемму к случайным величинам, рассматривающимся в теореме 10.1, получаем при $\lambda > 2\sqrt{2}$

$$\mathsf{P}\{\max_{i \leqslant n} |S_i| \geqslant \lambda\sigma\sqrt{n}\} \leqslant 2\mathsf{P}\left\{|S_n| \geqslant \frac{1}{2}\lambda\sigma\sqrt{n}\right\}.$$

По центральной предельной теореме

$$\mathsf{P}\left\{|S_n| \geqslant \frac{1}{2}\lambda\sigma\sqrt{n}\right\} \to \mathsf{P}\left\{|N| \geqslant \frac{1}{2}\lambda\right\} < \frac{8}{\lambda^3}\mathsf{E}\{|N|^3\}.$$

Поэтому, если ε положительно, мы имеем

$$\limsup_{n\to\infty} \mathsf{P}\left\{\max_{i\leqslant n}|S_i|\geqslant\lambda\sigma\sqrt{n}\right\}<\frac{\varepsilon}{\lambda^2} \qquad (10.10)$$

для достаточно больших λ. Плотность $\{X_n\}$ следует теперь из теоремы 8.4.

Приложение. Как указано во введении, теорема Донскера имеет следующую качественную интерпретацию: предельное соотношение $X_n \xrightarrow{\mathscr{D}} W$ означает, что если τ мало, то частица, подвергающаяся независимым смещениям $\xi_1, \xi_2, \ldots$ в последовательные моменты времени τ, $2\tau, \ldots$ и рассматриваемая издалека, будет совершать приближенно броуновское движение.

Более важным, нежели подобная качественная интерпретация, является использование теоремы Донскера для доказательства предельных теорем для различных функций от частичных сумм S_n. Во введении намечено, как использовать соотношение $X_n \xrightarrow{\mathscr{D}} W$ для вывода предельного распределения случайной величины $\max\limits_{i\leqslant n} S_i$; теперь мы осуществим это в деталях.

Так как $h(x)=\sup_t x(t)$ — непрерывная функция в C, то из соотношения $X_n \xrightarrow{\mathscr{D}} W$ в силу следствия 1 к теореме 5.1 получаем, что

$$\sup_t X_n(t) \xrightarrow{\mathscr{D}} \sup_t W_t.$$

Из очевидного равенства

$$\sup_{0\leqslant t\leqslant 1} X_n(t)=\max_{i\leqslant n}\frac{1}{\sigma\sqrt{n}}S_i$$

следует теперь, что

$$\frac{1}{\sigma\sqrt{n}}\max_{i\leqslant n} S_i \xrightarrow{\mathscr{D}} \sup_t W_t. \qquad (10.11)$$

Таким образом, мы получили бы предельное распределение величины $\max\limits_{i\leqslant n} S_i$ (надлежащим образом нормированной), если бы знали распределение $\sup_t W_t$. (Мы, разумеется, пока придерживаемся предположений теоремы 10.1). Для нахождения этого последнего распределения вычислим предельное распределение $\max\limits_{i \leqslant n} S_i$ в одном простом частном случае.

Допустим, что $S_0, S_1, \ldots$ — это случайные величины в симметричном случайном блуждании, выходящем из

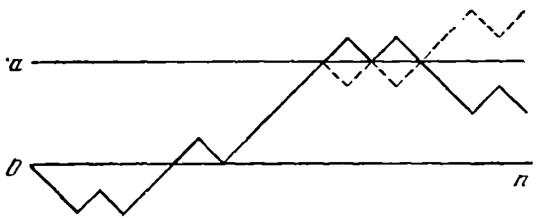

Рис. 5.

начала координат (рис. 5). Иными словами, предположим, что случайные величины ξ_n независимы и удовлетворяют условию

$$\mathsf{P}\{\xi_n = 1\} = \mathsf{P}\{\xi_n = -1\} = 1/2. \tag{10.12}$$

Покажем, что для любого неотрицательного целого числа a

$$\mathsf{P}\{\max_{0 \leqslant i \leqslant n} S_i \geqslant a\} = 2\mathsf{P}\{S_n > a\} + \mathsf{P}\{S_n = a\}. \tag{10.13}$$

Случай $a = 0$ очевиден. Пусть $a > 0$ и $M_i = \max\limits_{0 \leqslant j \leqslant i} S_j$. Так как

$$\mathsf{P}\{M_n \geqslant a\} - \mathsf{P}\{S_n = a\} = \\ = \mathsf{P}\{M_n \geqslant a,\ S_n < a\} + \mathsf{P}\{M_n \geqslant a,\ S_n > a\}$$

и

$$\mathsf{P}\{M_n \geqslant a,\ S_n > a\} = \mathsf{P}\{S_n > a\},$$

то равенство (10.13) будет выполнено, если мы покажем, что

$$\mathsf{P}\{M_n \geqslant a,\ S_n < a\} = \mathsf{P}\{M_n \geqslant a,\ S_n > a\}. \tag{10.14}$$

Ввиду условия (10.12) все 2^n возможные траектории $(S_1, S_2, \ldots, S_n)$ имеют одну и ту же вероятность 2^{-n}. Поэтому равенство (10.14) будет установлено, если мы покажем, что число траекторий, входящих в событие в левой части (10.14), равно числу траекторий, входящих в событие в правой части, а чтобы это показать, достаточно в свою очередь установить для траекторий взаимно однозначное соответствие*). Фиксируем траекторию $(S_1, S_2, \ldots, S_n)$, входящую в событие в левой части (10.14), и сопоставим ей траекторию, полученную отражением относительно a всех частичных сумм, начиная с той суммы, которая первой достигает уровня a. Так как соответствие взаимно однозначное, то (10.14) выполняется. Это рассуждение являет собой пример *принципа симметрии* (*отражения*).

Пусть α — произвольное неотрицательное число, и пусть $a_n = -[-\alpha n^{1/2}]$. Из (10.14) мы имеем

$$\mathsf{P}\left\{\max_{i\leqslant n}\frac{1}{\sqrt{n}}S_i\geqslant\alpha\right\}=2\mathsf{P}\{S_n>a_n\}+\mathsf{P}\{S_n=a_n\}. \quad (10.15)$$

По центральной предельной теореме

$$\mathsf{P}\{S_n\geqslant a_n\}\to\mathsf{P}\{N\geqslant\alpha\}$$

(ввиду условия (10.12) в этом случае $\sigma^2=1$). Так как наибольшая вероятность в симметричном биномиальном распределении стремится к 0, то слагаемое $\mathsf{P}\{S_n=a_n\}$ в (10.15) стремится к нулю. Таким образом,

$$\mathsf{P}\left\{\max_{i\leqslant n}\frac{1}{\sqrt{n}}S_i\geqslant\alpha\right\}\to 2\mathsf{P}\{N\geqslant\alpha\},\qquad \alpha\geqslant 0. \quad (10.16)$$

Объединяя (10.16) и (10.11) (по-прежнему в предположении (10.12)), мы выводим

$$\mathsf{P}\{\sup_t W_t\leqslant\alpha\}=\frac{2}{\sqrt{2\pi}}\int_0^\alpha e^{-\frac{1}{2}u^2}\,du,\qquad \alpha\geqslant 0. \quad (10.17)$$

Очевидно, что левая часть соотношения (10.17) равна 0, если $\alpha<0$. Итак, мы установили некоторый факт, касающийся броуновского движения, комбинируя теорему Донскера с прямым вычислением вероятностей для случайного блуждания; простота этого вычисления объясняется отчасти тем, что оно было сведено к подсчету числа траекторий, отчасти тем, что случайное блуждание

*) Новый прием (в подлиннике: The new mathematics).

не может оказаться выше некоторого уровня a (a — целое положительное число), не пересекая этот уровень.

Теперь мы исключим предположение (10.12). Если величины ξ_n независимы, одинаково распределены и удовлетворяют условию (10.12), то соотношение (10.11) выполняется и из (10.17) можно теперь вывести

$$\mathsf{P}\left\{\frac{1}{\sigma\sqrt{n}}\max_{i\leqslant n} S_i \leqslant \alpha\right\} \to \frac{2}{\sqrt{2\pi}}\int_0^\alpha e^{-\frac{1}{2}u^2}\,du, \quad \alpha \geqslant 0. \tag{10.18}$$

Таким образом, мы получили предельное распределение для $\max_{i\leqslant n} S_i$ в предположениях теоремы Линдеберга — Леви.

Это рассуждение имеет общий характер. Если функция h непрерывна на C или непрерывна всюду, за исключением множеств винеровской меры 0, то соотношение $X_n \xrightarrow{\mathscr{D}} W$ влечет

$$h(X_n) \xrightarrow{\mathscr{D}} h(W). \tag{10.19}$$

(В только что рассмотренном случае $h(x) = \sup_t x(t)$.) Мы можем найти предельное распределение $h(X_n)$, если известно распределение $h(W)$, и часто можем найти распределение для $h(W)$, отыскивая предельное распределение для $h(X_n)$ в каком-либо частном случае и затем используя соотношение (10.19) в обратном направлении. Следующий параграф содержит дальнейшие примеры подобных рассуждений.

Следовательно, если величины ξ_n независимы и одинаково распределены с $\mathsf{E}\{\xi_n\} = 0$ и $\mathsf{E}\{\xi_n^2\} = \sigma^2$, то предельное распределение $h(X_n)$ не зависит от любых добавочных свойств ξ_n. По этой причине теорема Донскера часто называется *принципом инвариантности*. Здесь мы будем употреблять другое название — *функциональная центральная предельная теорема*. Под этим названием в дальнейшем будут объединены несколько аналогичных теорем (точно так же, как и термин *центральная предельная теорема* используют для обозначения целого класса теорем).

Необходимое условие плотности. Исключим предположение о независимости величин ξ_n и непосредственно предположим, что $X_n \xrightarrow{\mathcal{D}} W$ (X_n по-прежнему определены формулой (10.1)). Мы покажем, что в этих условиях для любого положительного ε существует λ ($\lambda > 1$) такое, что неравенство

$$\mathsf{P}\{\max_{i \leqslant n} |S_i| \geqslant \lambda\sigma\sqrt{n}\} \leqslant \frac{\varepsilon}{\lambda^2} \tag{10.20}$$

выполняется для всех n, начиная с некоторого n_0. В случае, если последовательность $\{\xi_n\}$ стационарна, это утверждение в точности совпадает с предположением теоремы 8.4, которое мы проверяли при доказательстве теоремы Донскера.

Пусть $Y = \sup_t |W_t|$. Ввиду равенства (10.17) и симметрии W при отражении относительно 0 случайная величина Y имеет конечный второй момент. Из (10.19) при $h(x) = \sup_t |x(t)|$ следует, что для любого положительного числа λ

$$\mathsf{P}\{\max_{i \leqslant n} |S_i| \geqslant \lambda\sigma\sqrt{n}\} \to \mathsf{P}\{Y \geqslant \lambda\} \leqslant \frac{1}{\lambda^2} \int\limits_{\{Y \geqslant \lambda\}} Y^2\, d\mathsf{P}. \tag{10.21}$$

По заданному ε выберем λ столь большим, чтобы интеграл в последней формуле был меньше ε; из (10.21) следует теперь, что неравенство (10.20) выполняется для всех достаточно больших n.

Таким образом, мы видим теперь, что (10.20) является именно тем условием, которое нужно проверять при доказательстве плотности в теореме Донскера. В главе 4 мы используем это же условие при доказательстве функциональных центральных предельных теорем для последовательностей зависимых случайных величин.

Предыдущее рассуждение в действительности приводит к следующему факту: предположим, что последовательность ξ_n стационарна и что конечномерные распределения последовательности X_n, определенной в (10.1), сходятся к конечномерным распределениям случайной функции X, причем $\sup_t |X(t)|$ имеет конечный второй момент. При этих условиях, если $\{X_n\}$ плотна, то предположения теоремы 8.4 выполняются. В этом смысле условие теоремы 8.4 является необходимым (см. абзац, следующий за формулой (8.9)).

Другое доказательство теоремы Донскера *). Нижеследующая часть этого параграфа посвящена второму доказательству теоремы 10.1. Это доказательство интересно тем, что оно мало использует предыдущую теорию. В действительности, считая доказанным существование винеровской меры **), мы нуждаемся лишь в теореме 2.2 и центральной предельной теореме.

Нам нужно доказать, что $P_n \Rightarrow W$, где P_n — распределение X_n. Согласно следствию 2 из теоремы 2.2 нам достаточно доказать, что $P_n(A) \to W(A)$ всякий раз, когда множество A является пересечением открытых сфер и удовлетворяет условию $W(\partial A) = 0$. Далее, если A есть пересечение сфер с центрами $x_1, \ldots, x_k$ и радиусами $\varepsilon_1, \ldots, \varepsilon_k$ и если

$$y(t) = \max_{1 \leqslant i \leqslant k} (x_i(t) - \varepsilon_i)$$

и

$$z(t) = \min_{1 \leqslant i \leqslant k} (x_i(t) + \varepsilon_i),$$

то A имеет вид

$$A = \{x\colon\ y(t) < x(t) < z(t),\ 0 \leqslant t \leqslant 1\}. \tag{10.22}$$

Нам нужно доказать, что

$$P_n(A) \to W(A) \tag{10.23}$$

в предположении, что A определяется равенством (10.22) с $y, z \in C$ и что $W(\partial A) = 0$. Мы можем также предположить, что $y(t) < z(t)$ для всех t, так как в ином случае A пусто. (Если $y(0) \neq 0$ и $z(0) \neq 0$, то автоматически $W(\partial A) = 0$, но этот случай нас не интересует.) С этого момента y и z предполагаются фиксированными.

В первом доказательстве теоремы Донскера мы показали, что конечномерные распределения P_n слабо сходятся к конечномерным распределениям W. Это утверждение, в котором никак не участвует пространство C, будем считать известным. Таким образом,

$$\lim_{n\to\infty} P_n\{x\colon\ (x(t_1), \ldots, x(t_k)) \in H\} = \\ = W\{x\colon\ (x(t_1), \ldots, x(t_k)) \in H\} \tag{10.24}$$

---

*) Остальная часть этого параграфа может быть пропущена.

**) Что само по себе может быть установлено прямым рассуждением; см. задачу 1 § 9.

для любого k-мерного борелевского множества H, удовлетворяющего условию

$$W\{x:\ (x(t_1), \ldots, x(t_k)) \in \partial H\} = 0.$$

Если для целых значений v положить

$$A_v = \left\{x:\ y\left(\frac{i}{v}\right) < x\left(\frac{i}{v}\right) < z\left(\frac{i}{v}\right),\ i = 0, 1, \ldots, v\right\},$$

то при u, стремящемся к бесконечности, A_{2^u} сужается до множества тех x, для которых выполняется неравенство $y(t) < x(t) < z(t)$ при всех двоично рациональных числах t. Но если $W(\partial A) = 0$, то последнее множество отличается от A множеством W-меры 0. Поэтому для каждого положительного η существует v такое, что $W(A) \leqslant W(A_v) \leqslant W(A) + \eta$. Так как $P_n(A) \leqslant P_n(A_v)$ и так как по (10.24) $\lim_n P_n(A_v) = W(A_v)$, то мы имеем $\limsup_n P_n(A) \leqslant W(A) + \eta$. В силу произвольности η

$$\limsup_n P_n(A) \leqslant W(A). \tag{10.25}$$

Мы завершим доказательство, показав, что

$$\liminf_n P_n(A) \geqslant W(A). \tag{10.26}$$

Эта часть доказательства является наиболее трудной. По заданному η выберем ε так, что если

$$D(\varepsilon) = \{x:\ y(t) + 3\varepsilon < x(t) < z(t) - 3\varepsilon,\ 0 \leqslant t \leqslant 1\}, \tag{10.27}$$

то $W(D(\varepsilon)) > W(A) - \eta$. (Очевидно, что множество (10.27) возрастает до множества A, когда ε стремится к 0.) Определим теперь

$$D_v(\varepsilon) = \left\{x:\ y\left(\frac{k}{v}\right) + 3\varepsilon < x\left(\frac{k}{v}\right) < z\left(\frac{k}{v}\right) - 3\varepsilon,\ k = 0, 1, \ldots, v\right\}.$$

При каждом v мы имеем $D(\varepsilon) \subset D_v(\varepsilon)$ и, следовательно,

$$W(D_v(\varepsilon)) > W(A) - \eta. \tag{10.28}$$

Выберем и зафиксируем v столь большим, что

$$\frac{9}{\varepsilon^2 v} < \eta, \quad w_y\left(\frac{1}{v}\right) < \varepsilon, \quad w_z\left(\frac{1}{v}\right) < \varepsilon, \tag{10.29}$$

где w обозначает модуль непрерывности (8.1). Тогда при $n \geqslant v$

$$w_y\left(\frac{1}{n}\right) < \varepsilon, \qquad w_z\left(\frac{1}{n}\right) < \varepsilon. \tag{10.30}$$

Пусть L_n обозначает множество элементов C, которые линейны на каждом интервале $[(i-1)/n, i/n]$, $i = 1, \ldots, n$, и пусть G_n — множество x из C, для которых неравенство

$$y\left(\frac{i}{n}\right) + \varepsilon \leqslant x\left(\frac{i}{n}\right) \leqslant z\left(\frac{i}{n}\right) - \varepsilon \tag{10.31}$$

нарушается при некоторых $i = 0, 1, \ldots, n$. Если $x \in L_n$, $x \in G_n^c$ (так что (10.31) выполняется для всех $i = 0, 1, \ldots, n$) и если $n \geqslant v$, то в силу (10.30) $y(t) < x(t) < z(t)$ при всех $t \in [0, 1]$, так что $x \in A$. Таким образом, $L_n \cap G_n^c \subset A$. Пусть $G_{n,r}$ — множество x, для которых двустороннее неравенство (10.31) выполняется при $i < r$, но не выполняется при $i = r$. Так как $P_n(L_n) = 1$ и так как $G_{n,r}$ не пересекаются и в объединении составляют G, то мы имеем

$$P_n(A^c) \leqslant \sum_{r=0}^{n} P_n(G_{n,r}), \qquad n \geqslant v. \tag{10.32}$$

При $1 \leqslant r \leqslant n$ пусть $k_{n,r}$ обозначает такое целое k, $1 \leqslant k \leqslant v$, что $(k-1)/v < r/n \leqslant k/v$, и пусть $k_{n,0} = 0$. По формуле (10.32)

$$P_n(A^c) \leqslant \sum_{r=1}^{n} P_n\left(G_{n,r} \cap \left\{x\colon \left|x\left(\frac{r}{n}\right) - x\left(\frac{k_{n,r}}{v}\right)\right| < \varepsilon\right\}\right) + \\ + \sum_{r=1}^{n} P_n\left(G_{n,r} \cap \left\{x\colon \left|x\left(\frac{r}{n}\right) - x\left(\frac{k_{n,r}}{v}\right)\right| \geqslant \varepsilon\right\}\right). \tag{10.33}$$

Если $x \in G_{n,r}$ и $|x(r/n) - x(k_{n,r}/v)| < \varepsilon$, то в силу (10.30)

$$x\left(\frac{k_{n,r}}{v}\right) \notin \left[y\left(\frac{k_{n,r}}{v}\right) + 3\varepsilon,\ z\left(\frac{k_{n,r}}{v}\right) - 3\varepsilon\right],$$

так что $x \notin D_v(\varepsilon)$. Таким образом, первая сумма в правой части (10.33) не больше, чем $P_n(D_v(\varepsilon)^c)$. Поэтому при $n \geqslant v$

$$P_n(A^c) \leqslant$$
$$\leqslant P_n(D_v(\varepsilon)^c) + \sum_{r=1}^{n} P_n\left(G_{n,r} \cap \left\{x: \left|x\left(\frac{r}{n}\right) - x\left(\frac{k_{n,r}}{v}\right)\right| \geqslant \varepsilon\right\}\right). \tag{10.34}$$

Покажем теперь, что сумма в (10.34) мала. Заметим сначала, что при всех s, t и n

$$\int (x_t - x_s)^2\, dP_n \leqslant 9\,|t - s|. \tag{10.35}$$

В самом деле, если s и t имеют вид i/n и j/n, то левая часть (10.35) равна $\mathsf{E}\{(S_i - S_j)^2/(n\sigma^2)\} = |t - s|$; если s и t лежат в одном и том же подинтервале $[(i-1)/n, i/n]$, то левая часть (10.35) равна $n(t-s)^2$, что не превышает $|t - s|$; общий результат следует теперь из неравенства Минковского. В силу независимости величин ξ_n имеем, следовательно,

$$P_n\left(G_{n,r} \cap \left\{x: \left|x\left(\frac{r}{n}\right) - x\left(\frac{k_{n,r}}{v}\right)\right| \geqslant \varepsilon\right\}\right) =$$
$$= P_n(G_{n,r})\, P_n\left\{x: \left|x\left(\frac{r}{n}\right) - x\left(\frac{k_{n,r}}{v}\right)\right| \geqslant \varepsilon\right\} \leqslant$$
$$\leqslant P_n(G_{n,r})\left(\frac{9}{\varepsilon^2}\right)\left(\frac{r}{n} - \frac{k_{n,r}}{v}\right). \tag{10.36}$$

Так как $r/n - k_{n,r}/v \leqslant 1/v$, то из (10.29) следует, что последний член в (10.36) не превосходит $\eta P_n(G_{n,r})$. Используя неравенство (10.34) и тот факт, что сумма вероятностей $P(G_{n,r})$ не превышает 1, мы приходим к оценке

$$P_n(A^c) \leqslant P_n(D_v(\varepsilon)^c) + \eta, \quad n \geqslant v.$$

Но из (10.24) вытекает, что $\lim_n P_n(D_v(\varepsilon)^c) = W(D_v(\varepsilon)^c)$, и поэтому в силу (10.28)

$$\limsup_n P_n(A^c) \leqslant W(A^c) + 2\eta.$$

Поскольку η произвольно, то тем самым неравенство (10.26) установлено. В комбинации с оценкой (10.25) оно приводит к предельному соотношению (10.23) и тем самым теорема Донскера установлена еще раз.

Примечание. Эрдеш и Кац (1946) первые предложили принцип инвариантности как способ отыскания предельных распределений величин, подобных $\max_{k \leqslant n} S_k$: сначала предельное распределение находится для специального случая, а затем производится переход к общему случаю посредством установления связи этого результата с броуновским движением (именно так и возникло это направление). Еще раньше в этой области появилась работа Колмогорова (1931).

Первоначальное утверждение Донскера (1951) состояло не в том, что $X_n \xrightarrow{\mathscr{D}} W$ или что $P_n \Rightarrow W$, а в том, что $h(X_n) \xrightarrow{\mathscr{D}} h(W)$ для всех действительных непрерывных фукций h из C (в силу теоремы 5.2 это то же самое). Аргументы, приводящие к утверждению (10.23), принадлежат Донскеру. Колмогоров и Прохоров (1954) впервые указали, что слабая сходимость имеет прямое сюда отношение и что $P_n \Rightarrow W$ следует из (10.23) и теоремы 2.2. Прохоров (1953 и 1956) стал применять понятие плотности. О совершенно ином подходе к данным проблемам см. Найт (1962) и о совсем ином их применении см. Страссен (1964). Ламперти (1962а) получил результат, отличающийся от теоремы 10.1 тем, что предельным процессом является не броуновское движение.

Задачи

1. Пусть случайные величины $\xi_{n1}, \ldots, \xi_{nk_n}$ независимы со средним 0 и дисперсиями σ_{ni}^2; положим $S_{ni} = \xi_{n1} + \ldots + \xi_{ni}$, $s_{ni}^2 = \sigma_{n1}^2 + \ldots + \sigma_{ni}^2$ и $s_n^2 = s_{nk_n}^2$. Пусть случайная функция X_n линейна на каждом интервале $\left[s_{n,\,i-1}^2/s_n^2,\ s_{ni}^2/s_n^2\right]$ и принимает значения $X_n\left(s_{ni}^2/s_n^2\right) = S_{ni}/s_n$ в точках деления. Полагая что условие Линдеберга (7.3) выполнено, обобщите теорему Донскера: используя теорему Линдеберга, неравенство (10.7) и следствие теоремы 8.3, покажите, что $X_n \xrightarrow{\mathscr{D}} W$. (Этот результат принадлежит Прохорову (1956).)

§ 11. Функции от траекторий броуновского движения

В предыдущем параграфе мы использовали теорему Донскера и свойства случайного блуждания для нахождения распределения $\sup_t W_t$. Найдя это распределение, мы еще раз использовали теорему Донскера для нахождения предельного распределения $\max_{i \leqslant n} S_i$ в общем случае. В данном параграфе мы используем те же самые приемы для других функций от траекторий броуновского движения и частичных сумм. Мы вычислим также некоторые распределения, связанные с броуновским мостом *).

Будем говорить, что имеет место случай Линдеберга — Леви, если рассматриваются частичные суммы S_n независимых одинаково распределенных случайных величин ξ_n со средним 0 и конечной положительной дисперсией σ^2. В этом случае X_n всегда будет обозначать случайную функцию (10.1). Будем говорить, что мы имеем дело со случайным блужданием, если каждая из величин ξ_n принимает значения $+1$ и -1 с вероятностью 1/2 каждое. В этом случае $\sigma^2 = 1$.

Максимум и минимум. Пусть $m = \inf_t W_t$, $M = \sup_t W_t$, и пусть

$$m_n = \min_{0 \leqslant i \leqslant n} S_i, \qquad M_n = \max_{0 \leqslant i \leqslant n} S_i \tag{11.1}$$

— соответствующие величины для частичных сумм. Отображение, переводящее точку x из C в точку $(\inf_t x(t), \sup_t x(t), x(1))$ пространства R^3, всюду непрерывно, так что по теоремам 10.1 и 5.1 в случае Линдеберга — Леви

$$\frac{1}{\sigma\sqrt{n}}(m_n, M_n, S_n) \xrightarrow{\mathscr{D}} (m, M, W_1). \tag{11.2}$$

Мы найдем сначала явную формулу для вероятности

$$p_n(a, b, v) = \mathsf{P}\{a < m_n \leqslant M_n < b, S_n = v\} \tag{11.3}$$

для случайного блуждания. Покажем, что если

$$p_n(j) = \mathsf{P}\{S_n = j\}, \tag{11.4}$$

то

$$p_n(a, b, v) = \\ = \sum_{k=-\infty}^{\infty} p_n(v + 2k(b-a)) - \sum_{k=-\infty}^{\infty} p_n(2b - v + 2k(b-a)) \tag{11.5}$$

для любых целых a, b и v, удовлетворяющих условиям

$$a \leqslant 0 \leqslant b, \qquad a < b, \qquad a \leqslant v \leqslant b. \tag{11.6}$$

---

*) Результаты о слабой сходимости в C имели бы малую ценность, если бы было невозможно провести вычисления, подобные приводимым в этом параграфе. Однако вычисления сами по себе несущественны для понимания общей теории.

Если $a < b$, то ряды в (11.5) в действительности являются конечными суммами. Отметим, что обе части (11.5) обращаются в 0, если n и v имеют противоположную четность.

Для частных значений n, a, b и v обозначим равенство (11.5) через $[n, a, b, v]$. Мы докажем, с помощью индукции по n, что $[n, a, b, v]$ справедливо, если выполняется (11.6). При $n = 0$ это следует из непосредственного рассмотрения возможных случаев. Предположим в соответствии с методом индукции, что $[n-1, a, b, v]$ выполняется для a, b, v, удовлетворяющих (11.6). Если $a = 0$, то (11.3) обращается в 0 (заметим, что i изменяется от 0 в (11.1)) и суммы справа в (11.5) сокращаются, поскольку $p_n(j) = p_n(-j)$. Таким образом, $[n, a, b, v]$ справедливо, если (11.6) выполняется и $a = 0$; случай $b = 0$ разбирается аналогичным способом. Для завершения рассуждения индукции мы должны убедиться в справедливости $[n, a, b, v]$ при том, что $a < 0 < b$ и $a \leqslant v \leqslant b$. Но в этом случае $a + 1 \leqslant 0$ и $b - 1 \geqslant 0$, так что оба равенства $[n-1,\ a-1,\ b-1,\ v-1]$ и $[n-1,\ a+1,\ b+1,\ v+1]$ удовлетворяют предположению индукции и, следовательно, оба справедливы. Равенство $[n, a, b, v]$ следует теперь из вероятностно очевидных рекуррентных формул

$$p_n(j) = \frac{1}{2} p_{n-1}(j-1) + \frac{1}{2} p_{n-1}(j+1)$$

и

$$p_n(a, b, v) = \frac{1}{2} p_{n-1}(a-1,\ b-1,\ v-1) + \\ + \frac{1}{2} p_{n-1}(a+1,\ b+1,\ v+1)\,^{*)}.$$

Из равенства (11.5) суммированием по v получаем, что если

$$a \leqslant 0 \leqslant b, \qquad a \leqslant u < v \leqslant b, \tag{11.7}$$

*) В задаче 2 указано, как можно прийти к соотношению (11.5) (вместо того, чтобы его проверять).

то

$$\mathsf{P}\{a < m_n \leqslant M_n < b,\ u < S_n < v\} = \sum_{k=-\infty}^{\infty} \mathsf{P}\{u + 2k(b-a) < S_n < v + 2k(b-a)\} - \sum_{k=-\infty}^{\infty} \mathsf{P}\{2b - v + 2k(b-a) < S_n < 2b - u + 2k(b-a)\}. \tag{11.8}$$

Положив в этой формуле $a = -n-1$, мы придем к соотношению

$$\mathsf{P}\{M_n < b,\ u < S_n < v\} = \mathsf{P}\{u < S_n < v\} - \mathsf{P}\{2b - v < S_n < 2b - u\}, \tag{11.9}$$

которое справедливо при $-n-1 \leqslant u < v \leqslant b,\ b \geqslant 0$ *). С помощью (11.9) можно снова получить равенство (10.13).

Итак, (11.8) выполняется для случайного блуждания, и в силу (11.2) мы можем найти распределение случайного вектора (m, M, W_1), переходя к пределу. Если a, b, u и v — суть действительные числа, удовлетворяющие (11.7), заменим их в (11.8) целыми числами $[an^{1/2}]$, $-[-bn^{1/2}]$, $[un^{1/2}]$, $-[-vn^{1/2}]$, соответственно. В силу центральной предельной теоремы и непрерывности нормального распределения почленный переход к пределу в (11.8) приводит к равенству

$$\mathsf{P}\{a < m \leqslant M < b,\ u < W_1 < v\} = \sum_{k=-\infty}^{\infty} \mathsf{P}\{u + 2k(b-a) < N < v + 2k(b-a)\} - \sum_{k=-\infty}^{\infty} \mathsf{P}\{2b - v + 2k(b-a) < N < 2b - u + 2k(b-a)\}. \tag{11.10}$$

Изменение порядка перехода к пределу и суммирования по k может быть обосновано вариантом теоремы Шеффе (стр. 307), приспособленным для рядов.

---

*) См. другое доказательство в задаче 1.

Совместное распределение M и W_1 можно получить, устремляя a к $-\infty$ в (11.10), но проще вернуться к случайному блужданию и перейти к пределу в (11.9), что дает равенство

$$P\{M<b,\ u<W_1<v\}=$$
$$=P\{u<N<v\}-P\{2b-v<N<2b-u\}, \qquad (11.11)$$

справедливое при $u\leqslant v\leqslant b$, $b\geqslant 0$. Полагая $v=b$ и устремляя $u\to-\infty$, возвращаемся к (10.17).

Из (11.10) с $u=a$ и $v=b$ вытекает, что

$$P\{a<m\leqslant M<b\}=$$
$$=\sum_{k=-\infty}^{\infty}(-1)^k P\{a+k(b-a)<N<b+k(b-a)\} \qquad (11.12)$$

при $a\leqslant 0\leqslant b$. Отсюда при $a=-b$ получаем для всех $b\geqslant 0$

$$P\{\sup_t|W_t|<b\}=$$
$$=\sum_{k=-\infty}^{\infty}(-1)^k P\{(2k-1)b<N<(2k+1)b\}. \qquad (11.13)$$

По непрерывности строгие неравенства во всех этих формулах могут быть дополнены равенствами. Правые части могут быть все переписаны как суммы нормальных интегралов. Можно *) преобразовать ряды в (11.13) к виду

$$1-\frac{4}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^k}{2k+1}e^{-\pi^2(2k+1)^2/(8b^2)}.$$

Хотя мы вывели равенства (11.10) — (11.13) переходом к пределу от случайного блуждания, мы получаем предельные распределения для (m_n, M_n, S_n), (M_n, S_n), (m_n, M_n) и $\max_{i\leqslant n}|S_i|$ (при соответствующей нормировке) в более общем случае Линдеберга — Леви, поскольку в этом случае выполняется соотношение (11.2).

Закон арксинуса. Для $x\in C$ пусть $h_1(x)$ есть супремум тех t из [0, 1], для которых $x(t)=0$; пусть $h_2(x)$ —

*) См. Феллер (1967b, стр. 406 и 710).

мера Лебега множества тех t из $[0, 1]$, для которых $x(t) > 0$; пусть, наконец, $h_3(x)$ — мера Лебега множества тех t из $[0, h_1(x)]$, для которых $x(t) > 0$. Тогда $T = h_1(W)$ есть момент времени, когда винеровская траектория в последний раз проходит через 0, $U = h_2(W)$ есть полное время, которое W проводит выше 0, и $V = h_3(W)$ — время, которое W проводит выше 0 в интервале $[0, T]$. Мы найдем совместное распределение (T, U, V, W_1).

В Добавлении II (стр. 316 и далее) показано, что каждое из отображений h_1, h_2 и h_3 измеримо и непрерывно всюду, за исключением множества винеровской меры 0. Поэтому в случае Линдеберга — Леви

$$(h_1(X_n),\ h_2(X_n),\ h_3(X_n),\ X_n(1)) \xrightarrow{\mathcal{D}} (T,\ U,\ V,\ W_1). \tag{11.14}$$

В схеме случайного блуждания вектор слева в (11.14) имеет простую интерпретацию: $T_n = nh_1(X_n)$ есть максимум i, $1 \leqslant i \leqslant n$, для которых $S_i = 0$; $U_n = nh_2(X_n)$ есть число i, $1 \leqslant i \leqslant n$, для которых обе суммы S_{i-1} и S_i неотрицательны; $V_n = nh_3(X_n)$ есть число i, $1 \leqslant i \leqslant T_n$, для которых обе суммы S_{i-1} и S_i неотрицательны и, конечно, $X_n(1) = S_n/\sqrt{n}$.

В введенных обозначениях в схеме случайного блуждания имеем, следовательно,

$$\left(\frac{1}{n} T_n,\ \frac{1}{n} U_n,\ \frac{1}{n} V_n,\ \frac{1}{\sqrt{n}} S_n\right) \xrightarrow{\mathcal{D}} (T,\ U,\ V,\ W_1). \tag{11.15}$$

Мы найдем распределение (T, U, V, W_1), переходя к пределу. В общем случае Линдеберга — Леви левая часть (11.14) является более сложной функцией частичных сумм $S_1, \ldots, S_n$. После того, как мы найдем распределение (T, U, V, W_1), мы покажем, что соотношение (11.14) даже в общем случае приводит к предельным теоремам для величин, естественным образом связанных с частичными суммами.

Так как компоненты случайного вектора (T, U, V, W_1) связаны соотношением

$$U = \begin{cases} 1 - T + V & \text{при } W_1 \geqslant 0, \\ V & \text{при } W_1 \leqslant 0, \end{cases} \tag{11.16}$$

то достаточно рассмотреть (T, V, W_1) и соответствующий вектор (T_n, V_n, S_n). Распределение последней величины в схеме случайного блуждания мы получим, опираясь на три факта, допускающих элементарные доказательства, которые мы здесь приводить не будем.

Во-первых, мы используем локальную предельную теорему для случайного блуждания: если m стремится к бесконечности и j меняется вместе с m таким образом, что j и m имеют одну и ту же четность и $j/\sqrt{m} \to y$, то*)

$$\frac{\sqrt{m}}{2} p_m(j) \to \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}. \tag{11.17}$$

Во-вторых, мы используем тот факт, что

$$\mathsf{P}\{S_1 > 0, \ldots, S_{m-1} > 0, S_m = j\} = \frac{j}{m} p_m(j) \tag{11.18}$$

для положительных j **). Если $S_{2m} = 0$, то величина $U_{2m} = V_{2m}$ принимает одно из значений $0, 2, \ldots, 2m$. Третий факт ***), который мы будем использовать, заключается в том, что эти $m+1$ значения имеют одну и ту же условную вероятность:

$$P\{V_{2m} = 2i \mid S_{2m} = 0\} = \frac{1}{m+1}, \quad i = 0, 1, \ldots, m. \tag{11.19}$$

Для вычисления вероятности того, что $T_n = 2k$, $V_n = 2i$ и $S_n = j$, возьмем в качестве условия событие $S_{2k} = 0$. При этом условии последовательности $(S_0, \ldots, S_{2k})$ и $(S_{2k+1}, \ldots, S_n)$ независимы, V_n зависит только от первой последовательности, а $T_n = 2k$ и $S_n = j$ тогда и только тогда, когда элементы второй последовательности не равны 0 и последний равен j. Из (11.18) и (11.19) заключаем, что

$$\mathsf{P}\{T_n = 2k, V_n = 2i, S_n = j\} = \\ = p_{2k}(0) \frac{1}{k+1} \frac{j}{n-2k} p_{n-2k}(j), \tag{11.20}$$

*) У Феллера (1967а, стр. 187) имеется локальная предельная теорема для биномиального распределения, из которой вытекает (11.17), так как величина $\frac{1}{2}(S_m + m)$ биномиально распределена.

**) См. Феллер (1967а, стр. 85). Это следует также из (11.9).

***) См. Феллер (1967а, стр. 87).

если

$$0 \leqslant 2i \leqslant 2k < n, \quad j > 0. \tag{11.21}$$

И левая, и правая части равенства (11.20) обращаются в 0, когда n и j противоположны по четности. Если j отрицательно, то (11.20) выполняется с $|j|$ вместо j в правой части.

Мы применим теорему 7.8 к решетке, образуемой точками $(2k/n,\ 2i/n,\ j/\sqrt{n})$, где j и n имеют одну и ту же четность. Допустим, что k, i и j стремятся к бесконечности вместе с n таким образом, что

$$\frac{2k}{n} \to t, \quad \frac{2i}{n} \to v, \quad \frac{j}{\sqrt{n}} \to x,$$

где $0 < v < t < 1$ и $x > 0$. Тогда (11.21) выполняется при больших n и из (11.20) и (11.17) следует, что

$$\left(\frac{2}{n} \cdot \frac{2}{n} \cdot \frac{2}{\sqrt{n}}\right)^{-1} \mathsf{P}\{T_n = 2k,\ V_n = 2i,\ S_n = j\} \to g(t, x),$$

где

$$g(t, x) = \frac{1}{2\pi} \frac{|x|}{[t(1-t)]^{3/2}} e^{-\frac{1}{2} x^2/(1-t)}, \qquad 0 < t < 1. \tag{11.22}$$

Тот же самый результат имеет место в силу симметрии для отрицательных x. Поэтому случайный вектор (в схеме случайного блуждания)

$$\left(\frac{1}{n} T_n,\ \frac{1}{n} V_n,\ \frac{1}{\sqrt{n}} S_n\right) \tag{11.23}$$

имеет предельное распределение в R^3, определяемое плотностью

$$f(t, v, x) = \begin{cases} g(t, x) & \text{при } 0 < v < t < 1, \\ 0 & \text{в остальных случаях.} \end{cases} \tag{11.24}$$

В силу соотношения (11.15) эту плотность имеет случайный вектор (T, V, W_1). Как следует из формулы (11.16), распределение вектора (T, U, V, W_1) можно теперь тоже явно выписать.

Из (11.24) вытекает, что условное распределение V при фиксированных T и W_1 равномерно на $[0, T]$; это соответствует (11.19). В силу (11.16), если $T = t$ и $W_1 = x$, то величина U равномерно распределена на $[1-t, 1]$ при $x>0$ и на $[0, t]$ при $x<0$. Учитывая возможные значения t и x с помощью формулы (11.24), мы находим выражение для безусловной плотности U:

$$\iint\limits_{\substack{x>0\\ 1-u<t}} g(t, x)\,dt\,dx + \iint\limits_{\substack{x<0\\ u<t}} g(t, x)\,dt\,dx. \tag{11.25}$$

Интеграл от $g(t, x)$ по области $x>0$ равен $1/[2\pi t^{3/2}(1-t)^{1/2}]$, а это есть производная по t от $-((1-t)/t)^{1/2}/\pi$; следовательно, выражение (11.25) равно $1/[\pi u^{1/2}(1-u)^{1/2}]$. Поэтому

$$\mathsf{P}\{U\leqslant u\}=\frac{1}{\pi}\int\limits_0^u \frac{ds}{\sqrt{s(1-s)}}=\frac{2}{\pi}\arcsin\sqrt{u}\,, \quad 0<u<1. \tag{11.26}$$

Это — распределение арксинуса Леви. Похожая выкладка показывает, что T также подчиняется закону распределения арксинуса:

$$P\{T\leqslant t\}=\frac{2}{\pi}\arcsin\sqrt{t}\,, \quad 0<t<1. \tag{11.27}$$

Соединим теперь (11.14) с только что полученными результатами для вывода предельной теоремы в общем случае Линдеберга — Леви. Условимся говорить, что пересечение нулевого уровня имеет место в момент i. если происходит событие

$$E_i=\{S_i=0\}\cup\{S_{i-1}>0>S_i\}\cup\{S_{i-1}<0<S_i\} \tag{11.28}$$

(что в схеме случайного блуждания означает $S_i=0$). Пусть T'_n есть максимальное из значений i, $1\leqslant i\leqslant n$, которые являются моментами пересечения нулевого уровня; пусть U'_n — число тех i, $1\leqslant i\leqslant n$, для которых $S_i>0$, и пусть V'_n — число тех i, $1\leqslant i\leqslant T'_n$, для которых $S_i>0$. Мы докажем, что

$$\left(\frac{1}{n}T'_n,\ \frac{1}{n}U'_n,\ \frac{1}{n}V'_n,\ \frac{1}{\sigma\sqrt{n}}S_n\right)\xrightarrow{\mathscr{D}}(T, U, V, W_1), \tag{11.29}$$

установив, что величина, стоящая в левой части, аппроксимирует левую часть в соотношении (11.14).

Очевидно, что T'_n/n отличается не более чем на $1/n$ от $h_1(X_n)$. Если γ_n есть число тех i, $1 \leqslant i \leqslant n$, для которых событие E_i происходит (число пересечений нулевого уровня), то U'_n/n и V'_n/n отличаются не более чем на γ_n/n от $h_2(X_n)$ и $h_3(X_n)$, соответственно. Поэтому (11.29) будет следовать из (11.14) и теоремы 4.1, если доказать, что $\gamma_n/n \xrightarrow{P} 0$, а для этого достаточно показать, что

$$\mathsf{E}\left\{\frac{\gamma_n}{n}\right\} = \frac{1}{n}\sum_{i=1}^{n}\mathsf{P}(E_i) \to 0. \tag{11.30}$$

Но для любого положительного ε

$$\mathsf{P}(E_i) \leqslant \mathsf{P}\{|\xi_i| \geqslant \varepsilon\sqrt{i}\} + \mathsf{P}\{|S_{i-1}| \leqslant \varepsilon\sqrt{i}\}$$

и, следовательно, по центральной предельной теореме $\mathsf{P}(E_i) \to 0$. Соотношение (11.30) есть теперь следствие теоремы о средних арифметических значениях сходящихся последовательностей.

Из (11.29) мы можем, например, сделать вывод о том, что предельным распределением U'_n/n и T'_n/n является арксинус-распределение.

Броуновский мост. Броуновский мост W° ведет себя как броуновская траектория W, подчиняющаяся условию $W_1 = 0$. Это замечание может быть использовано для вывода распределений, связанных с W°, при надлежащем переходе к пределу с учетом того, что событие $\{W_1 = 0\}$ имеет вероятность 0.

Пусть P_ε — вероятностная мера на C, определенная равенством

$$P_\varepsilon(A) = \mathsf{P}\{W \in A \mid 0 \leqslant W_1 \leqslant \varepsilon\}, \qquad A \in \mathscr{C}.$$

Докажем, что при ε, стремящемся к 0 *),

$$P_\varepsilon \Rightarrow W^\circ. \tag{11.31}$$

Рассмотрим случайную функцию W, определенную на некотором вероятностном пространстве, и случайную

*) Мы можем предположить, что ε стремится к 0 непрерывно, и воспользоваться определением (2.4). Мы можем также устремить ε к 0 и по некоторой фиксированной подпоследовательности (например, последовательности величин, обратных целым числам).

функцию W°, определенную на том же самом вероятностном пространстве соотношением $\overset{\circ}{W}_t = W_t - tW_1$ (см. конструкцию в § 9). Если мы докажем, что

$$\limsup_{\varepsilon\to 0} \mathsf{P}\{W \in F \mid 0 \leqslant W_1 \leqslant \varepsilon\} \leqslant \mathsf{P}\{W^\circ \in F\} \qquad (11.32)$$

для каждого замкнутого множества F из C, то (11.31) будет следовать из теоремы 2.1.

Для произвольных $t_1, \ldots, t_k$ величина W_1 не зависит от вектора $\left(\overset{\circ}{W}_{t_1}, \ldots, \overset{\circ}{W}_{t_k}\right)$, так как она некоррелирована с каждой компонентой. Поэтому

$$\mathsf{P}\{W^\circ \in A,\ W_1 \in B\} = \mathsf{P}\{W^\circ \in A\}\,\mathsf{P}\{W_1 \in B\} \qquad (11.33)$$

для конечномерного множества A из C и B, принадлежащего $\mathscr{R}^1$. Но при фиксированном B множество тех A из $\mathscr{C}$, которые удовлетворяют (11.33), является монотонным классом и, следовательно *), совпадает с $\mathscr{C}$. Таким образом, (11.33) выполняется для $A \in \mathscr{C}$ и $B \in \in \mathscr{R}^1$. В частности,

$$\mathsf{P}\{W^\circ \in A \mid 0 \leqslant W_1 \leqslant \varepsilon\} = \mathsf{P}\{W^\circ \in A\}.$$

Так как $\rho(W, W^\circ) = |W_1|$, где ρ — метрика в C, то $|W_1| \leqslant \delta$, и из того, что $W \in F$, следует, что $W^\circ \in F_\delta = \{x\colon \rho(x, F) \leqslant \delta\}$. Поэтому, если $\varepsilon < \delta$, то

$$\mathsf{P}\{W \in F \mid 0 \leqslant W_1 \leqslant \varepsilon\} \leqslant \mathsf{P}\{W^\circ \in F_\delta \mid 0 \leqslant W_1 \leqslant \varepsilon\} = \\ = \mathsf{P}\{W^\circ \in F_\delta\}.$$

Верхний предел в (11.32) не превосходит, таким образом, величины $\mathsf{P}\{W^\circ \in F_\delta\}$, которая, убывая, стремится к $\mathsf{P}\{W^\circ \in F\}$ при $\delta \downarrow 0$, если F замкнуто. Тем самым (11.32), а вместе с ним и (11.31) доказаны **).

Допустим теперь, что h есть измеримое отображение C в R^k и что W° принадлежит множеству D_h разрывов h с вероятностью 0. Из (11.31) и теоремы 5.1 следует, что равенство

$$\mathsf{P}\{h(W^\circ) \leqslant \alpha\} = \lim_{\varepsilon\to 0} \mathsf{P}\{h(W) \leqslant \alpha \mid 0 \leqslant W_1 \leqslant \varepsilon\} \qquad (11.34)$$

---

*) См. Халмош (1953, стр. 32).

**) Эта часть рассуждений просто повторяет доказательство теоремы 4.1.

выполняется для любого α, при котором левая часть непрерывна (как функция α, пробегающего R^k). При помощи (11.34) мы можем найти явные формулы для некоторых распределений, связанных с W°. Иногда более удобна другая форма соотношения (11.34):

$$\mathsf{P}\{h(W^\circ)\leqslant \alpha\}=\lim_{\varepsilon\to 0}\mathsf{P}\{h(W)\leqslant \alpha\,|-\varepsilon\leqslant W_1\leqslant 0\}. \qquad (11.35)$$

Это равенство устанавливается тем же путем, что и (11.34) (вместо $[0,\varepsilon]$ мы могли бы взять любое подмножество $[-\varepsilon,\varepsilon]$ положительной меры).

Пусть

$$m^\circ=\inf_t W_t^\circ, \qquad M^\circ=\sup_t W_t^\circ.$$

Допустим, что $a<0<b$, $0<\varepsilon<b$; из (11.10) при $c=b-a$ имеем

$$\mathsf{P}\{a<m\leqslant M<b,\ 0<W_1<\varepsilon\}=$$
$$=\sum_{k=-\infty}^{\infty}\mathsf{P}\{2kc<N<2kc+\varepsilon\}-$$
$$-\sum_{k=-\infty}^{\infty}\mathsf{P}\{2kc+2b-\varepsilon<N<2kc+2b\}. \qquad (11.36)$$

Так как

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathsf{P}\{x<N<x+\varepsilon\}=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}, \qquad (11.37)$$

то если мы перейдем к пределу под знаками сумм в (11.36), из (11.34) будет следовать, что

$$\mathsf{P}\{a<m^\circ\leqslant M^\circ\leqslant b\}=$$
$$=\sum_{k=-\infty}^{\infty}e^{-2(kc)^2}-\sum_{k=-\infty}^{\infty}e^{-2(b+kc)^2}. \qquad (11.38)$$

Поскольку ряд сходится равномерно по ε, переход к пределу под знаком суммы законен. Таким образом, мы нашли распределение (m°, M°). Если положить здесь $-a=b$, то получим

$$\mathsf{P}\{\sup_t|W_t^\circ|\leqslant b\}=1+2\sum_{k=1}^{\infty}(-1)^k e^{-2k^2b^2}, \quad b>0. \qquad (11.39)$$

С помощью точно такого же рассуждения, примененного к (11.11), получаем

$$\mathsf{P}\{M^\circ < b\} = 1 - e^{-2b^2}, \qquad b > 0. \tag{11.40}$$

Пусть U° — мера Лебега множества тех точек в [0, 1], для которых $W_t^\circ > 0$. Мы покажем, что величина U° равномерно распределена [0, 1], установив, что

$$\lim_{\varepsilon\to 0} \mathsf{P}\{U \leqslant \alpha \mid -\varepsilon \leqslant W_1 \leqslant 0\} = \alpha, \qquad 0 < \alpha < 1. \tag{11.41}$$

В силу (11.16) эта условная вероятность равна

$$\mathsf{P}\{V \leqslant \alpha \mid -\varepsilon \leqslant W_1 \leqslant 0\}.$$

Используя формулу для плотности (11.24), мы вывели, что распределение V при данных T и W_1 равномерно на $(0, T)$. Другими словами, если $L = V/T$, то величина L равномерно распределена на $(0, 1)$ и не зависит от (T, W_1). Поэтому условная вероятность в (11.41) равна

$$\mathsf{P}\{TL \leqslant \alpha \mid -\varepsilon \leqslant W_1 \leqslant 0\} = \int_0^1 \mathsf{P}\left\{T \leqslant \frac{\alpha}{s} \,\middle|\, -\varepsilon < W_1 \leqslant 0\right\} ds,$$

и (11.41) будет следствием теоремы об ограниченной сходимости, если мы докажем интуитивно очевидное соотношение

$$\mathsf{P}\{T \leqslant \theta \mid -\varepsilon \leqslant W_1 \leqslant 0\} \to 0, \qquad \theta < 1.$$

Но это соотношение является следствием (11.37) и формы плотности (11.24). Таким образом,

$$\mathsf{P}\{U^0 \leqslant \alpha\} = \alpha, \qquad 0 < \alpha < 1. \tag{11.42}$$

Примечание. Дальнейшие примеры вычисления распределений функций от броуновских траекторий см. в работах Ито, Маккин (1968), Карлин (1971, гл. 10), Эрдеш, Кац (1946 и 1947), Марк (1949), Дарлинг, Эрдеш (1956); см. также § 16 этой книги.

Задачи

1. Покажите при помощи принципа отражения в схеме случайного блуждания, что

$$\mathsf{P}\{M_n \geqslant b,\ S_n = v\} = \begin{cases} p_n(v) & \text{при } v \geqslant b, \\ p_n(2b - v) & \text{при } v \leqslant b, \end{cases} \tag{11.43}$$

где $b \geqslant 0$. Выведите отсюда (11.9).

2. Для неотрицательных целых чисел c_i пусть $\pi(c_1, \ldots, c_k; v)$ обозначает вероятность того, что n-шаговое случайное блуждание (n фиксировано) пересекает уровень c_1 (один или больше число раз), затем пересекает $-c_2$, затем пересекает $c_3, \ldots$, затем пересекает $(-1)^{k+1}c_k$ и заканчивается в v. Используя (11.43) и индукцию по k, покажите, что

$$\pi(c_1, \ldots c_k; v) = \begin{cases} p_n\left(2c_1 + \ldots + 2c_{k-1} + (-1)^{k+1}v\right) & \text{при} \quad (-1)^{k+1} v \geqslant c_k, \\ p_n\left(2c_1 + \ldots + 2c_k - (-1)^{k+1} v\right) & \text{при} \quad (-1)^{k+1} v \leqslant c_k. \end{cases}$$

[Отразите относительно $(-1)^k c_{k-1}$ часть траектории, находящуюся справа от первой точки прохождения через этот уровень, следующий за последовательными прохождениями уровней $c_1, -c_2, \ldots$ $\ldots (-1)^{k-2}c_{k-2}$.] Выведите (11.5), доказав для этого равенство

$$p_n(a, b, v) = p_n(v) - \pi(b; v) + \pi(b, a; v) - \\ - \pi(b, a, b; v) + \ldots - \pi(a; v) + \pi(a, b; v) - \pi(a, b, a; v) + \ldots$$

3. Для x из C пусть $h(x)$ есть наименьшее t, для которого $x(t) = \sup_s x(s)$. Покажите, что h измерима и непрерывна на множестве W-меры 1. Пусть τ_n есть наименьшее k, для которого $S_k = \max\limits_{i \leqslant n} S_i$. Докажите, что

$$\mathsf{P}\left\{\frac{\tau_n}{n} \leqslant \alpha\right\} \to \frac{2}{\pi} \arcsin \sqrt{\alpha} \qquad 0 < \alpha < 1,$$

в случае Линдеберга — Леви. [Для случайного блуждания см. Феллер (1967а, стр. 95).]

4. Выведите совместное предельное распределение максимума и минимума величин $S_i - in^{-1}S_n$, $0 < i < n$, в случае Линдеберга — Леви. [Рассмотрите $Y_n(t) = X_n(t) - tX_n(1)$, где X_n определено в (10.1).]

§ 12. Флуктуации частичных сумм

В §§ 9 и 10 мы установили плотность последовательности случайных функций путем нахождения оценок для распределения максимума некоторых частичных сумм. Здесь мы получим подобные оценки при весьма общих условиях; в результате мы придем к практическому критерию плотности, который будет использоваться на протяжении всей оставшейся части книги.

Максимум. Рассмотрим случайные величины $\xi_1, \ldots$ $\ldots, \xi_m$, которые не обязаны быть независимыми или одинаково распределенными. Пусть $S_k = \xi_1 + \ldots + \xi_k$ ($S_0 = 0$), и положим

$$M_m = \max_{0 \leqslant k \leqslant m} |S_k|. \tag{12.1}$$

Мы получим верхние оценки для вероятности $\mathsf{P}\{M_m \leqslant \lambda\}$, используя некий косвенный прием.

Если

$$M'_m = \max_{0 \leqslant k \leqslant m} \min\{|S_k|, |S_m - S_k|\}, \tag{12.2}$$

то

$$M'_m \leqslant M_m. \tag{12.3}$$

Если $\xi_1, \ldots, \xi_{k-1}, \xi_{k+1}, \ldots, \xi_m$ тождественно равны нулю, то $M'_m = 0$ и $M_m = |\xi_k|$. Таким образом, невозможно неравенство, противоположное неравенству (12.3).

С другой стороны, мы все же имеем

$$|S_k| \leqslant \min\{|S_m| + |S_k|, |S_m| + |S_m - S_k|\} = \\ = |S_m| + \min\{|S_k|, |S_m - S_k|\},$$

так что

$$M_m \leqslant M'_m + |S_m| \tag{12.4}$$

(здесь имеет место равенство, если $S_m = 0$). Поэтому

$$\mathsf{P}\{M_m \geqslant \lambda\} \leqslant \mathsf{P}\{M'_m \geqslant \lambda/2\} + \mathsf{P}\{|S_m| \geqslant \lambda/2\}. \tag{12.5}$$

Если мы оценим в отдельности каждое слагаемое в правой части (12.5), то получим оценку для выражения слева.

Оценивая $\mathsf{P}\{M_m \geqslant \lambda\}$ посредством (12.5), мы можем проиграть в точности по сравнению с какими-либо прямыми вычислениями. С другой стороны, так как правая часть (12.5) не превосходит $2\mathsf{P}\{M_m \geqslant \lambda/2\}$, то эта потеря не должна быть большой: добавочный множитель 2 несуществен для наших целей, а полученные оценки будут убывать с возрастанием λ достаточно медленно, так что переход от λ к $\lambda/2$ тоже не будет существенным.

Существует другой путь вывода оценок для распределения M_n из оценок для распределения M'_m. Мы покажем, что

$$M_m \leqslant 3M'_m + \max_{1 \leqslant l \leqslant m} |\xi_l|, \tag{12.6}$$

откуда будет следовать, что

$$\mathsf{P}\{M_m \geqslant \lambda\} \leqslant \mathsf{P}\{M'_m \geqslant \lambda/4\} + \mathsf{P}\{\max_{1 \leqslant l \leqslant m} |\xi_l| \geqslant \lambda/4\}. \qquad (12.7)$$

Опять, оценивая слагаемые справа, придем к оценке вероятности слева. Так как правая часть (12.7) не превосходит $2\mathsf{P}\{M_m \geqslant \lambda/8\}$, то использование (12.7) не приводит к значительной потере точности.

Чтобы доказать (12.6), рассмотрим множество I, состоящее из таких k, $0 \leqslant k \leqslant m$, для которых $|S_k| \leqslant \leqslant |S_m - S_k|$; разумеется, $0 \in I$. Если $S_m = 0$, то $M_m = M'_m$, так что (12.6) выполняется. Если $S_m \neq 0$, то $m \notin I$ и, следовательно, существует k, $0 < k \leqslant m$, для которого $k - 1 \in I$ и $k \notin I$:

$$|S_{k-1}| \leqslant |S_m - S_{k-1}|, \quad |S_{k-1}| \leqslant M'_m, \quad |S_m - S_k| < |S_k|$$

и

$$|S_m - S_k| \leqslant M'_m.$$

Для такого k мы имеем

$$|S_m| \leqslant |S_{k-1}| + |\xi_k| + |S_m - S_k| \leqslant 2M'_m + |\xi_k|, \qquad (12.8)$$

откуда, если учесть (12.4), следует (12.6). (Если $m = 5$ и $\xi_1 = \xi_2 = \xi_3 = \xi_4 = -\xi_5$, то в (12.6) мы имеем равенство.)

Смешанные моменты. Неравенства (12.5) и (12.7) указывают на необходимость оценить величину $\mathsf{P}\{M'_m \geqslant \lambda\}$. Предположим, что существуют неотрицательные числа $u_1, \ldots, u_m$ такие, что

$$\mathsf{E}\{|S_j - S_i|^\gamma \cdot |S_k - S_j|^\gamma\} \leqslant \Big(\sum_{i < l \leqslant j} u_l\Big)^\alpha \Big(\sum_{j < l \leqslant k} u_l\Big)^\alpha,$$
$$0 \leqslant i \leqslant j \leqslant k \leqslant m, \qquad (12.9)$$

где γ и α положительны. Это неравенство выполняется, например, с $\gamma = 2$ и $\alpha = 1$, если ξ_i независимы, $\mathsf{E}\{\xi_i\} = 0$, и если в качестве u_i берется дисперсия величины ξ_i.

Так как $xy \leqslant (x + y)^2$ для неотрицательных x и y, то (12.9) влечет неравенство

$$\mathsf{E}\{|S_j - S_i|^\gamma \cdot |S_k - S_j|^\gamma\} \leqslant \Big(\sum_{i < l \leqslant k} u_l\Big)^{2\alpha},$$
$$0 \leqslant i \leqslant j \leqslant k \leqslant m. \qquad (12.10)$$

Кроме того, если $|S_j - S_i| \geqslant \lambda$ и $|S_k - S_j| \geqslant \lambda$, где λ — положительное число, то $|S_j - S_i|^\gamma \cdot |S_k - S_j|^\gamma \geqslant \lambda^{2\gamma}$, так что из (12.10) следует, что

$$\mathsf{P}\{|S_j - S_i| \geqslant \lambda, |S_k - S_j| \geqslant \lambda\} \leqslant \frac{1}{\lambda^{2\gamma}} \Big(\sum_{i<l\leqslant k} u_l\Big)^{2\alpha},$$

$$0 \leqslant i \leqslant j \leqslant k \leqslant m. \quad (12.11)$$

(Неравенство тривиально, если $i = j$ или $j = k$.) Отметим, что (12.9) влечет (12.10), а (12.10) влечет (12.11) также в случае $\gamma = 0$ при условии, что мы положим $|x|^0$ равным 1, если $x \neq 0$, и равным 0, если $x = 0$.

Теорема 12.1. *Если $\gamma \geqslant 0$ и $\alpha > 1/2$ и если* (12.11) *выполняется для всех положительных λ, то для всех положительных λ*

$$\mathsf{P}\{M'_m \geqslant \lambda\} \leqslant \frac{K_{\gamma,\alpha}}{\lambda^{2\gamma}} (u_1 + \ldots + u_m)^{2\alpha}, \quad (12.12)$$

где $K_{\gamma,\alpha}$ — константа, зависящая только от γ и α.

Хотя конкретное значение константы не будет иметь для нас значения, отметим, что в качестве $K_{\gamma,\alpha}$ можно выбрать величину

$$K_{\gamma,\alpha} = \left[\frac{1}{2^{1/(2\gamma+1)}} - \left(\frac{1}{2^{1/(2\gamma+1)}}\right)^{2\alpha}\right]^{-(2\gamma+1)}. \quad (12.13)$$

Теорема 12.1 имеет сугубо теоретическое значение, так как постоянная $K_{2,1}$, например, равна 55 021 с точностью до целого числа.

Приложения. Прежде чем доказывать теорему, разберем несколько ее следствий. Если ξ_i независимы и $\mathsf{E}\{\xi_i\} = 0$, $\mathsf{E}\{\xi_i^2\} = \sigma_i^2$, то (12.9) выполняется с $\gamma = 2$ и $\alpha = 1$ при условии, что $u_i = \sigma_i^2$. Далее, $\sigma_1^2 + \ldots + \sigma_m^2 = s_m^2$ есть дисперсия S_m и (12.12) влечет неравенство

$$\mathsf{P}\{M'_m \geqslant \lambda\} \leqslant \frac{K s_m^4}{\lambda^4} \quad (12.14)$$

с $K = K_{2,1}$. Заменяя λ на λs_m и используя (12.5) и (12.7), мы получаем

$$\mathsf{P}\{M_m \geqslant \lambda s_m\} \leqslant \frac{2^4 K}{\lambda^4} + \mathsf{P}\left\{|S_m| \geqslant \frac{1}{2}\lambda s_m\right\} \quad (12.15)$$

и

$$\mathsf{P}\{M_m \geqslant \lambda s_m\} \leqslant \frac{4^4 K}{\lambda^4} + \mathsf{P}\left\{\max_{1\leqslant i\leqslant m} |\xi_i| \geqslant \frac{1}{4}\lambda s_m\right\}. \quad (12.16)$$

Согласно (10.7)

$$\mathsf{P}\{M_m \geqslant \lambda s_m\} \leqslant 2\mathsf{P}\{|S_m| \geqslant (\lambda - \sqrt{2})s_m\}. \qquad (12.17)$$

Множитель 2 справа не существен, и если λ велико, то $\lambda/2$ и $\lambda - \sqrt{2}$ суть величины одного порядка. Таким образом, в случае независимости ξ_i (12.15) не имеет преимущества перед (12.17), а (12.16) имеет, так как

$$\mathsf{P}\left\{\max_{1 \leqslant i \leqslant m} |\xi_i| \geqslant \frac{1}{4}\lambda s_m\right\} \leqslant \sum_{i=1}^{m} \mathsf{P}\left\{|\xi_i| \geqslant \frac{1}{4}\lambda s_m\right\} \leqslant$$
$$\leqslant \frac{4^2}{\lambda^2}\frac{1}{s_m^2}\sum_{i=1}^{m}\int\limits_{\left\{|\xi_i| \geqslant \frac{1}{4}\lambda s_m\right\}} \xi_i^2\, d\mathsf{P}, \qquad (12.18)$$

и поэтому (12.16) влечет

$$\mathsf{P}\{M_m \geqslant \lambda s_m\} \leqslant \frac{4^4 K}{\lambda^4} + \frac{4^2}{\lambda^2}\frac{1}{s_m^2}\sum_{i=1}^{m}\int\limits_{\left\{|\xi_i| \geqslant \frac{1}{4}\lambda s_m\right\}} \xi_i^2\, d\mathsf{P}. \quad (12.19)$$

Сумма в правой части входит в условие Линдеберга (7.3).

Можно получить и другие результаты. Например, из (12.14) следует (см. (3) на стр. 305), что

$$\int\limits_{\{(M'_m)^2 \geqslant a s_m^2\}} \left(\frac{M'}{s_m}\right)^2 d\mathsf{P} \leqslant 2\frac{K}{a}.$$

Так как ξ_i независимы, мы имеем (суммирование происходит в пределах $1 \leqslant i, k \leqslant m$)

$$\int\limits_{\{\max_k \xi_k^2 \geqslant a s_m^2\}} \max_i \frac{\xi_i^2}{s_m^2}\, d\mathsf{P} \leqslant \sum_{k,\,i}\int\limits_{\{\xi_k^2 \geqslant a s_m^2\}} \frac{\xi_i^2}{s_m^2}\, d\mathsf{P} =$$
$$= \sum_i \int\limits_{\{\xi_i^2 \geqslant a s_m^2\}} \frac{\xi_i^2}{s_m^2}\, d\mathsf{P} + \sum_{k \neq i} \mathsf{P}\{\xi_k^2 \geqslant a s_m^2\}\frac{\sigma_i^2}{s_m^2} \leqslant$$
$$\leqslant \left(1 + \frac{1}{a}\right)\sum_i \frac{1}{s_m^2}\int\limits_{\{\xi_i^2 \geqslant a s_m^2\}} \xi_i^2\, d\mathsf{P}.$$

Для неотрицательных U и V имеем

$$\int\limits_{\{U+V\geqslant a\}} (U+V)\,d\mathsf{P} \leqslant 2 \int\limits_{\left\{U\geqslant \frac{1}{2}a\right\}} U\,d\mathsf{P} + 2 \int\limits_{\left\{V\geqslant \frac{1}{2}a\right\}} V\,d\mathsf{P}.$$

Неравенство

$$\int\limits_{\left\{M_m^2\geqslant as_m^2\right\}} \left(\frac{M_m^2}{s_m^2}\right) d\mathsf{P} \leqslant K'\left[\frac{1}{a} + \sum_{i=1}^{m} \frac{1}{s_m^2} \int\limits_{\left\{|\xi_i|\geqslant \frac{1}{4}as_m\right\}} \xi_i^2\,d\mathsf{P}\right] \tag{12.20}$$

(где K' — некоторая абсолютная постоянная) теперь является следствием (12.6) и того факта, что сумма справа здесь не превосходит 1. Неравенство сохраняется, если M_m^2 заменить на S_m^2. Это доказывает, в частности, равномерную интегрируемость квадратов нормированных сумм по строкам в схеме серий, удовлетворяющей условию Линдеберга.

Допустим теперь, что ξ_i независимы и одинаково распределены с $\mathsf{E}\{\xi_i\}=0$ и $\mathsf{E}\{\xi_i^2\}=\sigma^2$. Тогда (12.15) влечет

$$\mathsf{P}\{M_m\geqslant\lambda\sigma\sqrt{m}\} \leqslant \frac{2^4K}{\lambda^4} + \mathsf{P}\left\{|S_m|\geqslant\frac{1}{2}\lambda\sigma\sqrt{m}\right\}; \tag{12.21}$$

(12.16) влечет

$$\mathsf{P}\{M_m\geqslant\lambda\sigma\sqrt{m}\} \leqslant \frac{4^4K}{\lambda^4} + \mathsf{P}\left\{\max_{1\leqslant i\leqslant m}|\xi_i|\geqslant\frac{1}{4}\lambda\sigma\sqrt{m}\right\} \tag{12.22}$$

и (12.19) влечет

$$\mathsf{P}\{M_m\geqslant\lambda\sigma\sqrt{m}\} \leqslant \frac{4^4K}{\lambda^4} + \frac{4}{\lambda^2\sigma^2} \int\limits_{\left\{|\xi_1|\geqslant\frac{1}{4}\lambda\sigma\sqrt{m}\right\}} \xi_1^2\,d\mathsf{P}. \tag{12.23}$$

Так как интеграл в последнем неравенстве стремится к 0 при $\lambda\to\infty$, то для достаточно больших λ имеем

$$\mathsf{P}\{M_m\geqslant\lambda\sigma\sqrt{m}\} \leqslant \frac{\varepsilon}{\lambda^2}, \quad m=1,\ 2,\ \ldots$$

Таким образом, нами получено другое доказательство плотности случайных функций, используемых в теореме

Донскера (см. (10.10)). Отметим, что новое доказательство не использует центральную предельную теорему.

Хотя в приведенных приложениях случайные величины ξ_i независимы, в самой теореме 12.1 независимость не предполагается. Это существенное преимущество: большинство последующих приложений будут связаны с зависимыми последовательностями.

Доказательство теоремы 12.1. Предположение теоремы 12.1 состоит в том, что

$$\mathsf{P}\{|S_j - S_i| \geqslant \lambda,\ |S_k - S_j| \geqslant \lambda\} \leqslant \leqslant \frac{1}{\lambda^{2\gamma}} \Big(\sum_{i<l\leqslant k} u_l \Big)^{2\alpha}, \quad 0 \leqslant i \leqslant i \leqslant k \leqslant m, \tag{12.24}$$

для всех положительных λ; нам нужно найти постоянную K, зависящую только от γ и α, для которой

$$\mathsf{P}\{M'_m \geqslant \lambda\} \leqslant \frac{K}{\lambda^{2\gamma}} (u_1 + \ldots + u_m)^{2\alpha}. \tag{12.25}$$

Обозначим

$$\delta = \frac{1}{2\gamma + 1}, \tag{12.26}$$

так что $0 < \delta \leqslant 1$. Для больших K имеем

$$2^{\delta} \left[\frac{1}{2^{2\alpha\delta}} + \frac{1}{K^{\delta}} \right] \leqslant 1, \tag{12.27}$$

поскольку выражение в левой части приближается к $1/2^{(2\alpha-1)\delta}$ при $K \to \infty$ и ($2\alpha > 1$, $\delta > 0$) этот предел меньше 1. Мы докажем, что (12.25) выполняется, если K удовлетворяет (12.27) и если

$$K \geqslant 1. \tag{12.28}$$

(В действительности (12.27) влечет (12.28). Наименьшее K, удовлетворяющее (12.27), дается формулой (12.13).)

Доказательство ведется индукцией по m. При $m = 1$ результат тривиален; рассмотрим случай $m = 2$. Так как $M' = \min\{|S_1|, |S_2 - S_1|\}$, то из (12.24) и (12.28) следует, что

$$\mathsf{P}\{M'_2 \geqslant \lambda\} \leqslant \frac{1}{\lambda^{2\gamma}} (u_1 + u_2)^{2\alpha} \leqslant \frac{K}{\lambda^{2\gamma}} (u_1 + u_2)^{2\alpha}.$$

Допустим теперь (предположение индукции), что нужный нам результат выполняется для любого целого числа, меньшего чем m. Докажем, что он выполняется для m. Положим $u = u_1 + \ldots + u_m$; мы можем считать, что $u > 0$. Существует целое h, $1 \leqslant h \leqslant m$, такое, что

$$\frac{u_1 + \ldots + u_{h-1}}{u} \leqslant \frac{1}{2} \leqslant \frac{u_1 + \ldots + u_h}{u}, \tag{12.29}$$

где сумма слева равна 0, если $h = 1$.

Рассмотрим четыре величины

$$U_1 = \max_{0 \leqslant i \leqslant h-1} \min\{|S_i|,\ |S_{h-1} - S_i|\},$$

$$U_2 = \max_{h \leqslant j \leqslant m} \min\{|S_j - S_h|,\ |(S_m - S_h) - (S_j - S_h)|\} =$$

$$= \max_{h \leqslant j \leqslant m} \min\{|S_j - S_h|,\ |S_m - S_j|\},$$

$$D_1 = \min\{|S_{h-1}|,\ |S_m - S_{h-1}|\},$$

$$D_2 = \min\{|S_h|,\ |S_m - S_h|\}.$$

Так как (12.24) выполняется, если m заменено на $h - 1$, и так как $h - 1 < m$, то можно применить предположение индукции к случайным величинам $\xi_1, \ldots, \xi_{h-1}$ и величинам $u_1, \ldots, u_{h-1}$ и заключить, что

$$\mathsf{P}\{U_1 \geqslant \lambda\} \leqslant \frac{K}{\lambda^{2\gamma}} (u_1 + \ldots + u_{h-1})^{2\alpha} \leqslant \frac{u^{2\alpha}}{\lambda^{2\gamma}} \frac{K}{2^{2\alpha}}; \tag{12.30}$$

последнее неравенство здесь следует из (12.29). (При $h = 1$ (12.30) тривиально.)

Если индексы в (12.24) ограничить пределами $h \leqslant i \leqslant j \leqslant k \leqslant m$, то в рассмотрении будут участвовать лишь случайные величины $\xi_{h+1}, \ldots, \xi_m$. Так как $m - h < m$, то предположение индукции применимо к $\xi_{h+1}, \ldots, \xi_m$ и $u_{h+1}, \ldots, u_m$; следовательно,

$$\mathsf{P}\{U_2 \geqslant \lambda\} \leqslant \frac{K}{\lambda^{2\gamma}} (u_{h+1} + \ldots + u_m)^{2\alpha} \leqslant \frac{u^{2\alpha}}{\lambda^{2\gamma}} \frac{K}{2^{2\alpha}}; \tag{12.31}$$

последнее неравенство снова является следствием (12.29). (При $h = m$ (12.31) тривиально.)

Далее, из (12.24) получаем

$$\mathsf{P}\{D_1 \geqslant \lambda\} \leqslant \frac{1}{\lambda^{2\gamma}}(u_1 + \ldots + u_m)^{2\alpha} = \frac{u^{2\alpha}}{\lambda^{2\gamma}} \tag{12.32}$$

и

$$\mathsf{P}\{D_2 \geqslant \lambda\} \leqslant \frac{u^{2\alpha}}{\lambda^{2\gamma}}. \tag{12.33}$$

(Если $h = 1$, то (12.32) тривиально; если $h = m$, то (12.33) тривиально.)

Мы покажем, что

$$\min\{|S_i|, |S_m - S_i|\} \leqslant U_1 + D_1, \quad \text{если} \quad 0 \leqslant i \leqslant h - 1. \tag{12.34}$$

Обозначим через μ_i минимум в левой части (12.34). Если $|S_i| \leqslant U_1$, то

$$\mu_i \leqslant |S_i| \leqslant U_1 \leqslant U_1 + D_1.$$

Допустим, что $|S_{h-1} - S_i| \leqslant U_1$; если $|S_{h-1}| = D_1$, то

$$\mu_i \leqslant |S_i| \leqslant |S_{h-1} - S_i| + |S_{h-1}| \leqslant U_1 + D_1.$$

Допустим, что $|S_{h-1} - S_i| \leqslant U_1$ и $|S_m - S_{h-1}| = D_1$; тогда

$$\mu_i \leqslant |S_m - S_i| \leqslant |S_{h-1} - S_i| + |S_m - S_{h-1}| \leqslant U_1 + D_1.$$

Тем самым (12.34) доказано. Подобным же способом можно показать, что

$$\min\{|S_j|, |S_m - S_j|\} \leqslant U_2 + D_2, \quad \text{если} \quad h \leqslant j \leqslant m. \tag{12.35}$$

Согласно (12.34) и (12.35)

$$M'_m \leqslant \max\{U_1 + D_1,\ U_2 + D_2\}, \tag{12.36}$$

и следовательно,

$$\mathsf{P}\{M'_m \geqslant \lambda\} \leqslant \mathsf{P}\{U_1 + D_1 \geqslant \lambda\} + \mathsf{P}\{U_2 + D_2 \geqslant \lambda\}. \tag{12.37}$$

Если λ_0 и λ_1 положительны и $\lambda_0 + \lambda_1 = \lambda$, то в силу (12.30) и (12.32)

$$\mathsf{P}\{U_1 + D_1 \geqslant \lambda\} \leqslant \mathsf{P}\{U_1 \geqslant \lambda_0\} + \mathsf{P}\{D_1 \geqslant \lambda_1\} \leqslant$$
$$\leqslant \frac{u^{2\alpha}}{\lambda_0^{2\gamma}} \frac{K}{2^{2\alpha}} + \frac{u^{2\alpha}}{\lambda_1^{2\gamma}}. \tag{12.38}$$

Вычисление показывает, что если C_0, C_1 и λ положительны, то

$$\min_{\substack{\lambda_0,\,\lambda_1>0\\ \lambda_0+\lambda_1=\lambda}}\left[\frac{C_0}{\lambda_0^{2\gamma}}+\frac{C_1}{\lambda_1^{2\gamma}}\right]=\frac{1}{\lambda^{2\gamma}}\left[C_0^{\delta}+C_1^{\delta}\right]^{1/\delta}, \tag{12.39}$$

где δ определено в (12.26). Минимизируя самую правую часть неравенства (12.38), мы приходим к соотношению

$$\mathsf{P}\{U_1+D_1\geqslant\lambda\}\leqslant\frac{u^{2\alpha}}{\lambda^{2\gamma}}\left[\left(\frac{K}{2^{2\alpha}}\right)^{\delta}+1\right]^{1/\delta}. \tag{12.40}$$

Точно такое же неравенство выполняется для U_2+D_2 (нужно использовать (12.31) и (12.33)). В силу (12.37) имеем поэтому

$$\mathsf{P}\{M'_m\geqslant\lambda\}\leqslant\frac{u^{2\alpha}}{\lambda^{2\gamma}}2\left[\left(\frac{K}{2^{2\alpha}}\right)^{\delta}+1\right]^{1/\delta}. \tag{12.41}$$

В соответствии с условием (12.27) на K, правая часть здесь не превосходит $Ku^{2\alpha}/\lambda^{2\gamma}$. Этим завершается индукция; теорема 12.1 доказана.

Моменты *). Заменим теперь (12.9) другим предположением:

$$\mathsf{E}\{|S_j-S_i|^{\gamma}\}\leqslant\Big(\sum_{i<l\leqslant j}u_l\Big)^{\alpha},\qquad 0\leqslant i\leqslant j\leqslant m. \tag{12.42}$$

Из этого предположения следует, что если λ — положительное число, то

$$\mathsf{P}\{|S_j-S_i|\geqslant\lambda\}\leqslant\frac{1}{\lambda^{\gamma}}\Big(\sum_{i<l\leqslant j}u_l\Big)^{\alpha},\qquad 0\leqslant i\leqslant j\leqslant m. \tag{12.43}$$

Вместо $\alpha>1/2$ мы сделаем более сильное допущение $\alpha>1$.

Теорема 12.2. *Если $\gamma\geqslant 0$ и $\alpha>1$ и если* (12.43) *выполняется при всех положительных λ, то при всех положительных λ*

$$\mathsf{P}\{M_m\geqslant\lambda\}\leqslant\frac{K'_{\gamma,\alpha}}{\lambda^{\gamma}}(u_1+\ldots+u_m)^{\alpha}, \tag{12.44}$$

где $K'_{\gamma,\alpha}$ зависит только от γ и α.

---

*) Все последующие результаты этого параграфа потребуются в главах 3 и 4, но не в § 13 настоящей главы.

В качестве $K'_{\gamma, \alpha}$ можно взять

$$K'_{\gamma, \alpha} = 2^{\gamma}\left(1 + K_{\frac{1}{2}\gamma, \frac{1}{2}\alpha}\right). \tag{12.45}$$

Доказательство. По неравенству Шварца $\mathrm{P}(E_1 \cap E_2) \leqslant \mathrm{P}^{1/2}(E_1)\,\mathrm{P}^{1/2}(E_2)$. Поэтому из (12.43) получаем (напомним, что $xy \leqslant (x+y)^2$)

$$\mathrm{P}\{|S_j - S_i| \geqslant \lambda,\ |S_k - S_j| \geqslant \lambda\} \leqslant$$

$$\leqslant \frac{1}{\lambda^{\gamma/2}}\left(\sum_{i<l\leqslant j} u_l\right)^{\alpha/2} \frac{1}{\lambda^{\gamma/2}}\left(\sum_{j<l\leqslant k} u_l\right)^{\alpha/2} \leqslant \frac{1}{\lambda^{\gamma}}\left(\sum_{i<l\leqslant k} u_l\right)^{\alpha}.$$

Таким образом, (12.11) выполняется с γ и α, замененными на $\frac{1}{2}\gamma$ и $\frac{1}{2}\alpha$, так что согласно теореме 12.1

$$\mathrm{P}\{M'_m \geqslant \lambda\} \leqslant \frac{K}{\lambda^{\gamma}}(u_1 + \ldots + u_m)^{\alpha}, \tag{12.46}$$

где $K = K_{\frac{1}{2}\gamma, \frac{1}{2}\alpha}$. В силу (12.43) мы имеем

$$\mathrm{P}\{|S_m| \geqslant \lambda\} \leqslant \frac{1}{\lambda^{\gamma}}(u_1 + \ldots + u_m)^{\alpha} \tag{12.47}$$

и (12.44) следует теперь из неравенства (12.5). (Отметим, что так как правые части в последних двух неравенствах имеют один и тот же порядок, то (12.44) не может быть существенно улучшено при использовании дополнительной информации о распределении $|S_m|$.)

В качестве применения теоремы 12.2 рассмотрим вновь независимые одинаково распределенные случайные величины ξ_i со средним 0 и дисперсией σ^2; на этот раз предположим существование конечных четвертых моментов $\tau^4 = \mathrm{E}\{\xi_i^4\}$. Далее,

$$\mathrm{E}\{S_r^4\} = \sum \mathrm{E}\{\xi_i \xi_j \xi_k \xi_l\},$$

где индексы меняются независимо друг от друга в пределах от 1 до r. Так как ξ_i независимы и их средние равны 0, то, если значение одного индекса в некотором слагаемом отличается от значений трех остальных, это слагаемое исчезает. Так как $\sigma^4 \leqslant \tau^4$, то

$$\mathrm{E}\{S_r^4\} = r\tau^4 + 3r(r-1)\sigma^4 \leqslant 4r^2\tau^4,$$

и следовательно, (12.42) выполняется с $\gamma = 4$, $\alpha = 2$ и $u_1 = \ldots = u_m = 2\tau^2$.

Теорема 12.2 утверждает теперь, что

$$\mathsf{P}\{M_m \geqslant \lambda\} \leqslant 4\tau^4 K \frac{m^2}{\lambda^4}, \tag{12.48}$$

где $K = K'_{4,2}$. Заменяя λ на $\lambda\sigma\sqrt{m}$, получаем

$$\mathsf{P}\{M_m \geqslant \lambda\sigma\sqrt{m}\} \leqslant 4\tau^4 K \frac{1}{\lambda^4\sigma^4}. \tag{12.49}$$

Отсюда и из теоремы 8.4 сразу же следует плотность случайных функций (10.1), входящих в утверждение теоремы Донскера. Рассуждение, использующее (12.23), является более мощным, чем настоящее, поскольку оно требует существования лишь вторых моментов.

Критерий плотности. Теорема 12.2 приводит к критерию плотности последовательности $\{X_n\}$ случайных элементов C.

Теорема 12.3. *Последовательность $\{X_n\}$ плотна, если она удовлетворяет следующим двум условиям:*

(i) *Последовательность $\{X_n(0)\}$ плотна.*

(ii) *Существуют постоянные $\gamma \geqslant 0$ и $\alpha > 1$ и неубывающая непрерывная функция F на $[0, 1]$ такие, что неравенство*

$$\mathsf{P}\{|X_n(t_2) - X_n(t_1)| \geqslant \lambda\} \leqslant \frac{1}{\lambda^\gamma} |F(t_2) - F(t_1)|^\alpha \tag{12.50}$$

выполняется при всех t_1, t_2, n и всех положительных λ.

Неравенство (12.50) является следствием условия для моментов

$$\mathsf{E}\{|X_n(t_2) - X_n(t_1)|^\gamma\} \leqslant |F(t_2) - F(t_1)|^\alpha. \tag{12.51}$$

Доказательство. В силу теоремы 8.2 достаточно по заданным ε и η найти δ, $0 < \delta < 1$, для которого

$$\mathsf{P}\{w(X_n, \delta) \geqslant 3\varepsilon\} \leqslant \eta \tag{12.52}$$

при всех n. В силу следствия к теореме 8.3 неравенство (12.52) выполнено, если δ^{-1} — целое и

$$\sum_{j<\delta^{-1}} \mathsf{P}\{\sup_{j\delta \leqslant s \leqslant (j+1)\delta} |X_n(s) - X_n(j\delta)| \geqslant \varepsilon\} \leqslant \eta. \tag{12.53}$$

Фиксируем на время n, δ и j, и для целого положительного числа m рассмотрим случайные величины

$$\xi_i = X_n\left(j\delta + \frac{i}{m}\delta\right) - X_n\left(j\delta + \frac{i-1}{m}\delta\right), \quad i = 1, 2, \ldots, m. \tag{12.54}$$

В силу (12.50) эти случайные величины удовлетворяют (12.43) с

$$u_i = F(j\delta + i\delta m^{-1}) - F(j\delta + (i-1)\delta m^{-1}).$$

Поэтому по теореме 12.2

$$\mathbf{P}\left\{\max_{0 \leqslant i \leqslant m}\left|X_n\left(j\delta + \frac{i}{m}\delta\right) - X_n(j\delta)\right| \geqslant \varepsilon\right\} \leqslant \leqslant \frac{K}{\varepsilon^\gamma}[F((j+1)\delta) - F(j\delta)]^\alpha, \tag{12.55}$$

где $K = K'_{\gamma, \alpha}$. Так как X_n принадлежит C, то, устремляя m к ∞, приходим к неравенству

$$\mathbf{P}\{\sup_{j\delta \leqslant s \leqslant (j+1)\delta} |X_n(s) - X_n(j\delta)| \geqslant \varepsilon\} \leqslant \leqslant \frac{K}{\varepsilon^\gamma}[F((j+1)\delta) - F(j\delta)]^\alpha. \tag{12.56}$$

Поэтому

$$\sum_{j < \delta^{-1}} \mathbf{P}\{\sup_{j\delta \leqslant s \leqslant (j+1)\delta} |X_n(s) - X_n(j\delta)| \geqslant \varepsilon\} \leqslant$$

$$\leqslant \frac{K}{\varepsilon^\gamma}[F(1) - F(0)]\,[\max_{j < \delta^{-1}} |F((j+1)\delta) - F(j\delta)|]^{\alpha-1}, \tag{12.57}$$

если δ^{-1} — целое. Так как функция F непрерывна и $\alpha > 1$, то мы можем сделать последнюю величину в (12.57) малой, выбирая δ равным обратной величине большого целого числа. Теорема доказана.

Идеи, положенные в основу этого доказательства, приводят к условию существования в C случайного элемента с заданными конечномерными распределениями. Для каждого набора k точек из $[0, 1]$ пусть $\mu_{t_1 \ldots t_k}$ обозначает вероятностную меру на $(R^k, \mathscr{R}^k)$. Предположим, что эти меры удовлетворяют условиям согласованности теоремы Колмогорова (в ее общей форме; см. стр. 316).

Теорема 12.4. *В пространстве C существует случайный элемент с конечномерными распределениями $\mu_{t_1 \ldots t_k}$ при условии, что эти распределения согласованы и что существуют постоянные $\gamma \geqslant 0$ и $\alpha > 1$ и неубывающая непрерывная функция F на $[0, 1]$ такие, что*

$$\mu_{t_1, t_2}\{(\beta_1, \beta_2):\ |\beta_1 - \beta_2| \geqslant \lambda\} \leqslant \frac{1}{\lambda^\gamma} |F(t_2) - F(t_1)|^\alpha \quad (12.58)$$

при всех t_1 и t_2 и всех положительных λ.

Если $\mu_{t_1 \ldots t_k}$ удовлетворяют неравенству

$$\int_{R^2} |\beta_1 - \beta_2|^\gamma \, d\mu_{t_1, t_2}(\beta_1, \beta_2) \leqslant |F(t_2) - F(t_1)|^\alpha, \quad (12.59)$$

то (12.58) тоже справедливо.

Доказательство. Для каждого n построим кусочно-линейную случайную функцию X_n, которая линейна на каждом интервале $[(i-1)2^{-n}, i2^{-n}]$ и для которой совместное распределение величин

$$X_n(0),\ X_n\left(\frac{1}{2^n}\right),\ \ldots,\ X_n\left(\frac{2^n - 1}{2^n}\right),\ X_n(1)$$

есть $\mu_{t_0, \ldots, t_{2^n}}$ с $t_i = i/2^n$. Если t_1 и t_2 являются целыми кратными 2^{-n}, то в силу (12.58) имеем

$$\mathsf{P}\{|X_n(t_2) - X_n(t_1)| \geqslant \lambda\} \leqslant \frac{1}{\lambda^\gamma} |F(t_2) - F(t_1)|^\alpha. \quad (12.60)$$

Как и в предыдущем доказательстве рассмотрим фиксированные n, δ и j, причем δ^{-1} будем считать целым. Если числа

$$j\delta + i\delta m^{-1}, \qquad i = 0, 1, \ldots, m, \quad (12.61)$$

входящие в формулу (12.54), являются все целыми кратными 2^{-n}, то (12.55) выводится, как и раньше. Допустим, теперь, что $\delta 2^n$ — целое число. Если мы положим $m = \delta 2^n$, то точки (12.61) действительно являются целыми кратными 2^{-n}, так что (12.55) выполняется. Более того, совокупность (12.61) в этом случае в точности совпадает с совокупностью целых кратных 2^{-n}, лежащих в интервале $[j\delta, (j+1)\delta]$. Отсюда вытекает вследствие кусочной линейности X_n, что (12.56) и (12.57) снова выполняются.

Таким образом, (12.57) выполняется, если и δ^{-1}, и $\delta 2^n$ являются целыми числами. Если мы положим $\delta = 2^{-v}$, где целое число v достаточно велико, чтобы правая часть (12.57) была меньше η, то (12.52) выполняется при всех $n \geqslant v$. Этого достаточно для плотности, так как все $X_n(0)$ имеют одно и то же распределение.

По теореме Прохорова существует, следовательно, случайный элемент X в C и подпоследовательность $\{X_{n'}\}$ такие, что $X_{n'} \xrightarrow{\mathcal{D}} X$. Если $t_1, \ldots, t_k$ являются двоично рациональными числами, то в силу предположения о согласованности мер вектор $(X_n(t_1), \ldots, X_n(t_k))$ имеет распределение $\mu_{t_1 \ldots t_k}$ для всех достаточно больших n. Следовательно, распределение вектора $(X(t_1), \ldots$ $\ldots, X(t_n))$ есть $\mu_{t_1 \ldots t_k}$. Для произвольных точек $t_1, \ldots, t_k$ из $[0, 1]$ существуют двоично рациональные числа $s_1^{(v)}, \ldots, s_k^{(v)}$ такие, что $\lim\limits_v s_i^{(v)} = t_i$; так как $\left(X\left(s_1^{(v)}\right), \ldots, X\left(s_k^{(v)}\right)\right)$ сходится по распределению к $(X(t_1), \ldots, X(t_k))$ и так как (вследствие (12.58) и непрерывности F) $\mu_{s_1^{(v)} \ldots s_k^{(v)}}$ слабо сходится к $\mu_{t_1 \ldots t_k}$, то $\left(X\left(t_1\right), \ldots, X\left(t_k\right)\right)$ имеет $\mu_{t_1 \ldots t_k}$ своим распределением. Таким образом, X представляет собой искомую случайную функцию.

Существование W и W° снова следует из теоремы 12.4 *).

Другие неравенства. В главе 3 нам понадобится усиление теоремы 12.1. Если

$$M''_m = \max_{0 \leqslant i \leqslant j \leqslant k \leqslant m} \min\{|S_j - S_i|, |S_k - S_j|\}, \tag{12.62}$$

то

$$M'_m \leqslant M''_m. \tag{12.63}$$

(Если $m = 3$ и $\xi_1 = -\xi_2 = \xi_3$, то $M'_m = 0$ и $M''_m = |\xi_1|$.) Поэтому оценки для хвостов распределения M''_m сильнее,

*) Теорема 12.4 в комбинации с теоремой 9.2 дает условие непрерывности выборочных траекторий сепарабельных процессов. Обобщения и другие варианты теорем 12.3 и 12.4 см. в задачах 7 и 8 этого параграфа и в задаче 4 § 15.

чем оценки для хвостов распределения M'_m. Следующий результат уточняет теорему 12.1.

Теорема 12.5. *Если* $\gamma \geqslant 0$ *и* $\alpha > 1/2$ *и если для всех положительных* λ

$$\mathsf{P}\{|S_j - S_i| \geqslant \lambda,\ |S_k - S_j| \geqslant \lambda\} \leqslant \leqslant \frac{1}{\lambda^{2\gamma}}\Big(\sum_{i<l\leqslant k} u_l\Big)^{2\alpha}, \quad 0 \leqslant i \leqslant j \leqslant k \leqslant m, \tag{12.64}$$

то для всех положительных λ

$$\mathsf{P}\{M''_m \geqslant \lambda\} \leqslant \frac{K''_{\gamma,\alpha}}{\lambda^{2\gamma}}(u_1 + \ldots + u_m)^{2\alpha}, \tag{12.65}$$

где $K''_{\gamma,\alpha}$ — *постоянная, зависящая только от* γ *и* α.

Доказательство. Положим

$$N_m = \min_{1\leqslant l\leqslant m} \max\{\max_{0\leqslant i<l} |S_i|,\ \max_{l\leqslant i\leqslant m} |S_m - S_i|\}. \tag{12.66}$$

Если l — индекс, при котором здесь достигается минимум, и $i \leqslant j \leqslant k$, то или $i, j < l$ или же $l \leqslant j, k$, откуда вытекает, что

$$M''_m \leqslant 2N_m. \tag{12.67}$$

Следовательно, достаточно найти постоянную K, зависящую только от γ и α такую, что (12.64) влечет

$$\mathsf{P}\{N_m \geqslant \lambda\} \leqslant \frac{K}{\lambda^{2\gamma}}(u_1 + \ldots + u_m)^{2\alpha}. \tag{12.68}$$

(Так как $N_m \leqslant 2M''_m$, то (12.68) и (12.65) равносильны.)

Пусть K, $K \geqslant 1$, столь велико, что

$$\left(\frac{2K}{2^{2\alpha}}\right)^{\delta} + (4\cdot 2^{2\gamma})^{\delta} \leqslant K^{\delta}, \tag{12.69}$$

где $\delta = 1/(2\gamma + 1)$, как и раньше. Мы докажем индукцией по m, что такое значение K годится для нашей цели. Случаи $m = 1$ и $m = 2$ разбираются просто.

Следуя принципу индукции, предположим, что результат имеет место для целых чисел, меньших m. Как и при доказательстве теоремы 12.1, выберем h, $1 \leqslant h \leqslant m$, такое, что

$$\frac{u_1 + \ldots + u_{h-1}}{u} \leqslant \frac{1}{2} \leqslant \frac{u_1 + \ldots + u_h}{u},$$

где $u = u_1 + \ldots + u_m$.

Введем обозначение A_1 для N_{h-1}:

$$A_1 = \min_{1 \leqslant l_1 < h} \max \{ \max_{0 \leqslant i < l_1} |S_i|, \max_{l_1 \leqslant i < h} |S_{h-1} - S_i| \}. \tag{12.70}$$

Введем обозначение A_2 для величины N_{m-h}, построенной по случайным величинам $\xi_{h+1}, \ldots, \xi_m$:

$$A_2 = \min_{h < l_2 \leqslant m} \max \{ \max_{h \leqslant i < l_2} |S_i - S_h|, \max_{l_2 \leqslant i \leqslant m} |S_m - S_i| \}. \tag{12.71}$$

Применяя предположение индукции к $\xi_1, \ldots, \xi_{h-1}$, имеем

$$\mathsf{P}\{A_1 \geqslant \lambda\} \leqslant \frac{K}{\lambda^{2\gamma}} (u_1 + \ldots + u_{h-1})^{2\alpha} \leqslant \frac{u^{2\alpha}}{\lambda^{2\gamma}} \frac{K}{2^{2\alpha}}. \tag{12.72}$$

Применяя предположение индукции к $\xi_{h+1}, \ldots, \xi_m$, имеем

$$\mathsf{P}\{A_2 \geqslant \lambda\} \leqslant \frac{K}{\lambda^{2\gamma}} (u_{h+1} + \ldots + u_m)^{2\alpha} \leqslant \frac{u^{2\alpha}}{\lambda^{2\gamma}} \frac{K}{2^{2\alpha}}. \tag{12.73}$$

Обозначим

$$\mu(i, j, k) = \min\{ |S_j - S_i|, |S_k - S_j| \}$$

и определим

$$B = \max\{\mu(0, h-1, m);\ \mu(0, h-1, h);\ \mu(h-1, h, m);\ \mu(0, h, m)\}. \tag{12.74}$$

В силу (12.64)

$$\mathsf{P}\{B \geqslant \lambda\} \leqslant 4 \frac{u^{2\alpha}}{\lambda^{2\gamma}}. \tag{12.75}$$

Покажем теперь, что

$$N_m \leqslant \max\{A_1, A_2\} + 2B. \tag{12.76}$$

Пусть l_1 и l_2 — те значения индексов, при которых достигается минимум в (12.70) и (12.71). Для того чтобы доказать (12.76), мы должны указать такое l, $1 \leqslant l \leqslant m$, для которого

$$\max_{0 < i < l} |S_i| \leqslant \max\{A_1, A_2\} + 2B \tag{12.77}$$

и

$$\max_{l\leqslant i\leqslant m} |S_m - S_i| \leqslant \max\{A_1, A_2\} + 2B. \qquad (12.78)$$

Допустим сначала, что

$$|S_{h-1}| \leqslant B \qquad (12.79)$$

и

$$|S_m - S_h| \leqslant B. \qquad (12.80)$$

Мы покажем, что значение $l = h$ удовлетворяет (12.77) и (12.78). Действительно, из $0 \leqslant i \leqslant l_1$ следует, что $|S_i| \leqslant A_1$, и из $l_1 \leqslant i < l = h$ следует в силу (12.79), что

$$|S_i| \leqslant |S_i - S_{h-1}| + |S_{h-1}| \leqslant A_1 + B,$$

так что (12.77) выполняется; соотношение (12.78) устанавливается тем же самым путем.

Допустим теперь, что какое-либо из неравенств (12.79), (12.80) не выполняется. Для определенности предположим, что не выполняется неравенство (12.79). Покажем, что значение $l = l_1$ удовлетворяет (12.77) и (12.78). Так как (12.79) не выполняется, то из определения величины B в (12.74) следует, что $|S_m - S_{h-1}| \leqslant B$ и $|\xi_h| \leqslant B$. Если $0 \leqslant i < l = l_1$, то $|S_i| \leqslant A_1$. Следовательно, неравенство (12.77) справедливо. Если $l = l_1 \leqslant \leqslant i < h$, то

$$|S_m - S_i| \leqslant |S_{h-1} - S_i| + |S_m - S_{h-1}| \leqslant A_1 + B;$$

если $h \leqslant i < l_2$, то

$$|S_m - S_i| \leqslant |S_i - S_h| + |\xi_h| + |S_m - S_{h-1}| \leqslant A_2 + 2B;$$

если $l_2 \leqslant i \leqslant m$, то $|S_m - S_i| \leqslant A_2$. Следовательно, неравенство (12.78) тоже выполняется. Если вместо (12.79) не выполняется (12.80), то (12.77) и (12.78) справедливы с $l = l_2$. Тем самым (12.76) доказано.

Для положительных чисел λ_0 и λ_1, в сумме равных λ, (12.76) влечет

$$\mathsf{P}\{N_m \geqslant \lambda\} \leqslant \mathsf{P}\{A_1 \geqslant \lambda_0\} + \mathsf{P}\{A_2 \geqslant \lambda_0\} + \mathsf{P}\left\{B \geqslant \frac{1}{2}\lambda_1\right\}. \qquad (12.81)$$

Применяя неравенства (12.72), (12.73) и (12.75), мы получаем

$$\mathsf{P}\{N_m \geqslant \lambda\} \leqslant \frac{u^{2\alpha}}{\lambda_0^{2\gamma}} \frac{2K}{2^{2\alpha}} + 4 \cdot 2^{2\gamma} \frac{u^{2\alpha}}{\lambda_1^{2\gamma}}.$$

Используя теперь (12.39), находим

$$\mathsf{P}\{N_m \geqslant \lambda\} \leqslant \frac{u^{2\alpha}}{\lambda^{2\gamma}} \left[\left(\frac{2K}{2^{2\alpha}}\right)^{\delta} + (4 \cdot 2^{2\gamma})^{\delta}\right]^{1/\delta},$$

откуда в силу (12.69) следует (12.68). Этим завершается доказательство.

Предположения теоремы 12.5 удовлетворены, если ξ_i независимы, $\mathsf{E}\{\xi_i\}=0$, $\mathsf{E}\{\xi_i^2\}=u_i$, $\gamma=2$ и $\alpha=1$. В этом случае ξ_i и u_i удовлетворяют также более сильным предположениям следующей теоремы и, следовательно, справедливы и ее более сильные выводы. Если все дисперсии u_i, кроме одной, равны 0, то $\mathsf{P}\{M''_m \geqslant \lambda\} = 0$, когда $\lambda > 0$; правая часть (12.65) при этом положительна, а правая часть (12.83) равна 0.

Теорема 12.6. *Если $\gamma \geqslant 0$ и $\alpha > 1/2$ и если для всех положительных λ*

$$\mathsf{P}\{|S_j - S_i| \geqslant \lambda,\ |S_k - S_j| \geqslant \lambda\} \leqslant$$
$$\leqslant \frac{1}{\lambda^{2\gamma}} \Big(\sum_{i<l\leqslant j} u_l\Big)^{\alpha} \Big(\sum_{j<l\leqslant k} u_l\Big)^{\alpha}, \tag{12.82}$$

то для всех положительных λ

$$\mathsf{P}\{M''_m \geqslant \lambda\} \leqslant$$
$$\leqslant \frac{K'''_{\gamma,\alpha}}{\lambda^{2\gamma}} (u_1 + \ldots + u_m)^{2\alpha} \min_{1\leqslant h\leqslant m} \left[1 - \frac{u_h}{u_1 + \ldots + u_m}\right]^{\alpha}, \tag{12.83}$$

где $K'''_{\gamma,\alpha}$ зависит только от γ и α.

Моментной формой неравенства (12.82) является первое из рассмотренных нами неравенств (12.9), которое влечет (12.82).

Доказательство. Выберем h, минимизирующее последний множитель в (12.83); обозначим

$$u = u_1 + \ldots + u_m,$$
$$p = (u_1 + \ldots + u_{h-1})/u, \quad p_h = u_h/u$$

и

$$q=(u_{h+1}+\ \dots\ +u_m)/u.$$

Как и в предыдущем доказательстве, определим A_1 формулой (12.70), A_2 — формулой (12.71) и B — формулой (12.74). В силу (12.82)

$$\mathsf{P}\{\mu(0,\,h-1,\,m)\geqslant\lambda\}\leqslant\frac{u^{2\alpha}}{\lambda^{2\gamma}}\,p^\alpha\,(p_h+q)^\alpha\leqslant\frac{u^{2\alpha}}{\lambda^{2\gamma}}\,(1-p_h)^\alpha.$$

Подобная оценка имеется для каждой из остальных величин μ в (12.74). Поэтому

$$\mathsf{P}\{B\geqslant\lambda\}\leqslant 4\,\frac{u^{2\alpha}}{\lambda^{2\gamma}}\,(1-p_h)^\alpha.$$

Обозначим $K=K''_{\gamma,\,\alpha}$. Так как из (12.82) следует (12.64), то из теоремы 12.5 (в усиленном варианте (12.68)) вытекает, что

$$\mathsf{P}\{A_1\geqslant\lambda\}\leqslant\frac{Ku^{2\alpha}}{\lambda^{2\gamma}}\,p^{2\alpha}\leqslant\frac{Ku^{2\alpha}}{\lambda^{2\gamma}}\,(1-p_h)^\alpha$$

и

$$\mathsf{P}\{A_2\geqslant\lambda\}\leqslant\frac{Ku^{2\alpha}}{\lambda^{2\gamma}}\,q^{2\alpha}\leqslant\frac{Ku^{2\alpha}}{\lambda^{2\gamma}}\,(1-p_h)^\alpha.$$

Далее, неравенство (12.76), как и раньше, справедливо и, следовательно,

$$\mathsf{P}\{N_m\geqslant\lambda\}\leqslant$$
$$\leqslant\mathsf{P}\left\{A_1\geqslant\tfrac{1}{2}\,\lambda\right\}+\mathsf{P}\left\{A_2\geqslant\tfrac{1}{2}\,\lambda\right\}+\mathsf{P}\left\{B\geqslant\tfrac{1}{4}\,\lambda\right\}.$$

Комбинируя это неравенство с тремя предыдущими, мы приходим к оценке

$$\mathsf{P}\{N_m\geqslant\lambda\}\leqslant 4^{2\gamma}(2K+4)\,\frac{u^{2\alpha}}{\lambda^{2\gamma}}\,(1-p_h)^\alpha.$$

Справедливость (12.83) при некотором $K'''_{\gamma,\,\alpha}$ легко теперь следует из (12.67).

Примечание. Результаты этого параграфа являются новыми. В их основе лежат работы Колмогорова (см. Слуцкий (1937) и Ченцов (1956)).

Задачи

1. Сделайте задачу 1 § 10 еще раз, используя (12.19) вместо (10.7).

2. Если мы ослабим (12.42), предположив, что это неравенство справедливо только для пары индексов $i = 0$, $j = m$, но компенсируем это предположением, что $S_1, \ldots, S_m$ образуют мартингал и $\gamma \geqslant 1$, то [Дуб (1956), стр. 302)] $\mathsf{P}\{M_m \geqslant \lambda\} \leqslant (u_1 + \ldots + u_m)^{\alpha}/\lambda^{\gamma}$ (эта оценка по существу одинакова по силе с (12.44)).

3. Пусть $\xi_1, \ldots, \xi_m$ независимы со средним 0 и конечными моментами порядка $2k$. Используйте полиномиальную теорему для нахождения постоянной C_k, зависящей только от k и такой, что

$$\mathsf{E}\left\{S_m^{2k}\right\} \leqslant C_k \left[\sum_{i=1}^{m} \mathsf{E}^{1/k}\left\{\xi_i^{2k}\right\}\right]^k.$$

Обобщите (12.48).

4. Используя (12.12), покажите, что для положительного ε

$$\mathsf{E}\left\{\left(M_m'\right)^{2\gamma-\varepsilon}\right\} \leqslant \frac{2\gamma}{\varepsilon} K_{\gamma,\alpha}(u_1 + \ldots + u_m)^{2\alpha(1-\varepsilon/2\gamma)}.$$

5. Преобразуя доказательство неравенства Меньшова [Дуб (1956), стр. 144)], покажите, что если (12.10) справедливо с $\alpha \geqslant 1/2$ и $\gamma \geqslant 1/2$, то

$$\mathsf{E}\left\{\left(M_m'\right)^{2\gamma}\right\} \leqslant (\log_2 2m)^{2\gamma}(u_1 + \ldots + u_m)^{2\alpha}.$$

Затем выведите, что если справедливо (12.42) с $\alpha \geqslant 1$ и $\gamma \geqslant 1$, то

$$\mathsf{E}\left\{M_m^{\gamma}\right\} \leqslant (\log_2 4m)^{\gamma}(u_1 + \ldots + u_m)^{\alpha}.$$

(При $\gamma = 2$ и $\alpha = 1$ снова получается неравенство Меньшова.)

6. Используйте теорему 12.2 для доказательства того, что если $\xi_1, \xi_2, \ldots$ удовлетворяют неравенству

$$\mathsf{E}\left\{\left|\sum_{i<l\leqslant j} \xi_l\right|^{\gamma}\right\} \leqslant \left(\sum_{i<l\leqslant j} u_l\right)^{\alpha} \qquad 0 \leqslant i \leqslant j < \infty,$$

с $\gamma \geqslant 0$, $\alpha > 1$ и $\sum u_t < \infty$, то ряд $\sum \xi_t$ сходится с вероятностью 1. Получите результат, связанный аналогичным образом с теоремой 12.6.

7. Теорема 12.3 остается верной, если заменить (12.50) предположением, что неравенство

$$\mathsf{P}\{|X_n(t) - X_n(t_1)| \geqslant \lambda,\ |X_n(t_2) - X_n(t)| \geqslant \lambda\} \leqslant \\ \leqslant \frac{1}{\lambda^{2\gamma}}|F(t_2) - F(t_1)|^{2\alpha} \qquad (12.84)$$

справедливо при $t_1 \leqslant t \leqslant t_2$ и всех n, где $\alpha > 1/2$.

8. Теорема 12.4 станет неверной, если (12.58) ослабить до условия, аналогичного (12.84); она также станет неверной, если условие

$\alpha > 1$ заменить условием $\alpha \geqslant 1$. [Предположите, что μ_t есть единичная масса в 0 или в 1, если $t < 1/2$ или $t \geqslant 1/2$.] Посмотрите, однако, задачу 4 в § 15.

9. Пусть

$$x_{nk}(t) = \begin{cases} 2n\left(t - \frac{k-1}{n}\right) & \text{при } \frac{k-1}{n} \leqslant t \leqslant \frac{k-1}{n} + \frac{1}{2n}, \\ 2n\left(\frac{k}{n} - t\right) & \text{при } \frac{k-1}{n} + \frac{1}{2n} \leqslant t \leqslant \frac{k}{n}, \\ 0 & \text{в остальных случаях} \end{cases}$$

и пусть X_n принимает значение x_{nk} с вероятностью $1/n$, $k = 1, \ldots, n$. Тогда

$$\mathsf{E}\{|X_n(t_2) - X_n(t_1)|\} \leqslant 4|t_2 - t_1|$$

для всех n, t_1 и t_2. Так как последовательность $\{X_n\}$ не является плотной, мы не можем положить $\alpha = 1$ в теореме 12.3. Подобным же образом, мы не можем положить $\alpha = 1/2$ в задаче 7.

§ 13. Эмпирические функции распределения

Пусть $\xi_1, \xi_2 \ldots$ — случайные величины на некотором пространстве $(\Omega, \mathscr{B}, \mathsf{P})$. Будем предполагать, что

$$0 \leqslant \xi_n(\omega) \leqslant 1 \tag{13.1}$$

(этого всегда можно добиться с помощью соответствующего преобразования).

Эмпирическая (или выборочная) *функция распределения* $F_n(t, \omega)$, соответствующая точкам $\xi_1(\omega), \ldots, \xi_n(\omega)$, определяется при $0 \leqslant t \leqslant 1$ как умноженное на $1/n$ число индексов $i \leqslant n$, для которых $\xi_i(\omega) \leqslant t$.

Если ξ_n независимы и имеют одну и ту же функцию распределения $F(t)$ (ввиду (13.1) существенны только значения t из $[0, 1]$), то при больших n $F_n(t, \omega)$ должна аппроксимировать $F(t)$, иными словами, разность

$$F_n(t, \omega) - F(t) \tag{13.2}$$

должна быть мала. Согласно теореме Гливенко — Кантелли величина

$$\sup_{0 \leqslant t \leqslant 1} |F_n(t, \omega) - F(t)| \tag{13.3}$$

сходится к 0 с вероятностью 1 *). Если функция F не-

*) Лоэв (1962, стр. 28).

прерывна, то можно найти предельное распределение (надлежащим образом нормированной) величины (13.3); эта предельная теорема, принадлежащая Колмогорову, соотносится с теоремой Гливенко — Кантелли так же, как центральная предельная теорема с законом больших чисел.

Мы рассмотрим в этом параграфе только случай $F(t) \equiv t$; так что ξ_n независимы и имеют одно и то же равномерное распределение на $[0, 1]$. Результат Колмогорова заключается в том, что

$$P\{\omega: \sup_t |\sqrt{n}(F_n(t, \omega) - t)| \leqslant \alpha\} \to$$
$$\to 1 - 2\sum_{k=1}^{\infty}(-1)^{k+1} e^{-2k^2\alpha^2}, \quad \alpha \geqslant 0. \tag{13.4}$$

Мы получим (13.4), используя теорию слабой сходимости в C.

Рассмотрим функцию $Y_n(\omega)$, принимающую значение

$$Y_n(t, \omega) = \sqrt{n}(F_n(t, \omega) - t) \tag{13.5}$$

в точке t. Хотя $Y_n(\omega)$ и является заданной на $[0, 1]$ случайной функцией, она, будучи разрывной, не является элементом C. Если бы мы установили сходимость по распределению Y_n как случайного элемента соответствующего пространства разрывных функций с соответствующей метрикой, мы были бы в состоянии получить (13.4) методами §§ 10 и 11, так как

$$\sup_t |\sqrt{n}(F_n(t, \omega) - t)| = h(Y_n(\omega)),$$

где $h(x) = \sup_t |x(t)|$. Хотя в следующей главе мы будем анализировать Y_n как случайный элемент метрического пространства разрывных функций, здесь мы обойдем проблемы разрывности, принимая другое определение эмпирической функции распределения. (Мы сможем тем не менее вывести именно соотношение (13.4), а не какое-либо его видоизменение.)

Пусть $G_n(t, \omega)$ как функция $t \in [0, 1]$ есть функция распределения, соответствующая равномерному распределению массы $(n+1)^{-1}$ на каждом из $(n+1)$ интервалов с концами $[\xi_{(i-1)}(\omega), \xi_{(i)}(\omega)]$, где $\xi_{(0)} = 0$, $\xi_{(n+1)} = 1$ и $\xi_{(1)}, \ldots, \xi_{(n)}$ есть величины $\xi_1, \ldots, \xi_n$, расположенные в порядке возрастания. (Отметим, что $F_n(t, \omega)$ соответствует распределению массы n^{-1} в каждой из n точек $\xi_1(\omega), \ldots, \xi_n(\omega)$.) Функции $F_n(t, \omega)$ и $G_n(t, \omega)$ близки между собой (рис. 6):

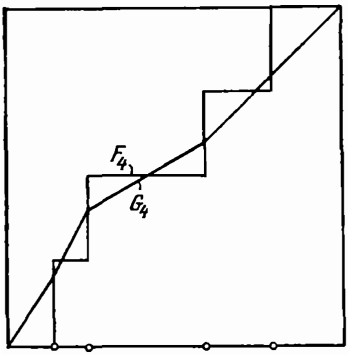

Рис. 6.

$$|F_n(t, \omega) - G_n(t, \omega)| \leqslant 1/n, \quad 0 \leqslant t \leqslant 1. \tag{13.6}$$

Пусть теперь $Z_n(\omega)$ — элемент пространства C со значением

$$Z_n(t, \omega) = \sqrt{n}\,(G_n(t, \omega) - t) \tag{13.7}$$

в точке t. Так как при каждом t функция $Z_n(t)$ представляет собой случайную величину, Z_n является случайным элементом в C $\left(Z_n^{-1}\mathscr{C} \subset \mathscr{B}\right)$. В силу (13.6) имеем

$$\sup_t |Y_n(t, \omega) - Z_n(t, \omega)| \leqslant 1/\sqrt{n}. \tag{13.8}$$

Теорема 13.1. *Если ξ_n независимы и одинаково распределены на* $[0, 1]$ *и если случайная функция Z_n определена формулой* (13.7), *то*

$$Z_n \xrightarrow{\mathscr{D}} W^\circ, \tag{13.9}$$

где W° является броуновским мостом.

Прежде чем доказывать теорему, выведем из нее (13.4). Так как функция $h(x) = \sup_t |x(t)|$ непрерывна на C, то из (13.9) вытекает, что $h(Z_n) \xrightarrow{\mathscr{D}} h(W^\circ)$. В силу (11.39) имеем

$$\mathsf{P}\{\sup_t |Z_n(t)| \leqslant \alpha\} \to 1 - 2\sum_{k=1}^{\infty} (-1)^{k+1} e^{-2k^2\alpha^2}, \qquad \alpha \geqslant 0,$$

так что (13.4) следует из (13.8) и теоремы 4.1.

Используя (11.40), мы можем тем же самым путем получить соотношение

$$\mathsf{P}\left\{\sqrt{n}\,\sup_t (F_n(t, \omega) - t) \leqslant \alpha\right\} \to 1 - e^{-2\alpha^2}, \quad \alpha \geqslant 0. \tag{13.10}$$

Если $h(x)$ представляет собой лебегову меру множества тех t, для которых $x(t) > 0$, то в силу (11.42)

$$\mathsf{P}\{h(Z_n) \leqslant \alpha\} \to \alpha, \qquad 0 \leqslant \alpha \leqslant 1; \tag{13.11}$$

$h(Z_n)$ приближенно равно произведению n^{-1} на число индексов i, $1 \leqslant i \leqslant n$, для которых $\xi_{(i)} > i/n$.

Доказательство. Для доказательства (13.9) мы сначала покажем, что конечномерные распределения Z_n сходятся к конечномерным распределениям W°. Пусть $U_n(t, \omega) = nF_n(t, \omega)$ — число тех точек среди $\xi_1(\omega), \ldots$ $\ldots, \xi_n(\omega)$, которые удовлетворяют условию $\xi_i(\omega) \leqslant t$. Если

$$0 = t_0 < t_1 < \ldots < t_k = 1,$$

то случайные величины

$$U_n(t_i) - U_n(t_{i-1}), \qquad i = 1, 2, \ldots, k, \tag{13.12}$$

имеют полиномиальное распределение с параметрами n и $p_i = t_i - t_{i-1}$, $i = 1, 2, \ldots, k$. Из центральной предельной теоремы для полиномиальной схемы следует, что случайный вектор с компонентами

$$Y_n(t_i) - Y_n(t_{i-1}) = \frac{1}{\sqrt{n}}((U_n(t_i) - U_n(t_{i-1})) - np_i), \tag{13.13}$$
$$i = 1, 2, \ldots, k,$$

сходится по распределению к случайному вектору с компонентами $W^\circ_{t_i} - W^\circ_{t_{i-1}}$ (по формулам (9.14) и (9.15) эти нормально распределенные случайные величины имеют нужные дисперсии $p_i(1 - p_i)$ и ковариации $-p_ip_j$). В силу (13.8) сказанное остается верным при замене (13.13) на

$$Z_n(t_i) - Z_n(t_{i-1}), \qquad i = 1, 2, \ldots, k.$$

Таким образом, конечномерные распределения случайной функции Z_n сходятся надлежащим образом. Если мы докажем, что последовательность $\{Z_n\}$ плотна, то (13.9) будет выполняться. По теореме 8.3 для этого достаточно показать, что для любых положительных ε и η существуют δ, $0 < \delta < 1$, и n_0 такие, что как только $n \geqslant n_0$, то при всех t

$$\mathsf{P}\{\sup_{t \leqslant s \leqslant t+\delta} |Z_n(s) - Z_n(t)| \geqslant \varepsilon\} \leqslant \delta\eta. \tag{13.14}$$

В силу (13.8) мы можем заменить (13.14) на

$$\mathsf{P}\{\sup_{t\leqslant s\leqslant t+\delta}|Y_n(s)-Y_n(t)|\geqslant\varepsilon\}\leqslant\delta\eta. \tag{13.15}$$

Легко показать, что событие, вероятность которого оценивается в (13.15), принадлежит σ-полю, порожденному $\xi_1, \ldots, \xi_n$; неравенство (13.15) — это всего лишь неравенство, касающееся $\xi_1, \ldots, \xi_n$ (мы не анализируем Y_n как случайный элемент пространства разрывных функций). Для удобства обозначений положим в (13.15) $t=0$:

$$\mathsf{P}\{\sup_{s\leqslant\delta}|Y_n(s)|\geqslant\varepsilon\}\leqslant\delta\eta. \tag{13.16}$$

(Это не является дополнительным ограничением, поскольку распределения приращений $Y_n(t)$ стационарны.)

Оценим вероятность в (13.16), используя теорему 12.1. Покажем, что

$$\mathsf{E}\{|Y_n(s+p_1)-Y_n(s)|^2\cdot|Y_n(s+p_1+p_2)-Y_n(s+p_1)|^2\}\leqslant \leqslant 6p_1p_2. \tag{13.17}$$

Пусть α_i, $1\leqslant i\leqslant n$, равно $1-p_1$ или $-p_1$ в зависимости от того, лежит ξ_i в интервале $(s, s+p_1]$ или нет, и пусть β_i равно $1-p_2$ или $-p_2$ в зависимости от того, лежит ξ_i в интервале $(s+p_1, s+p_1+p_2]$ или нет. Тогда (13.17) эквивалентно неравенству

$$\mathsf{E}\left\{\left(\sum_{i=1}^n\alpha_i\right)^2\left(\sum_{i=1}^n\beta_i\right)^2\right\}\leqslant 6n^2p_1p_2. \tag{13.18}$$

Так как ξ_i независимы, то случайные векторы (α_i, β_i) также независимы. Так как ξ_i равномерно распределены, то векторы (α_i, β_i) принимают значения $(1-p_1, -p_2)$, $(-p_1, 1-p_2)$ и $(-p_1, -p_2)$ с вероятностями p_1, p_2 и $p_3=1-p_1-p_2$, соответственно. В силу равенства $\mathsf{E}\{\alpha_i\}=\mathsf{E}\{\beta_i\}=0$ и по соображениям симметрии

$$\mathsf{E}\left\{\left(\sum_{i=1}^n\alpha_i\right)^2\left(\sum_{i=1}^n\beta_i\right)^2\right\}=n\mathsf{E}\{\alpha_1^2\beta_1^2\}+ +n(n-1)\mathsf{E}\{\alpha_1^2\}\mathsf{E}\{\beta_2^2\}+2n(n-1)\mathsf{E}\{\alpha_1\beta_1\}\mathsf{E}\{\alpha_2\beta_2\}.$$

Теперь (13.18) следует из соотношений

$$\mathsf{E}\{\alpha_1^2\beta_1^2\}=p_1(1-p_1)^2p_2^2+p_2p_1^2(1-p_2)^2+p_3p_1^2p_2^2\leqslant 3p_1p_2,$$
$$\mathsf{E}\{\alpha_1^2\}E\{\beta_2^2\}=p_1(1-p_1)p_2(1-p_2)\leqslant p_1p_2$$

и

$$\mathsf{E}\{\alpha_1\beta_1\}E\{\alpha_2\beta_2\}=p_1^2p_2^2\leqslant p_1p_2.$$

Рассмотрим при фиксированном δ величины

$$Y_n\left(\frac{i}{m}\delta\right)-Y_n\left(\frac{i-1}{m}\delta\right),\qquad i=1,\ \ldots,\ m. \tag{13.19}$$

Если $\gamma=2$ и $\alpha=1$ и если $u_i=6^{1/2}\delta/m$, то в силу (13.17) предположения теоремы 12.1 выполнены с величинами (13.19) в роли ξ_i. Для

$$M'_m=\max_{1\leqslant i\leqslant m}\min\left\{\left|Y_n\left(\frac{i}{m}\delta\right)\right|,\ \left|Y_n(\delta)-Y_n\left(\frac{i}{m}\delta\right)\right|\right\},$$

мы получаем, таким образом,

$$\mathsf{P}\{M'_m\geqslant\varepsilon\}\leqslant\frac{6K}{\varepsilon^4}\delta^2 \tag{13.20}$$

с $K=K_{2,1}$.

Если

$$M_m=\max_{1\leqslant i\leqslant m}\left|Y_n\left(\frac{i}{m}\delta\right)\right|,$$

то согласно (12.4)

$$M_m\leqslant M'_m+|Y_n(\delta)|.$$

Отсюда и из (13.20) имеем

$$\mathsf{P}\{M_m\geqslant\varepsilon\}\leqslant\frac{2^4\cdot 6K}{\varepsilon^4}\delta^2+\mathsf{P}\left\{|Y_n(\delta)|\geqslant\frac{\varepsilon}{2}\right\}. \tag{13.21}$$

Далее, при каждом ω функция $Y_n(s,\omega)$ непрерывна справа в точке s. Поэтому M_m сходится к $\sup\limits_{s\leqslant\delta}|Y_n(s,\omega)|$ при каждом ω, когда $m\to\infty$. Следовательно, (13.21) влечет

$$\mathsf{P}\{\sup_{s\leqslant\delta}|Y_n(s)|>\varepsilon\}\leqslant\frac{2^4\cdot 6K}{\varepsilon^4}\delta^2+\mathsf{P}\left\{|Y_n(\delta)|\geqslant\frac{\varepsilon}{2}\right\}. \tag{13.22}$$

Вследствие асимптотической нормальности вектора (13.13) $Y_n(\delta) \xrightarrow{\mathscr{D}} \sqrt{\delta(1-\delta)} \cdot N$ при $n \to \infty$ (δ фиксировано) и, следовательно,

$$\mathsf{P}\left\{|Y_n(\delta)| \geqslant \frac{\varepsilon}{2}\right\} \to \mathsf{P}\left\{N \geqslant \frac{\varepsilon/2}{\sqrt{\delta(1-\delta)}}\right\} \leqslant \leqslant \frac{2^4\delta^2}{\varepsilon^4}\mathsf{E}\{N^4\} = \frac{3 \cdot 2^4}{\varepsilon^4}\delta^2.$$

Поэтому при n, превосходящем некоторое n_δ,

$$\mathsf{P}\left\{|Y_n(\delta)| \geqslant \frac{\varepsilon}{2}\right\} < \frac{6 \cdot 2^4}{\varepsilon^4}\delta^2,$$

откуда в силу (13.22)

$$\mathsf{P}\{\sup_{s \leqslant \delta} |Y_n(s)| > \varepsilon\} \leqslant \frac{6 \cdot 2^4(K+1)}{\varepsilon^4}\delta^2. \tag{13.23}$$

По заданным ε и η выберем δ так, что $6 \cdot 2^4(K+1) \times \delta^2/\varepsilon^4 < \delta\eta$. При $n > n_\delta$ (13.16) следует из (13.23).

Этим завершается доказательство теоремы 13.1. Изложение несколько затемняется тем, что, желая в действительности анализировать Y_n, мы вынуждены заменить Y_n на Z_n только для того, чтобы остаться в пространстве C. В § 16 мы рассмотрим функцию Y_n в естественном для нее пространстве и докажем предельное соотношение $Y_n \xrightarrow{\mathscr{D}} W^\circ$ и некоторые его обобщения в этом пространстве. Хотя доказательства в § 16 опираются на солидную теорию, они много более прозрачны, чем данные в настоящем параграфе.

Примечание. Дубу (1956) принадлежит идея доказательства (13.4), использующего переход от случайной функции (13.5) к броуновскому мосту. Донскер (1952) обосновал этот переход. Доказательство, помещенное здесь, восходит к Ченцову (1956). Кац (1949) первоначально доказал (13.11). Обзор предельных теорем, связанных с (13.5), и обширную библиографию см. у Дарлинга (1957). См. также более поздние статьи: Бикель (1967 и 1969), Бирнбаум, Пайк (1958), Пайк (1965 и 1968), Пайк, Шорак (1968).

Глава 3

ПРОСТРАНСТВО D

§ 14. Геометрия пространства D

Пространство C непригодно для описания процессов, которые, подобно пуассоновскому процессу, но в противоположность броуновскому движению, могут содержать скачки. В этой главе изучается слабая сходимость в пространстве, которое содержит некоторые разрывные функции.

Пространство D. Пусть $D = D[0, 1]$ — пространство функций x, определенных на $[0, 1]$, непрерывных справа и имеющих левосторонние пределы:

(i) При $0 \leqslant t < 1$ предел $x(t+) = \lim\limits_{s \downarrow t} x(s)$ существует и $x(t+) = x(t)$.

(ii) При $0 < t \leqslant 1$ предел $x(t-) = \lim\limits_{s \uparrow t} x(s)$ существует.

Говорят, что функция x имеет *разрыв первого рода* в точке t, если пределы $x(t-)$ и $x(t+)$ существуют, но различны и $x(t)$ заключается между ними. Все разрывы элемента пространства D — это разрывы первого рода; предположение о том, что $x(t) = x(t+)$, представляет собой удобную нормировку. Разумеется, C является подмножеством пространства D.

Определим для $x \in D$ и $T_0 \subset [0, 1]$

$$w_x(T_0) = \sup\{|x(s) - x(t)|: \ s, t \in T_0\}. \qquad (14.1)$$

Модуль непрерывности функции x, определенный формулой (8.1), можно выразить с помощью (14.1) следующим образом:

$$w_x(\delta) = \sup_{0 \leqslant t \leqslant 1-\delta} w_x[t, t+\delta]. \qquad (14.2)$$

Функция, непрерывная на $[0, 1]$, равномерно непрерывна на этом отрезке. Следующая лемма дает соответствующее представление о равномерности для элементов D.

Лемма 1. *Для любой функции x из D и любого положительного числа ε существуют точки $t_0, t_1, \ldots, t_r$ такие, что*

$$0 = t_0 < t_1 < \ldots < t_r = 1 \tag{14.3}$$

и

$$w_x[t_{i-1}, t_i) < \varepsilon, \qquad i = 1, 2, \ldots, r. \tag{14.4}$$

Доказательство. Определим τ как точную верхнюю грань тех точек t из $[0, 1]$, для которых интервал $[0, t)$ можно разбить на конечное множество подинтервалов $[t_{i-1}, t_i)$, удовлетворяющих условию (14.4). Поскольку $x(0) = x(0+)$, то $\tau > 0$; так как предел $x(\tau-)$ существует, то интервал $[0, \tau)$ сам может быть разбит аналогичным образом на подинтервалы; τ не может быть меньше 1, так как в этом случае $x(\tau) = x(\tau+)$.

Из этой леммы следует, что может существовать лишь конечное множество точек t, в которых скачок $|x(t) - x(t-)|$ превосходит заданное положительное число, поэтому x имеет не более чем счетное множество точек разрыва. Из леммы следует также, что функция x ограничена:

$$\sup_t |x(t)| < \infty. \tag{14.5}$$

Наконец, из леммы следует, что функция x может быть равномерно приближена простыми функциями, принимающими постоянное значение на интервалах, поэтому x измерима по Борелю.

Нам понадобится модуль, играющий в D ту же роль, что и модуль непрерывности в C. Положим при $0 < \delta < 1$

$$w'_x(\delta) = \inf_{\{t_i\}} \max_{0 < i \leqslant r} w_x[t_{i-1}, t_i), \tag{14.6}$$

где нижняя грань берется по всем конечным множествам $\{t_i\}$ точек, удовлетворяющих условиям

$$\begin{cases} 0 = t_0 < t_1 < \ldots < t_r = 1, \\ t_i - t_{i-1} > \delta, \quad i = 1, 2, \ldots, r. \end{cases} \tag{14.7}$$

Лемма 1 эквивалентна утверждению о том, что для любого элемента x пространства D справедливо предельное соотношение

$$\lim_{\delta\to 0} w'_x(\delta) = 0. \tag{14.8}$$

Определение модуля $w'_x(\delta)$ имеет смысл, даже если x не принадлежит D. Подобно тому, как условие $\lim\limits_{\delta\to 0} w'_x(\delta) = 0$ необходимо и достаточно для того, чтобы произвольная функция x, заданная на $[0, 1]$, принадлежала C, условие (14.8) необходимо и достаточно для того, чтобы функция x принадлежала D.

Так как при любом $\delta < 1/2$ интервал $[0, 1)$ можно разбить на подинтервалы $[t_{i-1}, t_i)$ длины $\delta < t_i - t_{i-1} \leqslant \leqslant 2\delta$, то

$$w'_x(\delta) \leqslant w_x(2\delta) \tag{14.9}$$

при $\delta < 1/2$. Противоположное неравенство, вообще говоря, не имеет места; это следует из соотношения (14.8) и того факта, что $w_x(\delta)$ не стремится к 0 вместе с δ, если x имеет разрывы.

Допустим, однако, что $x \in C$. При данном ε выберем точки $\{t_i\}$, удовлетворяющие условиям (14.7) и

$$\max_{0 < i \leqslant r} w_x[t_{i-1}, t_i) < w'_x(\delta) + \varepsilon. \tag{14.10}$$

Если $|s - t| < \delta$, то s и t лежат в одном и том же подинтервале $[t_{i-1}, t_i)$ или в граничащих подинтервалах. Из условия (14.10) и предположения о непрерывности x следует поэтому, что $w_x(\delta) \leqslant 2w'_x(\delta) + 2\varepsilon$. В силу произвольности ε имеем

$$w_x(\delta) \leqslant 2w'_x(\delta), \quad \text{если} \quad x \in C. \tag{14.11}$$

Согласно соотношениям (14.9) и (14.11) модули $w_x(\delta)$ и $w'_x(\delta)$ для непрерывных функций x по существу совпадают.

Топология Скорохода. Две функции x и y близки друг другу в равномерной топологии (используемой в пространстве C), если график $x(t)$ может быть преобразован в график $y(t)$ посредством равномерно малого возмущения ординат при остающихся неизменными абсциссах. В пространстве D мы разрешим также равномерно малую деформацию шкалы времени. Физически это равнозначно допущению о невозможности измерить время с большей точностью, чем положение в пространстве. Следующая топология, предложенная Скороходом, воплощает эту идею.

Пусть Λ обозначает класс строго возрастающих, непрерывных отображений отрезка $[0, 1]$ на себя. Если $\lambda \in \Lambda$, то $\lambda(0) = 0$ и $\lambda(1) = 1$. Определим расстояние $d(x, y)$ между x и y из D как нижнюю грань таких положительных ε, для которых существует $\lambda \in \Lambda$ такое, что

$$\sup_t |\lambda t - t| \leqslant \varepsilon \tag{14.12}$$

и

$$\sup_t |x(t) - y(\lambda t)| \leqslant \varepsilon. \tag{14.13}$$

В силу (14.5) функция $d(x, y)$ конечна (положим $\lambda t \equiv t$). Ясно, что $d(x, y) \geqslant 0$. Далее, из равенства $d(x, y) = 0$ следует, что для любого t либо $x(t) = y(t)$, либо $x(t) = y(t-)$, откуда в свою очередь вытекает, что $x = y$. Если отображение λ принадлежит Λ, то отображение λ^{-1} тоже принадлежит Λ; равенство $d(x, y) = d(y, x)$ является следствием двух соотношений:

$$\sup_t |\lambda^{-1} t - t| = \sup_t |\lambda t - t|$$

и

$$\sup_t |x(\lambda^{-1} t) - y(t)| = \sup_t |x(t) - y(\lambda t)|.$$

Если λ_1 и λ_2 принадлежат классу Λ, то и их композиция $\lambda_2 \lambda_1$ также принадлежит Λ; неравенство треугольника выполняется в силу неравенств

$$\sup_t |\lambda_2 \lambda_1 t - t| \leqslant \sup_t |\lambda_1 t - t| + \sup_t |\lambda_2 t - t|$$

и

$$\sup_t |x(t) - z(\lambda_2 \lambda_1 t)| \leqslant \\ \leqslant \sup_t |x(t) - y(\lambda_1 t)| + \sup_t |y(t) - z(\lambda_2 t)|.$$

Таким образом, доказано, что d — метрика.

Эта метрика определяет топологию Скорохода. Равномерное расстояние между x и y можно определить как нижнюю грань тех положительных ε, для которых $\sup_t |x(t) - y(t)| \leqslant \varepsilon$. Отображение λ в (14.12) и (14.13) осуществляет именно ту равномерно малую деформацию шкалы времени, о которой упоминалось выше.

Последовательность элементов x_n пространства D сходится к пределу x в топологии Скорохода тогда и только тогда, когда в Λ существуют функции λ_n такие, что равномерно по t

$$\lim_n x_n(\lambda_n t) = x(t)$$

и

$$\lim_n \lambda_n t = t.$$

Если x_n равномерно сходится к x, то имеет место сходимость в топологии Скорохода (достаточно положить $\lambda_n t \equiv t$). С другой стороны, последовательность $x_n = I_{[0,\ 1/2+1/n)}$ сходится к $x = I_{[0,\ 1/2)}$ в топологии Скорохода, но $x_n(t) \nrightarrow x(t)$ при $t = 1/2$. Так как

$$|x_n(t) - x(t)| \leqslant |x_n(t) - x(\lambda_n^{-1}t)| + |x(\lambda_n^{-1}t) - x(t)|,$$

то сходимость в топологии Скорохода влечет сходимость $x_n(t) \to x(t)$ в точках t непрерывности функции x, и если x (равномерно) непрерывна во всех точках отрезка $[0,\ 1]$, то сходимость в топологии Скорохода влечет равномерную сходимость. В частности, *топология в C, индуцируемая топологией Скорохода, совпадает с равномерной топологией.*

Пространство D сепарабельно; в качестве счетного всюду плотного множества можно взять множество, образованное такими элементами x, которые имеют рациональное значение при $t = 1$ и принимают постоянное рациональное значение при некотором целом k на каждом подинтервале $[(i-1)/k,\ i/k)$, $0 < i \leqslant k$ (это следует из леммы 1 и определения d). Однако пространство D не является полным относительно метрики d: если

$$x_n = I_{[1/2,\ 1/2+1/n)}, \tag{14.14}$$

то

$$d(x_n,\ x_m) = |1/n - 1/m|, \tag{14.15}$$

и следовательно, последовательность $\{x_n\}$ фундаментальна в метрике d, хотя и не сходится. Мы введем в D другую метрику d_0, которая эквивалентна d в том смысле, что порождает также топологию Скорохода, но относительно которой D является полным пространством. Полнота пространства облегчит нам характеризацию компактных множеств.

Идея определения d_0 такова: требуется, чтобы преобразование λ шкалы времени, которое фигурирует в определении d, было бы близким к функции, тождественно равной 1, в некотором, на первый взгляд, более строгом, чем в (14.12), смысле; именно, мы требуем, чтобы тангенс угла наклона любой хорды $(\lambda t - \lambda s)/(t - s)$ был близок к 1, или, что равносильно, но более удобно аналитически, чтобы его логарифм был близок к 0.

Если λ — неубывающая функция на $[0, 1]$, $\lambda 0 = 0$ и $\lambda 1 = 1$, то положим

$$\|\lambda\| = \sup_{s \neq t} \left| \log \frac{\lambda t - \lambda s}{t - s} \right|. \tag{14.16}$$

Если величина $\|\lambda\|$ конечна, то тангенсы углов наклона хорд лежат между двумя положительными константами; поэтому функция λ непрерывна и строго возрастает и, следовательно, является элементом класса Λ. Хотя величина $\|\lambda\|$ может быть бесконечной даже для функций λ, принадлежащих Λ, такие элементы не войдут в следующее определение.

Определим расстояние $d_0(x, y)$ как нижнюю грань таких положительных чисел ε, для которых Λ содержит некоторую функцию λ такую, что

$$\|\lambda\| \leqslant \varepsilon \tag{14.17}$$

и

$$\sup_t |x(t) - y(\lambda t)| \leqslant \varepsilon. \tag{14.18}$$

В силу (14.5) величина $d_0(x, y)$ конечна (положим $\lambda t \equiv t$). Из соотношений

$$\|\lambda^{-1}\| = \|\lambda\| \tag{14.19}$$

и

$$\|\lambda_2\lambda_1\| \leqslant \|\lambda_1\| + \|\lambda_2\| \tag{14.20}$$

следует, что d_0 — метрика.

Мы покажем теперь, что метрики d и d_0 эквивалентны. Будет показано также, что пространство D полно относительно метрики d_0; заметим пока, что если x_n определить равенством (14.14), то

$$d_0(x_n, x_m) = \min\left\{1, \left|\log \frac{m}{n}\right|\right\}, \quad m, n > 3, \tag{14.21}$$

так что во всяком случае не сходящаяся ни к какому пределу последовательность $\{x_n\}$ не является фундаментальной в метрике d_0.

Если $d_0(x, y) < \varepsilon$, то условия (14.17) и (14.18) выполняются для некоторой функции $\lambda \in \Lambda$. Если $\varepsilon < 1/4$, то, поскольку $\lambda 0 = 0$,

$$\log(1-2\varepsilon) < -\varepsilon \leqslant \log\frac{\lambda t}{t} \leqslant \varepsilon < \log(1+2\varepsilon)\ ^{*)},$$

так что $|\lambda t - t| \leqslant 2\varepsilon$. Поэтому

$$d(x, y) \leqslant 2d_0(x, y), \quad \text{если} \quad d_0(x, y) < 1/4. \tag{14.22}$$

Из сравнения выражений (14.15) и (14.21) можно вывести, что неравенство, противоположное неравенству (14.22), вообще говоря, невозможно. Следующая лемма, однако, показывает, что расстояние $d_0(x, y)$ мало в том случае, когда $d(x, y)$ и $w'_x(\delta)$ (или $w'_y(\delta)$) малы одновременно.

Лемма 2. *Если* $d(x, y) < \delta^2$, *где* $0 < \delta < 1/4$, *то* $d_0(x, y) \leqslant 4\delta + w'_x(\delta)$.

Доказательство. Выберем точки t_i, удовлетворяющие условиям (14.7) и

$$w_x[t_{i-1}, t_i) < w'_x(\delta) + \delta, \quad i = 1, 2, \ldots, r. \tag{14.23}$$

Затем выберем в Λ такой элемент μ, что

$$\sup_t |x(t) - y(\mu t)| = \sup_t |x(\mu^{-1}t) - y(t)| < \delta^2 \tag{14.24}$$

и

$$\sup_t |\mu t - t| < \delta^2. \tag{14.25}$$

Мы хотим найти в Λ такую функцию λ, которая была бы близка к μ, но не имела бы хорд (которые, возможно, имеет μ), тангенсы углов наклона которых сильно отличаются от 1. Пусть λ совпадает с μ в точках t_i и линейна в промежутках между этими точками. Так как композиция $\mu^{-1}\lambda$ является возрастающей функцией и $\mu^{-1}\lambda t_i = t_i$ при всех i, то t и $\mu^{-1}\lambda t$ всегда принадлежат

*) Если $|s| \leqslant 1/2$, то $s - s^2 \leqslant \log(1+s)$, тогда как $\log(1+s) \leqslant s$ для всех $s > -1$.

одному и тому же подинтервалу $[t_{i-1}, t_i)$. Поэтому в силу (14.23) и (14.24)

$$|x(t)-y(\lambda t)| \leqslant |x(t)-x(\mu^{-1}\lambda t)| + |x(\mu^{-1}\lambda t)-y(\lambda t)| <$$
$$< w'_x(\delta)+\delta+\delta^2 < 4\delta + w'_x(\delta).$$

Теперь достаточно показать, что $\|\lambda\| \leqslant 4\delta$. Так как λ совпадает с μ в точках t_i, из неравенства (14.25) и неравенства $t_i - t_{i-1} > \delta$ следует, что

$$|(\lambda t_i - \lambda t_{i-1}) - (t_i - t_{i-1})| < 2\delta^2 < 2\delta(t_i - t_{i-1}).$$

Поскольку функция λ является ломаной, получаем

$$|(\lambda t - \lambda s) - (t-s)| \leqslant 2\delta |t-s|$$

при всех s и t. Поэтому

$$\log(1-2\delta) \leqslant \log \frac{\lambda t - \lambda s}{t-s} \leqslant \log(1+2\delta),$$

и следовательно, $\|\lambda\| \leqslant 4\delta$, поскольку $\delta < 1/4$.

Теорема 14.1. *Метрики d и d_0 эквивалентны.*

Доказательство. Обозначим открытый d-шар через $S_d(x, \varepsilon)$ и открытый d_0-шар через $S_{d_0}(x, \varepsilon)$. Из неравенства (14.22) следует, что внутри произвольного шара $S_d(x, \varepsilon)$ можно найти шар $S_{d_0}(x, \delta)$. (Выбор нового радиуса не зависит от центра x.)

Из леммы 2 следует, что если

$$\delta < 1/4, \quad 4\delta + w'_x(\delta) < \varepsilon, \tag{14.26}$$

то $S_d(x, \delta^2) \subset S_{d_0}(x, \varepsilon)$. По заданным x и ε мы можем в силу (14.8) найти δ, удовлетворяющее условию (14.26). Таким образом, внутри произвольного d_0-шара мы можем найти d-шар с тем же самым центром. (На этот раз выбор нового радиуса зависит от центра x, что, конечно, должно быть так, если пространство D полно в метрике d_0 — если метрики d и d_0 задают различные классы фундаментальных последовательностей.) Таким образом, метрики d и d_0 эквивалентны.

Полнота пространства D.

Теорема 14.2. *Пространство D полно в метрике d_0.*

Доказательство. Достаточно показать, что каждая фундаментальная в метрике d_0 последовательность содержит сходящуюся в этой же метрике подпоследовательность. Если последовательность $\{x_k\}$ фундаментальна в метрике d_0, то она содержит подпоследовательность $\{y_n\}=\{x_{k_n}\}$ такую, что

$$d_0(y_n, y_{n+1}) < 1/2^n. \tag{14.27}$$

Мы покажем, что $\{y_n\}$ сходится в метрике d_0 к некоторому пределу.

В силу (14.27) Λ содержит функции μ_n такие, что

$$\sup_t |y_n(t) - y_{n+1}(\mu_n t)| < 1/2^n \tag{14.28}$$

и

$$\|\mu_n\| < 1/2^n. \tag{14.29}$$

Отсюда следует, что при $m \geqslant 1$

$$\sup_t |\mu_{n+m+1}\mu_{n+m} \ldots \mu_{n+1}\mu_n t - \mu_{n+m} \ldots \mu_{n+1}\mu_n t| = \\ = \sup_s |\mu_{n+m+1}s - s| \leqslant 1/2^{n+m}$$

(здесь $\mu_{n+m} \ldots \mu_{n+1}\mu_n$ обозначает многократную композицию). Таким образом, при фиксированном n функции

$$\mu_{n+m} \ldots \mu_{n+1}\mu_n t \tag{14.30}$$

равномерно фундаментальны при $m \to \infty$. Поэтому последовательность функций (14.30) равномерно сходится к пределу

$$\lambda_n t = \lim_m \mu_{n+m} \ldots \mu_{n+1}\mu_n t. \tag{14.31}$$

Функция λ_n непрерывна, не убывает и $\lambda_n 0 = 0$, $\lambda_n 1 = 1$. Если мы покажем, что $\|\lambda_n\|$ конечна, то отсюда будет следовать, что функция λ_n строго возрастающая и, следовательно, является элементом Λ. В силу (14.20)

$$\left|\log \frac{\mu_{n+m} \ldots \mu_{n+1}\mu_n t - \mu_{n+m} \ldots \mu_{n+1}\mu_n s}{t-s}\right| \leqslant \\ \leqslant \|\mu_{n+m} \ldots \mu_{n+1}\mu_n\| \leqslant \|\mu_n\| + \\ + \|\mu_{n+1}\| + \ldots + \|\mu_{n+m}\| \leqslant 1/2^{n-1}.$$

Устремляя $m \to \infty$ в левой части последнего неравенства, находим, что $\|\lambda_n\| \leqslant 1/2^{n-1}$; в частности, мы получаем, что $\|\lambda_n\|$ конечна, и поэтому $\lambda_n \in \Lambda$. Согласно (14.31) имеем $\lambda_n = \lambda_{n+1}\mu_n$. Поэтому в силу (14.28)

$$\sup_t \left|y_n\left(\lambda_n^{-1}t\right) - y_{n+1}\left(\lambda_{n+1}^{-1}t\right)\right| = \\ = \sup_s |y_n(s) - y_{n+1}(\mu_n s)| < 1/2^n.$$

Таким образом, функции $y_n(\lambda_n^{-1}t)$, являющиеся элементами пространства D, равномерно фундаментальны и, следовательно, равномерно сходятся к предельной функции $x(t)$. Легко показать, что функция x также будет элементом D. Поскольку

$$\sup_t \left| y_n(\lambda_n^{-1}t) - x(t) \right| \to 0 \quad \text{и} \quad \|\lambda_n\| \to 0,$$

то

$$d_0(y_n, x) \to 0,$$

что и завершает доказательство.

Компактность в D. Займемся теперь задачей характеризации компактных подмножеств D. В терминах модуля $w_x(\delta)$, определенного формулой (14.6), мы имеем следующий аналог теоремы Арцела — Асколи (стр. 302):

Теорема 14.3. *Множество A имеет компактное замыкание в топологии Скорохода тогда и только тогда, когда*

$$\sup_{x \in A} \sup_t |x(t)| < \infty \tag{14.32}$$

и

$$\lim_{\delta \to 0} \sup_{x \in A} w'_x(\delta) = 0. \tag{14.33}$$

Отметим, что ввиду неравенства (14.9) условие (14.33) слабее, чем

$$\lim_{\delta \to 0} \sup_{x \in A} w_x(\delta) = 0. \tag{14.34}$$

Доказательство. Поскольку пространство D полно в метрике d_0, то для доказательства достаточности условий теоремы нужно лишь показать, что их следствием является полная ограниченность множества A относительно метрики d_0. Покажем сначала, что в условиях (14.32) и (14.33) множество A вполне ограничено относительно метрики d.

По заданному положительному ε выберем целое число k так, что $1/k < \varepsilon$ и $w'_x(1/k) < \varepsilon$ при всех x из A. Обозначим через H конечную ε-сеть в интервале $[-\alpha, \alpha]$, где

$$\alpha = \sup_{x \in A} \sup_t |x(t)|.$$

Пусть B — конечное множество таких функций y в D, которые на каждом интервале $[(u-1)/k, u/k]$ принимают постоянное значение из H и удовлетворяют условию $y(1)\in H$. Мы покажем, что множество B является 2ε-сетью относительно метрики d.

Пусть x принадлежит A. В силу неравенства $w'_x(1/k)<\varepsilon$ существуют точки t_i, $0=t_0<t_1<\ldots<t_r=1$, такие, что

$$t_i-t_{i-1}>1/k \tag{14.35}$$

и

$$w'_x[t_{i-1}, t_i)<\varepsilon, \qquad 0<i\leqslant r. \tag{14.36}$$

Выберем целые числа u_i так, чтобы $u_i/k\leqslant t_i<(u_i+1)/k$, $i=0, 1, \ldots, r$. Поскольку в силу (14.35) числа u_i должны быть различны, то в Λ существует функция λ, которая переводит точки u_i/k в t_i и линейна в промежутке между этими точками. Выберем в множестве B точку y так, чтобы

$$\left|y\left(\frac{u}{k}\right)-x\left(\lambda\frac{u}{k}\right)\right|<\varepsilon, \quad 0\leqslant u\leqslant k. \tag{14.37}$$

Поскольку интервал $\left[\lambda\frac{u}{k}, \lambda\frac{u+1}{k}\right)$ должен лежать внутри одного из интервалов $\left[\lambda\frac{u_i}{k}, \lambda\frac{u+1}{k}\right)=[t_i, t_{i+1})$, то функция $x(\lambda t)$ не может в силу (14.36) изменяться больше, чем на ε, когда t меняется в интервале $[u/k, (u+1)/k)$. Так как функция y постоянна на интервалах этого вида, то из неравенства (14.37) следует, что $|y(t)-x(\lambda t)|<2\varepsilon$ для всех t. По построению

$$\left|\lambda\frac{u_i}{k}-\frac{u_i}{k}\right|=\left|t_i-\frac{u_i}{k}\right|<\frac{1}{k}<\varepsilon,$$

откуда в силу линейности $\sup_t|\lambda t-t|<\varepsilon$. Таким образом, $d(x, y)<2\varepsilon$, так что множество B является 2ε-сетью относительно метрики d.

При заданном положительном η выберем δ так, чтобы $0<\delta<1/4$ и $4\delta+w'_x(\delta)<\eta$ для всех x из A. Затем выберем ε так, чтобы выполнялось неравенство $0<2\varepsilon<\delta^2$. Если B представляет собой только что построенное множество — конечную 2ε-сеть для множества A относительно метрики d — то в силу леммы 2 множество B является η-сетью для A относительно метрики d_0. Таким образом, множество A вполне ограничено в метрике d_0, и следовательно, поскольку D полно в d_0, замыкание A^- множества A компактно.

Мы доказали достаточность условий (14.32) и (14.33). Если множество A^- компактно, то оно ограничено, и поэтому (14.32) выполняется (отметим, что $\sup_t |x(t)|$ есть расстояние в смысле метрики d или метрики d_0 от x до функции, тождественно равной 0). Настоящая теорема отличается от теоремы Арцела — Асколи тем, что ни для какого отдельного значения t из условия

$$\sup_{x \in A} |x(t)| < \infty \tag{14.38}$$

и условия (14.33), вместе взятых, не следует (14.32). Нетрудно, однако, доказать, что (14.32) выполняется, если справедливо (14.33) и если условие (14.38) справедливо для каждого значения t (или даже для множества значений t, включающего 1 и плотного в $[0, 1]$).

Остается доказать необходимость условия (14.33). В силу (14.8) для любого x модуль $w'_x(\delta)$ стремится к 0 вместе с δ. Если мы докажем, что при любом δ функция $w'_x(\delta)$ полунепрерывна сверху в точке x, то (см. стр. 298) эта сходимость будет равномерной на компактных множествах и условие (14.33) будет выполнено.

Пусть x, δ и ε заданы. Для того чтобы доказать полунепрерывность сверху, мы должны найти такое η, что из соотношения $d(x, y) < \eta$ следовало бы

$$w'_y(\delta) < w'_x(\delta) + \varepsilon. \tag{14.39}$$

Выберем сначала точки t_i, удовлетворяющие условию (14.7) и условию

$$w_x[t_{i-1}, t_i) < w'_x(\delta) + \varepsilon/2, \quad i = 1, 2, \ldots, r. \tag{14.40}$$

Возьмем затем η столь малым, что

$$t_i - t_{i-1} > \delta + 2\eta, \quad i = 1, 2, \ldots, r,$$

и

$$\eta < \frac{1}{4}\varepsilon. \tag{14.41}$$

Если $d(x, y) < \eta$, то для некоторой функции λ из Λ имеем

$$\sup_t |\lambda t - t| = \sup_t |\lambda^{-1} t - t| < \eta$$

и

$$\sup_t |y(t) - x(\lambda t)| < \eta. \tag{14.42}$$

Пусть $s_i = \lambda^{-1} t_i$. Тогда

$$s_i - s_{i-1} > t_i - t_{i-1} - 2\eta > \delta. \tag{14.43}$$

Далее, если s и t одновременно принадлежат интервалу $[s_{i-1}, s_i)$, то λs и λt одновременно принадлежат интервалу $[t_{i-1}, t_i)$, и поэтому в силу (14.40), (14.41) и (14.42) $|y(s) - y(t)| < w'_x(\delta) + \varepsilon$. Таким образом,

$$w_y[s_{i-1}, s_i) < w'_x(\delta) + \varepsilon,$$

откуда, ввиду соотношения (14.43), следует (14.39). Этим завершается доказательство теоремы 14.3.

Вторая характеризация компактности. Хотя использование модуля $w'(\delta)$ естественно в том отношении, что с его участием достигается полная характеризация компактных множеств, иногда более удобно иметь дело с другими модулями, например, с

$$w''_x(\delta) = \sup \min \{ |x(t) - x(t_1)|, \ |x(t_2) - x(t)| \}, \tag{14.44}$$

где верхняя грань берется по всем t_1, t и t_2, удовлетворяющим неравенствам

$$t_1 \leqslant t \leqslant t_2, \qquad t_2 - t_1 \leqslant \delta. \tag{14.45}$$

Пусть заданы числа δ и ε. Разобьем интервал $[0, 1)$ на подинтервалы $[s_{i-1}, s_i)$ такие, что $s_i - s_{i-1} > \delta$ и $w_x[s_{i-1}, s_i) < w'_x(\delta) + \varepsilon$. Если условие (14.45) выполняется, то либо t_1 и t_2 лежат в одном и том же подинтервале $[s_{i-1}, s_i)$ и в этом случае

$$|x(t) - x(t_1)| < w'_x(\delta) + \varepsilon, \quad |x(t_2) - x(t)| < w'_x(\delta) + \varepsilon,$$

либо они лежат в смежных интервалах $[s_{i-1}, s_i)$ и $[s_i, s_{i+1})$, и в этом случае $|x(t) - x(t_1)| < w'_x(\delta) + \varepsilon$ при $t_1 \leqslant t < s_i$ и $|x(t_2) - x(t)| < w'_x(\delta) + \varepsilon$ при $s_i \leqslant t \leqslant t_2$. Следовательно,

$$w''_x(\delta) \leqslant w'_x(\delta). \tag{14.46}$$

Противоположное неравенство невозможно. Если, к примеру,

$$x_n(t)=\begin{cases}1 & \text{при} \quad 0\leqslant t<n^{-1},\\ 0 & \text{при} \quad n^{-1}<t\leqslant 1\end{cases} \tag{14.47}$$

или если

$$x_n(t)=\begin{cases}0 & \text{при} \quad 0\leqslant t\leqslant 1-n^{-1},\\ 1 & \text{при} \quad 1-n^{-1}\leqslant t\leqslant 1,\end{cases} \tag{14.48}$$

то $w''_{x_n}(\delta)=0$, тогда как

$$w'_{x_n}(\delta)=\begin{cases}0, & \text{если} \quad \delta<n^{-1},\\ 1, & \text{если} \quad \delta\geqslant n^{-1}.\end{cases}$$

Так как последовательность $\{x_n\}$ не имеет компактного замыкания ни в том, ни в другом случае, то не существует никакого условия компактности, включающего (в добавление к (14.32)) ограничение только на $w''_x(\delta)$. Можно, однако, сформулировать условие компактности в терминах $w''_x(\delta)$ и поведения x вблизи 0 и 1.

Теорема 14.4. *Множество A имеет компактное замыкание в топологии Скорохода тогда и только тогда, когда*

$$\sup_{x\in A}\sup_t|x(t)|<\infty \tag{14.49}$$

и

$$\begin{cases}\lim\limits_{\delta\to 0}\sup\limits_{x\in A} w''_x(\delta)=0,\\ \lim\limits_{\delta\to 0}\sup\limits_{x\in A} w_x[0,\delta)=0,\\ \lim\limits_{\delta\to 0}\sup\limits_{x\in A} w_x[1-\delta,1)=0.\end{cases} \tag{14.50}$$

Доказательство. Достаточно, ввиду утверждения теоремы 14.3, показать, что условие (14.50) эквивалентно соотношению

$$\lim_{\delta\to 0}\sup_{x\in A} w'(\delta)=0. \tag{14.51}$$

Тот факт, что (14.51) влечет (14.50), является следствием неравенства (14.46) и определения $w'(\delta)$. Мы докажем обратную импликацию.

Для данного положительного ε выберем положительное δ так, чтобы для всех x из A

$$w''_x(\delta) < \varepsilon, \quad w_x[0, \delta) < \varepsilon, \quad w_x[1-\delta, 1) < \varepsilon. \tag{14.52}$$

Предположим, что x принадлежат A; покажем, что

$$w'_x\left(\frac{1}{2}\delta\right) \leqslant 6\varepsilon. \tag{14.53}$$

Этого, очевидно, достаточно для доказательства теоремы.

Сначала покажем, что если

$$t_1 \leqslant s \leqslant t \leqslant t_2, \quad t_2 - t_1 \leqslant \delta, \tag{14.54}$$

то

$$\min\{|x(s) - x(t_1)|, |x(t_2) - x(t)|\} < 2\varepsilon. \tag{14.55}$$

Действительно, если $|x(s) - x(t_1)| > \varepsilon$, то в силу (14.52) $|x(t) - x(s)| < \varepsilon$ и $|x(t_2) - x(s)| < \varepsilon$, так что $|x(t_2) - x(t)| < 2\varepsilon$.

Допустим, что x в каждой из двух точек τ_1 и τ_2 имеет скачок, превосходящий 2ε. Если $0 < \tau_2 - \tau_1 < \delta$, то существуют точки t_1, s, t, t_2, удовлетворяющие неравенствам (14.54), такие, что $t_1 < \tau_1 = s$ и $t < \tau_2 = t_2$; если t_1 достаточно близко к τ_1, а t достаточно близко к τ_2, то неравенство (14.55) нарушается, что, как мы видели, невозможно. Таким образом, отрезок $[0, 1]$ не может содержать две точки, находящиеся на расстоянии не более δ одна от другой, в каждой из которых функция x имеет скачки, превосходящие 2ε. Кроме того, в силу (14.52) ни один из интервалов $[0, \delta)$ и $[1-\delta, 1)$ не может содержать точку, в которой скачок функции x превышает 2ε.

Таким образом, существуют точки s_i, $0 = s_0 < s_1 < \ldots < s_r = 1$, такие, что $s_i - s_{i-1} \geqslant \delta$, и такие, что любая точка, в которой x имеет скачок, больший чем 2ε, является одной из точек s_i. Если $s_j - s_{j-1} > \delta$ для пары соседних точек s_{j-1}, s_j, то мы расширим систему $\{s_i\}$, включив в нее точку, лежащую посередине между ними. Этот путь приводит к новой расширенной системе $s_0, \ldots, s_r$ (с новым r), которая удовлетворяет условию

$$\frac{1}{2}\delta < s_i - s_{i-1} \leqslant \delta, \quad i = 1, 2, \ldots, r.$$

Если мы теперь покажем, что при любом значении i

$$w_x[s_{i-1}, s_i) \leqslant 6\varepsilon, \tag{14.56}$$

то получим (14.53). Предположим, что $s_{i-1} \leqslant t_1 < t_2 < s_i$. Тогда $t_2 - t_1 < \delta$. Обозначим через σ_1 верхнюю грань тех σ из $[t_1, t_2]$, для которых $\sup\limits_{t_1 \leqslant u \leqslant \sigma} |x(u) - x(t_1)| \leqslant 2\varepsilon$; пусть σ_2 — нижняя грань тех значений σ из $[t_1, t_2]$, для которых $\sup\limits_{\sigma \leqslant u \leqslant t_2} |x(t_2) - x(u)| \leqslant 2\varepsilon$. Если $\sigma_1 < \sigma_2$, то существуют точки s справа от точки σ_1 и сколь угодно близкие к ней, для которых $|x(s) - x(t_1)| > 2\varepsilon$, и существуют точки t слева от точки σ_2 и сколь угодно близкие к ней, для которых $|x(t_2) - x(t)| > 2\varepsilon$; так как мы можем выбрать s, t так, что $s < t$, то мы получаем противоречие с тем фактом, что (14.54) влечет (14.55). Поэтому $\sigma_2 \leqslant \sigma_1$, и отсюда следует, что

$$|x(\sigma_1 -) - x(t_1)| \leqslant 2\varepsilon \quad \text{и} \quad |x(t_2) - x(\sigma_1)| \leqslant 2\varepsilon.$$

Поскольку $t_1 < \sigma_1 \leqslant t_2$, то σ_1 является внутренней точкой интервала (s_{i-1}, s_i), поэтому скачок в σ_1 не превосходит 2ε. Таким образом, $|x(t_2) - x(t_1)| \leqslant 6\varepsilon$. Отсюда следует неравенство (14.56), которое, как отмечалось, доказывает (14.53), а вместе с тем и всю теорему*).

Конечномерные множества. Конечномерные множества играют в пространстве D ту же роль, что и в пространстве C. Для набора точек $t_1, \ldots, t_k$ из $[0, 1]$ определим естественную проекцию $\pi_{t_1 \ldots t_k}$ из D в R^k обычным способом:

$$\pi_{t_1 \ldots t_k}(x) = (x(t_1), \ldots, x(t_k)). \tag{14.57}$$

Очевидно, π_0 и π_1 — непрерывные всюду функции. Пусть $0 < t < 1$. Если последовательность x_n сходится к x в топологии Скорохода и функция x непрерывна в точке t, то (стр. 157) $x_n(t) \to x(t)$. Предположим, с другой стороны, что x имеет разрыв в точке t. Если λ_n является функцией из Λ, линейной на отрезках $[0, t]$ и $[t, 1]$ и удовлетворяющей условию $\lambda_n t = t - 1/n$, и если $x_n(s) \equiv x(\lambda_n s)$, то x_n сходится к x в топологии Скорохода, но $x_n(t)$ не сходится к $x(t)$. Таким образом: *если*

---

*) Точки 0 и 1 играют особую роль в теории пространства D вследствие того, что каждая функция λ из Λ оставляет их неподвижными. Разумным расширением класса Λ можно было бы упростить теорию.

$0 < t < 1$, *то* π_t *непрерывна в* x *тогда и только тогда, когда* x *непрерывна в* t.

Мы должны доказать, что проекция $\pi_{t_1 \ldots t_k}$ измерима относительно σ-поля $\mathcal{D}$ борелевских множеств в топологии Скорохода. Достаточно рассмотреть какое-либо отдельное значение t (так как отображение в пространстве R^k измеримо, если измерима каждая его компонента), и мы можем допустить, что $t < 1$ (поскольку π_1 непрерывна). Если x_n сходится к x в топологии Скорохода, то $x_n(s) \to x(s)$ в точках непрерывности s функции x и, следовательно, во всех точках s, за исключением множества лебеговой меры 0. Поскольку x_n равномерно ограничены, то для любого положительного ε

$$\frac{1}{\varepsilon} \int\limits_t^{t+\varepsilon} x_n(s)\, ds \to \frac{1}{\varepsilon} \int\limits_t^{t+\varepsilon} x(s)\, ds, \qquad n \to \infty. \tag{14.58}$$

Таким образом, функция $h_\varepsilon(x) = \varepsilon^{-1} \int\limits_t^{t+\varepsilon} x(s)\, ds$ непрерывна в топологии Скорохода. В силу непрерывности x справа $h_\varepsilon(x) \to \pi_t(x)$ для любого x при $\varepsilon \to 0$. Таким образом, функция $\pi_t(x)$ измерима.

Установив измеримость $\pi_{t_1 \ldots t_k}$, мы можем так же, как в случае пространства C, определить конечномерные множества как множества вида $\pi^{-1}_{t_1 \ldots t_k} H$, где $k \geqslant 1$ и $H \in \mathcal{R}^k$. По определению измеримости любое конечномерное множество принадлежит $\mathcal{D}$.

Пусть T_0 — подмножество отрезка $[0, 1]$. Обозначим $\mathcal{F}_{T_0}$ класс множеств вида $\pi^{-1}_{t_1 \ldots t_k} H$, где k произвольно, t_i суть произвольные точки множества T_0 и $H \in \mathcal{R}^k$. Класс $\mathcal{F}_{T_0}$ есть (конечно аддитивное) поле. Разумеется, $\mathcal{F}_{[0, 1]}$ является классом конечномерных множеств.

Теорема 14.5. *Если множество* T_0 *содержит* 1 *и плотно в* $[0, 1]$, *то* $\mathcal{F}_{T_0}$ *порождает* $\mathcal{D}$.

Эти условия являются также и необходимыми для того, чтобы класс $\mathcal{F}_{T_0}$ порождал $\mathcal{D}$, однако нам этот факт не понадобится. Выбор в качестве T_0 отрезка $[0, 1]$ не облегчает доказательства теоремы.

Доказательство. Поскольку пространство D сепарабельно, достаточно показать, что любой открытый d_0-шар $S_{d_0}(x, r)$ принадлежит σ-полю, порожденному $\mathcal{F}_{T_0}$. Фиксируем центр x и радиус r.

Выберем в T_0 последовательность точек $t_1, t_2, \ldots$, плотную в $[0, 1]$, так, чтобы $t_1 = 1$. При $0 < \varepsilon < r$ и $k \geqslant 1$ пусть множество $A_k(\varepsilon)$ состоит из таких y, для которых в Λ существует функция λ, удовлетворяющая условиям

$$\|\lambda\| < r - \varepsilon$$

и

$$\max_{1 \leqslant i \leqslant k} |y(t_i) - x(\lambda t_i)| < r - \varepsilon.$$

Достаточно доказать, что

$$S_{d_0}(x, r) = \bigcup_{\varepsilon} \bigcap_{k=1}^{\infty} A_k(\varepsilon), \tag{14.59}$$

где суммирование происходит по рациональным ε из $(0, r)$, и что каждое $A_k(\varepsilon)$ принадлежит $\mathcal{F}_{T_0}$.

Мы докажем сначала второй факт. При фиксированных ε и k пусть H_1 обозначает множество точек вида $(x(\lambda t_1), \ldots, x(\lambda t_k))$ в R^k, где λ является произвольной функцией из Λ, удовлетворяющей условию $\|\lambda\| < r - \varepsilon$. Пусть H_2 — множество точек $(\alpha_1, \ldots, \alpha_k)$ в R^k таких, что $|\alpha_i - \beta| < r - \varepsilon$, $i = 1, \ldots, k$, для некоторой точки $(\beta_1, \ldots, \beta_k)$ из H_1. Множество H_2 открыто и, следовательно, принадлежит $\mathcal{R}^k$, и $A_k(\varepsilon) = \pi^{-1}_{t_1 \ldots t_k} H_2$. Таким образом, $A_k(\varepsilon) \in \mathcal{F}_{T_0}$.

Легко обнаружить, что левая часть (14.59) содержится в правой части. Мы завершим доказательство, если покажем, что

$$\bigcap_{k=1}^{\infty} A_k(\varepsilon) \subset S_{d_0}(x, r). \tag{14.60}$$

Пусть y принадлежит пересечению в левой части (14.60). Выберем для каждого k функцию λ_k в Λ, удовлетворяющую условиям

$$\|\lambda_k\| < r - \varepsilon \tag{14.61}$$

и

$$\max_{1 \leqslant i \leqslant k} |y(t_i) - x(\lambda_k t_i)| < r - \varepsilon. \tag{14.62}$$

По теореме Хелли о выборе (стр. 311) существуют подпоследовательность $\{\lambda_{k'}\}$ и неубывающая функция λ такие, что

$$\lim_{k'} \lambda_{k'}t = \lambda t \tag{14.63}$$

для всех точек непрерывности t функции λ. Мы покажем, что $\lambda \in \Lambda$, $\|\lambda\| \leqslant r - \varepsilon$ и $\sup_t |y(t) - x(\lambda t)| \leqslant r - \varepsilon$. Отсюда будет вытекать, что $d_0(x, y) < r$, и следовательно, (14.60) будет доказано.

Если s и t — две различные точки непрерывности функции λ, то в силу (14.61)

$$\left| \log \frac{\lambda t - \lambda s}{t - s} \right| = \lim_{k'} \left| \log \frac{\lambda_{k'}t - \lambda_{k'}s}{t - s} \right| \leqslant r - \varepsilon. \tag{14.64}$$

Это соотношение показывает, что λ не имеет скачков (в частности, (14.63) справедливо при всех t), и отсюда следует, что λ — строго возрастающая функция; таким образом, $\lambda \in \Lambda$. Из неравенства (14.64), кроме того, вытекает, что $\|\lambda\| \leqslant r - \varepsilon$. В силу (14.62) $|y(t_1) - x(t_1)| < r - \varepsilon$ (напомним, что $t_1 = 1$). Если $i > 1$, то согласно (14.62) $|y(t_i) - x(\lambda_{k'} t_i)| < r - \varepsilon$ при $k' \geqslant i$. Поскольку $\lambda_{k'}t \to \lambda t$, мы получаем, что либо $|y(t_i) - x(\lambda t_i)| \leqslant r - \varepsilon$, либо $|y(t_i) - x((\lambda t_i) -)| \leqslant r - \varepsilon$. Отсюда, так как последовательность $\{t_i\}$ всюду плотна, вытекает, что $\sup_t |y(t) - x(\lambda t)| \leqslant r - \varepsilon$. Это доказывает теорему 14.5.

Так как класс $\mathcal{F}_T$ является полем, то из доказанной теоремы вытекает, что если множество T_0 содержит 1 и плотно в $[0, 1]$, то класс $\mathcal{F}_{T_0}$ является определяющим классом. Но даже $\mathcal{F}_{[0,1]}$ не является определяющим сходимость классом — противоречащие примеры для C годятся и в рассматриваемом случае.

Если P — вероятностная мера на $(D, \mathscr{D})$, то ее конечномерными распределениями являются меры $P\pi^{-1}_{t_1 \ldots t_k}$. Если множество T_0 содержит 1 и всюду плотно в $[0, 1]$, то мера P полностью определяется своими конечномерными распределениями для точек из T_0.

Мы будем использовать в D те же соглашения о координатных величинах, которые мы использовали в C (см. конец § 8).

Примечание. Изучаемая здесь топология представляет собой J_1-топологию Скорохода (1956); см. также Колмогоров (1956), Прохоров (1956) и Скороход (1961).

Задачи

1. Пусть D^+ — класс функций, определенных на отрезке [0, 1], которые могут иметь разрывы только первого рода в интервале (0, 1) и имеют правосторонние пределы в 0 и левосторонние пределы в 1. Превратите D^+ в псевдометрическое пространство таким образом, что (i) расстояние между x и y равно 0 тогда и только тогда, когда x и y совпадают в их общих точках непрерывности, и (ii) D изометрично стандартному факторпространству [Келли (1968, стр. 136)].

2. Топология Скорохода сильнее топологии, которая задается метрикой $\int |x(t) - y(t)|\,dt$.

3. По отношению к топологии Скорохода и обычному сложению функций пространство D не является топологической группой.

4. Множество C нигде не плотно в D.

5. Положим

$$w_x'''(\delta) = \sup_{t_2 - t_1 < \delta} \sup_{t_1 < t < t_2} \min\{w_x(t_1, t),\ w_x(t, t_2)\}.$$

Покажите, что

$$w_x'''(\delta) = \sup_{t_2 - t_1 < \delta} \inf_{t_1 < t < t_2} \max\{w_x(t_1, t),\ w_x(t, t_2)\}$$

и

$$w_x''(\delta) \leqslant w_x'''(\delta) \leqslant 2w_x''(\delta).$$

§ 15. Слабая сходимость и плотность в D

Доказательство слабой сходимости в функциональном пространстве в два этапа: во-первых, доказательство слабой сходимости конечномерных распределений и, во-вторых, доказательство плотности — вот путь, который оправдал себя в пространстве C. Желательно было бы приспособить эту технику к пространству D. Поскольку D сепарабельно и полно относительно метрики d_0, то семейство вероятностных мер на $(D, \mathscr{D})$ относительно компактно тогда и только тогда, когда оно плотно, так что этот вопрос не вызывает трудностей. С другой стороны, тот факт, что естественные проекции не являются непрерывными, несколько осложняет дело.

Конечномерные распределения. Пусть для вероятностной меры P на $(D, \mathscr{D})$ множество T_P состоит из таких точек t отрезка [0, 1], для которых проекция π_t непрерывна всюду, за исключением точек, составляющих множество P-меры 0. Точки 0 и 1 всегда принадлежат T_P. Если $0 < t < 1$, то $t \in T_P$ тогда и только тогда, когда $P(J_t) = 0$, где множество

$$J_t = \{x\colon\ x(t) \neq x(t-)\} \tag{15.1}$$

представляет собой множество тех x, которые разрывны в t [напомним (стр. 169), что при $0 < t < 1$ проекция π_t непрерывна в x тогда и только тогда, когда функция x непрерывна в t].

Любой элемент пространства D имеет самое бо́льшее счетное множество скачков. Докажем аналогичный факт, а именно, что неравенство $P(J_t) > 0$ выполняется для не более чем счетного множества точек t. Пусть $J_t(\varepsilon)$ — множество x, имеющих в t скачок $|x(t) - x(t-)|$, превосходящий ε. Для фиксированных положительных ε и δ может существовать самое бо́льшее конечное множество точек t, для которых $P(J_t(\varepsilon)) \geqslant \delta$, так как если бы это неравенство выполнялось для последовательности различных точек t_1, t_2, ..., то множество $\limsup_n J_{t_n}(\varepsilon)$ имело бы меру, не превосходящую δ, и следовательно, было бы непустым, но это противоречит тому факту, что у всякого элемента x скачок может превосходить ε только в конечном числе точек. Поэтому при фиксированном положительном ε $P(J_t(\varepsilon))$ может превосходить 0 не более, чем в счетном множестве точек t. Поскольку $P(J_t(\varepsilon)) \uparrow P(J_t)$ при $\varepsilon \downarrow 0$, то мы получаем нужный результат.

Таким образом, *множество* T_P *содержит* 0 *и* 1, *и его дополнение в* [0, 1] *не более чем счетно.* Если $t_1, \ldots, t_k$ все принадлежат T_P, то $\pi_{t_1 \ldots t_k}$ непрерывна всюду, за исключением множества P-меры 0.

Предположим, что

$$P_n \Rightarrow P, \tag{15.2}$$

где P_n и P — вероятностные меры на $(D, \mathscr{D})$; по теореме 5.1 имеем

$$P_n \pi^{-1}_{t_1 \ldots t_k} \Rightarrow P \pi^{-1}_{t_1 \ldots t_k}, \tag{15.3}$$

если все t_i лежат в T_P. В общем случае (15.3) не будет следовать из (15.2), если некоторые из точек t_i лежат вне T_P. Например, если P — единичная масса, сосредоточенная на функции $I_{[0, 1/2]}$, а P_n — единичная масса, сосредоточенная на функции $I_{[0, 1/2+n^{-1}]}$, то $P_n \Rightarrow P$, в то время как $P_n\pi^{-1}$ не сходится к $P\pi^{-1}_{1/2}$.

Докажем теперь аналог теоремы 8.1.

Теорема 15.1. *Если семейство $\{P_n\}$ плотно и если $P_n\pi^{-1}_{t_1 \ldots t_k} \Rightarrow P\pi^{-1}_{t_1 \ldots t_k}$ всякий раз, когда $t_1, \ldots, t_k$ все принадлежат T_P, то $P_n \Rightarrow P$.*

Доказательство. Ввиду плотности семейства $\{P_n\}$ каждая подпоследовательность $\{P_{n'}\}$ содержит другую подпоследовательность $\{P_{n''}\}$, слабо сходящуюся к некоторому пределу Q. Достаточно (теорема 2.3) показать, что все такие пределы Q совпадают с P.

Если точки $t_1, \ldots, t_k$ все принадлежат множеству $T_P \cap T_Q$, то по предположению теоремы $P_{n''}\pi^{-1}_{t_1 \ldots t_k} \Rightarrow$ $\Rightarrow P\pi^{-1}_{t_1 \ldots t_k}$ и, кроме того, $P_{n''}\pi^{-1}_{t_1 \ldots t_k} \Rightarrow Q\pi^{-1}_{t_1 \ldots t_k}$, так как $P_{n''} \Rightarrow Q$. Таким образом,

$$P\pi^{-1}_{t_1 \ldots t_k} = Q\pi^{-1}_{t_1 \ldots t_k}, \quad t_1, \ldots, t_k \in T_P \cap T_Q. \qquad (15.4)$$

Так как каждое из множеств T_P и T_Q содержит все точки отрезка $[0, 1]$, за исключением, самое большее, счетного их числа, то же самое верно для множества $T_P \cap T_Q$; в частности, это пересечение является всюду плотным множеством в $[0, 1]$. Поскольку $T_P \cap T_Q$ содержит 1, из теоремы 14.5 следует, что $\mathscr{D}$ порождается конечномерными множествами, построенными на точках из $T_P \cap T_Q$. Равенство $P = Q$ вытекает теперь из (15.4), что и завершает доказательство.

Плотность. Исследование проблемы плотности в C исходным пунктом имело утверждение теоремы 8.2, которое заменяло компактность ее характеризацией Арцела — Асколи. Теорема 14.3, содержащая характеризацию компактности в D, приводит тем же самым путем к следующему результату. Пусть $\{P_n\}$ — последовательность вероятностных мер на $(D, \mathscr{D})$.

Теорема 15.2. *Последовательность $\{P_n\}$ плотна тогда и только тогда, когда выполняются следующие два условия:*

(i) *Для любого положительного* η *существует такое* a, *что*

$$P_n\{x:\ \sup_t |x(t)| > a\} \leqslant \eta, \qquad n \geqslant 1. \tag{15.5}$$

(ii) *Для любых положительных* ε *и* η *существуют* δ, $0 < \delta < 1$, *и целое число* n_0 *такие, что*

$$P_n\{x:\ w'_x(\delta) \geqslant \varepsilon\} \leqslant \eta, \qquad n \geqslant n_0. \tag{15.6}$$

Используя теорему 14.4 вместо теоремы 14.3, получаем еще одну совокупность условий для плотности.

Теорема 15.3. *Последовательность* $\{P_n\}$ *плотна тогда и только тогда, когда выполняются следующие два условия:*

(i) *Для любого положительного* η *существует такое* a, *что*

$$P_n\{x:\ \sup_t |x(t)| > a\} \leqslant \eta, \qquad n \geqslant 1.$$

(ii) *Для любых положительных* ε *и* η *существуют* δ, $0 < \delta < 1$, *и целое число* n_0 *такие, что*

$$P_n\{x:\ w''_x(\delta) \geqslant \varepsilon\} \leqslant \eta, \qquad n \geqslant n_0, \tag{15.7}$$

$$P_n\{x:\ w_x[0,\ \delta) \geqslant \varepsilon\} \leqslant \eta, \qquad n \geqslant n_0 \tag{15.8}$$

и

$$P_n\{x:\ w_x[1-\delta,\ 1) \geqslant \varepsilon\} \leqslant \eta, \qquad n \geqslant n_0. \tag{15.9}$$

Как показывает следующая теорема, в наиболее интересных случаях условия (15.8) и (15.9) автоматически выполняются. Пусть P_n и P — вероятностные меры на $(D, \mathscr{D})$; согласно определению (15.1) $J_1 = \{x:\ x(1) \neq x(1-)\}$.

Теорема 15.4. *Предположим, что*

$$P_n\pi^{-1}_{t_1 \ldots t_k} \Rightarrow P\pi^{-1}_{t_1 \ldots t_k} \tag{15.10}$$

всякий раз, когда точки $t_1, \ldots, t_k$ *все принадлежат* T_P. *Допустим далее, что* $P(J_1) = 0$. *Предположим, наконец, что для любых положительных* ε *и* η *существуют* δ, $0 < \delta < 1$, *и целое число* n_0 *такие, что*

$$P_n\{x:\ w''_x(\delta) \geqslant \varepsilon\} \leqslant \eta, \qquad n \geqslant n_0. \tag{15.11}$$

Тогда $P_n \Rightarrow P$.

Доказательство. В силу теоремы 15.1 достаточно показать, что последовательность $\{P_n\}$ плотна.

Проверим сначала условие (i) предыдущей теоремы. По заданному положительному η выберем δ и n_0, удовлетворяющие (15.11) с $\varepsilon = 1$ (например). Поскольку множество T_P всюду плотно в $[0, 1]$, оно содержит точки $t_1, \ldots$ $\ldots, t_k$, $0 = t_1 < \ldots < t_k = 1$, такие, что $t_i - t_{i-1} < \delta$. Из (15.10) вытекает, что последовательность $\{P_n \pi_{t_1 \ldots t_k}^{-1}\}$ является плотной в R^k, и следовательно, при надлежащим образом выбранном a_0

$$P_n\{x\colon \max_{1 \leqslant i \leqslant k} |x(t_i)| > a_0\} \leqslant \eta, \qquad n \geqslant 1. \tag{15.12}$$

Если $|x(t_i)| \leqslant a_0$, $i = 1, \ldots, k$, и если $w''_x(\delta) < 1$, то $|x(t)| < a_0 + 1$ при всех t. Если $a = a_0 + 1$, то из (15.12) и (15.11) (с $\varepsilon = 1$) мы имеем $P_n\{x\colon \sup_t |x(t)| > a\} \leqslant 2\eta$ при $n \geqslant n_0$. Для конечного множества значений n, предшествующих n_0, мы можем добиться выполнения этого неравенства надлежащим увеличением a, так как каждое отдельное распределение P_n плотно. Тем самым доказано выполнение условия (i).

Для доказательства условия (ii) предыдущей теоремы по заданным ε и η нужно найти δ и n_0, удовлетворяющие (15.8) и (15.9). Рассмотрим сначала (15.8). Так как каждый элемент x из D непрерывен справа, для всех достаточно малых δ мы имеем

$$P\{x\colon |x(\delta) - x(0)| \geqslant \varepsilon\} < \eta. \tag{15.13}$$

Выберем в плотном множестве T_P достаточно малое δ такое, чтобы выполнялось (15.13) и чтобы при соответствующем n_0 выполнялось (15.11). Так как и 0, и δ принадлежат T_P, то мы получаем соотношение $P_n\pi_{0,\delta}^{-1} \Rightarrow P\pi_{0,\delta}^{-1}$ как частный случай (15.10) и из (15.13) следует, что для всех достаточно больших n

$$P_n\{x\colon |x(\delta) - x(0)| \geqslant \varepsilon\} < \eta. \tag{15.14}$$

Если $w''_x(\delta) < \varepsilon$ и $|x(\delta) - x(0)| < \varepsilon$, то $|x(s) - x(0)| < 2\varepsilon$ для всех s из $[0, \delta]$, и следовательно, $w_x[0, \delta] < 4\varepsilon$. Таким образом, из (15.11) и (15.14) вытекает неравенство $P_n\{x\colon w_x[0, \delta) \geqslant 4\varepsilon\} \leqslant 2\eta$. Таким образом, (15.8) выполнено.

С одним лишь изменением те же доводы годятся по симметрии и для вывода (15.9). Вместо (15.13) нам теперь требуется неравенство

$$P\{x\colon |x(1)-x(1-\delta)|\geqslant\varepsilon\}<\eta \tag{15.15}$$

для малых δ. Так как непрерывность слева элементов пространства D не предполагается, то неравенство (15.15) в общем случае не выполняется. Так как мы предположили, что $P(J_1)=0$, то функция x непрерывна слева в точке $t=1$, за исключением x, принадлежащих множеству P-меры 0, так что (15.15) действительно выполняется при всех достаточно малых δ.

Этим завершается доказательство теоремы 15.4. Условие $P(J_1)=0$ существенно: если P — единичная масса, сосредоточенная на функции $I_{\{1\}}$, и P_n — единичная масса, сосредоточенная на функции $I_{[1-n^{-1},\,1]}$, то все конечномерные распределения сходятся и условие, связанное с (15.11), справедливо, но P_n тем не менее не сходится слабо к P.

Поскольку

$$\lim_{t\to 1} P\{x\colon |x(1-)-x(t)|>\varepsilon\}=0,\quad \varepsilon>0,$$

то $P(J_1)=0$ тогда и только тогда, когда

$$\lim_{t\to 1} P\{x\colon |x(1)-x(t)|>\varepsilon\}=0,\quad \varepsilon>0. \tag{15.16}$$

Обычно условие $P(J_1)=0$ легко проверить с помощью (15.16).

Наш следующий результат показывает, что происходит, когда вместо $w'_x(\delta)$ или $w''_x(\delta)$ используется модуль $w_x(\delta)$, соответствующий пространству C.

Теорема 15.5. *Допустим, что при каждом положительном* η *существует* a *такое, что*

$$P_n\{x\colon |x(0)|>a\}\leqslant\eta,\qquad n\geqslant 1. \tag{15.17}$$

Допустим далее, что при любых положительных ε *и* η *существуют* δ, $0<\delta<1$, *и целое число* n_0 *такие, что*

$$P_n\{x\colon w_x(\delta)\geqslant\varepsilon\}\leqslant\eta,\qquad n\geqslant n_0. \tag{15.18}$$

Тогда последовательность $\{P_n\}$ *плотна, и если* P — *слабый предел подпоследовательности* $\{P_{n'}\}$, *то* $P(C)=1$, Доказательство. Поскольку $w'_x(\delta) \leqslant w_x(2\delta)$ при $\delta < 1/2$ (см. (14.9)), то выполняется условие (ii) теоремы 15.2. Справедливость условия (i) этой же теоремы легко следует из предположений доказываемой теоремы и неравенства

$$|x(t)| \leqslant |x(0)| + \sum_{i=1}^{k} |x(it/k) - x((i-1)t/k)|.$$

Следовательно, последовательность $\{P_n\}$ плотна.

Если $w_y\left(\frac{1}{2}\delta\right) \geqslant 2\varepsilon$, то y является внутренней точкой множества $\{x\colon w_x(\delta) \geqslant \varepsilon\}$; поэтому из сходимости $P_n \Rightarrow P$ вытекает, что

$$P\left\{y\colon w_y\left(\frac{1}{2}\delta\right) \geqslant 2\varepsilon\right\} \leqslant \liminf_{n'} P_{n'}\{x\colon w_x(\delta) \geqslant \varepsilon\}. \quad (15.19)$$

По данным ε и η выберем δ и n_0 так, чтобы выполнялось (15.18); в силу (15.19) тогда $P\left\{y\colon w_y\left(\frac{1}{2}\delta\right) \geqslant 2\varepsilon\right\} \leqslant \leqslant \eta$. Для каждого k существует, следовательно, положительное δ_k такое, что если $A_k = \{y\colon w_y(\delta_k) \geqslant 1/k\}$, то $P(A_k) < 1/k$. Положим $A = \liminf_k A_k$; тогда $P(A) = 0$, и если $x \notin A$, то $\lim\limits_{\delta \to 0} w_x(\delta) = 0$, откуда $P(C) = 1$, что и доказывает теорему.

В приложениях условие (15.18) проверяется обычно с помощью следствия к теореме 8.3, которое переносится на случай пространства D без изменений. Заметим, что если n_0 в (15.18) равно 1 для любых ε и η, то $P_n(C) = 1$ при всех n.

Случайные элементы пространства D. Случайный элемент пространства D мы будем часто называть случайной функцией. При этом из контекста всегда будет ясно, является ли рассматриваемая случайная функция случайным элементом C или D.

Если X отображает пространство $(\Omega, \mathscr{B}, \mathrm{P})$ в D, то при любом ω $X(\omega)$ является элементом D, значение которого в точке t мы обозначаем через $X(t, \omega)$. При каждом t обозначим $X(t)$ действительную функцию $\pi_t X$, определенную на Ω; значение этой функции в точке ω равно $X(t, \omega)$. Точно так же, как в случае пространства C, X является случайным элементом $D(X^{-1}\mathscr{D}\subset\mathscr{B})$ тогда и только тогда, когда при каждом t $X(t)$ является случайной величиной (это следует из теоремы 14.5).

Последовательность $\{X_n\}$ случайных элементов пространства D называется плотной, если последовательность соответствующих распределений является плотной. Каждая из теорем 15.1—15.5 может быть легко переформулирована в терминах случайных функций.

Критерий сходимости. Объединим результаты этого параграфа с результатами § 12 для получения конкретного критерия сходимости по распределению.

Пусть X_n и X — случайные элементы пространства D. Будем писать T_X вместо T_P, где P — распределение X. Таким образом, T_X содержит 0 и 1, и если $0<t<1$, то $t\in T_X$ тогда и только тогда, когда $\mathsf{P}\{X(t)\neq X(t-)\}=0$. В дальнейшем будем писать $w''(X,\delta)$ вместо $w''_x(\delta)$.

Теорема 15.6. *Пусть*

$$(X_n(t_1),\ \ldots,\ X_n(t_k))\xrightarrow{\mathscr{D}}(X(t_1),\ \ldots,\ X(t_k))\qquad(15.20)$$

всякий раз, когда $t_1,\ \ldots,\ t_k$ все принадлежат T_X; пусть $\mathsf{P}\{X(1)\neq X(1-)\}=0$, и пусть

$$\mathsf{P}\{|X_n(t)-X_n(t_1)|\geqslant\lambda,\ |X_n(t_2)-X_n(t)|\geqslant\lambda\}\leqslant\leqslant\frac{1}{\lambda^{2\gamma}}[F(t_2)-F(t_1)]^{2\alpha}\qquad(15.21)$$

при $t_1\leqslant t\leqslant t_2$ и $n\geqslant1$, где $\gamma\geqslant0$, $\alpha>1/2$ и F — неубывающая непрерывная функция на $[0,1]$. Тогда $X_n\xrightarrow{\mathscr{D}}X$.

Можно указать более ограничительный вариант условия (15.21) в терминах моментов, а именно,

$$\mathsf{E}\{|X_n(t)-X_n(t_1)|^{\gamma}|X_n(t_2)-X_n(t)|^{\gamma}\}\leqslant[F(t_2)-F(t_1)]^{2\alpha}.$$

Доказательство. В силу теоремы 15.4 достаточно показать, что для любых положительных ε и η найдется положительное δ такое, что

$$\mathsf{P}\{w''(X_n,\ \delta)\geqslant\varepsilon\}\leqslant\eta\qquad(15.22)$$

для всех n. При фиксированных τ, δ и n для всякого положительного целого m рассмотрим случайные величины

$$\xi_i = X_n\left(\tau + \frac{i}{m}\delta\right) - X_n\left(\tau + \frac{i-1}{m}\delta\right), \qquad (15.23)$$
$$i = 1, 2, \ldots, m.$$

Пусть

$$M''_m = \max\min\left\{\left|X_n\left(\tau+\frac{j}{m}\delta\right) - X_n\left(\tau+\frac{i}{m}\delta\right)\right|, \left|X_n\left(\tau+\frac{k}{m}\delta\right) - X_n\left(\tau+\frac{j}{m}\delta\right)\right|\right\},$$

где максимум берется по i, j, k таким, что $0 \leqslant i \leqslant j \leqslant \leqslant k \leqslant m$. В силу (15.21) и теоремы 12.5

$$\mathsf{P}\{M''_m \geqslant \varepsilon\} \leqslant \frac{K}{\varepsilon^{2\gamma}}[F(\tau+\delta) - F(\tau)]^{2\alpha}, \qquad (15.24)$$

где $K = K''_{\gamma,\alpha}$ зависит только от γ и α.

Положим

$$w''(X_n, [\tau, \tau+\delta]) = \sup\min\{|X_n(t) - X_n(t_1)|, |X_n(t_2) - X_n(t)|\}, \qquad (15.25)$$

где верхняя грань берется по t_1, t, t_2, удовлетворяющим соотношению $\tau \leqslant t_1 \leqslant t \leqslant t_2 \leqslant \tau+\delta$. Поскольку X_n является функцией, непрерывной справа на $[0,1]$, то, устремляя m к ∞ в (15.24), получаем

$$\mathsf{P}\{w''(X_n, [\tau, \tau+\delta]) > \varepsilon\} \leqslant \frac{K}{\varepsilon^{2\gamma}}[F(\tau+\delta) - F(\tau)]^{2\alpha}. \qquad (15.26)$$

Допустим теперь, что $\delta = 1/(2u)$ есть величина, обратная четному целому числу, и допустим, что

$$w''(X_n, [2i\delta, (2i+2)\delta]) \leqslant \varepsilon, \quad 0 \leqslant i \leqslant u-1, \qquad (15.27)$$

и

$$w''(X_n, [(2i+1)\delta, (2i+3)\delta]) \leqslant \varepsilon, \quad 0 \leqslant i \leqslant u-2. \qquad (15.28)$$

Если $t_1 \leqslant t \leqslant t_2$ и $t_2 - t_1 \leqslant \delta$, то обе точки t_1 и t_2 лежат в некотором из $2u-1$ интервалов, указанных в (15.27) и (15.28), так что

$$\min\{|X_n(t) - X_n(t_1)|, |X_n(t_2) - X_n(t)|\} \leqslant \varepsilon.$$

Таким образом, из (15.27) и (15.28) вытекает, что $w''(X_n, \delta) \leqslant \varepsilon$. Отсюда в силу (15.26)

$$\mathsf{P}\{w''(X_n, \delta) \geqslant \varepsilon\} \leqslant \frac{K}{\varepsilon^{2\gamma}} \left(\sum{}' + \sum{}''\right), \tag{15.29}$$

где каждая из сумм $\sum'$ и $\sum''$ имеет вид

$$\sum_{k=1}^{r} [F(t_k) - F(t_{k-1})]^{2\alpha}$$

с $0 \leqslant t_1 \leqslant \ldots \leqslant t_r \leqslant 1$ и $t_k - t_{k-1} \leqslant 2\delta$. Но тогда

$$\mathsf{P}\{w''(X_n, \delta) \geqslant \varepsilon\} \leqslant \frac{2K}{\varepsilon^{2\gamma}} [F(1) - F(0)] [w_F(2\delta)]^{2\alpha-1}. \tag{15.30}$$

Поскольку $2\alpha > 1$ и функция F непрерывна, то правая часть последнего неравенства стремится к 0 вместе с δ, что и доказывает (15.22).

Критерий существования*). Пользуясь развитыми выше идеями, можно также получить условие существования в D случайного элемента с заданными конечномерными распределениями. Пусть, так же как в теореме 12.4, $\mu_{t_1 \ldots t_k}$ для каждого набора $t_1, \ldots, t_k$ точек из $[0, 1]$ есть вероятностная мера на $(R^k, \mathscr{R}^k)$, и предположим, что эти меры удовлетворяют условиям согласованности теоремы Колмогорова (о существовании меры).

Теорема 15.7. *В пространстве D существует случайный элемент с конечномерными распределениями $\mu_{t_1 \ldots t_k}$ при условии, что, во-первых, эти распределения согласованы; во-вторых,*

$$\mu_{t_1 t t_2}\{(\beta_1, \beta, \beta_2) : |\beta - \beta_1| \geqslant \lambda, \ |\beta_2 - \beta| \geqslant \lambda\} \leqslant$$
$$\leqslant \frac{1}{\lambda^{2\gamma}} [F(t_2) - F(t_1)]^{2\alpha} \tag{15.31}$$

для $t_1 \leqslant t \leqslant t_2$, где $\gamma \geqslant 0$, $\alpha > 1/2$ и F — неубывающая непрерывная функция на $[0, 1]$; и, в-третьих,

$$\lim_{h \downarrow 0} \mu_{t, t+h}\{(\beta_1, \beta_2) : |\beta_1 - \beta_2| \geqslant \varepsilon\} = 0, \quad 0 \leqslant t < 1. \tag{15.32}$$

---

*) Эта тема может быть пропущена.

Можно указать более ограничительный вариант условия (15.31):

$$\int\limits_{R^3} |\beta - \beta_1|^{\gamma} |\beta_2 - \beta|^{\gamma} d\mu_{t_1 t t_2}(\beta_1, \beta, \beta_2) \leqslant [F(t_2) - F(t_1)]^{2\alpha}.$$

Доказательство. Для произвольного числа n легко построить случайную функцию X_n, которая постоянна на каждом интервале $[(i-1)2^{-n}, i2^{-n})$ и для которой распределение случайного вектора

$$\left(X_n(0), X_n\left(\frac{1}{2^n}\right), \ldots, X_n\left(\frac{2^n-1}{2^n}\right), X_n(1)\right)$$

есть $\mu_{t_0 \ldots t_k}$ с $k = 2^n$ и $t_i = i2^{-n}$. Мы докажем, что последовательность $\{X_n\}$ плотна. Сначала покажем, что при заданных ε и η существует δ такое, что для достаточно больших n

$$P\{w''(X_n, \delta) \geqslant \varepsilon\} \leqslant \eta. \tag{15.33}$$

Если значения t_1, t и t_2 являются целыми кратными 2^{-n} и $t_1 \leqslant t \leqslant t_2$, то в силу (15.31) имеем

$$P\{|X_n(t) - X_n(t_1)| \geqslant \lambda, |X_n(t_2) - X_n(t)| \geqslant \lambda\} \leqslant$$
$$\leqslant \frac{1}{\lambda^{2\gamma}} [F(t_2) - F(t_1)]^{2\alpha}.$$

Рассмотрим снова величины (15.23). Если точки $\tau + j\delta/m$, $j = 0, 1, \ldots, m$, все являются целыми кратными 2^{-n}, то неравенство (15.24) получается так же, как и раньше. Если τ и δ являются целыми кратными 2^{-n}, тогда то же самое справедливо для точек $\tau + j\delta/m$ при условии, что $m = \delta 2^n$. Однако поскольку функция X_n постоянна на каждом интервале $[(i-1)2^{-n}, i2^{-n})$, то величина M''_m при $m = \delta 2^n$ равна в точности величине, определенной формулой (15.25). Таким образом, неравенство (15.26) выполняется, если и τ, и δ являются целыми кратными 2^{-n}.

Предположим теперь, что $\delta = 1/2^v$ для некоторого целого $v > 1$. Если $n \geqslant v$, то δ является целым кратным 2^{-n} и таковыми же являются концевые точки всех интервалов, входящих в формулы (15.27) и (15.28). Отсюда, как и раньше, следует, что (15.30) выполняется при $n \geqslant v$. Таким образом, существует такое δ, что (15.33) выполняется при всех больших n.

Для доказательства плотности последовательности элементов X_n достаточно показать, что соответствующие распределения P_n удовлетворяют предположениям теоремы 15.3. Условие (15.7) выполнено вследствие того, что справедливо (15.33). Если $2^{-k} < \delta$, то

$$\sup_t |X_n(t)| \leqslant \max_{i \leqslant 2^k} \left| X_n\left(\frac{i}{2^k}\right)\right| + w''(X_n, \delta).$$

Поскольку распределения первого члена в правой части совпадают при всех $n \geqslant k$, то из (15.33) следует, что условие (i) теоремы 15.3 выполняется.

Для доказательства (15.8) и (15.9) мы временно предположим, что если для некоторого положительного числа δ_0 имеет место неравенство $h \leqslant \delta_0$, то

$$\mu_{0,h}\{(\beta_1, \beta_2) : \beta_1 = \beta_2\} = 1, \quad \mu_{1,1-h}\{(\beta_1, \beta_2) : \beta_1 = \beta_2\} = 1. \tag{15.34}$$

При этом условия (15.8) и (15.9), конечно, выполняются, так что последовательность $\{X_n\}$ плотна. Если X — предел сходящейся по распределению некоторой подпоследовательности, то $(X(t_1), \ldots, X(t_k))$ имеет распределение $\mu_{t_1 \ldots t_k}$ для двоично рациональных чисел $t_1, \ldots, t_k$, а общий случай получается с помощью (15.32) посредством аппроксимации сверху.

Остается снять ограничение (15.34). Возьмем $\delta_0 < 1/2$ и положим

$$f(t) = \begin{cases} 0 & \text{при} \quad 0 \leqslant t \leqslant \delta_0, \\ \dfrac{t-\delta_0}{1-2\delta_0} & \text{при} \quad \delta_0 \leqslant t \leqslant 1-\delta_0, \\ 1 & \text{при} \quad 1-\delta_0 \leqslant t \leqslant 1. \end{cases}$$

Пусть $\nu_{t_1 \ldots t_k} = \mu_{s_1 \ldots s_k}$, где $s_i = f(t_i)$. Тогда $\nu_{t_1 \ldots t_k}$ удовлетворяет условиям теоремы (с новой функцией F), а также специальному условию (15.34), так что в пространстве D существует случайный элемент Y с этими конечномерными распределениями. Требуемый случайный элемент X теперь получим, положив $X(t) = Y(\delta_0 + t(1 - 2\delta_0))$ при $0 \leqslant t \leqslant 1$.

Для иллюстрации применений теоремы 15.7 докажем с ее помощью существование в D случайного элемента, описывающего цепь Маркова с непрерывным временем. Пусть $p_{ij}(t)$, i, $j=1, 2, \ldots$, — неотрицательные числа, определенные при $t>0$ и удовлетворяющие соотношениям $\sum\limits_j p_{ij}(t)=1$ и $\sum\limits_j p_{ij}(s)\,p_{jk}(t)=p_{ik}(s+t)$. Пусть p_i, $i=1, 2, \ldots$, — неотрицательные числа, удовлетворяющие условию $\sum\limits_i p_i=1$. Положим $p_j(0)=p_j$ и $p_j(t)=\sum\limits_i p_i p_{ij}(t)$ при $t>0$.

В предположении, что соотношение

$$\lim_{t\to 0} p_{ii}(t)=1 \tag{15.35}$$

выполняется равномерно по i, мы докажем существование в D случайного элемента X с конечномерными распределениями, заданными равенствами

$$\mathsf{P}\{X(t_u)=i_u,\ u=0,\ 1,\ \ldots,\ m\}= \\ =p_{i_0}(t_0)\prod_{u=1}^{m} p_{i_{u-1}i_u}(t_u-t_{u-1}) \tag{15.36}$$

при $0\leqslant t_0<t_1<\ldots<t_m\leqslant 1$. Согласованность этих конечномерных распределений устанавливается без труда.

Докажем сначала существование положительного K такого, что для всех i и всех положительных t

$$0\leqslant 1-p_{ii}(t)\leqslant Kt. \tag{15.37}$$

Вследствие равномерности в предельном соотношении (15.35), каково бы ни было ε, найдется такое δ, что $p_{ii}(s)>1-\varepsilon$ при всех i, если $s\leqslant 2\delta$. Если $t\leqslant\delta$, то $\delta\leqslant mt\leqslant 2\delta$ при некотором положительном целом m, и мы имеем

$$1-\varepsilon\leqslant p_{ii}(mt)\leqslant \\ \leqslant[p_{ii}(t)]^m+\sum_{l=0}^{m-2}[p_{ii}(t)]^l\sum_{j\neq i}p_{ij}(t)\,p_{ji}((m-l-1)\,t)\leqslant \\ \leqslant[p_{ii}(t)]^m+\sum_{l=0}^{m-2}[p_{ii}(t)]^l\sum_{j\neq i}p_{ij}(t)\,\varepsilon\leqslant \\ \leqslant[p_{ii}(t)]^m+\varepsilon\{1-[p_{ii}(t)]^m\}.$$

При $\varepsilon = 1/3$ из последнего неравенства следует (так как $\log u \leqslant u - 1$), что $m(1 - p_{ii}(t)) \leqslant \log 2 \leqslant 1$. Возьмем $K = \delta^{-1}$. Поскольку $mt \geqslant \delta$, то мы получаем, что (15.37) выполняется при $t \leqslant \delta$. Но при $t \geqslant \delta$ неравенство (15.37) тривиально.

Если соотношение (15.36) справедливо, то

$$P\{X(t) \neq X(t_1),\ X(t_2) \neq X(t)\} = \sum p_i(t_1)\, p_{ij}(\delta_1)\, p_{jk}(\delta_2), \tag{15.38}$$

где $\delta_1 = t - t_1$, $\delta_2 = t_2 - t$ и где суммирование производится по всем i и по всем таким j и k, для которых $i \neq j$ и $j \neq k$. В силу (15.37) сумма в (15.38) не превосходит $K^2\delta_1\delta_2 \geqslant K^2(\delta_1 + \delta_2)^2$. Таким образом, конечномерные распределения, заданные в (15.36), удовлетворяют условию (15.31) с $\gamma = 0$, $\alpha = 1$ и $F(t) = Kt$. Так как из (15.36) и (15.37) следует также неравенство $P\{X(t+h) \neq X(t)\} \leqslant Kh$, то условие (15.32) также выполняется. Таким образом, некоторый случайный элемент X пространства D удовлетворяет условию (15.36). (Так как различные значения $X(t)$ отличаются друг от друга не меньше чем на 1, то с вероятностью 1 X является ступенчатой функцией, постоянной на интервалах.)

Другие критерии*). В теореме 15.6 предполагалось, что функция F неубывающая и непрерывная. Можно заменить предположение о непрерывности предположением о непрерывности справа, если при этом усилить условия (15.21) следующим образом:

$$P\{|X_n(t) - X_n(t_1)| \geqslant \lambda,\ |X_n(t_2) - X_n(t)| \geqslant \lambda\} \leqslant \\ \leqslant \frac{1}{\lambda^{2\gamma}} [F(t) - F(t_1)]^{\alpha} [F(t_2) - F(t)]^{\alpha}. \tag{15.39}$$

В самом деле, неравенство (15.39) и теорема 12.6 позволяют заменить (15.24) более сильным неравенством

$$P\{M''_m \geqslant \varepsilon\} \leqslant \frac{K}{\varepsilon^{2\gamma}} [F(\tau + \delta) - F(\tau)]^{2\alpha} \times \\ \times \min_{1 \leqslant i \leqslant m} \left[1 - \frac{F\left(\tau + \frac{i}{m}\delta\right) - F\left(\tau - \frac{i-1}{m}\delta\right)}{F(\tau + \delta) - F(\tau)}\right]^{\alpha}. \tag{15.40}$$

*) Эта тема может быть пропущена.

Пусть $J(\tau, \tau+\delta)$ обозначает максимальный скачок F в интервале $(\tau, \tau+\delta]$. В силу (15.40) имеем

$$\mathsf{P}\{M''_m \geqslant \varepsilon\} \leqslant$$
$$\leqslant \frac{K}{\varepsilon^{2\gamma}}[F(\tau+\delta)-F(\tau)]^{2\alpha}\left[1-\frac{J(\tau, \tau+\delta]}{F(\tau+\delta)-F(\tau)}\right]^{\alpha}. \quad (15.41)$$

(Если $F(\tau)=F(\tau+\delta)$, то правая часть этого неравенства полагается равной нулю.)

Точно так же, как и раньше, мы получаем неравенство (15.29), в котором каждая из сумм $\sum'$ и $\sum''$ теперь имеет вид

$$\sum_{k=1}^{r}[F(t_k)-F(t_{k-1})]^{2\alpha}\left[1-\frac{J(t_{k-1}, t_k]}{F(t_k)-F(t_{k-1})}\right]^{\alpha}, \quad (15.42)$$

где $0 \leqslant t_1 \leqslant \ldots \leqslant t_r \leqslant 1$ и $t_k - t_{k-1} \leqslant 2\delta$. Для доказательства (15.22) достаточно поэтому показать, что при достаточно малом δ все суммы (15.42) становятся меньше, чем $\eta_0 = \eta\varepsilon^{2\gamma}/(2K)$. Обозначим разность $F(t_k)-F(t_{k-1})$ через Δ_k и $J(t_{k-1}, t_k]$ через J_k. Поскольку $\alpha > 1/2$, а величины Δ_k в сумме составляют самое большее $F(1)-F(0)$, то

$$\sum_k \Delta_k^{\alpha}(\Delta_k - J_k)^{\alpha} \leqslant [F(1)-F(0)]^{\alpha+1/2}\max_k(\Delta_k-J_k)^{\alpha-1/2}.$$

Левая часть этого неравенства есть не что иное, как сумма (15.42), и следовательно, достаточно лишь показать, что если $0 \leqslant t-s \leqslant 2\delta$ для достаточно малого δ, то имеет место оценка

$$F(t)-F(s)-J(s, t] < \eta_1 =$$
$$= \frac{\eta_0^{1/(\alpha-1/2)}}{[F(1)-F(0)]^{(\alpha+1/2)/(\alpha-1/2)}}. \quad (15.43)$$

Поскольку F является элементом пространства D, то можно указать (см. (14.8)) точки s_i, $0=s_0<s_1<\ldots$ $\ldots<s_l=1$, такие, что $w_F[s_{i-1}, s_i) < \frac{1}{2}\eta_1$. Пусть 2δ меньше, чем наименьшая из величин $s_i - s_{i-1}$. Если $t-s \leqslant 2\delta$, то либо s и t лежат в одном и том же интервале $[s_{i-1}, s_i)$, и в этом случае левая часть (15.43) меньше, чем $\frac{1}{2}\eta_1$, либо s и t лежат в смежных интервалах $[s_{i-1}, s_i)$ и $[s_i, s_{i+1})$, и в этом случае левая часть (15.43) не превосходит $F(t)-F(s_i)+F(s_i-)-F(s)\leqslant\eta_1$.

Таким образом, из неравенства (15.39) для непрерывной справа функции F следует (15.22). Доказательство теоремы завершается теперь, как и прежде, и мы заключаем, что $X_n \xrightarrow{\mathscr{D}} X$. Аналогично, в теореме 15.7 можно считать функцию F непрерывной справа, если заменить (15.31) на

$$\mu_{t_1tt_2}\{(\beta_1, \beta, \beta_2): |\beta-\beta_1|\geqslant\lambda,\ |\beta_2-\beta|\geqslant\lambda\}\leqslant$$
$$\leqslant\frac{1}{\lambda^{2\gamma}}[F(t)-F(t_1)]^{\alpha}[F(t_2)-F(t)]^{\alpha}. \qquad (15.44)$$

Сепарабельные стохастические процессы*). Согласно теореме 9.2, если конечномерные распределения сепарабельного стохастического процесса могут быть реализованы как конечномерные распределения вероятностной меры, заданной на $(C, \mathscr{C})$, то выборочные траектории процесса непрерывны с вероятностью 1. Мы докажем подобный результат для пространства D.

Пусть $\{\xi_t\colon 0\leqslant t\leqslant 1\}$ — стохастический процесс, и пусть P — вероятностная мера на $(D, \mathscr{D})$. Предположим, что процесс и мера имеют одни и те же конечномерные распределения. Предположим, кроме того, что процесс сепарабелен, так что условие (9.16) выполняется для любых α и β и любого открытого интервала I. Не ограничивая общности, можно считать, что $0\in T_0$; в этом случае (9.16) выполняется также и тогда, когда I имеет вид $[0, s)$.

Теперь мы построим множества, служащие тем же целям, что и множества (9.20) в рассуждениях § 9. Пусть V обозначает теперь систему чисел:

$$V\colon k;\ r_1, \ldots, r_k;\ s_1, \ldots, s_k;\ \alpha_1, \ldots, \alpha_k, \qquad (15.45)$$

где k — произвольное целое число, r_i, s_i и α_i — рациональные числа и

$$0=r_1<s_1<r_2<s_2<\ \ldots<r_k<s_k=1.$$

*) Эта тема может быть пропущена.

Положим

$$\Omega_{T_0}(V, \varepsilon) = \bigcap_{i=1}^{k} \{\omega\colon a_i \leqslant \xi_t(\omega) \leqslant a_i + \varepsilon,\ t \in (r_i, s_i) \cap T_0\} \cap$$
$$\cap \bigcap_{i=2}^{k} \{\omega\colon \min\{a_{i-1}, a_i\} \leqslant \xi_t(\omega) \leqslant$$
$$\leqslant \max\{a_{i-1}, a_i\} + \varepsilon,\ t \in (s_{i-1}, r_i) \cap T_0\}; \qquad (15.46)$$

в первом пересечении мы заменяем (r_1, s_1) на $[r_1, s_1)$ (для $i = 1$). Пусть $\mathscr{V}_{k\delta}$ обозначает класс систем (15.45), которые имеют фиксированное значение k и удовлетворяют условиям $r_i - s_{i-1} < \delta$, $i = 2, \ldots, k$. Определим теперь множество

$$\Omega_{T_0} = \bigcap_{\varepsilon} \bigcup_{k} \bigcap_{\delta} \bigcup_{V \in \mathscr{V}_{k\delta}} \Omega_{T_0}(V, \varepsilon), \qquad (15.47)$$

где ε и δ пробегают положительные рациональные значения. Наконец, определим $\Omega_T(V, \varepsilon)$ и Ω_T, заменяя T_0 на $T = [0, 1]$ в соотношениях (15.46) и (15.47).

Нетрудно показать, что если $\omega \in \Omega_T$, то выборочная траектория процесса, отвечающая ω, непрерывна справа в точке $t = 0$, имеет предел слева в точке $t = 1$ и на $(0, 1)$ может иметь разрывы только первого рода. С другой стороны, если выборочная траектория обладает этим свойством, то $\omega \in \Omega_T$, что следует из леммы 1 § 14 (примененной к выборочным траекториям, нормированным так, что они непрерывны справа).

Докажем, что $\Omega_T \in \mathscr{B}$ и $\mathsf{P}(\Omega_T) = 1$. Справедливость соотношений (9.22) и $\Omega_{T_0} \in \mathscr{B}$ получается с помощью тех же самых рассуждений, что и в § 9; следовательно, $\Omega_T \in \mathscr{B}$ и $\mathsf{P}(\Omega_T) = \mathsf{P}(\Omega_{T_0})$, и остается лишь показать, что

$$\mathsf{P}(\Omega_{T_0}) = 1. \qquad (15.48)$$

Для занумерованных в каком-либо порядке точек $(t_1, t_2, \ldots)$, образующих множество T_0, определим отображения $\varphi\colon \Omega \to R^\infty$ и $\psi\colon D \to R^\infty$, положив $\varphi(\omega) = (\xi_{t_1}(\omega), \xi_{t_2}(\omega), \ldots)$ и $\psi(x) = (x(t_1), x(t_2), \ldots)$. Точно также, как и раньше, равенство (9.25) справедливо для любого множества H из $\mathscr{R}^\infty$, где теперь P обозначает вероятностную меру на $(D, \mathscr{D})$, а не на $(C, \mathscr{C})$, обладающую конечномерными распределениями процесса $\{\xi_t\}$.

Для системы (15.45) и произвольного $\varepsilon > 0$ пусть множество $H_{T_0}(V, \varepsilon)$ состоит из таких точек $z = (z_1, z_2, \ldots)$ из R^∞, которые обладают следующими свойствами: во-первых, для любого $i = 1, 2, \ldots, k$ неравенство

$$\alpha_i \leqslant z_u \leqslant \alpha_i + \varepsilon$$

выполняется для всякого u, для которого $t_u \in (r_i, s_i)$ (или $t_u \in [r_1, s_1)$ в случае $i = 1$), и, во-вторых, для каждого $i = 2, \ldots, k$ неравенство

$$\min\{\alpha_{i-1}, \alpha_i\} \leqslant z_u \leqslant \max\{\alpha_{i-1}, \alpha_i\} + \varepsilon$$

выполняется для любого u, для которого $t_u \in (s_{i-1}, r_i)$. Определим теперь

$$H_{T_0} = \bigcap_\varepsilon \bigcup_k \bigcap_\delta \bigcup_{V \in \mathcal{V}_{k\delta}} H_{T_0}(V, \varepsilon).$$

Множества $H_{T_0}(V, \varepsilon)$ и H_{T_0} принадлежат $\mathcal{R}^\infty$, $\varphi^{-1}H_{T_0}(V, \varepsilon) = \Omega_{T_0}(V, \varepsilon)$ и $\varphi^{-1}H_{T_0} = \Omega_{T_0}$. В силу (9.25) имеем $\mathsf{P}(\Omega_{T_0}) = P(\psi^{-1}H_{T_0})$, откуда, поскольку $\psi^{-1}H_{T_0} = D$, вытекает (15.48).

Мы доказали следующий результат:

Теорема 15.8. *Если $\{\xi_t\colon 0 \leqslant t \leqslant 1\}$ — сепарабельный стохастический процесс и если на $(D, \mathscr{D})$ существует вероятностная мера, обладающая теми же конечномерными распределениями, что и $\{\xi_t\}$, то выборочные траектории процесса с вероятностью* 1 *непрерывны справа при $t = 0$, имеют пределы слева при $t = 1$ и на* $(0, 1)$ *могут иметь разрывы лишь первого рода.*

Следующий пример показывает, что невозможно продвинуться дальше и доказать, что траектории непрерывны справа и, следовательно, лежат в D (с вероятностью 1). Положим

$$\xi_t(\omega) = \begin{cases} 0, & \text{если } 0 \leqslant t \leqslant \omega, \\ 1, & \text{если } \omega < t \leqslant 1, \end{cases}$$

где точка ω равномерно распределена на $\Omega = (0, 1)$. Этот пример лишний раз подчеркивает тот факт, что, полагая элементы D непрерывными справа, мы просто приняли удобное соглашение.

Задачи

1. Покажите, что теорема 15.4 остается верной, если потребовать лишь, чтобы (15.10) выполнялось для всех $t_1, \ldots, t_k$ из T_0, где множество T_0 всюду плотно в $[0, 1]$ и содержит 0 и 1. Покажите, что теорема неверна, если $0 \notin T_0$ или если $1 \notin T_0$.

2. Условие (15.32) в теореме 15.7 является необходимым.

3. Если случайный элемент X пространства D обладает тем свойством, что

$$\lim_{\delta \to 0} \sup_{0 \leqslant t \leqslant 1-\delta} \frac{1}{\delta} \mathsf{P}\{|X(t+\delta) - X(t)| \geqslant \varepsilon\} = 0$$

для любого положительного ε, то $\mathsf{P}\{X \in C\} = 1$.

4. Используя теорему 15.7 и задачу 3, покажите, что распределения $\mu_{t_1 \ldots t_k}$ могут быть реализованы как конечномерные распределения случайного элемента пространства C, если выполняется (15.31) (где $\gamma \geqslant 0$, $\alpha > 1/2$ и F не убывает и непрерывна) и если для любого положительного ε

$$\lim_{\delta \to 0} \sup_{0 \leqslant t \leqslant 1-\delta} \frac{1}{\delta} \mu_{t,\, t+\delta} \{(\beta_1, \beta_2): |\beta_2 - \beta_1| \geqslant \varepsilon\} = 0.$$

Имеет место аналогичное утверждение с (15.44) вместо (15.31). Используйте теорему 9.2 для вывода условий непрерывности выборочных траекторий сепарабельного процеса. (Сравните с задачей 8 из § 12.)

5. Используйте теорему 15.7 для построения пуассоновского процесса в D. Сделайте также прямое построение, отправляясь от независимых показательно распределенных случайных величин.

§ 16. Приложения

Теорема Донскера. В некоторых отношениях пространство D более удобно для формулировки теоремы Донскера, чем пространство C. Пусть случайные величины $\xi_1, \xi_2, \ldots$ определены на $(\Omega, \mathscr{B}, \mathsf{P})$, и пусть $S_n = \xi_1 + \ldots + \xi_n$ — их частичные суммы. Определим функцию $X_n(\omega)$, принадлежащую пространству D, положив

$$X_n(t, \omega) = \frac{1}{\sigma \sqrt{n}} S_{[nt]}(\omega). \tag{16.1}$$

Поскольку $X_n(t)$ при каждом t является случайной величиной, то X_n является случайной функцией — случайным элементом пространства D.

Мы хотим доказать, что при подходящих условиях распределение функции X_n сходится слабо к винеровской мере W. Мера W определена на $(C, \mathscr{C})$, но ее легко продолжить на $(D, \mathscr{D})$: поскольку $C \in \mathscr{D}$, а относительная топология Скорохода в пространстве C совпадает с равномерной топологией в этом пространстве, то для любого A из $\mathscr{D}$ $A \cap C \in \mathscr{C}$. Мы можем поэтому продолжить меру W на $(D, \mathscr{D})$, полагая ее значение на $A \in \mathscr{D}$ равным $W(A \cap C)$. Разумеется, C служит носителем меры W. В дальнейшем мы будем интерпретировать W или как вероятностную меру на $(D, \mathscr{D})$, или как случайный элемент D с этой вероятностной мерой в качестве его распределения.

Теорема 16.1. *Предположим, что случайные величины ξ_n независимы и одинаково распределены со средним* 0 *и конечной положительной дисперсией σ^2:*

$$\mathsf{E}\{\xi_n\} = 0, \qquad \mathsf{E}\{\xi_n^2\} = \sigma^2. \tag{16.2}$$

Тогда случайные функции X_n, определенные формулой (16.1), *удовлетворяют предельному соотношению*

$$X_n \xrightarrow{\mathscr{D}} W. \tag{16.3}$$

Доказательство. Доказательство сходимости конечномерных распределений, данное в § 10, без трудностей переносится на рассматриваемый случай. Таким образом, нужно доказать лишь плотность.

Имеется несколько способов доказательства плотности. Один из них заключается в проверке выполнения предположений теоремы 15.5. С другой стороны, условие (ii) теоремы 8.3, сформулированной для случайных элементов пространства D, требует, чтобы для положительных ε и η существовали бы δ, $0 < \delta < 1$, и целое число n_0 такие, что при $0 \leqslant t \leqslant 1$

$$\frac{1}{\delta} \mathsf{P}\left\{\sup_{t \leqslant s \leqslant t+\delta} |X_n(s) - X_n(t)| \geqslant \varepsilon\right\} \leqslant \eta, \quad n \geqslant n_0$$

(если $t + \delta$ превосходит 1, то мы заменяем $t + \delta$ на 1). Далее это условие сводится с помощью рассуждений, использованных при доказательстве теоремы 8.4, к следующему условию: для любого положительного ε существуют $\lambda > 1$ и целое число n_0 такие, что

$$\mathsf{P}\left\{\max_{i \leqslant n} |S_i| \geqslant \lambda\sigma\sqrt{n}\right\} \leqslant \frac{\varepsilon}{\lambda^2}, \quad n \geqslant n_0.$$

Следовательно, достаточно проверить лишь это последнее условие. Это можно осуществить либо с помощью центральной предельной теоремы, как в § 10 (см. (10.7)), либо с помощью оценок из § 12 (см. (12.23)). Таким образом, методы предыдущей главы переносятся сюда без существенных изменений.

Используя результаты § 15, можно построить очень простое доказательство. В силу теоремы 15.6 достаточно установить неравенство

$$\mathsf{E}\{|X_n(t)-X_n(t_1)|^2\cdot|X_n(t_2)-X_n(t)|^2\}\leqslant \leqslant 4(t_2-t_1)^2,\quad t_1\leqslant t\leqslant t_2. \tag{16.4}$$

Ввиду (16.1), (16.2) и независимости величин ξ_n, для левой части неравенства (16.4) имеет место оценка

$$\frac{1}{\sigma^4 n^2}\mathsf{E}\{|S_{[nt]}-S_{[nt_1]}|^2\}\mathsf{E}\{|S_{[nt_2]}-S_{[nt]}|^2\}= =\frac{1}{n^2}([nt]-[nt_1])([nt_2]-[nt])\leqslant\left\{\frac{[nt_2]-[nt_1]}{n}\right\}^2. \tag{16.5}$$

Если $t_2-t_1\geqslant 1/n$, то (16.4) следует из (16.5). Если $t_2-t_1<1/n$, то либо t_1 и t, либо t и t_2 лежат в одном и том же интервале $[(i-1)/n,\ i/n)$; в любом из этих случаев левая часть в (16.4) обращается в 0. Этим устанавливается справедливость неравенства (16.4) в общем случае, и теорема доказана.

Так как значение $\sup_t X_n(t)$ одинаково для случайного элемента D, определенного формулой (16.1), и для случайного элемента C, определенного формулой (10.1), и равно $\max\limits_{i\leqslant n} S_i/(\sigma\sqrt{n})$, то мы можем, очевидно, использовать теорему 16.1 для еще одного вывода (10.18). (Нужно доказать, что отображение h: $D\to R^1$, определенное равенством $h(x)=\sup_t x(t)$, непрерывно, но это не составляет труда.) Подобным же образом мы можем заново получить результаты § 11.

Некоторые функции от частичных сумм S_i имеют простое выражение в терминах случайных элементов X_n из теоремы 16.1, но не имеют простого выражения в терминах случайных элементов X_n из теоремы 10.1. Такие функции легче анализировать в D, чем в C. К примеру, пусть $h(x)$ при $x\in D$ является лебеговой мерой множества таких t, для которых $x(t) > 0$. Функция h измерима ($h^{-1}\mathscr{R}^1 \subset \mathscr{D}$) и непрерывна всюду, за исключением множества винеровской меры 0 (стр. 318). Если случайная функция X_n определена формулой (16.1), то $h(X_n)$ в точности равно умноженному на n^{-1} числу положительных сумм среди $S_1, \ldots, S_{n-1}$, в то время как это, вообще говоря, неверно, если функция X_n определена формулой (10.1). Объединяя теорему 5.1, теорему 16.1 и (11.26), мы приходим к закону арксинуса, справедливому в предположениях теоремы Линдеберга — Леви *).

Доминированные меры. Случайные величины в теореме 16.1 определены на вероятностном пространстве $(\Omega, \mathscr{B}, \mathsf{P})$. Мы покажем, что теорема остается верной, если меру P заменить произвольной вероятностной мерой P_0 на $(\Omega, \mathscr{B})$, доминируемой мерой P (иначе говоря, абсолютно непрерывной относительно P). Например, предположим, что $\Omega = [0, 1]$, $\mathscr{B}$ состоит из линейных борелевских подмножеств отрезка $[0, 1]$ и $\xi_n(\omega) = 2\omega_n - 1$, где ω_n является n-м знаком в двоичном разложении ω. Тогда теорема 16.1 применима к последовательности $\{\xi_n\}$, если ω имеет равномерное распределение на $[0, 1]$. Согласно доказываемой ниже теореме утверждение теоремы 16.1 по-прежнему верно, если ω имеет произвольное распределение на $[0, 1]$, обладающее плотностью относительно лебеговой меры.

Нам понадобится следующий предварительный результат (ниже $\sigma(\mathscr{B}_0)$ обозначает σ-поле, порожденное $\mathscr{B}_0$):

Теорема 16.2. *Пусть $E_1, E_2, \ldots$ — измеримые множества в вероятностном пространстве $(\Omega, \mathscr{B}, \mathsf{P})$. Допустим, что существуют постоянная α и подполе $\mathscr{B}_0$ σ-поля $\mathscr{B}$ такие, что для любого множества E из $\mathscr{B}_0$*

$$\mathsf{P}(E_n \cap E) \to \alpha \mathsf{P}(E). \tag{16.6}$$

Допустим, кроме того, что все множества E_n принадлежат $\sigma(\mathscr{B}_0)$. Если P мажорирует другую вероятностную меру P_0 на $\mathscr{B}_0$, то

$$\mathsf{P}_0(E_n) \to \alpha. \tag{16.7}$$

*) Некоторые дальнейшие приложения см. в задачах.

Доказательство. Точно так же, как в доказательстве теоремы 4.5, из предположений следует, что

$$\int_{E_n} g\,d\mathsf{P} \to \alpha \int g\,d\mathsf{P} \tag{16.8}$$

для любой интегрируемой функции g. Остается заметить, что соотношения (16.8) и (16.7) означают одно и то же, если g является производной Радона — Никодима меры P_0 относительно меры P.

Теорема 16.3. *Теорема* 16.1 *остается справедливой, если меру* P *заменить произвольной доминируемой ею вероятностной мерой* P_0.

Доказательство. Определим случайную функцию X'_n:

$$X'_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} \sum_{p_n \leqslant i \leqslant nt} \xi_i(\omega), \tag{16.9}$$

где $\{p_n\}$ — последовательность целых чисел, сходящаяся к бесконечности столь медленно, что $p_n/\sqrt{n} \to 0$ ($X'_n(t, \omega) = 0$ при $t < p_n/n$). Если $\delta_n = \sup_t |X_n(t) - X'_n(t)|$, то $\delta_n \leqslant \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{p_n} |\xi_i|$. В силу неравенства Минковского и того факта, что $p_n/\sqrt{n} \to 0$, имеем

$$\mathsf{E}^{1/2}\{\delta_n^2\} \leqslant \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{p_n} \mathsf{E}^{1/2}\{\xi_i^2\} \to 0,$$

откуда по неравенству Чебышева

$$\delta_n \xrightarrow{P} 0, \tag{16.10}$$

где сходимость понимается в смысле меры P (т. е. предполагается, что P определяет распределение ξ_n). По теореме 16.1

$$X_n \xrightarrow{\mathscr{D}} W, \tag{16.11}$$

где сходимость рассматривается по отношению к мере P. Поскольку $d(X_n, X'_n) \leqslant \delta_n$, где d представляет собой метрику в D (любую из рассмотренных выше), то из теоремы 4.1 следует, что

$$X'_n \xrightarrow{\mathscr{D}} W \tag{16.12}$$

относительно P.

Возьмем и временно зафиксируем W-непрерывное множество A (в D); согласно (16.12)

$$\mathsf{P}\{X'_n \in A\} \to W(A). \tag{16.13}$$

Пусть $\mathscr{B}_0$ обозначает поле цилиндров: $\mathscr{B}_0$ состоит из множеств вида

$$\{\omega\colon (\xi_1(\omega), \ldots, \xi_k(\omega)) \in H\},$$

где $k \geqslant 1$ и $H \in \mathscr{R}^k$. Если $E \in \mathscr{B}_0$, то, так как $p_n \to \infty$, $\mathsf{P}(\{X'_n \in A\} \cap E) = \mathsf{P}\{X'_n \in A\}\,\mathsf{P}(E)$ для больших n, и из (16.13) следует, что

$$\mathsf{P}(\{X'_n \in A\} \cap E\} \to W(A)\,\mathsf{P}(E).$$

Поскольку множества $\{X'_n \in A\}$ все принадлежат σ-полю, порожденному $\mathscr{B}_0$, то из теоремы 16.2 (при $\alpha = W(A)$) вытекает, что если мера P_0 абсолютно непрерывна относительно P, то

$$\mathsf{P}_0\{X'_n \in A\} \to W(A). \tag{16.14}$$

Так как соотношение (16.14) выполняется для любого множества A непрерывности меры W, то соотношение (16.12) выполняется и в том случае, когда оно понимается в смысле меры P_0. Далее, соотношение (16.10) означает, что $\mathsf{P}\{\delta_n \geqslant \varepsilon\} \to 0$ при каждом положительном ε; если мера P_0 абсолютно непрерывна относительно P, то также $\mathsf{P}_0\{\delta_n \geqslant \varepsilon\} \to 0$, и соотношение (16.10) остается справедливым и тогда, когда оно понимается в смысле меры P_0. Применяя еще раз теорему 4.1, мы видим, что (16.11) выполняется в смысле P_0, что и завершает доказательство.

Теорема 16.3 касается сходимости по распределению. Заметим, что при переходе от P к доминируемой мере P_0 тривиальным образом сохраняются такие свойства последовательности $\{S_n\}$, как закон повторного логарифма, который выполняется с вероятностью 1.

Эмпирические функции распределения. Пусть $\xi_1, \xi_2, \ldots$ — случайные величины такие, что

$$0 \leqslant \xi_n(\omega) \leqslant 1,$$

и пусть $F_n(t, \omega)$ — доля точек из множества $\xi_1(\omega), \ldots$ $\ldots, \xi_n(\omega)$, не превосходящих t, так что $F_n(\cdot, \omega)$ является эмпирической функцией распределения. Пусть $Y_n(\omega)$ — элемент пространства D, принимающий в точке $t \in [0, 1]$ значение

$$Y_n(t, \omega) = \sqrt{n}\,(F_n(t, \omega) - F(t)), \tag{16.15}$$

где F — общая функция распределения величин ξ_n. Так как $Y_n(t)$ при каждом t есть случайная величина, то Y_n является случайным элементом пространства D. В § 13 мы изучали сходный случайный элемент (13.7) пространства C; здесь мы будем исследовать Y_n непосредственно и откажемся от принимавшегося ранее ограничения, что ξ_n распределены равномерно. Стохастический процесс $\{Y_n(t)\colon 0 \leqslant t \leqslant 1\}$ называется иногда эмпирическим процессом.

Теорема 16.4. *Предположим, что случайные величины ξ_n независимы и имеют общую функцию распределения $F(t)$. Если Y_n определяется формулой* (16.15), *то*

$$Y_n \xrightarrow{\mathscr{D}} Y, \tag{16.16}$$

где Y — гауссовский случайный элемент пространства D такой, что

$$\begin{cases} \mathsf{E}\{Y(t)\} = 0, \\ \mathsf{E}\{Y(s)Y(t)\} = F(s)(1 - F(t)), \quad s \leqslant t. \end{cases} \tag{16.17}$$

Доказательство. Подобно винеровской мере, мера W° продолжима с $(C, \mathscr{C})$ на $(D, \mathscr{D})$. Обозначим через W° случайный элемент пространства D, имеющий в качестве своего распределения эту продолженную меру. Предполагая сначала, что величины ξ_n распределены равномерно на $[0, 1]$, мы покажем, что $Y_n \xrightarrow{\mathscr{D}} W^\circ$.

Пусть $U_n(t, \omega)$ — число точек множества $\xi_1(\omega), \ldots$ $\ldots, \xi_n(\omega)$, которые не превосходят t. При $0 = t_0 <$ $< t_1 < \ldots < t_k = 1$ случайные величины $U_n(t_i) -$ $- U_n(t_{i-1})$, $i = 1, \ldots, k$, имеют полиномиальные распределения с параметрами n и $p_i = t_i - t_{i-1}$, и так же, как в § 13, из центральной предельной теоремы для полиномиальных испытаний следует, что конечномерные распределения элемента Y_n сходятся слабо к конечномерным распределениям элемента W°. В силу теоремы 15.6 достаточно доказать, что при $t_1 \leqslant t \leqslant t_2$

$$\mathsf{E}\{|Y_n(t) - Y_n(t_1)|^2 \cdot |Y_n(t_2) - Y_n(t)|^2\} \leqslant 6(t - t_1)(t_2 - t).$$

Но последнее неравенство в точности совпадает с неравенством (13.17), установленным в § 13. Следовательно, $Y_n \xrightarrow{\mathscr{D}} W^\circ$.

Допустим теперь, что величины ξ_n имеют произвольную функцию распределения F на $[0, 1]$. Определим функцию, «обратную» к F, равенством

$$\varphi(s) = \inf\{t\colon\ s \leqslant F(t)\}.$$

Неравенство $s \leqslant F(t)$ справедливо тогда и только тогда, когда $\varphi(s) \leqslant t$, так что если величина η_n равномерно распределена на $[0, 1]$, то $\mathsf{P}\{\varphi(\eta_n) \leqslant t\} = F(t)$. Так как в теореме используется лишь совместное распределение величин ξ_n, то мы можем представить эти величины в виде $\xi_n = \varphi(\eta_n)$, где величины η_n независимы и равномерно распределены.

Если $G_n(\cdot, \omega)$ — эмпирическая функция распределения, соответствующая величинам $\eta_1(\omega), \ldots, \eta_n(\omega)$ и $Z_n(t, \omega) = \sqrt{n}\ (G_n(t, \omega) - t)$, то, как установлено выше, $Z_n \xrightarrow{\mathscr{D}} W^\circ$. Но эмпирическое распределение, отвечающее величинам $\xi_1(\omega), \ldots, \xi_n(\omega)$, равно $F_n(t, \omega) = G_n(F(t), \omega)$, так что если Y_n определяется формулой (16.15), то $Y_n(t, \omega) = Z_n(F(t), \omega)$. Определим отображение $\psi\colon D \to D$ равенством $(\psi x)(t) = x(F(t))$. Если x_n сходится к x в топологии Скорохода и $x \in C$, то эта сходимость равномерная, так что ψx_n сходится к ψx равномерно и, следовательно, сходится в топологии Скорохода. Поэтому из сходимости $Z_n \xrightarrow{\mathscr{D}} W^\circ$ и теоремы 5.1 ледует, что $Y_n = \psi(Z_n) \xrightarrow{\mathscr{D}} \psi(W^\circ)$. Поскольку $Y = \psi(W^\circ)$ является гауссовским случайным элементом, удовлетворяющим условию (16.17), то теорема доказана.

Если бы мы знали распределение величины

$$\sup_t Y(t) = \sup_t W^\circ(F(t)), \tag{16.18}$$

мы могли бы найти предельное распределение величины

$$\sqrt{n}\ \sup_t (F_n(t, \omega) - F(t)). \tag{16.19}$$

Величина (16.18) имеет то же самое распределение, что и $\sup_t W_t^\circ$ (а именно, распределение, задаваемое формулой (11.40)), в случае, когда функция F непрерывна; в общем случае это неверно. Подобные замечания относятся и к величине

$$\sup_t |Y(t)| = \sup_t |W^\circ(F(t))|. \tag{16.20}$$

Примечание. Случайная функция (16.1) имеет независимые приращения, так же как и W; об общих результатах, касающихся слабой сходимости в этом случае, см. Кимм (1957 и 1960) и Скороход (1957). О сходимости диффузионных процессов см. Стоун (1963). Теорема 16.2 принадлежит Реньи (1958).

По поводу теоремы 16.4 см. примечания и ссылки в конце § 13 См. Шмид (1958) о распределениях величин (16.18) и (16.20) при произвольной функции распределения F.

Задачи

1. В предположениях теоремы Линдеберга — Леви существуют предельные распределения для случайных величин:

а) $\dfrac{1}{n^{3/2}} \sum\limits_{k=1}^{n} |S_k|$,

б) $\dfrac{1}{n^2} \sum\limits_{k=1}^{n} S_k^2$,

в) $\dfrac{1}{\sqrt{n}} \min\limits_{\beta n \leqslant k \leqslant n} S_k \quad 0 < \beta < 1$.

Постройте соответствующие три отображения из D в R^1 и докажите, что они измеримы и непрерывны на множестве W-меры 1. О форме предельных распределений для (а) и (б) см. Донскер (1951) и для (в) Марк (1949).

2. Теорема 16.1 (имеет смысл и) справедлива, когда n стремится к бесконечности непрерывным образом [см. (2.4)].

3. При каждом n пусть $\xi_{n1}, \ldots, \xi_{nn}$ — независимые случайные величины такие, что $\mathsf{P}\{\xi_{nk} = 1\} = p_n$ и $\mathsf{P}\{\xi_{nk} = 0\} = 1 - p_n$. Определим Y_n, положив $Y_n(t) = \sum\limits_{k \leqslant nt} \xi_{nk}$. Предполагая, что

$np_n \to \lambda$, докажите, что $Y_n \xrightarrow{\mathscr{D}} Y$, где Y — соответствующий пуассоновский процесс (с траекториями в D).

4. Покажите, что теорема 16.4 остается справедливой, если меру P, отвечающую величинам ξ_n, заменить вероятностной мерой, абсолютно непрерывной относительно P.

5. Докажите теорему 16.4, применяя теорему 15.6 с (15.39) вместо (15.21) и избегая, таким образом, представления $\xi_n = \varphi(\eta_n)$.

6. Пусть P — мера Лебега на единичном интервале Ω и $\xi_n(\omega) = 2\omega_n - 1$, где ω_n — n-й знак в двоичном разложении ω. Покажите, что в теореме 16.3 требование, заключающееся в том, что мера P_0 доминируется мерой P, существенно: например, теорема неверна, если P_0 — масса, сосредоточенная в одной точке. Теорема также перестает быть верной, если P_0 — канторовская мера (в смысле определения в книге Биллингсли (1969, стр. 48)).

7. Рассмотрите задачу 1 из § 10 применительно к D. (См. также задачу 1 из § 12.) Обобщите результаты задачи 1 настоящего параграфа.

§ 17. Случайная замена времени

Иногда возникает необходимость аппроксимировать распределения частичных сумм $S_\nu = \xi_1 + \ldots + \xi_\nu$, где индекс ν сам является случайной величиной. В этом параграфе мы докажем несколько функциональных центральных предельных теорем для таких сумм случайного числа случайных величин.

Суммы случайного числа случайных величин. Допустим, что частичные суммы $S_n = \xi_1 + \ldots + \xi_n$ при неслучайном индексе n удовлетворяют условиям центральной предельной теоремы, скажем, с нормирующими множителями $\sigma\sqrt{n}$, так что

$$\frac{1}{\sigma\sqrt{n}} S_n \xrightarrow{\mathscr{D}} N \tag{17.1}$$

при $n \to \infty$. Если ν — случайное целое число, принимающее большие значения с большой вероятностью, то имеется определенная надежда, что величина $S_\nu/(\sigma\sqrt{\nu})$ будет распределена приближенно нормально.

Для того чтобы сформулировать предельную теорему, рассмотрим последовательность $\{\nu_n\}$ случайных целых чисел. Мы ищем условия, при которых

$$\frac{1}{\sigma\sqrt{\nu_n}} S_{\nu_n} \xrightarrow{\mathscr{D}} N, \tag{17.2}$$

когда $n \to \infty$. Недостаточно, оказывается, предположить, что ν_n стремится к бесконечности по вероятности в том смысле, что при любом α

$$\mathsf{P}\{\nu_n \leqslant \alpha\} \to 0. \tag{17.3}$$

Например, допустим, что ξ_n независимы и принимают значения $+1$ и -1 с вероятностью $1/2$ каждое, так что соотношение (17.1) выполняется с $\sigma = 1$. Если ν_n есть время n-го возвращения в 0 (n-е значение k, для которого $S_k = 0$), то сходимость (17.3) имеет место (так как $\nu_n \geqslant \nu_{n-1} + 1$), но соотношение (17.2) не выполняется (поскольку $S_{\nu_n} = 0$).

Для того чтобы получить (17.2), мы должны предположить нечто большее, нежели (17.3). Наша основная цель будет состоять в том, чтобы доказать, что (17.2) выполняется, если ν_n/n сходится по вероятности к положительной конечной случайной величине, а величины ξ_n независимы и одинаково распределены со средним 0 и дисперсией σ^2. Мы получим этот результат из функциональной центральной предельной теоремы, которая дает значительно большую информацию.

Определим случайный элемент X_n пространства D равенством

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega)$$

и определим также Y_n, положив

$$Y_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{\nu_{[nt]}(\omega)}(\omega)$$

(очевидно, Y_n является случайным элементом D). Так как $Y_n(1) = S_{\nu_n}/(\sigma\sqrt{n})$, то мы будем иметь возможность аппроксимировать распределение S_{ν_n} и получить результаты, подобные (17.2), если покажем, что случайные элементы Y_n сами сходятся по распределению.

Распределение такой случайной функции, как Y_n, проще всего исследовать, используя тот факт, что Y_n получается из X_n в результате случайной замены времени. Пусть $\Phi_n(t, \omega) = \nu_{[nt]}(\omega)/n$. Так как

$$Y_n(t, \omega) = X_n(\Phi_n(t, \omega), \omega) \tag{17.4}$$

(мы допустим временно, что $\nu_n \leqslant n$, так что $\Phi_n(t, \omega) \leqslant \leqslant 1$), то Y_n получается из X_n заменой времени посредством случайной функции Φ_n. Чтобы пояснить рассуждения, используемые в этом параграфе, мы сначала рассмотрим подобные случайные изменения времени в общей ситуации. Мы предположим, что обе функции X_n и Φ_n сходятся по распределению, и найдем условия, при которых сходятся по распределению функции Y_n, определяемые формулой (17.4).

Случайная замена времени. Пусть D_0 состоит из таких элементов φ пространства D, которые являются неубывающими и удовлетворяют ограничению $0 \leqslant \varphi(t) \leqslant \leqslant 1$ при всех t. Такие функции φ представляют собой преобразование интервала времени $[0, 1]$. Мы введем в D_0 топологию, индуцированную топологией Скорохода в D. Поскольку, как легко показать, $D_0 \in \mathscr{D}$, то σ-поле $\mathscr{D}_0$ борелевских множеств в D_0 состоит из подмножеств D_0, которые принадлежат $\mathscr{D}$ (см. стр. 307).

Обозначим через $x \circ \varphi$, где $x \in D$ и $\varphi \in D_0$, композицию x и φ, т. е. функцию на $[0, 1]$, значения которой в точке t равно

$$(x \circ \varphi)(t) = x(\varphi(t)). \tag{17.5}$$

Далее, функция $x \circ \varphi$ принадлежит D, и если отображение ψ: $D \times D_0 \to D$ задано равенством

$$\psi(x, \varphi) = x \circ \varphi, \tag{17.6}$$

то, как показано на стр. 318, ψ является измеримым отображением ($\psi^{-1}\mathscr{D} \subset \mathscr{D} \times \mathscr{D}_0$).

Пусть X — случайный элемент D, и пусть Φ — случайный элемент D_0. Предположим, что X и Φ имеют одинаковую область определения, так что (X, Φ) представляет собой случайный элемент пространства $D \times D_0$ с топологией произведения (см. стр. 308). Если $X \circ \Phi$ принимает в точке ω значение $X(\omega) \circ \Phi(\omega)$ (т. е. если $X \circ \Phi = \psi(X, \Phi)$), то $X \circ \Phi$ является случайным элементом D; $X \circ \Phi$ получается из X с помощью случайной замены времени, определяемой элементом Φ.

Предположим, что помимо X и Φ при каждом n имеются случайные элементы X_n и Φ_n пространств D и D_0, соответственно, причем X_n и Φ_n имеют одну и ту же область определения (которая может меняться вместе с n). Мы ищем условия, при которых из сходимости $(X_n, \Phi_n) \xrightarrow{\mathscr{D}} (X, \Phi)$ (сходимость по распределению в топологии произведения) следует сходимость $X_n \circ \Phi_n \xrightarrow{\mathscr{D}} X \circ \Phi$.

Допустим, что

$$(X_n, \Phi_n) \xrightarrow{\mathscr{D}} (X, \Phi) \tag{17.7}$$

и

$$\mathsf{P}\{X \in C\} = \mathsf{P}\{\Phi \in C\} = 1. \tag{17.8}$$

Если мы покажем, что функция ψ, определенная формулой (17.6), непрерывна в (x, φ) при $x \in C$ и $\varphi \in C \cap D_0$, то в силу следствия 1 из теоремы 5.1 получим

$$X_n \circ \Phi_n \xrightarrow{\mathscr{D}} X \circ \Phi. \tag{17.9}$$

Если x_n сходится к x и φ_n сходится к φ в топологии Скорохода и если x и φ принадлежат C, то обе эти сходимости равномерны. Но

$$\sup_t |x_n(\varphi_n(t)) - x(\varphi(t))| \leqslant$$
$$\leqslant \sup_t |x_n(t) - x(t)| + \sup_t |x(\varphi_n(t)) - x(\varphi(t))|;$$

отсюда вытекает, поскольку функция x равномерно непрерывна, что $x_n \circ \varphi_n$ сходится к $x \circ \varphi$ равномерно и, следовательно, в топологии Скорохода, что и доказывает соотношение (17.9).

Остается, разумеется, еще вопрос о том, когда (X_n, Φ_n) сходится по распределению. Здесь могут быть использованы теоремы 4.4 и 4.5.

Приложения. Вернемся теперь к рассмотрению сумм $S_n = \xi_1 + \ldots + \xi_n$. Пусть ν_n при каждом n — положительная целочисленная случайная величина, определенная на том же самом вероятностном пространстве, что и ξ_n. Определим X_n и Y_n, положив

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega), \tag{17.10}$$

$$Y_n(t, \omega) = \frac{1}{\sigma\sqrt{\nu_n(\omega)}} S_{[\nu_n(\omega)t]}(\omega) = X_{\nu_n(\omega)}(t, \omega). \tag{17.11}$$

Теорема 17.1. *Если*

$$\frac{\nu_n}{a_n} \xrightarrow{P} \theta, \tag{17.12}$$

где θ — положительная постоянная и a_n — постоянные, стремящиеся к бесконечности, то сходимость

$$X_n \xrightarrow{\mathscr{D}} W \tag{17.13}$$

влечет сходимость

$$Y_n \xrightarrow{\mathscr{D}} W. \tag{17.14}$$

Доказательство. Без ограничения общности можно предположить, что

$$0 < \theta < 1, \tag{17.15}$$

поскольку к этому случаю всегда можно перейти, выбрав новые постоянные a_n. Кроме того, не ограничивая общности, можно считать, что a_n являются целыми числами. Определим

$$\Phi_n(t, \omega) = \begin{cases} t\dfrac{\nu_n(\omega)}{a_n} & \text{при } \dfrac{\nu_n(\omega)}{a_n} \leqslant 1, \\ t\theta & \text{в ином случае.} \end{cases} \tag{17.16}$$

Так как

$$\sup_t |\Phi_n(t) - t\theta| \leqslant \left|\frac{\nu_n}{a_n} - \theta\right| \xrightarrow{P} 0, \tag{17.17}$$

то Φ_n сходится по вероятности в топологии Скорохода к элементу $\varphi(t) = \theta t$ пространства D_0. В силу (17.13) и предположения о том, что $a_n \to \infty$, из теоремы 4.4 вытекает, что $(X_{a_n}, \Phi_n) \xrightarrow{\mathscr{D}} (W, \varphi)$. Следовательно, так как из (17.7) и (17.8) вытекает (17.9), имеем $X_{a_n} \circ \Phi_n \to W \circ \varphi$.

Если положить

$$Y'_n(t, \omega) = \frac{1}{\sigma\sqrt{a_n}} S_{[\nu_n(\omega) t]}(\omega),$$

то $X_{a_n} \circ \Phi_n$ и Y'_n имеют одинаковое значение в точке ω, если $\nu_n(\omega)/a_n \leqslant 1$. Вероятность последнего события стремится к 1 в силу (17.12) и (17.15). Следовательно, $Y'_n \xrightarrow{\mathscr{D}} W \circ \varphi$. Далее,

$$\sup_t \left|\frac{1}{\sqrt{\theta}} Y'_n(t, \omega) - Y_n(t, \omega)\right| =$$

$$= \left|\frac{1}{\sqrt{\theta}} - \sqrt{\frac{a_n}{\nu_n(\omega)}}\right| \sup_t |Y'_n(t, \omega)| \xrightarrow{P} 0$$

и, следовательно, Y_n сходится по распределению к $\theta^{-1/2}(W \circ \varphi)$. Поскольку $\theta^{-1/2}(W \circ \varphi)$ имеет то же самое распределение, что и W, то мы получаем (17.14).

Разумеется, из (17.14) следует сходимость $S_{\nu_n}/(\sigma\sqrt{\nu_n}) \xrightarrow{\mathscr{D}} N$, равно как закон арксинуса, предельные законы для максимума и тому подобное. Отметим, что мы не делали в отношении ξ_n никаких предположений помимо предположения (17.13), которое выполняется, например, в ситуации, когда действует теорема Линдеберга — Леви. В главе 4 мы докажем сходимость (17.13) для разного рода последовательностей зависимых случайных величин $\{\xi_n\}$ — к любой из них будет применима теорема 17.1.

Если предел в (17.12) не является постоянной величиной, то мы должны сделать более специальные предположения относительно ξ_n.

Теорема 17.2. *Предположим, что случайные величины $\xi_1, \xi_2, \ldots$ независимы и одинаково распределены, причем* $\mathsf{E}\{\xi_n\} = 0$ *и* $\mathsf{E}\{\xi_n^2\} = \sigma^2$. *Если*

$$\nu_n/a_n \xrightarrow{P} \theta, \tag{17.18}$$

где θ — положительная случайная величина и a_n — стремящиеся к бесконечности постоянные, то $Y_n \xrightarrow{\mathscr{D}} W$.

Доказательство. Предположим вначале, что случайная величина θ ограничена, так что существует постоянная K такая, что

$$0 < \theta \leqslant K$$

с вероятностью 1. Мы можем выбрать постоянные a_n так, чтобы они были целыми и чтобы $K < 1$.

Если мы определим Φ_n формулой (17.16), то, как и раньше, $\Phi_n \xrightarrow{P} \Phi$, где Φ — случайный элемент D_0, определенный равенством

$$\Phi(t) = \theta t.$$

Пусть $\mathscr{B}_0$ — поле цилиндрических множеств. Определим X_n' формулой (16.9), где $p_n \to \infty$ и $p_n/\sqrt{n} \to 0$. Тогда, так же как в § 16, имеем

$$\sup_t |X_n(t) - X_n'(t)| \xrightarrow{P} 0 \tag{17.19}$$

и

$$\mathsf{P}(\{X'_n \in A\} \cap E) \to W(A)\,\mathsf{P}(E)$$

для любого множества A непрерывности меры W и любого множества E из $\mathscr{B}_0$. Далее, по теореме 4.4 $(\Phi_n, \nu_n/a_n) \xrightarrow{P} (\Phi, \theta)$ в смысле топологии произведения на $D_0 \times R^1$; так как каждая функция X'_n измерима относительно $\sigma(\mathscr{B}_0)$, то из теоремы 4.5 следует, что

$$\left(X'_{a_n}, \Phi_n, \frac{\nu_n}{a_n}\right) \xrightarrow{\mathscr{D}} (W, \Phi_0, \theta_0)$$

в смысле топологии произведения в $D \times D_0 \times R^1$, где θ_0 не зависит от W и $\Phi_0(t) = \theta_0 t$. В силу (17.19)

$$\left(X_{a_n}, \Phi_n, \frac{\nu_n}{a_n}\right) \xrightarrow{\mathscr{D}} (W, \Phi_0, \theta_0).$$

Далее, отображение, которое переводит точку (x, φ, α) в $\alpha^{-1/2}(x \circ \varphi)$, непрерывно в этой точке, если $x \in C$ $\varphi \in C \cap D_0$ и $\alpha > 0$. В силу следствия 1 из теоремы 5.1 поэтому

$$(\nu_n/a_n)^{-1/2}\left(X_{a_n} \circ \Phi_n\right) \xrightarrow{\mathscr{D}} \theta_0^{-1/2}(W \circ \Phi_0).$$

Так как θ_0 и W независимы, то $\theta_0^{-1/2}(W \circ \Phi_0)$ имеет то же самое распределение, что и W. Кроме того, $(\nu_n/a_n)^{-1/2}\left(X_{a_n} \circ \Phi_n\right)$ совпадает с Y_n, если $\nu_n/a_n < 1$, а вероятность этого неравенства стремится к 1, так как $K < 1$. Таким образом, если случайная величина θ ограничена, то $Y_n \xrightarrow{\mathscr{D}} W$.

Откажемся теперь от ограниченности θ. Пусть $u > 0$; положим $\theta_u = \theta$ и $\nu_{un} = \nu_n$, если $\theta \leqslant u$, и $\theta_u = u$ и $\nu_{un} = a_n u$, если $\theta > u$. Для любого u имеем $\nu_{un}/a_n \xrightarrow{P} \theta_u$ при $n \to \infty$, так что, в силу уже доказанного, если обозначить

$$Y_{un}(t) = \frac{1}{\sigma\sqrt{\nu_{un}(\omega)}} S_{[\nu_{un}(\omega)t]}(\omega),$$

то $Y_{un} \xrightarrow{\mathscr{D}} W$ при $n \to \infty$. Поскольку $\mathsf{P}\{Y_{un} \neq Y_n\} \leqslant \leqslant \mathsf{P}\{\theta > u\}$, то $Y_n \xrightarrow{\mathscr{D}} W$ следует теперь из теоремы 4.2.

Таким образом, теорема 17.2 доказана. Приведенное доказательство применимо ко многим зависимым последовательностям $\{\xi_n\}$; см. главу 4.

Теория восстановления. Развитые выше идеи могут быть использованы для вывода центральной предельной теоремы, связанной с теорией восстановления. Пусть $\eta_1, \eta_2, \ldots$ — последовательность положительных случайных величин. Определим

$$\nu_t = \max\left\{k\colon \sum_{i=1}^{k} \eta_i \leqslant t\right\}, \quad t \geqslant 0, \tag{17.20}$$

и положим $\nu_t = 0$ при $\eta_1 > t$. Если η_k означает промежуток времени между $(k-1)$-м и k-м событиями в некоторой последовательности событий, то ν_t является числом событий, наступающих до момента времени t.

Мы предположим существование положительных постоянных μ и σ таких, что если

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{[nt]} (\eta_i(\omega) - \mu),$$

то $X_n \xrightarrow{\mathcal{D}} W$. Это предположение будет выполнено, если мы допустим, как это обычно принято в теории восстановления, что η_n независимы и одинаково распределены со средним μ и дисперсией σ^2. Определим Z_n равенством

$$Z_n(t, \omega) = \frac{\nu_{nt}(\omega) - nt/\mu}{\sigma\mu^{-3/2}\sqrt{n}}.$$

Теорема 17.3. *Если* $X_n \xrightarrow{\mathcal{D}} W$, *то* $Z_n \xrightarrow{\mathcal{D}} W$.

Доказательство. Мы будем предполагать ниже, что $\mu > 1$ (это лишь вопрос выбора шкалы). Покажем сначала, что

$$\sup_{0 \leqslant v \leqslant u} \left| \frac{\nu_v}{u} - \frac{1}{\mu}\frac{v}{u} \right| \xrightarrow{P} 0 \tag{17.21}$$

при $u \to \infty$. Из условия $X_n \xrightarrow{\mathcal{D}} W$ следует, что

$$\sup_{0 \leqslant t \leqslant s} \frac{1}{s} \left| \sum_{i=1}^{[t]} \eta_i - \mu t \right| \xrightarrow{P} 0 \tag{17.22}$$

при $s \to \infty$ (так как замена s^{-1} на $s^{-1/2}$ приводит к сходимости по распределению). В силу определения (17.20) из неравенства $\nu_v > t$ вытекает, что $\sum_{i=0}^{[t]} \eta_i \leqslant v$. Поэтому из неравенства

$$\sup_{0 \leqslant v \leqslant u} \left(\frac{\nu_v}{u} - \frac{1}{\mu} \frac{v}{u} \right) > \varepsilon$$

следует, что

$$\sup_{0 \leqslant t \leqslant u(\mu^{-1}+\varepsilon)} \left| \sum_{i=1}^{[t]} \eta_i - \mu t \right| \geqslant \mu u \varepsilon. \tag{17.23}$$

Аналогично, неравенство $\inf\limits_{0 \leqslant v \leqslant u} \left(\frac{\nu_v}{u} - \frac{1}{\mu} \frac{v}{u} \right) < -\varepsilon$ влечет (при $\varepsilon < \mu^{-1}$)

$$\sup_{0 \leqslant t \leqslant u(\mu^{-1}-\varepsilon)} \left| \sum_{i=1}^{[t]} \eta_i - \mu t \right| \geqslant \mu u \varepsilon. \tag{17.24}$$

В силу (17.22) вероятности неравенств (17.23) и (17.24) стремятся к 0 при $u \to \infty$, что и доказывает соотношение (17.21).

Положим

$$\Phi_n(t, \omega) = \begin{cases} \dfrac{\nu_{tn}(\omega)}{n} & \text{при } \dfrac{\nu_n(\omega)}{n} \leqslant 1, \\ \dfrac{t}{\mu} & \text{в ином случае.} \end{cases}$$

В силу (17.21) $\Phi_n \xrightarrow{P} \varphi$, где $\varphi(t) = t/\mu$, откуда стандартным путем получаем, что $X_n \circ \Phi_n \xrightarrow{\mathscr{D}} W \circ \varphi$. Пусть

$$Y_n(t) = \frac{1}{\sigma \sqrt{n}} \sum_{i=1}^{\nu_{tn}(\omega)} (\eta_i(\omega) - \mu).$$

Имеем $Y_n = X_n \circ \Phi_n$, если $\nu_n/n \leqslant 1$, причем вероятность последнего неравенства стремится к 1 в силу (17.21) и предположения $\mu > 1$. Следовательно, $Y_n \xrightarrow{\mathscr{D}} W \circ \varphi$.

Из определения (17.20) имеем

$$Y_n(t) \leqslant \frac{nt - \mu \nu_{nt}}{\sigma \sqrt{n}} \leqslant Y_n(t) + \frac{\eta_{\nu_{nt}+1}}{\sigma \sqrt{n}}.$$

Так как $\max_{i\leqslant n}|\eta_i|/\sqrt{n}\xrightarrow{P}0$, то $\sup_{t\leqslant 1}|\eta_{\nu_{nt}}|/\sigma\sqrt{n}\xrightarrow{P}0$, откуда в свою очередь вытекает, что если

$$Z'_n(t)=\frac{nt-\mu\nu_{nt}}{\sigma\sqrt{n}},$$

то $Z'_n\xrightarrow{\mathcal{D}}W\circ\varphi$. Поэтому $\mu^{1/2}Z'_n\xrightarrow{\mathcal{D}}W$ (напомним, что $\varphi(t)=t/\mu$), откуда вследствие симметрии W вытекает, что $Z_n\xrightarrow{\mathcal{D}}W$.

В теореме 17.3 можно устремлять n к бесконечности непрерывным образом.

Примечание. Результаты этого параграфа являются новыми и обобщают результаты работы Биллингсли (1962). О центральной предельной теореме для случайного числа случайных величин см. Реньи (1960) и Блюм, Хэнсон и Розенблатт (1963). См. Ламперти (1962) по поводу другой теоремы о слабой сходимости в теории восстановления.

§ 18. Равномерная топология

Рассмотрим кратко равномерную топологию на D, т. е. топологию, задаваемую равномерной метрикой

$$\rho(x, y)=\sup_t|x(t)-y(t)|. \tag{18.1}$$

Относительно этой метрики D является полным метрическим пространством. Однако оно не сепарабельно: например, точки x_θ, $0<\theta<1$, где

$$x_\theta(t)=\begin{cases}0, & \text{если } 0\leqslant t<\theta,\\ 1, & \text{если } \theta\leqslant t\leqslant 1,\end{cases} \tag{18.2}$$

образуют несчетное множество с расстоянием между любыми двумя его элементами, равными 1 (см. стр. 295).

Условимся, что любое понятие, имеющее префикс U, будет относиться к равномерной топологии: U-открытое, U-непрерывное и т. д. Таким же образом префикс S будет использоваться для топологии Скорохода в D. Если d — любая из двух метрик, которые задают топологию Скорохода (см. § 14), и если ρ — равномерная метрика (18.1), то $d(x, y)\leqslant\rho(x, y)$. Поэтому равномерная топология сильней, чем топология Скорохода: U-сходимость влечет S-сходимость. С другой стороны, как отмечалось в § 14, если x_n S-сходится к x и $x \in C$, то x_n U-сходится к x. В частности, равномерная и скороходовская топологии на D редуцируют одну и ту же топологию на C.

Пусть $\mathscr{D}$ обозначает, как обычно, класс S-борелевских множеств, и пусть $\mathscr{U}$ обозначает класс U-борелевских множеств. Поскольку S-открытое множество является в то же время U-открытым, то

$$\mathscr{D} \subset \mathscr{U}. \tag{18.3}$$

Пусть P_n и P — вероятностные меры, заданные на $\mathscr{U}$, *более широком* из двух классов $\mathscr{U}$ и $\mathscr{D}$. Будем обозначать $P_n \Rightarrow P[U]$ слабую сходимость в смысле равномерной топологии и $P_n \Rightarrow P[S]$ слабую сходимость в смысле топологии Скорохода. Сходимость $P_n \Rightarrow P[U]$ равносильна сходимости

$$\int f\,dP_n \to \int f\,dP \tag{18.4}$$

для всех ограниченных, U-непрерывных функций f, тогда как сходимость $P_n \Rightarrow P[S]$ равносильна (18.4) лишь для всех ограниченных, S-непрерывных f. Следовательно, из $P_n \Rightarrow P[U]$ вытекает $P_n \Rightarrow P[S]$. Обратное утверждение неверно, так как из S-сходимости не следует U-сходимость (достаточно рассмотреть точечные распределения).

Предположим теперь, что P_n и P определены на $\mathscr{U}$, что $P_n \Rightarrow P[S]$ и что $P(C) = 1$. Покажем, что $P_n \Rightarrow P[U]$. Для любого множества A из $\mathscr{U}$ обозначим через A_U его U-замыкание и через A_S — его S-замыкание. Заметим, что $A_U \subset A_S$. Так как $P_n \Rightarrow P[S]$ и $P(C) = 1$, то $\limsup_n P_n(A) \leqslant \limsup_n P_n(A_S) \leqslant P(A_S) = P(A_S \cap C)$. Поскольку всякая последовательность, S-сходящаяся к некоторому пределу в C, в то же время U-сходится к этому же пределу, то $A_S \cap C \subset A_U$. Следовательно, $\limsup_n P_n(A) \leqslant P(A_U)$, откуда в силу теоремы 2.1 вытекает, что $P_n \Rightarrow P[U]$.

Нами доказаны следующие утверждения. *Предположим, что P_n и P определены на $\mathscr{U}$.* (i) *Если $P_n \Rightarrow P[U]$, то $P_n \Rightarrow P[S]$.* (ii) *Если $P_n \Rightarrow P[S]$ и $P(C) = 1$, то $P_n \Rightarrow$*

$\Rightarrow \mathsf{P}[U]$ *). Ввиду (i) лучше по возможности доказывать $P_n \Rightarrow P[U]$, чем $P_n \Rightarrow P[S]$, так как, например, при этом мы получаем асимптотические распределения для большего числа функций на D. С другой стороны, ввиду (ii), если $P(C) = 1$, то доказательство сходимости $P_n \Rightarrow \Rightarrow P[S]$ автоматически является доказательством лучшего результата: $P_n \Rightarrow P[U]$.

Рассмотрим распределения P_n случайных функций X_n, о которых идет речь в теореме Донскера, сформулированной для пространства D (см. (16.1)). Согласно теореме 16.1 $P_n \Rightarrow W[S]$, где W — винеровская мера. Поскольку $W(C) = 1$, мы можем вывести более сильное заключение: $P_n \Rightarrow W[U]$.

В этом рассуждении, однако, имеется пробел: меры P_n и W определены только на $\mathscr{D}$. Пока остается открытым вопрос о том, можно ли продолжать эти меры на более широкое σ-поле $\mathscr{U}$. Разумеется, мера W может быть продолжена — рассуждения, используемые в § 16 для продолжения W с $(C, \mathscr{C})$ на $(D, \mathscr{D})$, позволяют нам продолжить также W с $(C, \mathscr{C})$ на $(D, \mathscr{U})$. Далее, мера P_n определена как $\mathsf{P}X_n^{-1}$ (областью определения X_n служит вероятностное пространство $(\Omega, \mathscr{B}, \mathsf{P})$), что возможно в силу включения

$$X_n^{-1}\mathscr{D} \subset \mathscr{B}. \tag{18.5}$$

Если нам нужно определить P_n на более широком σ-поле $\mathscr{U}$, то мы должны заменить (18.5) более сильным условием

$$X_n^{-1}\mathscr{U} \subset \mathscr{B}. \tag{18.6}$$

Если A — конечномерное множество в пространстве D, то $X_n^{-1}A \in \mathscr{B}$, и, следовательно, это верно и в том случае, когда A является U-сферой. Однако так как равномерная топология не сепарабельна, то соотношение (18.6) не вытекает отсюда без дополнительных рассуждений.

---

*) Эти утверждения справедливы вследствие того, что U-сходимость всегда влечет S-сходимость, а S-сходимость к пределу в C влечет U-сходимость; они не затрагивают никаких других свойств C, D и двух рассматриваемых топологий.

Из определения (16.1) функций X_n легко вытекает, что область значений $X_n\Omega$ сепарабельна в равномерной топологии. Поэтому если множество A U-открыто, то существует счетная совокупность U-сфер A_i такая, что $A \cap X_n\Omega = \bigcup_i (A_i \cap X_n\Omega)$. При этом $X_n^{-1}A = \bigcup_i X_n^{-1}A_i$, и так как каждое множество $X_n^{-1}A_i$ принадлежит $\mathscr{B}$, то и $X_n^{-1}A$ принадлежит $\mathscr{B}$, что и доказывает (18.6).

Таким образом, распределение P_n определяется (или может быть определено) на $\mathscr{U}$ естественным способом:

$$P_n(A) = \mathsf{P}\{\omega\colon\ X_n(\omega) \in A\}, \quad A \in \mathscr{U},$$

и мы можем усилить теорему 16.1, понимая под сходимостью $X_n \xrightarrow{\mathscr{D}} W$ слабую сходимость в равномерной топологии: $P_n \Rightarrow W[U]$.

Рассмотрим далее случайные функции Y_n, определенные равенством

$$Y_n(\omega, t) = \sqrt{n}\,(F_n(t, \omega) - t),$$

где $F_n(t, \omega)$ — эмпирическая функция распределения для величин $\xi_1(\omega), \ldots, \xi_n(\omega)$, которые предполагаются независимыми и равномерно распределенными на единичном интервале. Из теоремы 16.4 следует, что $Y_n \xrightarrow{\mathscr{D}} W^\circ$. Поскольку мера W° точно так же, как мера W, может быть продолжена на $\mathscr{U}$, то мы можем понимать сходимость $Y_n \xrightarrow{\mathscr{D}} W^\circ$ в равномерной топологии, а не в топологии Скорохода, если только сможем продолжить распределение элемента Y_n с $\mathscr{D}$ на $\mathscr{U}$, т. е. если мы сможем усилить включение

$$Y_n^{-1}\mathscr{D} \subset \mathscr{B} \tag{18.7}$$

до

$$Y_n^{-1}\mathscr{U} \subset \mathscr{B}. \tag{18.8}$$

Но в этом месте наша программа не проходит: мы покажем, что соотношение (18.8) неверно.

Чтобы убедиться в том, что включение (18.8) не выполняется, рассмотрим случай $n = 1$. Пусть h — отображение Ω в D, которое сопоставляет ω функцию распределения с единичным скачком в точке $\xi_1(\omega)$; имеем $h(\omega)=x_{\xi_1(\omega)}$, где x_θ определяется формулой (18.2). Очевидно, что (18.8) эквивалентно соотношению

$$h^{-1}\mathcal{U}\subset\mathcal{B}, \tag{18.9}$$

и мы покажем, что (18.9) неверно, если величина $\xi_1(\omega)$ равномерно распределена.

Пусть A_θ обозначает U-шар с центром x_θ и радиусом 1/2. Так как точки x_θ находятся на расстоянии 1 друг от друга в равномерной топологии, то мы имеем

$$h^{-1}\Big(\bigcup_{\theta\in H}A_\theta\Big)=\{\omega:\ \xi_1(\omega)\in H\} \tag{18.10}$$

для любого подмножества H единичного интервала. Если бы включение (18.9) было верно, то, так как множество $\bigcup_{\theta\in H}A_\theta$ U-открыто, множество (18.10) принадлежало бы $\mathcal{B}$, и поэтому для *каждого* подмножества H единичного интервала мы имели бы $\{\omega:\ \xi_1(\omega)\in H\}\in\mathcal{B}$. Таким образом, мера $\mu=\mathsf{P}\xi_1^{-1}$ была бы мерой, заданной на каждом подмножестве единичного интервала и обладающей тем свойством, что $\mu(a,b)=b-a$ для любого подинтервала (a,b). Однако такой меры μ не существует.

Несколько развив эти идеи, можно показать, что (18.8) неверно при любом n, если только распределение, общее для всех ξ_i, не является чисто атомическим. Поэтому при рассмотрении эмпирических функций распределения как случайных элементов пространства D мы должны придерживаться топологии Скорохода — и это исключительно из соображений измеримости *). Как

---

*) О, знал бы человек, сколь тяжки хвори
Приносит невоздержанность обжоре,
В пирах он стал бы более умерен ...

Перевод отрывка из «Кентерберийских рассказов» Дж. Чосера (Chaucer)—«Рассказ монаха» (The Pardoner's Tale):

O, wiste a man how manye maladyes
Folwen of excesse and of glotonyes,
He wolde been the moore mesurable ...

правило, мы можем интерпретировать сходимость $X_n \xrightarrow{\mathscr{D}} X$ в равномерной топологии лишь в двух случаях: (i) когда X принадлежит C с вероятностью 1 и (ii) когда скачки X_n появляются не в случайные, а в фиксированные моменты времени.

Примечание. Чибисов (1965) отметил, что (18.8) неверно. Другой путь в обход проблем, порожденных несепарабельностью пространства D в равномерной топологии, состоит в том, чтобы рассматривать σ-поле $\mathscr{U}_0$, порожденное U-шарами. Для этого требуется иная теория слабой сходимости; см. Дадли (1966 и 1967).

Глава 4

ЗАВИСИМЫЕ ВЕЛИЧИНЫ

§ 19. Диффузия

Настоящий параграф содержит теорему, характеризующую броуновское движение, теорему, характеризующую более общие диффузионные процессы, и асимптотические варианты этих двух теорем. В § 20 полученные результаты используются при доказательстве варианта теоремы Донскера для величин, образующих стационарную последовательность, которая удовлетворяет условию равномерного перемешивания; в § 21 этот результат обобщается на функции от таких последовательностей, а в § 22 доказываются функциональные центральные предельные теоремы для эмпирических функций распределения в обоих этих случаях. В §§ 23 и 24 результаты настоящего параграфа используются для вывода предельных теорем для мартингалов и симметрично зависимых случайных величин *).

Несмотря на то, что многие результаты этой главы могли бы быть сформулированы и доказаны в пространстве C, мы будем последовательно использовать пространство D с введенной в нем топологией Скорохода.

Характеризация броуновского движения. Случайная функция W принадлежит C с вероятностью 1 и имеет независимые приращения. Кроме того, $\mathsf{E}\{W_t\}=0$ и $\mathsf{E}\{W_t^2\}=t$. Эти свойства полностью характеризуют W:

Теорема 19.1. *Пусть X — случайный элемент в D, который имеет независимые приращения и удовлетворяет*

---

*) Все последующие параграфы зависят от настоящего. §§ 20, и 22 следует читать подряд, но эти три параграфа в совокупности, § 23 и § 24 представляют собой самостоятельные части.

условиям

$$\mathsf{P}\{X \in C\} = 1, \quad \mathsf{E}\{X(t)\} = 0 \quad и \quad \mathsf{E}\{X^2(t)\} = t.$$

Тогда X имеет распределение, совпадающее с W.

Доказательство. Из независимости и условий, налагаемых на моменты, следует, что

$$\mathsf{E}\{(X(t) - X(s))\} = |t - s|$$

и $\mathsf{P}\{X(0) = 0\} = 1$. Мы должны показать, что $X(t) - X(s)$ имеет распределение $N(0, t-s)$ при $s \leqslant t$. Так как приращения независимы, достаточно доказать это при $s = 0$, а в силу непрерывности мы можем считать, что $t < 1$.

Пусть $\varphi(t, \cdot)$ — характеристическая функция величины $X(t)$:

$$\varphi(t, u) = \mathsf{E}\{e^{iuX(t)}\}.$$

Поскольку $\mathsf{P}\{X \in C\} = 1$, функция $\varphi(t, u)$ непрерывна и по t, и по u. Мы покажем, что $\varphi(t, u)$ удовлетворяет дифференциальному уравнению

$$\frac{\partial}{\partial t}\varphi(t, u) = -\frac{1}{2}u^2\varphi(t, u), \quad 0 \leqslant t < 1, \quad u \in R^1. \tag{19.1}$$

Отсюда будет следовать *), что

$$\varphi(t, u) = \varphi(0, u)\, e^{-\frac{1}{2}tu^2};$$

так как $X(0) = 0$ с вероятностью 1, это повлечет за собой нормальность. При выводе уравнения (19.1) мы можем понимать частную производную, как производную

*) Если f удовлетворяет уравнению $f'(t) = A(t)f(t)$ с непрерывной функцией A, то производная отношения f к необращающейся в 0 функции $f_0(t) = \exp\left\{\int_0^t A(\tau)\,d\tau\right\}$ равна 0, так что $f(t)/f_0(t) = f(0)/f_0(0) = f(0)$.

справа *). Таким образом, мы должны доказать

$$\lim_{h\downarrow 0}\frac{1}{h}\left[\varphi(t+h,\,u)-\varphi(t,\,u)\right]=-\frac{1}{2}u^2\varphi(t,\,u),\quad 0\leqslant t<1. \tag{19.2}$$

Используя стандартную оценку **), имеем

$$e^{iv}=1+iv-\frac{1}{2}v^2+c(v), \tag{19.3}$$

где

$$|c(v)|\leqslant|v|^3 \tag{19.4}$$

для действительных v. Обозначим

$$\Delta_{s,t}=\Delta(s,\,t)=X(t)-X(s). \tag{19.5}$$

Из независимости этих приращений и из соотношений $\mathsf{E}\{\Delta_{s,t}\}=0$ и $\mathsf{E}\{\Delta_{s,t}^2\}=t-s$ имеем

$$\varphi(t+h,\,u)-\varphi(t,\,u)=\mathsf{E}\{e^{iuX(t)}[e^{iu\Delta(t,\,t+h)}-1]\}=$$
$$=\mathsf{E}\left\{e^{iuX(t)}\left[iu\Delta_{t,\,t+h}-\frac{1}{2}u^2\Delta_{t,\,t+h}^2+c(u\Delta_{t,\,t+h})\right]\right\}=$$
$$=-\frac{1}{2}u^2h\varphi(t,\,u)+\mathsf{E}\{e^{iuX(t)}c(u\Delta_{t,\,t+h})\}.$$

Ввиду оценки (19.4) соотношение (19.2) будет иметь место, если мы докажем, что

$$\lim_{h\downarrow 0}\frac{1}{h}\mathsf{E}\{|\Delta_{t,\,t+h}|^3\}=0. \tag{19.6}$$

Снова используя независимость приращений, мы получаем

$$\mathsf{E}\{\Delta_{t_1,\,t}^2\Delta_{t,\,t_2}^2\}=(t-t_1)(t_2-t)\leqslant(t_2-t_1)^2,\quad t_1\leqslant t\leqslant t_2. \tag{19.7}$$

*) Чтобы доказать, что непрерывная функция f с непрерывной правосторонней производной f^+ на $[0,\,1)$ имеет и двустороннюю производную на $(0,\,1)$, достаточно предположить f действительной и показать, что функция $F(t)=f(t)-f(0)-\int\limits_0^t f^+(\tau)\,d\tau$ тождественно равна нулю. Если, например, $F(t_0)<0$, то функция $G(t)=F(t)-tF(t_0)/t_0$ удовлетворяет условиям $G(0)=G(t_0)=0$ и (так как F^+ обращается в нуль) $G^+(t)>0$, так что на $[0,\,t_0]$ G должна иметь положительный максимум в некоторой внутренней точке s; но тогда $G^+(s)\leqslant 0$, и мы получаем противоречие.

**) Феллер (1967b стр. 584).

Зафиксируем t и $t+h$, где $h>0$. Если

$$M'_m = \max_{0\leqslant i\leqslant m} \min\left\{\left|\Delta\left(t, t+\frac{i}{m}h\right)\right|, \left|\Delta\left(t+\frac{i}{m}h, t+h\right)\right|\right\}, \tag{19.8}$$

то в силу формулы (19.7) и теоремы 12.1

$$\mathsf{P}\{M'_m \geqslant \lambda\} \leqslant K_{2,1}\frac{h^2}{\lambda^4}. \tag{19.9}$$

Далее, согласно (12.6)

$$|\Delta_{t,\,t+h}| \leqslant 3M'_m + \max_{1\leqslant i\leqslant m}\left|\Delta\left(t+\frac{i-1}{m}h, t+\frac{i}{m}h\right)\right|. \tag{19.10}$$

Так как X принадлежит C с вероятностью 1, максимум в правой части (19.10) стремится к нулю с вероятностью 1 при $m\to\infty$ и, используя неравенство (19.9), получаем

$$\mathsf{P}\{|\Delta_{t,\,t+h}| \geqslant \lambda\} \leqslant K\frac{h^2}{\lambda^4}, \tag{19.11}$$

где $K = 4^4 K_{2,1}$.

Из (19.11) следует (см. (3) на стр. 305) оценка

$$\mathsf{E}\{|\Delta_{t,\,t+h}|^3\} \leqslant a + \int_a^\infty \frac{Kh^2}{\alpha^{4/3}}\,d\alpha = a + 3Kh^2\frac{1}{a^{1/3}},$$

справедливая при $a>0$. Выбирая $a = K^{3/4}h^{3/2}$ (это значение минимизирует правую часть), мы получаем

$$\mathsf{E}\{|\Delta_{t,\,t+h}|^3\} \leqslant 4K^{3/4}h^{3/2},$$

что влечет за собой (19.6) и этим завершает доказательство.

Условие $\mathsf{P}\{X\in C\} = 1$ здесь существенно, так как остальные предположения теоремы выполнены, если

$$X(t) = Y(t) - t, \tag{19.12}$$

где $Y(t)$ — пуассоновский процесс с параметром 1.

Теперь мы выведем из теоремы 19.1 одну предельную теорему. Пусть X_n — случайные элементы в D. Будем говорить, что X_n имеет *асимптотически независимые приращения*, если для

$$0\leqslant s_1\leqslant t_1 < s_2 \leqslant t_2 < \ldots < s_r \leqslant t_r \leqslant 1 \tag{19.13}$$

разность

$$\mathsf{P}\{X_n(t_i)-X_n(s_i)\in H_i,\ i=1,\ \ldots,\ r\}-\\ -\prod_{i=1}^{r}\mathsf{P}\{X_n(t_i)-X_n(s_i)\in H_i\} \qquad (19.14)$$

стремится к 0 при $n\to\infty$ для любых линейных борелевских множеств $H_1,\ \ldots,\ H_r$. Отметим, что мы требуем выполнения (19.14) лишь только для строго разделенных интервалов $[s_i, t_i]$, т. е. при $t_{i-1}<s_i$.

Напомним, что модуль непрерывности $w_x(\delta)=w(x, \delta)$, определенный формулой (8.1), имеет смысл и для x, принадлежащих D. Модуль непрерывности появляется в нижеследующей теореме в силу того, что носителем предельного распределения на D служит C.

Теорема 19.2. *Пусть X_n имеет асимптотически независимые приращения, последовательность $\{X_n^2(t);\ n\geqslant 1\}$ равномерно интегрируема при каждом t, $\mathsf{E}\{X_n(t)\}\to 0$ и $\mathsf{E}\{X_n^2(t)\}\to t$ при $n\to\infty$. Пусть, далее, для любых положительных чисел ε и η существует положительное число δ такое, что при всех достаточно больших n*

$$\mathsf{P}\{w(X_n, \delta)\geqslant\varepsilon\}\leqslant\eta. \qquad (19.15)$$

Тогда $X_n\xrightarrow{\mathscr{D}}W$.

Доказательство. Так как $\mathsf{E}\{X_n^2(0)\}\to 0$, то последовательность $\{X_n(0)\}$ является плотной. Из соотношения (19.15) и теоремы 15.5 следует, что последовательность $\{X_n\}$ плотна, и если X есть предел сходящейся по распределению подпоследовательности X_n, то $\mathsf{P}\{X\in C\}=1$. Теперь достаточно показать, что каждая такая функция распределена как W. Но из соотношений $\mathsf{E}\{X_n(t)\}\to 0$, $\mathsf{E}\{X_n^2(t)\}\to t$ и равномерной интегрируемости последовательности $\{X_n^2(t);\ n\geqslant 1\}$ следует, что $\mathsf{E}\{X(t)\}=0$ и $\mathsf{E}\{X^2(t)\}=t$ (теорема 5.4). Далее, если имеет место (19.13), то приращения $X(t_i)-X(s_i)$, $i=1,\ \ldots,\ r$, независимы, а так как $\mathsf{P}\{X\in C\}=1$, то же самое верно, если допустить равенства $t_i=s_{i+1}$ в (19.13). Таким образом, утверждение теоремы следует из теоремы 19.1.

В (19.15) нельзя заменить w на модуль непрерывности w', определенный формулой (14.6) (чтобы в этом убедиться, достаточно положить X_n тождественно равным X из формулы (19.12)).

Условие равномерной интегрируемости последовательности $\{X_n^2(t);\ n \geqslant 1\}$ при каждом t в доказанной теореме существенно; в самом деле, в силу теоремы 5.4 это условие является следствием соотношений $\mathsf{E}\{X_n^2(t)\} \to t$ и $X_n \xrightarrow{\mathcal{D}} W$. В качестве конкретного примера, когда это условие не выполняется, возьмем $X_n(t) = \sqrt{t}\,\xi_n$, где ξ_n принимает значения $\sqrt{n}$, $-\sqrt{n}$ и 0 с соответствующими вероятностями $1/(2n)$, $1/(2n)$ и $1 - 1/n$; здесь, несмотря на то, что X_n не сходится по распределению к W, выполнены все предположения теоремы, кроме требования равномерной интегрируемости.

Теорема 19.2 вместе с результатами § 12 приводит к еще одному доказательству теоремы Донскера. Пусть X_n — случайная функция (10.1), участвующая в этой теореме. Тогда в силу соотношения (12.23) и теоремы 8.4 (перенесенной на D) выполняется условие (19.15). Далее, в силу (12.20)

$$\int\limits_{\{S_n^2/(n\sigma^2) \geqslant a\}} S_n^2/n\, d\mathsf{P} \leqslant K' \left[\frac{1}{a} + \frac{1}{\sigma^2} \int\limits_{\left\{|\xi_1| \geqslant \frac{1}{4} a\sigma\sqrt{n}\right\}} \xi_1^2\, d\mathsf{P} \right],$$

где K' — абсолютная постоянная. Отсюда следует равномерная интегрируемость последовательности $\{X_n^2(t),\ n \geqslant 1\}$. Остальные условия теоремы 19.2 проверяются без труда, и мы получаем $X_n \xrightarrow{\mathcal{D}} W$. Заметим, что проведенное доказательство не опирается на центральную предельную теорему; в частности, мы получаем независимое доказательство теоремы Линдеберга — Леви *). Как мы увидим в § 20, этот метод применим также к некоторым последовательностям зависимых случайных величин.

---

*) Рассуждение легко приспособить и к случаю Линдеберга; см. задачу 1 § 10 и задачу 1 § 12.

Другие диффузионные процессы*). Если X удовлетворяет условиям теоремы 19.1, то при $0 \leqslant t_1 \leqslant \ldots \leqslant t_k < 1$ и положительном h

$$\begin{aligned} \mathsf{E}\{X(t_k+h)-X(t_k)\,\|\,X(t_1),\ \ldots,\ X(t_k)\} &= 0, \\ \mathsf{E}\{(X(t_k+h)-X(t_k))^2\,\|\,X(t_1),\ \ldots,\ X(t_k)\} &= h. \end{aligned} \tag{19.16}$$

Если $\mathsf{P}(X \in C) = 1$ и $\mathsf{P}\{X(0)=0\} = 1$, то, как мы увидим, из уравнений (19.16) следует, что X описывает броуновское движение и, следовательно, можно обойтись без предположения о независимости приращений. То же самое верно, если в добавление к некоторым условиям регулярности потребовать лишь *приближенное* (в уточняемом ниже смысле) выполнение соотношений (19.16) при малых h.

Диффузионные процессы, отличные от броуновского движения, можно охарактеризовать подобным же способом, заменив правые части уравнений (19.16) функциями от h, t_k и $X(t_k)$. Мы предположим, что для малых h

$$\mathsf{E}\{X(t_k+h)-X(t_k)\,\|\,X(t_1),\ \ldots,\ X(t_k)\} \approx h\rho(t_k)X(t_k) \tag{19.17}$$

и

$$\mathsf{E}\{(X(t_k+h)-X(t_k))^2\,\|\,X(t_1),\ \ldots,\ X(t_k)\} \approx h\sigma^2(t_k) \tag{19.18}$$

(смысл этих предположений будет уточнен ниже), где $\rho(t)$ и $\sigma^2(t)$ — заданные функции, и докажем (см. ниже теорему 19.3), что X — гауссовская случайная функция, удовлетворяющая условиям

$$\mathsf{E}\{X(t)\} = 0, \qquad 0 \leqslant t \leqslant 1, \tag{19.19}$$

и

$$\mathsf{E}\{X(s)X(t)\} = \int_0^s \sigma^2(r)\exp\left[2\int_r^s \rho(\tau)\,d\tau + \int_s^t \rho(\tau)\,d\tau\right]dr, \tag{19.20}$$

$$0 \leqslant s \leqslant t \leqslant 1.$$

---

*) Остальная часть настоящего параграфа используется только в §§ 23 и 24.

При $\rho(t)=0$ и $\sigma^2(t)=1$ мы получаем опять броуновское движение, а при $\rho(t)=-1/(1-t)$ и $\sigma^2(t)=1$ — броуновский мост.

Мы будем предполагать, что функции $\sigma^2(t)$ и $\rho(t)$ непрерывны на $[0, 1]$, функция $\sigma^2(t)$ неотрицательна и предел

$$\lim_{t\to 1} g(t)\,G(t), \tag{19.21}$$

где

$$g(t)=\exp\left[\int_0^t \rho(\tau)\,d\tau\right]$$

и

$$G(t)=\int_0^t \sigma^2(r)\,g^{-2}(r)\,dr,$$

существует и конечен. При $t\to 1$ неубывающая функция $G(t)$ имеет предел (конечный или бесконечный) и, следовательно, $g(t)$ имеет конечный предел (который равен нулю, если $G\to\infty$). Таким образом, правая часть (19.20), определенная при $0\leqslant s\leqslant t<1$, имеет при $t\to 1$ предел, который мы будем считать ее значением при $t=1$.

Если определить $X(t)$ равенством

$$X(t)=g(t)\,(1+G(t))\,W^\circ(G(t)/(1+G(t)))$$

для $0\leqslant t<1$, где W° — броуновский мост, то свойство (19.20) выполняется при $0\leqslant s\leqslant t<1$. Поскольку все функции $g(t)$, $g(t)\,G(t)$ и $G(t)/(1+G(t))$ имеют предел при $t\to 1$, мы можем определить $X(1)$, переходя к пределу при $t\to 1$. Таким образом, действительно существует непрерывная гауссовская случайная функция, удовлетворяющая условиям (19.19) и (19.20). Очевидно, $X(0)=0$ с вероятностью 1. Мы будем рассматривать X как случайный элемент пространства D, принадлежащий с вероятностью 1 пространству C.

Сформулируем три условия, которые характеризуют X как такую случайную функцию. Каждое условие будет дано в двух вариантах: первый — в форме, удобной для большинства приложений, второй — в значительно более слабой форме, которая достаточна для характеризации. Первое условие уточняет, в каком смысле понимаются аппроксимационные формулы (19.17) и (19.18).

Условие 1. *Если* $0 \leqslant t_1 \leqslant \ldots \leqslant t_k < 1$, *то*

$$\lim_{h \downarrow 0} \frac{1}{h} \mathsf{E}\{|\mathsf{E}\{X(t_k+h) - X(t_k) \,\|\, X(t_1), \ldots, X(t_k)\} - \\ - h\rho(t_k) X(t_k)|\} = 0. \quad (19.22)$$

и

$$\lim_{h \downarrow 0} \frac{1}{h} \mathsf{E}\{|\mathsf{E}\{(X(t_k+h) - X(t_k))^2 \,\|\, X(t_1), \ldots, X(t_k)\} - \\ - h\sigma^2(t_k)|\} = 0. \quad (19.23)$$

Условие 1а. *Если* $0 \leqslant t_1 \leqslant \ldots \leqslant t_k < 1$, *то для всех действительных* $u_1, \ldots, u_k$

$$\lim_{h \downarrow 0} \frac{1}{h} \mathsf{E}\left\{\exp\left[\sum_{j=1}^{k} iu_j X(t_j)\right] \times \right. \\ \left. \times [X(t_k+h) - X(t_k) - h\rho(t_k) X(t_k)]\right\} = 0 \quad (19.24)$$

и

$$\lim_{h \downarrow 0} \frac{1}{h} \mathsf{E}\left\{\exp\left[\sum_{j=1}^{k} iu_j X(t_j)\right] \times \right. \\ \left. \times [(X(t_k+h) - X(t_k))^2 - h\sigma^2(t_k)]\right\} = 0. \quad (19.25)$$

Условие 2. *Справедливо неравенство*

$$\sup_t \mathsf{E}\{X^2(t)\} < \infty. \quad (19.26)$$

Условие 2а. *Случайные величины* $X(t)$, $0 \leqslant t \leqslant 1$, *равномерно интегрируемы.*

Условие 3. *Существует константа* K *такая, что*

$$\mathsf{E}\{(X(t) - X(t_1))^2 (X(t_2) - X(t))^2\} \leqslant K(t_2 - t_1)^2, \quad (19.27) \\ t_1 \leqslant t \leqslant t_2.$$

Условие 3а. *При* $t < 1$

$$\lim_{\alpha \to \infty} \limsup_{h \downarrow 0} \frac{1}{h} \int\limits_{\{(X(t+h) - X(t))^2 \geqslant \alpha h\}} (X(t+h) - X(t))^2 \, d\mathsf{P} = 0. \quad (19.28)$$

Очевидно, что условия 1 и 2 влекут, соответственно, условия 1а и 2а. Для того чтобы убедиться, что условие 3 влечет условие 3а (которое по существу является условием равномерной интегрируемости), заметим, что при переходе к обозначению (19.5) неравенство (19.27) превращается в неравенство (19.7) с дополнительным множителем K в правой части, и если X принадлежит C с вероятностью 1, то неравенство (19.11) можно вывести также, как и раньше (правда, с новой константой K, значение которой здесь несущественно). Используя соотношение (3) со стр. 305, получаем, что интеграл в (19.28) не может превосходить $2Kh/\alpha$, так что условие 3а действительно является следствием условия 3. (Можно получить условия, промежуточные между 2 и 2а и между 3 и 3а, заменив показатель степени 2 в неравенствах (19.26) и (19.27) на $1+\varepsilon$.)

Теорема 19.3. *Пусть X — случайный элемент в пространстве D, для которого* $\mathsf{P}\{X \in C\} = 1$ *и* $\mathsf{P}\{X(0) = 0\} = 1$. *Предположим, что функции* $\sigma^2(t)$ *и* $\rho(t)$ *непрерывны на* $[0, 1)$, *а предел* (19.21) *существует и конечен. Если X удовлетворяет условиям* 1 (*или* 1а), 2 (*или* 2а) *и* 3 (*или* 3а), *то X — непрерывная гауссовская случайная функция, определенная соотношениями* (19.19) *и* (19.20).

Прежде чем доказывать теорему, рассмотрим ее в связи с двумя хорошо известными нам примерами. Если

$$\rho(t) = 0, \qquad \sigma^2(t) = 1, \tag{19.29}$$

то (19.20) принимает вид $\mathsf{E}\{X(s)X(t)\} = s$, $s \leqslant t$, так что X — броуновское движение. Если с вероятностью 1

$$\mathsf{E}\{X(t_k+h) - X(t_k) \,\|\, X(t_1), \ldots, X(t_k)\} = 0 \tag{19.30}$$

и

$$\mathsf{E}\{(X(t_k+h) - X(t_k))^2 \,\|\, X(t_1), \ldots, X(t_k)\} = h, \tag{19.31}$$

то условия (19.22) и (19.23), конечно, выполняются и, кроме того, имеют место неравенства (19.26) и (19.27) (последнее даже с произведением $(t - t_1)\,(t_2 - t_1)$ в правой части). Следовательно, если $X \in C$ и $X(0) = 0$ с вероятностью 1, то X будет броуновским движением. Поскольку из предположений теоремы 19.1 вытекают соотношения (19.30) и (19.31), то теорема 19.1 является следствием настоящей теоремы.

Если

$$\rho(t) = -\frac{1}{1-t}, \qquad \sigma^2(t) = 1, \tag{19.32}$$

то (19.20) принимает вид $\mathsf{E}\{X(s)X(t)\} = s(1-t)$, $s \leqslant t$, и распределение X должно совпадать с распределением броуновского моста W°.

Доказательство теоремы. Зафиксируем действительные числа $u_1, \ldots, u_k$ и точки t_k такие, что $0 \leqslant t_1 \leqslant \ldots \leqslant t_k < 1$. Пусть t и u изменяются в полосе:

$$t_k \leqslant t < 1, \quad -\infty < u < \infty. \tag{19.33}$$

Мы выведем дифференциальное уравнение для функции

$$\psi(t, u) = \mathsf{E}\{\exp[i(u_1X(t_1) + \ldots + u_kX(t_k) + uX(t))]\}. \tag{19.34}$$

Для удобства обозначим $Z = u_1X(t_1) + \ldots + u_kX(t_k)$, так что функция (19.34) примет вид

$$\psi(t, u) = \mathsf{E}\{e^{iZ}e^{iuX(t)}\}. \tag{19.35}$$

Поскольку $X \in C$ с вероятностью 1, ψ непрерывна (по паре переменных t и u) в полосе (19.33). Так как $\mathsf{E}\{X(t)\}$ существует, мы имеем

$$\frac{\partial}{\partial u}\psi(t, u) = \mathsf{E}\{iX(t)e^{iZ}e^{iuX(t)}\}, \tag{19.36}$$

и из условия 2а следует, что эта функция также непрерывна в полосе (19.33).

В обозначениях (19.3) и (19.5) имеем

$$\psi(t+h, u) - \psi(t, u) = \mathsf{E}\{e^{iZ}e^{iuX(t)}[e^{iu\Delta(t,\,t+h)} - 1]\} =$$
$$= \mathsf{E}\left\{e^{iZ}e^{iuX(t)}\left[iu\Delta_{t,\,t+h} - \frac{1}{2}u^2\Delta^2_{t,\,t+h} + c(u\Delta_{t,\,t+h})\right]\right\}.$$

В силу (19.35) и (19.36)

$$\Bigg|\frac{1}{h}[\psi(t+h, u) - \psi(t, u)] -$$
$$- u\rho(t)\frac{\partial}{\partial u}\psi(t, u) + \frac{1}{2}u^2\sigma^2(t)\psi(t, u)\Bigg| \leqslant$$
$$\leqslant \frac{|u|}{h}\left|\mathsf{E}\{e^{iZ}e^{iuX(t)}[\Delta_{t,\,t+h} - h\rho(t)X(t)]\}\right| +$$
$$+ \frac{u^2}{2h}\left|\mathsf{E}\{e^{iZ}e^{iuX(t)}[\Delta^2_{t,\,t+h} - h\sigma^2(t)]\}\right| +$$
$$+ \frac{1}{h}\mathsf{E}\{|c(u\Delta_{t,\,t+h})|\}. \tag{19.37}$$

Поскольку $|e^{iv}-1-iv|\leqslant\frac{1}{2}v^2$, мы, в дополнение к (19.4), получим неравенство $|c(v)|\leqslant v^2$. Поэтому

$$\frac{1}{h}\mathsf{E}\{|c(u\Delta_{t,\,t+h})|\}\leqslant|u|^3\alpha^{3/2}h^{1/2}+\frac{u^2}{h}\int\limits_{\{\Delta^2_{t,\,t+h}>\alpha h\}}\Delta^2_{t,\,t+h}\,d\mathsf{P}$$

и из условия 3а теперь следует, что третий член правой части (19.37) стремится к 0 вместе с h. Но по условию Iа (с $(t_1, \ldots, t_k, t)$) вместо $(t_1, \ldots, t_k)$ и $(u_1, \ldots, u_k, u)$ вместо $(u_1, \ldots, u_k)$) остальные два члена тоже стремятся к 0. Следовательно,

$$\frac{\partial}{\partial t}\psi(t, u)=u\rho(t)\frac{\partial}{\partial u}\psi(t, u)-\frac{1}{2}u^2\sigma^2(t)\psi(t, u). \qquad (19.38)$$

Поскольку правая часть (19.38) непрерывна в полосе (19.33), то это справедливо и для левой части *).

Решим теперь уравнение (19.38). Для произвольного v определим функцию

$$\lambda_v(s)=v\exp\left[-\int\limits_{t_k}^{s}\rho(\tau)\,d\tau\right], \qquad t_k\leqslant s<1, \qquad (19.39)$$

и положим

$$g_v(s)=\psi(s, \lambda_v(s)), \qquad t_k\leqslant s<1. \qquad (19.40)$$

По правилу дифференцирования сложной функции (если обозначить через ψ_1 и ψ_2 частные производные по первому и второму аргументу функции ψ)

$$g'_v(s)=\psi_1(s, \lambda_v(s))+\psi_2(s, \lambda_v(s))\lambda'_v(s),$$

откуда в силу (19.38), (19.39) и (19.40) получаем

$$g'_v(s)=-\frac{1}{2}\lambda^2_v(s)\,\sigma^2(s)\,g_v(s).$$

Следовательно **),

$$\psi(s, \lambda_v(s))=\psi(t_k, v)\exp\left[-\frac{1}{2}\int\limits_{t_k}^{s}\sigma^2(r)\lambda^2_v(r)\,dr\right], \qquad (19.41)$$

$$t_k\leqslant s<1.$$

*) Здесь производная по t понимается как двусторонняя, хотя h и стремится к 0 справа; см. первую сноску на стр. 216.

**) См. сноску на стр. 215.

Для произвольной точки (t, u) из полосы (19.33) положим $v = u\exp\left[\int\limits_{t_k}^{t} \rho(\tau)\,d\tau\right]$. Тогда $\lambda_v(t) = u$ и

$$\lambda_v^2(r) = u^2 \exp\left[2\int\limits_{r}^{t} \rho(\tau)\,d\tau\right], \qquad t_k \leqslant r \leqslant t,$$

так что в силу (19.41) с $s = t$

$$\psi(t, u) = \psi(t_k, ua)\,e^{-\frac{1}{2}u^2b^2}, \tag{19.42}$$

где

$$a = \exp\int\limits_{t_k}^{t} \rho(\tau)\,d\tau$$

и

$$b^2 = \int\limits_{t_k}^{t} \sigma^2(r)\exp\left[2\int\limits_{r}^{t} \rho(\tau)\,d\tau\right]dr.$$

Возвращаясь к определению функции ψ, из (19.42) получаем

$$\mathsf{E}\{\exp[i(u_1X(t_1) + \ldots + u_kX(t_k) + uX(t))]\} =$$
$$= \mathsf{E}\{\exp[i(u_1X(t_1) + \ldots + u_kX(t_k) + uaX(t_k))]\}\,e^{-\frac{1}{2}u^2b^2}.$$

Мы установили это для $0 \leqslant t_1 \leqslant \ldots \leqslant t_k \leqslant t < 1$; по непрерывности доказанное остается верным и при $t = 1$. Отсюда следует, что для произвольного v

$$\mathsf{E}\{\exp[i(u_1X(t_1) + \ldots + u_kX(t_k) + v(X(t) - aX(t_k)))]\} =$$
$$= \mathsf{E}\{\exp[i(u_1X(t_1) + \ldots + u_kX(t_k))]\}\,e^{-\frac{1}{2}v^2b^2}.$$

Таким образом, случайный вектор $(X(t_1), \ldots, X(t_k))$ и случайная величина $X(t) - aX(t_k)$ независимы, причем последняя имеет распределение $N(0, b^2)$ *).

Полагая $k = 1$, $t_1 = 0$ и используя тот факт, что $\mathsf{P}\{X(0) = 0\} = 1$, мы заключаем, что случайная величина $X(t)$ распределена нормально со средним 0. Пола-

*) Отсюда следует, что $\{X(t)\}$ — марковский процесс.

гая $k=1$ и беря t_1 произвольным, мы видим, что (19.20) справедливо и что величины $X(t_1)$ и $X(t)-aX(t_1)$ имеют совместное нормальное распределение (при подходящем выборе a), а потому совместно нормальны $X(t_1)$ и $X(t)$. Полагая далее $k=2$, мы убеждаемся, что величины $X(t_1)$, $X(t_2)$ и $X(t)-aX(t_2)$ имеют совместное нормальное распределение, и следовательно, то же справедливо и для $X(t_1)$, $X(t_2)$ и $X(t)$. Продолжая это рассуждение, мы видим, что все конечномерные распределения, отвечающие $\{X(t)\}$, нормальны, и теорема доказана.

Для получения асимптотического варианта теоремы 19.3 нам понадобятся три новых условия, соответствующих условиям 1, 2 и 3. Как и прежде, каждое условие будет дано в двух формах. Пусть X_n — случайные элементы в D.

Условие 1° *Если* $0\leqslant t_1\leqslant \ldots \leqslant t_k<1$, *то*

$$\lim_{h\downarrow 0}\limsup_{n\to\infty}\frac{1}{h}\mathsf{E}\{|\mathsf{E}\{X_n(t_k+h)-X_n(t_k)\,\|\,X_n(t_1),\ \ldots,\ X_n(t_k)\}- -h\rho(t_k)X_n(t_k)|\}=0 \tag{19.43}$$

и

$$\lim_{h\downarrow 0}\limsup_{n\to\infty}\frac{1}{h}\mathsf{E}\{|\mathsf{E}\{(X_n(t_k+h)- -X_n(t_k))^2\,\|\,X_n(t_1),\ \ldots,\ X_n(t_k)|\}-h\sigma^2(t_k)|\}=0. \tag{19.44}$$

Условие 1°а. *Если* $0\leqslant t_1\leqslant \ldots \leqslant t_k<1$, *то для всех действительных* $u_1,\ \ldots,\ u_k$

$$\lim_{h\downarrow 0}\limsup_{n\to\infty}\frac{1}{h}\left|\mathsf{E}\left\{\exp\left[\sum_{j=1}^{k}iu_jX_n(t_j)\right]\times \times[X_n(t_k+h)-X_n(t_k)-h\rho(t_k)X_n(t_k)]\right\}\right|=0 \tag{19.45}$$

и

$$\lim_{h\downarrow 0}\limsup_{n\to\infty}\frac{1}{h}\left|\mathsf{E}\left\{\exp\left[\sum_{j=1}^{k}iu_jX_n(t_j)\right]\times \times[(X_n(t_k+h)-X_n(t_k))^2-\sigma^2(t_k)]\right\}\right|=0. \tag{19.46}$$

Условие 2°. *Справедливо неравенство*

$$\sup_t\limsup_{n\to\infty}\mathsf{E}\{X_n^2(t)\}<\infty. \tag{19.47}$$

Условие 2°а. *Справедливо равенство*

$$\lim_{a\to\infty}\sup_t \limsup_{n\to\infty} \int\limits_{\{|X_n(t)|\geqslant a\}} |X_n(t)|\,d\mathsf{P}=0. \qquad (19.48)$$

Условие 3°. *Существует постоянная* K, *такая, что* $\mathsf{E}\{(X_n(t)-X_n(t_1))^2(X_n(t_2)-X_n(t))^2\}\leqslant$

$$\leqslant K(t_2-t_1)^2, \quad t_1\leqslant t\leqslant t_2, \qquad (19.49)$$

для всех n.

Условие 3°а. *При* $t<1$

$$\lim_{a\to\infty}\limsup_{h\downarrow 0}\limsup_{n\to\infty}\frac{1}{h}\int\limits_A (X_n(t+h)-X_n(t))^2\,d\mathsf{P}=0, \qquad (19.50)$$

где $A=\{(X_n(t+h)-X_n(t))^2\geqslant ah\}$.

Легко показать, что из условий 1° и 2° следуют, соответственно, условия 1°а и 2°а. Несмотря на то, что из условия 3° не следует 3°а, каждое из этих условий достаточно для справедливости утверждения теоремы.

Теорема 19.4. *Пусть последовательность* $\{X_n^2(t); n\geqslant 1\}$ *равномерно интегрируема при каждом* t, *пусть* $X_n(0)\xrightarrow{P}0$ *и для любых положительных* ε *и* η *существует такое число* δ, *что для всех достаточно больших* n

$$\mathsf{P}\{w(X_n,\delta)\geqslant\varepsilon\}\leqslant\eta. \qquad (19.51)$$

Пусть, далее, $\sigma^2(t)$ *и* $\rho(t)$ *непрерывны на* $[0,1)$, *а предел* (19.21) *существует и конечен. Если* $\{X_n\}$ *удовлетворяет условиям* 1° (*или* 1°а), 2° (*или* 2°а) *и* 3° (*или* 3°а), *то* $X_n\xrightarrow{\mathcal{D}}X$, *где* X — *непрерывная гауссовская случайная функция, для которой выполняются* (19.19) *и* (19.20).

Доказательство. Согласно теореме 15.5 любая подпоследовательность последовательности $\{X_n\}$ в свою очередь содержит последовательность, сходящуюся по распределению к некоторому случайному элементу пространства D, для которого $\mathsf{P}\{X(0)=0\}=1$ и $\mathsf{P}\{X\in C\}=1$. Достаточно доказать, что каждый такой предел X должен удовлетворять остальным предположениям теоремы 19.3. Из теоремы 5.4 и предполагаемой равномерной интегрируемости последовательности $\{X_n^2(t); n\geqslant 1\}$ следует, что условие 1°а влечет условие 1а;

кроме того, по теореме 5.3 2°а влечет 2а, 3°а влечет 3а и 3° влечет 3 (как мы уже видели, 3 влечет 3а). Таким образом, X удовлетворяет условиям 1а, 2а, 3а, что и завершает доказательство.

З а м е ч а н и е. Предположим, что выполнено условие 3°. Пусть с вероятностью 1

$$X_n(0)=0, \tag{19.52}$$

и пусть максимальный скачок X_n с вероятностью 1 удовлетворяет условию

$$\max_t |X_n(t)-X_n(t-)| \leqslant \varepsilon_n, \tag{19.53}$$

где

$$\varepsilon_n \to 0. \tag{19.54}$$

Тогда, используя теорему 12.1 и неравенство (12.6), мы имеем

$$\mathsf{P}\{\sup_{t\leqslant s\leqslant t+\delta} |X_n(s)-X_n(t)| \geqslant \lambda\} \leqslant K' \frac{\delta^2}{\lambda^4} \tag{19.55}$$

при $K'=4^4KK_{2,1}$, где $\varepsilon_n<\frac{1}{4}\lambda$, откуда в силу следствия из теоремы 8.3 получаем (19.51). Из (19.55) теперь следует, что

$$\mathsf{P}\{|X_n(t)|\geqslant\lambda\} \leqslant \frac{K'}{\lambda^4},$$

если $\varepsilon_n<\frac{1}{4}\lambda$; таким образом (см. (3) на стр. 305),

$$\int\limits_{\{X_n^2(t)\geqslant\alpha\}} X_n^2(t)\,d\mathsf{P} \leqslant 2\frac{K'}{\alpha} \tag{19.56}$$

при $\varepsilon_n^2<\frac{1}{16}\alpha$, откуда вытекает, что последовательность $\{X_n^2(t);\ n\geqslant 1\}$ равномерно интегрируема. Наконец, (19.56) влечет (19.48). Следовательно, если соотношения (19.52), (19.53) и (19.54) выполнены, то утверждение теоремы 19.4 будет справедливым при выполнении лишь условий 1° (или 1°а) и 3°.

П р и м е ч а н и е. Приведенные здесь результаты принадлежат Розену (1967а). Условия изменены так, чтобы использовать теорему 12.1. Идея вывода центральной предельной теоремы из дифференциального уравнения исходит от Хинчина (1936); теорема 19.3 при выполнении условий (19.30) и (19.31) принадлежит Леви (1972) и Дубу (1956).

§ 20. Процессы с перемешиванием

φ-перемешивание. Пусть

$$\ldots, \xi_{-1}, \xi_0, \xi_1, \ldots \tag{20.1}$$

— стационарная в узком смысле последовательность случайных величин, определенных на вероятностном пространстве $(\Omega, \mathscr{B}, \mathsf{P})$. При $a \leqslant b$ обозначим $\mathscr{M}_a^b$ σ-поле, порожденное случайными величинами $\xi_a, \ldots, \xi_b$. Пусть $\mathscr{M}_{-\infty}^a$ — σ-поле, порожденное последовательностью $\ldots, \xi_{a-1}, \xi_a$, а $\mathscr{M}_a^\infty$ — σ-поле, порожденное последовательностью $\xi_a, \xi_{a+1}, \ldots$

Рассмотрим неотрицательную функцию φ от натурального аргумента. Будем говорить, что последовательность $\{\xi_n\}$ является последовательностью с φ-*перемешиванием*, если для $E_1 \in \mathscr{M}_{-\infty}^k$ и $E_2 \in \mathscr{M}_{k+n}^\infty$ при любых k, $-\infty < k < \infty$ и n, $n \geqslant 1$,

$$|\mathsf{P}(E_1 \cap E_2) - \mathsf{P}(E_1)\mathsf{P}(E_2)| \leqslant \varphi(n)\mathsf{P}(E_1). \tag{20.2}$$

Это условие является совместным условием на последовательность $\{\xi_n\}$ и на функцию φ. Мы рассматриваем лишь те функции φ, для которых

$$\lim_{n\to\infty} \varphi(n) = 0, \tag{20.3}$$

и обычно будем еще накладывать дополнительные ограничения на скорость сходимости $\varphi(n)$ к нулю. Если мы говорим, что $\{\xi_n\}$ — последовательность с φ-перемешиванием, и не делаем никаких ограничений на φ, то мы имеем в виду, что (20.2) выполняется для некоторой функции φ, удовлетворяющей условию (20.3).

Если $\mathsf{P}(E_1) > 0$, то условие (20.2) равносильно условию

$$|\mathsf{P}(E_2 | E_1) - \mathsf{P}(E_2)| \leqslant \varphi(n), \tag{20.4}$$

а при $\mathsf{P}(E_1) = 0$ условие (20.2), очевидно, выполнено. Будем считать, что левая часть (20.4) равна нулю при $\mathsf{P}(E_1) = 0$, так что условие φ-перемешивания принимает вид

$$\sup\{|\mathsf{P}(E_2 | E_1) - \mathsf{P}(E_2)| : E_1 \in \mathscr{M}_{-\infty}^k,\ E_2 \in \mathscr{M}_{k+n}^\infty\} \leqslant \varphi(n). \tag{20.5}$$

Если $\varphi(n)$ мало, то (20.2) и (20.4) означают, что E_2 почти не зависит от E_1. В процессе с φ-перемешиванием отдаленное будущее почти не зависит от прошлого и настоящего *). Это свойство позволяет нам доказывать различные центральные предельные теоремы и функциональные центральные предельные теоремы.

Часто вместо $\varphi(n)$ мы будем использовать обозначение φ_n. Если при каждом n φ_n принимает минимальное значение, удовлетворяющее (20.5), то

$$1 \geqslant \varphi_1 \geqslant \varphi_2 \geqslant \ldots \tag{20.6}$$

Если заменить φ_n на

$$\min\{1, \varphi_1, \ldots, \varphi_n\}, \tag{20.7}$$

то $\{\xi_n\}$ по-прежнему будет последовательностью с φ-перемешиванием и (20.6) будет выполнено. Следовательно, без ограничения общности можно считать, что (20.6) выполнено. Для дальнейшего удобно положить $\varphi_0 = 1$. При этом (20.5) становится тривиальным, когда $n = 0$.

Для фиксированного E_2 класс

$$\{E_1: |\mathsf{P}(E_2|E_1) - \mathsf{P}(E_2)| \leqslant \varepsilon\} \tag{20.8}$$

является монотонным классом **) (т. е. замкнутым относительно операций объединения возрастающих и пересечения убывающих последовательностей множеств); при фиксированном E_1 то же самое верно и для класса

$$\{E_2: |\mathsf{P}(E_2|E_1) - \mathsf{P}(E_2)| \leqslant \varepsilon\}. \tag{20.9}$$

Отсюда следует, что если (20.4) имеет место для E_1 из $\mathscr{A}$ и E_2 из $\mathscr{B}$, где $\mathscr{A}$ и $\mathscr{B}$ — поля, порождающие $\mathscr{M}_{-\infty}^{k}$ и $\mathscr{M}_{k+n}^{\infty}$, соответственно, то (20.5) тоже справедливо; это замечание полезно при проверке того, является ли данный процесс процессом с φ-перемешиванием.

---

*) Так как (20.2) и (20.4) не симметричны относительно E_1 и E_2, прошлое и будущее нельзя поменять местами; в качестве примера процесса с φ-перемешиванием, который перестает быть таковым при обращении времени, можно взять $\xi_n = \sum_{k=1}^{\infty} \eta_{n-k}/2^k$, где η_k независимы и принимают значения ± 1, каждое с вероятностью $1/2$.

**) См. Халмош (1953, стр. 31).

Пример 1. Последовательность (20.1) называется m-зависимой, если случайные векторы $(\xi_i, \ldots, \xi_k)$ и $(\xi_{k+n}, \ldots, \xi_j)$ независимы при $n > m$. (В этой терминологии последовательность независимых случайных величин является 0-зависимой.) m-зависимая последовательность является последовательностью с φ-перемешиванием, где φ такова, что $\varphi_n = 0$ при $n > m$. Примером m-зависимой последовательности является последовательность

$$\xi_n = a_0\zeta_n + a_1\zeta_{n-1} + \ldots + a_m\zeta_{n-m}, \tag{20.10}$$

где a_i — постоянные, а ζ_n независимы и одинаково распределены.

Пример 2. Пусть $\{\zeta_n\}$ — стационарный марковский процесс с конечным пространством состояний, и пусть $\xi_n = f(\zeta_n)$, где f — некоторая действительная функция. определенная на этом пространстве состояний. Если $\mathcal{M}_a^b$ определена, как и прежде, и $\mathcal{N}_a^b$ — σ-поле, порожденное случайными величинами $\zeta_a, \ldots, \zeta_b$, то $\mathcal{M}_a^b \subset \mathcal{N}_a^b$.

Пусть p_u — стационарные вероятности, а p_{uv} — переходные вероятности марковского процесса $\{\zeta_n\}$. Для каждой конечной последовательности состояний u_0, $u_1, \ldots, u_l$ имеем

$$\mathsf{P}\{\zeta_t = u_0, \ldots, \zeta_{t+l} = u_l\} = p_{u_0} p_{u_0 u_1} \ldots p_{u_{l-1} u_l}.$$

Предположим, что все вероятности p_u положительны, так что функция

$$\varphi_n = \max_{u,\, v} \left| \frac{p_{uv}^{(n)}}{p_v} - 1 \right| \tag{20.11}$$

конечна (здесь $p_{uv}^{(n)}$ — — вероятности перехода за n шагов). Пусть H_1 — некоторое множество последовательностей из $i + 1$ состояний, а H_2 — некоторое множество последовательностей из $j + 1$ состояний. Для элемента

$$E_1 = \{(\zeta_{k-i}, \ldots, \zeta_k) \in H_1\} \tag{20.12}$$

из $\mathcal{N}_{-\infty}^k$ и для элемента

$$E_2 = \{(\zeta_{k+n}, \ldots, \zeta_{k+n+j}) \in H_2\} \tag{20.13}$$

из $\mathcal{N}_{k+n}^{\infty}$ имеем

$$|\mathsf{P}(E_1 \cap E_2) - \mathsf{P}(E_1)\mathsf{P}(E_2)| \leqslant$$
$$\leqslant \sum p_{u_0} p_{u_0 u_1} \ldots p_{u_{i-1} u_i} \left| p_{u_i v_0}^{(n)} - p_{v_0} \right| p_{v_0 v_1} \ldots p_{v_{j-1} v_j},$$

где сумма распространяется на все $(u_0, \ldots, u_i)$ из H_1 и все $(v_0, \ldots, v_j)$ из H_2. Отсюда, используя соотношение $\sum\limits_{v} p_{uv} = 1$ и определение (20.11), получаем (20.2). Используя теперь тот факт, что (20.8) и (20.9) — монотонные классы, выводим, что (20.2) имеет место для всех E_1 из $\mathcal{N}_{-\infty}^{k}$ и E_2 из $\mathcal{N}_{k+n}^{\infty}$ и, следовательно, для всех E_1 из $\mathcal{M}_{-\infty}^{k}$ и E_2 из $\mathcal{M}_{k+n}^{\infty}$.

Если переходная матрица (p_{uv}) неприводима и апериодична, то*)

$$p_{uv}^{(n)} \to p_v \tag{20.14}$$

для всех u и v. Так как пространство состояний по предположению конечно, сходимость здесь должна быть равномерной по u и v, так что функция φ_n, определенная соотношением (20.11), сходится к 0. Следовательно, $\{\xi_n\}$ — последовательность с φ-перемешиванием. Более того, в рассматриваемой ситуации известна и скорость сходимости в (20.14) — она показательная. Поэтому существуют положительные постоянные a и ρ, $\rho < 1$, такие, что если

$$\varphi_n = a\rho^n, \tag{20.15}$$

то $\{\xi_n\}$ является последовательностью с φ-перемешиванием. Это также верно для некоторых последовательностей $\xi_n = f(\zeta_n)$, где $\{\zeta_n\}$ — марковский процесс с бесконечным пространством состояний, например, если $\{\zeta_n\}$ удовлетворяет условию Дёблина, имеет один эргодический класс и апериодичен **).

В некоторых приложениях мы имеем лишь одностороннюю (стационарную в узком смысле) последовательность

$$\xi_1, \xi_2, \ldots, \tag{20.16}$$

*) Дуб (1956, стр. 159 и далее).
**) Дуб (1956, стр. 174 и далее).

В таких случаях мы будем говорить, что (20.16) является последовательностью с φ-перемешиванием, если

$$\sup\{|\mathsf{P}(E_2|E_1)-\mathsf{P}(E_2)|:\ E_1\in\mathcal{M}_1^k,\ E_2\in\mathcal{M}_{k+n}^\infty\}\leqslant\varphi_n \quad (20.17)$$

для целых положительных чисел k и n. Если задана односторонняя последовательность (20.16), то можно построить двустороннюю последовательность (20.1) с теми же самыми конечномерными распределениями. Новая последовательность, вообще говоря, будет определена на новом вероятностном пространстве (в качестве этого пространства всегда можно взять произведение бесконечной в обе стороны последовательности действительных прямых). Новую двустороннюю последовательность (20.1) мы будем называть двусторонне бесконечным расширением последовательности (20.16).

Если (20.16) есть последовательность с φ-перемешиванием, то в силу стационарности ее двусторонне бесконечное расширение удовлетворяет условию (20.4) для $E_1\in\mathcal{M}_{k-l}^k$ и $E_2\in\mathcal{M}_{k+n}^\infty$. Так как (20.8) есть монотонный класс, отсюда следует, что двусторонне бесконечное расширение есть последовательность с φ-перемешиванием с той же самой функцией φ. Нетрудно проверить, что справедливо и обратное. Если расширение (20.1) является последовательностью с φ-перемешиванием, то исходная последовательность (20.16) тоже является последовательностью с φ-перемешиванием с той же самой функцией φ.

Если мы докажем, например, что

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i \xrightarrow{\mathcal{D}} N,$$

где случайные величины ξ_i являются элементами последовательности (20.1), то этот результат будет верен также, если под ξ_i понимать элементы последовательности (20.16), так как здесь существенную роль играют лишь конечномерные распределения. Все наши теоремы являются утверждениями такого же типа и для их справедливости несущественно, являются ли ξ_i элементами последовательности (20.16) или (20.1). В доказательствах мы всегда будем считать, что мы имеем дело с последовательностью (20.1). Это более удобно, поскольку при этом прошлое всегда является бесконечной последовательностью, а не конечной последовательностью переменной длины.

Пример 3. Пусть Ω — единичный интервал, $\mathscr{B}$ — класс борелевских подмножеств Ω, и пусть P — мера Лебега на $\mathscr{B}$. Положим

$$\xi_n(\omega) = g(\omega_n, \omega_{n+1}, \ldots, \omega_{n+m}), \quad n = 1, 2, \ldots, \tag{20.18}$$

где ω имеет двоичное разложение $\omega = \frac{\omega_1}{2} + \frac{\omega_2}{2^2} + \ldots$ и g — некоторая функция от последовательности длины $m+1$, состоящей из нулей и единиц. Последовательность $\{\xi_n\}$ стационарна и m-независима. Эта последовательность имеет вид (20.16) и к ней применима развиваемая здесь теория, поскольку, как было указано, ее можно расширить до двусторонне бесконечной последовательности (на некотором новом вероятностном пространстве).

Пример 4. Пусть Ω и $\mathscr{B}$ означают то же, что и в примере 3. Для построения еще одного примера с односторонней последовательностью $\{\xi_n\}$ предположим, что P — гауссова мера

$$\mathsf{P}(A) = \frac{1}{\log 2} \int_A \frac{dx}{1+x}, \qquad A \in \mathscr{B}, \tag{20.19}$$

а $\xi_n(\omega)$ есть n-й коэффициент $a_n(\omega)$ в разложении ω в непрерывную дробь. Последовательность $\{a_1(\omega), a_2(\omega), \ldots\}$ стационарна и удовлетворяет условию φ-перемешивания с функцией φ вида

$$\varphi_n = a\rho^n, \quad a > 0, \quad 0 < \rho < 1\,^{*}). \tag{20.20}$$

Неравенства для моментов. Пусть случайная величина ξ $\mathscr{M}_{-\infty}^{k}$-измерима, а случайная величина η $\mathscr{M}_{k+n}^{\infty}$-измерима. Если ξ и η являются индикаторами множеств, то согласно (20.2) величина

$$|\mathsf{E}\{\xi\eta\} - \mathsf{E}\{\xi\}\mathsf{E}\{\eta\}| \tag{20.21}$$

не превосходит $\varphi_n \mathsf{E}\{\xi\}$. Нам потребуется оценка величины (20.21) для случайных величин ξ и η более общих,

---

*) Леви (1937, глава 9).

чем индикаторы. Ниже всюду предполагается, если не оговорено противное, что $\{\xi_n\}$ есть стационарная последовательность с φ-перемешиванием.

Лемма 1. *Если случайная величина* ξ $\mathcal{M}_{-\infty}^{k}$-*измерима, а случайная величина* η $\mathcal{M}_{k+n}^{\infty}$-*измерима* ($n \geqslant 0$), *то из*

$$\mathsf{E}\{|\xi|^r\} < \infty, \quad \mathsf{E}\{|\eta|^s\} < \infty, \quad r, s > 1, \quad \frac{1}{r} + \frac{1}{s} = 1, \tag{20.22}$$

следует, что

$$|\mathsf{E}\{\xi\eta\} - \mathsf{E}\{\xi\}\mathsf{E}\{\eta\}| \leqslant 2\varphi_n^{1/r}\mathsf{E}^{1/r}\{|\xi|^r\}\mathsf{E}^{1/s}\{|\eta|^s\}. \tag{20.23}$$

Доказательство. Чтобы лучше понять эту лемму, рассмотрим сначала два экстремальных случая. Если $\varphi_n = 0$, так что ξ и η независимы, то неравенство (20.23) справедливо, так как обе его части равны нулю. Если $\varphi_n = 1$, т. е. при отсутствии ограничений, в силу неравенств Гёльдера и Минковского имеем

$$|\mathsf{E}\{\xi\eta\} - \mathsf{E}\{\xi\}\mathsf{E}\{\eta\}| = |\mathsf{E}\{\xi(\eta - \mathsf{E}\{\eta\})\}| \leqslant$$
$$\leqslant \mathsf{E}^{1/r}\{|\xi|^r\}\mathsf{E}^{1/s}\{|\eta - \mathsf{E}\{\eta\}|^s\} \leqslant 2\mathsf{E}^{1/r}\{|\xi|^r\}\mathsf{E}^{1/s}\{|\eta|^s\}.$$

Так как мы можем аппроксимировать ξ и η простыми случайными величинами, при рассмотрении общего случая можно предполагать, что

$$\xi = \sum_i u_i I_{A_i} \tag{20.24}$$

и

$$\eta = \sum_j v_j I_{B_j}, \tag{20.25}$$

где $\{A_i\}$ $[\{B_j\}]$ — конечное разбиение выборочного пространства Ω на элементы σ-поля $\mathcal{M}_{-\infty}^{k}$ $[\mathcal{M}_{k+n}^{\infty}]$. Для таких случайных величин по неравенству Гёльдера имеем

$$|\mathsf{E}\{\xi\eta\} - \mathsf{E}\{\xi\}\mathsf{E}\{\eta\}| =$$
$$= \left|\sum_i u_i \mathsf{P}(A_i)^{1/r}\left[\mathsf{P}(A_i)^{1/s}\sum_j v_j(\mathsf{P}(B_j|A_i) - \mathsf{P}(B_j))\right]\right| \leqslant$$
$$\leqslant \mathsf{E}^{1/r}\{|\xi|^r\}\left\{\sum_i \mathsf{P}(A_i)\left|\sum_j v_j(\mathsf{P}(B_j|A_i) - \mathsf{P}(B_j))\right|^s\right\}^{1/s}.$$

и следовательно, достаточно доказать, что

$$\sum_i \mathsf{P}(A_i)\left|\sum_j v_j(\mathsf{P}(B_j\,|\,A_i)-\mathsf{P}(B_j))\right|^s \leqslant 2^s\varphi_n^{s/r}\mathsf{E}\{|\eta|^s\}. \tag{20.26}$$

Для каждого i из неравенства Гёльдера следует, что

$$\left|\sum_j v_j(\mathsf{P}(B_j\,|\,A_i)-\mathsf{P}(B_j))\right| \leqslant$$
$$\leqslant \left\{\sum_j |v_j|^s|\mathsf{P}(B_j\,|\,A_i)-\mathsf{P}(B_j)|\right\}^{1/s}\left\{\sum_j|\mathsf{P}(B_j\,|\,A_i)-\mathsf{P}(B_j)|\right\}^{1/r}.$$

Так как

$$\sum_i \mathsf{P}(A_i)\sum_j |v_j|^s|\mathsf{P}(B_j\,|\,A_i)-\mathsf{P}(B_j)| \leqslant 2\mathsf{E}\{|\eta|^s\},$$

то неравенство (20.26) будет установлено, если мы покажем, что для всех i

$$\sum_j |\mathsf{P}(B_j\,|\,A_i)-\mathsf{P}(B_j)| \leqslant 2\varphi_n. \tag{20.27}$$

Если $C_i^+\,[C_i^-]$ есть объединение тех B_j, для которых разность $\mathsf{P}(B_j\,|\,A_i)-\mathsf{P}(B_j)$ положительна [неположительна], то C_i^+ и C_i^- принадлежат $\mathscr{M}_{k+n}^{\infty}$ и, следовательно,

$$\sum_j |\mathsf{P}(B_j\,|\,A_i)-\mathsf{P}(B_j)| =$$
$$=[\mathsf{P}(C_i^+\,|\,A_i)-\mathsf{P}(C_i^+)]+[\mathsf{P}(C_i^-)-\mathsf{P}(C_i^-\,|\,A_i)] \leqslant 2\varphi_n.$$

Таким образом, (20.27) справедливо, и доказательство закончено.

Полагая в лемме 1 $r=1$ и $s=\infty$, получаем формально следующий результат: если $\mathsf{E}\{|\xi|\}<\infty$ и $\mathsf{P}\{|\eta|>C\}=0$ (ξ $\mathscr{M}_{-\infty}^{k}$-измерима, а η $\mathscr{M}_{k+n}^{\infty}$-измерима), то

$$|\mathsf{E}\{\xi\eta\}-\mathsf{E}\{\xi\}\mathsf{E}\{\eta\}| \leqslant 2\varphi_n\mathsf{E}\{|\xi|\}C. \tag{20.28}$$

Это неравенство легко установить и непосредственно: достаточно рассмотреть простые случайные величины, заданные соотношениями (20.24) и (20.25), где $|v_j|\leqslant C$.

Имеем

$$|\mathsf{E}\{\xi\eta\} - \mathsf{E}\{\xi\}\mathsf{E}\{\eta\}| \leqslant \leqslant C\sum_i |u_i| \mathsf{P}(A_i) \sum_j |\mathsf{P}(B_j \mid A_i) - \mathsf{P}(B_j)|,$$

и поэтому (20.28) следует из (20.27).

Нам понадобится лишь тот частный случай (20.27), когда обе случайные величины ξ и η ограничены.

Лемма 2. *Если случайная величина* ξ $\mathscr{M}_{-\infty}^{k}$-*измерима и* $|\xi| \leqslant C_1$ *и если случайная величина* η $\mathscr{M}_{k+n}^{\infty}$-*измерима* ($n \geqslant 0$) *и* $|\eta| \leqslant C_2$, *то*

$$|\mathsf{E}\{\xi\eta\} - \mathsf{E}\{\xi\}\mathsf{E}\{\eta\}| \leqslant 2C_1C_2\varphi_n. \tag{20.29}$$

Пусть

$$S_n = \xi_1 + \ldots + \xi_n \tag{20.30}$$

и $S_0 = 0$. Временно откажемся от предположения о том, что $\{\xi_n\}$ является последовательностью с φ-перемешиванием (оставляя, однако, предположение о стационарности).

Лемма 3. *Пусть* $\mathsf{E}\{\xi_0\} = 0$ *и*

$$\sum_{k=0}^{\infty} |\mathsf{E}\{\xi_0\xi_k\}| < \infty. \tag{20.31}$$

Тогда при всех n

$$\frac{1}{n}\mathsf{E}\{S_n^2\} \leqslant 2\sum_{k=0}^{\infty} |\mathsf{E}\{\xi_0\xi_k\}| \tag{20.32}$$

и

$$\frac{1}{n}\mathsf{E}\{S_n^2\} \to \sigma^2, \tag{20.33}$$

где

$$\sigma^2 = \mathsf{E}\{\xi_0^2\} + 2\sum_{k=1}^{\infty} \mathsf{E}\{\xi_0\xi_k\}. \tag{20.34}$$

Доказательство. Обозначим $\rho_k = \mathsf{E}\{\xi_0\xi_k\}$. В силу стационарности

$$\mathsf{E}\{S_n^2\} = n\rho_0 + 2\sum_{k=1}^{n-1}(n-k)\rho_k;$$

отсюда сразу следует (20.32), а (20.33) вытекает из неравенства

$$\left|\sigma^2-\frac{1}{n}\mathsf{E}\{S_n^2\}\right|\leqslant 2\sum_{k=n}^{\infty}|\rho_k|+\frac{2}{n}\sum_{i=1}^{n-1}\sum_{k=i}^{\infty}|\rho_k|.$$

Предположим снова, что $\{\xi_n\}$ удовлетворяет условию φ-перемешивания. Если ξ_0 имеет нулевое среднее значение и конечную дисперсию, то по лемме 1 с $r=s=2$

$$|\mathsf{E}\{\xi_0\xi_k\}|\leqslant 2\varphi_k^{1/2}\mathsf{E}\{\xi_0^2\}. \tag{20.35}$$

Если $\sum_n \varphi_n^{1/2}<\infty$, то имеет место неравенство (20.31), так что дисперсия суммы S_n асимптотически равна $n\sigma^2$, где σ^2 определяется формулой (20.34). Величина σ^2 может быть равной нулю, даже если ξ_0 имеет положительную дисперсию. Это так, например, если $\xi_n=\zeta_n-\zeta_{n-1}$, где ζ_n независимы и одинаково распределены (см. (20.10)). Случай $\sigma=0$ является вырожденным и будет исключаться в большинстве нижеследующих теорем.

Лемма 4. *Если случайная величина ξ_0 ограничена постоянной C, $\mathsf{E}\{\xi_0\}=0$ и $\sum_n \varphi_k^{1/2}<\infty$, то*

$$\mathsf{E}\{S_n^4\}\leqslant K_\varphi C^4 n^2,$$

где K_φ зависит лишь от φ.

Доказательство. Мы покажем, что

$$\mathsf{E}\{S_n^4\}\leqslant 768C^4\left[\sum_{i=1}^{\infty}\varphi_i^{1/2}\right]^2 n^2. \tag{20.36}$$

Если заменить φ_i на величину (20.7), то сумма ряда в (20.36) не возрастет. Следовательно, можно считать, что

$$1=\varphi_0\geqslant\varphi_1\geqslant\varphi_2\geqslant\ldots \tag{20.37}$$

В силу стационарности

$$\mathsf{E}\{S_n^4\}\leqslant 4!n\sum|\mathsf{E}\{\xi_0\xi_i\xi_{i+j}\xi_{i+j+k}\}|, \tag{20.38}$$

где сумма берется по тем i, j, k, для которых

$$i,\ j,\ k\geqslant 0,\qquad i+j+k\leqslant n. \tag{20.39}$$

По лемме 2

$$|\mathsf{E}\{\xi_0(\xi_i\xi_{i+j}\xi_{i+j+k})\}| \leqslant 2C^4\varphi_i \tag{20.40}$$

и

$$|\mathsf{E}\{(\xi_0\xi_i\xi_{i+j})\,\xi_{i+j+k}\}| \leqslant 2C^4\varphi_k. \tag{20.41}$$

Применяя еще раз лемму 2 (и пользуясь стационарностью), имеем

$$|\mathsf{E}\{(\xi_0\xi_i)(\xi_{i+j}\xi_{i+j+k})\}| \leqslant |\mathsf{E}\{\xi_0\xi_i\}\,\mathsf{E}\{\xi_0\xi_k\}| + 2C^4\varphi_j.$$

Еще дважды применяя лемму 2, получаем

$$|\mathsf{E}\{\xi_0\xi_i\}| \leqslant 2C^2\varphi_i, \qquad |\mathsf{E}\{\xi_0\xi_k\}| \leqslant 2C^2\varphi_k;$$

подставляя эти два неравенства в предыдущее неравенство, находим

$$|\mathsf{E}\{(\xi_0\xi_i)(\xi_{i+j}\xi_{i+j+k})\}| \leqslant 4C^4\varphi_i\varphi_k + 2C^4\varphi_j. \tag{20.42}$$

Согласно (20.40), (20.41) и (20.42) слагаемое в (20.38) не превосходит произведения $4C^4$ на минимальную из величин

$$\varphi_i, \ \varphi_k, \ \varphi_i\varphi_k + \varphi_j,$$

и поэтому вся сумма не превосходит произведения $4C^4$ на

$$\sum_{j,\ k\leqslant i} \varphi_i + \sum_{i,\ j\leqslant k} \varphi_k + \sum_{i,\ k\leqslant j} (\varphi_i\varphi_k + \varphi_j) =$$

$$= \sum_{i,\ k\leqslant j} \varphi_i\varphi_k + 3\sum_{j,\ k\leqslant i} \varphi_i, \tag{20.43}$$

где индексы суммирования, кроме указанных ограничений, подчиняются и (20.39).

Поскольку $\varphi_i \leqslant 1$, то

$$\sum_{i,\ k\leqslant j} \varphi_i\varphi_k \leqslant \sum_{j=0}^{n} \sum_{i,\ k=0}^{\infty} \varphi_i\varphi_k \leqslant 2n\left[\sum_{i=0}^{\infty} \varphi_i^{1/2}\right]^2.$$

Кроме того, согласно (20.37)

$$\sum_{j,\ k\leqslant i} \varphi_i \leqslant \sum_{i=0}^{n} \sum_{j,\ k=0}^{i} \varphi_i \leqslant$$

$$\leqslant 2n\sum_{i=0}^{n} (i+1)\,\varphi_i = 2n\sum_{u=0}^{n} \sum_{v=u}^{n} \varphi_v \leqslant$$

$$\leqslant 2n\sum_{u=0}^{\infty} \varphi_u^{1/2} \sum_{v=u}^{\infty} \varphi_v^{1/2} \leqslant 2n\left[\sum_{i=0}^{\infty} \varphi_i^{1/2}\right]^2.$$

Из этих двух неравенств вытекает, что (20.43) не больше, чем

$$8n\left[\sum_{i=0}^{\infty}\varphi_i^{1/2}\right]^2, \tag{20.44}$$

и поэтому сумма в (20.38) не превосходит произведения $4C^4$ на эту величину. Умножая (20.44) на $4!n4C^4$, получаем (20.36).

Функциональная центральная предельная теорема. Определим случайный элемент X_n из D, положив

$$X_n(t,\omega)=\frac{1}{\sigma\sqrt{n}}S_{[nt]}(\omega),\qquad 0\leqslant t\leqslant 1. \tag{20.45}$$

Теорема 20.1. *Пусть $\{\xi_n\}$ является последовательностью с φ-перемешиванием, для которой $\sum_n \varphi_n^{1/2}<\infty$ и ξ_0 имеет нулевое среднее значение и конечную дисперсию. Тогда ряд*

$$\sigma^2=\mathsf{E}\{\xi_0^2\}+2\sum_{k=1}^{\infty}\mathsf{E}\{\xi_0\xi_k\} \tag{20.46}$$

абсолютно сходится; если $\sigma^2>0$ и случайный элемент X_n определен формулой (20.45), *то*

$$X_n\xrightarrow{\mathscr{D}}W. \tag{20.47}$$

Доказательство. Абсолютная сходимость ряда (20.46) следует из (20.35). Предположим, что $\sigma^2>0$. Мы докажем (20.47), применяя теорему 19.2, избегая тем самым использование центральной предельной теоремы.

Мы покажем, что последовательность $\{S_n^2/n\}$ равномерно интегрируема. Если воспользоваться сокращенной записью $\mathsf{E}_\alpha\{U\}$ для обозначения интеграла от U по множеству $\{U\geqslant\alpha\}$, то условие равномерной интегрируемости примет вид

$$\lim_{\alpha\to\infty}\sup_n\mathsf{E}_\alpha\left\{\frac{1}{n}S_n^2\right\}=0. \tag{20.48}$$

Докажем сначала, что утверждение теоремы справедливо, если выполнено условие (20.48). Прежде всего нам нужно показать, что X_n имеет асимптотически независимые приращения в том смысле, что (19.13) влечет сходимость к нулю разности (19.14) при $n \to \infty$.

Предположим, что u_i и v_j — целые числа такие, что $u_1 \leqslant v_1 < u_2 \leqslant v_2 < \ldots < u_r \leqslant v_r$ и $u_i - v_{i-1} \geqslant b$, $i = 2, \ldots, r$, и пусть E_i принадлежит $\mathcal{M}_{u_i}^{v_i}$, $i = 1, \ldots, r$. Из определения φ-перемешивания, применяя индукцию по r, получаем

$$\left| \mathsf{P}\left(\bigcap_{i=1}^{r} E_i\right) - \prod_{i=1}^{r} \mathsf{P}(E_i) \right| \leqslant r\varphi(b). \tag{20.49}$$

Положим, что имеет место (19.13), и пусть $E_i = \{X_n(t_i) - X_n(s_i) \in H_i\}$, где $H_i \in \mathcal{R}^1$. Тогда E_i принадлежит $\mathcal{M}_{[ns_i]+1}^{[nt_i]}$, и если δ есть наименьшая разность среди $s_i - t_{i-1}$, то $[ns_i] + 1 - [nt_{i-1}] > [n\delta]$, так что согласно (20.49) модуль разности (19.14) не превосходит $r\varphi([n\delta])$. Поскольку δ положительно, X_n действительно имеет асимптотически независимые приращения.

Так как $X_n^2(t) \leqslant S_{[nt]}^2/\sigma^2[nt]$, то из (20.48) следует равномерная интегрируемость последовательности $\{X_n^2(t)\}$ для каждого t. Очевидно, $\mathsf{E}\{X_n(t)\} = 0$ и из леммы 3 вытекает, что $\mathsf{E}\{X_n^2(t)\} \to t$. Согласно теореме 8.4 (в применении к D) условие плотности (19.15) будет выполнено, если мы покажем, что для любого положительного ε существует λ, $\lambda > \sigma$, и целое число n_0 такое, что при $n \geqslant n_0$

$$P\left\{\max_{i \leqslant n} |S_i| \geqslant \lambda\sqrt{n}\right\} \leqslant \frac{\varepsilon}{\lambda^2}. \tag{20.50}$$

Положим $S_n^* = \sum_{j=1}^{n} |\xi_j|$. Так как ξ_0 имеет конечный второй момент, то существует возрастающая последовательность целых чисел m_i такая, что $\mathsf{P}\{|\xi_0| \geqslant \sqrt{n}/i^2\} \leqslant 1/ni^2$ при $n \geqslant m_i$. Если положить $p_n = i$ для $m_i \leqslant n < m_{i+1}$ (и $p_n = 1$ для $n < m_1$), то p_n стремится к бесконечности, но так медленно, что при всех положительных λ

$$\lim_{n\to\infty} n\mathsf{P}\left\{S_{p_n}^* \geqslant \lambda\sqrt{n}\right\} = 0. \tag{20.51}$$

В то же время мы можем выбрать p_n таким образом, чтобы $p_n \leqslant n$.

Для заданного ε выберем λ так, чтобы $\lambda > \sigma$ и

$$\mathsf{P}\{|S_i| > \lambda\sqrt{i}\} < \frac{\varepsilon}{\lambda^2} \tag{20.52}$$

для всех i, что возможно в силу (20.48). Положив

$$E_i = \{\max_{j<i} |S_j| < 3\lambda\sqrt{n} \leqslant |S_i|\},$$

имеем

$$\mathsf{P}\{\max_{i\leqslant n} |S_i| \geqslant 3\lambda\sqrt{n}\} \leqslant$$
$$\leqslant \mathsf{P}\{|S_n| \geqslant \lambda\sqrt{n}\} + \sum_{i=1}^{n-1} \mathsf{P}(E_i \cap \{|S_n - S_i| \geqslant 2\lambda\sqrt{n}\}).$$

Последняя сумма не превосходит

$$\sum_{i=1}^{n-p-1} \mathsf{P}\{|S_i - S_{i+p}| \geqslant \lambda\sqrt{n}\} +$$
$$+ \sum_{i=1}^{n-p-1} \mathsf{P}(E_i \cap \{|S_n - S_{i+p}| \geqslant \lambda\sqrt{n}\}) +$$
$$+ \sum_{i=n-p}^{n-1} \mathsf{P}\{|S_n - S_i| \geqslant \lambda\sqrt{n}\},$$

где $p = p_n$. Каждый член в первой и третьей из этих сумм не превосходит $\mathsf{P}\{S_p^* \geqslant \lambda\sqrt{n}\}$, а вторую сумму можно оценить, используя тот факт, что $E_i \in \mathscr{M}_{-\infty}^{i}$:

$$\mathsf{P}\{\max_{i\leqslant n} |S_i| \geqslant 3\lambda\sqrt{n}\} \leqslant$$
$$\leqslant \mathsf{P}\{|S_n| \geqslant \lambda\sqrt{n}\} + (n+p)\mathsf{P}\{S_p^* \geqslant \lambda\sqrt{n}\} +$$
$$+ \sum_{i=1}^{n-p-1} \mathsf{P}(E_i)[\mathsf{P}\{|S_n - S_{i+p}| \geqslant \lambda\sqrt{n}\} + \varphi_p].$$

Неравенство (20.52) и то обстоятельство, что E_i не пересекаются, приводит теперь к неравенству

$$\mathsf{P}\{\max_{i\leqslant n} |S_i| \geqslant 3\lambda\sqrt{n}\} \leqslant$$
$$\leqslant \frac{\varepsilon}{\lambda^2} + 2n\mathsf{P}\{S_p^* \geqslant \lambda\sqrt{n}\} + \frac{\varepsilon}{\lambda^2} + \varphi_p.$$

Отсюда, из (20.51) и сходимости p_n к ∞ вытекает, что для больших n

$$\mathsf{P}\{\max_{i\leqslant n}|S_i|\geqslant 3\lambda\sqrt{n}\}<\frac{3\varepsilon}{\lambda^2},$$

что отличается от (20.50) лишь несущественными множителями.

Теперь остается доказать (20.48). Если случайная величина $|\xi_0|$ ограничена постоянной C, то по лемме 4

$$\mathsf{E}_\alpha\left\{\frac{1}{n}S_n^2\right\}\leqslant\frac{1}{\alpha}\mathsf{E}\left\{\frac{1}{n^2}S_n^4\right\}\leqslant\frac{K_\varphi C^4}{\alpha}, \qquad (20.53)$$

где K_φ зависит лишь от φ. Таким образом, (20.48) верно, если ξ_0 ограничена. Если же величина ξ_0 неограничена, то для всяких действительного x и положительного u определим

$$f_u(x)=\begin{cases}x, & \text{если } |x|\leqslant u,\\ 0, & \text{если } |x|>u,\end{cases} \qquad g_u(x)=\begin{cases}0, & \text{если } |x|\leqslant u,\\ x, & \text{если } |x|>u,\end{cases}$$

и положим

$$\bar{f}_u(x)=f_u(x)-\mathsf{E}\{f_u(\xi_0)\},\quad \bar{g}_u(x)=g_u(x)-\mathsf{E}\{g_u(\xi_0)\}.$$

Тогда $x=f_u(x)+g_u(x)=\bar{f}_u(x)+\bar{g}_u(x)$, так что если $S_{nu}=\sum_{j=1}^{n}\bar{f}_u(\xi_j)$ и $D_{nu}=\sum_{j=1}^{n}\bar{g}_u(\xi_j)$, то $S_n=S_{nu}+D_{nu}$ и, следовательно,

$$S_n^2\leqslant 2S_{nu}^2+2D_{nu}^2. \qquad (20.54)$$

Поскольку $\bar{f}_u(\xi_0)$ не превосходит $2u$, из (20.53) вытекает, что

$$\mathsf{E}_\alpha\left\{\frac{1}{n}S_{nu}^2\right\}\leqslant\frac{K_\varphi(2u)^4}{\alpha}. \qquad (20.55)$$

По лемме 1

$$|\mathsf{E}\{\bar{g}_u(\xi_0)\bar{g}_u(\xi_k)\}|\leqslant 2\varphi_k^{1/2}\mathsf{E}\{(\bar{g}_u(\xi_0))^2\}\leqslant$$
$$\leqslant 2\varphi_k^{1/2}\mathsf{E}\{(g_u(\xi_0))^2\}=2\varphi_k^{1/2}\mathsf{E}_{u^2}\{\xi_0^2\}$$

и согласно (20.32) имеем

$$\mathsf{E}\left\{\frac{1}{n}D_{nu}^2\right\}\leqslant 4\left[\sum_{k=0}^{\infty}\varphi_k^{1/2}\right]\mathsf{E}_{u^2}\{\xi_0^2\}. \qquad (20.56)$$

Из (20.54) и соотношения $\mathsf{E}_\alpha\{U+V\}\leqslant 2\mathsf{E}_{\frac{1}{2}\alpha}\{U\}+2\mathsf{E}\{V\}$ вытекает, что

$$\mathsf{E}_\alpha\left\{\frac{1}{n}S_n^2\right\}\leqslant 4\mathsf{E}_{\frac{1}{2}\alpha}\left\{\frac{1}{n}S_{nu}^2\right\}+4\mathsf{E}\left\{\frac{1}{n}D_{nu}^2\right\},$$

откуда согласно (20.55) и (20.56)

$$\mathsf{E}_\alpha\left\{\frac{1}{n}S_n^2\right\}\leqslant K'_\varphi\left[\frac{u^4}{\alpha}+\mathsf{E}_{u^2}\{\xi_0^2\}\right],$$

где K'_φ зависит лишь от φ. Выбирая теперь u так, чтобы $K'_\varphi\mathsf{E}_{u^2}\{\xi_0^2\}<\frac{1}{2}\varepsilon$, а затем α так, чтобы $K'_\varphi u^4/\alpha<\frac{1}{2}\varepsilon$, получаем (20.48). Это завершает доказательство теоремы (20.1).

При выполнении условий теоремы мы имеем, конечно,

$$\frac{1}{\sqrt{n}}S_n\xrightarrow{\mathscr{D}}N(0,\sigma^2), \tag{20.57}$$

если $\sigma^2>0$. Если же $\sigma=0$, то (20.57) равносильно соотношению $S_n/\sqrt{n}\xrightarrow{P}0$ и легко вытекает из неравенства Чебышева и леммы 3 *).

Нетрудно получить многомерный вариант соотношения (20.57). Если ξ_n есть случайный вектор

$$\xi_n=(\xi_n^{(1)},\ \ldots,\ \xi_n^{(r)}), \tag{20.58}$$

то понятие φ-перемешивания можно определить точно так же, как и прежде. Если $\sum\limits_n\varphi_n^{1/2}<\infty$ и $\xi_n^{(i)}$ имеют нулевое среднее значение и конечные дисперсии, то, применяя технику Крамера — Уолда (теорема 7.7) к (20.57), получаем, что вектор $n^{-\frac{1}{2}}\sum\limits_{k=1}^{n}\xi_k$ распределен асимптотически нормально с нулевым средним и ковариациями

$$\sigma_{ij}=E\{\xi_0^{(i)}\xi_0^{(j)}\}+\sum_{k=1}^{\infty}\mathsf{E}\{\xi_0^{(i)}\xi_k^{(j)}\}+\sum_{k=1}^{\infty}\mathsf{E}\{\xi_k^{(i)}\xi_0^{(j)}\}, \tag{20.59}$$

---

*) Если $\sigma^2=0$, модифицируя доказательство (20.50), можно показать даже то, что

$$\max_{i\leqslant n}|S_i|/\sqrt{n}\xrightarrow{P}0.$$

где ряды сходятся абсолютно. Матрица (σ_{ij}) может быть как вырожденной, так и невырожденной.

Полученные результаты применимы к каждому из четырех приведенных выше примеров. Отметим, что в примерах 1 (стр. 232) и 3 (стр. 235) ряд в (20.46) конечен. В примере 2, где $\xi_n = f(\zeta_n)$ и переходная матрица для $\{\zeta_n\}$ неприводима и апериодична, асимптотическая дисперсия σ^2 может быть записана в виде

$$\sigma^2 = \sum_{uv} \gamma_{uv} f(u) f(v),$$

где

$$\gamma_{uv} = \delta_{uv} p_u - p_u p_v + p_u \sum_{k=1}^{\infty} \left(p_{uv}^{(k)} - p_v\right) + p_v \sum_{k=1}^{\infty} \left(p_{vu}^{(k)} - p_u\right).$$

Если матрица (γ_{uv}) переводит вектор $(f(u))$ в нулевой вектор, то $\sigma^2 = 0$.

Интегралы вместо сумм. Теорему 20.1 можно естественным образом распространить на процесс $\{v_t\}$ с непрерывным временем. Пусть

$$\{v_t(\omega)\colon\ -\infty < t < \infty\} \tag{20.60}$$

— стационарный стохастический процесс, удовлетворяющий условиям

$$\mathsf{E}\{v_0\} = 0, \qquad \mathsf{E}\{v_0^2\} < \infty. \tag{20.61}$$

Предположим, что процесс (20.60) измерим *), так что интегралы

$$\int_0^t v_u(\omega)\, du \tag{20.62}$$

определены и конечны с вероятностью 1. Пусть Y_n — случайный элемент из D, определенный равенством

$$Y_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} \int_0^{nt} v_s(\omega)\, ds,$$

где σ будет определено позже (конечно, Y_n принадлежит C). Наша цель — показать, что

$$Y_n \xrightarrow{\mathscr{D}} W. \tag{20.63}$$

*) Дуб (1956, стр. 63).

Назовем $\{v_t\}$ процессом с φ-перемешиванием, если

$$|\mathsf{P}(E_1 \cap E_2) - \mathsf{P}(E_1)\mathsf{P}(E_2)| \leqslant \varphi(t)\mathsf{P}(E_1) \tag{20.64}$$

при всех E_1, принадлежащих σ-полю, порожденному семейством $\{v_u\colon\ u \leqslant s\}$, и при всех E_2, принадлежащих σ-полю, порожденному семейством $\{v_u\colon u \geqslant s+t\}$. Функция $\varphi(t)$ определена теперь для всех положительных t. Предположим, что $\{v_t\}$ является процессом с φ-перемешиванием с функцией φ такой, что

$$\int\limits_0^\infty \varphi^{1/2}(t)\,dt < \infty. \tag{20.65}$$

Мы докажем, что при этих предположениях и при выполнении (20.61), выполняется предельное соотношение (20.63) с величиной σ (которую мы считаем положительной), определенной равенством

$$\sigma^2 = 2\int\limits_0^\infty \mathsf{E}\{v_0 v_t\}\,dt. \tag{20.66}$$

Можно было бы провести прямое доказательство, видоизменив соответствующим образом доказательство для дискретного времени, однако проще свести настоящий случай к предыдущему. Если

$$\xi_i(\omega) = \int\limits_{i-1}^{i} v_s(\omega)\,ds,$$

то, применяя дважды теорему Фубини, получаем, что

$$\mathsf{E}\{\xi_0\} = 0$$

и

$$\mathsf{E}\{\xi_0^2\} \leqslant$$

$$\leqslant \mathsf{E}\left\{\left[\int\limits_0^1 |v_s|\,ds\right]^2\right\} \leqslant \mathsf{E}\left\{\int\limits_0^1 v_s^2\,ds\right\} = \mathsf{E}\{v_0^2\} < \infty. \tag{20.67}$$

Нетрудно проверить, что (20.66) и (20.46) дают одно и то же значение σ. Так как из (20.64) вытекает, что $\{\xi_n\}$ удовлетворяет условию (20.2), а (20.65) влечет сходимость ряда $\sum\limits_n \varphi^{1/2}(n)$ (при условии, что φ имеет минимальное значение, удовлетворяющее (20.64)), то из теоремы 20.1 следует, что для случайной функции, определенной формулой (20.45), имеем

$$X_n \xrightarrow{\mathscr{D}} W. \tag{20.68}$$

Далее,

$$\delta_n = \sup_t |Y_n(t) - X_n(t)| \leqslant \frac{1}{\sigma\sqrt{n}} \max_{1 \leqslant i \leqslant n} \int_{i-1}^{i} |v_t|\, dt,$$

так что

$$\mathsf{P}\{\delta_n \geqslant \varepsilon\} \leqslant \frac{1}{\varepsilon^2\sigma^2} \int\limits_{\{|\zeta| \geqslant \varepsilon\sigma\sqrt{n}\}} \zeta^2\, d\mathsf{P},$$

где $\zeta = \int_0^1 |v_s|\, ds$. Поскольку в силу (20.67) математическое ожидание $\mathsf{E}\{\zeta^2\}$ конечно, то $\delta_n \xrightarrow{P} 0$ и (20.63) следует из соотношения (20.68) и теоремы 4.1.

Соотношение (20.63) остается справедливым и при непрерывном стремлении n к бесконечности. Если v_t интерпретировать как скорость частицы в момент времени t, то интеграл (20.62) равен ее перемещению в интервале времени (s, t). Соотношение (20.63) означает, что движение частицы близко к броуновскому движению.

Нестационарность. Возвращаясь к случаю дискретного времени, обобщим теорему 16.3 следующим образом. Покажем, что в теореме 20.1 вероятностную меру P, управляющую процессом $\{\xi_n\}$, можно заменить любой вероятностной мерой P_0, абсолютно непрерывной относительно P. Процесс, отвечающий мере P_0, не обязан быть стационарным.

Теорема 20.2. *Теорема 20.1 остается в силе при замене P произвольной вероятностной мерой P_0, абсолютно непрерывной относительно P.*

Доказательство. Определим X_n соотношением (20.45) и положим

$$X'_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} \sum_{p_n \leqslant i \leqslant nt} \xi_i(\omega),$$

где $p_n \to \infty$, но $p_n/\sqrt{n} \to 0$. Так же как и в § 16 (см. (16.10)), имеем

$$\sup_t |X_n(t) - X'_n(t)| \xrightarrow{P} 0.$$

Если E принадлежит $\mathcal{M}_{-\infty}^{k}$, то для каждого A из $\mathcal{D}$

$$|\mathsf{P}(\{X'_n \in A\} \cap E) - \mathsf{P}\{X'_n \in A\}\mathsf{P}(E)| \leqslant \varphi(p_n - k) \to 0.$$

Поскольку это верно для всех k, то дальнейшая часть доказательства проводится точно так же, как в теореме 16.3.

Таким образом, в примере 4 (стр. 235) гауссову меру можно заменить лебеговой.

Следующая теорема может быть доказана аналогичным видоизменением доказательства теоремы 17.2.

Теорема 20.3. *Если условия теоремы* 20.1 *выполнены и* ν_n — *случайные величины с целыми положительными значениями такие, что* $\nu_n/a_n \xrightarrow{P} \theta$, *где* θ — *положительная случайная величина и* $a_n \to \infty$, *то при* $\sigma > 0$

$$Y_n \xrightarrow{\mathcal{D}} W,$$

где

$$Y_n(t, \omega) = \frac{1}{\sigma\sqrt{\nu_n(\omega)}} S_{[\nu_n(\omega)t]}(\omega).$$

Если $\mathsf{P}\{\xi_0 \in H\} > 0$, то по теореме 20.2 все наши предельные теоремы останутся в силе, если вероятности заменить на условные вероятности относительно события $\{\xi_0 \in H\}$. Можно пойти дальше и рассмотреть условные вероятности относительно события $\{\xi_0 = \alpha\}$. Чтобы убедиться в этом, нам нужна еще одна лемма. Как обычно, предположим, что $\{\xi_n\}$ является последовательностью с φ-перемешиванием.

Лемма 5. *Если* $E_2 \in \mathcal{M}_{k+n}^{\infty}$, *то*

$$|\mathsf{P}\{E_2 \| \mathcal{M}_{-\infty}^{k}\} - \mathsf{P}(E_2)| \leqslant \varphi_n \tag{20.69}$$

с вероятностью 1.

Доказательство. Допустим, что неравенство $\mathsf{P}\{E_2 \| \mathcal{M}_{-\infty}^{k}\} - \mathsf{P}(E_2) > \varphi_n$ справедливо для некоторого множества E_1, имеющего положительную меру и принадлежащую $\mathcal{M}_{-\infty}^{k}$. Интегрирование по множеству E_1 приводит к противоречию с (20.2). Таким образом,

$\mathsf{P}\{E_2 \| \mathscr{M}^k_{-\infty}\} - \mathsf{P}(E_2) \leqslant \varphi_n$ с вероятностью 1. Аналогично устанавливается и обратное неравенство.

Предположим теперь, что последовательность $\{\xi_n\}$ регулярна в следующем смысле. Во-первых, ξ_n порождает σ-поле $\mathscr{B}$ (по существу это не является ограничением). Во-вторых, существует условная вероятность на $\mathscr{B}$ относительно $\mathscr{M}^0_{-\infty}$. Иными словами, предполагается, что существует функция $\mathsf{P}_\omega(M)$ такая, что для каждой точки ω из Ω $\mathsf{P}_\omega(\cdot)$ является вероятностной мерой на $\mathscr{B}$ и для каждого E из $\mathscr{B}$ функция $\mathsf{P}_.(E)$ является вариантом условной вероятности $\mathsf{P}\{E \| \mathscr{M}^0_{-\infty}\}$ (и, следовательно, $\mathscr{M}^0_{-\infty}$-измерима). Это условие регулярности выполнено, например, если $(\Omega, \mathscr{B})$ есть произведение бесконечной в обе стороны последовательности измеримых пространств $(R^1, \mathscr{R}^1)$, а ξ_n — координатные случайные величины *).

Теорема 20.4 *Пусть последовательность $\{\xi_n\}$ регулярна в описанном выше смысле. При выполнении условий теоремы* 20.1 *в $\mathscr{B}$ существует множество E^* такое, что $\mathsf{P}(E^*) = 0$, и если $\omega_0 \notin E^*$, то утверждение этой теоремы остается верным при замене P на P_{ω_0}.*

Доказательство. Пусть $S_n = \xi_1 + \ldots + \xi_n$, и пусть случайная функция X_n определена равенством (20.45). Для целых положительных u и n определим случайный элемент Z_{un} в D, положив

$$Z_{un}(t, \omega) = \frac{1}{\sigma\sqrt{n}} \sum_{u \leqslant i \leqslant nt} \xi_i(\omega).$$

Если

$$\delta_{un}(\omega) = \sup_t |X_n(t, \omega) - Z_{un}(t, \omega)|,$$

то

$$|\delta_{un}(\omega)| \leqslant \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{u} |\xi_i(\omega)|.$$

Обозначим E_0 множество тех ω, для которых $\xi_i(\omega) = \pm\infty$ для некоторого i. Имеем $\mathsf{P}(E_0) = 0$ и для любого u

$$\lim_{n\to\infty} \delta_{un}(\omega) = 0, \qquad \omega \notin E_0. \tag{20.70}$$

*) См. Лоэв (1962, стр. 379).

Так как по теореме (20.1) $X_n \xrightarrow{\mathscr{D}} W$, то, используя (20.70) и теорему 4.1, получаем, что для любого u

$$Z_{un} \xrightarrow{\mathscr{D}} W \tag{20.71}$$

при $n \to \infty$. В (20.71) последовательностью ξ_n управляет мера P.

Поскольку $\mathsf{P}(E_0) = 0$, то $\mathsf{P}_{\omega_0}(E_0) = 0$ при всех ω_0, лежащих вне некоторого исключительного множества E_1 P-меры 0. Если F — замкнутое множество в D и $F_\varepsilon = \{x\colon d(x, F) \leqslant \varepsilon\}$, где d — какая-либо метрика для топологии Скорохода, то из (20.70) следует, что

$$\limsup_{n\to\infty} \mathsf{P}_{\omega_0}\{\omega\colon X_n(\omega) \in F\} \leqslant \limsup_{n\to\infty} \mathsf{P}_{\omega_0}\{\omega\colon Z_{un}(\omega) \in F_\varepsilon\} \tag{20.72}$$

для всех ω_0, не принадлежащих E_1.

По лемме 5, если $E \in \mathscr{M}_u^\infty$, то

$$|\mathsf{P}_{\omega_0}(E) - \mathsf{P}(E)| \leqslant \varphi(u) \tag{20.73}$$

для ω_0, лежащих вне некоторого исключительного множества (P-меры 0), зависящего от E. Так как σ-поле $\mathscr{M}_u^\infty$ порождено семейством $\{\xi_u, \xi_{u+1}, \ldots\}$, то оно порождено счетным подполем. Поскольку $\mathsf{P}_{\omega_0}(\cdot)$ есть мера, от отсюда следует, что (20.73) верно для всех $u \geqslant 1$ и всех $E \in \mathscr{M}_u^\infty$, если ω_0 находится вне некоторого большого исключительного множества E_2. Далее, множество $\{\omega\colon Z_{un}(\omega) \in F_\varepsilon\}$ принадлежит $\mathscr{M}_u^\infty$, и поэтому

$$\mathsf{P}_{\omega_0}\{\omega\colon Z_{un}(\omega) \in F_\varepsilon\} \leqslant \mathsf{P}\{\omega\colon Z_{un}(\omega) \in F_\varepsilon\} + \varphi(u) \tag{20.74}$$

при $\omega_0 \notin E_2$.

Если $\omega_0 \notin E^* = E_1 \cup E_2$, то справедливы оба соотношения (20.72) и (20.74), так что

$$\limsup_{n\to\infty} \mathsf{P}_{\omega_0}\{\omega\colon X_n(\omega) \in F\} \leqslant \leqslant \limsup_{n\to\infty} \mathsf{P}\{\omega\colon Z_{un}(\omega) \in F_\varepsilon\} + \varphi(u)$$

для любых F и ε. Поскольку соотношение (20.71) имеет место в смысле меры P и множество F_ε замкнуто, то верхний предел в правой части последнего неравенства не превосходит $\mathsf{P}\{W \in F_\varepsilon\}$. Полагая $\varepsilon \to 0$ и $u \to \infty$, заключаем, что

$$\limsup_{n\to\infty} \mathsf{P}_{\omega_0}\{\omega\colon X_n(\omega) \in F\} \leqslant \mathsf{P}\{W \in F\}$$

для всех замкнутых множеств F, если $\omega_0 \notin E^*$. Для завершения доказательства остается теперь применить теорему 2.1

Для односторонней последовательности теорема 20.4 остается верной при замене вероятностей условными вероятностями относительно события $\{\xi_1 = \alpha\}$.

§ 21. Функции от процессов с перемешиванием

В этом параграфе мы обобщим результаты предыдущего параграфа, рассматривая стационарный процесс $\{\eta_n\}$, где каждая случайная величина η_n является функцией от всей последовательности $\{\xi_n\}$. Как и прежде, предполагается, что последовательность $\{\xi_n\}$ стационарна и удовлетворяет условию φ-перемешивания.

Предварительные замечания. Пусть f — измеримое отображение пространства бесконечных в обе стороны последовательностей действительных чисел в действительную прямую

$$f(\ldots, \alpha_{-1}, \alpha_0, \alpha_1, \ldots) \in R^1.$$

Определим случайные величины

$$\eta_n = f(\ldots, \xi_{n-1}, \xi_n, \xi_{n+1}, \ldots), \quad n = 0, \pm 1, \pm 2, \ldots, \tag{21.1}$$

где ξ_n занимает нулевое место в аргументе функции f. Мы докажем предельные теоремы для процесса $\{\eta_n\}$. Возможные значения величин ξ_n предполагаются действительными, но ξ_n может принимать значения из R^k и даже из общего измеримого пространства.

Если функция f зависит лишь от конечного числа координат ее аргумента, то $\{\eta_n\}$ тоже является процессом с φ-перемешиванием (с новой функцией φ) и, следовательно, применимы теоремы из § 20. (Заметим, что процесс $\{\eta_n\}$ в любом случае стационарен.) Если же f существенно зависит от всех координат ее аргумента, то $\{\eta_n\}$ не обязательно будет процессом с φ-перемешиванием. Предельные теоремы для $\{\eta_n\}$ мы получим в предположении, что величины η_n хорошо аппроксимируются функциями, зависящими лишь от конечного числа величин ξ_n.

Пусть f_l — измеримое отображение пространства R^{2l+1} в R^1. Для точек пространства R^{2l+1} примем обозначение $(\alpha_{-l}, \ldots, \alpha_0, \ldots, \alpha_l)$, так что значение f_l запишется в виде

$$f_l(\alpha_{-l}, \ldots, \alpha_0, \ldots, \alpha_l).$$

Положим

$$\eta_{ln} = f_l(\xi_{n-l}, \ldots, \xi_n, \ldots, \xi_{n+l}) \tag{21.2}$$

(здесь индекс ln — пара, а не произведение). Процесс $\{\eta_{ln}\}$ стационарен для каждого l. Будем считать, что функция f и процесс $\{\xi_n\}$ удовлетворяют условиям

$$\mathsf{E}\{\eta_0\} = 0, \quad \mathsf{E}\{\eta_0^2\} < \infty.$$

Предположим далее, что существуют функции f_l такие, что величины η_{ln} (определенные равенством (21.2)) имеют конечный второй момент и

$$\sum_{l=1}^{\infty} \nu_l^{1/2} < \infty,$$

где

$$\nu_l = \nu(l) = \mathsf{E}\{|\eta_0 - \eta_{l0}|^2\}. \tag{21.3}$$

Хорошая аппроксимация η_n функциями лишь от конечного числа ξ_n понимается именно в этом смысле.

Если вместо бесконечной в обе стороны последовательности (20.1) имеется односторонняя последовательность (20.16), то вместо (21.1) нужно предположить, что η_n имеет вид

$$\eta_n = f(\xi_n, \xi_{n+1}, \ldots), \tag{21.4}$$

где f — теперь действительная измеримая функция от односторонних последовательностей $(\alpha_1, \alpha_2, \ldots)$. В этом случае в качестве f_l мы берем измеримое отображение R^l в R^1 и получаем

$$\eta_{ln} = f_l(\xi_n, \ldots, \xi_{n+l-1}). \tag{21.5}$$

Предполагается, что, как и прежде, $\sum\limits_l \nu_l^{1/2} < \infty$, где

$$\nu_l = \mathsf{E}\{|\eta_1 - \eta_{l1}|^2\}. \tag{21.6}$$

Заменим последовательность $\{\xi_n\}$ ее двусторонне бесконечным расширением (см. обсуждение, следующее за (20.16)). Если мы запишем $f(\alpha_1, \alpha_2, \ldots)$ в виде функции $f(\ldots, \alpha_{-1}, \alpha_0, \alpha_1, \ldots)$, в действительности зависящей лишь от координат с положительными индексами, и проделаем то же самое с функциями f_l, то последовательности (21.1) и (21.2) будут иметь те же конечномерные распределения, что и последовательности (21.4) и (21.5). Следовательно, $\sum\limits_l v_l^{1/2} < \infty$ (см. (21.3)), и поэтому каждой предельной теореме для функций от (20.1) соответствует предельная теорема для функций от (20.16). Таким образом, как и раньше, мы можем ограничиться рассмотрением двусторонне бесконечного случая.

Поскольку величина η_{l0}, определенная формулой (21.2), $\mathscr{M}_{-l}^{l}$-измерима, из неравенства Йенсена для условных математических ожиданий *) следует, что

$$\mathsf{E}\left\{\left|\eta_{l0} - \mathsf{E}\left\{\eta_0 \| \mathscr{M}_{-l}^{l}\right\}\right|^2\right\} \leqslant \mathsf{E}\left\{\mathsf{E}\left\{\left|\eta_{l0} - \eta_0\right|^2 \| \mathscr{M}_{-l}^{l}\right\}\right\} = v_l.$$

Поэтому, в силу неравенства Минковского,

$$\mathsf{E}^{1/2}\left\{\left|\eta_0 - \mathsf{E}\left\{\eta_0 \| \mathscr{M}_{-l}^{l}\right\}\right|^2\right\} \leqslant 2v_l^{1/2}.$$

Так как ряд $\sum\limits_l 2v_l^{1/2}$ сходится, если сходится ряд $\sum\limits_l v_l^{1/2}$, то без ограничения общности можно считать, что

$$\eta_{ln} = \mathsf{E}\left\{\eta_n \| \mathscr{M}_{n-l}^{n+l}\right\} \text{ **)}. \qquad (21.7)$$

Следующая лемма показывает, что если величины η_{ln} заданы формулой (21.7), то последовательность v_l не убывает.

---

*) Дуб (1956) или Биллингсли (1969). Всюду ниже мы считаем $\mathsf{E}\{\xi \| \mathscr{F}\}$ $\mathscr{F}$-измеримым (а не совпадающим только почти всюду с некоторой $\mathscr{F}$-измеримой функцией).

**) Согласно Дубу (1956, стр. 542) условное математическое ожидание $\mathsf{E}\left\{\eta_n \| \mathscr{M}_{n-l}^{n+l}\right\}$ представимо в виде (21.2), так как оно измеримо $\mathscr{M}_{n-l}^{n+l}$. Однако нам это неважно: в доказательствах используется лишь тот факт, что последовательность $\{\eta_{ln}\}$ стационарна при каждом l и что величины η_{ln} измеримы относительно $\mathscr{M}_{n-l}^{n+l}$.

Лемма 1. *Пусть $\mathscr{F}$ и $\mathscr{G}$ — σ-поля такие, что $\mathscr{F} \subset \mathscr{G}$. Если* $\mathsf{E}\{\xi^2\} < \infty$, *то*

$$\mathsf{E}\{|\xi - \mathsf{E}\{\xi \| \mathscr{G}\}|^2\} \leqslant \mathsf{E}\{|\xi - \mathsf{E}\{\xi \| \mathscr{F}\}|^2\}. \tag{21.8}$$

Доказательство. Следующие равенства и неравенства верны с вероятностью 1. Если $\eta = \xi - \mathsf{E}\{\xi \| \mathscr{F}\}$, то

$$\xi - \mathsf{E}\{\xi \| \mathscr{G}\} = \eta - \mathsf{E}\{\xi \| \mathscr{G}\}.$$

Стало быть, нужно доказать, что $\mathsf{E}\{|\eta - \mathsf{E}\{\eta \| \mathscr{G}\}|^2\} \leqslant \mathsf{E}\{\eta^2\}$. Беря математическое ожидание от обеих сторон неравенства

$$\mathsf{E}\{|\eta - \mathsf{E}\{\eta \| \mathscr{G}\}|^2 \| \mathscr{G}\} = \mathsf{E}\{\eta^2 \| \mathscr{G}\} - \mathsf{E}^2\{\eta \| \mathscr{G}\} \leqslant \mathsf{E}\{\eta^2 \| \mathscr{G}\},$$

получаем нужный результат.

Функциональная центральная предельная теорема. Обозначим

$$S_n = \eta_1 + \ldots + \eta_n, \tag{21.9}$$

и пусть

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega). \tag{21.10}$$

Теорема 21.1. *Пусть $\{\xi_n\}$ — процесс с φ-перемешиванием, причем $\sum_n \varphi_n^{1/2} < \infty$, и случайная величина η_n, определенная соотношением* (21.1), *имеет нулевое среднее значение и конечную дисперсию. Пусть, далее, существуют случайные величины вида* (21.2) *такие, что $\sum_l \nu_l^{1/2} < \infty$, где ν_l определяется соотношением* (21.3). *Тогда ряд*

$$\sigma^2 = \mathsf{E}\{\eta_0^2\} + 2\sum_{k=1}^{\infty} \mathsf{E}\{\eta_0\eta_k\} \tag{21.11}$$

абсолютно сходится; если $\sigma^2 > 0$ и X_n — случайная функция, определенная формулой (21.10), *то*

$$X_n \xrightarrow{\mathscr{D}} W. \tag{21.12}$$

Доказательство. Покажем сначала, что ряд в (21.11) абсолютно сходится. Для каждого k и i

$$|\mathsf{E}\{\eta_0\eta_k\}| \leqslant |\mathsf{E}\{\mathsf{E}\{\eta_0 \| \mathscr{M}_{-i}^{i}\}\mathsf{E}\{\eta_k \| \mathscr{M}_{k-i}^{k+i}\}\}| +$$
$$+ 2\mathsf{E}^{1/2}\{|\mathsf{E}\{\eta_0 \| \mathscr{M}_{-i}^{i}\}|^2\}\mathsf{E}^{1/2}\{|\eta_0 - \mathsf{E}\{\eta_0 \| \mathscr{M}_{-i}^{i}\}|^2\} +$$
$$+ \mathsf{E}\{|\eta_0 - \mathsf{E}\{\eta_0 \| \mathscr{M}_{-i}^{i}\}|^2\}.$$

Если $i=\left[\frac{1}{3}k\right]$, то по лемме 1 предыдущего параграфа

$$|\mathsf{E}\{\mathsf{E}\{\eta_0\|\mathcal{M}_{-i}^{i}\}\mathsf{E}\{\eta_k\|\mathcal{M}_{k-i}^{k+i}\}\}|\leqslant$$
$$\leqslant 2\varphi_i^{1/2}\mathsf{E}\{\mathsf{E}^2\{\eta_0\|\mathcal{M}_{-i}^{i}\}\}\leqslant 2\varphi_i^{1/2}\mathsf{E}\{\eta_0^2\},$$

и следовательно,

$$|\mathsf{E}\{\eta_0\eta_k\}\leqslant 2\varphi^{1/2}\left(\left[\tfrac{k}{3}\right]\right)\mathsf{E}\{\eta_0^2\}+$$
$$+2\mathsf{E}^{1/2}\{\eta_0^2\}\,\nu^{1/2}\left(\left[\tfrac{k}{3}\right]\right)+\nu\left(\left[\tfrac{k}{3}\right]\right). \qquad (21.13)$$

Таким образом, ряд в (21.11) действительно сходится, так как по предположению сходятся оба ряда $\sum\limits_n \varphi_n^{1/2}$ и $\sum\limits_l \nu_l^{1/2}$.

Соотношение (21.12) можно было бы вывести из теоремы 19.2, однако более удобно для этой цели использовать результаты предыдущего параграфа. Докажем сначала, что

$$\frac{1}{\sqrt{n}}S_n\xrightarrow{\mathcal{D}}N(0,\ \sigma^2). \qquad (21.14)$$

Можно предположить, что величины η_{ln} заданы посредством соотношения (21.7). Положим

$$\delta_{ln}=\eta_n-\eta_{ln}, \qquad (21.15)$$

так что $\nu(l)=\mathsf{E}\{\delta_{ln}^2\}$ и $\mathsf{E}\{\delta_{ln}\}=\mathsf{E}\{\eta_{ln}\}=0$. Имеем

$$\frac{1}{\sqrt{n}}S_n=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\eta_{lj}+\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\delta_{lj}. \qquad (21.16)$$

Чтобы доказать (21.14), мы, используя теорему 20.1, покажем, что при больших n первая сумма в правой части (21.6) приблизительно нормальна, а затем убедимся, что вторая сумма мала при больших l.

Прежде всего, имеем

$$\mathsf{E}\{|\eta_{l0}-\mathsf{E}\{\eta_{l0}\|\mathcal{M}_{-i}^{i}\}|^2\}\leqslant\nu_l.$$

В самом деле, при $i \geqslant l$ левая часть здесь равна нулю, а при $i \leqslant l$ она совпадает с величиной $\mathsf{E}\{\mathsf{E}^2\{\eta_i - \mathsf{E}\{\eta_0 \| \mathscr{M}_{-i}^{i}\} \| \mathscr{M}_{-l}^{l}\}\}$, которая по теореме Йенсена не превосходит ν_i. Следовательно, в рассуждении, которое привело нас к (21.13), мы можем заменить η_0 на η_{l0}, а η_k на η_{lk}. Поскольку $\mathsf{E}\{\eta_{l0}^2\} \leqslant \mathsf{E}\{\eta_0^2\}$, то мы можем заключить теперь, что $|\mathsf{E}\{\eta_{l0}\eta_{lk}\}|$ не превосходит правой части неравенства (21.13). Таким образом, ряд

$$\sigma_l^2 = \mathsf{E}\{\eta_{l0}^2\} + 2\sum_{k=1}^{\infty} \mathsf{E}\{\eta_{l0}\eta_{lk}\}$$

сходится абсолютно и равномерно по l. Далее, поскольку этот ряд почленно сходится к ряду (21.11), имеем

$$\lim_{l\to\infty} \sigma_l^2 = \sigma^2. \tag{21.17}$$

Нетрудно проверить, что для каждого l $\{\eta_{ln}\}$ есть процесс с $\varphi^{(l)}$-перемешиванием, где

$$\varphi^{(l)}(n) = \begin{cases} 1, & \text{если } n \leqslant 2l, \\ \varphi(n-2l), & \text{если } n > 2l. \end{cases}$$

Сходимость $\sum_n (\varphi(n))^{1/2}$ влечет сходимость $\sum_n (\varphi^{(l)}(n))^{1/2}$ и по теореме 20.1 для каждого фиксированного l имеем

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \eta_{lj} \xrightarrow{\mathscr{D}} N(0, \sigma_l^2) \tag{21.18}$$

при $n \to \infty$.

Поскольку согласно (21.17) $N(0, \sigma_l^2) \xrightarrow{\mathscr{D}} N(0, \sigma^2)$, то в силу (21.16) и (21.18) соотношение (21.14) будет следовать из теоремы 4.2, если мы покажем, что

$$\lim_{l\to\infty} \limsup_{n\to\infty} \mathsf{P}\left\{\left|\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \delta_{lj}\right| \geqslant \varepsilon\right\} = 0 \tag{21.19}$$

для любого положительного ε. С этой целью оценим дисперсию $\sum_{j=1}^{n} \delta_{lj}$.

Если m означает наибольшее из чисел l и i, то по определению (21.15) $\delta_{l0} - \mathsf{E}\{\delta_{l0} \| \mathcal{M}_{-i}^{i}\} = \delta_{m0}$. Так как $\nu(m) \leqslant \nu(i)$, то

$$\mathsf{E}\{|\delta_{l0} - \mathsf{E}\{\delta_{l0} \| \mathcal{M}_{-i}^{i}\}|^2\} \leqslant \nu_l.$$

Следовательно, в рассуждении, приводящем к неравенству (21.13), мы можем η_0 заменить на δ_{l0} и η_k на δ_{lk}, поскольку по лемме 1 $\mathsf{E}\{\delta_{l0}^2\} \leqslant \mathsf{E}\{\eta_0^2\}$, то $|\mathsf{E}\{\delta_{l0}\delta_{lk}\}|$ не превосходит правой части (21.13). Поэтому ряд

$$\tau_l^2 = \mathsf{E}\{\delta_{l0}^2\} + 2\sum_{k=1}^{\infty} \mathsf{E}\{\delta_{l0}\delta_{lk}\}$$

сходится абсолютно и равномерно по l. Так как каждый член этого ряда стремится к 0 при $l \to \infty$, то

$$\lim_{l\to\infty} \tau_l^2 = 0. \tag{21.20}$$

Применяя к $\{\delta_{ln}\}$ лемму 3 из предыдущего параграфа ($\{\delta_{ln}\}$ не является последовательностью с φ-перемешиванием, но лемма этого и не требует), имеем

$$\lim_{n\to\infty} \frac{1}{n} \mathsf{E}\left\{\left|\sum_{j=1}^{n} \delta_{lj}\right|^2\right\} = \tau_l^2. \tag{21.21}$$

Неравенство Чебышева и (21.20) приводят теперь к соотношению (21.19), что и завершает доказательство (21.14).

Из абсолютной сходимости ряда (21.11) и леммы 3 из предыдущего параграфа (напомним еще раз, что эта лемма требует лишь стационарности — φ-перемешивание не обязательно) следует, что

$$\frac{1}{n} \mathsf{E}\{S_n^2\} \to \sigma^2. \tag{21.22}$$

При $\sigma^2 = 0$ соотношение (21.14) равносильно соотношению $S_n/\sqrt{n} \xrightarrow{P} 0$. Начиная с этого места, мы будем предполагать, что $\sigma^2 > 0$. Чтобы доказать (21.12), мы установим сначала сходимость конечномерных распределений, а затем — плотность. Пусть p_n — целые положительные числа, стремящиеся к бесконечности с некоторой скоростью, которая будет уточнена ниже. Положим

$$U_{ni} = \mathsf{E}\left\{S_{i-2p_n} \,\middle\|\, \mathcal{M}_{-\infty}^{i-p_n}\right\} \tag{21.23}$$

и

$$V_{ni} = \mathsf{E}\left\{S_n - S_{i+2p_n} \,\middle\|\, \mathcal{M}_{i+p_n}^{\infty}\right\}. \tag{21.24}$$

В этих формулах мы полагаем $S_{i-2p_n} = 0$ при $i < 2p_n$ и $S_n - S_{i+2p_n} = 0$ при $i + 2p_n > 0$. В дальнейшем вместо p_n мы будем часто писать просто p.

Используя неравенство Минковского и лемму 1, имеем

$$\mathsf{E}^{1/2}\{|S_k - \mathsf{E}\{S_k \| \mathcal{M}_{-\infty}^{k+p}\}|^2\} \leqslant$$
$$\leqslant \sum_{j=1}^{k} \mathsf{E}^{1/2}\{|\eta_j - \mathsf{E}\{\eta_j \| \mathcal{M}_{-\infty}^{k+p}\}|^2\} \leqslant \sum_{j=1}^{k} \nu^{1/2}(k+p-j).$$

Если

$$\mu(p) = \sum_{i=p}^{\infty} \nu^{1/2}(i), \tag{21.25}$$

то

$$\mathsf{E}\{|S_k - \mathsf{E}\{S_k \| \mathcal{M}_{-\infty}^{k+p}\}|^2\} \leqslant \mu^2(p), \tag{21.26}$$

и поэтому

$$\mathsf{E}\{|U_{ni} - S_i|^2\} \leqslant 2\mathsf{E}\{S_{2p_n}^2\} + 2\mu^2(p_n) \tag{21.27}$$

для всех i. Аналогично получаем

$$\mathsf{E}\{|V_{ni} - (S_n - S_i)|^2\} \leqslant 2\mathsf{E}\{S_{2p_n}^2\} + 2\mu^2(p_n). \tag{21.28}$$

Так как величины U_{ni}, V_{ni} измеримы относительно σ-полей $\mathcal{M}_{-\infty}^{i-p}$, $\mathcal{M}_{i+p}^{\infty}$, соответственно, то

$$|\mathsf{P}\{U_{ni} \in H_1,\ V_{ni} \in H_2\} - \mathsf{P}\{U_{ni} \in H_1\}\mathsf{P}\{V_{ni} \in H_2\}| \leqslant$$
$$\leqslant \varphi(2p_n) \tag{21.29}$$

для любых линейных борелевских множеств H_1 и H_2.

Рассмотрим теперь конечномерные распределения. Из (21.14) следует, что

$$X_n(t) - X_n(s) \xrightarrow{\mathcal{D}} W_t - W_s. \tag{21.30}$$

Мы докажем, что

$$(X_n(t),\ X_n(1)-X_n(t)) \xrightarrow{\mathcal{D}} (W_t,\ W_1-W_t); \qquad (21.31)$$

рассуждение легко переносится и на случай размерности большей, чем 2. Пусть p_n стремится к бесконечности настолько медленно, что $p_n/n \to 0$. Поскольку согласно (21.22) $\mathsf{E}\{S_k^2\} \sim k\sigma^2$, используя неравенство (21.27) и неравенство Чебышева, получаем

$$\frac{1}{\sigma\sqrt{n}} U_{n,\,[nt]} - X_n(t) \xrightarrow{P} 0. \qquad (21.32)$$

Аналогично,

$$\frac{1}{\sigma\sqrt{n}} V_{n,\,[nt]} - (X_n(1)-X_n(t)) \xrightarrow{P} 0 \qquad (21.33)$$

и (21.31) будет установлено, если показать, что

$$\left(\frac{1}{\sigma\sqrt{n}} U_{n,\,[nt]},\ \frac{1}{\sigma\sqrt{n}} V_{n,\,[nt]}\right) \xrightarrow{\mathcal{D}} (W_t,\ W_1-W_t).$$

В силу (21.29) достаточно доказать (теорема 3.1) сходимость по распределению каждой из двух координат; но это следует из (21.30), поскольку имеют место соотношения (21.32) и (21.33).

Вернемся теперь к вопросу о плотности и покажем, что для любого положительного ε существует $\lambda > \sigma$ такое, что

$$\mathsf{P}\left\{\max_{i \leqslant n} |S_i| \geqslant \lambda\sqrt{n}\right\} \leqslant \frac{\varepsilon}{\lambda^2} \qquad (21.34)$$

при всех достаточно больших n. Положим

$$S_n^* = \sum_{j=1}^{n} |\eta_j|.$$

Рассуждая так же, как при доказательстве (20.51), можно построить последовательность p_n, стремящуюся к бесконечности настолько медленно, что если

$$\beta_n(\lambda) = \mathsf{P}\left\{S_{2p_n}^* \geqslant \lambda\sqrt{n}\right\}, \qquad (21.35)$$

то

$$\lim_{n\to\infty} n\beta_n(\lambda) = 0 \qquad (21.36)$$

для любого положительного λ.

Рассмотрим величины (21.23) и (21.24) для этой последовательности p_n. Используя обозначение (21.35), имеем

$$\mathsf{P}\{|S_i-U_{ni}|\geqslant\lambda\sqrt{n}\}\leqslant$$
$$\leqslant\mathsf{P}\left\{\left|S_{i-2p_n}-\mathsf{E}\left\{S_{i-2p_n}\|\mathcal{M}_{-\infty}^{i-p_n}\right\}\right|\geqslant\right.$$
$$\left.\geqslant\frac{1}{2}\lambda\sqrt{n}\right\}+\beta_n\left(\frac{1}{2}\lambda\right),$$

так что согласно (21.26)

$$\mathsf{P}\{|U_{ni}-S_i|\geqslant\lambda\sqrt{n}\}\leqslant\frac{4\mu^2(p_n)}{\lambda^2n}+\beta_n\left(\frac{1}{2}\lambda\right) \quad (21.37)$$

для всех i. Аналогично,

$$\mathsf{P}\{|V_{ni}-(S_n-S_i)|\geqslant\lambda\sqrt{n}\}\leqslant\frac{4\mu^2(p_n)}{\lambda^2n}+\beta_n\left(\frac{1}{2}\lambda\right). \quad (21.38)$$

Из соотношений (21.14), (21.22) и теоремы 5.4 следует, что величины S_n^2/n равномерно интегрируемы. Следовательно, существует λ, $\lambda>\sigma$, такое, что

$$\mathsf{P}\{|S_k|\geqslant\lambda\sqrt{k}\}\leqslant\frac{\varepsilon}{\lambda^2} \quad (21.39)$$

для всех k. Зафиксируем такое λ. Согласно (21.37)

$$\mathsf{P}\{\max_{i\leqslant n}|S_i|\geqslant 6\lambda\sqrt{n}\}\leqslant\mathsf{P}\{\max_{i\leqslant n}|U_{ni}|\geqslant 5\lambda\sqrt{n}\}+$$
$$+4\frac{\mu^2(p_n)}{\lambda^2}+n\beta_n\left(\frac{1}{2}\lambda\right). \quad (21.40)$$

Рассмотрим множества

$$E_i=\{\max_{j\leqslant i}|U_{nj}|<5\lambda\sqrt{n}\leqslant|U_{ni}|\}.$$

Имеем

$$\mathsf{P}\{\max_{i\leqslant n}|U_{ni}|\geqslant 5\lambda\sqrt{n}\}\leqslant\mathsf{P}\{|S_n|\geqslant\lambda\sqrt{n}\}+$$
$$+\sum_{i=1}^{n-1}\mathsf{P}(E_i\cap\{|S_n-U_{ni}|\geqslant 4\lambda\sqrt{n}\}). \quad (21.41)$$

Так как

$$|S_n-U_{ni}|\leqslant|S_n-S_i-V_{ni}|+|V_{ni}|+|S_i-U_{ni}|,$$

то i-е слагаемое в (21.41) не превосходит

$$\mathsf{P}\{|S_n - S_i - V_{ni}| \geqslant \lambda\sqrt{n}\} + \mathsf{P}(E_i \cap \{|V_{ni}| \geqslant 2\lambda\sqrt{n}\}) + \\ + \mathsf{P}\{|S_i - U_{ni}| \geqslant \lambda\sqrt{n}\}. \qquad (21.42)$$

Поскольку E_i принадлежит $\mathscr{M}_{-\infty}^{i-p}$ и V_{ni} измеримо относительно $\mathscr{M}_{i+p}^{\infty}$,

$$\mathsf{P}(E_i \cap \{|V_{ni}| \geqslant 2\lambda\sqrt{n}\}) \leqslant \\ \leqslant \mathsf{P}(E_i)\mathsf{P}\{|V_{ni}| \geqslant 2\lambda\sqrt{n}\} + \mathsf{P}(E_i)\varphi(2p_n).$$

Согласно (21.38) $\mathsf{P}\{|V_{ni}| \geqslant 2\lambda\sqrt{n}\} \leqslant \mathsf{P}\{|S_{n-i}| \geqslant \lambda\sqrt{n-i}\} + 4\mu^2(p_n)/\lambda^2 n + \beta_n\left(\frac{1}{2}\lambda\right)$, так что

$$\mathsf{P}(E_i \cap \{|V_{ni}| \geqslant 2\lambda\sqrt{n}\}) \leqslant \frac{4\mu^2(p_n)}{\lambda^2 n} + \beta_n\left(\frac{1}{2}\lambda\right) + \\ + \mathsf{P}(E_i)\mathsf{P}\{|S_{n-i}| \geqslant \lambda\sqrt{n-i}\} + \mathsf{P}(E_i)\varphi(2p_n).$$

Комбинируя эту оценку второго слагаемого в сумме (21.42) с оценками (21.37) и (21.38) для остальных двух слагаемых, мы видим, что i-е слагаемое в (21.41) не превосходит

$$\mathsf{P}(E_i)\mathsf{P}\{|S_{n-i}| \geqslant \lambda\sqrt{n-i}\} + \mathsf{P}(E_i)\varphi(2p_n) + \\ + \frac{12\mu^2(p_n)}{\lambda^2 n} + 3\beta_n\left(\frac{1}{2}\lambda\right).$$

Так как множества E_i не пересекаются, то согласно (21.39) имеем

$$\mathsf{P}\{\max_{i \leqslant n} |U_{ni}| \geqslant 5\lambda\sqrt{n}\} \leqslant \\ \leqslant \frac{2\varepsilon}{\lambda^2} + \varphi(2p_n) + \frac{12\mu^2(p_n)}{\lambda^2} + 3n\beta_n\left(\frac{1}{2}\lambda\right).$$

Здесь три последних члена в правой части стремятся к 0 (см. (21.25) и (21.36)), и поэтому из (21.40) следует, что

$$\mathsf{P}\{\max_{i \leqslant n} |S_i| \geqslant 6\lambda\sqrt{n}\} \leqslant \frac{3\varepsilon}{\lambda^2}$$

для всех достаточно больших n, что и доказывает (21.34).

Теорема 21.1, таким образом, доказана. В частности, мы получаем центральную предельную теорему (21.14), справедливую даже при $\sigma^2 = 0$. Эта центральная предельная теорема переносится и на многомерный случай. Именно, пусть

$$\eta_n = (\eta_n^{(1)}, \ldots, \eta_n^{(r)}), \tag{21.43}$$

где

$$\eta_n^{(i)} = f^{(i)}(\ldots, \xi_{n-1}, \xi_n, \xi_{n+1}, \ldots).$$

Предположим, что величины $\eta_n^{(i)}$ имеют нулевое среднее значение и конечную дисперсию. Предположим далее, что эти величины аппроксимируются величинами

$$\eta_{ln}^{(i)} = f_l^{(i)}(\xi_{n-l}, \ldots, \xi_n, \ldots, \xi_{n+l}).$$

Если $\sum_l \nu_l^{1/2}(i) < \infty$ для каждого i, где

$$\nu_l(i) = \mathsf{E}\{|\eta_0^{(i)} - \eta_{l0}^{(i)}|^2\},$$

то в силу неравенства

$$\mathsf{E}^{1/2}\left\{\left|\sum_{i=1}^r a_i\eta_0^{(i)} - \sum_{i=1}^r a_i\eta_{l0}^{(i)}\right|^2\right\} \leqslant \sum_{i=1}^r |a_i|\,\nu_l^{1/2}(i),$$

любая линейная комбинация компонент вектора η_n удовлетворяет условиям теоремы 21.1, так что применим метод Крамера — Уолда. Таким образом, вектор $n^{-1/2}\sum_{k=1}^n \eta_k$ оказывается асимптотически нормальным. Ковариации предельного нормального распределения равны

$$\mathsf{E}\{\eta_0^{(i)}\eta_0^{(j)}\} + \sum_{k=1}^\infty \mathsf{E}\{\eta_0^{(i)}\eta_k^{(j)}\} + \sum_{k=1}^\infty \mathsf{E}\{\eta_k^{(i)}\eta_0^{(j)}\}, \tag{21.44}$$

где ряды сходятся абсолютно.

Приложения. Пример 1. Пусть величины ξ_n независимы и одинаково распределены со средним 0 и дисперсией 1. Положим

$$\eta_n = \sum_{i=-\infty}^{\infty} a_i\xi_{n+i},$$

где a_i таковы, что $\sum_{i=1}^\infty a_i^2 < \infty$. Случайная величина (21.7) в данном случае имеет вид

$$\eta_{ln} = \sum_{i=-l}^{l} a_i\xi_{n+i},$$

а требование $\sum_l \nu_l^{1/2} < \infty$ сводится к требованию

$$\sum_{l=1}^{\infty}\Big(\sum_{|i|>l} a_i^2\Big)^{1/2} < \infty.$$

Пример 2. Пусть $(\Omega, \mathscr{B}, \mathsf{P})$, как и в примере 3 предыдущего параграфа (стр. 235), означает единичный интервал с лебеговой мерой, и пусть $\xi_n(\omega)$ — n-й знак в двоичном разложении ω $(n \geqslant 1)$. Тогда последовательность $\{\xi_1, \xi_2, \ldots\}$ стационарна и независима. Случайную величину η_n вида (21.4) можно рассматривать как функцию $\eta_n(\omega) = f\ (T^{n-1}\omega)$, где f — функция на единичном интервале, а T — преобразование $T\omega = 2\omega \pmod 1$. Предположим, что интеграл $\int_0^1 f^2(\omega)\,d\omega$ конечен и ряд $\sum_l \nu_l^{1/2}$ сходится, где

$$\nu_l = \int_0^1 |\, f(\omega) - f_l(\omega)\,|^2\, d\omega,$$

а f_l для каждого l есть функция лишь от первых l знаков $\xi_1(\omega), \ldots, \xi_l(\omega)$ в двоичном разложении ω. Тогда из теоремы 21.1 следует, что

$$\frac{1}{\sqrt{n}}\Big(\sum_{l=1}^{n} f(T^l\omega) - \mu n\Big) \xrightarrow{\mathscr{D}} N(0, \sigma^2), \tag{21.45}$$

где $\mu = \int_0^1 f(\omega)\,d\omega$, а σ^2 — предельная дисперсия. Это и есть соответствующая функциональная центральная предельная теорема. Если вообще существуют указанные выше функции f_l, то их можно взять равными

$$f_l(\omega) = 2^l \int_{(i-1)/2^l}^{i/2^l} f(s)\,ds, \qquad \omega \in \Big(\frac{i-1}{2^l}, \frac{i}{2^l}\Big] \tag{21.46}$$

(см. (21.7)).

Если $f(\omega) = I_{(0,t)}(\omega)$ и f_l определена формулой (21.46), то $\nu_l \leqslant 2^{-l}$, так что (21.45) выполняется (с $\mu = t$).

В этом случае $\sum_{i=1}^{n} f(T^i\omega)$ — количество тех i, $1 \leqslant i \leqslant n$, для которых $T^{i-1}\omega < t$. Если t — двоично-рациональное число, то f совпадает с f_l при некотором l и (21.45) вытекает из центральной предельной теоремы для $l-1$ зависимых случайных величин. Если t не является двоично-рациональным, то $f(\omega)$ зависит от всего двоичного разложения ω.

Если f непрерывна и f_l определена формулой (21.46), то $v_l \leqslant 4w_f^2(2^{-l})$, где $w_f(\delta)$ есть модуль непрерывности f. Следовательно, (21.45) выполняется, если ряд $\sum_l w_f(2^{-l})$ сходится. Последнее условие выполнено, если, например, f удовлетворяет равномерному условию Гёльдера с некоторым положительным показателем

$$|f(\omega) - f(\omega')| \leqslant K|\omega - \omega'|^{\theta}, \quad K,\ \theta > 0. \tag{21.47}$$

Можно, например, взять $f(\omega) = \omega$. В этом случае $\eta_n(\omega) = T^n(\omega)$. (Можно проверить, что последовательность $\{\eta_n\}$ не удовлетворяет условию φ-перемешивания, какова бы ни была стремящаяся к 0 функция φ, поскольку знать число $T^n\omega$ и знать последовательность $T^{n+1}\omega$, $T^{n+2}\omega$, ... — это одно и то же.)

Пример 3. Рассмотрим пространство $(\Omega, \mathscr{B}, \mathsf{P})$ и последовательность $\{a_1(\omega), a_2(\omega), \ldots\}$ из примера 4 предыдущего параграфа. Если обозначить T_1 преобразование

$$T_1\omega = \frac{1}{\omega} - \left[\frac{1}{\omega}\right] \tag{21.48}$$

и если $a(\omega) = [1/\omega]$, то $a_n(\omega) = a(T_1^{n-1}\omega)$, $n \geqslant 1$ *). Как и в предыдущем примере, (21.4) можно рассматривать как функцию $\eta_n(\omega) = f(T_1^{n-1}\omega)$, где функция f определена на единичном интервале. Соотношение

$$\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} f(T_1^i\omega) - \mu n\right) \xrightarrow{\mathscr{D}} N(0, \sigma^2)$$

*) В связи с этим и другими используемыми здесь фактами о T_1 см. § 4 в книге Биллингсли (1969).

имеет место, если

$$\mu = \frac{1}{\log 2}\int_0^1 \frac{f(\omega)}{1+\omega}\,d\omega,$$

если

$$\frac{1}{\log 2}\int_0^1 \frac{f^2(\omega)}{1+\omega}\,d\omega < \infty \tag{21.49}$$

и если ряд $\sum_l \nu_l^{1/2}$ сходится. Здесь

$$\nu_l = \frac{1}{\log 2}\int_0^1 \frac{(f(\omega) - f_l(\omega))^2}{1+\omega}\,d\omega \tag{21.50}$$

и при каждом l f_l есть функция лишь от $a_1(\omega), \ldots$ $\ldots, a_l(\omega)$. Поскольку $1 \leqslant 1+\omega \leqslant 2$, то (21.49) равносильно неравенству

$$\int_0^1 f^2(\omega)\,d\omega < \infty$$

и требование $\sum_l \nu_l^{1/2} < \infty$ от величин (21.50) равносильно такому же требованию от величин

$$\nu_l = \int_0^1 |f(\omega) - f_l(\omega)|^2\,d\omega.$$

Так как множество $\{\omega\colon a_i(\omega) = a_i,\ 1 \leqslant i \leqslant l\}$ есть интервал длины, не превосходящей 2^{-l+1}, то так же, как и выше в примере 2, в качестве $f(\omega)$ можно взять $I_{(0,t)}(\omega)$ или любую функцию, удовлетворяющую условию Гёльдера (21.47). Рассмотрим еще один пример. Поскольку *)

$$\left|\log\omega - \log\frac{p_l(\omega)}{q_l(\omega)}\right| \leqslant \frac{1}{2^{l-2}},$$

где $p_l(\omega)/q_l(\omega)$ — l-я подходящая дробь, то можно взять $f(\omega) = -\log\omega$. Для этой функции $\mu = \pi^2/(12\log 2)$. Более того,

$$\log q_n(\omega) = -\sum_{k=0}^{n-1}\log T_1^k\omega + 4\theta, \quad |\theta| \leqslant 1, \tag{21.51}$$

*) Биллингсли (1969, стр. 54).

и поэтому

$$\frac{1}{\sqrt{n}}\left(\log q_n(\omega) - \frac{n\pi^2}{12\log 2}\right) \xrightarrow{\mathscr{D}} N(0,\ \sigma^2) \qquad (21.52)$$

для соответствующего положительного σ^2.

Легко можно привести и функциональные варианты этих предельных теорем.

Диофантово приближение. Полученные результаты приводят к предельным теоремам, связанным с диофантовым приближением. Дробь p/q является наилучшим приближением числа ω, если форма

$$|q' \cdot \omega - p'|, \qquad (21.53)$$

где $0 < q' \leqslant q$, достигает минимума при $p' = p$, $q' = q$. Последовательные наилучшие приближения совпадают *) с подходящими дробями $p_n(\omega)/q_n(\omega)$, $n = 1, 2, \ldots$, и поэтому значение формы (21.53) для n-го наилучшего приближения равно

$$d_n(\omega) = |q_n(\omega) \cdot \omega - p_n(\omega)|. \qquad (21.54)$$

Поскольку **)

$$|-\log d_n(\omega) - \log q_{n+1}(\omega)| \leqslant \log 2, \qquad (21.55)$$

то из (21.52) следует, что

$$\frac{1}{\sqrt{n}}\left(-\log d_n(\omega) - \frac{n\pi^2}{12\log 2}\right) \xrightarrow{\mathscr{D}} N(0,\ \sigma^2). \qquad (21.56)$$

Эта предельная теорема имеет и функциональную форму. Из (21.51) и теоремы 4.1 следует, что если Y_n есть случайный элемент из D, определенный соотношением

$$Y_n(t,\ \omega) = \frac{1}{\sigma\sqrt{n}}\left(\log q_{[nt]}(\omega) - \frac{[nt]\,\pi^2}{12\log 2}\right)$$

(с тем же значением σ, что в (21.52)), то $Y_n \xrightarrow{\mathscr{D}} W$. Положим

$$Z_n(t,\ \omega) = \frac{1}{\sigma\sqrt{n}}\left(-\log d_{[nt]}(\omega) - \frac{[nt]\,\pi^2}{12\log 2}\right).$$

*) Хинчин (1961).

**) См. (4.6) на стр. 54 в книге Биллингсли (1969).

Из соотношения (21.55) теперь вытекает функциональная центральная предельная теорема, отвечающая (21.56)

$$Z_n \xrightarrow{\mathscr{D}} W. \tag{21.57}$$

Из (21.57) можно вывести закон арксинуса для наилучших приближений. В силу (21.56) имеем $k^{-1}\log d_k \xrightarrow{P}$ $\xrightarrow{P} \pi^2/(12\log 2)$ (можно показать, что сходимость здесь с вероятностью 1), так что мера расхождения $d_k(\omega)$ имеет нормальный порядок $e^{-k\pi^2/(12\log 2)}$. Условимся называть k-е наилучшее приближение $p_k(\omega)/q_k(\omega)$ «верхним», если

$$d_k(\omega) < e^{-k\pi^2/(12\log 2)},$$

и «нижним» в противном случае. Если $\theta_n(\omega)$ есть доля верхних приближений среди подходящих дробей

$$\frac{p_1(\omega)}{q_1(\omega)}, \ldots, \frac{p_n(\omega)}{q_n(\omega)}, \tag{21.58}$$

то согласно (21.57) и (11.26)

$$\mathsf{P}\{\theta_n \leqslant \alpha\} \to \frac{2}{\pi}\arcsin\sqrt{\alpha}, \quad 0 < \alpha < 1.$$

Например, если n велико и $p_n(\omega)/q_n(\omega)$ — верхнее приближение, то шансы на то, что более 85% подходящих дробей (21.58) также являются верхними, весьма невелики.

Нестационарность. Положим

$$X'_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} \mathsf{E}\left\{S_{[nt]} \| \mathscr{M}_{p_n}^{\infty}\right\},$$

где $p_n \to \infty$, но $p_n/\sqrt{n} \to 0$, и пусть случайная функция X_n определена соотношением (21.10). Если положить

$$\delta_n = \sup_t | X_n(t) - X'_n(t) |,$$

то, очевидно,

$$\delta_n \leqslant \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^{n} \left|\eta_j - \mathsf{E}\left\{\eta_j \| \mathscr{M}_{p_n}^{\infty}\right\}\right|.$$

В силу неравенства Минковского и стационарности

$$\mathsf{E}^{1/2}\{\delta_n^2\} \leqslant \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^{p_n} \mathsf{E}^{1/2}\left\{\left|\eta_j - \mathsf{E}\{\eta_j \| \mathscr{M}_{-p_n}^{\infty}\}\right|^2\right\} + \\ + \frac{1}{\sigma\sqrt{n}} \sum_{l=1}^{\infty} \mathsf{E}^{1/2}\left\{\left|\eta_0 - \mathsf{E}\{\eta_0 \| \mathscr{M}_{-l}^{\infty}\}\right|^2\right\}.$$

Если $\sum_l \nu_l^{1/2} < \infty$, то согласно лемме 1 и выбору p_n $\mathsf{E}\{\delta_n^2\} \to 0$. Следовательно, рассуждая так же, как и выше, можно распространить теоремы 20.2 и 20.3 на процессы $\{\eta_n\}$. Отсюда, например, следует, что в приложениях к диофантовым приближениям гауссову меру можно заменить лебеговой.

Теорема 10.4, однако, не переносится. Чтобы убедиться в этом, рассмотрим процесс $\{\xi_n\}$ из примера 2 настоящего параграфа (стр. 264) и положим $\eta_n(\omega) = \sum_{k=n}^{\infty} \xi_k(\omega)/2^{l-n+1}$. Хотя сумма $\sum_{l=1}^{n} \eta_l(\omega)$ асимптотически нормальна (при подходящей нормировке), это перестает быть верным, если рассматривать условные вероятности относительно $\eta_1(\omega)$. В самом деле, поскольку $\eta_1(\omega) = \omega$, то если $\eta_1(\omega) = \alpha$, $0 < \alpha < 1$, то условное распределение суммы $\sum_{i=1}^{n} \eta_l(\omega)$ сосредоточено в точке $\sum_{i=1}^{n} \eta_i(\alpha)$.

Примечание. Результаты этого и предыдущего параграфов являются новыми и обобщают результаты автора (1956 и 1962). Центральные предельные теоремы при выполнении условий теорем 20.1 и 21.1 были доказаны Ибрагимовым (1962); см. его статью также по поводу ссылок на более раннюю литературу (см. также Дёблин (1940), Форте (1940) и Кац (1946)).

§ 22. Эмпирические функции распределения

Мы изучим теперь эмпирические функции распределения для последовательностей $\{\xi_n\}$ и $\{\eta_n\}$ типа тех, которые рассматривались в предыдущих двух параграфах.

Процессы с φ-перемешиванием. На протяжении настоящего параграфа предполагается, что $\{\xi_n\}$ — стационарный процесс с φ-перемешиванием. Нам потребуется следующая оценка, связанная с леммой 4 § 20.

Лемма 1. *Пусть* $|\xi_0| \leqslant 1$ *с вероятностью* 1, $\mathsf{E}\{\xi_0\}=0$ *и* $\sum_k k^2\varphi_k^{1/2} < \infty$. *Тогда*

$$\mathsf{E}\{S_n^4\} \leqslant 288\left[n^2\mathsf{E}^2\{\xi_0^2\} + n\mathsf{E}\{\xi_0^2\}\right]\left[\sum_{k=0}^{\infty}(k+1)^2\varphi_k^{1/2}\right]. \quad (22.1)$$

Доказательство. Имеем

$$\mathsf{E}\{S_n^4\} \leqslant 4!\, n \sum |\mathsf{E}\{\xi_0\xi_i\xi_{i+j}\xi_{i+j+k}\}|, \qquad (22.2)$$

где суммирование производится по i, j, k таким, что

$$i,\ j,\ k \geqslant 0, \quad i+j+k \leqslant n. \qquad (22.3)$$

Обозначим для краткости $\mathsf{E}\{\xi_0^2\}$ через p. Сначала докажем следующие три неравенства:

$$|\mathsf{E}\{\xi_0(\xi_i\xi_{i+j}\xi_{i+j+k})\}| \leqslant 2\varphi_i^{1/2}p, \qquad (22.4)$$

$$|\mathsf{E}\{(\xi_0\xi_i\xi_{i+j})\xi_{i+j+k}\}| \leqslant 2\varphi_k^{1/2}p, \qquad (22.5)$$

$$|\mathsf{E}\{(\xi_0\xi_i)(\xi_{i+j}\xi_{i+j+k})\}| \leqslant 4\varphi_i^{1/2}\varphi_k^{1/2}p^2 + 2\varphi_j^{1/2}p. \quad (22.6)$$

По лемме 1 § 20 левая часть (22.4) не превосходит

$$2\varphi^{1/2}\mathsf{E}^{1/2}\{\xi_0^2\}\,\mathsf{E}^{1/2}\{\xi_i^2\xi_{i+j}^2\xi_{i+j+k}^2\}.$$

Отсюда, поскольку $|\xi_n| \leqslant 1$, получаем (22.4). Аналогично устанавливается (22.5).

Применив еще раз лемму 1 § 20, получим, что левая часть (22.6) не превосходит

$$|\mathsf{E}\{\xi_0\xi_i\}\,\mathsf{E}\{\xi_0\xi_k\}| + 2\varphi_j^{1/2}\mathsf{E}^{1/2}\{\xi_0^2\xi_i^2\}\,\mathsf{E}^{1/2}\{\xi_0^2\xi_k^2\}. \quad (22.7)$$

Еще дважды применяя ту же лемму, имеем

$$|\mathsf{E}\{\xi_0\xi_i\}| \leqslant 2\varphi_i^{1/2}p, \quad |\mathsf{E}\{\xi_0\xi_k\}| \leqslant 2\varphi_k^{1/2}p.$$

Так как $|\xi_n| \leqslant 1$, то (22.7) не превосходит $4\varphi_i^{1/2}\varphi_k^{1/2}p^2 + 2\varphi_j^{1/2}p$, и мы получаем (22.6).

Из (22.2) и трех доказанных неравенств теперь вытекает, что

$$\mathsf{E}\{S_n^4\} \leqslant 4!\,4n\left(p^2\sum_{i,\,k\leqslant j}\varphi_i^{1/2}\varphi_k^{1/2} + 3p\sum_{j,\,k\leqslant i}\varphi_i^{1/2}\right),$$

где индексы суммирования удовлетворяют также условию (22.3). Так как

$$\sum_{i,k\leqslant j}\varphi_i^{1/2}\varphi_k^{1/2}\leqslant\sum_{j=0}^{n}\sum_{i,k=0}^{\infty}\varphi_i^{1/2}\varphi_k^{1/2}\leqslant 2n\left[\sum_{k=0}^{\infty}\varphi_k^{1/2}\right]^2$$

и

$$\sum_{j,k\leqslant i}\varphi_i^{1/2}\leqslant\sum_{i=0}^{n}\sum_{j,k=0}^{i}\varphi_i^{1/2}\leqslant\sum_{k=0}^{\infty}(k+1)^2\varphi_k^{1/2},$$

то неравенство (22.1) доказано.

Предположим теперь, что

$$0\leqslant\xi_n(\omega)\leqslant 1.$$

Пусть $F_n(t,\omega)$ — эмпирическая функция распределения для $\xi_1(\omega), \ldots, \xi_n(\omega)$. Определим Y_n соотношением

$$Y_n(t,\omega)=\sqrt{n}\,(F_n(t,\omega)-F(t)), \tag{22.8}$$

где F — функция распределения для ξ_0. Положим

$$g_t(\alpha)=I_{[0,t]}(\alpha)-F(t). \tag{22.9}$$

Теорема 22.1. *Пусть $\{\xi_n\}$ — процесс с φ-перемешиванием, причем $\sum\limits_n n^2\varphi_n^{1/2}<\infty$, и функция распределения F величины ξ_0 непрерывна на* $[0, 1]$. *Тогда*

$$Y_n\xrightarrow{\mathscr{D}}Y, \tag{22.10}$$

где Y_n — случайная функция (22.8), *а Y — гауссовская случайная функция, для которой*

$$\mathsf{E}\{Y(t)\}=0 \tag{22.11}$$

и

$$\mathsf{E}\{Y(s)Y(t)\}=\mathsf{E}\{g_s(\xi_0)g_t(\xi_0)\}+$$
$$+\sum_{k=1}^{\infty}\mathsf{E}\{g_s(\xi_0)g_t(\xi_k)\}+\sum_{k=1}^{\infty}\mathsf{E}\{g_s(\xi_k)g_t(\xi_0)\}. \tag{22.12}$$

Ряды в (22.12) *сходятся абсолютно и* $\mathsf{P}\{Y\in C\}=1$.

Доказательство. Покажем сначала, что можно ограничиться случаем, когда величина ξ_0 распределена равномерно на [0, 1]. Процесс $\{F(\xi_n)\}$ является процессом с φ-перемешиванием, и, поскольку функция F непрерывна, величина $F(\xi_0)$ распределена равномерно. Пусть

$F'_n(t, \omega)$ — эмпирическая функция распределения для $F(\xi_1(\omega)), \ldots, F(\xi_n(\omega))$ и

$$Y'_n(t) = \sqrt{n}\,(F'_n(t, \omega) - t).$$

С вероятностью 1 имеем

$$Y'_n(F(t)) = Y_n(t)$$

для всех t. Если теорема верна в случае равномерного распределения ξ_0, то $Y'_n \xrightarrow{\mathcal{D}} Y'$, где Y' — гауссовская случайная функция, для которой $\mathsf{E}\{Y'(t)\} = 0$,

$$\mathsf{E}\{Y'(s)\,Y'(t)\} = \mathsf{E}\{g'_s(F(\xi_0))\,g'_t(F(\xi_0))\} +$$

$$+ \sum_{k=1}^{\infty} \mathsf{E}\{g'_s(F(\xi_0))\,g'_t(F(\xi_k))\} + \sum_{k=1}^{\infty} \mathsf{E}\{g'_s(F(\xi_k))\,g'_t(F(\xi_0))\}$$

(здесь $g'_t(\alpha) = I_{[0,\,t]}(\alpha) - t$) и $\mathsf{P}\{Y' \in C\} = 1$. Пусть h — отображение D в себя, определенное равенством $(hx)(t) = x(F(t))$, и пусть $Y = h(Y')$. Так как $C \subset h^{-1}C$, то $\mathsf{P}\{Y \in C\} = 1$. Далее, Y — гауссовская функция, удовлетворяющая соотношениям (22.11) и (22.12). Поскольку отображение h непрерывно на C, то из следствия 1 к теореме 5.1 вытекает, что $Y'_n \xrightarrow{\mathcal{D}} Y'$ влечет $Y_n \xrightarrow{\mathcal{D}} Y$. Таким образом, достаточно рассмотреть лишь равномерный случай.

Предположим теперь, что величина ξ_0 распределена равномерно на $[0, 1]$, и положим в (22.9) $F(t) = t$. Так как $\{g_t(\xi_n)\}$ есть процесс с φ-перемешиванием и

$$Y_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} g_t(\xi_i),$$

то согласно теореме 20.1 величина $Y_n(t)$ асимптотически нормальна. Из многомерного варианта этого результата следует, что для каждого набора $t_1, \ldots, t_k$ распределение случайного вектора $(Y_n(t_1), \ldots, Y_n(t_k))$ стремится к центрированному нормальному распределению. В силу (20.59) ковариации этого предельного нормального распределения совпадают с соответствующими ковариациями (22.12) функции Y. Из теоремы 20.1 следует также, что ряды (22.12) абсолютно сходятся.

Мы докажем, что для любых положительных ε и η существует δ, $0 < \delta < 1$, такое, что

$$\mathsf{P}\{w(Y_n, \delta) \geqslant \varepsilon\} \leqslant \eta \tag{22.13}$$

при всех достаточно больших n. Из теоремы 15.5 будет тогда следовать, что последовательность $\{Y_n\}$ плотна, и если Y — предел по распределению некоторой подпоследовательности $\{Y_{n'}\}$, то $\mathsf{P}\{Y \in C\} = 1$, $Y_n \xrightarrow{\mathscr{D}} Y$ и Y есть гауссовская случайная функция, удовлетворяющая соотношениям (22.11) и (22.12). Тем самым теорема будет доказана.

Зафиксируем ε и η. Так как величина ξ_0 распределена равномерно, то

$$\mathsf{E}\{|g_t(\xi_0) - g_s(\xi_0)|^2\} \leqslant |t - s|.$$

Применяя лемму 1 к $\{g_t(\xi_n) - g_s(\xi_n)\}$, получим

$$\mathsf{E}\left\{\left|\sum_{i=1}^{n}(g_t(\xi_i) - g_s(\xi_i))\right|^4\right\} \leqslant K_1(n^2|t-s|^2 + n|t-s|),$$

где K_1 зависит лишь от φ. Следовательно, если

$$\frac{\varepsilon}{n} \leqslant t - s, \tag{22.14}$$

то (считая $\varepsilon < 1$)

$$\mathsf{E}\{|Y_n(t) - Y_n(s)|^4\} \leqslant \frac{2K_1}{\varepsilon}(t-s)^2. \tag{22.15}$$

Пусть теперь p — число, для которого $\varepsilon/n < p$, и рассмотрим случайные величины

$$Y_n(s + ip) - Y_n(s + (i-1)p), \qquad i = 1, 2, \ldots, m,$$

где m — целое положительное число. В силу (22.15) и теоремы 12.2 имеем

$$\mathsf{P}\{\max_{i \leqslant m}|Y_n(s + ip) - Y_n(s)| \geqslant \lambda\} \leqslant \frac{K_2}{\varepsilon\lambda^4}m^2p^2, \tag{22.16}$$

где $K_2 = 2K'_{4,2}K_1$ зависит лишь от φ.

Покажем теперь, что

$$|Y_n(t) - Y_n(s)| \leqslant |Y_n(s+p) - Y_n(s)| + p\sqrt{n}, \tag{22.17}$$
$$s \leqslant t \leqslant s + p.$$

Полагая для простоты записи $s=0$, мы видим, что (22.17) равносильно неравенству

$$|U_n(t)-nt|\leqslant|U_n(p)-np|+np, \qquad 0\leqslant t\leqslant p,$$

где $U_n(t)$ есть число тех ξ_i в множестве $\xi_1, \ldots, \xi_n$, для которых $\xi_i\leqslant t$. Но

$$U_n(t)-nt\leqslant U_n(p)-nt=U_n(p)-np+n(p-t)\leqslant$$
$$\leqslant|U_n(p)-np|+np \text{ и } nt-U_n(t)\leqslant nt\leqslant|U_n(p)-np|+np.$$

Из (22.17) вытекает теперь, что

$$\sup_{s\leqslant t\leqslant s+mp}|Y_n(t)-Y_n(s)|\leqslant$$
$$\leqslant 3\max_{i\leqslant m}|Y_n(s+ip)-Y_n(s)|+p\sqrt{n}. \tag{22.18}$$

Если

$$\frac{\varepsilon}{n}\leqslant p<\frac{\varepsilon}{\sqrt{n}}, \tag{22.19}$$

то применимо неравенство (22.16) и из (22.18) мы получаем

$$\mathsf{P}\{\sup_{s\leqslant t\leqslant s+mp}|Y_n(t)-Y_n(s)|\geqslant 4\varepsilon\}\leqslant\frac{K_2}{\varepsilon^5}m^2p^2. \tag{22.20}$$

Выберем δ так, чтобы $K_2\delta/\varepsilon^5<\eta$. Из (22.20) следует, что

$$\mathsf{P}\{\sup_{s\leqslant t\leqslant s+\delta}|Y_n(t)-Y_n(s)|\geqslant 4\varepsilon\}<\eta\delta, \tag{22.21}$$

если только существует p и целое число m такие, что верно (22.19) и $mp=\delta$. Но это равносильно существованию целого числа m, для которого $(\delta/\varepsilon)\sqrt{n}<m\leqslant\leqslant(\delta/\varepsilon)n$, что выполняется при всех достаточно больших n.

Следовательно, для заданных ε и η существует такое δ, что (22.21) выполняется (при всех $s\leqslant 1$ и с 1 вместо $s+\delta$, если $s+\delta>1$) для всех достаточно больших n. Но отсюда вытекает (см. следствие к теореме 8.3), что для больших n

$$\mathsf{P}\{w(Y_n, \delta)\geqslant 12\varepsilon\}\leqslant\eta,$$

что совпадает с (22.13) с точностью до множителя 12.

Таким образом, $Y_n\xrightarrow{\mathcal{D}}Y$. Предположение о непрерывности F служит лишь для упрощения доказательства.

Его можно избежать, если использовать теорему 12.5 там, где мы использовали теорему 12.2. Разумеется, если функция F разрывна, то случайная функция Y не будет принадлежать C с вероятностью 1.

Функции от процессов с φ-перемешиванием. Как и в предыдущем параграфе, рассмотрим функцию

$$\eta_n = f(\ldots, \xi_{n-1}, \xi_n, \xi_{n+1}, \ldots) \tag{22.22}$$

от процесса с φ-перемешиванием. Предположим, как и прежде, что η_n хорошо аппроксимируется функциями

$$\eta_{ln} = f_l(\xi_{n-l}, \ldots, \xi_n, \ldots, \xi_{n+l}), \tag{22.23}$$

зависящими лишь от конечного числа ξ_i. Отметим, что вместо двусторонних можно было бы рассматривать односторонние последовательности.

Предположим, что

$$0 \leqslant \eta_n(\omega) \leqslant 1,$$

и определим Y_n равенством (22.8), где теперь F — функция распределения величины η_0, а $F_n(t, \omega)$ — эмпирическая функция распределения для $\eta_1(\omega), \ldots, \eta_n(\omega)$. При анализе асимптотического распределения Y_n, помимо ограничения на величину φ_n, нам понадобится некоторое условие, выполнение которого должно обеспечить совпадение эмпирического распределения $F_{ln}(t, \omega)$ величин $\eta_{l1}(\omega), \ldots, \eta_{ln}(\omega)$ с $F_n(t, \omega)$ в точках t множества H_l, которое быстро становится плотным в $[0, 1]$ при $l \to \infty$.

Предположим прежде всего, что

$$0 \leqslant \eta_{ln}(\omega) \leqslant 1.$$

Потребуем, далее, чтобы существовали множества H_l в $[0, 1]$, удовлетворяющие следующим трем свойствам.

(i) Если $t \in H_l$, то

$$I_{[0, t]}(\eta_0) = I_{[0, t]}(\eta_{l0}) \tag{22.24}$$

с вероятностью 1.

(ii) Множество

$$J_l = \{F(t) \colon t \in H_l\} \tag{22.25}$$

является ρ_l-сетью в $[0, 1]$, где ρ_l стремится к нулю с показательной скоростью: существуют положительные числа A и ρ, $\rho < 1$, такие, что

$$\rho_l \leqslant A\rho^l, \quad l = 1, 2, \ldots \tag{22.26}$$

(iii) $H_l \subset H_{l+1}$.

Из (i) следует, что если $t \in H_l$, то $F_{ln}(t, \omega) = F_n(t, \omega)$ с вероятностью 1. Пусть, как и прежде, g_t определяется соотношением (22.9).

Теорема 22.2. *Пусть $\{\xi_n\}$ — процесс с φ-перемешиванием, причем $\sum_n n^2\varphi_n^{1/2} < \infty$, η_0 имеет непрерывную функцию распределения F на $[0, 1]$ и существуют множества H_l с описанными выше свойствами. Тогда*

$$Y_n \xrightarrow{\mathscr{D}} Y,$$

где Y — гауссовская случайная величина, для которой

$$\mathsf{E}\{Y(t)\} = 0 \tag{22.27}$$

и

$$\mathsf{E}\{Y(s)Y(t)\} = \mathsf{E}\{g_s(\eta_0)\,g_t(\eta_0)\} + \\ + \sum_{k=1}^{\infty} \mathsf{E}\{g_s(\eta_0)\,g_t(\eta_k)\} + \sum_{k=1}^{\infty} \mathsf{E}\{g_s(\eta_k)\,g_t(\eta_0)\}. \tag{22.28}$$

Ряды в (22.28) сходятся абсолютно и $\mathsf{P}\{Y \in C\} = 1$.

Прежде чем перейти к доказательству теоремы, мы приведем два примера ее применения.

Пример 1. Пусть, как и в примере 2 предыдущего параграфа (стр. 264), $(\Omega, \mathscr{B}, \mathsf{P})$ — единичный интервал с лебеговой мерой, а $\xi_n(\omega)$ — n-й знак в двоичном разложении ω. Если $T\omega = 2\omega \pmod 1$ и $\eta_n(\omega) = T^{n-1}\omega$, то η_n имеет вид (22.22) (односторонний вариант). Величины η_n распределены равномерно с функцией распределения $F(t) \equiv t$.

Определим η_{ln} соотношением

$$\eta_{ln}(\omega) = \frac{i-1}{2^l}, \quad \text{если} \quad \frac{i-1}{2^l} < \eta_n(\omega) \leqslant \frac{i}{2^l}, \quad 1 \leqslant i \leqslant 2^l.$$

Величина $\eta_{l0}(\omega)$ является функцией от первых l знаков в разложении ω и, следовательно, η_{ln} имеет вид (22.23) (односторонний вариант). Если t — число вида $i/2^l$, то выполняется (22.24). Так как $F(t) \equiv t$, то можно взять

$$H_l = J_l = \left\{0, \frac{1}{2^l}, \ldots, \frac{2^l-1}{2^l}, 1\right\}.$$

Множество H_l является 2^{-l}-сетью в $[0, 1]$. Таким образом, $\{\eta_n\}$ удовлетворяет предположениям теоремы.

Пример 2. Пусть, как в примере 3 предыдущего параграфа (стр. 265), $(\Omega, \mathscr{B}, \mathsf{P})$ — единичный интервал с гауссовой мерой (20.19). Пусть $\xi_n(\omega) = a_n(\omega)$ — n-й элемент в разложении ω в непрерывную дробь, и пусть $\eta_n(\omega) = T_1^{n-1}\omega$, где T_1 — преобразование, определенное формулой (21.48). Тогда η_n имеет вид (22.22) (односторонний вариант) и ее функция распределения равна

$$F(t) = \frac{1}{\log 2}\int_0^t \frac{dx}{1+x} = \frac{\log(1+t)}{\log 2}.$$

Пусть, далее, $\eta_{ln}(\omega)$ — l-я подходящая дробь для $T_1^{n-1}\omega$:

$$\eta_{ln}(\omega) = \frac{p_l\left(T_1^{n-1}\omega\right)}{q_l\left(T_1^{n-1}\omega\right)}.$$

Величина η_{ln} имеет вид (22.23) (односторонний вариант). Пусть H_l — множество всех l-х подходящих дробей (для всех ω). Далее, единичный интервал распадается на счетное число подинтервалов*), имеющих своими концами точки из H_l. При этом $\eta_1(\omega) = \omega$ и $\eta_{l1}(\omega) = p_l(\omega)/q_l(\omega)$ лежат в одном и том же подинтервале, так что (22.24) выполняется для t из H_l. Длины этих подинтервалов не превосходят 2^{-l+1}. Поскольку производная функции F всюду не превосходит $1/\log 2$, то множество (22.25) является ρ_l-сетью в $[0, 1]$ с

$$\rho_l \leqslant \frac{2}{\log 2}\,\frac{1}{2^l}.$$

Таким образом, $\{\eta_n\}$ удовлетворяет условиям теоремы.

---

*) Фундаментальные интервалы ранга l — см. Биллингсли (1969, стр. 54).

Доказательство теоремы 22.2. Как и при доказательстве теоремы 22.1, мы сначала покажем, что достаточно рассмотреть случай равномерного распределения η_0 на [0, 1]. Положим $\eta'_n = F(\eta_n)$ и $\eta'_{ln} = F(\eta_{ln})$. Величина η'_0 распределена равномерно, и если t принадлежит множеству J_l, определенному соотношением (22.25), то

$$I_{[0,\,t]}(\eta'_0) = I_{[0,\,t]}(\eta'_{l0})$$

с вероятностью 1. Если $H'_l = J'_l = J_l$, то $\{\eta'_n\}$ и $\{\eta'_{ln}\}$ удовлетворяют условиям теоремы относительно множеств H'_l и J'_l. Рассуждая далее так же, как в начале доказательства предыдущей теоремы, получаем, что теорема справедлива, если она справедлива в равномерном случае.

Пусть теперь величина η_0 распределена равномерно. Положим в (22.9) $F(t) = t$. Предположим, что при $t \in H_l$ (22.4) имеет место с вероятностью 1. Пусть $H_l \subset \subset H_{l+1}$. Наконец, предположим, что H_l есть ρ_l-сеть в [0, 1], где ρ_l удовлетворяет условию (22.26). Не ограничивая общности, можно также считать, что 0 и 1 принадлежат H_l.

Из сделанных допущений следует, что $|\eta_0 - \eta_{l0}| \leqslant \leqslant 2\rho_l$ с вероятностью 1. Следовательно, каково бы ни было t,

$$\mathsf{P}\{\eta_0 \leqslant t < \eta_{l0}\} + \mathsf{P}\{\eta_{l0} \leqslant t < \eta_0\} \leqslant \\ \leqslant \mathsf{P}\{t - 2\rho_l < \eta_0 \leqslant t\} + \mathsf{P}\{t < \eta_0 \leqslant t + 2\rho_l\},$$

так что, поскольку величина η_0 распределена равномерно,

$$\mathsf{E}\{|g_t(\eta_0) - g_t(\eta_{l0})|^2\} \leqslant 4\rho_l.$$

Так как $g_t(\eta_{l0})$ является функцией лишь от $\xi_{-l}, \ldots, \xi_l$ и так как из (22.26) следует, что $\sum_l \rho_l^{1/2} < \infty$, то, применяя теорему 21.1, получаем, что величина

$$Y_n(t) = \frac{1}{\sqrt{n}} \sum_{l=1}^{n} g_t(\eta_l)$$

распределена асимптотически нормально. Многомерный вариант этого утверждения (см. (21.43), (21.44)) точно так же, как многомерный вариант теоремы 20.1 в предыдущем доказательстве, показывает, что конечномерные распределения Y_n сходятся к конечномерным распределениям Y и что ряды в (22.28) абсолютно сходятся.

Опять, как и в предыдущем доказательстве, теперь достаточно показать, что по заданным положительным ε и η можно выбрать δ, $0 < \delta < 1$, такое, что

$$\mathsf{P}\{w(Y_n, \delta) \geqslant \varepsilon\} \leqslant \eta \tag{22.29}$$

для всех больших n.

Если s и t принадлежат H_l, то процесс

$$g_t(\eta_n) - g_s(\eta_n) = g_t(\eta_{ln}) - g_s(\eta_{ln}), \quad n = 0, \pm 1, \pm 2, \ldots,$$

является процессом с $\varphi^{(l)}$-перемешиванием, где

$$\varphi^{(l)}(n) = \begin{cases} 1, & \text{если } n \leqslant 2l, \\ \varphi(n - 2l), & \text{если } n > 2l. \end{cases}$$

Так как величина η_0 распределена равномерно и так как

$$\sum_{k=0}^{\infty} (k+1)^2 [\varphi_k^{(l)}]^{1/2} = O(l^3),$$

то по лемме 1

$$\mathsf{E}\left\{\left|\sum_{i=1}^{n} (g_t(\eta_i) - g_s(\eta_i))\right|^4\right\} \leqslant K_3 l^6 (n^6 |t-s|^2 + n|t-s|),$$

где K_3 зависит лишь от φ. Следовательно, из

$$s, t \in H_l, \quad \frac{\varepsilon}{n} \leqslant t - s \tag{22.30}$$

вытекает (при $\varepsilon < 1$), что

$$\mathsf{P}\{|Y_n(t) - Y_n(s)| \geqslant \lambda\} \leqslant \frac{2K_3}{\varepsilon} \frac{l^6}{\lambda^4} (t-s)^2. \tag{22.31}$$

Подберем и зафиксируем m такое, что

$$\rho_{mv} \leqslant \frac{1}{2^{2v+2}}, \quad v = 1, 2, \ldots \tag{22.32}$$

(Это возможно вследствие (22.26).) Применяя индукцию, определим множества

$$L_v:\ t_0^{(v)},\ t_1^{(v)},\ \ldots,\ t_{2^v}^{(v)}$$

для $v \geqslant 0$ такие, что

$$0 = t_0^{(v)} < t_1^{(v)} < \ldots < t_{2^v}^{(v)} = 1,$$

$$t_i^{(v)} \in H_{mv},\quad t_{2i}^{(v)} = t_i^{(v-1)},\quad v \geqslant 1,$$

и

$$\frac{1}{2^{v+1}}\left(1 + \frac{1}{2^v}\right) \leqslant \alpha_v \leqslant \beta_v \leqslant \frac{1}{2^{v+1}}\left(3 - \frac{1}{2^v}\right), \tag{22.33}$$

где

$$\alpha_v = \min_{1 \leqslant i \leqslant 2^v}\left(t_i^{(v)} - t_{i-1}^{(v)}\right) \leqslant \max_{1 \leqslant i \leqslant 2^v}\left(t_i^{(v)} - t_{i-1}^{(v)}\right) = \beta_v.$$

Если $t_0^{(0)} = 0$ и $t_1^{(0)} = 1$, то L_0 обладает этими свойствами. Предположим, что $L_0, \ldots, L_v$ уже построены и имеют указанные свойства. Определим $t_{2i}^{(v+1)} = t_i^{(v)}$, и пусть $t_{2i+1}^{(v+1)}$ — точка из $H_{m(v+1)}$, удовлетворяющая условию

$$\left|t_{2i+1}^{(v+1)} - \frac{1}{2}\left(t_i^{(v)} + t_{i+1}^{(v)}\right)\right| < \rho_{m(v+1)}.$$

Такая точка существует, поскольку $H_{m(v+1)}$ является $\rho_{m(v+1)}$-сетью в отрезке $[0, 1]$. Из (22.32) следует, что (22.33) по-прежнему выполняется при замене v на $v+1$. Таким образом, множество L_{v+1} удовлетворяет нужным требованиям.

Зафиксируем число θ такое, что

$$1/2 < \theta_0 = \theta^4 < 1. \tag{22.34}$$

Предположим, что числа ε, n и v удовлетворяют условию

$$\frac{\varepsilon}{n} \leqslant \frac{1}{2^{v+1}} \leqslant \frac{1}{2^{v-1}} \leqslant \frac{\varepsilon}{\sqrt{n}}. \tag{22.35}$$

Из этих условий согласно (22.33) следует, что

$$\frac{\varepsilon}{n} \leqslant \alpha_v \leqslant \beta_v \leqslant \frac{\varepsilon}{\sqrt{n}}. \tag{22.36}$$

Предположим далее, что $0 < u < v$ и $0 \leqslant h < 2^{v-u}$. Покажем, что

$$\mathsf{P}\left\{\max_{0 \leqslant i \leqslant 2^u} \left|Y_n\left(t^{(v)}_{h2^u+i}\right) - Y_n\left(t^{(v)}_{h2^u}\right)\right| \geqslant \varepsilon\right\} \leqslant \\ \leqslant \frac{K_5}{\varepsilon^5} \theta_0^{v-u} \left[t^{(v)}_{(h+1)2^u} - t^{(v)}_{h2^u}\right], \qquad (22.37)$$

где K_5 зависит лишь от φ и фиксированных чисел m и θ из условий (22.32) и (22.34). Чтобы не загромождать обозначений, мы проведем доказательство в случае $h = 0$.

Если $0 \leqslant i \leqslant 2^u$, то

$$\left|Y_n\left(t^{(v)}_i\right)\right| \leqslant \sum_{k=0}^{u} \max_{1 \leqslant j \leqslant 2^{u-k}} \left|Y_n\left(t^{(v)}_{j2^k}\right) - Y_n\left(t^{(v)}_{(j-1)2^k}\right)\right|$$

(рассмотрите представление целого числа i в двоичной системе) и, следовательно,

$$\mathsf{P}\left\{\max_{0 \leqslant i \leqslant 2^u} \left|Y_n\left(t^{(v)}_i\right)\right| \geqslant \varepsilon\right\} \leqslant \\ \leqslant \sum_{k=0}^{u} \sum_{j=1}^{2^{u-k}} \mathsf{P}\left\{\left|Y_n\left(t^{(v)}_{j2^k}\right) - Y_n\left(t^{(v)}_{(j-1)2^k}\right)\right| \geqslant \varepsilon(1-\theta)\theta^{u-k}\right\}.$$

Далее, $t^{(v)}_{j2^k} = t^{(v-k)}_j \in H_{m(v-k)}$. Так как (22.30) влечет (22.31), то из (22.36) следует, что

$$\mathsf{P}\left\{\max_{0 \leqslant i \leqslant 2^u} \left|Y_n\left(t^{(v)}_i\right)\right| \geqslant \varepsilon\right\} \leqslant \\ \leqslant \sum_{k=0}^{u} \sum_{j=1}^{2^{u-k}} \frac{2K_3}{\varepsilon} \frac{(m(v-k))^6}{(\varepsilon(1-\theta)\theta^{u-k})^4} \left(t^{(v-k)}_j - t^{(v-k)}_{j-1}\right)^2 \leqslant \\ \leqslant \frac{K_4}{\varepsilon^5} \sum_{k=0}^{u} \frac{(v-k)^6 \beta_{v-k}}{\theta_0^{u-k}} t^{(v)}_{2^u},$$

где $K_4 = 2K_3 m^6/(1-\theta)^4$. Поскольку $\beta_{v-k} \leqslant 1/2^{v-k-1}$, то мы получаем (22.37) с постоянной

$$K_5 = 2K_4 \sum_{i=0}^{\infty} \frac{i^6}{(2\theta_0)^i}$$

(напомним, что $2\theta_0 > 1$ и что постоянная K_3 в (22.31) зависит лишь от φ).

Чисто геометрическое неравенство (22.17) применимо, как и прежде, и мы получаем

$$\sup_{t_{h2^u}^{(v)} \leqslant t \leqslant t_{(h+1)2^u}^{(v)}} \left|Y_n(t) - Y_n\left(t_{h2^u}^{(v)}\right)\right| \leqslant$$
$$\leqslant 3 \max_{0 \leqslant i \leqslant 2^u} \left|Y_n\left(t_{h2^u+i}^{(v)}\right) - Y_n\left(t_{h2^u}^{(v)}\right)\right| + \sqrt{n}\beta_v.$$

Теперь из (22.36) и (22.37) следует, что

$$\sum_{h=0}^{2^{v-u}-1} \mathsf{P}\left\{\sup_{t_{h2^u}^{(v)} \leqslant t \leqslant t_{(h+1)2^u}^{(v)}} \left|Y_n(t) - Y_n\left(t_{h2^u}^{(v)}\right)\right| > 4\varepsilon\right\} \leqslant \frac{K_5}{\varepsilon^5}\theta_0^{v-u}.$$

Так как $t_{h2^u}^{(v)} = t_h^{(v-u)} \in L_{v-u}$ и $\alpha_{v-u} \geqslant 1/2^{v-u+1}$, то, используя следствие к теореме 8.3, приходим к неравенству

$$\mathsf{P}\left\{w\left(Y_n, \frac{1}{2^{v-u+1}}\right) > 12\varepsilon\right\} \leqslant \frac{K_5}{\varepsilon^5}\theta_0^{v-u}. \tag{22.38}$$

Это неравенство справедливо при выполнении условия (22.35) и при $0 < u < v$.

Если ε и η заданы, выберем целое число r такое, что $K_5\theta_0^r/\varepsilon^5 < \eta$ (напомним, что $\theta_0 < 1$), и положим $\delta = 1/2^{r+1}$. Для всех n, превосходящих некоторое n_0, существует v, удовлетворяющее условию (22.35). Если взять $u = v - r$, то (22.38) справедливо и мы приходим к (22.29) (с 12ε вместо ε). Теорема 22.2 полностью доказана.

Нетрудно доказать, что теорема остается в силе при замене вероятностной меры P, управляющей последовательностями ξ_n и η_n, произвольной вероятностной мерой, абсолютно непрерывной относительно P. В частности, в примере 2 можно заменить гауссову меру на лебегову.

Примечание. Теоремы 22.1 и 22.2 — новые; случай, рассмотренный в примере 1, изучался Цесельским и Кестеном (1962).

§ 23. Мартингалы

В § 20 мы доказали функциональную центральную предельную теорему для стационарных процессов $\{\xi_n\}$, удовлетворяющих условию равномерного перемешивания. Это условие перемешивания можно ослабить и предполагать лишь эргодичность в том случае, когда частичные суммы $S_n = \xi_1 + \ldots + \xi_n$ образуют мартингал *).

Теорема 23.1. *Пусть $\{\xi_n\}$ — стационарный, эргодический процесс, для которого*

$$\mathsf{E}\{\xi_n \| \xi_1, \ldots, \xi_{n-1}\} = 0 \tag{23.1}$$

с вероятностью 1 и $0 < \mathsf{E}\{\xi_n^2\} = \sigma^2 < \infty$. Если положить $X_n(t, \omega) = S_{[nt]}(\omega)/\sigma\sqrt{n}$, то $X_n \xrightarrow{\mathscr{D}} W$.

Доказательство. Без ограничения общности можно рассматривать двусторонне бесконечную стационарную и эргодическую последовательность $\ldots, \xi_{-1}, \xi_0, \xi_1, \ldots$ **). Если $\mathscr{F}_k$ означает σ-поле, порожденное величинами $\ldots, \xi_{k-1}, \xi_k$, то вследствие (23.1) и стационарности,

$$\mathsf{E}\{\xi_k \| \mathscr{F}_{k-1}\} = 0 \tag{23.2}$$

с вероятностью 1.

Мы проверим выполнение условий теоремы 19.4 с функциями $\rho(t) = 0$ и $\sigma^2(t) = 1$, отвечающими броуновскому движению (см. (19.29)), откуда и будет следовать $X_n \xrightarrow{\mathscr{D}} W$. Для проверки условия 1 этой теоремы достаточно показать, что при $t < 1$

$$\lim_{h \downarrow 0} \limsup_{n \to \infty} \frac{1}{h} \mathsf{E}\{|\mathsf{E}\{X_n(t+h) - X_n(t)\} \| \mathscr{F}_{[nt]}\}|\} = 0 \tag{23.3}$$

и

$$\lim_{h \downarrow 0} \limsup_{n \to \infty} \frac{1}{h} \mathsf{E}\{|\mathsf{E}\{(X_n(t+h) - X_n(t))^2 \| \mathscr{F}_{[nt]}\} - h|\} = 0. \tag{23.4}$$

Соотношение (23.3) непосредственно вытекает из (23.2). Для доказательства (23.4) заметим, что согласно (23.2)

$$\mathsf{E}\{(S_{k+l} - S_k)^2 \| \mathscr{F}_k\} = \sum_{i=1}^{l} \mathsf{E}\{\xi_{k+i}^2 \| \mathscr{F}_k\},$$

*) По поводу используемых здесь эргодической теории и теории мартингалов см. книгу Дуба (1956).

**) В качестве ξ_n можно взять координатные величины в произведении двусторонне бесконечной последовательности действительных прямых, имеющие те же конечномерные распределения, что и исходный процесс; эргодичность сохраняется, поскольку она вполне определяется конечномерными распределениями.

откуда в силу стационарности математическое ожидание в (23.4) равно

$$\frac{h}{\sigma^2} \mathsf{E} \left\{ \left| \frac{1}{nh} \sum_{i=1}^{k_n} \mathsf{E} \{\xi_i^2 \| \mathcal{F}_0\} - \sigma^2 \right| \right\},$$

где $k_n = [n(t+h)] - [nt]$. Поскольку по неравенству Иенсена, последнее выражение не превосходит

$$\frac{h}{\sigma^2} \mathsf{E} \left\{ \left| \frac{1}{nh} \sum_{i=1}^{k_n} \xi_i^2 - \sigma^2 \right| \right\},$$

то (23.4) следует теперь из интегральной эргодической теоремы.

Согласно (23.2) величины ξ_k ортогональны и, следовательно,

$$\mathsf{E}\{S_n^2\} = n\sigma^2, \tag{23.5}$$

откуда вытекает справедливость условия 2.

Предположим, что мы доказали равномерную интегрируемость последовательности $\{\max\limits_{k \leqslant n} S_k^2/n\}$, т. е. что

$$\lim_{\alpha \to \infty} \sup_n \mathsf{E}_\alpha \left\{ \frac{1}{n} \max_{k \leqslant n} S_k^2 \right\} = 0, \tag{23.6}$$

где под $\mathsf{E}_\alpha\{U\}$ понимается интеграл от U по множеству $\{U \geqslant \alpha\}$. Тогда, конечно, последовательность $\{X_n^2(t);\ n \geqslant 1\}$ тоже будет равномерно интегрируемой при каждом t, что является одним из условий теоремы 19.4 и выполнение условия 3°а теперь легко вытекает из стационарности. Наконец, поскольку

$$\mathsf{P}\left\{\max_{k \leqslant n} |S_k| \geqslant \lambda \sqrt{n}\right\} \leqslant \frac{1}{\lambda^2} \mathsf{E}_{\lambda^2} \left\{ \frac{1}{n} \max_{k \leqslant n} S_k^2 \right\},$$

то по теореме 8.4 выполнено условие плотности (19.51). Следовательно, достаточно доказать (23.6).

Если ξ_0 имеет четвертый момент, то

$$\mathsf{E}\{S_n^4\} = \sum \mathsf{E}\{\xi_i \xi_j \xi_k \xi_l\},$$

где индексы независимо друг от друга пробегают все значения от 1 до n. Согласно (23.2) слагаемое, у которого максимальное значение индекса строго больше остальных, равно нулю. Следовательно,

$$\mathsf{E}\{S_n^4\} = \sum_k \mathsf{E}\{\xi_k^4\} + 4\sum_{l<k} \mathsf{E}\{\xi_l \xi_k^3\} + 6\sum_{l,\,j<k} \mathsf{E}\{\xi_l \xi_j \xi_k^2\}.$$

Если $|\xi_0| \leqslant C$ с вероятностью 1, то сумма первых двух сумм в правой части не превосходит $3n^2C^4$. Последняя сумма, равная $6\sum_{k=2}^{n} \mathsf{E}\{S_{k-1}^2 \xi_k^2\}$, тоже не превосходит согласно (23.5) $3n^2C^4$. Таким образом,

$$\mathsf{E}\{S_n^4\} \leqslant 6n^2C^4, \tag{23.7}$$

если $|\xi_0| < C$ с вероятностью 1.

В силу неравенства для мартингалов *)

$$\mathsf{E}\{\max_{k\leqslant n} |S_k|^\gamma\} \leqslant \left(\frac{\gamma}{\gamma-1}\right)^\gamma \mathsf{E}\{|S_n|^\gamma\} \tag{23.8}$$

при $\gamma > 1$. Следовательно, согласно (23.5)

$$\mathsf{E}\{\max_{k\leqslant n} S_k^2\} \leqslant 4n\mathsf{E}\{\xi_0^2\}. \tag{23.9}$$

Если же $|\xi_0|$ не превосходит C, то согласно (23.7)

$$\mathsf{E}\{\max_{k\leqslant n} S_k^4\} \leqslant \left(\frac{4}{3}\right)^4 6n^2C^4. \tag{23.10}$$

При $u > 0$ определим

$$\xi_{lu} = \begin{cases} \xi_l, & \text{если } |\xi_l| \leqslant u, \\ 0, & \text{если } |\xi_l| > u, \end{cases}$$

и положим

$$\eta_{lu} = \xi_{lu} - \mathsf{E}\{\xi_{lu} \,\|\, \mathcal{F}_{l-1}\},$$
$$\delta_{lu} = \xi_l - \eta_{lu} = \xi_l - \xi_{lu} - \mathsf{E}\{\xi_l - \xi_{lu} \,\|\, \mathcal{F}_{l-1}\}.$$

Если $S_{ku} = \sum_{l=1}^{k} \eta_{lu}$ и $D_{ku} = \sum_{l=1}^{k} \delta_{lu}$, то

$$\frac{1}{n}\max_{k\leqslant n} S_k^2 \leqslant \frac{2}{n}\max_{k\leqslant n} S_{ku}^2 + \frac{2}{n}\max_{k\leqslant n} D_{ku}^2. \tag{23.11}$$

---

*) Дуб (1956, стр. 285).

Далее, последовательность $\{\eta_{nu};\, n \geqslant 1\}$ обладает мартингальным свойством (23.1) и $|\eta_{nu}| \leqslant 2u$. Поэтому в силу (23.10)

$$\mathsf{E}_\alpha\left\{\frac{1}{n}\max_{k\leqslant n} S_{ku}^2\right\} \leqslant \frac{1}{\alpha}\,\mathsf{E}\left\{\frac{1}{n^2}\max_{k\leqslant n} S_{ku}^4\right\} \leqslant \frac{1}{\alpha}\left(\frac{4}{3}\right)^4 6\,(2u)^4.$$

Последовательность $\{\delta_{nu};\, n \geqslant 1\}$ также обладает мартингальным свойством, так что согласно (23.9) и лемме 1 из § 21 (см. стр. 255).

$$\mathsf{E}\left\{\frac{1}{n}\max_{k\leqslant n} D_{ku}^2\right\} \leqslant 4\mathsf{E}\{\delta_{0u}^2\} \leqslant 4\mathsf{E}\{(\xi_0 - \xi_{0u})^2\}.$$

Из соотношения (23.11) и неравенства $\mathsf{E}_\alpha\{U+V\} \leqslant \leqslant 2\mathsf{E}_{\frac{1}{2}\alpha}\{U\} + 2\mathsf{E}_{\frac{1}{2}\alpha}\{V\}$ следует, что

$$\mathsf{E}_\alpha\left\{\frac{1}{n}\max_{k\leqslant n} S_k^2\right\} \leqslant K\left[\frac{u^4}{\alpha} + \mathsf{E}_{u'}(\xi_0^2)\right],$$

где K — абсолютная постоянная. Так как ξ_0 имеет конечный второй момент, то мы получаем нужное соотношение (23.6).

Примечание. Центральная предельная теорема в предположениях теоремы 23.1 была доказана независимо Биллингсли (1961) и Ибрагимовым (1963); доказательство теоремы 23.1 (для ограниченных ξ_n) принадлежит Розену (1967а).

§ 24. Симметрично зависимые случайные величины

Выбор. Пусть для каждого n

$$x_{n1},\ x_{n2},\ \ldots,\ x_{nk_n} \tag{24.1}$$

— последовательность (необязательно различных) действительных чисел, и пусть $\xi_{n1}(\omega), \ldots, \xi_{nk_n}(\omega)$ — случайная перестановка этих чисел. Предположим, что каждая из $k_n!$ таких перестановок имеет одинаковую вероятность $1/k_n!$. Определим случайный элемент X_n в D, положив

$$X_n(t, \omega) = \sum_{i=1}^{[k_n t]} \xi_{ni}(\omega). \tag{24.2}$$

При $0 \leqslant t < 1/k_n$ считаем $X_n(t, \omega) = 0$. Если производится выбор из конечной совокупности (24.1) до ее полного исчерпания, то X_n описывает ход выбора.

Теорема 24.1. *Если*

$$\sum_{i=1}^{k_n} x_{ni} = 0, \qquad \sum_{i=1}^{k_n} x_{ni}^2 = 1, \tag{24.3}$$

$$\max_{1 \leqslant i \leqslant k_n} |x_{ni}| \to 0 \tag{24.4}$$

и X_n — случайный элемент (24.2), *то*

$$X_n \xrightarrow{\mathscr{D}} W^\circ. \tag{24.5}$$

Доказательство. Для доказательства достаточно проверить выполнение условий теоремы 19.4 с функциями $\rho(t) = -1/(1-t)$ и $\sigma^2(t) = 1$, которые характеризуют броуновский мост W° (см. (19.32)).

Проведем сначала один предварительный подсчет. Пусть $y_1, \ldots, y_m$ — действительные числа, $1 \leqslant m \leqslant u$, и пусть $\eta_1, \ldots, \eta_m$ — упорядоченная выборка без возвращения объема m (имеется всего $\binom{u}{m}$ возможных выборок, и все они равновероятны). Из соображений симметрии среднее значение суммы $\sum_{i=1}^{m} \eta_i$ равно $m\mathsf{E}\{\eta_1\}$ и, следовательно,

$$\mathsf{E}\left\{\sum_{i=1}^{m} \eta_i\right\} = \frac{m}{u} \sum_{i=1}^{u} y_i. \tag{24.6}$$

Второй момент этой суммы, тоже из соображений симметрии, равен

$$m\mathsf{E}\{\eta_1^2\} + m(m-1)\,\mathsf{E}\{\eta_1\eta_2\}.$$

Производя далее простые вычисления, получаем

$$\mathsf{E}\left\{\left[\sum_{i=1}^{m} \eta_i\right]^2\right\} = \frac{m(u-m)}{u(u-1)} \sum_{i=1}^{u} y_i^2 + \frac{m(m-1)}{u(u-1)} \left[\sum_{i=1}^{u} y_i\right]^2. \tag{24.7}$$

Применяя эти результаты к сумме $\sum_{i=1}^{m} \xi_{ni}$ и используя (24.3), приходим к соотношениям

$$\mathsf{E}\{X_n(t)\} = 0, \quad \mathsf{E}\{X_n^2(t)\} = \frac{[k_n t]\,(k_n - [k_n t])}{k_n(k_n - 1)}.$$

Второй момент здесь сходится к $t(1-t)$ (заметим, что из соотношений (24.3) и (24.4) следует, что $k_n \to \infty$).

Предположим теперь, что $1 \leqslant m_1 < m_1 + m_2 \leqslant k_n$, и пусть известны значения величин $\xi_{n1}, \ldots, \xi_{nm_1}$. Тогда условное распределение величин $\xi_{n,\,m_1+1}, \ldots, \xi_{n,\,m_1+m_2}$ совпадает с распределением выборки объема m_2 из совокупности (24.1) с выброшенными значениями $\xi_{n1}, \ldots, \xi_{nm_1}$. Из (24.6) и (24.7) с $m = m_2$ и $u = k_n - m_1$ в силу (24.3) получаем

$$\mathsf{E}\left\{\sum_{i=m_1+1}^{m_1+m_2} \xi_{ni} \,\Big\|\, \xi_{n1}, \ldots, \xi_{nm_1}\right\} = -\frac{m_2}{k_n - m_1} \sum_{i=1}^{m_1} \xi_{ni} \tag{24.8}$$

и

$$\mathsf{E}\left\{\left[\sum_{i=m_1+1}^{m_1+m_2} \xi_{ni}\right]^2 \,\Bigg\|\, \xi_{n1}, \ldots, \xi_{nm_1}\right\} =$$

$$= \frac{m_2(k_n - m_1 - m_2)}{(k_n - m_1)(k_n - m_1 - 1)}\left[1 - \sum_{i=1}^{m_1} \xi_{ni}^2\right] +$$

$$+ \frac{m_2(m_2 - 1)}{(k_n - m_1)(k_n - m_1 - 1)}\left[\sum_{i=1}^{m_1} \xi_{ni}\right]^2. \tag{24.9}$$

Зафиксируем $t < 1$; полагая в (24.8) $m_1 = [k_n t]$ и $m_2 = [k_n(t+h)] - [k_n t]$, находим

$$\mathsf{E}\{X_n(t+h) - X_n(t) \,\|\, X_n(t)\} = -A_n(h)\, X_n(t),$$

где

$$A_n(h) = \frac{[k_n(t+h)] - [k_n t]}{k_n - [k_n t]}.$$

Следовательно,

$$\frac{1}{h}\mathsf{E}\left\{\left|\mathsf{E}\{X_n(t+h) - X_n(t) \,\|\, X_n(t)\} + \frac{h}{1-t} X_n(t)\right|\right\} =$$

$$= \frac{1}{h}\left|A_n(h) - \frac{h}{1-t}\right| \mathsf{E}\{|X_n(t)|\}.$$

Так как величина $\mathsf{E}\{|X_n(t)|\}$ не превосходит $\mathsf{E}^{1/2}\{X_n^2(t)\}$ и тем самым ограничена и так как $\lim_n A_n(h) = h/(1-t)$, то отсюда следует (19.43). (Для простоты обозначений мы рассматриваем условие на $X_n(t)$ лишь при одном значении t; общий случай также просто выводится из (21.8). Это замечание относится и к нижеследующему рассуждению.)

Из (24.9) следует, что

$$\mathsf{E}\{(X_n(t+h) - X_n(t))^2 \| X_n(t)\} = \\ = B_n(h)\left[1 - \sum_{i=1}^{[k_nt]} \xi_{ni}^2\right] + C_n(h)\, X_n^2(t),$$

где

$$B_n(h) = \frac{([k_n(t+h)] - [k_nt])(k_n - [k_n(t+h)])}{(k_n - [k_nt])(k_n - [k_nt] - 1)}$$

и

$$C_n(h) = \frac{([k_n(t+h)] - [k_nt])([k_n(t+h)] - [k_nt] - 1)}{(k_n - [k_nt])(k_n - [k_nt] - 1)}.$$

Таким образом,

$$\frac{1}{h}\mathsf{E}\{|\mathsf{E}\{(X_n(t+h) - X_n(t))^2 \| X_n(t)\} - h|\} \leqslant \\ \leqslant \frac{|B_n(h)|}{h}\mathsf{E}\left\{\left|\sum_{i=1}^{[k_nt]} \xi_{ni}^2 - \frac{[k_nt]}{k_n}\right|\right\} + \\ + \left|\frac{B_n(h)}{h}\left(1 - \frac{[k_nt]}{k_n}\right) - 1\right| + \frac{|C_n(h)|}{h}\mathsf{E}\{X_n^2(t)\}. \quad (24.10)$$

Сумма $\sum_{i=1}^{[k_nt]} \xi_{ni}^2$ имеет среднее значение $[k_nt]/k_n$ и второй момент (в силу (24.7))

$$\frac{[k_nt](k_n - [k_nt])}{k_n(k_n-1)}\sum_{i=1}^{k_n} x_{ni}^4 + \frac{[k_nt]([k_nt]-1)}{k_n(k_n-1)} \to t^2.$$

Следовательно, дисперсия этой суммы стремится к 0. Так как

$$\lim_n B_n(h) = h(1-t-h)/(1-t)^2,$$

то первый член в правой части неравенства (24.10) стремится к 0 при $n\to\infty$. Поскольку среднее значение

$\mathsf{E}\{X_n^2(t)\}$ ограничено и $\lim_n C_n(h) = h^2/(1-t)^2$, то отсюда следует (19.44).

Таким образом, мы проверили условие 1° теоремы 19.4. Так как максимальный скачок $\max_i |X_{ni}|$ случайной функции X_n стремится к 0, то согласно замечанию, сделанному после доказательства теоремы 19.4, нам надо проверить лишь равенство $X_n(0) = 0$ (которое очевидно) и условие 3. Следовательно, осталось только показать, что

$$\mathsf{E}\{(X_n(t) - X_n(t_1))^2 (X_n(t_2) - X_n(t))^2\} \leqslant K(t_2 - t_1)^2, \quad t_1 \leqslant t \leqslant t_2, \tag{24.11}$$

для некоторого K, не зависящего от n, t_1 и t_2. Левая часть (24.11) равна

$$\sum \mathsf{E}\{\xi_{ni}\xi_{nj}\xi_{nk}\xi_{nl}\}, \tag{24.12}$$

где суммирование производится по i, j, k, l таким, что $[k_n t_1] < i, j \leqslant [k_n t] < k, l \leqslant [k_n t_2]$. Положим $[k_n t] - [k_n t_1] = m_1$ и $[k_n t_2] - [k_n t] = m_2$; из соображений симметрии сумма (14.12) равна

$$m_1 m_2 \mathsf{E}\{\xi_{n1}^2 \xi_{n2}^2\} + m_1 m_2 (m_1 + m_2 - 2)\, \mathsf{E}\{\xi_{n1}^2 \xi_{n2}\xi_{n3}\} + \\ + m_1(m_1 - 1) m_2 (m_2 - 1)\, \mathsf{E}\{\xi_{n1}\xi_{n2}\xi_{n3}\xi_{n4}\}.$$

Обозначим $\tau_n = \sum_{i=1}^{k_n} x_{ni}^4$. Дальнейшие вычисления (хотя и громоздкие, но стандартные) приводят к следующему выражению для (24.12):

$$m_1 m_2 \frac{1 - \tau_n}{k_n(k_n - 1)} + m_1 m_2 (m_1 + m_2 - 2) \frac{2\tau_n - 1}{k_n(k_n - 1)(k_n - 2)} + \\ + m_1(m_1 - 1) m_2 (m_2 - 1) \frac{3(1 - 2\tau_n)}{k_n(k_n - 1)(k_n - 2)(k_n - 3)}.$$

Поскольку согласно (24.3) $0 \leqslant \tau_n \leqslant 1$, то при $k_n \geqslant 6$ последнее выражение не превосходит

$$\frac{2m_1 m_2}{k_n^2} + \frac{8 m_1 m_2 (m_1 + m_2)}{k_n^3} + \frac{24 m_1^2 m_2^2}{k_n^4} \leqslant 34 \left(\frac{m_1 + m_2}{k_n}\right)^2.$$

Если $t_2 - t_1 \geqslant 1/k_n$, то $(m_1 + m_2)/k_n \leqslant 3(t_2 - t_1)$, и поэтому (24.11) имеет место с $K = 306$. Если $t_2 - t_1 <$

$<1/k_n$, то (24.11) выполняется, так как один из двух сомножителей под знаком математического ожидания обращается в нуль. Тем самым теорема 24.1 доказана.

Нам понадобится следующее следствие из теоремы 24.1. Для любого множества A непрерывности меры W° и любого положительного ε существует положительное $\delta(\varepsilon, A)$ такое, что если числа $x_{n1}, \ldots, x_{nk_n}$ удовлетворяют условию (24.3) и

$$\max_{1\leqslant i\leqslant k_n} |x_{ni}| \leqslant \delta(\varepsilon, A), \tag{24.13}$$

то

$$|\mathsf{P}\{X_n \in A\} - W^\circ(A)| < \varepsilon \tag{24.14}$$

(здесь X_n определяется соотношением (24.2)). Действительно, если бы для некоторых ε и A не существовало такого $\delta(\varepsilon, A)$, то можно было бы построить последовательность множеств $\{x_{ni}\}$, удовлетворяющих условиям (24.3) и (24.4), для которых соотношение (24.5) неверно.

Симметрично зависимые величины. Рассмотрим теперь более общую ситуацию. Пусть при каждом n случайные величины

$$\xi_{n1}, \xi_{n2}, \ldots, \xi_{nk_n} \tag{24.15}$$

симметрично зависимы в том смысле, что их совместное распределение инвариантно относительно их перестановки. Определим X_n, как и прежде, равенством (24.2). Предположим, что величины удовлетворяют следующим трем условиям:

$$\sum_{i=1}^{k_n} \xi_{ni} \xrightarrow{P} 0, \quad \sum_{i=1}^{k_n} \xi_{ni}^2 \xrightarrow{P} 1, \quad \max_{1\leqslant i\leqslant k_n} |\xi_{ni}| \xrightarrow{P} 0 \tag{24.16}$$

при $n \to \infty$.

Если случайные величины (24.15) представляют собой случайную перестановку множества (24.1), то они симметрично зависимы и из условий (24.3) и (24.4) следует (24.16). Таким образом, следующий результат обобщает теорему 24.1:

Теорема 24.2. *Если случайные величины* (24.15) *симметрично зависимы и удовлетворяют условию* (24.16), *то* $X_n \xrightarrow{\mathcal{D}} W^\circ$.

Доказательство. Если $\alpha_n = k_n^{-1} \sum_{i=1}^{k_n} \xi_{ni}$ и $\beta_n^2 = \sum_{i=1}^{k_n} (\xi_{ni} - \alpha_n)^2$, то согласно (24.16)

$$k_n \alpha_n \xrightarrow{P} 0, \qquad \beta_n \xrightarrow{P} 1. \tag{24.17}$$

Случайные величины $\eta_{ni} = (\xi_{ni} - \alpha_n)/\beta_n$, $i = 1, \ldots, k_n$, симметрично зависимы и

$$\sum_{i=1}^{k_n} \eta_{ni} = 0, \quad \sum_{i=1}^{k_n} \eta_{ni}^2 = 1, \quad \max_{1 \leqslant i \leqslant k_n} |\eta_{ni}| \xrightarrow{P} 0. \tag{24.18}$$

Определим случайный элемент Y_n в D, положив

$$Y_n(t, \omega) = \sum_{i=1}^{[k_n t]} \eta_{ni}(\omega).$$

Поскольку

$$X_n(t) = \beta_n Y_n(t) + [k_n t]\, \alpha_n,$$

то, в силу (24.17) и теоремы 4.1, соотношения $X_n \xrightarrow{\mathscr{D}} W^\circ$ и $Y_n \xrightarrow{\mathscr{D}} W^\circ$ равносильны. Мы докажем, что $Y_n \xrightarrow{\mathscr{D}} W^\circ$.

Пусть $\zeta_{n1}, \ldots, \zeta_{nk_n}$ — случайная перестановка множества $\eta_{n1}, \ldots, \eta_{nk_n}$. (Случайная перестановка случайных величин η_{ni} независима от них самих. Чтобы случайная перестановка была вполне определена, вероятностное пространство, на котором заданы величины η_{ni}, возможно, должно быть расширено.) Величины ζ_{ni} имеют то же самое совместное распределение, что и величины η_{ni}, так как последние симметрично зависимы. Следовательно, распределение Y_n совпадает с распределением случайной функции

$$Z_n(t, \omega) = \sum_{i=1}^{[k_n t]} \zeta_{ni}(\omega).$$

Для доказательства $Y_n \xrightarrow{\mathscr{D}} W^\circ$ достаточно, таким образом, установить, что

$$\mathsf{P}\{Z_n \in A\} \to W^\circ(A) \tag{24.19}$$

для каждого множества A непрерывности меры W°, По заданному множеству A непрерывности меры W° и положительному ε выберем $\delta(\varepsilon, A)$ таким образом, чтобы из (24.13) следовало (24.14) (при этом в (24.14) X_n есть случайная функция (24.2), определенная посредством случайной перестановки чисел (24.1), удовлетворяющих условию (24.3)). Событие

$$E_n = \{\max_{1 \leqslant l \leqslant k_n} |\eta_{nl}| < \delta(A, \varepsilon)\}$$

согласно (24.18) удовлетворяет условию $\mathsf{P}(E_n) \to 1$. По определению условной вероятности

$$\mathsf{P}\{Z_n \in A\} = \int_{E_n} \mathsf{P}\{Z_n \in A \| \eta_{n1}, \ldots, \eta_{nk_n}\} d\mathsf{P} + \\ + \int_{E_n^c} \mathsf{P}\{Z_n \in A \| \eta_{n1}, \ldots, \eta_{nk_n}\} d\mathsf{P}.$$

В силу (24.18) на множестве E_n имеем

$$\left|\mathsf{P}\{Z_n \in A \| \eta_{n1}, \ldots, \eta_{nk_n}\} - W^\circ(A)\right| < \varepsilon.$$

Следовательно,

$$|\mathsf{P}(Z_n \in A\} - W^\circ(A)| \leqslant 2\varepsilon + 2\mathsf{P}(E_n^c).$$

Поскольку $\mathsf{P}(E_n) \to 1$, отсюда вытекает (24.19) и теорема доказана.

Комбинируя теорему 24.2 с вычислениями из § 11 (см. (11.39), (11.40) и (11.42)), получаем предельные распределения для $\max\limits_{k \leqslant k_n} \sum\limits_{l=1}^{k} \xi_{nl}$, $\max\limits_{k \leqslant k_n} \left|\sum\limits_{l=1}^{k} \xi_{nl}\right|$ и для доли тех k, $1 \leqslant k \leqslant k_n$, которые удовлетворяют условию $\sum\limits_{l=1}^{k} \xi_{nl} > 0$.

Примечание. В связи с теоремой 24.1 и ее обобщениями на другие выборочные процедуры см. Розен (1964, 1967a, 1967c). Теорема 24.2 новая. Чернов и Тейчер (1958) доказали, что в предположениях этой теоремы $X_n(t) \xrightarrow{\mathscr{D}} N(0, t(1-t))$ (t фиксировано); см. их работу по поводу примеров величин, удовлетворяющих этим предположениям. Родственные приведенным предельные теоремы имеют дело с условными вероятностями, вычисленными при услови $X_n(1) = 0$, где X_n — некоторая случайная функция. См. Дуосс и Карлин (1963) (и ссылки там) и Трюмбо (1965), а также Лигетт (1968) и Уичура (1969).

Добавление I

МЕТРИЧЕСКИЕ ПРОСТРАНСТВА

Мы приведем здесь некоторые сведения о метрических пространствах, предполагая уже известными их первичные свойства и относящиеся к ним начальные определения (открытые и замкнутые множества, плотность, предельная точка, непрерывность и т. д.) *).

Мы будем обозначать метрическое пространство буквой S, а саму метрику символом $\rho(x, y)$. Будем обозначать замыкание подмножества A пространства S через A^-, его внутреннюю часть через A°, а его границу через ∂A $(\partial A = A^- - A^\circ)$. Определим расстояние от x до A равенством

$$\rho(x, A) = \inf\{\rho(x, y)\colon\ y \in A\};$$

легко проверить, что расстояние $\rho(x, A)$ равномерно непрерывно по x. Обозначим через $S(x, \varepsilon)$ открытый шар с центром x и радиусом ε: $S(x, \varepsilon) = \{y\colon \rho(x, y) < \varepsilon\}$. Под «шаром» мы будем понимать «открытый шар» и будем называть $S(x, \varepsilon)$ ε-шаром с центром x.

Две метрики ρ_1 и ρ_2 на S называются эквивалентными, если для любых x и ε существует такое δ, что $S_1(x, \delta) \subset S_2(x, \varepsilon)$ и $S_2(x, \delta) \subset S_1(x, \varepsilon)$ (здесь $S_i(x, \varepsilon) = \{y\colon \rho_i(x, y) < \varepsilon\}$), так что пространство S с метрикой ρ_1 гомеоморфно пространству S с метрикой ρ_2.

Сепарабельность. Пространство S по определению сепарабельно, если оно содержит счетное всюду плотное подмножество. База пространства S есть такой класс открытых множеств, что всякое открытое подмножество S является объединением элементов этого класса. От-

*) Более полное изложение см., например, в книгах Дьедонне (1964), Колмогоров и Фомин (1972).

крытым покрытием множества A называется класс открытых множеств, объединение которых содержит A. Множество A дискретно, если каждая его точка лежит в некотором шаре, не содержащем никаких других точек из A, иными словами, если каждая точка множества A является изолированной в относительной топологии. Если пространство S само дискретно, то его исходная метрика эквивалентна метрике, в которой расстояние между любыми двумя различными точками равно 1.

Следующие три условия эквивалентны:

(i) *S сепарабельно.*

(ii) *S имеет счетную базу.*

(iii) *Любое открытое покрытие любого подмножества S имеет счетное подпокрытие.*

Кроме того, из сепарабельности следует, что

(iv) *S не содержит несчетного дискретного множества, а отсюда в свою очередь вытекает, что*

(v) *S не содержит несчетного множества A такого, что*

$$\inf\{\rho(x, y): x, y \in A, x \neq y\} > 0. \tag{1}$$

Доказательство импликации (i) $\to$ (ii). Пусть D — счетное всюду плотное множество в S. Обозначим через $\mathscr{V}$ счетный класс сфер с рациональными радиусами и с центрами в D. Пусть множество G открыто; для доказательства того, что $\mathscr{V}$ является базой, мы должны показать, что если G_1 представляет собой объединение элементов $\mathscr{V}$, которые содержатся в G, то $G = G_1$. Очевидно, что $G_1 \subset G$; чтобы доказать включение $G \subset G_1$, достаточно для данного элемента x из G найти d в D и рациональное r такие, что $x \in S(d, r) \subset G$. Но если $x \in G$, то $S(x, \varepsilon) \subset G$ для некоторого положительного ε. Поскольку D плотно, то в D найдется элемент d такой, что $\rho(x, d) < \frac{1}{2}\varepsilon$. Выберем теперь рациональное r так, чтобы $\rho(x, d) < r < \frac{1}{2}\varepsilon$.

Доказательство импликации (ii) $\to$ (iii). Пусть $\{V_1, V_2, \ldots\}$ — счетная база. Предположим, что $\{G_\alpha\}$ — открытое покрытие множества A (α пробегает произвольное множество индексов). Для любого V_k, для которого существует такое G_α, что $V_k \subset G_\alpha$, пусть G_{α_k} — одно из таких G_α, содержащих V_k. Тогда $A \subset \bigcup_k G_{\alpha_k}$.

Доказательство импликации (iii) $\to$ (i). При каждом n $\{S(x, n^{-1}): x \in S\}$ является открытым покрытием S. Если (iii) выполняется, то существует счетное подпокрытие $\{S(x_{n,k}, n^{-1}): k = 1, 2, \ldots\}$. Счетное множество $\{x_{n,k}: n, k = 1, 2, \ldots\}$ является всюду плотным в S.

Доказательство импликации (iii) $\to$ (iv). Если A дискретно, то для каждой точки x из A существует положительное ε_x такое, что $S(x, \varepsilon_x)$ не содержит других точек A. Так как $\{S(x, \varepsilon_x): x \in A\}$ является открытым покрытием A, не обладающим каким бы то ни было подпокрытием, то (iii) не может выполняться, если A несчетно.

Поскольку множество, обладающее свойством (1), дискретно, то, разумеется, (iv) влечет (v). Хотя в дальнейшем нам это и не понадобится, отметим все же, что из (v) следует сепарабельность, так что все предположения (i) — (v) являются эквивалентными. (Для любого положительного ε можно, используя лемму Цорна, указать максимальное множество A_ε точек, отдаленных одна от другой на расстояние не менее чем ε; объединение множеств A_ε по всем рациональным ε плотно, а если (v) выполняется, то и счетно.)

Компактность. Множество A в S по определению компактно, если каждое открытое покрытие A содержит конечное подпокрытие. ε-сетью для A называется множество точек $\{x_k\}$ с тем свойством, что для любой точки x из A существует точка x_k такая, что $\rho(x, x_k) < \varepsilon$ (точки x_k не обязаны принадлежать A). Множество называется вполне ограниченным, если для любого положительного ε оно имеет конечную ε-сеть. Множество A назовем полным, если каждая фундаментальная последовательность в A сходится к некоторой точке из A.

Для произвольного множества A из S следующие четыре условия эквивалентны:

(i) *Множество A^- компактно.*

(ii) *Любое счетное открытое покрытие множества A^- имеет конечное подпокрытие.*

(iii) *Любая последовательность в A имеет предельную точку (т. е. содержит подпоследовательность, сходящуюся к пределу, который необходимо лежит в A^-).*

(iv) *A вполне ограничено, а A^- полно.*

Легко показать, что (iii) выполняется тогда и только тогда, когда любая последовательность в A^- имеет предельную точку (необходимо принадлежащую A^-), и что A является вполне ограниченным тогда и только тогда, когда вполне ограничено A^-. Поэтому мы можем предположить при доказательстве, что $A = A^-$, т. е. что множество A замкнуто. Импликация (i) $\to$ (ii) тривиальна.

Доказательство импликации (ii) $\to$ (iii). Фиксируя последовательность $\{x_n\}$ в A, определим F_n как замыкание множества $\{x_k: k \geqslant n\}$. Если $\bigcap_n F_n = 0$, то открытые множества F_n^c покрывают A и, следовательно, если (ii) выполняется, то $A \subset F_1^c \cup \ldots \cup F_n^c$ для некоторого n, откуда вытекает невозможное соотношение $F_n \cap A = 0$. Таким образом, $\bigcap_n F_n$ содержит некоторую точку x, которая должна быть предельной точкой последовательности $\{x_n\}$.

Доказательство импликации (iii) $\to$ (iv). Если A не является вполне ограниченным множеством, то существует некоторое положительное ε и некоторая последовательность $\{x_n\}$ такие, что $\rho(x_m, x_n) \geqslant \varepsilon$ при $m \neq n$; $\{x_n\}$ не может иметь предельную точку. Следовательно, из (iii) вытекает полная ограниченность. Кроме того, из (iii) следует полнота, так как если последовательность $\{x_n\}$ фундаментальна и имеет предельную точку x, то $\{x_n\}$ сходится к x.

Доказательство импликации (iv) $\to$ (i). Пусть (iv) выполняется. Допустим, что $\{G_\alpha\}$, где α пробегает произвольное множество индексов, является открытым покрытием A, не имеющим конечного подпокрытия. Покажем, что это приводит к противоречию.

Поскольку A вполне ограничено, оно может быть при любом n покрыто конечным множеством открытых шаров $B_{n1}, \ldots, B_{nk_n}$ радиуса 2^{-n}. По крайней мере одно из множеств B_{ni} должно обладать тем свойством, что никакое конечное подсемейство семейства $\{G_\alpha\}$ не покрывает $A \cap B_{ni}$ (последнее множество должно быть поэтому непустым); пусть C_n — одно из таких множеств B_{ni}. Поскольку система $B_{ni} \cap C_{n-1} \cap A$ покрывает $C_{n-1} \cap A$ (при $n > 1$), мы вправе также считать, что никакое конечное подсемейство семейства $\{G_\alpha\}$ не покрывает $C_n \cap C_{n-1} \cap A$, так что, в частности, $C_n \cap C_{n-1} \neq 0$.

Пусть x_n — одна из точек, которые принадлежат одновременно и C_n, и A. Поскольку $C_n \cap C_{n-1} \neq 0$ и C_n имеет радиус 2^{-n}, то $\rho(x_n, x_{n-1}) < 6 \cdot 2^{-n}$, откуда следует, что $\rho(x_n, x_{n+k}) < 6 \cdot 2^{-n}$. Таким образом, последовательность $\{x_n\}$ фундаментальна и имеет в силу предположенной полноты предел x, который должен принадлежать A.

Далее, $x \in G_\alpha$ при некотором α и $S(x, \varepsilon) \subset G_\alpha$ при некотором положительном ε. Но тогда $\rho(x, x_n) < \varepsilon/3$ при некотором n, удовлетворяющем неравенству $2^{-n} < \varepsilon/3$, откуда $C_n \subset G_\alpha$, что противоречит сделанному выше предположению *).

Подмножество k-мерного евклидова пространства R^k (с обычной метрикой) имеет компактное замыкание тогда и только тогда, когда оно ограничено.

Если h — непрерывное отображение S в другое метрическое пространство S' и если K — компактное подмножество S, то hK является компактным подмножеством S'.

Доказательство. Если $\{G'_\alpha\}$ является открытым покрытием hK, то $\{h^{-1}G'_\alpha\}$ является открытым покрытием K и, следовательно, имеет конечное подпокрытие $\{h^{-1}G'_{\alpha_i}\colon i = 1, \ldots, n\}$. Очевидно, что $\{G'_{\alpha_i}\colon i = 1, \ldots, n\}$ покрывает hK.

Полунепрерывность сверху. Функция f называется полунепрерывной сверху в точке x, если для любого положительного ε существует положительное δ такое, что из $\rho(x, y) < \delta$ вытекает $f(y) < f(x) + \varepsilon$. Легко видеть, что f всюду полунепрерывна сверху тогда и только тогда, когда для каждого действительного α множество $\{x\colon f(x) < \alpha\}$ является открытым.

Пусть f_n — действительные функции на S такие, что в каждой точке x последовательность $f_n(x)$ является невозрастающей и

$$\lim_{n\to\infty} f_n(x) = 0. \tag{2}$$

---

*) Доказательство заимствовано из книги Дьедоние (1964).

Если функции f_n полунепрерывны сверху, то сходимость в (2) равномерна на каждом компактном множестве.

Доказательство. При любом положительном ε открытые множества $G_n = \{x: f_n(x) < \varepsilon\}$ покрывают S. Если K — компакт, то $K \subset G_n$ для некоторого n, и равномерность доказана.

Пространство R^∞. Под R^∞ мы понимаем пространство последовательностей $x = (x_1, x_2, \ldots)$ действительных чисел. Функция $\rho_0(\alpha, \beta) = |\alpha - \beta|/(1 + |\alpha - \beta|)$ представляет собой метрику на прямой R^1, эквивалентную обычной метрике $|\alpha - \beta|$. Действительная прямая является полным пространством относительно метрики ρ_0. Функция $\rho(x, y) = \sum\limits_{k=1}^{\infty} \rho_0(x_k, y_k) 2^{-k}$ представляет собой метрику на R^∞. Множества

$$N_{k,\varepsilon}(x) = \{y: |y_i - x_i| < \varepsilon,\ i = 1, \ldots, k\} \tag{3}$$

открыты относительно метрики ρ. Более того, поскольку из $\rho(x, y) < \varepsilon \cdot 2^{-k}/(1 + \varepsilon)$ вытекает $y \in N_{k,\varepsilon}(x)$, откуда в свою очередь вытекает $\rho(x, y) < \varepsilon + 2^{-k}$, то множества (3) образуют базу топологии, порождаемой метрикой ρ.

Мы будем всегда рассматривать пространство R^∞ с этой топологией, которая называется топологией произведения или топологией покоординатной сходимости: $\lim_n x(n) = x$ тогда и только тогда, когда $\lim_n x_k(n) = x_k$ для каждого k. Пространство R^∞ сепарабельно; в качестве счетного всюду плотного множества в нем можно взять совокупность точек, все координаты которых рациональны и лишь в конечном числе отличны от 0.

Предположим, что $\{x(n)\}$ — фундаментальная последовательность в R^∞. Из соотношения

$$\rho_0(x_k(m), x_k(n)) \leqslant 2^k \rho(x(m), x(n))$$

легко вытекает, что при каждом k последовательность $\{x_k(1), x_k(2), \ldots\}$ фундаментальна на прямой R^1 с обычной метрикой, так что существует предел $x_k = \lim_n x_k(n)$. Исходная последовательность $x(n)$ сходится при этом к $x = (x_1, x_2, \ldots)$ в смысле R^∞. Таким образом, пространство R^∞ полно.

Подмножество A пространства R^∞ обладает компактным замыканием тогда и только тогда, когда множество

$\{x_k: x \in A\}$ при каждом k является ограниченным множеством на прямой.

Доказательство. Легко показать, что указанное условие необходимо для компактности. Достаточность этого условия мы можем доказать, используя классический диагональный метод. Исходя из заданной последовательности $\{x(n)\}$ в A, мы можем построить последовательность подпоследовательностей

$$\left\{\begin{array}{l} x(n_{11}),\ x(n_{12}),\ x(n_{13}),\ \ldots \\ x(n_{21}),\ x(n_{22}),\ x(n_{23}),\ \ldots \\ \ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots \end{array}\right. \tag{4}$$

следующим образом. Первая строка в (4) представляет собой подпоследовательность $\{x(n)\}$ такую, что существует $x_1 = \lim_i x_1(n_{1i})$; такая подпоследовательность найдется в силу того, что $\{x_1: x \in A\}$ является ограниченным множеством действительных чисел. Вторая строка в (4) представляет собой подпоследовательность первой строки, выбранную так, что существует $x_2 = \lim_i x_2(n_{2i})$; такая подпоследовательность найдется в силу того, что множество $\{x_2: x \in A\}$ ограничено.

Дальнейшие подпоследовательности строятся аналогично; k-я строка является подпоследовательностью $(k-1)$-й строки и предел $x_k = \lim_i x_k(n_{ki})$ существует. Пусть x — точка пространства R^∞ с координатами x_k. Если $n_i = n_{ii}$, то $\{x(n_i)\}$ есть подпоследовательность последовательности $\{x(n)\}$. При каждом k, кроме того, все $x(n_k)$, $x(n_{k+1})$, ... находятся в k-й строке (4), так что $\lim_i x_k(n_i) = x_k$. Таким образом, $\lim_i x(n_i) = x$, и отсюда вытекает, что A^- — компакт.

Наш последний результат относительно R^∞ является частным случаем теоремы вложения Урысона.

Любое сепарабельное метрическое пространство S гомеоморфно некоторому подмножеству R^∞.

Доказательство. Пусть $\{d_1, d_2, \ldots\}$ — последовательность точек, плотная в S. Определим отображение h из S в R^∞, положив

$$h(x) = (\rho(x, d_1),\quad \rho(x, d_2), \ldots),\quad x \in S,$$

где ρ — метрика на S. Если точки x_n пространства S сходятся к пределу x, то $\lim_n \rho(x_n, d_k) = \rho(x, d_k)$ при каждом k, так что, поскольку топология пространства R^∞ есть топология покоординатной сходимости, $h(x_n)$ сходится к $h(x)$. Таким образом, отображение h непрерывно.

Допустим, что x_n не сходится к x. Тогда для некоторого положительного ε и для любого элемента некоторой подпоследовательности $\{x_{n'}\}$ $\rho(x_{n'}, x) > \varepsilon$. Если $\rho(x, d_k) < \frac{1}{2}\varepsilon$, что должно быть справедливым для некоторого элемента плотной последовательности $\{d_k\}$, тогда $\rho(x_{n'}, d_k) > \frac{1}{2}\varepsilon$ для всех n'; таким образом, $\rho(x_n, d_k)$ не может сходиться к $\rho(x, d_k)$ при этом значении k и, следовательно, $h(x_n)$ не может сходиться к $h(x)$.

Таким образом, из $h(x_n) \to h(x)$ следует $x_n \to x$. Точно такое же рассуждение приводит к выводу, что из $h(x) = h(y)$ вытекает $x = y$ (нужно положить $x_n \equiv y$). Таким образом, h является взаимно однозначным, бинепрерывным отображением S на подмножество hS пространства R^∞.

Пространство C. Определим $C = C[0,1]$ как пространство непрерывных действительных функций на единичном интервале с равномерной метрикой. Расстояние между двумя элементами $x = x(t)$ и $y = y(t)$ пространства C равно

$$\rho(x, y) = \sup_t |x(t) - y(t)|;$$

легко проверить, что ρ — метрика. Сходимость в соответствующей топологии есть равномерная поточечная сходимость (непрерывных) функций.

Пространство C сепарабельно; одним из счетных всюду плотных множеств является множество всех многоугольных функций, которые при некотором целом k линейны на каждом подинтервале $[(i-1)/k, i/k]$, $i = 1, \ldots, k$, и принимают рациональные значения в точках i/k, $i = 0, 1, \ldots, k$.

Если $\{x_n\}$ — фундаментальная последовательность в C, то при каждом значении t последовательность $\{x_n(t)\}$ фундаментальна на прямой и, следовательно, имеет предел $x(t)$. Легко показать, что сходимость $x_n(t) \to x(t)$ равномерна по t, так что x принадлежит C и является пределом последовательности $\{x_n\}$. Таким образом, C полно.

Мы определим модуль непрерывности $w_x(\delta)$ элемента x пространства C, положив

$$w_x(\delta) = w(x, \delta) = \sup_{|s-t|<\delta} |x(s) - x(t)|, \quad 0 < \delta < 1. \tag{5}$$

Так как

$$|w_x(\delta) - w_y(\delta)| \leqslant 2\rho(x, y), \tag{6}$$

то модуль $w_x(\delta)$ при фиксированном положительном значении δ непрерывен по x. Отметим также, что, поскольку любой элемент пространства C равномерно непрерывен, то мы имеем

$$\lim_{\delta\to\infty} w_x(\delta) = 0, \qquad x \in C. \tag{7}$$

Теорема Арцела — Асколи. *Подмножество A пространства C имеет компактное замыкание тогда и только тогда, когда*

$$\sup_{x\in A} |x(0)| < \infty \tag{8}$$

и

$$\lim_{\delta\to 0} \sup_{x\in A} w_x(\delta) = 0. \tag{9}$$

Доказательство. Если A^- — компакт, то, очевидно, мы имеем (8). Поскольку модуль $w_x(1/n)$ непрерывен по x и не возрастает по n, то сходимость (7) равномерна на множестве A, если A^- — компакт (см. стр. 298), откуда следует (9).

Предположим теперь, что (8) и (9) выполняются. Выберем k столь большим, чтобы величина $\sup_{x\in A} w_x(1/k)$ была конечна. Поскольку

$$|x(t)| \leqslant |x(0)| + \sum_{i=1}^{k} \left| x\left(\frac{i}{k}t\right) - x\left(\frac{i-1}{k}t\right)\right|,$$

то отсюда следует, что

$$\sup_t \sup_{x\in A} |x(t)| < \infty. \tag{10}$$

Мы выведем из (9) и (10), что множество A вполне ограничено; отсюда будет вытекать, что замыкание A^-

компактно, так как пространство S полно. По заданному произвольному ε мы должны построить конечную ε-сеть для A. Пусть α обозначает конечную величину в формуле (10), и пусть H обозначает конечное множество точек

$$\frac{u}{v}\alpha, \quad u=0, \pm 1, \ldots, \pm v,$$

где v — целое число такое, что $\alpha/v < \varepsilon$ (H представляет собой ε-сеть для линейного интервала $[-\alpha, \alpha]$). Выберем теперь k столь большим, чтобы $w_x(1/k) < \varepsilon$ для всех x в A, и образуем множество B из таких элементов пространства C, которые линейны на каждом подинтервале $[(i-1)/k, i/k]$, $i=1, \ldots, k$, и принимают значения из H в точках i/k, $i=0, 1, \ldots, k$. Множество B конечно (оно содержит $(2v+1)^{k+1}$ точек); мы покажем, что оно является 2ε-сетью для A.

Если $x \in A$, то $|x(i/k)| \leqslant \alpha$. Поэтому существует точка y в B такая, что

$$|x(i/k) - y(i/k)| < \varepsilon, \quad i=0, 1, \ldots, k. \tag{11}$$

Так как $w_x(1/k) < \varepsilon$ и так как функция y линейна на каждом подинтервале $[(i-1)/k, i/k]$, то из (11) следует, что $\rho(x, y) < 2\varepsilon$. Это доказывает теорему.

Мы видели, что (8) и (10) эквивалентны при условии, что справедливо (9). Если (9) выполняется, то A называется равномерно равностепенно непрерывным множеством. Таким образом, теорема Арцела — Асколи утверждает, что A^- является компактом тогда и только тогда, когда A равномерно ограничено и равномерно равностепенно непрерывно.

Добавление II

РАЗНОЕ

Измеримость. Пусть $(\Omega, \mathscr{B})$ и $(\Omega', \mathscr{B}')$ — измеримые пространства; $\mathscr{B}[\mathscr{B}']$ является σ-полем подмножеств $\Omega[\Omega']$. Отображение h: $\Omega \to \Omega'$ из Ω в Ω' называется измеримым $(\mathscr{B}, \mathscr{B}')$, если прообраз $h^{-1}M'$ каждого M' из $\mathscr{B}'$ принадлежит $\mathscr{B}$. Если $h^{-1}\mathscr{B}'$ обозначает класс множеств $\{h^{-1}M'\colon M' \in \mathscr{B}'\}$, то это условие измеримости может быть кратко выражено как $h^{-1}\mathscr{B}' \subset \mathscr{B}$. Так как класс $\{M'\colon h^{-1}M' \in \mathscr{B}\}$ является σ-полем, то *если $\mathscr{B}'_0$ содержится в $\mathscr{B}'$ и порождает его, тогда из $h^{-1}\mathscr{B}'_0 \subset \mathscr{B}$ следует $h^{-1}\mathscr{B}' \subset \mathscr{B}$.*

Пусть $(\Omega'', \mathscr{B}'')$ — третье измеримое пространство, пусть j: $\Omega' \to \Omega''$ отображает Ω' в Ω''; обозначим через jh композицию отображений h и j: $(jh)(\omega) = j(h(\omega))$. Легко показать, что *если $h^{-1}\mathscr{B}' \subset \mathscr{B}$ и $j^{-1}\mathscr{B}'' \subset \mathscr{B}'$, то $(jh)^{-1}\mathscr{B}'' \subset \mathscr{B}$.*

Если $\Omega = S$ и $\Omega' = S'$ — метрические пространства, то отображение h называется непрерывным в том случае, когда для любого открытого множества G' из S' множество $h^{-1}G'$ открыто в S. Обозначим через $\mathscr{S}$ и $\mathscr{S}'$ σ-поля борелевских множеств в S и S'. Если h непрерывно, то поскольку $h^{-1}G' \subset \mathscr{S}$ для G' открытого в S' и поскольку открытые множества в S' порождают $\mathscr{S}'$, то $h^{-1}\mathscr{S}' \subset \mathscr{S}$, так что отображение h измеримо.

Если Ω' — k-мерное евклидово пространство R^k, то мы всегда берем в качестве $\mathscr{B}'$ класс $\mathscr{R}^k$ k-мерных борелевских множеств (σ-поле, порожденное открытыми подмножествами R^k с евклидовой метрикой), и мы говорим, что h $\mathscr{B}$-измеримо, если $h^{-1}\mathscr{R}^k \subset \mathscr{B}$; h $\mathscr{B}$-измеримо, если $h^{-1}\{\alpha \in R^k\colon \alpha_i < a\} \in \mathscr{B}$ для любого $i = 1, \ldots, k$ и любого действительного a. В частности, если $\Omega = S$ —

метрическое пространство, $k=1$ и h полунепрерывна сверху (см. стр. 298), то h $\mathcal{S}$-измерима.

Замена переменной. Если P — вероятностная мера на $(\Omega, \mathcal{B})$ и если $h^{-1}\mathcal{B}' \subset \mathcal{B}$, то $\mathsf{P}h^{-1}$ будет обозначать вероятностную меру на $(\Omega', \mathcal{B}')$, определяемую соотношением $(\mathsf{P}h^{-1})(M') = \mathsf{P}(h^{-1}M')$ для $M' \in \mathcal{B}'$. Если f, действительная функция на Ω', $\mathcal{B}'$-измерима, то действительная функция fh на Ω будет $\mathcal{B}$-измеримой.

При выполнении этих условий) функция f интегрируема относительно меры $\mathsf{P}h^{-1}$ тогда и только тогда, когда функция fh интегрируема относительно меры P, и в случае интегрируемости имеет место равенство*

$$\int_{h^{-1}M'} f(h(\omega))\,\mathsf{P}(d\omega) = \int_{M'} f(\omega')\,\mathsf{P}h^{-1}(d\omega') \tag{1}$$

для любого множества M' из $\mathcal{B}'$.

Если X переводит Ω в метрическое пространство S и $X^{-1}\mathcal{S} \subset \mathcal{B}$, так что X является случайным элементом S (см. § 4), мы обычно вместо $\int f(X(\omega))\,\mathsf{P}(d\omega)$ пишем $\mathsf{E}\{f(X)\}$. Если $P = \mathsf{P}X^{-1}$ есть распределение X в S, то из (1) следует, что

$$\mathsf{E}\{f(X)\} = \int_S f(x)\,P(dx). \tag{2}$$

«Хвосты» распределений. Пусть X — случайная величина на вероятностном пространстве $(\Omega, \mathcal{B}, \mathsf{P})$. Если X неотрицательна и интегрируема и если $\alpha \geqslant 0$, то

$$\int_{\{X \geqslant \alpha\}} X\,d\mathsf{P} = \alpha\mathsf{P}\{X \geqslant \alpha\} + \int_\alpha^\infty \mathsf{P}\{X \geqslant t\}\,dt. \tag{3}$$

Мы докажем этот факт, если установим равенство

$$\mathsf{E}\{X\} = \int_0^\infty \mathsf{P}\{X \geqslant t\}\,dt, \tag{4}$$

так как X в нем можно заменить произведением X и индикатора множества $\{X \geqslant \alpha\}$. Если X имеет конечную

---

*) Доказательство см., например, у Халмоша (1953, стр. 161).

область значений, то (4) получается посредством суммирования по частям. Общий результат является следствием того факта, что любая неотрицательная случайная величина X может быть представлена как предел неубывающей последовательности случайных величин с конечными областями значений. (Равенство остается справедливым и при его расширенном толковании: если одна часть (4) бесконечна, то бесконечна и другая часть.)

При тех же предположениях, что X неотрицательна и интегрируема и что $\alpha > 0$, мы имеем

$$\mathsf{P}\{X \geqslant \alpha\} \leqslant \frac{1}{\alpha} \int\limits_{\{X \geqslant \alpha\}} X\, d\mathsf{P} \leqslant \frac{1}{\alpha} \mathsf{E}\{X\}. \tag{5}$$

Теорема Шеффе. Пусть λ — мера (не обязательно конечная) на пространстве $(\Omega, \mathscr{B})$, и пусть $p(\omega)$ и $p_n(\omega)$ — плотности распределений вероятностей относительно λ; p и p_n суть неотрицательные функции на Ω, $\mathscr{B}$-измеримые и такие, что $\int p\, d\lambda = \int p_n\, d\lambda = 1$.

Теорема Шеффе*). *Если $p_n(\omega) \to p(\omega)$ всюду, кроме ω из некоторого множества λ-меры 0, то*

$$\sup_{E \in \mathscr{B}} \left| \int\limits_E p\, d\lambda - \int\limits_E p_n\, d\lambda \right| = \frac{1}{2} \int |p - p_n|\, d\lambda \to 0. \tag{6}$$

Доказательство. Положив $\delta_n = p - p_n$, имеем $\int \delta_n\, d\lambda = 0$. Следовательно, для любого множества E из $\mathscr{B}$

$$2 \left| \int\limits_E \delta_n\, d\lambda \right| = \left| \int\limits_E \delta_n\, d\lambda \right| + \left| \int\limits_{E^c} \delta_n\, d\lambda \right| \leqslant \int |\delta_n|\, d\lambda;$$

если же $E = \{\delta_n \geqslant 0\}$, то последнее неравенство становится равенством. Отсюда следует равенство в (6).

Далее, если δ_n^+ — положительная часть δ_n, то $\delta_n^+ \to 0$ всюду, за исключением множества λ-меры 0. Поскольку $0 \leqslant \delta_n^+ \leqslant p$, то из теоремы Лебега о сходимости

---

*) Шеффе (1947).

мажорируемой последовательности следует

$$\int |\delta_n| \, d\lambda = 2 \int \delta_n^+ \, d\lambda \to 0.$$

Частный случай: если $\sum_i p_n(i) = \sum_i p(i) = 1$, причем слагаемые неотрицательны, и если $\lim_n p_n(i) = p(i)$ при каждом i, то $\sum_i |p(i) - p_n(i)| \to 0$. Отсюда, если числа $a(i)$ ограничены, то $\sum_i a(i) p_n(i) \to \sum_i a(i) p(i)$.

Подпространства. Рассмотрим метрическое пространство S, и пусть $\mathcal{S}$ — σ-поле его борелевских подмножеств. Подмножество S_0 (не обязательно принадлежащее $\mathcal{S}$) само является метрическим пространством в относительной топологии. Пусть $\mathcal{S}_0$ — σ-поле борелевских множеств в S_0. Мы докажем, что

$$\mathcal{S}_0 = \{S_0 \cap A: \ A \in \mathcal{S}\}. \tag{7}$$

Если $h(x) = x$ для $x \in S_0$, то h является непрерывным отображением из S_0 в S и, следовательно, $h^{-1}\mathcal{S} \subset \mathcal{S}_0$, так что $S_0 \cap A \in \mathcal{S}_0$, если $A \in \mathcal{S}$. Поскольку $\{S_0 \cap A; \ A \in \mathcal{S}\}$ есть σ-поле в S_0, которое содержит все множества вида $S_0 \cap G$, где G — открытое множество из S (т. е. содержит все открытые множества из S_0), то мы получаем (7).

Если множество S_0 принадлежит $\mathcal{S}$, то (7) превращается в

$$\mathcal{S}_0 = \{A: \ A \subset S_0, \ A \in \mathcal{S}\}. \tag{8}$$

Произведение пространств. Рассмотрим метрические пространства S' и S'' с метриками ρ' и ρ'' и σ-полями $\mathcal{S}'$ и $\mathcal{S}''$ борелевских множеств. Прямоугольники

$$A' \times A'', \tag{9}$$

где A' — открытое множество из S' и A'' — открытое множество из S'', образуют базу топологии произведения в $S = S' \times S''$. Эта топология может быть также описана как топология, в которой сходимость $(x'_n, x''_n) \to (x', x'')$ имеет место тогда и только тогда, когда $x'_n \to x'$ и $x''_n \to x''$. Наконец, эта топология порождается различными метриками, как, например,

$$\rho((x', x''), (y', y'')) = \sqrt{[\rho'(x', y')]^2 + [\rho''(x'', y'')]^2} \tag{10}$$

и

$$\rho((x', x''), (y', y'')) = \max\{\rho'(x', y'), \rho''(x'', y'')\}. \tag{11}$$

Пусть $\mathcal{S}' \times \mathcal{S}''$ — σ-поле, порожденное измеримыми прямоугольниками (множествами (9) с $A' \in \mathcal{S}'$ и $A'' \in \mathcal{S}''$), и пусть $\mathcal{S}$ — σ-поле борелевских множеств в S относительно топологии произведения. Мы покажем сначала, что

$$\mathcal{S}' \times \mathcal{S}'' \subset \mathcal{S}. \tag{12}$$

Положим $\pi'(x', x'') = x'$ и $\pi''(x', x'') = x''$. Оба отображения π': $S \to S'$ и π'': $S \to S''$ непрерывны и, следовательно, измеримы $((\pi')^{-1}\mathcal{S}' \subset \mathcal{S}$ и $(\pi'')^{-1}\mathcal{S}'' \subset \mathcal{S})$. Отсюда следует, что если $A' \in \mathcal{S}'$ и $A'' \in \mathcal{S}''$, то $A' \times A'' = \pi'^{-1}A' \cap \pi''^{-1}A''$ принадлежит S. Так как $\mathcal{S}$ содержит все измеримые прямоугольники, то мы получаем включение (12).

Далее, пространство S сепарабельно тогда и только тогда, когда пространства S' и S'' оба сепарабельны. Покажем, что если S сепарабельно, то

$$\mathcal{S}' \times \mathcal{S}'' = \mathcal{S}. \tag{13}$$

Ввиду (12) достаточно показать, что если G открыто в S, то $G \in \mathcal{S}' \times \mathcal{S}''$. Но G можно представить в виде объединения прямоугольников (9) с открытым $A' \subset S'$ и открытым $A'' \subset S''$ (так что $A' \times A'' \in \mathcal{S}' \times \mathcal{S}''$), и если S сепарабельно, то G является счетным таким объединением. Это доказывает (13) *).

Допустим, что X' и X'' — случайные элементы S' и S'', соответственно, и имеют общую область определения $(\Omega, \mathcal{B})$. Тогда $(X'(\omega), X''(\omega))$ определяет отображение (X', X'') из Ω в $S = S' \times S''$, и очевидно, что $(X', X'')^{-1}(\mathcal{S}' \times \mathcal{S}'') \subset \mathcal{B}$. Если пространство S сепара-

*) Без предположения о сепарабельности равенство (13) может не выполняться: если $S' = S''$ — дискретное пространство, мощность которого превышает мощность континуума, то диагональ $\{(x, y)\colon x = y\}$ принадлежит $\mathcal{S}$, но не принадлежит $\mathcal{S}' \times \mathcal{S}''$ (см. задачу 2 на стр. 253 книги Халмоша (1953)).

бельно, то в силу (13) (X', X'') представляет собой случайный элемент S.

Если $S' = S''$, то $\rho'(x', y')$ определяет непрерывное отображение из $S' \times S'$ в R^1; если S' сепарабельно, то композиция $\rho(X', X'')$ измерима и, следовательно, является случайной величиной *).

Измеримость D_h. Пусть h отображает метрическое пространство S в другое метрическое пространство S', и пусть D_h — множество точек разрыва. Обозначим метрики в S и S' через ρ и ρ', соответственно. Обозначим через $A_{\varepsilon, \delta}$ множество точек x в S, для которых существуют точки y и z в S, удовлетворяющие неравенствам $\rho(x, y) < \delta$, $\rho(x, z) < \delta$ и $\rho'(hy, hz) \geqslant \varepsilon$. Множество $A_{\varepsilon, \delta}$ открыто. Так как

$$D_h = \bigcup_{\varepsilon} \bigcap_{\delta} A_{\varepsilon, \delta},$$

где ε и δ принимают лишь положительные рациональные значения, то D_h является борелевским множеством в S. Этот факт верен, даже если h неизмеримо (например, если h является индикатором множества A, то $D_h = \partial A$ замкнуто независимо от того, каково A).

Рассмотрим теперь множество E, фигурирующее в теореме 5.5. Мы будем предполагать, что отображение h измеримо и что S' сепарабельно, и докажем, что $E \subset \mathscr{S}$. Если $B_{\varepsilon, \delta, i}$ — множество таких x, что $\rho'(hx, hy) \geqslant \varepsilon$ для некоторой точки y, удовлетворяющей неравенству $\rho(x, y) < \delta$, то

$$E = \bigcup_{\varepsilon} \bigcap_{\delta} \bigcap_{k \geqslant 1} \bigcup_{i \geqslant k} B_{\varepsilon, \delta, i},$$

где ε и δ пробегают положительные рациональные значения. Достаточно, таким образом, найти множества $C_{\varepsilon, \delta, i}$, которые принадлежат $\mathscr{S}$ и удовлетворяют условию

$$B_{\varepsilon, \delta, i} \subset C_{\varepsilon, \delta, i} \subset B_{\frac{1}{2}\varepsilon, \delta, i}.$$

Если последовательность u_m плотна в S' и если $H_{\varepsilon, m} = \left\{x\colon \rho'(hx, u_m) < \frac{1}{4}\varepsilon\right\}$, то $H_{\varepsilon, m} \in \mathscr{S}$ и $S = \bigcup_m H_{\varepsilon, m}$.

---

*) Противоречащий пример в предыдущей сноске показывает, что это может быть неверно, если S' не является сепарабельным.

Требованиям, наложенным на $C_{\varepsilon, \delta, i}$, удовлетворяют множества

$$C_{\varepsilon, \delta, i} = \bigcup_m (H_{\varepsilon, m} \cap J_{\varepsilon, \delta, i, m}),$$

где $J_{\varepsilon, \delta, i, m}$ обозначает множество таких x, что $\rho'(hy, hz) \geqslant \varepsilon$ для некоторой пары точек y и z, удовлетворяющих условиям $\rho(x, y) < \delta$, $\rho(x, z) < \delta$ и $z \in H_{\varepsilon, m}$ ($J_{\varepsilon, \delta, i, m}$ — открытое множество). Отсюда вытекает, что $E \in \mathcal{S}$ *). Доказательство теоремы 5.5 остается в силе даже в том случае, когда E не принадлежит $\mathcal{S}$, но имеет внешнюю P-меру 0.

Теорема Хелли. Функцией распределения называется функция $F(x) = F(x_1, \ldots, x_k)$, заданная на R^k, со следующими тремя свойствами (мы используем терминологию и обозначения § 3):

(i) F всюду непрерывна справа;

(ii) $0 \leqslant F(x) \leqslant 1$ для всех x, F не убывает по каждой переменной и для любого k-мерного прямоугольника $(a, b]$

$$\sum \pm F(a_1 + \theta_1 d_1, \ldots, a_k + \theta_k d_k) \geqslant 0, \tag{14}$$

где $d_i = b_i - a_i$, суммирование производится по всем 2^k последовательностям $(\theta_1, \ldots, \theta_k)$ из нулей и единиц и знак $+$ или $-$ ставится в зависимости от того, является ли число нулей в последовательности четным или нечетным;

(iii) $F(x) \to 0$, если какая-нибудь из координат x стремится к $-\infty$, и $F(x) \to 1$, если все координаты x стремятся к ∞.

Если P — вероятностная мера на $(R^k, \mathcal{R}^k)$ и

$$F(x) = P\{y\colon y \leqslant x\}, \quad x \in R^k, \tag{15}$$

то F является функцией распределения. С другой стороны, если F — функция распределения, то существует ровно одна вероятностная мера P, удовлетворяющая соотношению (15) (это известный факт в теории функций действительной переменной). Если F обладает вышеуказанными свойствами (i) и (ii) (но, возможно, не обла-

*) Это доказательство принадлежит Ф. Топсё.

дает свойством (iii)), то существует конечная мера μ на $(R^k, \mathscr{R}^k)$ такая, что

$$F(x) = \mu\{y: y \leqslant x\}, \qquad x \in R^k; \tag{16}$$

при этом мера μ удовлетворяет условию $\mu(R^k) \leqslant 1$, но не обязана быть вероятностной мерой.

Теорема Хелли о выборе. *Если $\{F_n\}$ — последовательность функций распределения на R^k, то существуют подпоследовательность $\{F_{n'}\}$ и функция F, удовлетворяющая вышеуказанным свойствам* (i) *и* (ii) (*но, возможно, не удовлетворяющая свойству* (iii)) *такие, что*

$$\lim_{n'} F_{n'}(x) = F(x) \tag{17}$$

для всех точек x непрерывности функции F.

Доказательство. Обозначим через R_0^k множество рациональных точек в R^k (т. е. множество точек, координаты которых рациональны), и пусть $R_0^k = \{r(1), r(2), \ldots\}$. Для каждой функции F_n из заданной последовательности функций распределения последовательность

$$(F_n(r(1)),\ F_n(r(2)),\ \ldots) \tag{18}$$

представляет собой точку пространства R^∞. Так как $0 \leqslant F_n(x) \leqslant 1$, то из критерия компактности для R^∞ (см. стр. 299) следует, что некоторая подпоследовательность точек (18) сходится в смысле R^∞ к некоторой точке $(z_1, z_2, \ldots)$ пространства R^∞. Определим функцию F_0 на R_0^k равенством $F_0(r(k)) = z_k$. Мы видим, стало быть, что существуют функция F_0 на R_0^k и подпоследовательность $\{F_{n'}\}$ последовательности $\{F_n\}$ такие, что

$$\lim_{n'} F_{n'}(r) = F_0(r), \qquad r \in R_0^k. \tag{19}$$

Так как каждая функция F_n является функцией распределения, то из (19) следует, что F_0 удовлетворяет условию (ii) на R_0^k: $0 \leqslant F_0(r) \leqslant 1$, F_0 — не возрастает, когда каждая координата пробегает рациональные значения, и если a и b принадлежат R_0^k, то (14) выполняется. Если мы определим F на R^k равенством

$$F(x) = \inf\{F_0(r): x < r,\ r \in R_0^k\}, \qquad x \in R^k,$$

то F обладает свойствами (i) и (ii) (но не обязательно свойством (iii)).

Если F непрерывна в точке x, то, задаваясь произвольным $\varepsilon > 0$, мы можем найти в R_0^k такие точки r' и r'', что $r' < x < r''$ и

$$F(x) - \varepsilon < F_0(r') \leqslant F_0(r'') < F(x) + \varepsilon.$$

При каждом n'

$$F_{n'}(r') \leqslant F_{n'}(x) \leqslant F_{n'}(r'').$$

Из этих соотношений и (19) следует, что

$$F(x) - \varepsilon \leqslant \liminf_{n'} F_{n'}(x) \leqslant \limsup_{n'} F_{n'}(x) \leqslant F(x) + \varepsilon.$$

Так как ε здесь произвольно, то (17), следовательно, выполняется. Это доказывает теорему Хелли.

Теорема Колмогорова. Мы докажем теорему Колмогорова (или Даниеля — Колмогорова) о существовании меры. Определим проекцию π_k из R^∞ в R^k равенством $\pi_k(x) = (x_1, \ldots, x_k)$, как описано в § 3. Определим также при $k > 1$ проекцию ψ_k из R^k в R^{k-1} равенством $\psi_k(x_1, \ldots, x_k) = (x_1, \ldots, x_{k-1})$. Поскольку отображения π_k и ψ_k непрерывны, они измеримы $\left(\pi_k^{-1}\mathcal{R}^k \subset \mathcal{R}^\infty\right.$ и $\left.\psi_k^{-1}\mathcal{R}^{k-1} \subset \mathcal{R}^k\right)$.

Пусть P — вероятностная мера на $(R^\infty, \mathcal{R}^\infty)$, определим вероятностные меры μ_k на $(R^k, \mathcal{R}^k)$, полагая

$$\mu_k = P\pi_k^{-1}. \tag{20}$$

Так как

$$\pi_{k-1} = \psi_k\pi_k, \qquad k > 1, \tag{21}$$

то меры μ_k удовлетворяют соотношению

$$\mu_{k-1} = \mu_k\psi_k^{-1}, \qquad k > 1. \tag{22}$$

Задача заключается в том, чтобы пройти обратным путем и построить по заданным мерам μ_k, удовлетворяющим условиям согласованности (22), меру P, удовлетворяющую условию (20).

Теорема Колмогорова о существовании меры. *Если вероятностные меры μ_k на $(R^k, \mathcal{R}^k)$, $k \geqslant 1$, удовлетворяют условию* (22), *то на $(R^\infty, \mathcal{R}^\infty)$ существует единственная вероятностная мера P, удовлетворяющая условию* (20).

Доказательство. Определим при $k > i \geqslant 1$ непрерывное отображение $\psi_{k,i}$ из R^k в R^i равенством $\psi_{k,i}(x_1, \ldots, x_k) = (x_1, \ldots, x_i)$. Заметим, что $\psi_{k,k-1} = \psi_k$ и что

$$\psi_{k,i} = \psi_{i+1} \ldots \psi_{k+1}\psi_k, \quad \pi_i = \psi_{k,i}\pi_k. \tag{23}$$

Отсюда и из (22) следует, что

$$\mu_i = \mu_k\psi_{k,i}^{-1}, \quad k > i \geqslant 1. \tag{24}$$

Пусть $\mathcal{F}$ — совокупность конечномерных множеств, определенных так же, как в § 3; класс $\mathcal{F}$ состоит из множеств вида

$$A = \pi_i^{-1}H, \qquad H \in \mathcal{R}^i, \tag{25}$$

и является полем, порождающим σ-поле $\mathcal{R}^\infty$. Множество вида (25) можно также представить в форме

$$A = \pi_k^{-1}H', \quad H' \in \mathcal{R}^k, \tag{26}$$

при $k > i$; чтобы в этом убедиться, достаточно положить $H' = \psi_{k,i}^{-1}H$ и использовать (23).

Допустим теперь, что множество A задано как соотношением (25), так и соотношением (26), где $i < k$. Тогда $\alpha = (\alpha_1, \ldots, \alpha_k)$ лежит в H' в том и только том случае, когда $(\alpha_1, \ldots, \alpha_k, 0, 0, \ldots)$ лежит в $\pi_k^{-1}H' = \pi_i^{-1}H$, что в свою очередь справедливо в том и только том случае, когда $(\alpha_1, \ldots, \alpha_i)$ лежит в H. Таким образом, $H' = \psi_{k,i}^{-1}H$ и из (24) вытекает, что

$$\mu_i H = \mu_k H'. \tag{27}$$

Поскольку (25) и (26) вместе приводят к (27), мы можем согласованно определить на $\mathcal{F}$ функцию P, полагая $P(A) = \mu_i(H)$, если A задано равенством (25). Очевидно, $P(0) = 0$, $P(R^\infty) = 1$ и $0 \leqslant P(A) \leqslant 1$. Предположим, что A задано соотношением (25), а B — соотношением

$$B = \pi_k^{-1}J, \qquad J \in \mathcal{R}^k,$$

причем $k \geqslant i$. Множество A можно представить также в виде (26), и если $A \cap B = 0$, то тогда $H' \cap J = 0$, так что

$$P(A \cup B) = \mu_k(H' \cup J) = \mu_k(H') + \mu_k(J) = P(A) + P(B).$$

Таким образом, P является конечно аддитивной вероятностной мерой на $\mathcal{F}$.

Докажем, что P вполне аддитивна на $\mathcal{F}$. Для этого мы покажем, что если $A_k \in \mathcal{F}$,

$$A_1 \supset A_2 \supset \ldots \tag{28}$$

и $\bigcap_i A_i = 0$, то $\lim_i P(A_i) = 0$. Поскольку A_i принадлежит $\mathcal{F}$, оно имеет вид

$$A_i = \pi_{n_i}^{-1} H_i, \qquad H_i \in \mathcal{R}^{n_i}. \tag{29}$$

Так как множество вида (25) может быть представлено в виде (26) при любом k, превышающем i, то мы не ограничим общность, предположив, что

$$n_1 < n_2 < n_3 < \ldots \tag{30}$$

Мы покажем, что если множества (29) удовлетворяют соотношениям (28) и (30) и если существует положительное ε такое, что

$$P(A_i) > \varepsilon, \quad i = 1, 2, \ldots, \tag{31}$$

то $\bigcap_i A_i \neq 0$.

Из (31) и (29) в силу определения P получаем, что $\mu_{n_i}(H_i) > \varepsilon$. По теоремам 1.1 и 1.4 в R^{n_i} существует компактное подмножество K_i множества H_i такое, что $\mu_{n_i}(H_i - K_i) < \varepsilon/2^{i+1}$. Если $B_i = \pi_{n_i}^{-1} K_i$, то $B_i \in \mathcal{F}$, $B_i \subset A_i$ и

$$P(A_i - B_i) < \frac{\varepsilon}{2^{i+1}}.$$

Отсюда и из (28) следует, что $C_i = B_1 \cap \ldots \cap B_i$ есть подмножество A_i, удовлетворяющее соотношению $P(A_i - C_i) \leqslant \sum_{j=1}^{i} P(A_j - B_j) < \varepsilon/2$. В силу (31) поэтому $P(C_i) > \varepsilon/2$, так что C_i непусто.

Мы построили непустые множества $C_i \subset \pi_{n_i}^{-1} K_i$, где K_i — компакт и

$$C_1 \supset C_2 \supset \ldots \tag{32}$$

Поскольку $C_i \subset A_i$, то если мы найдем точку, общую для всех C_i, то получим $\bigcap_i A_i \neq 0$. Пусть $x(j)$ — произвольный элемент C_j. Если $j \geqslant i$, то в силу (32) $x(j) \in C_i$ и, следовательно, $\pi_{n_i} x(j) \in K_i$. Так как K_i — компакт, то мы имеем $\sup_{j \geqslant i} |x_i(j)| < \infty$ и, следовательно, $\sup_j |x_i(j)| < < \infty$. Поэтому при каждом i последовательность $\{x_i(1), x_i(2), \ldots\}$ представляет собой ограниченную последовательность действительных чисел. Следовательно (см. стр. 299), некоторая подпоследовательность $\{x(j')\}$ последовательности $\{x(j)\}$ сходится в смысле R^∞ к некоторому пределу $x \in R^\infty$. Так как $x(j') \in C_{j'}$ и каждое множество C_i замкнуто, то из (32) следует, что x принадлежит $\bigcap_i C_i$.

Мы показали, что P — вполне аддитивная вероятностная мера на конечно аддитивном поле $\mathcal{F}$. Так как $\mathcal{F}$ порождает $\mathcal{R}^\infty$, то P может быть продолжена*) до вероятностной меры на R^∞. Очевидно, P удовлетворяет условию (20). Поскольку $\mathcal{F}$ является определяющим классом, то может существовать лишь одна мера P, удовлетворяющая условию (20), что доказывает теорему Колмогорова.

Нетрудно доказать и более общий вариант этой теоремы. Пусть R^T — пространство всех действительных функций $x = x(t)$, заданных на произвольном множестве T, которое мы можем также считать бесконечным. (Если T состоит из целых чисел, то $R^T = R^\infty$.) Для конечного упорядоченного множества $\sigma = (s_1, \ldots, s_{k_\sigma})$ различных элементов T определим отображение π_σ: $R^T \to R^{k_\sigma}$ равенством $\pi_\sigma(x) = (x(s_1), \ldots, x(s_{k_\sigma}))$; пусть $\mathcal{R}^T$ — σ-поле, порожденное множествами $\pi_\sigma^{-1} H$, где $H \in \mathcal{R}^{k_\sigma}$ (для всех σ и H). Пара $(R^T, \mathcal{R}^T)$ является измеримым пространством. Никакая топология не участвует в этой конструкции (и вообще здесь нет необходимости в топологии).

Совокупность вероятностных мер μ_σ на $(R^{k_\sigma}, \mathcal{R}^{k_\sigma})$ (одна мера для каждого σ) называется согласованной, если $\mu_{\sigma_2} = \mu_{\sigma_1} \psi^{-1}$ всякий раз, когда $\sigma_2 = (s_{i_1}, \ldots, s_{i_j})$

*) Халмош (1953, гл. 3).

есть перестановка j элементов множества $\sigma_1 = (s_1, \ldots, s_k)$ (здесь $j \leqslant k$) и отображение ψ: $R^k \to R^j$ определяется равенством $\psi(x_1, \ldots, x_k) = (x_{i_1}, \ldots, x_{i_j})$.

Если меры μ_σ являются согласованными, то существует единственная вероятностная мера P на $(R^T, \mathscr{R}^T)$ такая, что $P\pi_\sigma^{-1} = \mu_\sigma$ для всех σ.

Чтобы доказать это, определим для последовательности $\tau = (t_1, t_2, \ldots)$ элементов T отображение π_τ: $R^T \to R^\infty$ равенством $\pi_\tau(x) = (x(t_1), x(t_2), \ldots)$. Из теоремы существования, доказанной для пространства R^∞, следует, что на $(R^\infty, \mathscr{R}^\infty)$ существует вероятностная мера P_τ такая, что $P_\tau \pi_k^{-1} = \mu_{t_1 \ldots t_k}$ для любого k.

Класс множеств $\pi_\tau^{-1}H$, где $H \in \mathscr{R}^\infty$ (для всех τ и H), в точности совпадает с классом $\mathscr{R}^T$ (а не только порождает его), и отсюда легко выводится, что равенство $P(\pi_\tau^{-1}H) = P_\tau(H)$ согласованно определяет желаемую меру P.

Измеримость некоторых отображений. Пусть T обозначает единичный интервал $[0, 1]$, и пусть $\mathscr{T}$ — класс линейных борелевских подмножеств T. При каждом t проекция π_t из C в R^1 измерима относительно $\mathscr{C}$. Поскольку отображение

$$(x, t) \to x(t) \tag{33}$$

из $C \times T$ в R^1 непрерывно в топологии произведения и поскольку $\mathscr{C} \times \mathscr{T}$ есть σ-поле борелевских множеств в этой топологии (см. стр. 308), то (33) измеримо относительно $\mathscr{C} \times \mathscr{T}$.

Пусть $x \in C$. Определим $h(x)$ как меру Лебега множества тех t из T, для которых $x(t) > 0$. Мы хотим доказать, что функция h измерима относительно $\mathscr{C}$, и мы получим это из более общего результата. Определим действительную функцию v действительной переменной, полагая

$$v(\alpha) = \begin{cases} 1, & \text{если } \alpha > 0, \\ 0, & \text{если } \alpha \leqslant 0. \end{cases} \tag{34}$$

Тогда

$$h(x) = \int_0^1 v(x(t))\, dt. \tag{35}$$

Очевидно, что функция v: (i) измерима по Борелю, (ii) ограничена и (iii) непрерывна всюду, за исключением некоторого множества лебеговой меры 0. Используя эти три свойства функции v, мы покажем, что функция h, определенная в (35), $\mathscr{C}$-измерима и непрерывна всюду, за исключением некоторого множества винеровской меры 0.

Поскольку для каждого элемента x из C функция $v(x(t))$ является ограниченной измеримой функцией аргумента t, интеграл в (35) имеет смысл. Так как отображение (33) измеримо относительно $\mathscr{C} \times \mathscr{T}$ и так как функция v измерима, то отображение ψ: $C \times T \to R^1$, заданное равенством $\psi(x, t) = v(x(t))$, также измеримо. Поскольку отображение ψ ограничено, то функция $h(x) = \int\limits_0^1 \psi(x, t)\,dt$ измерима по x *). Таким образом, функция h измерима относительно $\mathscr{C}$.

Если D_v — множество разрывов v, то в силу (iii) $\lambda(D_v) = 0$, где λ обозначает лебегову меру на прямой. Пусть E — множество пар (x, t) таких, что $x(t) \in D_v$. Для каждого положительного t

$$W\{x\colon (x, t) \in E\} = \frac{1}{\sqrt{2\pi t}} \int\limits_{D_v} e^{-\frac{1}{2}u^2}\,du = 0.$$

Отсюда и из теоремы Фубини применительно к мере $W \times \lambda$ на $\mathscr{C} \times \mathscr{T}$ вытекает, что $\lambda\{t\colon (x, t) \in E\} = 0$ при $x \notin A$, где A — множество из $\mathscr{C}$ такое, что $W(A) = 0$. Далее, если x_n сходится к x поточечно, то $v(x_n(t)) \to v(x(t))$ при каждом t таком, что $x(t) \notin D_v$. Если $x \notin A$, то это справедливо для почти всех t, и следовательно, по теореме о сходимости мажорируемой последовательности

$$\int\limits_0^1 v(x_n(t))\,dt \to \int\limits_0^1 v(x(t))\,dt.$$

Тем самым доказано, что h непрерывна всюду, за исключением точек, образующих множество W-меры 0.

---

*) См. Халмош (1953, гл. 7).

Рассуждение остается справедливым и тогда, когда W заменяется произвольной мерой P с тем свойством, что $P\pi_t^{-1}$ абсолютно непрерывна относительно меры Лебега для почти всех t. Например, это верно для W°.

За исключением доказательства измеримости отображения (33), это рассуждение слово в слово сохраняется в том случае, когда вместо пространства C мы рассматриваем D. Измеримость отображения (33) можно при этом доказать, соответствующим образом видоизменяя доказательство измеримости отображения π_t в § 14.

Остается рассмотреть еще одну функцию, определенную на C, а именно супремум $h(x)$ тех t из $[0, 1]$, для которых $x(t) = 0$. Поскольку множество $\{x: h(x) < \alpha\}$ открыто, то функция h измерима. Если h разрывна в точке x, то $x(t)$ должна оставаться по одну сторону от 0 при $t \in (h(x), 1)$ и должна оставаться по ту же самую сторону от 0 в интервале $(h(x) - \varepsilon, h(x))$ при некотором ε. Непрерывность всюду, кроме множества винеровской меры 0, будет поэтому доказана, если мы покажем, что при любом t_0 верхняя и нижняя грань W на $[t_0, 1]$ имеют непрерывные распределения. Так как случайная функция $W_t - W_{t_0}$, $t \in [t_0, 1]$, распределена как винеровская траектория с линейно преобразованной шкалой времени, то $-W_{t_0} + \sup\limits_{t \geqslant t_0} W_t$ имеет непрерывное распределение (см. (10.17)). Последняя случайная величина и W_{t_0} независимы, и следовательно, их сумма также имеет непрерывное распределение. Нижняя грань рассматривается тем же способом.

Еще об измеримости. В § 17 мы определили отображение ψ: $D \times D_0 \to D$, положив $\psi(x, \varphi) = x \circ \varphi$ (см. (17.5) и (17.6)). Нам нужно доказать, что оно ψ-измеримо ($\psi^{-1}\mathscr{D} \subset \mathscr{D} \times \mathscr{D}_0$). Так как конечномерные множества в D порождают $\mathscr{D}$, то достаточно доказать, что при каждом t отображение

$$(x, \varphi) \to \pi_t(x \circ \varphi) = x(\varphi(t)) \tag{36}$$

измеримо относительно $\mathscr{D} \times \mathscr{D}_0$. Если $\varphi_k(t)$ есть наименьшее отношение i/k, которое больше или равно $\varphi(t)$, то $\varphi_k(t) \downarrow \varphi(t)$ для каждого t. Следовательно, отображение

$$(x, \varphi) \to x(\varphi_k(t)) \tag{37}$$

сходится поточечно к (36), и достаточно, стало быть, доказать, что это последнее отображение измеримо относительно $\mathscr{D} \times \mathscr{D}_0$. Множество $\{(x, \varphi): x(\varphi_k(t)) \leqslant \alpha\}$ является объединением множества

$$\{(x, \varphi): \varphi(t) = 0\} \cap \{(x, \varphi): x(0) \leqslant \alpha\} \tag{38}$$

с множествами

$$\left\{(x, \varphi): \frac{i-1}{k} < \varphi(t) \leqslant \frac{i}{k}\right\} \cap \left\{(x, \varphi): x\left(\frac{i}{k}\right) \leqslant \alpha\right\}, \tag{39}$$
$$i = 1, \ldots, k.$$

Если $H \in \mathscr{R}^1$, то $\{\varphi \in D_0: \varphi(t) \in H\}$ является пересечением D_0 и подмножества $\pi_t^{-1}H$ пространства D, и, следовательно, принадлежит $\mathscr{D}_0$. Поэтому $\{(x, \varphi): \varphi(t) \in H\} \in \mathscr{D} \times \mathscr{D}_0$. Аналогично, $\{(x, \varphi): x(t) \in H\} \in \mathscr{D} \times \mathscr{D}_0$. Таким образом, все множества (38) и (39) принадлежат $\mathscr{D} \times \mathscr{D}_0$, что и доказывает измеримость (37).

Добавление III

ТЕОРЕТИЧЕСКИЕ ДОПОЛНЕНИЯ

Это Добавление, использующее в своей бóльшей части понятия общей (неметрической) топологии, посвящено вопросам, которые естественным образом возникают в связи с теорией слабой сходимости, но не имеют отношения к ее приложениям.

Проблема меры. Существует ли определенная на классе всех подмножеств данного множества U вероятностная мера, которая не имеет атомов (которая приписывает меру 0 каждой отдельной точке)? Ответ, разумеется, зависит только от мощности U. Очевидно, что такая мера не может существовать, если U счетно; в предположении, что справедлива континуум-гипотеза, можно показать*), что такая мера не существует, если U имеет мощность континуума.

Если неатомическая вероятностная мера, определенная на классе всех подмножеств U, на самом деле существует, то кардинальное число U называется *измеримым*; в противном случае кардинальное число U называется *неизмеримым*. Таким образом, число $\aleph_0$ неизмеримо, и из континуум-гипотезы следует, что мощность континуума также неизмерима. (Среди множеств на прямой именно неизмеримые множества представляют собой отклонения от нормы. Здесь терминология меняется на обратную: среди кардинальных чисел именно измеримые кардинальные числа являются отклоняющимися от нормы.)

Существование измеримых кардинальных чисел — это знаменитая нерешенная проблема теории множеств,

*) См. Биркгоф (1952).

так называемая проблема меры*). Если измеримые кардинальные числа вообще существуют, то они должны быть так велики, чтобы никогда не возникать в математике естественным образом. Мы увидим, что некоторые иррегулярные явления в теории слабой сходимости столь же далеки от обычной математической практики, так как они могут возникать лишь в том случае, если существуют измеримые кардинальные числа.

Сепарабельные меры. Рассмотрим теперь вероятностную меру P, заданную на классе $\mathcal{S}$ борелевских множеств в метрическом пространстве S. Согласно теореме 1.4 мера P является плотной, если S сепарабельно и полно. Разумеется, требование полноты может быть ослаблено до требования топологической полноты (пространство S топологически полно, если существует эквивалентная метрика, относительно которой S полно в обычном смысле), а анализ доказательства этой теоремы показывает, что предположение о сепарабельности может быть ослаблено до предположения о том, что P обладает сепарабельным носителем. Будем называть саму меру P *сепарабельной*, если она обладает сепарабельным носителем**). Мы приходим, таким образом, к следующему результату:

Теорема 1. *Вероятностная мера P плотна, если она сепарабельна и если пространство S топологически полно.*

Примечание 1. Предположение теоремы о сепарабельности P не может быть исключено. Действительно, если P плотна, то она обладает σ-компактным носителем и, следовательно, сепарабельна. Но, разумеется, само пространство S не обязано быть сепарабельным.

Примечание 2. Предположение о топологической полноте также не может быть исключено. В самом деле, пусть S — массивное***) подмножество отрезка $[0, 1]$ (множество, внутренняя мера Лебега которого $\lambda_*(S)$

---

*) В связи с проблемой меры см. работу Кейслера и Тарского (1964), где содержится также обширная библиография.

**) Это понятие является специализацией к метрическому случаю более общего понятия τ-*гладкой* меры; см. Варадарайн (1958а и 1961а).

***) В оригинале «thick». См. также Халмош (1953, стр. 78). (*Прим. перев.*)

равна 0, а внешняя мера Лебега $\lambda^*(S)$ равна 1) с относительной топологией прямой. Здесь $\mathcal{S}$ состоит из множеств вида $S \cap A$, где A — обычное линейное борелевское множество (см. формулу (7) на стр. 307). Если P получена из λ^* сужением области определения до $\mathcal{S}$, то P вполне аддитивна; более того, так как $\lambda^*(S) = 1$, то P — вероятностная мера. Если K — компакт в S, то K является обычным борелевским множеством, и поэтому $P(K) = 0$, так как $\lambda_*(S) = 0$. Следовательно, вероятностная мера P не плотна. Так как P сепарабельна (пространство S само сепарабельно), то теорема 1 перестает быть верной именно из-за нарушения предположения о полноте.

Примечание 3. Если S — множество рациональных чисел с относительной топологией прямой, то ввиду того, что S представляет собой σ-компакт, любая мера P на S плотна. С другой стороны, по теореме Бэра о категории *) S не является топологически полным множеством. Проблема топологической характеризации тех метрических пространств, в которых все вероятностные меры плотны, является до сих пор открытой.

Теперь возникает следующий вопрос — существуют ли несепарабельные вероятностные меры?

Теорема 2. *Любая вероятностная мера на $\mathcal{S}$ сепарабельна тогда и только тогда, когда любое дискретное* **) *подмножество S имеет неизмеримое кардинальное число.*

Доказательство. Докажем сначала необходимость. Предположим, что S содержит дискретное подмножество A_0 с измеримым кардинальным числом. Мы построим на $\mathcal{S}$ несепарабельную меру P. Обозначим через Q неатомическую вероятностную меру, заданную на классе всех подмножеств A_0, и определим равенством $P(A) = Q(A \cap A_0)$ неатомическую вероятностную меру P на $\mathcal{S}$. Если A сепарабельно, то множество $A \cap A_0$ является одновременно и дискретным, и сепарабельным и, следовательно, является (стр. 295) счетным, так что $Q(A \cap A_0) = 0$. Таким образом, P не может обладать сепарабельным носителем.

---

*) См. Келли (1968, стр. 267).

**) См. стр. 295.

Доказательство достаточности требует более глубокого рассуждения. Предположим, что любое дискретное подмножество S имеет неизмеримое кардинальное число, и рассмотрим произвольную вероятностную меру P на $\mathcal{S}$. Мы должны показать, что P сепарабельно. Заметим сначала, что для этого достаточно показать, что каждое открытое покрытие $\mathcal{G}$ пространства S содержит счетный подкласс $\{G_1, G_2, \ldots\}$ такой, что

$$P\left(\bigcup_n G_n\right) = 1. \tag{1}$$

Действительно, в этом случае для каждого k найдется последовательность $A_{k1}, A_{k2}, \ldots$ открытых $1/k$-сфер таких, что $P\left(\bigcup_n A_{kn}\right) = 1$. Множество $\bigcap_k \bigcup_n A_{kn}$ является при этом сепарабельным носителем для P.

Рассмотрим теперь открытое покрытие $\mathcal{G}$ пространства S. По теореме о паракомпактности*) существует класс $\mathcal{H}$ со следующими свойствами: (i) $\mathcal{H}$ есть открытое покрытие S, (ii) каждый элемент класса $\mathcal{H}$ содержится в некотором элементе $\mathcal{G}$, (iii) класс $\mathcal{H}$ может быть представлен как счетное объединение

$$\mathcal{H} = \bigcup_n \mathcal{H}_n, \tag{2}$$

где для каждого n

$$\inf\{\rho(A, B)\colon\ A, B \in \mathcal{H}_n,\ A \neq B\} > 0 \tag{3}$$

(здесь ρ — метрика в пространстве S).

Зафиксируем на время n. Из каждого элемента класса $\mathcal{H}_n$ выберем по одной точке и образуем, таким образом, множество S_n. Вследствие (3) S_n дискретно. Для $A \subset S_n$ обозначим через $\lambda_n(A)$ P-меру объединения всех тех элементов класса $\mathcal{H}_n$, которые имеют с A непустое пересечение (это объединение открыто). Так как элементы класса $\mathcal{H}_n$ не пересекаются, то λ_n является конечной мерой, определенной на классе всех подмножеств S_n. Поскольку, по предположению, S_n имеет неизмеримое кардинальное число, то λ_n обладает счетным носителем. (В противном случае мы могли бы вычесть атомическую часть λ_n и получили бы на классе всех подмножеств S_n

*) См. Келли (1968, стр. 212).

неатомическую меру ν_n такую, что $0 < \nu_n(S_n) < \infty$, которую затем можно было бы нормировать так, чтобы получить вероятностную меру.) Пусть A_n — счетное подмножество S_n, для которого

$$\lambda_n(S_n - A_n) = 0, \tag{4}$$

и пусть класс $\mathcal{J}_n$ состоит из тех элементов $\mathcal{H}_n$, которые имеют непустое пересечение с A_n. Положим $\mathcal{J} = \bigcup\limits_n \mathcal{J}_n$. Так как каждый класс $\mathcal{J}_n$ является счетным, то таковым является и класс $\mathcal{J}$.

Если H_n — объединение множеств из $\mathcal{H}_n$, а J_n — объединение множеств из $\mathcal{J}_n$, то из (4) и определения λ_n следует, что $P(H_n - J_n) = 0$. Поскольку $\mathcal{H}$ покрывает S, то из (2) следует, что $\bigcup\limits_n H_n = S$, и, далее, если $J = \bigcup\limits_n J_n$, что $P(J) = 1$. Далее, J есть в точности объединение множеств из $\mathcal{J}$, которые мы можем занумеровать как G_1', G_2', ... Каждое множество G_n' принадлежит $\mathcal{H}$ и, следовательно, содержится в некотором элементе G_n класса $\mathcal{G}$. Поскольку $P(J) = 1$, отсюда следует, что множества G_n удовлетворяют условию (1), что и требовалось доказать.

Тот очевидный факт, что любая мера P, заданная на сепарабельном пространстве S, сама сепарабельна, вытекает из того, что каждое дискретное подмножество S имеет неизмеримое кардинальное число $\aleph_0$. Из теоремы 2, кроме того, следует, что любая мера P на S сепарабельна, если само S имеет неизмеримое кардинальное число, что справедливо в том случае, когда мощность S не превышает мощности континуума и континуум-гипотеза имеет место. По существу же поиски несепарабельной вероятностной меры относятся не к теории вероятностей, а к теории множеств.

Топология слабой сходимости. Рассмотрим теперь пространство $Z = Z(S)$ вероятностных мер, заданных на $(S, \mathcal{S})$. Превратим Z в хаусдорфово топологическое пространство, взяв в качестве базисных окрестностей P множества вида

$$\left\{Q: \left|\int f_i\, dQ - \int f_i\, dP\right| < \varepsilon, \quad i = 1, \ldots, k\right\}, \tag{5}$$

где ε положительно, а $f_1, \ldots, f_k$ суть элементы $C(S)$. Полученную топологию мы будем называть *топологией слабой сходимости* и обозначать через $\mathscr{W}$. Мы имеем $P_n \Rightarrow P$ тогда и только тогда, когда последовательность $\{P_n\}$ $\mathscr{W}$-сходится к P.

Покажем, что существуют три другие базы для $\mathscr{W}$, а именно, множества

$$\{Q\colon\ Q(F_i) < P(F_i) + \varepsilon,\quad i = 1, \ldots, k\}, \tag{6}$$

где F_i замкнуты, множества

$$\{Q\colon\ Q(G_i) > P(G) - \varepsilon,\quad i = 1, \ldots, k\}, \tag{7}$$

где G_i открыты, и множества

$$\{Q\colon\ |Q(A_i) - P(A_i)| < \varepsilon,\quad i = 1, \ldots, k\}, \tag{8}$$

где A_i — множества непрерывности меры P. Разумеется, каждый из этих трех классов множеств служит базой (для некоторой топологии). Теорема 2.1 является приспособленным для последовательностей вариантом следующей теоремы:

Теорема 3. *Все три указанные базы порождают* $\mathscr{W}$.

Доказательство. Поскольку при $G_i = F_i^c$ (6) и (7) совпадают, то две соответствующие базы идентичны. Мы покажем, что база (6) порождает ту же самую топологию, что и база (8), а затем, что база (6) порождает ту же самую топологию, что и база (5), т. е. $\mathscr{W}$.

Фиксируем меру P. Если A является множеством непрерывности меры P, то для меры Q из некоторого множества вида (6) имеем соотношение $Q(A) \leqslant Q(A^-) < P(A^-) + \varepsilon = P(A) + \varepsilon$, а также $Q(A) \geqslant Q(A^\circ) > P(A^\circ) - \varepsilon = P(A) - \varepsilon$ (последнее соотношение справедливо и для Q из некоторого множества вида (7)). Таким образом, каждое множество вида (8) содержит множество вида (6) (с той же самой мерой P). С другой стороны, если F замкнуто, то найдется некоторое число δ, для которого множество

$$F^\delta = \{x\colon\ \rho(x, F) < \delta\} \tag{9}$$

есть множество непрерывности меры P и

$$P(F^\delta) < P(F) + \frac{1}{2}\varepsilon. \tag{10}$$

Если $|Q(F^\delta)-P(F^\delta)|<\frac{1}{2}\varepsilon$, то $Q(F)<P(F)+\varepsilon$. Таким образом, каждое множество вида (6) содержит некоторое множество вида (8).

Сравним теперь (6) с (5). Фиксируя замкнутое множество F, выберем δ так, чтобы множество (9) удовлетворяло условию (10), а затем выберем в $C(S)$ функцию f, принимающую значение 1 на F, значение 0 вне F^δ и значение, заключенное между 0 и 1 во всех других точках (теорема 1.2). Если $\left|\int f\,dQ-\int f\,dP\right|<\frac{1}{2}\varepsilon$, то $Q(F)<$ $<P(F)+\varepsilon$. Таким образом, каждое множество (6) содержит множество вида (5).

Остается лишь найти внутри множества (5) множество вида (6). По существу достаточно рассмотреть только одну функцию f из $C(S)$, и мы можем предположить, что $0<f(x)<1$ для всех x. Выберем k так, что $1/k<\varepsilon$, и пусть $F_i=\{x\colon\ i/k\leqslant f(x)\}$. В силу (2.2) мы имеем $\int f\,dQ<\varepsilon+k^{-1}\sum_i Q(F_i)$ и $k^{-1}\sum_i P(F_i)\leqslant\int f\,dP$. Таким образом, для Q из (6) мы получаем $\int f\,dQ<\int f\,dP+2\varepsilon$. То же самое рассуждение применительно к $1-f$ завершает доказательство.

Под $N(P)$ мы будем понимать $\mathscr{W}$-окрестность P, которая может иметь любую из форм (5) — (6). Мы свободны в выборе той базы, которая является более удобной.

Теорема 4. *Вероятностные меры с конечным носителем являются $\mathscr{W}$-плотными в Z.*

Доказательство. Рассмотрим окрестность $N(P)$ вида (6). В каждом непустом множестве B из конечного разбиения, порожденного множествами $F_1, \ldots, F_k$, выберем одну точку и поместим в нее массу $P(B)$. Полученная мера обладает конечным носителем и принадлежит $N(P)$, так как она совпадает с P на каждом множестве F_i.

Теорема 4 применима даже к аппроксимации несепарабельных мер P (если таковые имеются). Поэтому, если подмножество Π пространства Z состоит исключительно из сепарабельных мер, нельзя сделать вывод, что каждый элемент $\mathscr{W}$-замыкания Π является сепарабельным. Если, однако, элементы П обладают *общим* сепарабельным носителем и мера P принадлежит $\mathscr{W}$-замыканию П, то P сепарабельно (так как, если A есть замыкание общего сепарабельного носителя, то A сепарабельно и из теоремы 3 следует, что $P(A) = 1$). Поскольку счетная совокупность сепарабельных мер обладает общим сепарабельным носителем, то P является сепарабельной, если сепарабельна каждая мера P_n и $P_n \Rightarrow P$.

Естественно задаться вопросом, когда топология $\mathscr{W}$ метризуема. Для P и Q из Z обозначим через $p(P, Q)$ нижнюю грань таких положительных ε, для которых при всех A из $\mathscr{S}$ выполняются неравенства $Q(A) \leqslant P(A^\varepsilon) + \varepsilon$ и $P(A) \leqslant Q(A^\varepsilon) + \varepsilon$, где $A^\varepsilon = \{x\colon \rho(x, A) < \varepsilon\}$. Если $p(P, Q) = 0$, то P и Q совпадают на замкнутых множествах и, следовательно (теорема 1.1), являются тождественными. Нетрудно проверить, что p обладает и всеми остальными свойствами метрики — ее называют *метрикой Прохорова*. Мы увидим далее, что если топология $\mathscr{W}$ вообще метризуема, то p — одна из порождающих ее метрик.

Фиксируем P. Если для каждой окрестности $N(P)$ существует p-шар с центром в P, содержащийся в $N(P)$, то мы говорим, что p по меньшей мере так же сильна как $\mathscr{W}$ в P. Если, кроме того, для каждого p-шара с центром P существует окрестность $N(P)$, содержащаяся в нем, то мы говорим, что p и $\mathscr{W}$ эквивалентны в P.

Теорема 5. *Для произвольной меры P топология p является по меньшей мере такой же сильной, как и $\mathscr{W}$ в P. Более того, p и $\mathscr{W}$ эквивалентны в P тогда и только тогда, когда $\mathscr{W}$ имеет счетную базу в P, что в свою очередь выполняется тогда и только тогда, когда P сепарабельна.*

Доказательство. По заданным P, F и ε выберем δ так, чтобы $\delta < \varepsilon$ и чтобы множество (9) удовлетворяло условию (10). Если $p(P, Q) < \frac{1}{2}\delta$, то $Q(F) < P(F^\delta) + \delta < P(F) + \varepsilon$. Таким образом, каждая окрестность $N(P)$ вида (6) содержит p-шар с центром в P, что доказывает первую часть теоремы.

Если p и $\mathscr{W}$ эквивалентны в P, то, очевидно, $\mathscr{W}$ имеет счетную базу в P. Остается теперь доказать, что из этого последнего условия вытекает сепарабельность P и что сепарабельность P влечет эквивалентность ρ и $\mathscr{W}$ в P.

Допустим, что $\mathscr{W}$ имеет в P счетную базу $N_n(P)$, $n = 1, 2, \ldots$. По теореме 4 каждое пересечение $\bigcap_{k=1}^{n} N_k(P)$ содержит меру P_n с сепарабельным (даже конечным) носителем. Но тогда $P_n \Rightarrow P$, откуда, как было отмечено после доказательства теоремы 4, следует, что мера P сепарабельна.

Предположим, наконец, что мера P сепарабельна. Нам нужно доказать, что ρ и $\mathscr{W}$ эквивалентны в P. Для этого в силу первой части теоремы достаточно показать, что ρ-шар с радиусом ε и центром в P должен содержать некоторую окрестность $N(P)$.

Выберем δ так, что $3\delta < \varepsilon$. Покроем сепарабельный носитель меры P шарами с диаметром, меньшим чем δ, и являющимися множествами непрерывности меры P, затем перейдем к счетному подпокрытию (стр. 295) и с помощью обычной процедуры построим непересекающиеся множества $A_1, A_2, \ldots$ непрерывности P, покрывающие носитель. Выберем затем k так, чтобы

$$P\left(\bigcup_{i=1}^{k} A_i\right) > 1 - \delta. \tag{11}$$

Каждое из множеств $A_1, \ldots, A_k$ имеет диаметр, меньший δ, и если $\mathscr{A}$ представляет собой (конечный) класс объединений этих множеств, то каждый элемент $\mathscr{A}$ является множеством непрерывности меры P.

По теореме 3 существует такая окрестность $N(P)$, что если $Q \in N(P)$, то

$$|Q(A) - P(A)| < \delta, \qquad A \in \mathscr{A}. \tag{12}$$

Из неравенств (12) и (11) следует, что

$$Q\left(\bigcup_{i=1}^{k} A_i\right) > 1 - 2\delta. \tag{13}$$

Мы покажем, что $\rho(P, Q) < \varepsilon$, если $Q \in N(P)$.

Пусть B — произвольное множество из $\mathscr{S}$. Рассмотрим объединение A тех множеств из последовательности $A_1, \ldots, A_k$, которые пересекаются с B. Множество A удовлетворяет неравенству (12). Из соотношений

$$B \subset A \cup \left(\bigcup_{i=1}^{k} A_i\right)^c \quad \text{и} \quad A \subset B^\delta$$

(последнее верно, так как диаметр A_i меньше δ), а также неравенств (11), (12) и (13) следует, что

$$P(B) \leqslant Q(B^\delta) + 2\delta \quad \text{и} \quad Q(B) \leqslant P(B^\delta) + 3\delta.$$

Поскольку $3\delta < \varepsilon$, то мы получаем $p(P, Q) < \varepsilon$.

Из теорем 2 и 5 следует, что $\mathscr{W}$ метризуема тогда и только тогда, когда каждое дискретное подмножество S имеет неизмеримое кардинальное число, и в этом случае $\mathscr{W}$ метризуема посредством p.

Можно также задаться вопросом — когда Z сепарабельно? Если S сепарабельно, то топологии p и $\mathscr{W}$ совпадают на Z и из теоремы 4 следует, что множество вероятностных мер, обладающих конечным носителем, всюду плотно в Z. Нетрудно пойти дальше и показать, что если A_0 — счетное всюду плотное множество в S и Z_0 состоит из таких мер P, которые обладают конечным носителем, содержащимся в A_0, и имеют рациональные массы, то Z_0 счетно и всюду плотно в Z относительно p. Таким образом, $\mathscr{W}$ метризуема и сепарабельна, если сепарабельно S.

С другой стороны, если $\mathscr{W}$ удовлетворяет второй аксиоме счетности, то $\mathscr{W}$ сепарабельна и может быть метризована посредством p. Так как пространство S с его метрикой гомеоморфно множеству единичных масс, принадлежащих Z, то отсюда следует, что S сепарабельно.

Теорема Прохорова. Пересмотрим теорему Прохорова в свете предыдущих результатов. Если Π есть подмножество Z, то пусть Π_p и $\Pi_{\mathscr{W}}$ обозначают его замыкания относительно p и $\mathscr{W}$, соответственно. Из теоремы 5 вытекает, что $\Pi_p \subset \Pi_{\mathscr{W}}$ (со строгим вложением для некоторого подмножества Π, если топология $\mathscr{W}$ неметризуема).

Множество в топологическом пространстве называется компактным, если каждое открытое покрытие его имеет конечное подпокрытие; множество называется секвенциально компактным, если каждая последовательность его точек содержит подпоследовательность, сходящуюся к пределу, который также принадлежит этому множеству.

Рассмотрим следующие три утверждения:

1°. Каждая последовательность из Π содержит $\mathscr{W}$-сходящуюся подпоследовательность.

2°. $\Pi_{\mathscr{W}}$ является секвенциально компактным.

3°. $\Pi_{\mathscr{W}}$ является компактным.

Предел сходящейся подпоследовательности в 1° не обязан принадлежать Π, но, разумеется, он будет принадлежать $\Pi_{\mathscr{W}}$. Очевидно, что 2° влечет 1°. Из этих трех утверждений можно составить пять других импликаций; является ли любая из них верной, неизвестно*). Если, однако, каждое дискретное подпространство S имеет неизмеримое кардинальное число, то $\mathscr{W}$-топология метризуема, так что все три утверждения равносильны, и мы увидим, что это верно также тогда, когда Π плотно. В этой книге (см. начало § 6) мы *называем* Π относительно компактным множеством, если Π удовлетворяет условию 1°.

Рассмотрим прямую часть теоремы Прохорова теорема (6.1).

Теорема 6. *Если* Π *плотно, то* $\Pi_p = \Pi_{\mathscr{W}}$, $\Pi_{\mathscr{W}}$ *плотно,* p- *и* $\mathscr{W}$-*топологии совпадают на* $\Pi_{\mathscr{W}}$, $\Pi_{\mathscr{W}}$ *одновременно компактно и секвенциально компактно* (*в любой топологии*).

Доказательство. Поскольку Π плотно, при каждом ε существует компакт K такой, что $Q(K) > 1 - \frac{1}{2}\varepsilon$ для всех Q из Π. Допустим, что $P \in \Pi_{\mathscr{W}}$. По теореме 3 существует такая окрестность $N(P)$, что $Q \in N(P)$ влечет $Q(K) < P(K) + \frac{1}{2}\varepsilon$. Выбирая Q из множества $N(P) \cap \Pi$, получаем $P(K) > 1 - \varepsilon$. Таким образом, $\Pi_{\mathscr{W}}$ плотно.

---

*) Мне.

В частности, каждый элемент множества $\Pi_{\mathscr{W}}$ является сепарабельным. Поэтому (теорема 5) p- и $\mathscr{W}$-топологии совпадают на $\Pi_{\mathscr{W}}$; поскольку в любом случае $\Pi_p \subset \Pi_{\mathscr{W}}$, имеем $\Pi_p = \Pi_{\mathscr{W}}$. Из плотности $\Pi_{\mathscr{W}}$ в силу теоремы 6.1 вытекает, что $\Pi_{\mathscr{W}}$ секвенциально компактно в смысле $\mathscr{W}$. Так как в метрическом случае компактность и секвенциальная компактность совпадают, то доказательство закончено.

Прямая теорема Прохорова может быть использована также для исследования вопроса о полноте Z. Допустим, что S полно и что каждый элемент Z сепарабелен (и, следовательно, является плотным). Тогда p- и $\mathscr{W}$-топологии совпадают на Z, и мы покажем, что Z p-полно. Для этого достаточно показать, что p-фундаментальная последовательность $\{P_n\}$ необходимо содержит слабо сходящуюся подпоследовательность, что в свою очередь будет следовать из прямой теоремы Прохорова, если мы докажем, что последовательность $\{P_n\}$ плотна. Наконец (см. доказательство теоремы 6.2 на стр. 63), поскольку S полно, то последовательность $\{P_n\}$ плотна, если для любых положительных ε и δ существует конечная совокупность δ-шаров $A_1, \ldots, A_k$ такая, что при всех n

$$P_n(A_1 \cup \ldots \cup A_k) > 1 - \varepsilon. \tag{14}$$

Выберем η так, чтобы $2\eta < \min\{\varepsilon, \delta\}$, а затем выберем n_0 так, чтобы $p(P_n, P_{n_0}) < \eta$ при $n \geqslant n_0$. Поскольку мера P_{n_0} сепарабельна, существует конечное множество η-шаров $B_1, \ldots, B_k$ таких, что $P_{n_0}\left(\bigcup_{i=1}^{k} B_i\right) > 1 - \eta$. Пусть $A_1, \ldots, A_k$ — 2η-шары с теми же центрами. Так как $\left(\bigcup_{i=1}^{k} B_i\right)^{\eta} \subset \bigcup_{i=1}^{k} A_i$ и $p(P_n, P_{n_0}) < \eta$ при $n \geqslant n_0$, то отсюда следует, что (14) выполняется при $n \geqslant n_0$. Поскольку каждая мера P_n сепарабельна, мы можем добиться выполнения (14) и при $n < n_0$, увеличив, если нужно, систему $\{A_i\}$.

Таким образом, пространство Z является p-полным, если S полно и каждая вероятностная мера на нем сепарабельна. Допустим, с другой стороны, что Z p-полно. Нетрудно показать, что множество Z_0 единичных масс является p-замкнутым, а следовательно, также p-полным, откуда можно вывести, что S также полно.

Обратимся теперь к обратной части теоремы Прохорова.

Теорема 7. *Если множество мер* Π $\mathscr{W}$-*компактно и каждая мера из* Π *сепарабельна и если* S *топологически полно, то* Π *плотно.*

Доказательство. По теореме 1 элементы Π, взятые в отдельности, являются плотными. Для того чтобы доказать, что само Π плотно, достаточно, как обычно, построить по данным ε и δ конечное множество δ-шаров $A_1, \ldots, A_k$ таких, что $P(A_1 \cup \ldots \cup A_k) > 1 - \varepsilon$ при всех P из Π.

Если P принадлежит Π, то, поскольку P плотна, существует компактное множество K такое, что $P(K) > > 1 - \frac{1}{2}\varepsilon$, при этом K может быть покрыто конечным множеством открытых δ-шаров $B_{P,i}$, $i = 1, \ldots, k_P$. По теореме 3 найдется $\mathscr{W}$-окрестность $N(P)$ элемента P такая, что объединение G_P шаров $B_{P,i}$ удовлетворяет при $Q \in N(P)$ неравенству $Q(G_P) > P(G_P) - \frac{1}{2}\varepsilon$ и, следовательно, неравенству $Q(G_p) > 1 - \varepsilon$. Так как Π есть $\mathscr{W}$-компакт, то Π можно покрыть конечным набором таких окрестностей $N(P)$. В качестве системы $A_1, \ldots, A_k$ мы можем теперь взять $B_{P,i}$, соответствующие выбранным окрестностям $N(P)$.

Используя теорему 1, можно сформулировать теорему 7 в иной форме: если множество мер Π $\mathscr{W}$-компактно и каждая мера из него плотна и если S топологически полно, то Π плотно. Предположение о том, что каждая мера P из Π плотна, нельзя отбросить, так как одна неплотная мера P образует компактное множество. С другой стороны, представляется правдоподобным, что условие полноты может быть исключено, т. е. что Π плотно, если оно $\mathscr{W}$-компактно и если все его элементы плотны *). Хотя проблема и нерешена, тем не менее в этом направлении существует следующий интересный результат:

---

*) Варадарайн (1958а и 1961а) получил результат такого рода, но его доказательство содержит пробел.

Теорема 8. *Предположим, что $P_n \Rightarrow P$, где P и все P_n являются плотными. Тогда совокупность $\{P, P_1, P_2, \ldots\}$ плотна.*

Доказательство. Так как рассматриваемые меры, взятые в отдельности, плотны, то достаточно показать, что для любого ε существует компакт K такой, что

$$P_n(K) > 1 - \varepsilon \tag{15}$$

при всех достаточно больших n.

Для любого ε в силу плотности меры P существует компакт K' такой, что $P(K') > 1 - \frac{1}{3}\varepsilon$. Положим $G_k = \{x: \rho(x, K') < 1/k\}$. Так как $P_n \Rightarrow P$, то существует возрастающая последовательность $\{n_k\}$ такая, что при $n \geqslant n_k$

$$P_n(G_k) > P(G_k) - \frac{1}{3}\varepsilon > 1 - \frac{2}{3}\varepsilon. \tag{16}$$

При $n_k \leqslant n \leqslant n_{k+1}$, используя плотность P_n, выберем компакт K'_n, удовлетворяющий соотношениям $K'_n \subset G_k$ и

$$P_n(G_k - K'_n) < \frac{1}{3}\varepsilon.$$

Множество

$$K_k = K' \cup \bigcup_{n=n_k}^{n_{k+1}} K'_n$$

компактно, $K' \subset K_k \subset G_k$ и при $n_k \leqslant n \leqslant n_{k+1}$ имеет место неравенство

$$P_n(G_k - K_k) < \frac{1}{3}\varepsilon.$$

Из (16) теперь следует, что при $n_k \leqslant n \leqslant n_{k+1}$

$$P_n(K_k) > 1 - \varepsilon. \tag{17}$$

Обозначим $K = \bigcup_k K_k$, и пусть $\{x_u\}$ — последовательность в K. Если существует какое-либо одно множество K_k, содержащее все точки x_u, то $\{x_u\}$ содержит сходящуюся подпоследовательность; в противном случае существуют подпоследовательность $\{x_{u_m}\}$ и последовательность целых чисел $\{v_m\}$ такие, что $x_{u_m} \in K_{v_m} \subset G_{v_m}$ и

$v_m \geqslant m$, так что $\rho(x_{u_m}, K') < 1/m$. Последовательность $\{x_{u_m}\}$ должна содержать другую подпоследовательность, которая сходится. Таким образом, K — компакт. Если $n \geqslant n_1$, то $n_k \leqslant n \leqslant n_{k+1}$ при некотором k, и из (17) и соотношения $K \supset K_k$ следует (15).

В этой теореме предположение о том, что мера P плотна, существенно: чтобы в этом убедиться, достаточно взять неплотную меру P на сепарабельном пространстве S, как во втором примечании к теореме 1, и применить теорему 4.

Примечание. Теорема 8 и пример в примечании 2 к теореме 1 принадлежат Ле Каму (1960); остальные результаты принадлежат по большей части Варадарайну (1958а и 1961а).

БИБЛИОГРАФИЯ

Александров А. Д.
(1940—1943) Additive set functions in abstract spaces, Матем. сб. 8, 307—348; 9, 563—628; 13, 169—238.
Андерсон (Anderson T. W.)
(1963) Asymptotic theory for principal component analysis, Ann. Math. Statist. 34, 122—148.
Андерсон, Дарлинг (Anderson T. W., Darling D. A.)
(1952) Asymptotic theory of certain «goodness of fit» criteria based on stochastic processes, Ann. Math. Statist. 23, 193—212.
Банах (Banach S.)
(1932) Théorie des Opérations Linéaires. Warsaw: Monografie Matematyczne.
Бартошинский (Bartoszyński R.)
(1961) A characterization of the weak convergence of measures, Ann. Math. Statist. 32, 561—576.
Бикель (Bickel P. J.)
(1967) Some contributions to the theory of order statistics, Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), vol. I, 575—591; Univ. California. Press, Berkeley, California.
(1969) A distribution free version of the Smirnov two sample test in the p-variate case, Ann. Math. Statist., 40, 1, 1—23.
Биллингсли (Billingsley P.)
(1956) The invariance principle for dependent random variables. Trans. Amer. Math. Soc. 83, 250—268.
(1961) The Lindeberg — Lévy theorem for martingales, Proc. Amer. Math. Soc. 12, 788—792.
(1962) Limit theorems for randomly selected partial sums, Ann. Math. Statist. 33, 85—92.
(1964) An application of Prohorov's theorem to probabilistic number theory, Ill. J. Math. 8, 697—704.
(1969) Эргодическая теория и информация (перев. с англ.), М., «Мир».
Биллингсли, Топсё (Billingsley P., Topsøe F.)
(1967) Uniformity in weak convergence, Z. Wahrscheinlichkeits — theorie und Verw. Gebiete 7, 1—16.
Биркгоф (Birkhoff G.)
(1952) Теория структур (перев. с англ.), М., ИЛ.
Бирнбаум, Пайк (Birnbaum Z. W., Pyke R.)
(1958) On some distributions related to the statistic D_n^+, Ann. Math. Statist. 29, 179—187.

Блюм, Хэнсон, Розенблатт (Blum J. R., Hanson D. L., Rosenblatt J. I.)
(1963) On the central limit theorem for the sum of a random number of independent random variables, Z. Wahrscheinlickheistheorie und Verw. Gebiete 1, 389—393.
Бор (Bohr H.)
(1934) Почти периодические функции (перев. с нем.), М. — Л.
Варадарайн (Varadarajan V. S.)
(1958a) Convergence in distribution of stochastic processes, Thesis, Calcutta University.
(1958b) Weak convergence of measures on separable metric spaces, Sankyā 19, 15—22.
(1958c) On the convergence of probability distributions, Sankyā, 19, 23—26.
(1961a) Меры на топологических пространствах, Матем. сб. 55 (97), 35—100.
(1961b) Convergence of stochastic processes, Bull. Amer. Math. Soc. 67, 276—280.
Гихман И. И., Скороход А. В.
(1965) Введение в теорию случайных процессов, М., «Наука».
Гросс (Gross L.)
(1963) Harmonic analysis on Hilbert space, Met. Amer. Math. Soc. 46.
Дадли (Dudley R. M.)
(1966) Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces, Ill. J. Math. 10, 109—126.
(1967) Measures on non-separable metric spaces, Ill. J. Math. 11, 449—453.
Данфорд, Шварц (Dunford N., Schwartz J. T.)
(1962) Линейные операторы. Общая теория, М., ИЛ (перев. с англ.)
Дарлинг (Darling D. A.)
(1957) The Kolmogorov — Smirnov, Cramér — von Mises tests, Ann. Math. Statist. 28, 823—838.
Дарлинг, Эрдеш (Dárling D. A., Erdös P. E.)
(1956) A limit theorem for the maximum of normalized sums of independent random variables, Duke Math. J. 23, 143—155.
Дёблин (Doeblin W.)
(1940) Remarques sur la théorie metrique des fractions continues, Compositio Math. 7, 353—371.
Донскер (Donsker M.)
(1951) An invariance principle for certain probability limit theorems, Met. Amer. Math. Soc. 6.
(1952) Justifications and extension of Doob's hueristic approach to the Kolmogorov — Smirnov theorems, Ann. Math. Statist. 23, 277—281.
Дуосс, Карлин (Dwass M., Karlin S.)
(1963) Conditional limit theorems, Ann. Math. Statist. 34, 1147—1167.
Дуб (Doob J. L.)
(1949) Heuristic approach to the Kolmogorov — Smirnov theorems, Ann. Math. Statist. 20, 393—403.
(1956) Вероятностные процессы (перев. с англ.), М., ИЛ.

Дьедонне (Dieudonné J.)
(1964) Основы современного анализа (перев. с фр.), М., ИЛ.
Ибрагимов И. А.
(1962) Некоторые предельные теоремы для стационарных процессов, Теория вероят. и ее примен. 7, вып. 4, 361—392.
(1963) Центральная предельная теорема для одного класса зависимых случайных величин, Теория вероят. и ее применен. 8, вып. 1, 89—94.
Ито, Маккин (Itô K., McKean H. P.)
(1968) Диффузионные процессы и их траектории (перев. с англ.), М., «Мир».
Каллианпур (Kallianpur G.)
(1961) The topology of weak convergence of probability measures, J. Math. Mech. 10, 947—969.
Карлин (Karlin S.)
(1971) Основы теории случайных процессов (перев. с англ.), М., «Мир».
Кац (Kac M.)
(1946) On the distributions of values of sums of type, Ann. of Math. 47, 33—49.
(1949) On deviations between theoretical and empirical distributions, Proc. Nat. Acad. Sci. U. S. A. 35, 252—257.
Кейслер, Тарский (Keisler H. J., Tarski A.)
(1946) From accessible to inaccessible cardinals, Fund. Math. 53, 225—308.
Келли (Kelley J. L.)
(1968) Общая топология (перев. с англ.), М., «Наука».
Кимм (Kimme E. G.)
(1957) On the convergence of sequences of stochastic processes, Trans. Amer. Math. Soc. 84, 208—229.
(1960) Some equivalence conditions for the convergence in distribution of sequences of stochastic processes, Trans. Amer. Math. Soc. 95, 495—515.
Колмогоров А. Н
(1931) Eine Verallgemeinierung des Laplace — Liapounoffschen Satzes, Изв. Акад. наук СССР, сер. физ.-мат., 959—962.
(1956) О сходимости А. В. Скорохода, Теория вероят. и ее примен. 1, вып. 2, 239—247.
(1974) Основные понятия теории вероятностей, 2-е изд., М., «Наука». (Первое оригинальное издание книги: Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin — Göttingen — Heidelberg: Springer — Verlag, 1933.)
Колмогоров А. Н., Прохоров Ю. В.
(1954) Zufällige Funktionen und Grenzverteilungssätze, Berich über die Tagung Wahrscheinlichkeitsrechnung und Mathematischen Statistik, 113—126. Berlin: Deutscher Verlag der Wissenschaften.
Колмогоров А. Н., Фомин С. В.
(1972) Элементы теории функций и функционального анализа, М., «Наука».
Крамер (Cramér H.)
(1975) Математические методы статистики (перев. с англ.), 2-е изд., М., «Мир».

Кремер, Уолд (Cramér H., Wold H.)
(1936) Some theorems on distribution function, J. London Math Soc, 11, 290—295.
Ламперти (Lamperti J.)
(1962) An invariance principle in renewal theory, Ann. Math. Statist. 33, 685—696.
(1962a) A new class of probability limit theorems, J. Math. Mech. 11, 749—772.
(1962b) On convergence of stochastic processes, Trans. Amer. Math. Soc. 104, 430—435.
Ле Кам (LeCam L.)
(1960) Сходимость по распределению случайных процессов (перев. с англ.), Математика, 4: 3, 107—142.
Леви (Lèvy P.)
(1925) Cacul des Probabilites. Paris: Gauthier — Villars.
(1937) Théorie de l'Addition des Variables Aléatoires. Paris: Gauthier — Villars.
(1972) Стохастические процессы и броуновское движение (перев. с фр.), М., «Наука».
Лигет (Liggett T. M.)
(1968) An invariance principle for conditioned sums of independent random variables, J. Math. Mech. 18, 559—570.
Лоэв (Loève M.)
(1962) Теория вероятностей (перев. с фр.), М., ИЛ.
Манн, Вальд (Mann H. B., Wald A.)
(1943) On stochastic limit and order relations, Ann. Math. Statist. 14, 217—226.
Марк (Mark A. M.)
(1949) Some probability theorems, Bull. Amer. Math. Soc. 55, 885—900.
Найт (Knight F. B.)
(1962) On the random walk and Brownian motion, Trans. Amer. Math. Soc. 103, 218—228.
Окстоби, Улам (Oxtoby J. C., Ulam S.)
(1939) On the existence of a measure invariant under a transformation, Ann. of Math. (2) 40, 560—566.
Пайк (Pyke R.)
(1965) Spacings, J. Roy. Stat. Soc. Ser. B27, 395—449.
(1968) The weak convergence of the empirical process with random sample size, Proc. Cambridge Phil. Soc. 64, 155—160.
Пайк, Шорак (Pyke R., Shorack G.)
(1968) Weak convergence of the two-sample empirical process and a new approach to Chernoff — Savage theorems, Ann. Math. Statist., 39, 3, 755—771.
Партасарати (Parthasarathy K. R.)
(1967) Probability Measures on Metric Spaces. New York: Academic Press.
Прохоров Ю. В.
(1953) Распределения вероятностей в функциональных пространствах, Усп. матем. наук 8, вып. 3, 165—167.
(1956) Сходимость случайных процессов и предельные теоремы теории вероятностей, Теория вероят. и ее примен. 1, вып. 2, 177—238.

(1960) The method of characteristic functionals, Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2, 403—429.

Ранга Рао (Ranga Rao R.)

(1962) Relations between weak and uniform convergence of measures with applications, Ann. Math. Statist. 33, 659—680.

Ранга Рао, Варадарайи (Ranga Rao R., Varadarajn V. S.)

(1958) On a theorem in metric spaces, Ann. Math. Statist. 29, 612—613.

Реньи (Rényi A.)

(1958) On mixing sequences of sets, Acta Math. Acad. Sci. Hung. 9, 215—228.

(1960) On the central limit theorem for the sum of a random number of independent random variables, Acta. Math. Acad. Sci. Hung. 11, 97—102.

Розеи (Rosén B.)

(1964) Limit theorems for sampling from a finite population, Arkiv för Mathematik 5, 383—424.

(1967a) On the central limit theorem for sums of dependent random variables, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 7, 48—82.

(1967b) On asymptotic normality of sums of dependent random variables, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 7, 95—102.

(1967c) On the central limit theorem för a class of sampling procedures, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 7, 103—115.

Сетурамаи (Sethuraman J.)

(1964) On the probability of large deviations of families of sample means, Ann. Math. Statist. 35, 1304—1316.

(1965) On the probability of large deviations of the mean for random variables in D [10, 1], Ann. Math. Statist. 36, 280—285.

Скороход А. В.

(1956) Предельиые теоремы для случайиых процессов, Теория вероят. и ее примеи. 1, вып. 3, 289—319.

(1957) Предельиые теоремы для случайиых процессов с иезависимыми приращеииями, Теория вероят. и ее примеи. 2, вып. 2, 145—177.

(1961) Исследоваиия по теории случайиых процессов, Киев, Изд-во КГУ.

Слуцкий (Slutsky E. E.)

(1925) Über Stochastische Asymptoten und Grenzwerte, Metron 5, 1—90.

(1937) Qualche proposizione relativa alla teoria delle funzioni aleatorie, Giorn. Inst. Ital. Attuair 8, 183—199.

Стоуи (Stone C.)

(1963) Weak convergence of stochastic processes defined on semi-infinite time intervals, Proc. Amer. Math. Soc. 14, 694—696.

(1963a) Limit theorems for random walks, birth and death processes, and diffusion processes, Ill. J. Math. 7, 638—660.

Штрассен (Strassen V.)
(1964) An invariance principle for the Law of the iterated logarithm, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 3, 211—226.
Титчмарш (Titchmarsh E. C.)
(1951) Теория функций (перев. с англ.), М., Гостехиздат.
Топсё (Topsøe F.)
(1967a) On the connection between P-continuity and P-uniformity in weak convergence, Теория вероят. и ее примен. 12, 279—288.
(1967b) Preservation of weak convergence under mappings, Ann. Math. Statist. 38, 1661—1665.
Трюмбо (Trumbo B. E.)
(1965) Sufficient conditions for the weak convergence of conditional probability distributions in a metric space, Thesis, University of Chicago.
Феллер (Feller W.)
(1967a) Введение в теорию вероятностей и ее приложения (перев. с англ.), том 1, М., «Мир».
(1967b) Введение в теорию вероятностей и ее приложения (перев. с англ.), том 2, М., «Мир».
Форте (Forter R.)
(1940) Sur une suite également répartie, Studia Math. 9, 54—69.
Уичура (Wichura M. J.)
(1969) Inequalities with applications to the weak convergence of random processes with multidimensional time parameter, Ann. Math. Statist. 40, 2, 681—687.
Халмош (Halmos P. R.)
(1953) Теория меры (перев. с англ.), М., ИЛ.
Харди, Райт (Hardy G. H., Wright E. M.)
(1960) An Introduction to the Theory of Numbers, fourth edition. Oxford: Clarendon Press.
Хеннекен, Тортра (Hennequin P. L., Tortrat A.)
(1974) Теория вероятностей и некоторые ее приложения (перев. с фр.), М., «Наука».
Хинчин А. Я.
(1936) Асимптотические законы теории вероятностей, М. — Л., ОНТИ.
(1961) Цепные дроби, 3-е изд., Физматгиз, М.
Цесельский, Кестен (Ciesielski Z., Kesten H.)
(1962) A limit theorem for the fractional parts of the sequence $2^k t$, Proc. Amer. Math. Soc. 13, 596—600.
Ченцов Н. Н.
(1956) Слабая сходимость случайных процессов с траекториями без разрыва второго рода и так называемый «эвристический» подход к критериям согласия типа Колмогорова — Смирнова, Теория вероят. и ее примен. 1, вып. 1, 155—161.
Чернов (Chernoff H.)
(1956) Large sample theory: parametric case, Ann. Math. Statist. 27, 1—22.
Чернов, Тейчер (Chernoff H., Teicher H.)
(1958) A central limit theorem for sequences of exchangeable random variables, Ann. Math. Statist. 29, 118—130.

Чибисов Д. М.
(1965) К исследованию асимптотической мощности критериев согласии, Теории вероит. и ее применеи. 10, вып. 3, 460—478.
Шеффе (Scheffé H.)
(1947) A useful convergence theorem for probability distributions, Ann. Math. Statist. 18, 434—438.
Шмид (Schmid P.)
(1958) On the Kolmogorov and Smirnov limit theorems for discontinuous distribution functions, Ann. Math. Statist. 29 (1011—1027).
Эрдеш, Кац (Erdõs P., Kac M.)
(1946) On certain limit theorems in the theory of probability, Bull. Amer. Math. Soc. 52, 292—302.
(1947) On the number of positive sums of independent random variables, Bull. Amer, Math, Soc, 53, 1011—1020,

СПИСОК ОБОЗНАЧЕНИЙ

Этот список включает лишь те специальные термины, которые систематически используются на всем протяжении книги. Хотя произвольному элементу пространства S предназначено обозначение x, фактически он может быть обозначен любой строчной буквой латинского алфавита (возможно, с индексами); это замечание касается обозначения произвольного элемента любого пространства.

Теория множеств

E^c — дополнение множества E; $E_1 - E_2 = E_1 \cap E_2^c$ — разность E_1 и E_2; $E_1 + E_2 = (E_1 - E_2) \cup (E_2 - E_1)$ — симметрическая разность E_1 и E_2; I_E — индикатор или характеристическая функция множества E.

Вероятностные пространства

$(\Omega, \mathscr{B})$ — измеримое пространство; ω и E обозначают, соответственно, элементы Ω и $\mathscr{B}$; P — вероятностная мера на $(\Omega, \mathscr{B})$; E обозначает математическое ожидание; ξ — случайная величина, а $\{\xi_n\}$ и $\{\xi_t\}$ — стохастические процессы.

Если h отображает $(\Omega, \mathscr{B})$ в $(\Omega', \mathscr{B}')$, то соотношение $h^{-1}\mathscr{B}' \subset \mathscr{B}$ означает, что $h^{-1}E' \in \mathscr{B}$ для любого $E' \in \mathscr{B}'$, т. е. что h — измеримое отображение, и $\mathsf{P}h^{-1}$ принимает значение $\mathsf{P}(h^{-1}E')$ в E' (стр. 304).

Произвольное метрическое пространство

S — метрическое пространство с произвольным элементом x, метрикой $\rho(x, y)$ (и $\rho(x, A)$ — см. стр. 296) и σ-полем $\mathscr{S}$ борелевских множеств; A°, A^{-1} и ∂A определяют, соответственно, внутренность, замыкание и границу произвольного множества A (которое является элементом $\mathscr{S}$, если не оговорено противное); $S(x, \varepsilon)$ — открытый ε-шар с центром в x; F — произвольное замкнутое множество (но этой буквой обозначается также функция распределения),

G — произвольное открытое множество, а K — произвольное компактное множество (но эта буква употребляется также для обозначения постоянной).

$C(S)$ с общим элементом f есть класс ограниченных непрерывных функций на C; если h отображает S в другое метрическое пространство, то D_h есть множество ее точек разрыва.

P и Q — вероятностные меры на $(S, \mathscr{S})$; соотношение $P_n \Rightarrow P$ обозначает слабую сходимость; Π — семейство мер P.

Специальные метрические пространства

σ-поле борелевских множеств в пространстве изображается рукописным вариантом (с любыми надлежащими индексами) той прописной буквы латинского алфавита, которой обозначено само пространство.

R^k — евклидово k-мерное пространство; H — произвольный элемент $\mathscr{R}^k$; F — функции распределения; $F_n \Rightarrow F$ обозначает слабую сходимость (стр. 31); обозначения $x \leqslant y$, $x < y$, $|x - y|$ и $(a, b]$ см. на стр. 30.

R^∞ — топологическое произведение счетной последовательности прямых (стр. 32 и 299) с общей точкой $x = (x_1, x_2, \ldots)$; $\pi_k(x) = (x_1, \ldots, x_k)$ — естественная проекция в R^k.

C — пространство $C[0, 1]$ (стр. 33, 82, 301) с метрикой ρ и общей точкой $x = x(t)$; $\pi_{t_1 \ldots t_k}(x) = (x(t_1), \ldots, x(t_k))$ — естественная проекция в R^k; о координатной случайной величине x_t и о понятии «распределение x_t относительно P» см. стр. 90.

D — пространство $D[0, 1]$ (стр. 153) с двумя эквивалентными метриками d (стр. 156) и d_0 (стр. 158) и общей точкой $x = x(t)$; проекции $\pi_{t_1 \ldots t_k}$ и координатные случайные величины x_t определяются так же, как для C; специальные символы: λ и Λ (стр. 156), $\|\lambda\|$ (стр. 158), T_P (стр. 172), J_t (стр. 173) и T_x (стр. 179).

Случайные элементы

X — случайный элемент (стр. 36) со значением $X(\omega)$ в ω; если область изменения X есть пространство функций, то $X(t) = \pi_t X$ и $X(t, \omega) = \pi_t(X(\omega))$ (стр. 86 для C; стр. 178 для D); $X_n \xrightarrow{\mathscr{D}} X$ и $X_n \xrightarrow{\mathscr{D}} P$ обозначают сходимость по распределению (стр. 38, 39); $X_n \xrightarrow{P} a$ и $X_n \xrightarrow{P} X$ обозначают сходимость по вероятности (стр. 40 и 42); о (X, Y) и $\rho(X, Y)$ см. стр. 40, 308, 309.

N и $N(\mu, \sigma^2)$ — нормальные распределения или величины (стр. 39)′ W — винеровская мера на C (стр. 91), случайная функции в C с винеровской мерой в качестве распределения (стр. 95) или соответствующая мера или случайная функции в D (стр. 190); W° — броуновский мост как мера или случайная функции в C или D (стр. 95, 96).

Модули

Модули непрерывности для C и соответственные модули для D:
$w_x(\delta) = w(x, \delta)$, стр. 83 и 302,
$w_X(\delta) = w(X, \delta)$ стр. 87,
$w_x(T_0)$, стр. 153,
$w'_x(\delta)$, стр. 154,
$w''_x(\delta)$,стр. 165,
$w''(X, \delta)$, стр. 179.

ИМЕННОЙ УКАЗАТЕЛЬ

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ

Патрик Биллингсли

СХОДИМОСТЬ ВЕРОЯТНОСТНЫХ МЕР

М., 1977 г., 352 стр. с илл.

Редактор *В. В. Абгарян*

Техн. редактор *И. Ш. Аксельрод*

Корректоры *З. В. Автонеева*, *М. Л. Медведская*

Сдано в набор 23.12.76. Подписано к печати 22.06.77. Бумага 84×108$^1/_{32}$ тип. № 1. Физ. печ. л. 11. Усл. печ. л. 18,48. Уч.-изд. л. 17,74. Тираж 9000 экз. Цена книги 1 р. 50 к. Заказ № 439.

Издательство «Наука»
Главная редакция
физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Ордена Трудового Красного Знамени
Ленинградская типография № 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.

4 253 151