С. В. СИМУШКИН

МЕТОДЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Учебное пособие

КРАСНОДАР

УДК 519.2
ББК 22.171я73

С 37 **Симушкин С. В.** Методы теории вероятностей

ISBN 978-5-8114-3442-8

В пособии в едином ключе представлены основные приёмы, используемые при анализе случайных явлений, которые редко включают в начальные курсы теории вероятностей, оставляя их для углублённых курсов магистратуры и аспирантуры в области стохастического анализа. Изложение материала замкнуто в себе и опирается на представленные в пособии классические основания теории вероятностей, связанные с понятиями измеримого пространства, измеримого отображения и со свойствами меры и интеграла Лебега.

Предназначено для магистрантов и аспирантов физико-математических специальностей, входящих в УГС: «Математика и механика», «Компьютерные и информационные науки», «Физика и астрономия».

УДК 519.2
ББК 22.171я73

Рецензент

И. Н. ВОЛОДИН — профессор кафедры математической статистики Казанского федерального университета.

Обложка
Е. А. ВЛАСОВА

СОДЕРЖАНИЕ

Обозначения и сокращения

(?!)	—	факт, доказательство которого доверено читателю;
⚠	—	замечание;
Упр.	—	упражнение для самостоятельного решения;
⤇ ⤆	—	знаки начала и завершения доказательства;
$\circledcirc$	—	знак окончания примера (серии примеров);
$\{x : \mathfrak{A}(x)\}$	—	совокупность элементов x, удовлетворяющих $\mathfrak{A}$;
$\langle O_\alpha \rangle_{\alpha\in\aleph}$	—	совокупность объектов O_α, индексированных α;
$j = \overline{m,n}$	—	индекс j пробегает значения $m, m+1, \ldots, n$;
$\biguplus$	—	объединение непересекающихся подмножеств;
$\forall$	—	любой, любого, любых и т.п.;
$\lim_\downarrow$, $\searrow$	—	предел убывающей последовательности;
$\lim_\uparrow$, $\nearrow$	—	предел возрастающей последовательности;
$S := D$	—	S определяется (впервые) посредством D;
$S \overset{\mathrm{df}}{=} D$	—	напоминание, что S определялось через D;
$\llbracket a \rrbracket$	—	целая часть a — наибольшее целое, не больше a;
$\#(A)$	—	количество элементов A;
$(a; c]$	—	полуоткрытый интервал;
$\lvert M \rvert$	—	определитель матрицы M;
$M^\flat$, $\vec{x}^\flat$	—	операция транспонирования;
$\mathbb{R}^k$	—	k-мерное евклидово пространство (иногда $\mathbb{R}^1 = \mathbb{R}$);
$\vec{x}$	—	вектор-столбец; строковая запись $\vec{x} = (x_1, \ldots, x_k)$ используется для сокращения;
$\vec{x} + c$	—	$(x_1 + c, \ldots, x_k + c)$ — сдвиг координат вектора;
$\vec{x} \lessdot \vec{y}$	—	$x_j < y_j$ для $\forall\, j$;
$\xi \backsim \ldots$	—	ξ имеет распределение $\ldots$;
сл.в.	—	случайная величина;
сл.вектор	—	случайный вектор;
т. т. т. когда	—	тогда и только тогда, когда;
усл.м.о.	—	условное математическое ожидание;
ф.вер.	—	функция вероятностей дискретной сл.в.;
ф.р. (ф.пл.)	—	функция распределения (функция плотности).

Политика использования символов

A, B, C	—	подмножества или действительные константы;
Γ, B	—	гамма-функция, бета-функция;
$\mathbf{P}$, $\mathbf{E}$, $\mathbf{D}$	—	вероятность, математическое ожидание, дисперсия;
$\mathbf{P}\{\,\cdot \mid \cdot\,\}$	—	условная вероятность;
$\mathbf{E}[\,\cdot \mid \cdot\,]$	—	условное математическое ожидание;
Ω	—	пространство исходов ω;
$\mathcal{K}$	—	полукольцо (интервалов);
$\mathfrak{C}$	—	кольцо, порождённое классом подмножеств;
σ	—	префикс, связанный со счётными операциями, или стандартное отклонение случайной величины;
$\mathscr{A}$, $\mathscr{F}$	—	σ-алгебра (сигма-алгебра) подмножеств;
$\mathscr{B}$	—	борелевская σ-алгебра;
a, b, c	—	произвольные константы;
$x, y, z,$ s, u, v, t	—	произвольные переменные;
ξ, η, ζ	—	случайные величины, функции на Ω;
$\varepsilon, \varepsilon, \delta$	—	малые величины;
$\mathbb{N}$	—	множество натуральных чисел $1, 2, \ldots$;
$\mathbb{N}_0$	—	множество целых неотрицательных чисел $0, 1, 2, \ldots$;
$\mathcal{N}(\mu, \sigma^2)$	—	нормальный (гауссовский) закон (см. стр. 107);
μ	—	мера или значение математического ожидания сл.в.;
λ	—	мера Лебега или параметр вероятностных законов;
φ	—	характеристическая функция (преобразование Фурье);
ϕ, Φ	—	плотность, функция распределения $\mathcal{N}(0, 1)$ закона;
f, F	—	плотность, ф. распределения сл. величины (вектора);
h, H, g, G	—	функции действительного (векторного) аргумента;
$\dot{\mathbb{I}}$	—	единичная матрица (размерность по контексту);
$\dot{\mathrm{I}}_A = \dot{\mathrm{I}}(A)$	—	индикаторная функция множества A;
$\boldsymbol{\Sigma}$	—	матрица ковариаций;
$\sum, \displaystyle\sum$	—	знаки суммирования; индекс суммирования опускается, если он очевиден из контекста.

Предисловие. В пособии рассматриваются основные теоретико-вероятностные методы, которыми должны, на наш взгляд, свободно владеть выпускники магистратуры и аспирантуры, специализирующиеся в области теории вероятностей. Понятно, что выбор тем, представленных в пособии, весьма субъективен. Краткое описание этих тем, возможно, даст некоторое обоснование их выбора. Во-первых, сразу отметим, что в пособии не рассматриваются комбинаторные задачи классической вероятностной схемы, поскольку литературе, посвящённой этим задачам, несть числа (см. [7, 18, 21, 27]). Во-вторых, по тем же причинам аксиоматика теории вероятностей представлена здесь, во введении, в краткой форме. Другими словами, подразумевается, что читатель уже знаком со структурой вероятностной модели случайного явления как варианта измеримого пространства, наделённого вероятностной мерой.

С другой стороны, в процессе написания пособия оказалось, что для построения более или менее замкнутого курса никак не обойтись без обращения к первоистокам — к математическим основаниям теории вероятностей, т.е. к измеримым пространствам, измеримым отображениям, мерам и интегралу Лебега. Этим вопросам посвящена глава IX, которая содержит подробное описание всех этапов построения меры, а затем и интеграла Лебега в произвольном абстрактном пространстве. Здесь приведены (с доказательствами) все основные факты, потребность в которых возникает при изучении методов теории вероятностей. В частности, теорема Каратеодори о продолжении меры, теорема Фубини о порядке интегрирования, теорема Радона–Никодима об абсолютно непрерывных мерах и многое другое. В связи с этим главу IX можно рекомендовать в качестве самостоятельного дополнительного пособия к курсу теории меры.

В первой главе представлены методы анализа вероятностных распределений систем случайных величин (случайных векторов). В обычных курсах теории вероятностей эти вопросы рассматриваются вскользь, при этом основной метод обработки многомерных величин может быть охарактеризован словами «по аналогии» (имеется в виду аналогия с одномерным случаем). Более тщательный взгляд на эти проблемы показал, что такой метод часто не работает. Например, в одномерном случае непрерывная функция распределения, у которой условие дифференцируемости нарушается в «узкой» области, чаще всего обладает плотностью. Поэтому «механическая» концепция плотности как производной функции распределения здесь редко расходится с истиной. В многомерном случае такая концепция в очень неожиданных ситуациях приводит к ложным выводам. В этой же главе на основе естественных интуитивных соображений вводится понятие условного распределения (условного математического ожидания) для дискретных и абсолютно непрерывных вероятностных моделей. Выделяется ключевое свойство этих конструкций, позволяющее распространить их на вероятностные модели общего вида. Построению условного математического ожидания для общих моделей посвящена глава III.

Далее, в литературе по теории вероятностей недостаточно описаны методы исследования преобразований случайных векторов, хотя это и есть основная задача теории вероятностей (как её понимали авторы многих классических монографий). Вторая глава данного пособия в некоторой степени восполняет этот пробел.

Методы главы VIII, посвящённой случайным процессам, в существенной части опираются на инструментарий, разработанный в первых главах пособия. Здесь рассмотрены вопросы существования процессов и способы их описания, изучаются свойства гауссовских (в частности винеровских), пуассоновских и марковских процессов, а также свойства мартингалов. Часть материала посвящена теории стохастического интегрирования.

Основные задачи теории вероятностей, как известно, связаны со сходимостью случайных величин. Возникающие в связи с этим понятия и некоторые способы их анализа представлены в главе VI. Ясно, что при рассмотрении вопросов слабой сходимости никак не обойтись без классического метода характеристических функций. Возможности применения этого метода изучаются в главе IV. Кроме того, обращение к центральной предельной теореме и гауссовским процессам подразумевает свободное владение аппаратом многомерной нормальной вероятностной модели, которой посвящена глава V.

Наконец, глава VII содержит описание методов доказательства некоторых предельных теорем (закона больших чисел и центральной предельной теоремы) для сумм случайных величин при различных предположениях об их распределениях. Кроме того, в конце этой главы приведены некоторые факты, относящиеся к оценке скорости сходимости в центральной предельной теореме. Часть из этих фактов (неравенство Берри–Эссеена, закон повторного логарифма) приводится с доказательствами.

Отметим, что доказательства многих утверждений, приведённых в пособии, предполагают активное участие читателя. Такие утверждения помечены знаком (?!); в конце каждой главы даются небольшие подсказки к их реализации. Кроме того, некоторые утверждения, важные для понимания материала, представлены по ходу изложения в виде задач (Упр.), указания к решению которых также даны в конце каждой главы.

Структура пособия не вполне соответствует классическому линейному построению учебников, когда каждый новый факт подкреплён фактами, приведёнными на предыдущих страницах учебника. Скажем, доказательства фактов, относящихся к случайным величинам (векторам, гл. I), предполагают свободное обращение с понятиями и свойствами различных классов множеств (гл. IX). Чтение пособия с соблюдением последовательности логических связей может состоять, например, из следующей цепочки: секции C, D, E, F главы IX, гл. IX, гл. I–VIII.

Нумерация всех утверждений в пособии сплошная, включая теоремы, леммы, следствия и задачи. Для удобства ссылки на утверждения содержат указания на соответствующую страницу текста, если это утверждение находится достаточно далеко от места обращения к нему, например 597, стр. 525; в противном случае ссылка содержит только номер утверждения, например формула (22) на одной из ближайших страниц внутри текущей главы. Если ссылка на близлежащую формулу больше нигде не потребуется, то обращение к ней оформляется в виде символов типа $(*)$ или $(\natural)$ и т.п.

В список литературы, приведённый в конце пособия, включены лишь те монографии и учебники, которые в той или иной степени повлияли на содержание данного пособия. Ссылки по тексту приводились лишь на источники тех утверждений, доказательства которых либо здесь не приведены, либо представляют собой прямую цитату. Последних автор по мере сил пытался избежать. В частности, идеи многих примеров и контрпримеров были позаимстовованы из [18, 20, 29] и использованы после некоторой переработки. Достаточно широкая (с образовательной точки зрения) библиография представлена в учебных пособиях А.Н. Ширяева [27, 28].

Автор приносит благодарность профессору Йордану (Данчо) Стоянову, который не просто прочёл первые главы рукописи, но и сделал большое число ценных замечаний, касающихся как стиля изложения, так и точности приводимых утверждений. Оставшиеся в книге огрехи полностью на совести автора.

Введение. Теоретико-вероятностное описание любого эксперимента с непредсказуемым исходом начинается с описания так называемого вероятностного пространства — вероятностной модели, которая должна строиться с учётом механизма случайности в этом эксперименте. Если таковой механизм отсутствует, как, например, в популярном вопросе о возможности встречи со слоном на улицах города или в вопросе о результате экзамена для конкретного студента ещё до начала обучения, то исход эксперимента просто непредсказуем (имеется в виду — со 100%-ной уверенностью), но не относится к области интересов теории вероятностей. Облекать в цифровую форму ответы на подобные вопросы с точной интерпретацией указанного значения нет никакого смысла. В таких ситуациях вместо термина «вероятность» наиболее уместен термин «шанс» и правильные ответы на подобные вопросы должны состоять из фраз типа «шансы небольшие», или «шансы значительные», или «шансы пятьдесят на пятьдесят», или даже «шансы 75%». В двух последних случаях, несмотря на цифровую форму, ответы указывают только на (субъективную) степень уверенности говорящего в возможности осуществления непредсказуемого события.

Отличительной особенностью случайного эксперимента является принципиальная возможность его многократного повторения приблизительно в одних и тех же условиях, при этом отличия условий как раз и определяют случайный механизм этого эксперимента. Кроме того, постулируется, что относительная частота осуществления любого события с ростом числа проведённых экспериментов будет стабилизироваться около некоторой константы, которая и представляет собой числовое выражение понятия «вероятность события».

Вероятностная модель случайного эксперимента начинается с описания всех мыслимых исходов ω этого эксперимента, т.е. с описания так называемого пространства элементарных исходов Ω $(\ni \omega)$. Элементарными они называются потому, что в рамках выбранной модели их дробление на более мелкие части не только не предполагается, но даже запрещено. Появление того или иного элементарного исхода описывается с помощью вероятностной меры $\mathbf{P}$. Любое утверждение об исходе эксперимента формализуется в виде некоторого подмножества $A \subset \Omega$ (мы предполагаем, что читатель знаком с основными понятиями теории множеств; обозначения, необходимые нам для дальнейшего изложения, см. на стр. 520, греческий алфавит и начертания готических букв приведены на стр. 537).

► Дискретные вероятностные модели. В дискретных пространствах Ω с не более чем счётным числом элементов вероятность $\mathbf{P}$ задаётся с помощью *функции вероятностей* $p : \Omega \mapsto \mathbb{R}^1_+$, приписывающей каждому элементарному исходу ω вероятность его осуществления $p(\omega)$. Единственное условие, которому должна удовлетворять функция p (кроме её неотрицательности), чтобы $\sum_{\omega\in\Omega} p(\omega) = 1$. Вероятность любого подмножества $A \subset \Omega$ в этом случае равна сумме вероятностей всех элементарных исходов, входящих в A:

$$\mathbf{P}\{A\} = \sum_{\omega\in A} p(\omega). \tag{$*$}$$

Функцию множеств $(*)$ называют распределением вероятностей на Ω.

Классики теории вероятностей занимались исследованиями вероятностных моделей с конечным пространством Ω, в котором все исходы имеют одинаковую вероятность: $p(\omega) = {}^1/_N$, $N = \#(\Omega) < \infty$. Такие дискретные вероятностные модели с равномерным распределением вероятностей принято называть классическими. На заре становления теории вероятностей все проблемы сводились в основном к умению правильно подсчитывать количество элементарных исходов, соответствующих тому или иному событию. Большинство парадоксов теории вероятностей объясняется либо ошибками в таких подсчётах, либо неоправданной уверенностью в равновозможности всех элементарных исходов (в книге Г. Секей [19] эти парадоксы весьма наглядно разобраны с точки зрения современных концепций случайности).

В рамках классической модели абсолютно естественно выглядит формула условной вероятности $\mathbf{P}\{A \,|\, B\} = \mathbf{P}\{AB\} \,/\, \mathbf{P}\{B\} = \#(AB)/\#(B)$. Если известно, что произошло событие B, то в этом случае все элементарные исходы, входящие в B, могут быть взяты за основу нового пространства элементарных исходов $\Omega' = B$, в котором в полном соответствии с классической схемой вероятность любого события $C \subset \Omega'$ вычисляется по формуле $\mathbf{P}'\{C\} = \#(C)/\#(\Omega')$. После введения понятия условной вероятности также естественным становится определение понятия независимости событий $\mathbf{P}\{AB\} = \mathbf{P}\{A\}\,\mathbf{P}\{B\}$. Независимость может быть как следствием способа построения вероятностной модели, так и выступать постулатом, используемым для такого построения. Например, если эксперимент с подбрасыванием двух монет представить в виде пространства $\Omega = \{(0,0),(0,1),(1,0),(1,1)\}$

с равновероятными элементарными исходами, то любое событие, описывающее исход второго бросания (второго элемента вектора), не зависит от любого исхода первого бросания. Обратно, если постулировать факт независимости двух бросаний (с равными вероятностями «герба» и «решётки»), то в результате мы придём к тому же классическому вероятностному пространству из четырёх элементов (см. далее схему Бернулли).

Дальнейшее обобщение понятия независимости связывает классы (совокупности) событий: говорят, что классы событий $\mathcal{A}_j = \langle A_{j,1}, A_{j,2}, \ldots \rangle$, $j = 1, 2 \ldots, J$, независимы, если для любых наборов $A_{j_m,k_m} \in \mathcal{A}_{j_m}$, $m = 1, \ldots, M$ ($\leqslant J$), справедлива формула умножения вероятностей:

$$\mathbf{P}\Big\{ \bigcap\nolimits_{m=1}^{M} A_{j_m,k_m} \Big\} = \prod\nolimits_{m=1}^{M} \mathbf{P}\{A_{j_m,k_m}\}.$$

Например, вполне естественно, что любой набор событий, описывающих результаты подбрасывания монеты в первых трёх экспериментах, не зависит от результатов пятого и седьмого подбрасывания, а также от наборов событий, связанных с результатами экспериментов, начиная с десятого подбрасывания. Отметим, что выбор классов независимых событий возможен только в пространстве элементарных исходов Ω достаточно сложной структуры.

► Многие вероятностные модели с различающимися вероятностями исходов могут рассматриваться как производные от некоторого классического вероятностного пространства. Большое количество задач, возникающих при анализе подобных моделей, изучены в первых главах монографий [21], [27].

Одна из таких моделей возникает в так называемой урновой схеме, когда из генеральной совокупности K объектов отбирается n объектов. Можно считать, что объекты генеральной совокупности занумерованы числами от 1 до K. Если объекты выбираются поочерёдно, то результат всего эксперимента (элементарный исход) можно представить в виде вектора $\omega = (x_1, \ldots, x_n)$, где x_j — идентификационный номер j-го элемента выборки. Рассматривают две схемы выбора — с возвращением и без возвращения.

В случае выбора без возвращения все элементы вектора различаются между собой. Легко понять, что количество таких векторов равно числу «размещений» $\mathrm{A}_K^n = K(K-1)\cdots(K-n+1) = K!/(K-n)!$. Если эксперимент организован так, что выбор каждого очередного объекта производится с равной вероятностью из всей совокупности имеющихся к этому моменту объектов, то мы с полным правом можем считать равновероятными все варианты отбора, т.е. вероятность $p(x_1, \ldots, x_n) = 1/\mathrm{A}_K^n$. Такое вероятностное пространство используется в основном для решения вопросов, связанных с порядком поступления объектов в выборку.

1| Упр. Докажите, что с вероятностью $1/n!$ будет получена выборка, в которой все элементы расположены в порядке возрастания, т.е. каждый последующий элемент больше предыдущего.

► Чаще всего рассматривается ситуация, когда объекты генеральной совокупности разделены на два сорта, — представляет интерес вопрос о количестве объектов первого сорта среди выбранных объектов. В этом случае порядок поступления элементов выборки не существенен, поэтому можно все выборки одинакового состава объединить в один элементарный исход $(x_{(1)}, \ldots, x_{(n)})$, где $x_{(1)} < \ldots < x_{(n)}$, т.е. перед фиксацией исхода всего эксперимента объекты выборки упорядочиваются. Подобная ситуация встречается также в случае, когда элементы выборки поступают не последовательно, а все сразу (одной «горстью»). Поскольку каждая выборка фиксированного состава может быть получена одинаковым количеством $n!$ способов из неупорядоченной выборки, то снова можно предполагать равную вероятность всех таких элементарных исходов. Другими словами, вероятность любого исхода $(x_{(1)}, \ldots, x_{(n)})$ равна $1/\mathrm{C}_K^n$, где биномиальный коэффициент $\mathrm{C}_K^n = \mathrm{A}_K^n/n!$.

Определение. Вероятностная модель $(\Omega, \mathbf{P})$, где $\Omega = \{0, 1, \ldots, n\}$ и

$$p(j) = \frac{\mathrm{C}_M^j\, \mathrm{C}_{K-M}^{n-j}}{\mathrm{C}_K^n}, \qquad j \in \Omega,$$

называется гипергеометрической моделью. Такая модель описывает количество объектов первого сорта в выборке объема n в схеме выбора без возвращения из совокупности, содержащей M объектов первого сорта и $K - M$ объектов второго сорта.

✓ В схеме выбора с возвращением объектов в генеральную совокупность элементы выборки могут повторяться, поэтому количество элементарных исходов (векторов), в которых важен порядок поступления, равно K^n .

Определение. Вероятностная модель $(\Omega, \mathbf{P})$, где $\Omega = \{0, 1, \ldots, n)$ и ф.вер.

$$p(j) = \mathrm{C}_n^j\, \frac{M^j (K-M)^j}{K^n} = \mathrm{C}_n^j\, q^j (1-q)^{n-j}, \qquad j \in \Omega,\ q = {}^M\!/_K,$$

называется биномиальной моделью $\mathcal{B}\mathtt{in}(n, q)$. Такая модель описывает количество объектов первого сорта в выборке объема n в схеме выбора с возвращением из совокупности, в которой относительная доля объектов первого сорта равна q.

Легко понять, что при большом объёме генеральной совокупности и относительно малом объёме выборки значения вероятностей исходов в гипергеометрической модели мало отличаются от соответствующих значений вероятностей биномиальной модели. Другими словами,

$$\text{(?!)} \qquad \lim_{K\to\infty, M/K\to q} \frac{\mathrm{C}_M^{\,j}\,\mathrm{C}_{K-M}^{\,n-j}}{\mathrm{C}_K^{\,n}} = \mathrm{C}_n^{\,j}\; q^j(1-q)^{n-j}, \qquad j \in \{0, 1, \ldots, n\}.$$

► Другой подход к определению биномиальной модели связан с понятием независимости экспериментов в схеме Бернулли. Предполагается, что проводится n испытаний, в каждом из которых:

а) наблюдается один из двух возможных исходов, обычно называемых успехом и неудачей и обозначаемых 1 и 0 соответственно;

б) вероятность успеха в каждом испытании одинакова и равна q;

в) исходы испытаний независимы, т.е. вероятность любого сочетания элементов выборочного вектора равна произведению соответствующих вероятностей: $\mathbf{P}\{(x_1, \ldots, x_n)\} = \prod_1^n q^{x_j}(1-q)^{1-x_j}$.

Случайное число S_n, равное количеству испытаний, закончившихся успехом (из общего количества n испытаний), подчинено биномиальному распределению. Если обозначить через ξ_j, $j = \overline{1, n}$, набор независимых бернуллиевских сл.в., принимающих с вероятностью q значение 1 и с вероятностью $(1-q)$ значение 0, то биномиальная сл.в. имеет методически важное представление:

$$S_n = \sum\nolimits_1^n \xi_j.$$

✓ Если количество испытаний в схеме Бернулли не фиксировать заранее, а эксперимент проводить до первого успеха, то получим так называемую геометрическую модель.

Определение. Вероятностная модель $(\Omega, \mathbf{P})$ с $\Omega = \mathbb{N} = \{1, 2 \ldots\}$ и ф.вер.

$$(\natural) \qquad p(j) = q(1-q)^{j-1}, \qquad j \in \mathbb{N},$$

предназначенная для описания количества проведённых испытаний в схеме

Бернулли до первого успеха (включительно), называется геометрической. Случай ожидания более одного успеха приводит к вероятностной модели Паскаля (другое название — отрицательно-биномиальная модель).

Среди распределений в пространстве натуральных чисел $\Omega = \mathbb{N}$ геометрическая модель полностью характеризуется свойством «отсутствия последействия»: если эксперимент не завершился при k-м испытании, то (условная) вероятность того, что он завершится на каком-то следующем ($(k+j)$-м) шаге, вычисляется по той же формуле $(\natural)$.

► Другая модель со счётным числом исходов возникает в ситуациях, соответствующих условиям теоремы Пуассона: для биномиальной модели с «большим» количеством испытаний $(n \to \infty)$ и «малой» вероятностью успеха $\big(q = q_n \to 0,\ nq \to \lambda \in (0;\infty)\big)$ при $\forall j \in \mathbb{N}_0 = \{0, 1, 2\ldots\}$ справедливо асимптотическое равенство (?!)

$$\lim_n \mathrm{C}_n^j\, q^j (1-q)^{n-j} = \frac{\lambda^j}{j!} e^{-\lambda}. \tag{1}$$

Определение. Модель $(\Omega, \mathbf{P})$ с $\Omega = \mathbb{N}_0 = \{0, 1, 2\ldots\}$ и ф.вер $p(j) = \frac{\lambda^j}{j!} e^{-\lambda}$, $j \in \mathbb{N}_0$, называется вероятностной моделью Пуассона.

► В ситуации, когда перед фиксацией результата эксперимента в схеме выбора с возвращением компоненты вектора упорядочиваются (порядок поступления объектов в выборку игнорируется), мы имеем меньше оснований считать все элементарные исходы равновероятными, поскольку в этом случае различные элементарные исходы могут быть получены путём упорядочивания различного количества неупорядоченных выборок. Например, элементарный исход $(1, 1, 1)$ может быть получен только одним способом, исход $(1, 2, 2)$ — тремя способами из векторов $(1, 2, 2), (2, 1, 2), (2, 2, 1)$, а исход $(1, 2, 3)$ — шестью способами. В реальной жизни классическая вероятностная модель с таким пространством элементарных исходов встречается нечасто. В статистической физике, где вероятностная модель строится на основе экспериментальных наблюдений, эта модель оказалась пригодной для описания поведения некоторых элементарных частиц ([27]).

► Вероятностные модели с несчётным числом исходов. В пространствах более сложной структуры, например на числовой прямой, способ задания вероятности через функцию вероятностей приводит к моделям, эквивалентным, по существу, дискретным моделям, поскольку равенство $\sum_{\omega\in\Omega} p(\omega) = 1$ возможно только тогда, когда функция p отлична от нуля на не более чем счётном подмножестве $\omega \in \Omega$ (каковое и можно выбрать тогда в качестве пространства Ω, если нас не интересуют остальные точки исходного Ω). Расцвет теории вероятностей начался с того момента, когда было понято, что задание вероятности удобнее проводить в рамках теории меры, т.е. определять вероятность не просто для каждого элементарного исхода, а для всего набора так называемых измеримых событий — подмножеств Ω, для которых мы можем вычислить эту вероятность.

Меру, определённую для всех подмножеств исследуемого пространства Ω, можно задать только способом, подобным способу описания дискретных моделей. Так, например, считающая мера для любого подмножества $A \subset \Omega$ полагается равной количеству точек в A из некоторого заранее выбранного подмножества $\mathcal{S}$: $\mu(A) = \#(A \cap \mathcal{S})$. Заметим, что если количество точек $\mathcal{S}$ бесконечно, то эта мера может принимать значение $+\infty$. При конечном $\mathcal{S}$ считающая мера описывает эксперимент в классической схеме. В теории наиболее востребована мера, считающая целые точки числовой прямой (точки целочисленной решётки какого-либо конечномерного пространства). Мера Дира́ка есть считающей мера с подмножеством $\mathcal{S} = \{x_0\}$, состоящим только из одной точки, например точки $x_0 = 0$. Другими словами, мера Дира́ка $\delta_{x_0}(A) = 1$, если $A \ni x_0$, и $\delta_{x_0}(A) = 0$ в противном случае.

Если выделенный набор $\mathcal{S}$ содержит не более чем счётное число элементов, то, приписав каждому элементу $\omega \in \mathcal{S}$ вес $p(\omega) \geqslant 0$, можно определить меру для любого подмножества $A \subset \Omega$ с помощью равенства

$$\mu(A) = \sum_{\omega\in A\cap\mathcal{S}} p(\omega).$$

Нормированная весовая функция, т.е. функция p с условием $\sum_{\omega\in S} p(\omega) = 1$, задаёт уже вероятностную меру — вероятность. Такой способ переноса вероятности с дискретного $\mathcal{S}$ на более широкое пространство Ω удобен с теоретической точки зрения, поскольку позволяет исследовать соотношения между формально различными вероятностными моделями. Например, теорема Муавра–Лапласа, изначально воспринимаемая как аппроксимация биномиальных вероятностей с помощью некоторых интегральных выражений, теперь может трактоваться в терминах близости биномиальной и нормальной (гауссовской) меры.

► **Общая схема** построения меры на подмножествах пространства Ω обычно включает в себя несколько этапов (см. главу IX). Сначала мера определяется на некотором «естественном» классе подмножеств, после чего осуществляется расширение этого определения на другие подмножества Ω. Например, на числовой прямой можно определить меру сначала на классе (полукольце, стр. 521) $\mathcal{K}$ интервалов вида $(a; b]$, при этом мера должна удовлетворять ключевому свойству сигма-аддитивности (стр. 424); далее мера естественным способом распространяется с сохранением свойства сигма-аддитивности на класс (кольцо, стр. 522) $\mathcal{A}$ подмножеств, представимых в виде конечного объединения непересекающихся интервалов из $\mathcal{K}$. На следующем шаге мера продолжается на подмножества, которые могут быть получены из элементов $\mathcal{A}$ с помощью некоторого «предельного перехода».

Способ продолжения, предложенный М. Жорданом, опирается на конечные объединения (или пересечения) элементов $\mathcal{A}$. При таком способе класс измеримых подмножеств не замкнут относительно счётных объединений, что неудобно для целей теории вероятностей. Подход А. Лебега позволяет сохранить свойство сигма-аддитивности меры и определяет эту меру на подмножествах некоторой сигма-алгебры (стр. 522), содержащей кольцо $\mathcal{A}$. Таким образом, лебегово продолжение любой меры (теорема Каратеодори 475, стр. 433), заданной вначале на каком-либо «естественном» полукольце $\mathcal{K}$ (или кольце $\mathcal{A}$), будет определено на минимальной сигма-алгебре (593, стр. 524), содержащей $\mathcal{K}$ (или $\mathcal{A}$).

На числовой прямой (в пространстве $\mathbb{R}^k$) минимальная сигма-алгебра, содержащая класс интервалов (параллелепипедов), совпадает с минимальной сигма-алгеброй $\mathscr{B}(\mathbb{R}^k)$, порождённой всеми открытыми подмножествами Ω (см. 481, стр. 437). Такая сигма-алгебра называется борелевской в честь Э. Бореля, который первым распространил определение меры на эту сигма-алгебру подмножеств. Стало быть, если не задаваться вопросами построения меры, то можно сразу считать, что мера задана на некоторой сигма-алгебре $\mathscr{F}$ подмножеств Ω (борелевской сигма-алгебре, если Ω метрическое или топологическое пространство, или сигма-алгебре всех подмножеств $\mathscr{F} = \mathscr{P}(\Omega)$, если Ω дискретное пространство) и удовлетворяет свойству сигма-аддитивности. Вероятностная мера, кроме всего прочего, должна удовлетворять свойству нормированности: $\mathbf{P}\{\Omega\} = 1$. Тройка объектов $(\Omega, \mathscr{F}, \mathbf{P})$ называется вероятностным пространством. Вероятностное пространство представляет собой формализованное описание стохастического (случайного) эксперимента.

Один из способов построения меры на числовой прямой использует неубывающую непрерывную справа функцию $F(x)$, $x \in \mathbb{R}^1$ (стр. 441). При этом для интервалов полагается $\mu_F(a;c] = F(c) - F(a)$ (если функция F непрерывна слева, то на первом шаге полагается $\mu_F[a;c) = F(c) - F(a)$). Мера μ_F, продолженная на борелевскую сигма-алгебру $\mathbb{R}^1$, называется мерой Лебега–Стилтьеса. Собственно мера Лебега λ есть мера Лебега–Стилтьеса с порождающей функцией $F(x) = x$, т.е. при её определении полагается $\lambda(a;c] = c - a$. Функция F, удовлетворяющая дополнительно условию $F(-\infty) = 1 - F(+\infty) = 0$, задаёт вероятностную меру, и в этом случае F называется функцией распределения, а мера Лебега–Стилтьеса μ_F — распределением вероятностей.

► **Случайные величины**. Чаще всего наблюдается не сам исход ω, а его преобразование с помощью некоторой действительной функции $\xi : \Omega \mapsto \mathbb{R}^1$. Вычислить вероятность (меру) какого-либо утверждения вида $\xi \in B$, где B — некоторое подмножество $\mathbb{R}^1$, можно только в том случае, когда прообраз этого подмножества при отображении ξ, т.е. подмножество $\xi^{-1}(B) := \{\omega \in \in \Omega : \ \xi(\omega) \in B\}$ исходного пространства Ω, измерим относительно рассматриваемой меры. Вместо того чтобы с каждым преобразованием ξ связывать класс подобных подмножеств числовой прямой, удобнее, наоборот, сначала выделить на $\mathbb{R}^1$ семейство естественных подмножеств (обычно это борелевская сигма-алгебра), и затем рассматривать только такие функции ξ, у которых измеримы прообразы $\xi^{-1}(B)$ любых подмножеств B из выделенного семейства. Такие функции называются измеримыми или случайными величинами, если исходное пространство Ω снабжено вероятностной мерой. Заметим, что если класс измеримых подмножеств Ω совпадает с совокупностью всех подмножеств $\mathscr{P}(\Omega)$, то, очевидно, любое преобразование измеримо.

Поскольку лебегово продолжение σ-аддитивной меры выделяет класс измеримых подмножеств, образующих σ-алгебру, понятие измеримости функции также связывают с σ-алгебрами. Функция $\xi : \Omega_1 \mapsto \Omega_2$ измерима относительно σ-алгебры $\mathscr{F}_1$ на Ω_1 и σ-алгебры $\mathscr{F}_2$ на Ω_2, кратко, $(\mathscr{F}_1, \mathscr{F}_2)$-измерима, если для $\forall B \in \mathscr{F}_2$ прообраз $\xi^{-1}(B) \in \mathscr{F}_1$. Если $\Omega_2 \subset \mathbb{R}^k$ и σ-алгебра $\mathscr{F}_2 = \mathscr{B}(\Omega_2)$, т.е. $(\Omega_2, \mathscr{F}_2)$ есть борелевское пространство, то ξ называется измеримым по Борелю. Можно показать, что для этого достаточно, чтобы $\{\xi \leqslant x\} := \{\omega : \ \xi(\omega) \leqslant x\} \in \mathscr{F}_1$ при $\forall x \in \mathbb{R}^k$ (стр. 465). Если, кроме того, и пространство $(\Omega_1, \mathscr{F}_1)$ борелевское, то отображение ξ называют борелевским.

▶ **Интеграл Лебега. Математическое ожидание**. С измеримой функцией ξ, заданной на пространстве с мерой $(\Omega, \mathscr{F}, \mu)$ и принимающей значения в $\mathbb{R}^1$, можно связать понятие интеграла (интеграла Лебега) относительно меры μ. Определение интеграла Лебега начинается с так называемых простых функций ξ, принимающих конечное число значений, скажем, $\xi \in \{x_1, \ldots, x_k\}$. Для таких функций интеграл полагается равным (стр. 472) среднему взвешенному значению:

$$\int_\Omega \xi \, d\mu \ \Big(= \int_\Omega \xi(\omega)\, \mu(d\omega)\Big) \ := \sum\nolimits_{j=1}^k x_j\, \mu(\omega : \xi = x_j).$$

Затем рассматриваются неотрицательные измеримые функции ξ (стр. 473). Показывается, что для любой монотонно возрастающей последовательности неотрицательных простых функций, приближающих ξ снизу, существует единственный предел (конечный или бесконечный) последовательности интегралов от этих простых функций, каковой и берётся в качестве интеграла от ξ. В соответствии со сказанным, интеграл Лебега от неотрицательной измеримой по Борелю функции ξ может быть определён как

$$\int_\Omega \xi \, d\mu := \lim_{n\to\infty} \Big[\sum_{j=1}^{n2^n} \frac{j-1}{2^n} \mu\Big(\frac{j-1}{2^n} < \xi \leqslant \frac{j}{2^n}\Big) + n\mu\big(\xi > n\big)\Big].$$

2| Упр. Докажите, что если интеграл конечен, то $\lim_n n\, \mu\big(\xi > n\big) = 0$.

Для произвольной измеримой функции используется представление в виде разности положительной и отрицательной частей: $\xi = \xi^+ - \xi^-$, $\xi^\pm = \max\{\pm\xi, 0\}$, и полагается

$$\int_\Omega \xi\, d\mu := \int_\Omega \xi^+\, d\mu - \int_\Omega \xi^-\, d\mu,$$

если здесь нет неопределённости вида $\infty - \infty$; в этом случае говорят, что интеграл существует.

3| Упр. Докажите, что интеграл $\int \xi\, d\mu$ от измеримой функции ξ конечен (функция интегрируема) т. т. т. когда $\int |\xi|\, d\mu < \infty$.

▶ Интеграл относительно вероятностной меры $\mathbf{P}$ называется математическим ожиданием. Операция вычисления математического ожидания обозначается символом $\mathbf{E}$:

$$\mathbf{E}\,\xi := \int_\Omega \xi(\omega)\ \mathbf{P}(d\omega).$$

Интеграл относительно меры Лебега–Стилтьеса μ_F на $\mathbb{R}^1$, порождённой функцией F, называется интегралом Лебега–Стилтьеса и обозначается

$$\int_{\mathbb{R}} \xi\, dF \,\Big(= \int_{\mathbb{R}} \xi(x)\, dF(x)\Big) \;:= \int_{\mathbb{R}} \xi\, d\mu_F.$$

Аналогия с записью интеграла Римана–Стилтьеса здесь не случайна. Для большого класса функций ξ и порождающих функций F интеграл Лебега–Стилтьеса совпадает с интегралом Римана–Стилтьеса (см. 566, стр. 497). Например, если F всюду непрерывна и её производная $f(x) = F'(x)$ существует и непрерывна всюду, то по теореме Радона–Никодима (558, стр. 491)

$$\int_{\mathbb{R}} \xi(x)\, dF(x) \;= \int_{-\infty}^{\infty} \xi(x) f(x)\, dx,$$

где интеграл уже можно понимать как несобственный интеграл Римана. Для того чтобы он был конечен (и совпадал с интегралом Лебега), необходимо требовать его абсолютную сходимость (см. предыдущее упражнение).

⚠ Определение математического ожидания через интеграл Лебега, а не через интеграл Римана представляется более естественным обобщением понятия среднего значения, поскольку и при вычислении среднего значения, и при построении интеграла Лебега суммирование ведётся по значениям самой функции, а не по значениям аргумента. Кроме того, как показано в замечании 302, стр. 261, конечность интеграла Римана не обеспечивает возможность применения закона больших чисел, связывающего понятия математического ожидания и среднего арифметического.

► Если задано вероятностное пространство $(\Omega, \mathscr{F}, \mathbf{P})$, то сл.величина $\xi : \Omega \mapsto \mathbb{R}^1$ порождает (индуцирует) вероятностную меру на измеримом борелевском пространстве $(\mathbb{R}^1, \mathscr{B})$ по формуле $\mathsf{P}_\xi(B) = \mathbf{P}\{\xi \in B\}$, $B \in \mathscr{B}$. Эта мера называется распределением случайной величины. В дальнейшем для идентификации распределения сл.в. будем использовать обозначение вида $\xi \backsim \mathsf{P}$. Часто символ распределения P заменяется его мнемоническим обозначением, например: $\xi \backsim \mathcal{P}\mathtt{ois}(\gamma)$ означает, что сл.в. ξ имеет распределение Пуассона с параметром γ (список популярных распределений см. на стр. 107). Кроме того, запись вида $\xi \backsim \eta$ будет означать, что индуцированные распределения сл.в. ξ, η совпадают.

Распределение ξ можно описать с помощью чуть меньшего класса событий из $\mathbb{R}^1$, а именно класса интервалов вида $(-\infty; x]$, $x \in \mathbb{R}^1$. Непрерывная справа функция $F(x) = \mathsf{P}\{(-\infty; x]\} = \mathbf{P}\{\xi \leqslant x\}$, $x \in \mathbb{R}^1$, называется функцией распределения (ф.р.) случайной величины ξ.

4| Упр. Пусть $\Omega = [0;1]$, $\mathscr{B}$ — борелевская σ-алгебра, $\mathbf{P}$ — мера Лебега. Покажите, что сл.величины $\xi = \omega$ и $\eta = 1 - \omega$ имеют одинаковое распределение: $\xi \backsim \eta$.

Таким образом, случайную величину можно определить как реализацию случайного эксперимента с действительным исходом, значения которой появляются в соответствии с вероятностным распределением P (или в соответствии с ф.р. F). Подобная интерпретация возможна ещё и потому, что любые числовые характеристики случайной величины, как то: вероятность некоторого события или математическое ожидание по теореме о замене переменных (стр. 483) могут быть вычислены либо как интеграл Лебега от некоторой функции $h(x)$ относительно меры P, либо как интеграл Лебега от функции $h(\xi(\omega))$ относительно меры $\mathbf{P}$:

$$\mathbf{E}\, h(\xi) := \int_\Omega h(\xi(\omega))\ \mathbf{P}(d\omega) = \int_{\mathbb{R}} h(x)\ \mathsf{P}(dx) \ \stackrel{\text{df}}{=}\ \int_{\mathbb{R}} h(x)\, dF(x).$$

Если математическое ожидание вычисляется по некоторому подмножеству $A \subset \Omega$, то обычно применяется запись интеграла Лебега следующего вида:

$$\int_A h(\xi(\omega))\, d\,\mathbf{P} := \mathbf{E}\left[h(\xi)\, \dot{\mathrm{I}}_A\right].$$

Для событий вида $A = \{\omega : \xi(\omega) \in B\}$, $B \in \mathscr{B}$,

$$\int_B h(x)\, dF(x) := \mathbf{E}\left[h(\xi)\, \dot{\mathrm{I}}_A\right]$$

(о нюансах интерпретации такой записи см. замечание 569, стр. 498). В частности, вероятность попадания сл.в. в любое борелевское подмножество B может быть представлена в виде интеграла: $\mathbf{P}\{\xi \in B\} = \int_B dF(x)$.

⚠ Распределение сл.в. может быть описано и с помощью открытых интервалов вида $(-\infty; x)$; в этом случае ф.р. определяется как (непрерывная слева) функция $F(x) = \mathbf{P}\{\xi < x\}$, $x \in \mathbb{R}^1$.

В этом пособии мы будем придерживаться способа определения с непрерывной справа ф.р., что связано только с удобством интерпретации многомерных неравенств вида $\vec{x} \leqslant \vec{y}$ и $\vec{x} < \vec{y}$ (см. замечание на стр. 442).

Отметим ещё, что привычную с университетской скамьи запись интеграла вида $\int_a^b h(x)\, dF(x)$ можно использовать без опасения ошибки для обозначения интеграла Лебега–Стилтьеса, только если ф.р. F непрерывна на обоих концах области интегрирования. В противном случае значение интеграла зависит от того, включены крайние точки в область интегрирования или нет. Для того чтобы исключить возможность двоякой интерпретации, в сомнительных ситуациях мы будем использовать более информативную запись вида $\int_{(a;b]} h\,dF$.

► Чаще всего рассматривают распределения двух основных типов: дискретного и абсолютно непрерывного. Дискретный тип определяется не более чем счётным набором точек $\mathcal{S} \subset \mathbb{R}^1$ и функцией вероятностей $p(x)$, $x \in \mathcal{S}$ (см. описание выше). Для абсолютно непрерывного типа существует функция (плотность) $f(x) \geqslant 0$, $x \in \mathbb{R}^1$, такая, что математическое ожидание

$$\mathbf{E}\left[h(\xi)\right] = \int_{\mathbb{R}} h(x)f(x)\,dx$$

для любой неотрицательной измеримой функции h, где интеграл, вообще говоря, понимается как интеграл Лебега относительно меры Лебега (с генерирующей функцией $G(x) = x$). В частности, функция распределения

$$F(z) = \int\limits_{(-\infty;z]} f(x)\,dx = \int\limits_{-\infty}^{z} f(x)\,dx, \qquad z \in \mathbb{R}^1,$$

где последнее равенство символизирует возможность вычисления этого интеграла как несобственного интеграла Римана, если таковой существует (например, когда плотность кусочно-непрерывна). Плотность распределения абсолютно непрерывного типа почти во всех (по мере Лебега) точках $x \in \mathbb{R}^1$ совпадает с производной ф.р.: $f(x) = F'(x)$; в точках, где производная не существует (мера Лебега таких точек у монотонной функции F равна нулю), значение плотности можно выбрать произвольным образом, исходя, например, из удобства записи.

Заметим, что и для сл.в. с дискретным типом распределения математическое ожидание можно записать в виде интеграла относительно ф.вер. ξ:

$$\mathbf{E}\left[h(\xi)\right] = \int\limits_{\mathbb{R}} h(x)p(x)\,d\mu_{\mathcal{S}} \overset{\mathrm{df}}{=} \sum_{x\in\mathcal{S}} h(x)p(x),$$

где $\mu_{\mathcal{S}}$ — мера, считающая точки множества $\mathcal{S}$ возможных значений ξ. Такая запись означает, что ф.вер. $p(x)$ есть плотность (производная Радона–Никодима) распределения ξ относительно меры $\mu_{\mathcal{S}}$; $p(x) = 0$, $x \notin \mathcal{S}$.

► Выбор той или иной модели осуществляется либо с привлечением неких «физических» (не вероятностных) соображений, либо на основе чисто вероятностного анализа. Например, если рассмотреть дискретное равномерное распределение среди чисел, представимых в виде $k/10^n$, $k = 0, 1, \ldots, 10^n$ (такие числа генерируют стандартные датчики случайных чисел), то легко видеть, что ф.р. этого вероятностного закона при $n \to \infty$ приближается к функции $F(x) = x$, $x \in [0; 1]$.

Определение. Распределение абсолютно непрерывного типа на числовой прямой, определяемое функцией плотности, равной 1 при $x \in [0; 1]$ и 0 при $x \notin [0; 1]$, называется равномерным распределением в отрезке $[0; 1]$.

Равномерное распределение можно рассматривать как аналог классического дискретного распределения в классе абсолютно непрерывных моделей. Характеризационное свойство этого распределения состоит в том, что любые два интервала равной длины равновероятны. В несколько вольной интерпретации это можно переформулировать в том смысле, что равномерная сл.в. принимает все значения с равной вероятностью.

✓ Перенеся характеризационное свойство «отсутствия последействия» геометрического распределения с пространства целых чисел на положительную часть числовой прямой, получим так называемую модель показательного распределения, которая применяется для описания срока службы (момента отказа) ξ приборов, для которых выход из строя происходит не в результате старения, а в силу неких «форс-мажорных» обстоятельств.

Определение. Модель $(\Omega, \mathscr{B}, \mathbf{P})$ с $\Omega = \mathbb{R}^1_+$, борелевской σ-алгеброй $\mathscr{B}$ и абсолютно непрерывным распределением $\mathbf{P}$ с плотностью

$$f(x) = \frac{1}{\lambda} e^{-x/\lambda}, \quad x \geqslant 0,$$

называется показательной (экспоненциальной) вероятностной моделью.

<u>**5**</u>| Упр. Покажите, что для показательного распределения $\xi \backsim \mathcal{E}\mathrm{x}(\lambda)$ условная вероятность $\mathbf{P}\big\{\xi > x + t \,\big|\, \xi > t\big\}$ выхода из строя после момента времени $x + t$ при условии, что объект дослужил до момента t, не зависит от t. Обратно, если условная вероятность обладает таким свойством и $\mathbf{P}\{\xi > 0\} = 1$, то $\xi \backsim \mathcal{E}\mathrm{x}(\lambda)$ с некоторым $\lambda > 0$.

Заметим, что в некотором смысле показательная модель является предельной для геометрической модели (см. <u>279</u>, стр. 238).

<u>**6**</u>| Упр. Докажите, что если $\xi \backsim \mathcal{E}\mathrm{x}(\lambda)$, то целая часть $[\![\xi]\!]$ (с округлением вверх) имеет геометрическое распределение.

В теории надёжности коэффициент

$$\gamma(t) = \lim_{\delta\to 0} \frac{\mathbf{P}\{t < \xi \leqslant t+\delta\}}{\delta\,\mathbf{P}\{\xi > t\}}, \qquad t \geqslant 0,$$

называется интенсивностью отказов среди объектов, дослуживших до момента t. Для показательной модели $\xi \backsim \mathcal{E}\mathrm{x}(\lambda)$ интенсивность отказов не зависит от срока службы (объекты «не стареют»): $\gamma(t) \equiv \frac{1}{\lambda}$. В этом случае $\lambda = \mathbf{E}\xi$, т.е. среднее время службы обратно пропорционально интенсивности отказов.

<u>7</u>| Упр. Докажите, что показательный закон полностью характеризуется свойством постоянства интенсивности отказов:

$$\xi \backsim \mathcal{E}\mathrm{x}(\lambda) \quad \Leftrightarrow \quad \gamma(t) = \frac{1}{\lambda},\ \forall\, t > 0, \quad \text{и} \quad \mathbf{P}\{\xi > 0\} = 1.$$

Если распределение срока службы имеет плотность и по некоторым соображениям интенсивность отказов можно описать некоторой функцией γ, то, решая дифференциальное уравнение $\gamma(t) = F'(t)/(1 - F(t)$, легко получаем, что функция надежности срока службы

$$1 - F(t) = \exp\Big\{ -\int_0^t \gamma(x)\,dx \Big\}, \qquad t \geqslant 0.$$

Понятно, что только при условии $\int_0^\infty \gamma(x)\,dx = +\infty$ распределение будет собственным, описывающим сл.в. ξ, принимающую с вероятностью 1 конечные значения: $\mathbf{P}\{\xi < +\infty\} = 1$.

✓ Знаменитая теорема Муавра–Лапласа утверждает, что для биномиальной модели $\xi \backsim \mathcal{B}\mathtt{in}(n,q)$ вероятность того, что среднее количество успехов ${}^{\xi}/_{n}$ в n испытаниях будет принадлежать интервалу $J_n = [q + \frac{a}{\sqrt{n}}\sigma\,; q + \frac{b}{\sqrt{n}}\sigma\,]$, где $\sigma = \sqrt{q(1-q)}$, удовлетворяет асимптотическому равенству

$$\lim_n \mathbf{P}\Big\{ \frac{\xi}{n} \in J_n \Big\} = \lim_n \sum_{j/n \in J_n} \mathbf{P}\{j\} \;=\; \Phi(b) - \Phi(a),$$

где Φ (греческое «фи» или старорусское «ферт») есть ф.р. так называемого стандартного нормального закона:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{1}{2}t^2}\,dt.$$

Определение. Модель $(\Omega, \mathscr{B}, \mathbf{P})$ с $\Omega = \mathbb{R}^1$, борелевской σ-алгеброй

$\mathscr{B}$ и абсолютно непрерывным распределением $\mathbf{P}$, определяемым плотностью

$$\phi(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}x^2}, \quad x \in \mathbb{R}^1,$$

называется стандартной нормальной (гауссовской) вероятностной моделью.

Характеристики рассмотренных выше законов, а также ряда других популярных законов распределения вероятностей приведены в конце главы I.

► Указания к решению задач.

✦ 1, стр. 16. Из фиксированной выборки различных элементов, упорядоченных по возрастанию, можно получить $n!$ различных неупорядоченных выборок одинакового состава.

✦ (?!), стр. 17. Заметить, что, например, биномиальный коэффициент

$$\mathrm{C}\,{}^{n-j}_{K-M} = \frac{K^{n-j}}{(n-j)!} \prod_{i=0}^{n-j-1} \left(1 - \frac{M}{K} + \frac{i}{K}\right) \asymp \frac{K^{n-j}}{(n-j)!}(1-q)^{n-j}.$$

Аналогично для других биномиальных коэффициентов.

✦ (1), стр. 18. Воспользоваться представлением для биномиальной вероятности:

$$\mathrm{C}\,{}^{j}_{n}\, q^j (1-q)^{n-j} = \frac{1}{j!} n^j \prod_{i=0}^{j-1} \left(1 - \frac{i}{n}\right) \frac{\lambda^j}{n^j} \left(1 - \frac{\lambda}{n}\right)^{n-j}.$$

✦ 2, стр. 22. Поскольку значение интеграла не зависит от аппроксимирующей снизу последовательности простых функций, то это значение не должно зависеть от наличия (или отсутствия) слагаемого $n\mu\{\xi > n\}$.

✦ 3, стр. 22. $\int_\Omega |\xi|\, d\mu = \int_\Omega \xi^+\, d\mu + \int_\Omega \xi^-\, d\mu$.

✦ 4, стр. 24. Ф.р. $F_\eta(x) = \mathbf{P}\{1 - \omega \leqslant x\} = \mathbf{P}\{\omega \geqslant 1 - x\} = 1-(1-x) = x$; аналогично для ф.р. F_ξ сл.в. ξ.

✦ 5, стр. 26. По условию задачи для функции $\overline{F}(x) = \mathbf{P}\{\xi > x\}$ справедливо равенство $\overline{F}(x+t) = \overline{F}(x)\overline{F}(t)$ для $\forall x, t \geqslant 0$. Для любого натурального m отсюда получаем $\overline{F}(mx) = \big(\overline{F}(x)\big)^m$, $x \geqslant 0$. В частности, $\overline{F}\big({}^1\!/_m\big) = \big(\overline{F}(1)\big)^{1/m}$ и, следовательно, $\overline{F}\big({}^k\!/_m\big) = \big(\overline{F}(1)\big)^{k/m}$. Так как функция $\overline{F}$ непрерывна справа, то $\overline{F}(x) = \big(\overline{F}(1)\big)^x$, $x > 0$. Так как $\lim_{x\to\infty} \overline{F}(x) = 0$, то $\overline{F}(1) \neq 1$, кроме того, $\overline{F}(1) \neq 0$, поскольку по условию $\lim_{x\to 0} \big(\overline{F}(1)\big)^x = 1$.

✦ 6, стр. 26. Вероятность $\mathbf{P}\{\lceil\xi\rceil > k\} = \mathbf{P}\{\xi > k\} = \big(e^{-1/\lambda}\big)^k$, $k = 1, 2, \ldots$

✦ 7, стр. 27. Условие постоянства интенсивности отказов эквивалентно дифференциальному уравнению $\overline{F}(t)' = -\gamma \overline{F}(t)$ для функции надёжности $\overline{F}(t) = \mathbf{P}\{\xi > t\}$.

I Многомерные случайные величины

Случайный вектор обычно определяется как измеримое отображение элементарных исходов вероятностного пространства $(\Omega, \mathscr{F}, \mathbf{P})$ в борелевское пространство $(\mathbb{R}^k, \mathscr{B}^k)$. Поскольку в зависимости от тех или иных потребностей (одну и ту же) борелевскую σ-алгебру можно описать посредством различных классов подмножеств $\mathbb{R}^k$, скажем, как минимальную σ-алгебру, порождённую классом открытых множеств, или классом полуоткрытых параллелепипедов, или классом всех прямых произведений борелевских подмножеств числовой прямой и т.д. (см. стр. 445), определение измеримости функций со значениями в $\mathbb{R}^k$ также может быть, в зависимости от ситуации, связано с различными классами подмножеств. Дальнейшее изложение предполагает владение читателем почти всем материалом главы IX.

Определение. *Случайным вектором* (сл.вектором) $\vec{\xi}$ называется *измеримое отображение* $\vec{\xi} : (\Omega, \mathscr{F}, \mathbf{P}) \mapsto (\mathbb{R}^k, \mathscr{B}^k)$,[‡] т.е. такое отображение, что

a) $\vec{\xi}^{-1}(B) \overset{\text{df}}{=} \{\omega : \vec{\xi}(\omega) \in B\} \in \mathscr{F}, \quad \forall B \in \mathscr{B}^k,$ или

b) $\{\omega : \vec{\xi}(\omega) \leqslant \vec{x}\} \in \mathscr{F}, \quad \forall \vec{x} \in \mathbb{R}^k.$

Измеримое отображение $\xi : (\Omega, \mathscr{F}, \mathbf{P}) \mapsto (\mathbb{R}^1, \mathscr{B}^1)$ в борелевскую прямую называется *случайной величиной* (сл.в.).

[‡] Отображение $\vec{\xi}$ определено на Ω и принимает значения в $\mathbb{R}^k$; в представленной записи подчёркивается наличие в пространствах соответствующих σ-алгебр; при отсутствии вероятностной меры $\mathbf{P}$ нет причин называть отображение «случайным», $\vec{\xi}$ просто измеримо.

Из леммы 507, стр. 464, и теоремы 488, стр. 445, следует

8| Лемма. (?!) *Условия определений* a–b) *эквивалентны между собой и, кроме того, эквивалентны каждому из следующих условий:*

c) $\vec{\xi}$ *есть упорядоченный набор сл.в.* $(\xi_1, \ldots, \xi_k)$*;*

d) *прообраз* $\forall$ *открытого множества* $B \subset \mathbb{R}^k$ *измерим:* $\vec{\xi}^{-1}(B) \in \mathscr{F}$*;*

e) $\{\omega : \vec{a} \lessdot \vec{\xi}(\omega) \leqslant \vec{b}\} \in \mathscr{F}, \quad \forall \vec{a} \lessdot \vec{b}\ (\in \mathbb{R}^k)$.(‡)

⊫ c) (Вариант доказательства.) Отображение $\vec{\xi}$ порождает на исходном пространстве Ω σ-алгебру $\sigma(\vec{\xi}) := \vec{\xi}^{-1}(\mathscr{B}^k) := \langle \vec{\xi}^{-1}(B) : \ B \in \mathscr{B}^k \rangle$. Если $\vec{\xi}$ измеримо, то $\sigma(\vec{\xi}) \subset \mathscr{F}$. Ясно, что $\sigma(\vec{\xi})$ есть наименьшая σ-алгебра, относительно которой измеримо $\vec{\xi}$. Кроме того, выбирая множества $B \in \mathscr{B}^k$ вида $B = A \times \mathbb{R}^1 \times \cdots \times \mathbb{R}^1$, видим, что в эту σ-подалгебру входят все σ-подалгебры $\sigma(\xi_j)$, $j = \overline{1,k}$, порождённые компонентами вектора $\vec{\xi}$, т.е. $\sigma(\vec{\xi}) \supset \sigma\big(\bigcup_1^k \sigma(\xi_j)\big)$. Обратно, прообразы любых борелевских «прямоугольников» $\vec{\xi}^{-1}(B_1 \times \cdots \times B_k) \in \sigma\big(\bigcup_1^k \sigma(\xi_j)\big)$. Поскольку «прямоугольники» порождают борелевскую σ-алгебру $\mathscr{B}^k$, то $\sigma(\vec{\xi}) \subset \sigma\big(\bigcup_1^k \sigma(\xi_j)\big)$. Стало быть, σ-подалгебра, порождённая сл.вектором, совпадает с минимальной σ-подалгеброй, относительно которой измеримы все компоненты вектора. ⫤

§ 1. Распределение случайных векторов

Наблюдаемая в эксперименте реализация вектора $\vec{\xi}$ есть результат применения k-мерной функции к элементарному исходу ω, значения которого поступают в соответствии с вероятностным законом $\mathbf{P}$, т.е. вероятность попадания ω в любое подмножество $A \subset \Omega$ равна $\mathbf{P}\{A\}$. Чаще всего пространство исходов Ω и тем более вероятностная мера на нём недоступны исследователю. Однако не всё так плохо. Заметим, что все интересующие нас характеристики $\vec{\xi}$ могут быть вычислены как математическое ожидание некоторой функции от этого вектора:

$$\mathbf{E}\, h(\vec{\xi}) := \int_\Omega h \circ \vec{\xi}(\omega)\, \mathbf{P}(d\omega) = \int_{\mathbb{R}^k} h(\vec{x})\, \mathrm{P}_{\vec{\xi}}(d\vec{x}), \tag{1}$$

где в последнем равенстве мера $\mathrm{P}_{\vec{\xi}}$ на борелевском пространстве $(\mathbb{R}^k, \mathscr{B}^k)$ индуцирована (порождена) отображением $\vec{\xi}$.

(‡) Под неравенством $\vec{x} \leqslant \vec{y}$ для векторов из $\mathbb{R}^k$ мы понимаем их покомпонентное сравнение: $x_j \leqslant y_j$, $\forall j = \overline{1,k}$; неравенство $\vec{x} < \vec{y}$, как правило, будет означать, что $\vec{x} \leqslant \vec{y}$ и $x_j < y_j$ для некоторого $j \leqslant k$. Если $x_j < y_j$ при $\forall j \leqslant k$, то будем писать $\vec{x} \lessdot \vec{y}$.

Определение. Вероятностная мера

$$\mathsf{P}_{\vec{\xi}}\{B\} := \mathbf{P}\{\,\vec{\xi}^{\,-1}(B)\,\} = \mathbf{P}\{\,\omega : \vec{\xi}(\omega) \in B\,\}, \qquad B \in \mathscr{B}^k,$$

называется *распределением сл.вектора* $\vec{\xi}$ в $\mathbb{R}^k$.

Справедливость (1) следует из теоремы 542, стр. 483, о замене переменных в интеграле Лебега.

Так как распределение $\mathsf{P}_{\vec{\xi}}$ позволяет вычислять любые вероятностные характеристики сл.вектора, то, в принципе, сл.вектор можно определить как случайную точку в пространстве $\mathbb{R}^k$, реализации которой поступают в соответствии с распределением $\mathsf{P}_{\vec{\xi}}$. Другими словами, задание сл.вектора эквивалентно заданию некоторого вероятностного закона (распределения) на борелевских подмножествах евклидова пространства $\mathbb{R}^k$. Если на борелевской σ-алгебре пространства $\mathbb{R}^k$ задано некоторое вероятностное распределение P, то с ним можно (искусственно) сопоставить пространство $\Omega = \mathbb{R}^k$ и сл.вектор $\vec{\xi}(\vec{x}) = \vec{x}$; в этом случае компоненты этого сл.вектора $\xi_j(\vec{x}) = x_j$. Такой способ задания сл.вектора называется координатным и часто применяется для доказательства существования семейства сл.в. с тем или иным свойством. Обращение к единому источнику в виде пространства $(\Omega, \mathscr{F}, \mathbf{P})$ упрощает формулировку и интерпретацию свойств сл.векторов. Так, независимость сл.в. удобнее определять как свойство функций на вероятностном пространстве без какой-либо отсылки на распределение сл.векторов, хотя такой способ также вполне допустим (см. ниже).

Мы постоянно будем обращаться к понятию *носителя* сл.вектора, интерпретируя его как множество *значений*, которые сл.вектор *может принять*. Строгое определение носителя как наименьшего замкнутого подмножества $\mathscr{X} \subset \mathbb{R}^k$, для которого $\mathsf{P}_{\vec{\xi}}\{\mathscr{X}\} = \mathbf{P}\{\,\vec{\xi} \in \mathscr{X}\,\} = 1$, часто не совсем удобно с описательной точки зрения. Мы будем по мере надобности изменять вид носителя, сохраняя за ним только свойство единичной вероятности.

✧ **Дискретные распределения.** Проще всего задаются распределения случайных векторов с конечным или счётным числом значений.

Определение. Сл.вектор $\vec{\xi}$ имеет *дискретное* распределение на (не более чем счётном) носителе $\mathscr{X} = \{\vec{x}_j,\ j = 1, 2, \ldots, N\} \subset \mathbb{R}^k,\ N \leqslant \infty$, если

$$p(\vec{x}_j) := \mathbf{P}\{\,\vec{\xi} = \vec{x}_j\,\} > 0 \qquad \text{и} \qquad \mathsf{P}_{\vec{\xi}}\{\mathscr{X}\} = \sum\nolimits_1^N p(\vec{x}_j) = 1.$$

Функция $p(\vec{x}) = \mathbf{P}\{\,\vec{\xi} = \vec{x}\,\}$, $\vec{x} \in \mathscr{X}$, называется *функцией вероятностей* $\vec{\xi}$ (ф.вер. или p.d.f. — от probability distribution function).

Запись дискретного распределения упрощается, если, вопреки строгому определению, в его носитель включить точки с нулевой вероятностью так, чтобы можно было представить $\mathscr{X}$ в виде k-мерной таблицы. При такой записи носитель вектора равен прямому произведению носителей компонент: $\mathscr{X}_{\vec{\xi}} = \mathscr{X}_{\xi_1} \times \ldots \times \mathscr{X}_{\xi_k}$. Например, распределение сл.вектора (ξ_1, ξ_2), сосредоточенного в точках $(-1,0), (0,1), (1,0)$, можно записать в виде таблицы $2 \cdot 3$:

$\xi_2 \backslash \xi_1$	-1	0	1
1	0	$^1/_3$	0
0	$^1/_3$	0	$^1/_3$.

9 Примеры. Наиболее востребованные в приложениях многомерные дискретные распределения — многомерное гипергеометрическое и мультиномиальное (полиномиальное).

1) *Многомерное гипергеометрическое распределение* возникает в ситуациях, когда производится отбор части элементов конечной группы (генеральной совокупности), в которой объекты разделены на $k \geqslant 2$ категорий. Как пример — частичный опрос перед выборами из k кандидатов, если этот опрос производится методом случайного, равновероятного отбора опрашиваемых из обезличенного списка избирателей. Пусть ξ_j — количество объектов из j-й категории, попавших в выборку. Если генеральная совокупность состоит из N_j объектов j-й категории, $N = \sum_{j=1}^{k} N_j$ — общее число объектов совокупности и производится выбор без возвращения n объектов, то ф.вер. сл.вектора $\vec{\xi} = (\xi_1, \ldots, \xi_k)$, т.е. набор вероятностей того, что выборка содержит n_j объектов j-й категории, $j = \overline{1,k}$, вычисляется по формуле[‡]

$$p_{\text{MH}}(n_1, \ldots, n_k) = \Bigl(\prod_{j=1}^{k} \mathrm{C}_{N_j}^{n_j}\Bigr) \frac{1}{\mathrm{C}_N^n}, \qquad n_1 + \ldots + n_k = n. \tag{2}$$

Кратко будем писать $\vec{\xi} \backsim \mathcal{H}\mathrm{g}_k(n; N_1, \ldots, N_k)$.

Компоненты $\vec{\xi}$ линейно связаны: $\xi_1 + \ldots + \xi_k = n$, поэтому, вообще говоря, его распределение нужно описывать в пространстве меньшей размерности ($\mathbb{R}^{k-1}$), как это обычно и делается для $k = 2$, когда в эксперименте фик-

[‡] Как всегда, $\mathrm{C}_A^a = 0$ при $a > A$.

сируется только количество объектов одной категории, при этом ко второй категории (в общем случае к k-й) относится «всё остальное».

2) Ситуацию, подпадающую под описание *мультиномиальной модели* $\mathcal{M}\mathtt{ult}_k(n; q_1, \ldots, q_k)$, хорошо иллюстрируют соревнования по стрельбе из лука. При каждом выстреле стрелок может или промахнуться мимо мишени, или поразить одно из пяти концентрических колец мишени, учитывая центральное «яблочко». Таким образом, при одном выстреле происходит один из $k = 6$ исходов. Пусть q_j — вероятность j-го исхода (промаха или поражения соответствующего кольца) при одном выстреле: $q_1 + \ldots + q_k = 1$. Соревнования проводятся сериями по $n = 12$ выстрелов. Если компонента ξ_j вектора $\vec{\xi} = (\xi_1, \ldots, \xi_k)$ указывает на количество реализаций j-го исхода в серии, то ф.вер. сл.вектора $\vec{\xi}$

$$p_{\mathsf{M}}(n_1, \ldots, n_k) \;=\; \mathbf{P}\big\{\, \xi_1 = n_1, \ldots, \xi_k = n_k \,\big\} \;=\; \frac{n!}{n_1! \cdots n_k!}\; q_1^{n_1} \cdots q_k^{n_k} \tag{3}$$

для всех тех $(n_1, \ldots, n_k)$, для которых $n_1 + \ldots + n_k = n$. Косвенно справедливость этой формулы подтверждает известная формула для полинома:

$$1 \;=\; (q_1 + \ldots + q_k)^n \;=\; \sum_{\substack{(n_1, \ldots, n_k)\,: \\ n_1 + \ldots + n_k = n}} \frac{n!}{n_1! \cdots n_k!}\; q_1^{n_1} \cdots q_k^{n_k}.$$

Отсюда и второе название для модели — *полиномиальная*. Заметим, что при $k = 2$ мультиномиальная модель $\mathcal{M}\mathtt{ult}_2(n; q, 1-q)$ эквивалентна биномиальной модели $\mathcal{B}\mathtt{in}(n; q)$ с единственным отличием: в последней фиксируется количество исходов только одного типа, скажем, успехов ξ_1, а количество альтернативных исходов, если надо, просто вычисляется: $\xi_2 = n - \xi_1$.

Мультиномиальное распределение будет иметь и сл.вектор из предыдущего примера, если отбор производится с возвращением. В этом случае $q_j = {}^{N_j}\!/_{N}$ — относительная доля j-й категории в генеральной совокупности. Если объём выборки n значительно меньше полного объёма генеральной совокупности N, то вероятности любых событий гипергеометрической модели с достаточной точностью могут быть найдены по формулам мультиномиальной модели. ◎

10| Упр. а) Докажите формулы для ф.вер. многомерной гипергеометрической p_{MH} и мультиномиальной p_{M} моделей.

б) Покажите, что для фиксированных $q_j = N_j/N, j = \overline{1, k}$, вероятность $p_{\mathsf{MH}}(n_1, \ldots, n_k) \to p_{\mathsf{M}}(n_1, \ldots, n_k)$ при $N \to \infty$ для $\forall\,(n_1, \ldots, n_k)$.

11| Пример. Предположим, что 1-го кандидата поддерживают 40% всех избирателей, 2-го — 45%, 3-го — 10%. Оставшиеся 5% избирателей входят в категорию «все остальные» (против всех или не примут участие в выборах). Тогда вероятность получения при опросе $n = 1000$ респондентов результата, совпадающего с состоянием всего общества, приблизительно равна $p_{\text{М}}(400, 450, 100, 50) = 6.67404 \cdot 10^{-5}$. Если считать ту же вероятность по точной формуле гипергеометрической модели (скажем, с числом всех выборщиков $N = 100 \cdot 10^6$), то получим $p_{\text{МН}} = 6.67414 \cdot 10^{-5}$.

О точности прогноза результатов выборов, основанного на частичном опросе, можно судить по вероятности попадания вектора относительных частот в некоторые области. Так, в описанной ситуации вероятность попадания вектора частот $\vec{\xi}/_n \cdot 100\%$ в область $[37; 43] \times [42; 48] \times [7; 13] \times [2; 8]$ (в %) равна 0.916. Вероятность того, что первый кандидат при опросе 1000 респондентов получит больше голосов, чем второй, равна всего 0.041. Поэтому если такое произошло, можно усомниться в истинности предположений о распределении голосов избирателей или в беспристрастности (равновероятности) отбора при опросе. ◎

✧ **Функция распределения случайного вектора**. Распределение сл.вектора может быть описано с помощью функции распределения.

Определение. Функция

$$F(\vec{x}) \;:=\; \mathbf{P}\big\{\, \omega : \vec{\xi}(\omega) \leqslant \vec{x} \,\big\} \;=\; \mathbf{P}\big\{\, \xi_1 \leqslant x_1 \,, \,\dots\, , \xi_k \leqslant x_k \,\big\}, \quad \vec{x} \in \mathbb{R}^k,$$

называется *функцией распределения* сл.вектора $\vec{\xi}$ (ф.р. или c.d.f.)(‡).

Заметим, что ф.р. однозначно определяет вероятностное распределение сл.вектора. Точнее, справедлива

12| **Лемма.** *Если ф.р. случайных векторов $\vec{\xi}, \vec{\eta}$ совпадают: $F_{\vec{\xi}}(\vec{x}) = F_{\vec{\eta}}(\vec{x})$, $\forall\, \vec{x} \in \mathbb{R}^k$, то совпадают и их распределения: $\mathbf{P}\big\{\, \vec{\xi} \in B \,\big\} = \mathbf{P}\big\{\, \vec{\eta} \in B \,\big\}$, $\forall\, B \in \mathscr{B}^k$.*

⊨⇒ Воспользуемся приведённой ниже леммой 13. По условию настоящей леммы распределения $\vec{\xi}$ и $\vec{\eta}$ совпадают на подмножествах вида $\mathbb{U}(\vec{x}) = \{\vec{u} \in \mathbb{R}^k : \; \vec{u} \leqslant \vec{x}\}$. Класс таких подмножеств, очевидно, замкнут относительно пересечений и к тому же порождает борелевскую σ-алгебру $\mathscr{B}^k$ (см.

(‡) Как всегда, запятая между событиями под знаком вероятности заменяет знак пересечения (союз И).

488, стр. 445), поэтому распределения $\vec{\xi}, \vec{\eta}$ совпадают на $\mathscr{B}^k$. $\Lleftarrow$

13| **Лемма.** *Если две вероятностные меры μ_1, μ_2 на измеримом пространстве $(\Omega, \mathscr{F})$ совпадают на классе подмножеств $\mathcal{X} \subset \mathscr{F}$, замкнутом относительно пересечений, то эти меры совпадают на σ-алгебре $\sigma(\mathcal{X})$, порождённой $\mathcal{X}$.*

$\Rrightarrow$ Обозначим через $\mathcal{E} = \langle B \in \mathscr{F} : \ \mu_1(B) = \mu_2(B) \rangle$ семейство измеримых подмножеств Ω, на которых меры равны. По условию $\mathcal{E} \supset \mathcal{X}$. Очевидно, $\varnothing$, $\Omega \in \mathcal{E}$. Если $B_1 \subset B_2$ — два подмножества $\mathcal{E}$, то

$$\mu_1(B_2 \setminus B_1) \ = \ \mu_1(B_2) - \mu_1(B_1) \ = \ \mu_2(B_2 \setminus B_1).$$

Кроме того, если монотонная последовательность $B_n \searrow B$ (или $B_n \nearrow B$) и все $B_n \in \mathcal{E}$, то по свойству непрерывности меры

$$\mu_1(B) \ = \ \lim_n \mu_1(B_n) \ = \ \lim_n \mu_2(B_n) \ = \ \mu_2(B),$$

т.е. $\mathcal{E}$ есть λ-система, содержащая Ω. По теореме 598, стр. 526, класс $\mathcal{E}$ содержит σ-алгебру $\sigma(\mathcal{X})$. $\Lleftarrow$

Ф.р. можно определить и с помощью строгих неравенств в виде $F(\vec{x}) = \mathbf{P}\{ \vec{\xi} < \vec{x} \}$. Выбранный здесь способ обусловлен удобством интерпретации векторных неравенств $\vec{\xi} \leqslant \vec{x}$ (см. замечание на стр. 442).

Вообще говоря, ф.р. сл.вектора зависит от порядка расположения компонент вектора. Например, равенство $F_{\vec{\xi}}(x_1, x_2) = F_{\vec{\eta}}(x_1, x_2)$ для ф.р. $\vec{\xi} = (\xi_1, \xi_2)$ и $\vec{\eta} = (\xi_2, \xi_1)$ имеет место только в специальных случаях (*перестановочные* сл.в.). Однако очевидно, что $F_{\vec{\xi}}(x_1, x_2) = F_{\vec{\eta}}(x_2, x_1)$. Другими словами, если $\pi = (j_1, \ldots, j_k)$ — какая-либо перестановка индексов $(1, \ldots, k)$ и $\pi\vec{a}$ — соответствующая перестановка компонент вектора $\vec{a}$, то справедливо равенство

$$F_{\vec{\xi}}(\vec{x}) = F_{\pi\vec{\xi}}(\pi\vec{x}), \quad \forall\, \vec{x} \in \mathbb{R}^k.$$

Для формулировки свойств ф.р. сл.векторов нам понадобится разностный оператор $\Delta_{\vec{a},\vec{c}}$, определяемый формулами

$$\begin{aligned} \Delta^{(j)}_{a,c} F(\vec{x}) \ &= F(x_1, \ldots, x_{j-1}, c, x_{j+1}, \ldots, x_k) - F(x_1, \ldots, x_{j-1}, a, x_{j+1}, \ldots, x_k), \\ \Delta_{\vec{a},\vec{c}} F \ &= \Delta^{(1)}_{a_1,c_1} \ldots \Delta^{(k)}_{a_k,c_k} F(\vec{x}), \end{aligned} \tag{4}$$

который используется для вычисления вероятности попадания сл.вектора $\vec{\xi}$ в

k-мерный параллелепипед $(\vec{a};\vec{c}] = (a_1;c_1]\times\ldots\times(a_k;c_k]$. Кроме того, определим область $\mathbb{U}(\vec{x}) = \{\vec{u}\in\mathbb{R}^k:\ \vec{u}\leqslant\vec{x}\}$; очевидно, ф.р. $F(\vec{x}) = \mathsf{P}_{\vec{\xi}}\{\mathbb{U}(\vec{x})\} = \mathbf{P}\{\,\vec{\xi}\in\mathbb{U}(\vec{x})\,\}$. В одномерном случае $\mathbb{U}(x) = (-\infty;x]$ и

$$\Delta_{a,c}F \;=\; F(c)-F(a) \;=\; \mathbf{P}\{a<\xi\leqslant c\}\,.$$

14| **Теорема.** *Пусть $F(\vec{x})$ — ф.р. сл.вектора $\vec{\xi}$. Тогда:*

(F_1) *ф.р. $F(\vec{x})$ непрерывна справа*

a) *по каждой переменной: для $\forall\, j=\overline{1,k}$*

$$\lim_{\varepsilon\to+0}F(x_1,\ldots,x_j+\varepsilon,\ldots,x_k)=F(\vec{x});$$

b) *по совокупности переменных: $\lim\limits_{\vec{\varepsilon}\to+0}F(\vec{x}+\vec{\varepsilon})=F(\vec{x})$* (‡)*;*

(F_2) *ф.р. $F(\vec{x})$ не убывает: $F(\vec{x})\leqslant F(\vec{y})$, если $\vec{x}\leqslant\vec{y}$;*

(F_3) *вероятность попадания $\vec{\xi}$ в параллелепипед $(\vec{a};\vec{c}]$, $\vec{a}\lessdot\vec{c}$,*

$$\mathbf{P}\{\,\vec{\xi}\in(\vec{a};\vec{c}]\,\} \;=\; \Delta_{\vec{a},\vec{c}}F(\vec{x});$$

(F_4) a) $\lim\limits_{\vec{x}\to\vec{y}}F(\vec{x}) \;=\; 1$, *если все координаты $\vec{y}$ равны $+\infty$, т.е. для $\forall\,\varepsilon>0$ найдётся $M>0$ такое, что $F(\vec{x})>1-\varepsilon$ для $\forall\,\vec{x}\geqslant M$* (‡)*;*

b) $\lim\limits_{\vec{x}\to\vec{y}}F(\vec{x})=0$, *если хотя бы одна координата $y_j=-\infty$; точнее, для $\forall\,\varepsilon>0$ найдётся $M>0$ такое, что $F(\vec{x})<\varepsilon$, если $x_j\leqslant -M$:*

$$F(+\infty,\ldots,+\infty)=1,\qquad F(x_1,\ldots,-\infty,\ldots,x_k)=0.$$

⊫ Ясно, что $\mathbb{U}(\vec{x})\subset\mathbb{U}(\vec{y})$ при $\forall\,\vec{x}\leqslant\vec{y}$. Отсюда, ввиду монотонности вероятности, легко следует свойство (F_2).

В силу (F_2) для доказательства (F_1) достаточно рассмотреть только монотонные последовательности $\vec{\varepsilon}_n\searrow\vec{0}$. В этом случае $\mathbb{U}(\vec{x})=\lim\limits_{n}{}_{\downarrow}\,\mathbb{U}(\vec{x}+\vec{\varepsilon}_n)$, и в силу непрерывности вероятности $F(\vec{x})=\mathbf{P}\{\mathbb{U}(\vec{x})\}=\lim\limits_{n}{}_{\downarrow}\,\mathbf{P}\{\mathbb{U}(\vec{x}+\vec{\varepsilon}_n)\} = \lim\limits_{n}{}_{\downarrow}\,F(\vec{x}+\vec{\varepsilon}_n)$.

Аналитически свойство (F_3) (к примеру, в трехмерном случае) можно доказать так: искомая вероятность

$$\mathbf{P}\{a_1<\xi_1\leqslant c_1,\ a_2<\xi_2\leqslant c_2,\ a_3<\xi_3\leqslant c_3\} \;=\; G_1(c_1)-G_1(a_1) \;=\; \Delta^{(1)}_{a_1,c_1}G_1(x_1),$$

(‡) $\vec{\varepsilon}\to+0 \Leftrightarrow \vec{\varepsilon}\to\vec{0},\ \vec{\varepsilon}\geqslant 0$; сравнение $\vec{x}\geqslant M$ означает, что $x_j\geqslant M$ для $\forall\,j=\overline{1,k}$.

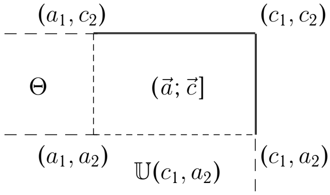
$$\text{где } G_1(x_1) \quad = \mathbf{P}\{\xi_1 \leqslant x_1,\ a_2 < \xi_2 \leqslant c_2,\ a_3 < \xi_3 \leqslant c_3\} = \Delta^{(2)}_{a_2,c_2} G_2(x_1,x_2),$$
$$\text{где } G_2(x_1,x_2) = \mathbf{P}\{\xi_1 \leqslant x_1,\ \xi_2 \leqslant x_2,\ a_3 < \xi_3 \leqslant c_3\} = \Delta^{(3)}_{a_3,c_3} F(x_1,x_2,x_3).$$

В двумерном случае доказательство свойства (F_3) можно представить геометрически. Область $\mathbb{U}(a,c)$ есть «угловая» область на плоскости $\mathbb{R}^2$ с правым верхним углом в точке (a,c). Как видно из рисунка, прямоугольник

$$(\vec{a};\vec{c}] = \big[\mathbb{U}(c_1,c_2) \setminus \mathbb{U}(c_1,a_2)\big] \setminus \Theta,$$

где, в свою очередь, область

$$\Theta = \mathbb{U}(a_1,c_2) \setminus \mathbb{U}(a_1,a_2).$$



Таким образом, ввиду вложенности соответствующих областей

$$\mathbf{P}\big\{\,\vec{\xi} \in (\vec{a};\vec{c}\,]\,\big\} = \underbrace{\big(F(c_1,c_2) - F(c_1,a_2)\big)}_{\Delta^{(2)}_{a_2,c_2}F(c_1,x_2)} - \underbrace{\big(F(a_1,c_2) - F(a_1,a_2)\big)}_{\Delta^{(2)}_{a_2,c_2}F(a_1,x_2)}.$$

Для доказательства (F_4) рассмотрим вектор $\vec{n} = (n,\ldots,n)$, все координаты которого равны n, тогда, как легко понять, $\lim\limits_{n}{}_{\uparrow}\,\mathbb{U}(\vec{n}) = \mathbb{R}^k$, т.е. $\bigcup_{n=1}^{\infty}\mathbb{U}(\vec{n}) = \mathbb{R}^k$. В силу непрерывности вероятности $\mathbf{P}\big\{\,\vec{\xi} \in \mathbb{U}(\vec{n})\,\big\} \nearrow 1$. С другой стороны,

$$F(\vec{x}) \leqslant \mathbf{P}\{\xi_j \leqslant x_j\} \leqslant \mathbf{P}\{\xi_j \in (-\infty;-n]\},$$

если $x_j \leqslant -n$. Область $(-\infty;-n] \searrow \varnothing$ при $n\to\infty$, поэтому ввиду непрерывности вероятности $\mathbf{P}\{\xi_j \in (-\infty;-n]\} \searrow 0$, что и доказывает $(\mathrm{b},\mathrm{F}_4)$. ⫤

15| Упр. Докажите справедливость следующих соотношений для вероятностей попадания сл.в. в интервалы:

$$\mathbf{P}\{a<\xi\leqslant c\} = F(c)-F(a), \qquad \mathbf{P}\{a<\xi<c\} = F(c-0)-F(a),$$
$$\mathbf{P}\{a\leqslant\xi<c\} = F(c-0)-F(a-0), \qquad \mathbf{P}\{a\leqslant\xi\leqslant c\} = F(c)-F(a-0),$$
$$\mathbf{P}\{\xi=c\} = F(c)-F(c-0). \tag{5}$$

Как будут выглядеть эти соотношения, если ф.р. определена посредством равенства $F(x) = \mathbf{P}\big\{\,\xi < x\,\big\}$?

В одномерном случае условий непрерывности справа (F_1), монотонности (F_2) и нормированности (F_4) достаточно, чтобы функция F определяла некоторую вероятностную меру. Для многомерных сл.векторов этих условий может быть мало.

16| **Теорема.** *Если функция* $F:\mathbb{R}^1\mapsto[0;1]$ *непрерывна справа, не убывает,* $F(-\infty)=0$ *и* $F(+\infty)=1$, *то на борелевском пространстве* $(\mathbb{R}^1,\mathscr{B}^1)$ *существует единственная вероятностная мера* P *такая, что* $\forall\, a<c\ \ (\in\mathbb{R}^1)$

$$\mathsf{P}\{(-\infty;c]\}=F(c)\quad \text{и}\quad \mathsf{P}\{(a;c]\}=\Delta_{a,c}F=F(c)-F(a). \tag{6}$$

В многомерном случае существуют функции F, удовлетворяющие свойствам (F_1), (F_2), (F_4), у которых, однако, значение разностного оператора Δ на некоторых параллелепипедах может быть отрицательным.

17| Упр. Функция $G(x_1,x_2)=\dot{\mathrm{I}}\,(x_1+x_2>0)$ удовлетворяет свойствам (F_1), (F_2), (F_4), но не может быть ф.р.

18| **Теорема.** *Если функция* $F:\mathbb{R}^k\mapsto[0;1]$ *обладает свойствами* (F_1), (F_4) *и для* $\forall\,\vec a\lessdot\vec c\ \ (\in\mathbb{R}^k)$ *оператор* $\Delta_{\vec a,\vec c}F\geqslant 0$, *то на борелевском пространстве* $(\mathbb{R}^k,\mathscr{B}^k)$ *существует единственная вероятностная мера* P *такая, что при* $\forall\,\vec a\lessdot\vec c\ \ (\in\mathbb{R}^k)$

$$\mathsf{P}\{\,\mathbb{U}(\vec c)\,\}=F(\,\vec c\,)\quad \text{и}\quad \mathsf{P}\{(\vec a,\vec c]\}=\Delta_{\vec a,\vec c}F. \tag{7}$$

$\Rrightarrow$ 16. В одномерном случае полукольцо образует совокупность $\mathcal{K}$ всех интервалов вида $(a;c]$, где $-\infty<a<c<+\infty$. В главе IX, стр. 440 (см. также многомерный случай ниже), доказано, что функция множеств $\mathsf{P}\,(a;c]=F(c)-F(a)$ обладает свойствами аддитивности и σ-полуаддитивности на этом полукольце, что в силу 464, стр. 426, гарантирует её σ-аддитивность. По теореме Каратеодори существует единственная мера на борелевской σ-алгебре $\mathscr{B}^1$ (порождённой полукольцом $\mathcal{K}$), совпадающая с P на $\mathcal{K}$. По условию $F(a)\to 0$ при $a\to-\infty$, следовательно, P удовлетворяет обоим свойствам (6). Так как $(-\infty;c]\nearrow\mathbb{R}^1$, $c\to+\infty$, то по свойству непрерывности меры $\mathsf{P}\{\mathbb{R}^1\}=F(+\infty)-F(-\infty)=1$, т.е. P — вероятностная мера.

18. В многомерном случае определим функцию множеств P на полукольце $\mathcal{K}^k$ параллелепипедов вида $(\vec a,\vec c]$, $\vec a\lessdot\vec c$, с помощью второго соотношения (7). Доказательство аддитивности такой меры проведём только для случая

$k = 2$ (общий случай аналогичен). Пусть прямоугольник $\Pi = (a_1; b_1] \times (a_2; b_2]$ разбит на два прямоугольника $\Pi_1 = (a_1; c] \times (a_2; b_2]$ и $\Pi_2 = (c; b_1] \times (a_2; b_2]$. Тогда

$$\begin{aligned}
\mu_F(\Pi) &= F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2),\\
\mu_F(\Pi_1) &= F(c, b_2) - F(a_1, b_2) - F(c, a_2) + F(a_1, a_2),\\
\mu_F(\Pi_2) &= F(b_1, b_2) - F(c, b_2) - F(b_1, a_2) + F(c, a_2).
\end{aligned}$$

Очевидно, что $\mu_F(\Pi) = \mu_F(\Pi_1) + \mu_F(\Pi_2)$. Отсюда по индукции следует свойство аддитивности для любых разбиений в виде сетки

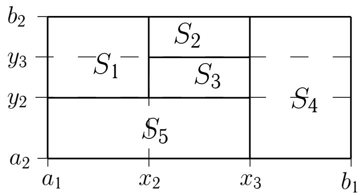
$$(a_1; b_1] \times (a_2; b_2] = \biguplus_{i=1}^{k-1} \biguplus_{j=1}^{m-1} (x_i; x_{i+1}] \times (y_j; y_{j+1}],$$

где $a_1 = x_1 < x_2 < \ldots < x_k = b_1, \quad a_2 = y_1 < y_2 < \ldots < y_m = b_2$.

При произвольном разбиении $(a_1; b_1] \times (a_2; b_2] = \biguplus_1^N S_n$, где $S_n = (a_{1n}; b_{1n}] \times (a_{2n}; b_{2n}]$, зададим сетку с координатами

$$a_1 = x_1 < x_2 < \ldots < x_k = b_1,$$

$$a_2 = y_1 < y_2 < \ldots < y_m = b_2,$$



где каждое x_i совпадает с одним из a_{1n} или b_{1n}, а каждое y_j — с одним из a_{2n} или b_{2n}.

Для доказательства аддитивности теперь достаточно заметить, что разбиение всего прямоугольника $(a_1; b_1] \times (a_2; b_2]$ с помощью выбранной сетки порождает разбиение (также сеточное) каждого из прямоугольников S_n.

Докажем σ-полуаддитивность μ_F. Пусть $(\vec{A}\,; \vec{C}\,] \subset \biguplus_1^\infty (\vec{a}_j; \vec{c}_j]$ и $0 < \varepsilon < \mathsf{P}\{(\vec{A}; \vec{C}\,]\}$. В силу непрерывности F справа по совокупности переменных можно подобрать $\vec{A} \lessdot \vec{A}' \lessdot \vec{C}$ и $\vec{c}_j^{\,\prime} > \vec{c}_j$ так, чтобы

$$\mathsf{P}\{(\vec{A}\,; \vec{C}\,]\} < \mathsf{P}\{(\vec{A}'; \vec{C}\,]\} + \varepsilon \quad \text{и} \quad \mathsf{P}\{(\vec{a}_j; \vec{c}_j^{\,\prime}]\} < \mathsf{P}\{(\vec{a}_j; \vec{c}_j]\} + \frac{1}{2^j}\varepsilon, \ j \geqslant 1.$$

Таким образом, имеем

$$[\vec{A}'; \vec{C}\,] \subset (\vec{A}; \vec{C}\,] \subset \biguplus_1^\infty (\vec{a}_j; \vec{c}_j] \subset \bigcup_1^\infty (\vec{a}_j; \vec{c}_j^{\,\prime}).$$

По лемме Гейне–Бореля из покрытия ограниченного замкнутого множества (компакта) открытыми множествами можно выделить конечное подпокрытие:

$(\vec{A}'; \vec{C}] \subset [\vec{A}'; \vec{C}] \subset \bigcup_1^n (\vec{a}_j; \vec{c}_j') \subset \bigcup_1^n (\vec{a}_j; \vec{c}_j']$.

Для аддитивной функции множеств на полукольце справедливо свойство (конечной) полуаддитивности, следовательно,

$$\mathsf{P}\{(\vec{A}; \vec{C}]\} < \mathsf{P}\{(\vec{A}'; \vec{C}]\} + \varepsilon \leqslant \sum\nolimits_1^n \mathsf{P}\{(\vec{a}_j; \vec{c}_j']\} + \varepsilon \leqslant$$
$$\leqslant \sum\nolimits_1^\infty \mathsf{P}\{(\vec{a}_j; \vec{c}_j]\} + \varepsilon \sum\nolimits_1^\infty \frac{1}{2^j} + \varepsilon = \sum\nolimits_1^\infty \mathsf{P}\{(\vec{a}_j; \vec{c}_j]\} + 2\varepsilon,$$

что и доказывает σ-полуаддитивность P ввиду произвольности ε. По лемме 464, стр. 426, отсюда следует σ-аддитивность P, а по теореме Каратеодори — существование на борелевской σ-алгебре $\mathscr{B}^k$ единственной меры, совпадающей с P на полукольце $\mathcal{K}^k$. ⫷

Теорема Каратеодори, на основе которой делается вывод о существовании меры Лебега–Стилтьеса в $\mathbb{R}^k$, устанавливает факт существования меры, удовлетворяющей только второму равенству (7).

19| **Лемма.** (?!) *Мера* P, *построенная в теореме* 18, *действительно есть вероятность с функцией распределения* F, *т.е.*

a) *она нормированная:* $\mathsf{P}\{\mathbb{R}^k\} = 1$;

b) *для неё справедливы оба соотношения* (7).

В частности, функция F *удовлетворяет свойству* (F_2) *для ф.р.*

20| Примеры. 1) Если $F = F_1(x_1) F_2(x_2)$, где F_1, F_2 — некоторые одномерные ф.р., то $\Delta_{\vec{a},\vec{c}} F = \big(F_1(c_1) - F_1(a_1)\big)\big(F_2(c_2) - F_2(a_2)\big)$. В частности, если все ф.р. $F_j(x) = x,\ 0 \leqslant x \leqslant 1$ [‡], то ф.р. $F_U(x_1, x_2) = x_1 x_2$ определяет сл.вектор $\vec{\xi} \backsim \mathfrak{Un}([0;1]^2)$ с равномерным распределением (мерой Лебега) в единичном квадрате.

Задавать с помощью ф.р. равномерное распределение в области, отличной от прямоугольника со сторонами, параллельными осям координат, неудобно — проще использовать для этого функцию плотности (см. ниже).

2) Функция $F(x, y) = F_1(\min(x, y))$, где F_1 — одномерная ф.р., задаёт распределение, сосредоточенное на биссектрисе $y = x$ (т.е. сл.вектор $(\xi, \eta) \backsim (\xi, \xi) \backsim (\eta, \eta)$). Действительно, для любого прямоугольника $(\vec{a}; \vec{c}]$, у которого все вершины лежат по одну сторону от биссектрисы, значение разностного оператора $\Delta_{\vec{a},\vec{c}} F = 0$. Поскольку любое открытое множество $\mathbb{R}^2$ есть объединение счётного числа подобных прямоугольников, то вероятность попадания ниже или выше биссектрисы равна нулю.

[‡] По свойству ф.р. обязательно $F_j(x) = 0$ при $x \leqslant 0$ и $F_j(x) = 1$ при $x \geqslant 1$.

з) Способ описания многомерных ф.р. через копулы (функции распределения на гиперкубе $[0;1]^k$ специального вида) приведён ниже после рассмотрения понятия частного распределения. ◎

✧ Маргинальные распределения. Каждая компонента сл.вектора представляет собой сл.в. и может быть описана с помощью ф.р. Аналогично, любой подвектор сл.вектора может быть описан с помощью соответствующей ф.р.

Определения. Функция распределения $F_{\vec{\xi}}(x_1,\ldots,x_k)$, $\vec{x}\in\mathbb{R}^k$, называется *совместной* функцией распределения $\vec{\xi}$.

Маргинальной (частной) функцией распределения называется функция распределения $F_{\vec{\eta}}(x_{j_1},\ldots,x_{j_m})$ любого подвектора $\vec{\eta}=(\xi_{j_1},\ldots,\xi_{j_m})$, $1\leqslant j_1<\ldots<j_m\leqslant k$, вектора $\vec{\xi}$ (с соблюдением порядка).

Функции $F_{\xi_j}(x_j)$, $j=\overline{1,k}$, называются одномерными маргинальными ф.р.

Слово marginal можно перевести как «предельный, крайний». В следующей лемме даётся способ нахождения частных ф.р. (для наглядности только одномерной и двумерной) как предела совместной ф.р., когда все переменные, не входящие в описание подвектора, устремляются к $+\infty$, т.е. на края (сравните со свойством (F_4)).

21| Лемма. (?!) *Если* $(\xi_1,\ldots,\xi_k)\backsim F(x_1,\ldots,x_k)$, $\vec{x}\in\mathbb{R}^k$, *то*

$$(\checkmark)\quad F_{\xi_1}(x_1)=F(x_1,+\infty,\ldots,+\infty),$$

$$(\checkmark)\quad F_{(\xi_1,\xi_2)}(x_1,x_2)=F(x_1,x_2,+\infty,\ldots,+\infty).$$

Другое объяснение термина «маргинальный» восходит к естественному способу отыскания одномерных распределений по таблице двумерного дискретного распределения. Внутренняя часть такой таблицы содержит вероятности для каждого сочетания значений сл.в. ξ_1 и ξ_2. Распределение ξ_1, получаемое сложением вероятностей в каждом столбце, записывается в нижний край таблицы напротив ячейки p_{ξ_1}:

$\xi_2\backslash\xi_1$	-1	1	3	p_{ξ_2}	
7	0.1	0.2	0.3	0.6	$=\mathbf{P}\{\,\xi_2=7\,\}$
4	0.2	0	0.2	0.4	$=\mathbf{P}\{\,\xi_2=4\,\}$
p_{ξ_1}	0.3	0.2	0.5	1	

Таким образом, ф.вер. ξ_1 занимает крайнюю — «маргинальную» — строку таблицы вероятностного закона. Аналогично, с занесением в крайний столбец таблицы, находится ф.вер. ξ_2.

Если носитель дискретного распределения представлен в виде прямого произведения носителей компонент: $\mathscr{X} = \mathscr{X}_1 \times \mathscr{X}_2$, где

$$\mathscr{X}_1 = \left\langle\, x_i, \quad i = \overline{1, n_1}\,\right\rangle, \qquad \mathscr{X}_2 = \left\langle\, y_j, \quad j = \overline{1, n_2}\,\right\rangle,$$

и задана ф.вер. $p(x_i, y_j) = \mathbf{P}\{\xi_1 = x_i, \xi_2 = y_j\}$, $i = \overline{1, n_1}$, $j = \overline{1, n_2}$, то ф.вер. для маргинального распределения, скажем ξ_1,

$$p_{\xi_1}(x_i) = \sum\nolimits_{j=1}^{n_2} p(x_i, y_j), \qquad x_i \in \mathscr{X}_1. \tag{8}$$

Здесь числа n_1, n_2 могут принимать «значение» $+\infty$.

22| Примеры. 1) Легко понять из определения, что маргинальное распределение многомерной гипергеометрической модели (2), стр. 32, снова будет иметь тот же тип:

$$\begin{aligned} (\xi_1, \ldots, \xi_k) &\ \backsim\ \mathcal{H}\mathrm{g}_k(n; N_1, \ldots, N_k) \qquad \Rightarrow \\ (\xi_1, \ldots, \xi_{j-1}, \xi_j^*) &\ \backsim\ \mathcal{H}\mathrm{g}_j(n; N_1, \ldots, N_{j-1}, N_j^*), \end{aligned}$$

при $2 \leqslant j < k$, $N_j^* = N_k + \ldots + N_j$. Здесь сл.в. ξ_j^* описывает количество элементов выборки, попавших в одну из категорий $j, \ldots, k$, что соответствует новой категории «всё остальное». Впрочем, справедливость этого можно доказать и формально: при $n_j^* = n - (n_1 + \ldots + n_{j-1})$

$$\mathbf{P}\{\xi_1 = n_1, \ldots, \xi_{j-1} = n_{j-1}, \xi_j^* = n_j^*\} = \sum_{n_j + \ldots + n_k = n_j^*} p_{\text{мн}}(n_1, \ldots, n_j, \ldots, n_k) =$$

$$= \left(\prod_{i=1}^{j-1} \mathrm{C}_{N_i}^{n_i}\right) \frac{\mathrm{C}_{N_j^*}^{n_j^*}}{\mathrm{C}_N^n} \left[\sum_{n_j + \ldots + n_k = n_j^*} \left(\prod_{l=j}^{k} \mathrm{C}_{N_l}^{n_l}\right) \frac{1}{\mathrm{C}_{N_j^*}^{n_j^*}}\right] = \left(\prod_{i=1}^{j-1} \mathrm{C}_{N_i}^{n_i}\right) \mathrm{C}_{N_j^*}^{n_j^*} \frac{1}{\mathrm{C}_N^n},$$

где последнее равенство верно, т.к. сумма всех вероятностей гипергеометрической модели $\mathcal{H}\mathrm{g}_{(k-j+1)}(n_j^*; N_j, \ldots, N_k)$ (в квадратных скобках) равна 1.

В частности, $\xi_2^* = n - \xi_1$ и $(\xi_1, n - \xi_1) \backsim \mathcal{H}\mathrm{g}_2(n; N_1, N - N_1)$, что эквивалентно гипергеометрической модели распределения ξ_1.

2) Аналогичным образом можно убедиться в справедливости утверждения о маргинальном распределении мультиномиального закона. Пусть $q_j^* = 1 - (q_1 + \ldots + q_{j-1})$ — вероятность «неосуществления» ни одного из исходов с номерами $1, \ldots, j-1$, ξ_j^* — количество таких событий в серии.

Тогда

$$(\xi_1, \ldots, \xi_k) \backsim \mathtt{Mult}_k(n; q_1, \ldots, q_k) \quad \Rightarrow$$
$$(\xi_1, \ldots, \xi_{j-1}, \xi_j^*) \backsim \mathtt{Mult}_j(n; q_1, \ldots, q_{j-1}, q_j^*) .$$

В частности, $(\xi_1, n-\xi_1) \backsim \mathtt{Mult}_2(n; q_1, 1-q_2)$, что эквивалентно биномиальной модели: $\xi_1 \backsim \mathtt{Bin}(n; q_1)$ $(\xi_2 = n - \xi_1)$. ◎

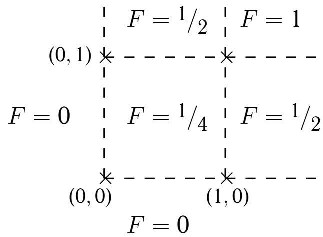
✧ **О точках непрерывности функции распределения**. В силу тождества (5), стр. 37, одномерная ф.р. терпит разрыв в точке x т. т. т. когда эта точка имеет положительную вероятность: $\mathbf{P}\{\xi = x\} = F(x+) - F(x-) > 0$. Так как количество точек x, для которых вероятность $\mathbf{P}\{\xi = x\} \geqslant {}^1/_n$, не более n, то для одномерной ф.р. справедлива

23| **Лемма.** (?!) *Множество точек разрыва любой одномерной функции распределения не более чем счётно.*

Для пространств $\mathbb{R}^k$ с $k > 1$ сказанное выше несправедливо. Следующие примеры, кроме всего прочего, показывают, что у многомерной ф.р. количество точек разрыва может быть и несчётным, а также что здесь существенным является не только поведение F в точках пространства $\mathbb{R}^k$, но и поведение частных ф.р. в соответствующих координатах точек.

24| Примеры. 1) Пусть сл.вектор (ξ, η) сосредоточен с вероятностями ${}^1/_4$ в четырех вершинах единичного квадрата $[0; 1] \times [0; 1]$. Точки разрыва его ф.р. расположены на четырех лучах, параллельных осям координат:

$a)$ $x = 0,\ y \geqslant 0,$ $\qquad b)$ $x = 1,\ y \geqslant 0,$
$c)$ $y = 0,\ x \geqslant 0,$ $\qquad d)$ $y = 1,\ x \geqslant 0.$



Точка $(0, -1)$ — точка непрерывности, хотя маргинальное распределение ξ терпит разрыв при $x = 0$. С другой стороны, скажем, точка $(0, 2)$ есть точка разрыва совместной ф.р., несмотря на то что вероятность попадания в эту точку равна нулю и маргинальная ф.р. η непрерывна при $y = 2$.

2) Пусть сл.вектор (ξ, η) имеет равномерное распределение на сторонах квадрата $[0; 1] \times [0; 1]$, т.е. вероятность попадания (ξ, η) в любую область $B \subset \mathbb{R}^2$ равна одной четвертой длины всех частей сторон квадрата, пересекаю-

щих B. Ясно, что у такого распределения нет ни одной точки сосредоточения массы, однако его ф.р. терпит разрывы на тех же лучах, что и в предыдущем примере, за исключением точки $(0,0)$.

Совместная ф.р. $F(x,y)=\frac{1}{4}(x+y)$, если $0\leqslant x,y<1$, $F(x,y)=0$, если $\min\{x,y\}<0$. Эта функция непрерывна в точке $(0,0)$, хотя обе частные ф.р. $F_\xi(t)=F_\eta(t)=\left(\frac{1}{2}t+\frac{1}{4}\right)\dot{\mathrm{I}}_{[0;1)}(t)+\dot{\mathrm{I}}_{[1;\infty)}(t)$ терпят разрыв в нуле.

3) Пример 62, стр. 66, демонстрирует ещё один аспект отличия одномерного случая от многомерного. Как показано в том примере, непрерывная ф.р. может иметь целую линию сосредоточения — вероятность осуществления, казалось бы, невозможного при непрерывной совместной ф.р. события $\xi=\eta$ равна $^1/_2$. ◎

25| Лемма. (?!) *Пусть $F(x_1,\ldots,x_k)$ — ф.р. сл.вектора $\vec{\xi}$, $F_j(x_j)$, $j=1,\ldots,k$, — одномерные маргинальные ф.р.*

I) *Функция F непрерывна в точке $\vec{x}_0=(x_{01},\ldots,x_{0k})$ т. т. т. когда вероятность попадания $\vec{\xi}$ на границу «угла» $\mathbb{U}(\vec{x}_0)=\left\{\vec{x}:\vec{x}\leqslant\vec{x}_0\right\}$ равна нулю.*

В частности, если $\mathbf{P}\left\{\vec{\xi}=\vec{x}_0\right\}>0$, то в точке $\vec{x}_0$ ф.р. $\vec{\xi}$ терпит разрыв.

II.а) *Ф.р. F непрерывна в точке $(x_{01},\ldots,x_{0k})$, если каждая одномерная маргинальная ф.р. F_j непрерывна в точке x_{0j}, $j=\overline{1,k}$;*

б) *совместная ф.р. F непрерывна всюду т. т. т. когда каждая одномерная маргинальная ф.р. непрерывна всюду.*

Количество точек разрыва многомерной ф.р. может быть и несчётным, однако на любом луче, не параллельном осям координат, таких точек разве лишь счётно. Точнее, справедлива

26| Лемма. *Подмножество $A\subset\mathbb{R}^k$, у которого любая пара векторов $\vec{x},\vec{y}\in A$ имеет различные координаты: $x_j\neq y_j$, $\forall\, j=\overline{1,k}$, содержит не более чем счётное число точек разрыва ф.р. F.*

Следовательно, в любом параллелепипеде $(\vec{a};\vec{c}]\subset\mathbb{R}^k$ таком, что $\vec{a}\lessdot\vec{c}$, найдётся бесконечное число точек непрерывности F.

⊨⇒ Пусть A — несчётное подмножество с указанным в лемме свойством координат. Так как любая одномерная ф.р. имеет не более чем счётное число точек разрыва, то в A найдётся не более чем счётное число векторов, в которых хотя бы по одной из координат соответствующая маргинальная ф.р. терпит разрыв. В остальных точках A все маргинальные ф.р. непрерывны, что по лемме 25 гарантирует непрерывность ф.р. в этих точках.

Во втором утверждении важна только выпуклость параллелепипеда и то, что при условии $\vec{a} \lessdot \vec{c}$ найдётся пара векторов $\vec{x}, \vec{y}$ из параллелепипеда, координаты которых различаются. Стало быть, отрезок $\lambda\vec{x} + (1-\lambda)\vec{y},\ \lambda \in [0;1]$, соединяющий эти два вектора, полностью попадает внутрь параллелепипеда и каждая пара точек этого отрезка имеет различные координаты. В силу предыдущего утверждения на этом отрезке имеется несчётное число точек непрерывности F. ⇚

Из лемм 23, 26 следует важное для дальнейшего утверждение.

27| **Теорема.** *Если ф.р. F_1, F_2 совпадают в точках непрерывности каждой из них, то эти функции совпадают всюду.*

⇛ Пусть $\vec{x}$ — какая-либо точка $\mathbb{R}^k$. Тогда на отрезке между точками $\vec{x}$ и $\vec{x} + z$, $z > 0$,(‡) найдётся последовательность точек $\vec{x}_n \searrow \vec{x}$, в каждой из которых функции F_1, F_2 непрерывны, т.е. $F_1(\vec{x}_n) = F_2(\vec{x}_n)$. Отсюда, ввиду непрерывности справа ф.р., следует, что и $F_1(\vec{x}) = F_2(\vec{x})$. ⇚

► **Обобщённая функция распределения.** Любая непрерывная справа функция $F(\vec{x})$, для которой $\Delta_{\vec{a},\vec{c}}F \geqslant 0$ при любых $\vec{a} < \vec{c}$, определяет меру Лебега–Стилтьеса. Чтобы мера была вероятностной, она должна быть нормированной — в теореме 18 это свойство обеспечивают условия (F_4). Отказ от этих условий может привести не только к тому, что мера Лебега–Стилтьеса перестанет быть вероятностной, но и к такому «неприятному» последствию, как отсутствие монотонности функции $F : F(\vec{a}) \nleqslant F(\vec{c})$ при некоторых $\vec{a} < \vec{c}$. В качестве примера можно привести вероятностную меру Лебега–Стилтьеса, задаваемую функцией $F(x_1, x_2) = G(x_1)G(x_2)$ $(\geqslant 0)$, где $G(x) = x$, если $x \in [-1;0]$, $G(x) = -1$ при $x \leqslant -1$, $G(x) = 0$ при $x \geqslant 0$. Этой мере соответствует ф.р. $(G(x_1)+1)(G(x_2)+1)$, которая не может быть получена с помощью линейного преобразования F — ещё одно отличие от одномерного случая, когда все генерирующие функции одной меры отличаются сдвигом на константу.

Определение. Функция $0 \leqslant F \leqslant 1$ называется *обобщённой функцией распределения* на $\mathbb{R}^k$, если (F_1) она непрерывна справа по совокупности переменных, (F_2) $F(\vec{a}) \leqslant F(\vec{c})$, (F_3) $\Delta_{\vec{a},\vec{c}}\,F \geqslant 0$ при любых $\vec{a} < \vec{c}$ $(\in \mathbb{R}^k)$.

(‡) $\vec{x} + z = (x_1 + z, \ldots, x_k + z)$.

► **Задание распределения через копулу**. По лемме 40, стр. 53, для сл.в. ξ с непрерывной ф.р. F справедливо соотношение $F(\xi) \backsim \mathcal{U}\mathtt{n}(0,1)$. Аналогично, если у сл.вектора $\vec{\xi}$ каждая одномерная маргинальная ф.р. непрерывна, то ф.р. $\mathcal{C}(y_1,\ldots,y_k)$ сл.вектора $(\eta_1,\ldots,\eta_k) = \big(F_{\xi_1}(\xi_1),\ldots,F_{\xi_k}(\xi_k)\big)$ $(\in [0;1]^k)$, имеет равномерные на $[0;1]$ одномерные маргинальные распределения. Совместную ф.р. сл.вектора $\vec{\xi}$ можно записать в виде

$$F(\vec{x}) \;=\; \mathcal{C}(F_{\xi_1}(x_1),\ldots,F_{\xi_k}(x_k)). \tag{9}$$

Определение. *Копулой* называется ф.р. $\mathcal{C} : [0;1]^k \mapsto [0;1]$

$$\mathcal{C}(x_1,\ldots,x_k) \;=\; \begin{cases} 0, & \text{когда хотя бы одна координата } x_i = 0, \\ x_i, & \text{когда } x_j = 1 \text{ при } \forall\, j \neq i. \end{cases}$$

Все одномерные маргинальные распределения копулы $\mathcal{C}$ равномерны $\mathcal{U}\mathtt{n}[0;1]$.

Теорема Sklar'a [33] утверждает, что функция $F(\vec{x})$ есть многомерная ф.р. т. т. т. когда она может быть представлена в виде (9) с некоторой копулой $\mathcal{C}$, причём если все одномерные маргинальные ф.р. непрерывны, то для данной ф.р. F существует единственная копула $\mathcal{C}$. Считается, что копула в концентрированной форме описывает структуру связей между компонентами сл.вектора.

28| Примеры. 1) Ф.р. равномерного распределения на гиперкубе задаёт равномерную копулу с независимыми компонентами.

2) Пусть $\Phi_{\boldsymbol{\Sigma}}$ — ф.р. k-мерного нормального закона с нулевым вектором средних, единичными дисперсиями и матрицей корреляций $\boldsymbol{\Sigma}$, Φ^{-1} — обратная ф.р. нормального $(0,1)$ закона на $\mathbb{R}^1$. Копула

$$\mathcal{C} = \Phi_{\boldsymbol{\Sigma}}(\Phi^{-1}(x_1),\ldots,\Phi^{-1}(x_k))$$

называется нормальной или гауссовской копулой.

3) Архимедова копула задаётся выпуклой, строго убывающей функцией (генератором) $\psi : (0;1] \mapsto \mathbb{R}^1_+$ такой, что $\lim\limits_{x\to+0} \psi(x) = \infty$, $\psi(1) = 0$:

$$\mathcal{C}(y_1,\ldots,y_k) = \psi^{-1}(\psi(y_1) + \ldots + \psi(y_k)).$$

Если генератор $\psi(y) = -\ln(y)$, то копула $\mathcal{C}$ совпадает с равномерной копулой из первого примера. ◎

§ 2. Независимость случайных величин

Совместная (в совокупности) независимость сл.в. определяется посредством условия на вероятность одновременного осуществления событий, относящихся к каждой из рассматриваемых сл.в. Чтобы формализовать такое определение, необходимо предполагать, что все сл.в. заданы на одном вероятностном пространстве. В дальнейшем это условие дополнительно не оговаривается.

Определения. Сл.в. $\xi_1, \ldots, \xi_k$ *независимы в совокупности*, если

$$\mathbf{P}\{\xi_1 \in B_1, \ldots, \xi_k \in B_k\} = \prod\nolimits_1^k \mathbf{P}\{\xi_j \in B_j\}$$

для любых измеримых подмножеств $B_j \subset \mathbb{R}^1$, $j = 1, \ldots, k$.

Сл.векторы $\vec{\xi}_j \in \mathbb{R}^{k_j}$, $j = 1, \ldots, m$, *независимы в совокупности*, если $\mathbf{P}\{\,\vec{\xi}_1 \in B_1, \ldots, \vec{\xi}_m \in B_m\,\} = \prod_1^m \mathbf{P}\{\,\vec{\xi}_j \in B_j\,\}$ для любых измеримых подмножеств $B_j \subset \mathbb{R}^{k_j}$, $j = 1, \ldots, m$.

Отметим, что независимость сл.в. в совокупности не следует из попарной независимости. Можно привести пример (см. 67 , стр. 72) сл.вектора, у которого любой подвектор (не обязательно двумерный) будет состоять из независимых в совокупности сл.в., однако весь вектор не будет удовлетворять условию совокупной независимости.

Подмножества вида $(-\infty; x]$ измеримы по Борелю, поэтому ф.р. сл.вектора с независимыми компонентами

$$F_{\vec{\xi}}(\vec{x}) = \prod\nolimits_1^k \mathbf{P}\{\xi_j \leqslant x_j\} = \prod\nolimits_1^k F_{\xi_j}(x_j). \tag{10}$$

Справедливо и обратное утверждение.

29| **Теорема.** *Компоненты сл.вектора* $\vec{\xi} = (\xi_1, \ldots, \xi_k)$ *независимы в совокупности т. т. т. когда совместная ф.р.* $F_{\vec{\xi}}$ *удовлетворяет* (10), *т.е. равна произведению одномерных частных ф.р.*

$\Rrightarrow$ Рассмотрим только случай $k = 2$ (общий случай — по индукции). При любом фиксированном x_2 (таком, что $F_{\xi_2}(x_2) > 0$) условная вероятность

$$\mathbf{P}\left\{\xi_1 \in B \mid \xi_2 \leqslant x_2\right\} =: \mu_1\{B\} \tag{11}$$

задаёт (?!) вероятностную меру на $B \in \mathscr{B}$. В силу (10) соответствующая этой мере ф.р. совпадает с ф.р. F_{ξ_1} : для $\forall\, x_1 \in \mathbb{R}^1$

$$\mu_1\{(-\infty; x_1]\} = \frac{\mathbf{P}\{\xi_1 \leqslant x_1, \xi_2 \leqslant x_2\}}{\mathbf{P}\{\xi_2 \leqslant x_2\}} = \frac{F_{\vec{\xi}}(x_1, x_2)}{F_{\xi_2}(x_2)} = F_{\xi_1}(x_1).$$

Так как совпадают ф.р., то в силу 12, стр. 34, совпадают и вероятностные меры, которые их определяют: $\mu_1\{B\} = \mathbf{P}\big\{\xi_1 \in B \,\big|\, \xi_2 \leqslant x_2\big\} = \mathbf{P}\{\xi_1 \in B\}$ для $\forall\, B \in \mathscr{B}$, т.е. $\mathbf{P}\{\xi_1 \in B, \xi_2 \leqslant x_2\} = \mathbf{P}\{\xi_1 \in B\}\,\mathbf{P}\{\xi_2 \leqslant x_2\}$. Для точек x_2 с $F_{\xi_2}(x_2) = \mathbf{P}\{\xi_2 \leqslant x_2\} = 0$ это равенство, очевидно, также справедливо.

Повторив эти же рассуждения для меры $\mu_2\{A\} := \mathbf{P}\big\{\xi_2 \in A \,\big|\, \xi_1 \in B_1\big\}$ с фиксированным множеством B_1, получим доказательство теоремы. ⫤

Аналогичный факт справедлив и для независимых сл.векторов.

30| **Теорема.** *Сл.векторы* $\vec{\eta}_j$ $(\in \mathbb{R}^{k_j})$, $j = \overline{1,n}$, *независимы в совокупности т. т. т. когда совместная ф.р. вектора* $\vec{\xi} = (\vec{\eta}_1, \ldots, \vec{\eta}_n)$ *равна произведению ф.р. составляющих его векторов:*

$$F_{\vec{\xi}}(x_{11}, \ldots, x_{1k_1}, \ldots, x_{n1}, \ldots, x_{nk_n}) = \prod\nolimits_1^n F_{\vec{\eta}_j}(x_{j1}, \ldots, x_{jk_j}).$$

Утверждения последних двух теорем часто кладут в основу определения независимости сл.векторов (величин). К сожалению, при этом всё равно приходится доказывать эквивалентность этих определений. Например, чтобы доказать следующее утверждение.

31| **Лемма.** (?!) *Пусть векторы* $\vec{\eta}_j$ $(\in \mathbb{R}^{k_j})$, $j = \overline{1,m}$, *независимы в совокупности, тогда для любых борелевских функций* $g_j : \mathbb{R}^{k_j} \to \mathbb{R}^{n_j}$, $j = \overline{1,m}$, *сл.векторы* $g_1(\vec{\eta}_1), \ldots, g_m(\vec{\eta}_m)$ *также независимы в совокупности.*

32| Пример. При сравнении истинных дисперсий по критерию Фишера, основанному на двух независимых выборках $(\xi_1, \ldots, \xi_n)$ и $(\eta_1, \ldots, \eta_m)$ из нормального распределения, существенно используется тот факт, что выборочные дисперсии $\frac{1}{n}\sum_{i=1}^n \big(\xi_i - \frac{1}{n}\sum_{j=1}^n \xi_j\big)^2$ и $\frac{1}{m}\sum_{i=1}^m \big(\eta_i - \frac{1}{m}\sum_{j=1}^m \eta_j\big)^2$ независимы. ◎

33| Упр. Будут ли независимы сл.в. ξ_1 и сл.вектор (ξ_2, ξ_3), если независимы попарно ξ_1 и ξ_2, а также ξ_1 и ξ_3?

✧ **Независимость классов событий**. Предыдущие рассуждения о независимости относились, вообще говоря, к распределению сл.векторов. Однако на эту проблему можно посмотреть в более общем контексте, отнеся соответствующие условия на события исходного вероятностного пространства $(\Omega, \mathscr{F}, \mathbf{P})$. Удобства такого подхода особенно ярко проявляются при анализе независимости сл.процессов.

Определение. Пусть $(\Omega, \mathscr{F}, \mathbf{P})$ — вероятностное пространство. Классы $\mathscr{L}_1, \ldots, \mathscr{L}_n$ множеств σ-алгебры $\mathscr{F}$ называются независимыми, если

$$\mathbf{P}\Big\{ \bigcap\nolimits_1^n A_j \Big\} = \prod\nolimits_1^n \mathbf{P}\{A_j\}$$

для любых наборов событий $A_j \in \mathscr{L}_j$, $j = \overline{1, n}$, где часть событий A_j может совпадать с Ω.

Напомним, что класс $\xi^{-1}(\mathscr{B})$ множеств вида $\xi^{-1}(B)$, когда множества B выбираются из борелевской σ-алгебры $\mathscr{B}$, есть σ-алгебра $\sigma\big(\xi\big)$, порождённая сл.в. ξ. Поэтому условие независимости сл.в. может быть сформулировано как условие независимости порождённых ими σ-алгебр.

34| **Лемма.** (?!) *Сл.векторы $\vec{\xi}_1, \ldots, \vec{\xi}_n$ независимы в совокупности т. т. т. когда независимы порождаемые ими σ-алгебры.*

Следующая теорема позволяет осуществлять проверку независимости σ-алгебр по более узким классам множеств. Такие классы должны порождать соответствующие σ-алгебры и, кроме того, определять меру. Говорят, что класс $\mathcal{F}$ *определяет меру*, если любые две меры, совпадающие на $\mathcal{F}$, совпадают и на σ-алгебре $\sigma\big(\mathcal{F}\big)$, порождённой $\mathcal{F}$. По лемме 13, стр. 35, класс $\mathcal{F}$ будет определять меру, если он замкнут относительно операции пересечения.

35| **Теорема.** *Для независимости σ-алгебр $\mathscr{L}_1 = \sigma\big(\mathcal{F}_1\big), \ldots, \mathscr{L}_n = \sigma\big(\mathcal{F}_n\big)$, порождённых классами $\mathcal{F}_1, \ldots, \mathcal{F}_n$ подмножеств σ-алгебры $\mathscr{F}$, достаточно, чтобы классы $\mathcal{F}_1, \ldots, \mathcal{F}_n$ были независимы и определяли меру, например были замкнуты относительно операции пересечения.*

⊫ Пусть $A_j \in \mathcal{F}_j, j = \overline{2, n}$, — произвольные подмножества. Зададим две меры на множествах $B \in \mathscr{F}$:

$$\mu_1(B) = \mathbf{P}\Big\{ B \bigcap\nolimits_2^n A_j \Big\}, \qquad \mu_2(B) = \mathbf{P}\{B\} \prod\nolimits_2^n \mathbf{P}\{A_j\}.$$

По условию теоремы меры совпадают на классе $\mathcal{F}_1$, следовательно, и на σ-алгебре $\mathscr{L}_1$, порождённой $\mathcal{F}_1$. Другими словами, для любых наборов множеств $B \in \mathscr{L}_1, A_j \in \mathcal{F}_j, j = \overline{2, n}$, имеет место совпадение правых частей μ_1, μ_2, т.е. классы $\mathscr{L}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$ независимы. Аналогичными рассуждениями по индукции показывается независимость σ-алгебр $\mathscr{L}_1, \ldots, \mathscr{L}_n$. ⫤

Утверждения теорем 29 и 30 являются, по существу, следствиями теоремы 35. Так, например, если рассмотреть полукольцо $\mathcal{K}$ всех интервалов вида $\langle (a;b],\ -\infty < a < b < \infty \rangle$, то по условию теоремы 29 независимы (?!) полукольца $\xi_1^{-1}(\mathcal{K}),\dots,\xi_k^{-1}(\mathcal{K})$, индуцированные отображениями $\xi_1,\dots,\xi_k$. Согласно лемме 13, стр. 35, эти классы множеств определяют меру. В силу леммы 507, стр. 464, прообразы борелевской σ-алгебры $\xi_j^{-1}(\mathscr{B}) = \sigma\big(\xi_j^{-1}(\mathcal{K})\big)$, $j=1,\dots k$. По теореме 35 σ-алгебры $\sigma\big(\xi_j^{-1}(\mathcal{K})\big)$, $j=1,\dots k$, независимы. Независимость сл.в. $\xi_1,\dots,\xi_k$ следует теперь из леммы 34.

Следующий факт также может быть получен из теоремы 35. Для краткости лемма сформулирована для четырёх сл.в., хотя вполне очевидно, что она справедлива для любого числа сл.в. и даже для сл.векторов. Совсем просто эта лемма доказывается с помощью критерия независимости для ф.р.

36| **Лемма.** (?!) *Пусть сл.в. ξ_1,ξ_2,ξ_3,ξ_4 независимы в совокупности, тогда независимы сл.векторы (ξ_1,ξ_2) и (ξ_3,ξ_4) и для любых борелевских функций $\vec{h}_j:\mathbb{R}^2\mapsto\mathbb{R}^{k_j}$, $j=1,2$, независимы сл.векторы $\vec{h}_1(\xi_1,\xi_2)$ и $\vec{h}_2(\xi_3,\xi_4)$.*

✧ **Независимость бесконечных систем сл.в.** Пусть $\Xi = \langle\, \xi_t \,\rangle_{t\in\mathbb{T}}$ — некоторое семейство сл.в. (векторов) на вероятностном пространстве $(\Omega,\mathscr{F},\mathbf{P})$. Если $\mathbb{T}$ — бесконечный набор параметров, то распределение всего семейства Ξ в пространстве функций $\mathbb{R}^{\mathbb{T}}$ может быть описано только с помощью конечномерных распределений, т.е. распределений всевозможных сл.векторов $(\xi_{t_1},\dots,\xi_{t_k})$, где $(t_1,\dots,t_k)\subset\mathbb{T}$, $k\geqslant 1$, (см. гл. VIII).

Определение. Случайные величины $\langle\, \xi_t \,\rangle_{t\in\mathbb{T}}$ независимы в совокупности т. т. т. когда для $\forall k\geqslant 2$ и любых наборов параметров $(t_1,\dots,t_k)\subset\mathbb{T}$ независимы в совокупности сл.в. $\xi_{t_1},\dots,\xi_{t_k}$.

Лемму 36 можно обобщить. Рассмотрим для подмножества $\mathbb{S}\subset\mathbb{T}$ σ-алгебру $\mathscr{F}_{\mathbb{S}} = \sigma\big(\bigcup\limits_{s\in\mathbb{S}}\xi_s^{-1}(\mathscr{B})\big)$ на Ω, порождённую семейством $\langle\, \xi_s \,\rangle_{s\in\mathbb{S}}$.

37| **Лемма.** *Пусть $\mathcal{F}_{\mathbb{S}}$ — класс подмножеств вида $\bigcap_1^m \xi_{s_j}^{-1}(A_j)$, $(s_1,\dots,s_m)\subset\mathbb{S}$, $A_j\in\mathscr{B}$, $j=\overline{1,m}$, $m\geqslant 1$. Тогда $\mathscr{F}_{\mathbb{S}} = \sigma\big(\mathcal{F}_{\mathbb{S}}\big)$.*

⊨ Если $B\in\bigcup\limits_{s\in\mathbb{S}}\xi_s^{-1}(\mathscr{B})$, то $B=\xi_s^{-1}(A)$ $(\in\mathcal{F}_{\mathbb{S}})$ при некоторых $s\in\mathbb{S}$ и $A\in\mathscr{B}$. Следовательно, $\bigcup\limits_{s\in\mathbb{S}}\xi_s^{-1}(\mathscr{B})\subset\mathcal{F}_{\mathbb{S}}$ и $\mathscr{F}_{\mathbb{S}}\subset\sigma\big(\mathcal{F}_{\mathbb{S}}\big)$. Обратно, т.к. $\mathscr{F}_{\mathbb{S}}$ замкнута относительно конечных пересечений, то $\mathscr{F}_{\mathbb{S}}\supset\mathcal{F}_{\mathbb{S}}$. ⫤

38| **Теорема.** *Пусть $\langle\, \xi_t \,\rangle_{t\in\mathbb{T}}$ — семейство независимых в совокупности сл.в., $\mathbb{S},\mathbb{U}$ — непересекающиеся подмножества $\mathbb{T}$. Тогда σ-алгебры $\mathscr{F}_{\mathbb{S}},\mathscr{F}_{\mathbb{U}}$ независимы.*

⊨ Пусть $Z = \bigcap_{j=1}^{k} \xi_{s_j}^{-1}(B_j)$, $(s_1, \ldots, s_k) \subset \mathbb{S}$, — подмножество $\mathscr{F}_{\mathbb{S}}$. Обозначим через $G_{\mathbb{U}}$ семейство подмножеств σ-алгебры $\mathscr{F}_{\mathbb{U}}$, которые не зависят от Z. Это семейство по определению содержит класс $\mathcal{F}_{\mathbb{U}}$ всех подмножеств вида $\bigcap_{j=1}^{m} \xi_{u_j}^{-1}(A_j)$ (см. лемму 37). Легко понять, что класс $\mathcal{F}_{\mathbb{U}}$ замкнут относительно пересечений, а семейство $G_{\mathbb{U}}$, в свою очередь, замкнуто относительно монотонных разностей и монотонных пределов, т.е. $G_{\mathbb{U}}$ образует λ-систему. Так как $\Omega \in G_{\mathbb{U}}$, то в силу 598, стр. 526, $G_{\mathbb{U}} \supset \sigma\big(\mathcal{F}_{\mathbb{U}}\big) = \mathscr{F}_{\mathbb{U}}$, т.е. любое $\mathscr{F}_{\mathbb{U}}$-измеримое подмножество не зависит от Z. ⫤

Пусть теперь $\mathbb{T} = \big\langle 1, 2, \ldots \big\rangle$ — счётный набор индексов. Рассмотрим σ-алгебру $\mathscr{F}_k^m = \sigma\big(\xi_k, \ldots, \xi_m\big) = \sigma\big(\bigcup_{j=k}^{m} \xi_j^{-1}(\mathscr{B})\big)$, порождённую сл.в. $\xi_k, \ldots, \xi_m$, $1 \leqslant k \leqslant m \leqslant \infty$. Так как $\mathscr{F}_k^m \subset \mathscr{F}_k^{m+1} \subset \mathscr{F}_k^{\infty}$, то класс подмножеств $\bigcup_{m \geqslant k} \mathscr{F}_k^m$ образует алгебру, причём $\bigcup_{m \geqslant k} \mathscr{F}_k^m \supset \bigcup_{j \geqslant k} \xi_j^{-1}(\mathscr{B})$. Поэтому минимальная σ-алгебра $\sigma\big(\bigcup_{m \geqslant k} \mathscr{F}_k^m\big) = \mathscr{F}_k^{\infty}$.

События σ-алгебры $\bigcap_k \mathscr{F}_k^{\infty}$ называют остаточными (или «хвостовыми» — от *англ.* tail events). Хвостовым событием является, например, множество тех $\omega \in \Omega$, для которых ряд $\sum_1^{\infty} \xi_j(\omega)$ сходится.

39| **Теорема.** [*Закон нуля или единицы. А.Н. Колмогоров.*] *Если событие A — остаточное относительно семейства независимых в совокупности сл.в., то вероятность $\mathbf{P}\{A\} = 0$ или $\mathbf{P}\{A\} = 1$.*

⊨ Так как остаточное событие $A \in \mathscr{F}_k^{\infty}$ при $\forall\, k \geqslant 1$, то по теореме 38 это событие не зависит от любой σ-алгебры $\mathscr{F}_1^m$, $m \geqslant 1$, следовательно, не зависит от алгебры $\bigcup_{m \geqslant 1} \mathscr{F}_1^m$. В силу 35 событие A не зависит от σ-алгебры $\sigma\big(\bigcup_{m \geqslant 1} \mathscr{F}_1^m\big) = \mathscr{F}_1^{\infty}$, в частности не зависит от самого себя:

$$\mathbf{P}\{A\} = \mathbf{P}\{A \cap A\} = \mathbf{P}\{A\}\,\mathbf{P}\{A\},$$

что, очевидно, доказывает теорему. ⫤

§ 3. Математическое ожидание

Математическое ожидание любой измеримой функции $h(\vec{\xi})$ от сл.вектора $\vec{\xi}$ на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$ определяется как интеграл Лебега $\mathbf{E}\, h(\vec{\xi}) = \int_{\Omega} h(\vec{\xi}(\omega))\, \mathbf{P}(d\omega)$ (см. § 4, стр. 471). По формуле замены переменных это математическое ожидание может быть вычислено как интеграл $\mathbf{E}\, h(\vec{\xi}) = \int_{\mathbb{R}^k} h(\vec{x})\, \mathrm{P}_{\vec{\xi}}(d\vec{x})$ относительно распределения $\vec{\xi}$.

По конструкции интеграла Лебега математическое ожидание любой функции от дискретного сл.вектора $\vec{\xi}$, принимающего значения $\vec{x}_j$ с вероятностями $p(\vec{x}_j)$, $j = 1, 2, \ldots, N$, $N \leqslant \infty$, находится по формуле

$$\mathbf{E}\, h(\vec{\xi}) \ = \ \sum\nolimits_1^N h(\vec{x}_j) p(\vec{x}_j).$$

Если ряд сходится абсолютно, то говорят, что математическое ожидание конечно; если $N = \infty$ и ряд не сходится (к конечному или бесконечному значению) или сходится условно, то математическое ожидание не существует. В случаях, когда для положительной части $h^+(\vec{x}) := \max\{h(\vec{x}), 0\}$

$$\mathbf{E}\left[h^+(\vec{\xi})\right] \ = \sum_{j:\, h(\vec{x}_j)>0} h(\vec{x}_j) p(\vec{x}_j) \ = \ +\infty,$$

а $\mathbf{E}\left[h^-(\vec{\xi})\right] < +\infty$ для отрицательной части $h^-(\vec{x}) := \max\{-h(\vec{x}), 0\}$ (или наоборот), математическое ожидание $\mathbf{E}\, h(\vec{\xi})$ существует и равно $+\infty$ $(-\infty)$.

Поскольку ф.р. полностью определяет распределение $\vec{\xi}$, то в целях повышения информативности в записи математического ожидания борелевской функции от $\vec{\xi}$ используют ф.р. этого сл.вектора:

$$\mathbf{E}\, h(\vec{\xi}\,) \ = \ \int_{\mathbb{R}^k} h(\vec{x})\, \mathrm{P}_{\vec{\xi}}(d\vec{x}) \ =: \ \int_{\mathbb{R}^k} h(\vec{x})\, dF_{\vec{\xi}}(\vec{x}). \tag{12}$$

При этом, кроме указания на ф.р., соответствующую сл.вектору $\vec{\xi}$, достигается ещё и в некотором смысле вычислительная цель, поскольку подчёркивается связь интеграла Лебега–Стилтьеса с интегралом Римана–Стилтьеса (см. 566, стр. 497), средства вычисления которого хорошо разработаны в анализе. Для этого, правда, каждый раз приходится делать оговорку об абсолютной интегрируемости $h(\vec{x})$, если математическое ожидание конечно:

$$\exists \left| \mathbf{E}\, h(\vec{\xi}\,) \right| < \infty \quad \Leftrightarrow \quad \int_{\mathbb{R}^k} |h(\vec{x})|\, dF_{\vec{\xi}}(\vec{x}) < \infty.$$

⚠ Эквивалентность понятий интегрируемости и абсолютной интегрируемости для интеграла Лебега позволяет формулировать утверждения интегрального типа только для неотрицательных сл.в.

Интеграл относительно меры Лебега на $\mathbb{R}^k$ будем обозначать $\int_{\mathbb{R}^k} h(\vec{x})\, d\vec{x}$. Как следует из теоремы Фубини (стр. 502), если этот интеграл конечен (или функция $h \geqslant 0$), то он равен повторному интегралу:

$$\int_{\mathbb{R}^k} h(\vec{x})\, d\vec{x} \ = \ \int_{\mathbb{R}^1} \Big(\int_{\mathbb{R}^1} \cdots \Big(\int_{\mathbb{R}^1} h(x_1, \ldots, x_k)\, dx_k \Big) \cdots dx_2 \Big)\, dx_1.$$

Если распределение сл.вектора сразу задаётся с помощью ф.р. $F_{\vec\xi}$, вероятности любых событий $B \subset \mathbb{R}^k$ можно записать в виде

$$\mathbf{P}\{\,\vec\xi \in B\,\} \;=\; \int_{\mathbb{R}^k} \dot{\mathrm{I}}_B(\vec x)\, dF_{\vec\xi}(\vec x) \;=:\; \int_B dF_{\vec\xi}(\vec x).$$

Обратно, любую ф.р. можно представить в виде интеграла по мере $\mathbf{P}$ от индикаторной функции множества $\{\omega : \vec\xi(\omega) \leqslant \vec x\}$:

$$F_{\vec\xi}(\vec x) \;\overset{\mathrm{df}}{=}\; \mathbf{P}\{\,\vec\xi \leqslant \vec x\,\} \;=\; \int_\Omega \dot{\mathrm{I}}\,(\omega; \vec\xi \leqslant \vec x)\; \mathbf{P}(d\omega).$$

✧ **Представление через функцию распределения**. Приведём некоторые способы вычисления математического ожидания, использующие только ф.р., без обращения к функции плотности (см. далее).

40| **Лемма.** *Пусть сл.в. ξ имеет непрерывную ф.р. F, тогда:*

а) *сл.в. $\eta := F(\xi) \backsim \mathfrak{Un}(0,1)$, т.е. распределена равномерно на $[0;1]$;*

б) *для любой неотрицательной борелевской функции $h : [0;1] \mapsto \mathbb{R}^1_+$*

$$\mathbf{E}\, h\big(F(\xi)\big) \;=\; \int_{\mathbb{R}} h\big(F(x)\big)\, dF(x) \;=\; \int_{[0;1]} h(t)\, dt.$$

⫤ а) В силу непрерывности ф.р. для $\forall\, t \in (0;1)$ найдётся хотя бы одна точка x_t такая, что $F(x_t) = t$. Если таких точек много, то выберем в качестве x_t наименьшую из них. Такая точка существует в силу непрерывности функции F. Очевидно, неравенство $F(x) < t$ эквивалентно неравенству $x < x_t$, поэтому $\mathbf{P}\{\eta < t\} = \mathbf{P}\{\xi < x_t\} = F(x_t) = t$ (опять в силу непрерывности F).

Утверждение б) следует из а), поскольку левая часть здесь равна математическому ожиданию $\mathbf{E}\, h\big(F(\xi)\big) = \mathbf{E}\, h(\eta) = \int_{[0;1]} h(t)\, dt$. ⫣

Для любой ф.р. можно определить псевдообратную функцию следующим образом:

$$\widetilde F(t) \;:=\; \inf\{x :\; F(x) \geqslant t\}$$

для $\forall\, t \in [0;1]$. При $t = 0$ вместо $\widetilde F(0) = -\infty$ удобнее положить $\widetilde F(0) = x_0$, если $F(x) > F(x_0) = 0$ при $\forall\, x > x_0$. Если $F(x) < 1$ при $\forall\, x \in \mathbb{R}^1$, то по определению минимума пустого множества $\widetilde F(1) = +\infty$.

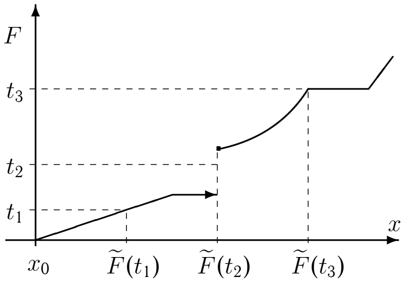


Рис. 1. К определению псевдообратной функции

Так как $F(x)$ непрерывна справа,

то, очевидно, $F(\widetilde{F}(t)) \geqslant t$. Если же F непрерывна в $x = \widetilde{F}(t)$, то $F(\widetilde{F}(t)) = t$.

Обратно, $\widetilde{F}(F(x)) \leqslant x$ при $\forall x \in \mathbb{R}^1$, т.к. $F(x) \geqslant t$ при $t = F(x)$, и $\widetilde{F}(t)$ — минимальная точка с таким свойством. Строгое неравенство $\widetilde{F}(F(x)) < x$ имеет место т. т. т. когда $F(y) = F(x)$ найдётся $y < x$, т.е. либо x принадлежит интервалу постоянства $(a; c)$ ф.р. F, где $F(a - \varepsilon) < F(a) = F(c - 0) < < F(c + \varepsilon)$, $\forall \varepsilon > 0$, либо есть правая граница такого интервала.

Для всюду непрерывной и строго монотонной ф.р. $\widetilde{F}(t) = F^{-1}(t)$, $\forall t \in \in (0; 1)$. Поэтому далее мы будем обозначать псевдообратную ф.р. как F^{-1}.

41| **Лемма.** *Пусть F — ф.р. сл.в. ξ, псевдообратная функция $F^{-1}(t) = \inf\{x: \ F(x) \geqslant t\}$, $\quad t \in (0; 1]$, тогда:*

а) *если $\eta \backsim \mathcal{U}\mathrm{n}(0, 1)$, то $F^{-1}(\eta) \backsim \xi$;*

б) *для любой неотрицательной борелевской функции $h : \mathbb{R}^1 \mapsto \mathbb{R}^1$*

$$\mathbf{E}\, h(\xi) \ = \ \mathbf{E}\, h\big(F^{-1}(F(\xi))\big) \ = \ \int_{(0;1)} h\big(F^{-1}(t)\big)\, dt.$$

$\models\!\!\Rightarrow$ а) При $\forall t \in (0; 1)$ неравенство $x \geqslant F^{-1}(t)$ эквивалентно неравенству $F(x) \geqslant t$. В самом деле, если $x < F^{-1}(t)$, то по построению $F(x) < t$. Обратно, если $x \geqslant F^{-1}(t)$, то ввиду монотонности F и отмеченного выше свойства, $F(x) \geqslant F\big(F^{-1}(t)\big) \geqslant t$. Таким образом, поскольку $\mathbf{P}\{\eta \leqslant t\} = t, \ \forall t \in [0; 1]$,

$$\mathbf{P}\left\{F^{-1}(\eta) \leqslant x\right\} = \mathbf{P}\{\eta \leqslant F(x)\} = F(x).$$

Второе равенство утверждения б) вытекает из а). Действительно, в правой части имеем математическое ожидание $\mathbf{E}\, h\big(F^{-1}(\eta)\big) = \mathbf{E}\, h(\xi)$.

Для доказательства первого равенства б) напомним, что строгое неравенство $F^{-1}\big(F(x)\big) < x$ возможно только для точек x, принадлежащих некоторому интервалу постоянства ф.р. F. В силу 53 мера (порождённая ф.р. F) таких точек равна нулю, стало быть, $\mathbf{P}\left\{F^{-1}\big(F(\xi)\big) = \xi\right\} = 1$. $\Leftarrow\!\!\dashv$

Математическое ожидание сл.в. можно вычислить с помощью формулы интегрирования по частям (теорема 577, стр. 504). Сначала рассмотрим сл.в. $\xi \geqslant 0$ с ф.р. $F(x) = 0$ для $\forall x < 0$, тогда $\mathbf{E}\xi = \int_{[0;\infty)} x\, dF$.

42| **Лемма.** *Если F — ф.р. неотрицательной с.в. ξ $(\geqslant 0)$, то*

$$\mathbf{E}\xi \ = \ \int_0^\infty (1 - F(x))\, dx \ = \ \int_0^\infty \mathbf{P}\{\xi > x\}\, dx,$$

где все части тождества конечны или бесконечны одновременно и интеграл *может пониматься как несобственный интеграл Римана.*

$\Rrightarrow$ По формуле интегрирования по частям 577, стр. 504, при $\forall\, a > 0$

$$\int_0^a x\, dF(x) \;=\; aF(a) - \int_0^a F(x)\, dx \;=\; \int_0^a \big(1 - F(x)\big)\, dx - a\big(1 - F(a)\big).$$

Чтобы положить здесь $a \to +\infty$ и получить требуемое, заметим, что если $\mathbf{E}\,\xi < \infty$, то по неравенству Маркова $a(1 - F(a)) = a\,\mathbf{P}\{\xi > a\} \leqslant \mathbf{E}\big[\xi\, \dot{\mathrm{I}}(\xi > a)\big] \to 0$ при $a \to \infty$ по теореме Лебега. Если же $\mathbf{E}\,\xi = \infty$, то левая часть $\int_0^a x\, dF(x) \to \infty$, поэтому и в правой части $\int_0^a (1 - F(x))\, dx \to \infty$, т.к. $a(1 - F(a)) \geqslant 0$. $\Lleftarrow$

Для отрицательной сл.в. аналогично имеем $\mathbf{E}\,\xi = -\int_{-\infty}^0 F(x)\, dx$.

43| **Теорема.** *Если F — ф.р. ξ, то*

$$\mathbf{E}\,\xi \;=\; \int_0^\infty \big(1 - F(x)\big)\, dx - \int_{-\infty}^0 F(x)\, dx,$$

$$\mathbf{E}\,|\xi| \;=\; \int_0^\infty \big(1 - F(x) + F(-x)\big)\, dx,$$

где правая и левая части обоих тождеств существуют или не существуют, конечны или бесконечны одновременно.

44| Упр. Пусть ξ — неотрицательная сл.в. Покажите, что

$$\mathbf{E}\,\xi \;=\; \int_0^\infty \mathbf{P}\{\xi > x\}\; dx \;=\; \int_0^\infty \mathbf{P}\{\xi \geqslant x\}\; dx.$$

Функция x^α, $\alpha > 0$, абсолютно непрерывна на любом конечном отрезке $[a; c] \subset (0; \infty)$, поэтому для вычисления $\mathbf{E}\,\xi^\alpha$ можно применить вариант формулы интегрирования по частям из теоремы 580, стр. 506.

45| **Лемма.** *Если ξ — неотрицательная с.в., то при $\alpha > 0$*

$$\mathbf{E}\,\xi^\alpha \;=\; \alpha \int_0^\infty x^{\alpha - 1}(1 - F(x))\, dx. \tag{13}$$

46| Упр. Покажите, что если $\lim\limits_{x\to\infty} x^p\, \mathbf{P}\{|\xi| > x\} = 0$ при некотором $p > 0$, то $\mathbf{E}\,|\xi|^q < \infty$ для $\forall\, q \in [0; p)$.

Как это принято, математическое ожидание $\mathbf{E}\,\xi^k$ при целом неотрицательном k будем называть моментом порядка k. В соответствии с предыдущими утверждениями существование конечных моментов порядка k у сл.в. напрямую зависит от скорости убывания её ф.р. $F(x)$ при $x \to -\infty$ и функции $1 - F(x)$ при $x \to +\infty$. Справедлива

47| **Теорема.** *Пусть ф.р.* F *сл.в.* ξ *такова, что функция* $T(x) = 1 - F(x) + F(-x)$ *положительна при* $\forall\, x > 0$ *и*

$$\delta = \varliminf_{x\to\infty} \frac{-\ln T(x)}{\ln x} > 0.$$

Тогда $\mathbf{E}\,|\xi|^a < \infty$ *при* $0 \leqslant a < \delta$ *и* $\mathbf{E}\,|\xi|^a = \infty$ *при* $a > \delta$.

$\Rrightarrow$ Пусть $0 < a < b < \delta$, тогда найдётся $M > 0$ такое, что $-\ln T(x) \geqslant b \ln x$ при всех $x \geqslant M$. Следовательно, при $\forall\, x \geqslant M$

$$1 - F(x) \leqslant x^{-b}, \quad F(-x) \leqslant x^{-b}.$$

Аналогично (13) получаем

$$\int_M^\infty x^a\, dF \;=\; a \int_M^\infty x^{a-1}(1 - F(x))\, dx + M^a(1 - F(M)) \;<\; \infty,$$

поскольку здесь $x^{a-1}(1-F(x)) \leqslant 1/x^{1+(b-a)}$, а разность $b-a > 0$. Аналогично, $\int_{-\infty}^{-M} |x|^a\, dF < \infty$.

Обратно, если $C = \int_{\mathbb{R}} |x|^a\, dF < \infty$, то для $\forall\, y > 0$ имеем $y^a T(y) \leqslant \int_{|x|>y} |x|^a dF \leqslant C$. Следовательно, $\delta = \varliminf_y \big(-\ln T(y)/\ln y\big) \geqslant a$. $\Lleftarrow$

Существуют примеры распределений как с конечным, так и с бесконечным значением математического ожидания $\mathbf{E}\,|\xi|^\delta$.

✧ Математическое ожидание произведения сл.в. Теорема 29, стр. 47, вместе с теоремой Фубини (стр. 502) позволяют упростить доказательство свойства математического ожидания произведения независимых сл.в.

48| **Теорема.** *Если сл.в.* $\langle\, \xi_j \,\rangle_1^J$ *независимы в совокупности и математические ожидания* $\mathbf{E}\big|\xi_j\big| < \infty,\ j = \overline{1, J}$, *то* $\mathbf{E}\big|\xi_1 \cdots \xi_J\big| < \infty$ *и*

$$\mathbf{E}\big[\xi_1 \cdots \xi_J\big] \;=\; \mathbf{E}\,\xi_1 \cdots \mathbf{E}\,\xi_J.$$

$\Rrightarrow$ Пусть сл.в. $\xi_j \backsim F_j(x),\ x \in \mathbb{R}^1,\ j = 1, 2$. Очевидно, функция $F(x_1, x_2) = F_1(x_1) F_2(x_2)$ задаёт ф.р. сл.вектора (ξ_1, ξ_2) в пространстве $\mathbb{R}^2$, причём соответствующая ей мера Лебега–Стилтьеса есть прямое произведение мер $\mu_1 \times \mu_2$, порождённых F_1 и F_2. По условию теоремы

$$\int_{\mathbb{R}} \Big(\int_{\mathbb{R}} |x_1 x_2|\, dF_2(x_2) \Big)\, dF_1(x_1) = \int_{\mathbb{R}} |x_1|\, dF_1(x_1) \int_{\mathbb{R}} |x_2|\, dF_2(x_2) < \infty.$$

Отсюда, применяя теорему Фубини–Тонелли 574, стр. 502, получаем, что математическое ожидание $\mathbf{E}\left|\xi_1\xi_2\right| < \infty$ и

$$\mathbf{E}\left[\xi_1\xi_2\right] = \int_{\mathbb{R}^2} x_1x_2\,dF(x_1,x_2) = \int_{\mathbb{R}} x_1\,dF_1(x_1)\int_{\mathbb{R}} x_2\,dF_2(x_2) = \mathbf{E}\left[\xi_1\right]\mathbf{E}\left[\xi_2\right].$$

Предположим, что утверждение теоремы верно для $n-1$ сл.в. Так как в силу леммы 36 сл.вектор $(\xi_1,\ldots,\xi_{n-1})$ не зависит от сл.в. ξ_n, то по лемме 31 произведение сл.в. $\xi_1\cdots\xi_{n-1}$ не зависит от ξ_n, причём в силу предположения индукции $\mathbf{E}\left|\xi_1\cdots\xi_{n-1}\right| < \infty$. По доказанному свойству для двух сл.в. имеем

$$\mathbf{E}\left[\left(\xi_1\cdots\xi_{n-1}\right)\cdot\xi_n\right] = \mathbf{E}\left[\left(\xi_1\cdots\xi_{n-1}\right)\right]\mathbf{E}\left[\xi_n\right] = \prod_1^n \mathbf{E}\left[\xi_j\right]$$

и $\mathbf{E}\left|\left(\xi_1\cdots\xi_{n-1}\right)\cdot\xi_n\right| < \infty$. ⋖|

Математическое ожидание $\mathbf{E}\left[\xi_1\xi_2\right]$ можно вычислить с помощью формулы Гёфдинга (24), стр. 74, обобщающей теорему 43.

49| Упр. (а) Покажите, что если ξ_1,ξ_2 — независимые сл.в. с $\mathbf{E}\xi_1 = \mathbf{E}\xi_2 = +\infty$, причём $\mathbf{E}\xi_1^- + \mathbf{E}\xi_2^- \neq 0$, то математическое ожидание $\mathbf{E}\left[\xi_1\xi_2\right]$ не существует.

(б) Приведите пример независимых сл.в., для которых $\mathbf{E}\xi_1 = +\infty$, $\mathbf{E}\xi_2 = 0$ и математическое ожидание $\mathbf{E}\left[\xi_1\xi_2\right]$ не существует.

§ 4. О типах распределений. Разложения Жордана и Лебега

По теореме Лебега о разложении мер 559, стр. 491, любая вероятностная мера $\mathbf{P}$ на пространстве с мерой $(\Omega,\mathscr{F},\lambda)$ может быть единственным способом представлена в виде суммы двух мер $\mathbf{P} = \mu_a + \mu_s$, где μ_s — сингулярная (относительно λ) часть меры, т.е. для некоторого множества N мера $\lambda(N) = 0$, в то время как мера $\mu_s(N^c) = 0$. Мера μ_a абсолютно непрерывна относительно меры λ, т.е. найдётся неотрицательная функция (плотность) f такая, что $\mu_a(B) = \int_B f(\omega)\,\lambda(d\omega)$, $B\in\mathscr{F}$. Произведя соответствующую нормировку (на $\gamma = \mu_a(\Omega)$), разложение Лебега для вероятностной меры можно представить в виде выпуклой комбинации вероятностных же мер:

$$\mathbf{P} = \gamma\,\mathbf{P}_a + (1-\gamma)\,\mathbf{P}_s\,.$$

Для распределений в пространстве $\mathbb{R}^1$ сингулярную часть разложения Лебега можно записать в виде суммы двух мер, одна из которых дискретная, а другая есть мера Лебега–Стилтьеса с непрерывной всюду ф.р.

Определение. Неотрицательная конечная функция $F : \mathbb{R}^1 \mapsto \mathbb{R}^1_+$ называется *дискретной* (или *функцией скачков*) т. т. т. когда найдётся не более чем счётный набор точек $\mathscr{X} = \{x_k \in \mathbb{R}^1,\ k = 1, \ldots, N\}$ и соответствующий набор положительных чисел $\{p_k > 0,\ k = 1, \ldots, N\}$, $N \leqslant \infty$, таких, что

$$F(c) = \sum_{x_k \leqslant c} p_k, \qquad \forall\, c \in \mathbb{R}^1.$$

Заметим, что требование конечности дискретной функции влечёт сходимость ряда, определяющего эту функцию. Поэтому $F(-\infty) = 0$.

50| **Лемма.** (?!) *Функция скачков не убывает, непрерывна во всех точках $x \notin \mathscr{X}$, непрерывна справа во всех $x = x_k \in \mathscr{X}$, причём скачок в точке x_k равен $F(x_k) - F(x_k - 0) = p_k$, $x_k \in \mathscr{X}$.*

51| **Теорема.** [*Разложение Жордана.*] *Пусть F — ф.р. на $\mathbb{R}^1$. Тогда существует единственное представление $F = F_c + F_d$, где F_c — всюду непрерывная функция, F_d — дискретная функция, причём $F_c(-\infty) = F_d(-\infty) = 0$.*

⊫ См. стр. 103. ⫣

⚠ Так же как и в разложении Лебега для мер, разложение Жордана можно представить в виде выпуклой комбинации ф.р.: $F = \alpha F_c + (1-\alpha) F_d$, где F_c, F_d — соответственно непрерывная и дискретная ф.р., $\alpha \in [0; 1]$.

Определения. Точка $x \in \mathbb{R}^1$ называется *точкой роста* неубывающей функции F, если $F(x - \varepsilon) < F(x + \varepsilon)$ для $\forall\, \varepsilon > 0$.

Неубывающая функция на $\mathbb{R}^1$ называется *сингулярной*, если она непрерывна и множество всех её точек роста имеет лебегову меру нуль.

52| Пример. (*Лестница Кантора.*) Рассмотрим представление числа $0 < x < 1$ в троичной системе счисления:

$$x \;=\; \{0, x_1 x_2 \ldots\}_3 \;=\; \frac{x_1}{3} + \frac{x_2}{3^2} + \cdots, \qquad x_k = 0, 1, 2;\ k \geqslant 1.$$

Положим $N_x = \min\{k : x_k = 1\}$ — первая позиция в троичном разложении x, где встречается 1; $N_x = \infty$, если в троичном разложении числа x отсутствуют единицы. Для чисел x с $J = N_x < \infty$ определим функцию

$$\mathcal{K}(x) \;:=\; \{0, \frac{x_1}{2}\frac{x_2}{2} \ldots \frac{x_{J-1}}{2} 1\}_2 \;\stackrel{\text{df}}{=}\; \sum\nolimits_{k=1}^{J-1} \frac{x_k}{2^{k+1}} + \frac{1}{2^J}$$

с очевидными изменениями в случае $N_x = \infty$. Например, в представлении числа $x = {}^3\!/_4 = \{0,2020\ldots\}_3 = \{0,(20)\}_3$ нет единиц, поэтому

$$\mathcal{K}\left(\frac{3}{4}\right) = \frac{1}{2} + \frac{1}{8} + \ldots = \frac{2}{3}.$$

Функция $\mathcal{K}(x)$, $x \in [0;1]$, называется *функцией Кантора* (канторовой лестницей). На рис. 2 приведён вариант графика функции $\mathcal{K}(x)$.

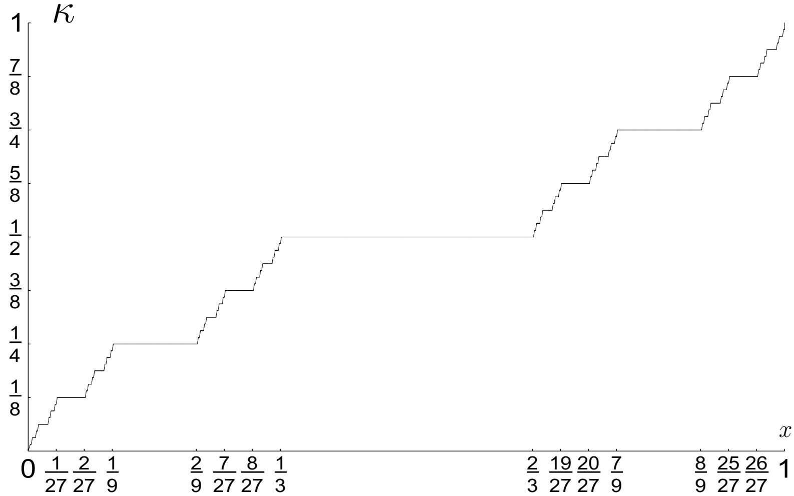


Рис. 2. График функции Кантора

Легко понять, что функция Кантора непрерывна и не убывает. Покажем её сингулярность. Пусть $x \in ({}^1\!/_3; {}^2\!/_3)$, тогда $x = \{0,1x_2x_3\ldots\}_3$ и значение функции $\mathcal{K}(x) = \{0,1\}_2 = {}^1\!/_2$ для всех таких x. Точки $x = {}^1\!/_3$ и $x = {}^2\!/_3$ могут быть представлены двумя способами, например: ${}^1\!/_3 = \{0,1\}_3 = \{0,022\ldots\}_3$. При любом таком представлении функция $\mathcal{K}(x) = \{0,1\}_2 = \{0,011\ldots\}_2 = {}^1\!/_2$. Таким образом, на отрезке $x \in [{}^1\!/_3; {}^2\!/_3]$ функция $\mathcal{K}(x) = {}^1\!/_2$, т.е. интервал $({}^1\!/_3; {}^2\!/_3)$ длины $\frac{1}{3}$ является интервалом постоянства функции.

Аналогично, $\mathcal{K}(x) = {}^1\!/_4$ при $x \in [{}^1\!/_9; {}^2\!/_9]$ и $\mathcal{K}(x) = {}^3\!/_4$ при $x \in [{}^7\!/_9; {}^8\!/_9]$. Суммарная длина последних двух интервалов равна $\frac{2}{9}$.

Продолжая подобным образом, получим семейство всё более измельчающихся интервалов постоянства функции $\mathcal{K}$, суммарная длина которых по формуле для суммы бесконечной геометрической прогрессии равна

$$\frac{1}{3} + 2 \cdot \frac{1}{9} + 4 \cdot \frac{1}{27} + \cdots = \frac{{}^1\!/_3}{1 - {}^2\!/_3} = 1.$$

Другими словами, мера Лебега множества точек роста равна нулю.

Множество точек роста функции Кантора, называемое *канторовым множеством*, состоит из точек отрезка $[0;1]$, для которых существует троичное представление, не содержащее единиц. Вместо доказательства — один пример:

$$\{0,021(2)\}_3 = \frac{1}{3} - \frac{1}{3^3} < \frac{1}{3} = \{0,0(2)\}_3 < \frac{1}{3} + \frac{1}{3^3} = \{0,100(2)\}_3$$

и $\mathcal{K}\Big(\frac{1}{3} - \frac{1}{3^3}\Big) = 0.375 < 0.5 = \mathcal{K}\Big(\frac{1}{3} + \frac{1}{3^3}\Big)$.

Канторово множество представляет собой пример несчётного множества лебеговой меры нуль. Это множество замкнуто и не имеет изолированных точек, т.е. оно совершенное. ◎

53| **Лемма.** *Пусть F — неубывающая функция, Q_F — множество всех её точек роста. Тогда*

а) *дополнение $Q_F^c = \biguplus_j (a_j;b_j)$ есть объединение не более чем счётного числа непересекающихся открытых интервалов, на каждом из которых функция F постоянна, т.е. $F(x) = F(y)$, $\forall\, x,y \in (a_j;b_j)$;*

б) *подмножество Q_F имеет полную меру относительно меры Лебега–Стилтьеса, порождённой F, т.е. мера дополнения $\mu_F(Q_F^c) = 0$.*

⊫ а) Из определения следует, что для любого $x \notin Q_F$, не являющегося точкой роста, найдётся $\varepsilon > 0$ такое, что $F(x-\varepsilon) = F(x) = F(x+\varepsilon)$, т.е. с каждым $x \notin Q_F$ можно связать открытый интервал $(l_x;r_x)$, где

$$l_x = \inf\{y:\ F(y) = F(x),\ y < x\}, \quad r_x = \sup\{y:\ F(y) = F(x),\ y > x\}.$$

Любая точка из интервала $(l_x;r_x)$ также не является точкой роста F, причём $F(y) = F(x)$ и $l_y = l_x$, $r_y = r_x$ для $\forall\, y \in (l_x;r_x)$. Для завершения доказательства заметим, что любой набор открытых интервалов числовой прямой содержит не более чем счётное число непересекающихся элементов.

б) По предыдущему множество $Q_F^c = \bigcup_1^N (a_j;b_j)$, $N \leqslant \infty$, т.е. это множество измеримо по Борелю. Кроме того, для каждого из интервалов мера

$$\mu_F(a_j;b_j) = \lim_{\varepsilon\to 0}{\uparrow}\, \mu_F(a_j+\varepsilon;b_j-\varepsilon) = \lim_{\varepsilon\to 0}(F(b_j+\varepsilon) - F(a_j-\varepsilon)) = 0,$$

что и требовалось. ⫤

54| **Следствие.** *Если мера Лебега множества точек роста ф.р. F равна нулю, то мера, порождённая F, сингулярна относительно меры Лебега.*

55| Пример. Любая дискретная функция (функция скачков) порождает меру Лебега–Стилтьеса, сосредоточенную в счётном числе точек. Поэтому такая мера сингулярна относительно меры Лебега.

С другой стороны, не всякая функция скачков имеет нулевое (по Лебегу) множество точек роста. Пусть

$$F(x) = \sum\nolimits_{q \leqslant x} p(q),$$

где сумма берётся по всем рациональным числам $q \leqslant x$, а строго положительные «веса» $p(q) > 0$ на множестве рациональных чисел выбраны произвольно, но так, чтобы сумма $\sum_q p(q) = 1$. Эта функция есть функция скачков, но множество её точек роста совпадает с $\mathbb{R}^1$. ◎

56| **Теорема.** [*Разложение Лебега для функций.*] *Любую ф.р. F можно единственным образом представить в виде*

$$F(x) = F_a(x) + F_d(x) + F_{cs}(x), \quad x \in \mathbb{R}^1,$$

где F_a, F_d — соответственно абсолютно непрерывная и дискретная функции, F_{cs} — непрерывная функция, порождающая сингулярную меру, $F_a(-\infty) = 0$.

⤇ Воспользовавшись разложением Лебега 559, стр. 491, для меры Лебега–Стилтьеса, представим (единственным образом) меру $\mu_F = \mu_a + \mu_s$. Тогда $F(x) = \mu_a(-\infty; x) + \mu_s(-\infty; x)$. По лемме 564, стр. 494, функция $F_a(x) = \mu_a(-\infty; x)$, генерирующая абсолютно непрерывную меру μ_a, есть абсолютно непрерывная функция, причём $F_a(-\infty) = 0$. Функция $\widetilde{F}(x) = \mu_s(-\infty; x)$ в соответствии с разложением Жордана представима в виде $\widetilde{F} = F_{cs} + F_d$ с функцией скачков F_d и непрерывной функцией F_{cs}, порождающей, очевидно, сингулярную меру. ⤆

Заметим, что дискретные функции не являются сингулярными, хотя меры Лебега–Стилтьеса, ими порождённые, сингулярны относительно меры Лебега.

57| Пример. Существуют сингулярные меры, у которых ф.р. непрерывна и всюду возрастает. Рассмотрим функцию Кантора $\mathcal{K}(x)$, генерирующую меру Лебега–Стилтьеса $\mu_{\mathcal{K}}$ на борелевской σ-алгебре $\mathbb{R}^1$, сосредоточенную на канторовом множестве $\mathcal{M} \subset [0; 1]$, и функцию скачков $F(x)$ из предыдущего примера, генерирующую меру Лебега–Стилтьеса μ_F, сосредоточенную на всех рациональных числах $Q \subset [0; 1]$. Определим свёртку этих функций (см. 112, стр. 117):

$$G(x) \overset{(a)}{=} \int_0^1 F(x-y)\, d\mathcal{K}(y) \overset{(b)}{=} \int_0^1 \mathcal{K}(x-y)\, dF(y), \qquad x \in [0; 2].$$

Из представления (a) следует строгая монотонность G, а из равенства (b) — её непрерывность. Рассмотрим множество всевозможных сумм $\mathcal{M} \oplus Q = \{x + r : x \in \mathcal{M}, r \in Q\}$. Поскольку мера Лебега λ инвариантна относительно сдвигов, то $\lambda(\mathcal{M} \oplus r) = \lambda(\mathcal{M}) = 0$, стало быть, $\lambda(\mathcal{M} \oplus Q) = \lambda(\cup_{r \in Q}(\mathcal{M} \oplus r)) = 0$. С другой стороны, по построению свёртки вероятностная мера

$$\mu_G(\mathcal{M} \oplus Q) \;=\; \mathbf{P}\{\xi + \eta \in \mathcal{M} \oplus Q\} \;\geqslant\; \mathbf{P}\{\xi \in \mathcal{M}, \eta \in Q\} \;=\; 1,$$

где независимые сл.в. $\xi \backsim \mathcal{K}$, $\eta \backsim F$. ◎

§ 5. Абсолютно непрерывные распределения

Напомним, что мера μ абсолютно непрерывна относительно меры Лебега λ ($\mu \ll \lambda$; допуская некоторую вольность речи, говорят просто абсолютно непрерывна), если мера $\mu(N) = 0$ для любого борелевского множества N, лебегова мера которого $\lambda(N) = 0$. Если распределение $\mathsf{P}_{\vec{\xi}}$ абсолютно непрерывно относительно меры Лебега, то по теореме Радона–Никодима 558, стр. 491,

$$\mathbf{P}\big\{\, \vec{\xi} \in B \,\big\} \;=\; \int\limits_B f(\vec{x})\, d\vec{x}, \qquad B \in \mathscr{B}^k, \tag{14}$$

где *плотность* распределения $f(\vec{x})$ можно выбрать так, чтобы

$$(\checkmark)\quad f(\vec{x}) \geqslant 0, \quad \forall\, \vec{x} \in \mathbb{R}^k, \qquad\qquad (\checkmark)\quad \int\limits_{\mathbb{R}^k} f(\vec{x})\, d\vec{x} \;=\; 1. \tag{15}$$

Часто плотность задают только на носителе $\mathscr{X} = \big\{\vec{x} : f(\vec{x}) > 0\big\}$ в виде $f = f(\vec{x})$, $\vec{x} \in \mathscr{X}$, подразумевая, что $f = 0$ при $\vec{x} \notin \mathscr{X}$. Если сл.вектор $\vec{\xi} \backsim f(\vec{x})$, $\vec{x} \in \mathscr{X}$, то математическое ожидание любой измеримой функции $h(\,\vec{\xi}\,)$ может быть найдено по формуле (20), стр. 493:

$$\mathbf{E}\, h(\,\vec{\xi}\,) \;=\; \int\limits_{\mathscr{X}} h(\vec{x}) f(\vec{x})\, d\vec{x} \;=\; \int\limits_{\mathbb{R}^k} \dot{\mathrm{I}}_{\mathscr{X}}(\vec{x}) h(\vec{x}) f(\vec{x})\, d\vec{x}. \tag{16}$$

Здесь $\big|\mathbf{E}\, h(\,\vec{\xi}\,)\big| < \infty$ т. т. т. когда интеграл сходится абсолютно, что всегда следует подчёркивать, т.к. он может вычисляться как римановский интеграл.

С другой стороны, для функций $F : \mathbb{R}^1\,(\mathbb{R}^k) \mapsto \mathbb{R}^1$ существует понятие абсолютной непрерывности, означающее, что для $\forall\, \varepsilon > 0\; \exists \delta > 0$ такое, что для любого не более чем счётного набора $\big\langle\, (a_j; c_j)\, \big\rangle_1^N$, $N \leqslant \infty$, непересекающихся открытых интервалов (параллелепипедов), суммарная длина (объём) которых не превосходит δ, суммарное приращение функции $\sum_1^N \bigl|F(c_j) - F(a_j)\bigr| \leqslant \varepsilon$ (в многомерном случае здесь сумма значений разностного оператора Δ). В соответствии с леммой 564, стр. 494, справедлива

58| Лемма. *Ф.р. F абсолютно непрерывна т. т. т. когда соответствующая ей вероятностная мера Лебега–Стилтьеса $\mu_F \ll \lambda$.*

59| Примеры. 1) Пусть ф.р. F дифференцируема всюду и её производная $f(x) = F'(x)$, $x \in \mathbb{R}^1$, ограничена, скажем, $\sup_x |f(x)| < M$. Тогда в силу формулы конечных приращений $\sum_1^N \bigl|F(c_j) - F(a_j)\bigr| \leqslant M \sum_1^N (c_j - a_j) < \varepsilon$, если $\sum_1^N (c_j - a_j) < \delta = \varepsilon / M$. Следовательно, ф.р. F абсолютно непрерывна.

2) Пусть ф.р. F непрерывна всюду и имеет кусочное представление с конечным числом интервалов $J_1, \ldots, J_n$, $\biguplus_1^n J_k = \mathbb{R}^1$:

$$F(x) = \sum\nolimits_1^n F_k(x)\,\dot{\mathrm{I}}\,(x; J_k).$$

Если каждая из функций F_k дифференцируема внутри «своего» интервала: $f_k(x) = F_k'(x)$, $x \in \overset{o}{J_k}$, и $\max_k \sup_{x \in \overset{o}{J_k}} |f_k(x)| \leqslant M$, то аналогично предыдущему можно установить, что F абсолютно непрерывна. Последнее условие выполняется, в частности, если все функции F_k непрерывно дифференцируемы на замыкании $[J_k]$ соответствующих интервалов (с левой или правой производной на краях интервалов). ◎

Условие абсолютной непрерывности можно связать с производной от ф.р. также следующим образом. Определим «правую» производную (‡):

$$F^{(k)}(\vec{x}) = \varliminf_{z \to +0} \frac{1}{z^k}\,\Delta_{\vec{x};\vec{x}+z} F.$$

Эта производная, как предел измеримых функций, измерима по Борелю на расширенной (точкой $+\infty$) полупрямой $\overline{\mathbb{R}}^1_+$ и, очевидно, неотрицательна. Кроме того, $F^{(k)}(\vec{x}) = \partial^k F(x_1, \ldots, x_k) \big/ \partial x_1 \cdots \partial x_k$ во всех точках $\vec{x} \in \mathbb{R}^k$, где существует смешанная производная (см. [23], т. 1, стр. 244–245).

60| Теорема. *Пусть $F(\vec{x})$ — функция распределения, непрерывная всюду на $\mathbb{R}^k$, P — распределение вероятностей, соответствующее F.*

I) *Мера P абсолютно непрерывна т. т. т. когда найдётся всюду неотрицательная функция (плотность) $f(\vec{x}) \geqslant 0$, $\vec{x} \in \mathbb{R}^k$, такая, что*

$$F(\vec{x}) \;=\; \int\limits_{\vec{u} \leqslant \vec{x}} f(\vec{u})\,d\vec{u} \;=\; \int\limits_{\mathbb{R}^k} \dot{\mathrm{I}}\,(\vec{u};\, \vec{u} \leqslant \vec{x})\, f(\vec{u})\,d\vec{u}. \tag{17}$$

(‡) $\vec{x} + z = (x_1 + z, \ldots, x_k + z)$.

II) *Если* $\int_{\mathbb{R}^k} F^{(k)}(\vec{x})\,d\vec{x} \geqslant 1$, *то мера* P *абсолютно непрерывна и можно выбрать плотность* $f(\vec{x}) = F^{(k)}(\vec{x})$; *при этом* $\int_{\mathbb{R}^k} F^{(k)}(\vec{x})\,d\vec{x} = 1$.

III) *Если функция* $f : \mathbb{R}^k \mapsto \mathbb{R}^1_+$ *удовлетворяет свойствам* (15), *то она есть плотность распределения с ф.р., определяемой* (17).

⊫ I) В одну сторону ($\Rightarrow$) этот факт следует из теоремы Радона–Никодима 558, стр. 491. Обратно, пусть для ф.р. справедливо представление (17). Определим с помощью функции f абсолютно непрерывную меру $\mu(B) = \int_B f(\vec{u})\,d\vec{u}$, $B \in \mathscr{B}^k$. Эта мера совпадает с P на множествах вида $B = \{\vec{u} : \vec{u} \leqslant \vec{x}\}$, что по лемме 12, стр. 34, гарантирует совпадение этих мер на всех подмножествах $\mathscr{B}^k$.

II) Определим функцию множеств

$$\mathsf{M}(B) = \int_B F^{(k)}(\vec{u})\,d\vec{u}, \qquad B \in \mathscr{B}^k.$$

По условию теоремы и по свойству интеграла Лебега эта функция есть абсолютно непрерывная мера на борелевских подмножествах. Покажем, что для любого конечного параллелепипеда $(\vec{a};\vec{c}]$ справедливо неравенство

$$\mathsf{P}\{(\vec{a};\vec{c}]\} \;=\; \Delta^{(k)}_{\vec{a};\vec{c}} F \;\geqslant\; \mathsf{M}(\vec{a};\vec{c}].$$

Заметим, что это соотношение гарантирует конечность меры M.

Так как ф.р. монотонна, то

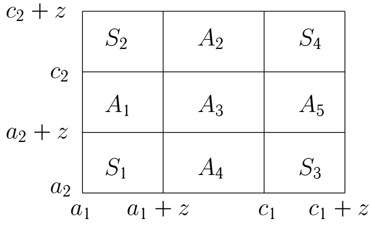
$$F(\vec{x}) \;\leqslant\; \frac{1}{z^k}\int_{(\vec{x},\vec{x}+z]} F(\vec{u})\,d\vec{u} \;\leqslant\; F(\vec{x}+z)$$

при $\forall\,\vec{x} \in \mathbb{R}^k$, $z > 0$. В силу непрерывности F отсюда получаем, что $z^{-k}\int_{(\vec{x},\vec{x}+z]} F(\vec{u})\,d\vec{u} \to F(\vec{x})$ при $z \to +0$. В целях сокращения записи дальнейшие построения будем производить в двумерном случае. Итак,

$$\begin{aligned}\mathsf{P}\{(\vec{a};\vec{c}]\} \;&=\; \Delta^{(k)}_{\vec{a};\vec{c}} F \;=\; F(c_1,c_2) - F(c_1,a_2) - F(a_1,c_2) + F(a_1,a_2) \;=\\ &=\; \lim_{z\to 0}\frac{1}{z^k}\Big(\int_{S_4} - \int_{S_2} - \int_{S_3} + \int_{S_1} F(\vec{u})\,d\vec{u}\Big)\end{aligned}$$

с областями

$$\begin{aligned} S_1 &= (a_1;a_1+z] \times (a_2;a_2+z],\\ S_2 &= (a_1;a_1+z] \times (c_2;c_2+z],\\ S_3 &= (c_1;c_1+z] \times (a_2;a_2+z],\\ S_4 &= (c_1;c_1+z] \times (c_2;c_2+z].\end{aligned}$$



Линейную комбинацию интегралов можно представить в виде интеграла от линейной комбинации индикаторных функций соответствующих областей:

$$\mathrm{I}_{S_4} - \mathrm{I}_{S_2} - \mathrm{I}_{S_3} + \mathrm{I}_{S_1} \; = \; \mathrm{I}_{\Theta_4} - \mathrm{I}_{\Theta_2} - \mathrm{I}_{\Theta_3} + \mathrm{I}_{\Theta_1},$$

где области

$$\begin{aligned}
\Theta_4 &= S_4 + A_2 + A_3 + A_5 = (a_1+z; c_1+z]\times(a_2+z; c_2+z],\\
\Theta_2 &= S_2 + A_1 + A_2 + A_3 = (a_1; c_1]\times(a_2+z; c_2+z],\\
\Theta_3 &= S_3 + A_3 + A_4 + A_5 = (a_1+z; c_1+z]\times(a_2; c_2],\\
\Theta_1 &= S_1 + A_1 + A_3 + A_4 = (a_1; c_1]\times(a_2; c_2].
\end{aligned}$$

В интегралах по $\Theta_4, \Theta_2, \Theta_3$ осуществим замену переменных (сдвиг) так, чтобы область интегрирования стала совпадать с Θ_1. Например,

$$\int\limits_{\Theta_2} F(u_1, u_2)\, du_1 du_2 \; = \; \int\limits_{a_1}^{c_1} \int\limits_{a_2+z}^{c_2+z} F(u_1, u_2)\, du_1 du_2 \; = \; \int\limits_{\Theta_1} F(u_1, u_2+z)\, du_1 du_2.$$

Следовательно, в силу леммы Фату 537, стр. 481,

$$\begin{aligned}
\mathrm{P}\{(\vec{a};\vec{c}]\} \; &= \; \lim_{z\to+0} \int\limits_{\Theta_1} \frac{\Delta_{\vec{u};\vec{u}+z}F}{z^k}\, d\vec{u} \; \geqslant \; \int\limits_{\Theta_1} \varliminf_{z\to+0} \frac{\Delta_{\vec{u};\vec{u}+z}F}{z^k}\, d\vec{u} \; = \\
&= \; \int\limits_{\Theta_1} F^{(k)}(\vec{u})\, d\vec{u} \; = \; \mathrm{M}(\vec{a};\vec{c}].
\end{aligned}$$

Ввиду непрерывности мер P, M аналогичные неравенства будут верны и для параллелепипедов, в которых некоторые из координат вершин равны $\pm\infty$.

Пусть $(a;c]$ — конечный интервал $\mathbb{R}^1$. Тогда объединение непересекающихся интервалов $(-\infty;a] \uplus (a;c] \uplus (c;\infty)$ задаёт разбиение $\mathbb{R}^1$. Прямое произведение двух таких разбиений

$$\begin{aligned}
\mathbb{R}^2 \; &= \; \big[(-\infty;a_1] \uplus (a_1;c_1] \uplus (c_1;\infty)\big]\times\big[(-\infty;a_2] \uplus (a_2;c_2] \uplus (c_2;\infty)\big] \; = \\
= \; &\big[(-\infty;a_1]\times(-\infty;a_2]\big] \; \uplus \; \big[(-\infty;a_1]\times(a_2;c_2]\big] \; \uplus \; \ldots \uplus \big[(c_1;\infty]\times(c_2;\infty]\big]
\end{aligned}$$

даёт разбиение $\mathbb{R}^2$ на 3^2 частей, одна из которых совпадает с прямоугольником $(a_1;c_1]\times(a_2;c_2]$. Аналогично строится разбиение пространства $\mathbb{R}^k$ на $N = 3^k$ непересекающихся параллелепипедов $Q_1, \ldots, Q_N$, один из которых, скажем Q_1, совпадает с конечным параллелепипедом $(\vec{a};\vec{c}]$, а у остальных координаты части вершин равны $\pm\infty$. Таким образом, по условию теоремы

$$1 \; = \; \mathrm{P}\{\mathbb{R}^k\} \; = \; \mathrm{P}\{Q_1\} + \textstyle\sum_2^N \mathrm{P}\{Q_j\} \; \geqslant$$

$$\geqslant \mathsf{M}\{Q_1\} + \textstyle\sum_2^N \mathsf{M}\{Q_j\} \;=\; \mathsf{M}\{\mathbb{R}^k\} \;\geqslant\; 1.$$

Следовательно, для любого конечного параллелепипеда $(\vec{a};\vec{c}\,]$ мера $\mathsf{M}(\vec{a};\vec{c}\,] = \mathsf{P}\{(\vec{a};\vec{c}\,]\}$. Отсюда по лемме 13, стр. 35, следует, что эти меры совпадают на всех борелевских множествах. Как уже отмечалось, мера M абсолютно непрерывна, а из её представления получаем по теореме Радона–Никодима, что в качестве плотности можно взять производную $F^{(k)}$.

Для доказательства III) проверим свойства ф.р. (см. 18, стр. 38). Свойства (F_1), (F_3), (F_4) следуют из теоремы Лебега, позволяющей переходить здесь к пределу под знаком интеграла. Например, если $\vec{x}_n \to \vec{y}$ и одна из координат $\vec{y}$ равна $-\infty$, то функция $\dot{\mathrm{I}}\,(\vec{u};\vec{u} \leqslant \vec{x}_n) \to 0$. Стало быть, $F(\vec{x}_n) \to 0$. Так же показывается непрерывность F. В силу тождества следующей леммы и по свойству линейности интеграла

$$\Delta_{\vec{a},\vec{c}} F = \int_{\mathbb{R}^k} \dot{\mathrm{I}}\,(\vec{u};(\vec{a};\vec{c}\,])f(\vec{u})\,d\vec{u} \geqslant 0. \qquad \Lleftarrow$$

61| **Лемма.** (?!) *Для индикатора* $\dot{\mathrm{I}}\,(\vec{u};\vec{u} \leqslant \vec{x})$, *как функции* $\vec{x}$ *при фиксированном* $\vec{u}$, *оператор* $\Delta_{\vec{a},\vec{c}}\,\dot{\mathrm{I}}\,(\vec{u};\vec{u} \leqslant \vec{x}) = \dot{\mathrm{I}}\,(\vec{u};\vec{u} \in (\vec{a};\vec{c}\,])$.

⚠ По построению интеграла Лебега от функции, принимающей бесконечные значения, равенство $\int_{\mathbb{R}^k} F^{(k)}(\vec{x})\,d\vec{x} = 1$ гарантирует, что мера Лебега тех точек $\vec{x}$, в которых $F^{(k)}(\vec{x}) = +\infty$, равна нулю. Поэтому для таких точек можно положить по определению $F^{(k)}(\vec{x}) = 0$.

62| Примеры. 1) Определим с помощью некоторой одномерной ф.р. $G(x)$ двумерную ф.р. $F(x,y) = G^2(y)$, если $x \geqslant y$, и $F(x,y) = G(x)G(y)$, если $x \leqslant y$. Эта функция удовлетворяет свойствам ф.р. Например, при $a_2 < a_1 < b_2 < b_1$, т.е. в ситуации, когда одна вершина (a_1,b_2) прямоугольника $A = (a_1;b_1] \times (a_2;b_2]$ лежит выше диагонали $y = x$ $(x < y)$, а три другие — ниже диагонали $(x > y)$,

$$\begin{aligned} \Delta_A F \;&=\; F(b_1,b_2) - F(b_1,a_2) - F(a_1,b_2) + F(a_1,a_2) \;= \\ =\; G^2(b_2) - G^2(a_2) - G(a_1)G(b_2) + G^2(a_2) \;&=\; G(b_2)(G(b_2) - G(a_1)) \;\geqslant\; 0. \end{aligned}$$

Остальные варианты расположения вершин исследуются аналогичным образом. В частности, когда все вершины прямоугольника лежат ниже диагонали, т.е. при $a_2 < b_2 < a_1 < b_1$, получаем, что вероятность этого прямоугольника равна нулю: $\mathsf{P}_F\{A\} = \Delta_A F = G^2(b_2) - G^2(a_2) - G^2(b_2) + G^2(a_2) = 0$. Поскольку любое открытое множество может быть представлено в виде счётного объединения прямоугольников подобного типа, то отсюда следует, что область строго ниже диагонали имеет нулевую вероятность.

Пусть G — ф.р. равномерного $\mathfrak{Un}(0,1)$ закона. Ясно, что ф.р. $F(x,y)$ непрерывна всюду и дифференцируема всюду, кроме границ единичного квадрата и его диагонали. Если положить производную $F''(x,y) = 1 \cdot \dot{\mathrm{I}}\,(0 < x < y < 1)$, то $\int_{\mathbb{R}^2} F''(x,y)\,dxdy = {}^1\!/_2 < 1$. Условие теоремы 60 не выполнено. Следовательно, мера P_F, генерируемая ф.р. F, не является абсолютно непрерывной. Функция $2F''$ есть не что иное, как плотность абсолютно непрерывной части $\mathbf{P}_a$ разложения Лебега меры Лебега–Стилтьеса μ_F; сингулярная часть $\mathbf{P}_s$ порождается ф.р. $F_s(x,y) = (\min(x,y))^2$. Как отмечалось в 20, стр. 40, распределение $\mathbf{P}_s$ сосредоточено на биссектрисе $y = x$.

Из выражения для ф.р. F не сразу видно, куда «пропала» вероятностная масса ${}^1\!/_2$ в результате вычисления производной. На самом деле эта ф.р. задаёт меру P_F, для которой вероятность диагонального множества $y = x$ отлична от нуля, в то время как мера Лебега этого множества равна нулю. Зафиксируем целое $N > 0$ и накроем диагональ квадратами $A_k = (a_{k-1}, a_k] \times (a_{k-1}, a_k]$, где $a_k = k2^{-N},\ k = 1, \ldots, 2^N$. Мера $\mathsf{P}_F(A_k) = k2^{-2N}$, поэтому

$$\mathsf{P}_F(y = x) \ = \ \lim_{N\to\infty} \sum\nolimits_1^{2^N} \mu_F(A_k) \ = \ \lim_{N\to\infty} \frac{2^N(2^N+1)}{2 \cdot 2^{2N}} = \frac{1}{2}\,.$$

В действительности этот факт справедлив для любой непрерывной ф.р. G. Легко понять, что функция F есть ф.р. вектора $(\xi, \max\{\xi, \eta\})$, где ξ, η — независимые сл.в. с общей ф.р. G. Поэтому вероятностную меру P_F тех точек, где $y = x$, можно записать следующим образом:

$$\mathsf{P}_F(y = x) \ = \ \mathbf{P}\{\xi = \max\{\xi, \eta\}\} \ = \ \mathbf{P}\{\xi \geqslant \eta\}\,. \tag{18}$$

Для сл.в. с непрерывной ф.р. вероятность (18) равна (?!) ${}^1\!/_2$.

2) Функция $F(x) = \sqrt{x},\ 0 \leqslant x \leqslant 1$, обладает неограниченной производной $f(x) = 1/2\sqrt{x},\ 0 < x \leqslant 1$. Однако эта функция также абсолютно непрерывна, поскольку интеграл от её производной, очевидно, равен 1. ◎

► Как уже отмечалось, дискретные распределения сингулярны относительно меры Лебега. Однако любое такое распределение абсолютно непрерывно относительно любой дискретной меры с более широким носителем. Большинство популярных дискретных сл.в. сосредоточены на подмножествах целых чисел $\mathbb{R}^1$. Поэтому все такие вероятностные распределения абсолютно непрерывны относительно меры $\mathcal{Z}$, считающей целые числа $\mathbb{R}^1$ (см. пример 462, стр. 425). Если $p(x)$, $x \in \langle 0, \pm 1, \dots \rangle$ — ф.вер. дискретной сл.в. ξ, то теорема Радона–Никодима в этом случае представляет собой тривиальное тождество

$$\mathbf{E}\, h(\xi) = \int_{\mathbb{R}} h(x)p(x)\, \mathcal{Z}(dx),$$

справедливое для любой неотрицательной функции h. Поэтому часто ф.вер. p называют функцией плотности с упоминанием, что она вычисляется относительно считающей меры. Ввиду этого, доказательства многих свойств плотностей относительно меры Лебега могут быть после элементарных модификаций перенесены для доказательства аналогичных свойств ф.вер.

✧ **Частная плотность. Независимость.** Если вектор $\vec{\xi}$ имеет абсолютно непрерывное распределение, то это справедливо и для любых частных распределений (относительно соответствующей меры Лебега).

63| **Теорема.** *Пусть распределение вектора* $\vec{\xi} = (\xi_1, \dots, \xi_k)$ *абсолютно непрерывно относительно меры Лебега в* $\mathbb{R}^k$ *с плотностью* $f(\vec{x})$, *тогда сл.вектор* (ξ_1, ξ_2) *и сл.в.* ξ_1 *также имеют абсолютно непрерывные распределения относительно мер Лебега в* $\mathbb{R}^2$ *и* $\mathbb{R}^1$ *соответственно с плотностями*

$$f_{(\xi_1,\xi_2)}(x_1, x_2) = \int\limits_{\mathbb{R}^{k-2}} f(x_1, x_2, \vec{x}^{(3)})\, d\vec{x}^{(3)},$$

$$f_{\xi_1}(x_1) = \int\limits_{\mathbb{R}^{k-1}} f(x_1, \vec{x}^{(2)})\, d\vec{x}^{(2)},$$

где $\vec{x}^{(j)} = (x_j, \dots, x_k)$, $j = 2, 3$.

⊫ Пусть $B \subset \mathbb{R}^2$ — произвольное борелевское множество, тогда прямое произведение (цилиндр) $B \times \mathbb{R}^{k-2}$ будет борелевским множеством в $\mathbb{R}^k$. По определению ф.пл. $f(\vec{x})$ сл.вектора $\vec{\xi}$

$$\mathbf{P}\{ (\xi_1, \xi_2) \in B \} = \int_{\mathbb{R}^k} \dot{\mathrm{I}}\left(\vec{x}; B \times \mathbb{R}^{k-2}\right) f(\vec{x})\, d\vec{x}.$$

Очевидно, значение индикаторной функции $\dot{\mathrm{I}}\left(\vec{x}; B \times \mathbb{R}^{k-2}\right) = \dot{\mathrm{I}}\left((x_1, x_2); B\right)$. Поэтому в силу теоремы Фубини

$$\mathbf{P}\{ (\xi_1, \xi_2) \in B \} = \int_{\mathbb{R}^2} \dot{\mathrm{I}}\left((x_1, x_2); B\right)\Big(\int_{\mathbb{R}^{k-2}} f(\vec{x})\, d\vec{x}^{(3)}\Big)\, d(x_1, x_2) =$$
$$= \int_B f_{(\xi_1,\xi_2)}(x_1, x_2)\, d(x_1, x_2).$$

Второе утверждение теоремы доказывается аналогично. ⫤

Другими словами, для нахождения частной плотности подвектора необходимо совместную плотность проинтегрировать по всем переменным, не относящимся к этому подвектору.

Обратное утверждение справедливо только при дополнительных предположениях о структуре связей между компонентами вектора.

64| **Теорема.** I) *Если сл.в.* $\xi_1, \dots, \xi_k$ *независимы в совокупности и распределение* ξ_j *абсолютно непрерывно с плотностью* f_j, $j = \overline{1,k}$, *то распределение сл.вектора* $\vec{\xi} = (\xi_1, \dots, \xi_k)$ *имеет плотность*

$$f_{\vec{\xi}}(\vec{x}) \;=\; \prod_{j=1}^{k} f_j(x_j), \qquad \vec{x} = (x_1, \dots, x_k) \in \mathbb{R}^k. \tag{19}$$

II) *Если плотность сл.вектора* $\vec{\xi}$ *представима в виде* (19), *то сл.в.* $\xi_1, \dots, \xi_k$ *независимы в совокупности и сл.в.* ξ_j *имеет плотность* f_j.

⫢ Схема доказательства теоремы весьма проста. Например, для I) необходимо: а) с учётом независимости, записать ф.р. $F_{\vec{\xi}}$ как произведение одномерных ф.р.; б) воспользовавшись абсолютной непрерывностью, представить каждую одномерную ф.р. в виде интеграла от плотности; в) применить теорему Фубини для представления ф.р. в виде (17). ⫤

⚠ Соображения, приведённые в предыдущем разделе на стр. 67, позволяют сделать вывод, что аналогичное утверждение (с заменой ф.пл. f_j на ф.вер. p_j) справедливо и для дискретных сл.в.

65| Примеры. 1) Легко видеть, что для сл.вектора (ξ_1, ξ_2) с ф.р. $F(x_1, x_2) = \min\{x_1, x_2\}$, $0 \leqslant x_1, x_2 \leqslant 1$, вероятность попадания в прямоугольник $(a_1; c_1] \times (a_2; c_2]$, не пересекающий диагональ $x_1 = x_2$ единичного квадрата $[0;1]^2$, равна нулю: если $a_2 > c_1$ или $a_1 > c_2$, то

$$\mathbf{P}\big\{ (\xi_1, \xi_2) \in (\vec{a}; \vec{c}\,] \big\} \;=\; F(c_1, c_2) - F(c_1, a_2) - F(a_1, c_2) + F(a_1, a_2) \;=\; 0.$$

Вектор (ξ_1, ξ_2) сосредоточен на диагонали. Распределение (ξ_1, ξ_2) не абсолютно непрерывно, т.к. вероятность диагонали $\mathbf{P}\big\{ \xi_1 = \xi_2 \big\} = 1$, в то время как её плоская мера Лебега равна нулю. Однако оба маргинальных распределения совпадают с абсолютно непрерывным равномерным распределением $\mathcal{U}\mathrm{n}(0,1)$ (например, $F_1(x_1) = F(x_1, 1) = x_1$).

2) Сл.вектор $\vec{\xi}$ имеет *равномерное* распределение в области $\mathscr{X} \subset$

$\subset \mathbb{R}^k$: $\vec{\xi} \backsim \mathtt{Un}(\mathscr{X})$, если вероятность попадания вектора в любое подмножество B этой области пропорциональна объёму B :

$$\mathbf{P}\{\,\vec{\xi} \in B\,\} \;=\; \frac{V(B \cap \mathscr{X})}{V(\mathscr{X})} \;=\; \frac{1}{V(\mathscr{X})} \int\limits_B \dot{\mathrm{I}}_{\mathscr{X}}(\vec{x})\, d\vec{x}, \qquad B \in \mathscr{B}^k, \tag{20}$$

где объём (мера Лебега) $0 < V(\mathscr{X}) = \int_{\mathscr{X}} d\vec{x} < \infty$. В силу (14) плотность равномерного распределения $f(\vec{x}) = 1/V(\mathscr{X})$, $\vec{x} \in \mathscr{X}$.

Плотность равномерного распределения в k-мерном кубе $\mathscr{X} = [0;1]^k$ равна 1 при $\vec{x} \in \mathscr{X}$. Очевидно, эта плотность удовлетворяет (19), т.е. компоненты соответствующего сл.вектора независимы в совокупности.

Если $V(\mathscr{X}) = 0$, то при описании равномерного распределения учитывают, что этот вектор всегда принадлежит $\mathscr{X}$, и вычисляют объём подмножеств посредством меры Лебега, определённой на $\mathscr{X}$. Например, вектор (ξ_1, ξ_2) имеет равномерное распределение на единичной окружности, если $\xi_1^2 + \xi_2^2 = 1$ и вероятность попадания на любую дугу равна длине дуги, делённой на 2π.

3) Другому популярному распределению — нормальному — будет посвящена целая глава. Здесь же приведём один частный вариант двумерной плотности с коэффициентом корреляции ρ $(|\rho| < 1)$:

$$f(x_1, x_2) \;=\; \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\Big\{ -\frac{1}{2(1-\rho^2)}\big(x_1^2 - 2\rho x_1 x_2 + x_2^2\big)\Big\}, \tag{21}$$

из которого видно, что плотность разбивается в произведение одномерных нормальных плотностей (по x_1 и x_2) т. т. т. когда $\rho = 0$, т.е. компоненты нормального вектора независимы, только если они не коррелируют.

4) Ещё одна популярная модель — *модель Дирихле*, или *многомерное бета-распределение*. Бета-распределение применяется для описания сл.в., принимающих значения в отрезке $[0;1]$, например для относительной частоты какого-либо события. Аналогично, если $(\xi_1, \ldots, \xi_{k+1})$ — набор вероятностей предпочтения одной из $(k+1)$ возможных категорий, например вероятностей выбора того или иного телеканала конкретным человеком, то изменчивость этого набора (между людьми) можно описать с помощью распределения Дирихле. Так как сумма $\xi_1 + \ldots + \xi_{k+1} = 1$, то удобнее перейти к вектору $(\xi_1, \ldots, \xi_k)$, отнеся, как всегда в такой ситуации, в $(k+1)$-ю категорию «всё остальное», и задавать его плотность распределения в пространстве $\mathbb{R}^k$, точ-

нее, в части $\mathscr{X}_{\mathfrak{D}}$ гиперкуба $[0;1]^k$:

$$f_{\mathfrak{D}}(\vec{x}) \;=\; \frac{1}{\mathrm{B}(a_1,\ldots,a_{k+1})}\prod_{j=1}^{k} x_j^{a_j-1}\Big(1-\Big(\sum_{j=1}^{k} x_j\Big)\Big)^{a_{k+1}-1}, \quad \vec{x}\in\mathscr{X}_{\mathfrak{D}}, \tag{22}$$

$$\mathscr{X}_{\mathfrak{D}} \;=\; \Big\langle\, \vec{x}\in\mathbb{R}^k:\;\; 0\leqslant x_1,\ldots,x_k\leqslant 1, \quad x_1+\ldots+x_k\leqslant 1 \,\Big\rangle,$$

$f_{\mathfrak{D}}(\vec{x}) = 0$ вне носителя $\mathscr{X}_{\mathfrak{D}}$; параметры модели $a_j > 0$ для $\forall j = 1,\ldots,k+1$. Кратко пишем $\vec{\xi} \backsim \mathcal{D}\mathtt{ir}_k(a_1,\ldots,a_{k+1})$. Нормирующая константа $\mathrm{B}(a_1,\ldots,a_{k+1})$, обеспечивающая выполнение условия $\int_{\mathscr{X}_{\mathfrak{D}}} f_{\mathfrak{D}}(\vec{x})\,d\vec{x} = 1$, вычисляется через многомерную бета-функции Эйлера

$$\mathrm{B}(a_1,\ldots,a_{k+1}) \;:=\; \frac{\Gamma(a_1)\cdots\Gamma(a_{k+1})}{\Gamma(a_1+\ldots+a_{k+1})}.$$

При $k=1$ распределение Дирихле совпадает с бета-распределением: $\mathcal{D}\mathtt{ir}_1(a_1,a_2) \backsim \mathcal{B}\mathtt{et}(a_1,a_2)$. При $k=2$ и фиксированном x_1

$$\int_{\mathbb{R}^1} f_{\mathfrak{D}}(x_1,x_2)\,dx_2 \;=\; \frac{1}{\mathrm{B}(a_1,a_2,a_3)}\int_0^{1-x_1} x_1^{a_1-1}x_2^{a_2-1}(1-x_1-x_2)^{a_3-1}\,dx_2.$$

Подстановка $x_2 \dashrightarrow (1-x_1)t,\ t\in[0;1]$, переводит правую часть в

$$\frac{1}{\mathrm{B}(a_1,a_2,a_3)}\; x_1^{a_1-1}(1-x_1)^{a_2+a_3-1}\int_0^1 t^{a_2-1}(1-t)^{a_3-1}\,dt,$$

где интеграл есть бета-функция $\mathrm{B}(a_2,a_3)=\Gamma(a_2)\Gamma(a_3)/\Gamma(a_2+a_3)$. Поэтому

$$\int_{\mathbb{R}} f_{\mathfrak{D}}(x_1,x_2)\,dx_2 \;=\; \frac{1}{\mathrm{B}(a_1,a_2+a_3)}\; x_1^{a_1-1}(1-x_1)^{a_2+a_3-1}, \qquad 0<x_1<1,$$

что совпадает с бета-плотностью $\mathcal{B}\mathtt{et}(a_1,a_2{+}a_3)$. Аналогичные выкладки справедливы и для $k>2$. Другими словами, проинтегрировав функцию (22) по какой-либо переменной x_j по всей области возможных значений (например, $0\leqslant x_2\leqslant 1-x_1-x_3-\ldots-x_k$), мы получим функцию того же вида.

Таким образом, сл.вектор $(\xi_1,\ldots,\xi_{k-1}) \backsim \mathcal{D}\mathtt{ir}_{(k-1)}(a_1,\ \ldots,a_{k-1},a_k+a_{k+1})$ в соответствии с теоремой 63. И аналогично для других подвекторов: все исключённые сл.в. присоединяются вместе со своими параметрами к категории «всё остальное», в частности $\xi_1 \backsim \mathcal{B}\mathtt{et}(a_1,a_2+\ldots+a_{k+1})$. Позднее мы покажем, что если часть сл.в. объединить (путем сложения) в новую сл.в., то новый вектор снова будет иметь распределение Дирихле. Например, $\big(\xi_1,\ldots,\xi_{k-2},(\xi_{k-1}+\xi_k)\big) \backsim \mathcal{D}\mathtt{ir}_{(k-1)}(a_1,\ldots,(a_{k-1}+a_k),a_{k+1})$.

Приведённые здесь свойства, как и многие другие свойства распределения

Дирихле, легко получаются из представления этого распределения через отношения независимых гамма сл.в. Доказательство следующей леммы можно осуществить, например, средствами главы II (см. Упр. B.7, стр. 133).

66| **Лемма.** (?!) *Пусть* $\left\langle \gamma_j \backsim \mathcal{G}\mathtt{am}(a_j, 1) \right\rangle_1^{k+1}$ *— независимые гамма сл.в. с параметром масштаба 1. Тогда*

$$\left(\frac{\gamma_1}{\gamma_1 + \ldots + \gamma_{k+1}}, \ldots, \frac{\gamma_k}{\gamma_1 + \ldots + \gamma_{k+1}}\right) \backsim \mathcal{D}\mathtt{ir}_k(a_1, \ldots, a_{k+1}).$$

◎

67| Пример. Рассмотрим ф.пл. $f(\vec{x}) = 1 + \prod_{j=1}^{k}(x_j - {}^1\!/_2)$, $\vec{x} \in [0;1]^k$. Любое m-мерное ($m < k$) маргинальное распределение есть равномерное распределение в кубе $[0;1]^m$, т.е. компоненты любого маргинального распределения независимы, однако ф.пл. f не удовлетворяет равенству (19), т.е. весь k-мерный вектор не является независимым в совокупности. ◎

§ 6. Числовые характеристики. Коэффициент корреляции

Для каждой компоненты ξ_j сл.вектора $\vec{\xi} \backsim F$ можно определить среднее как интеграл Лебега–Стилтьеса относительно соответствующей маргинальной ф.р.: $\mu_j = \mathbf{E}\,\xi_j = \int_{\mathbb{R}} x\,dF_{\xi_j}(x)$, или как интеграл Римана $\int_{-\infty}^{\infty} x f_{\xi_j}(x)\,dx$, если распределение ξ_j имеет плотность f_{ξ_j} и интеграл сходится абсолютно. Ранее, (12), стр. 52, математическое ожидание любой функции $h(\vec{\xi}\,)$ определялось через совместную ф.р. F.

68| **Лемма.** *Пусть* $h : \mathbb{R}^1 \mapsto \mathbb{R}^1$ *— борелевская функция,* $\widetilde{h} : \mathbb{R}^2 \mapsto \mathbb{R}^1$ *— её естественное продолжение на пространство* $\mathbb{R}^2$: $\widetilde{h}(x,y) = h(x)$, $\forall\, x \in \mathbb{R}^1, y \in \mathbb{R}^1$. *Тогда для сл.вектора* $\vec{\zeta} = (\xi, \eta)$ *с ф.р.* F *математическое ожидание*

$$\mathbf{E}\,\widetilde{h}(\xi,\eta) \overset{\mathrm{df}}{=} \int_{\mathbb{R}^2} \widetilde{h}(x,y)\,dF(x,y) = \int_{\mathbb{R}} h(x)\,dF_{\xi}(x) = \mathbf{E}\,h(\xi), \qquad (23)$$

где F_ξ *— (маргинальная) ф.р.* ξ *и интегралы существуют или не существуют, конечны или бесконечны одновременно.*

⚠ Утверждение леммы справедливо также для функций h на $\mathbb{R}^k$, $k \geqslant 1$, и сл.векторов $\vec{\zeta} = (\vec{\xi}, \vec{\eta}) \in \mathbb{R}^{k+m}$.

⤇ См. стр. 104 ⤆

⚠ Короткое доказательство этой теоремы с помощью варианта формулы Фубини с переходными (условными) ф.р. см. в примере 97, стр. 99.

✧ **Линейная регрессия**. Пусть теперь $\sigma_j^2 = \mathbf{D}\xi_j = \mathbf{E}(\xi_j - \mu_j)^2$ — дисперсия j-й компоненты сл.вектора, $j = \overline{1,k}$. Напомним одно из свойств дисперсии: если дисперсия конечна, то она доставляет минимум среднеквадратической ошибки прогноза сл.в. посредством константы, т.е. $\sigma_j^2 = \min_c \mathbf{E}(\xi_j - c)^2$. Найдём минимум среднеквадратической ошибки прогноза сл.в. ξ_1 с помощью линейной функции от ξ_2.

Определение. Линейная функция $x_1 = \alpha + \beta x_2$ есть *линейная среднеквадратическая регрессия* ξ_1 на ξ_2, если

$$\mathbf{E}(\xi_1 - \alpha - \beta\,\xi_2)^2 \;=\; \min_{b,c} \mathbf{E}(\xi_1 - c - b\,\xi_2)^2 \qquad (< \infty).$$

Задача минимизации здесь эквивалентна (?!) поиску минимума

$$\begin{aligned} S(b,c) \;&=\; \mathbf{E}\big[(\xi_1-\mu_1) - c - b(\xi_2-\mu_2)\big]^2 \;=\; 2bc\,\underline{\mathbf{E}(\xi_2-\mu_2)} - 2c\,\underline{\mathbf{E}(\xi_1-\mu_1)} + \\ &\quad + \mathbf{E}(\xi_1-\mu_1)^2 + c^2 + b^2\,\mathbf{E}(\xi_2-\mu_2)^2 - 2b\,\mathbf{E}\big[(\xi_1-\mu_1)(\xi_2-\mu_2)\big] \;= \\ &=\; \sigma_1^2 + c^2 + b^2\sigma_2^2 - 2b\sigma_{12}, \end{aligned}$$

где подчеркнутые слагаемые равны нулю и в последнем равенстве использовано обозначение $\sigma_{12} = \mathbf{E}\big[(\xi_1-\mu_1)(\xi_2-\mu_2)\big]$.

Очевидно, минимум по c достигается при $c = 0$, а минимум квадратичной функции (по b) достигается при $b = \sigma_{12}/\sigma_2^2$. Поэтому

$$\min_{b,c} S(b,c) \;=\; \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2} \;=\; \sigma_1^2\Big(1 - \Big(\frac{\sigma_{12}}{\sigma_1\sigma_2}\Big)^2\Big).$$

Таким образом, качество линейного прогноза, т.е. величина так называемой остаточной дисперсии, полностью (не считая дисперсию σ_1^2) характеризуется коэффициентом $\sigma_{12}/(\sigma_1\sigma_2)$. Заметим, что если дисперсии σ_1^2, σ_2^2 конечны, то и $|\sigma_{12}| < \infty$. Этот факт легко получить из неравенства $0 \leqslant S(b,0) = \sigma_1^2 + b^2\sigma_2^2 - 2b\,\sigma_{12}$, если выбрать в нём сначала $b = 1$, а затем $b = -1$.

⚠ Последний факт обычно доказывают, ссылаясь на неравенство Коши–Буняковского (Шварца) $\big|\mathbf{E}\,\xi\eta\big| \leqslant \sqrt{\mathbf{E}\,\xi^2\,\mathbf{E}\,\eta^2}$ (см. 553, 554, стр. 489).

Определения. *Коэффициентом ковариации* называется величина $\sigma_{12} = \mathbf{Cov}(\xi_1,\xi_2) := \mathbf{E}\big[(\xi_1-\mu_1)(\xi_2-\mu_2)\big]$, где $\mu_j = \mathbf{E}\,\xi_j$.

Если дисперсии $0 < \sigma_j^2 = \mathbf{D}\xi_j < \infty,\ j = 1, 2$, то величина

$$\rho \ := \ \mathbf{Corr}(\xi_1, \xi_2) \ = \ \frac{\mathbf{Cov}(\xi_1, \xi_2)}{\sqrt{\mathbf{D}\xi_1 \mathbf{D}\xi_2}} \ = \ \frac{\sigma_{12}}{\sigma_1 \sigma_2}$$

называется *коэффициентом корреляции* (между) сл.в. ξ_1, ξ_2. Когда коэффициент корреляции равен нулю, говорят, что сл.в. *не коррелируют*.

69| Лемма. (?!) $\mathbf{Cov}(\xi_1, \xi_2) = \mathbf{E}[\xi_1 \xi_2] - \mu_1 \mu_2$.

Первое утверждение теоремы 43, стр. 55, обобщается на случай произведения двух сл.в. (доказательство см. стр. 104).

70| Лемма. [*Формула Гёфдинга.*] *Пусть $F(x, y)$ — совместная ф.р. (ξ_1, ξ_2), F_1, F_2 — маргинальные ф.р. ξ_1, ξ_2 соответственно и пусть дисперсии $\mathbf{D}\xi_1, \mathbf{D}\xi_2 < \infty$. Тогда ковариация*

$$\mathbf{Cov}(\xi_1, \xi_2) \ = \ \iint_{\mathbb{R}^2} \big(F(x, y) - F_1(x) F_2(y)\big)\, dxdy. \tag{24}$$

Слово *relation* (связь, зависимость) с приставкой *co* даёт право называть коэффициент корреляции коэффициентом взаимосвязи, взаимозависимости сл.в. Принимая во внимание свойства этого коэффициента (ниже), его следует называть коэффициентом линейной взаимозависимости. Аналогично, коэффициент ковариации можно назвать коэффициентом соизменчивости.

Таким образом, справедлива

71| Теорема. *Уравнение $x_1 = \mu_1 + \rho \dfrac{\sigma_1}{\sigma_2}(x_2 - \mu_2)$ определяет наилучший линейный прогноз сл.в. ξ_1 по значениям $\xi_2 = x_2$ (линию регрессии ξ_1 на ξ_2). Дисперсия ошибки прогноза равна $\sigma_1^2(1 - \rho^2)$.*

✧ **Свойства коэффициента корреляции.** Остаточная дисперсия линейной регрессии всегда неотрицательна и по теореме 71 равна $(1 - \rho^2)\mathbf{D}\xi_1$. Отсюда легко выводятся свойства коэффициента корреляции.

72| Теорема. *Пусть ρ — коэффициент корреляции ξ_1, ξ_2.*

(R_1) $-1 \leqslant \rho \leqslant 1$.

(R_2) $\rho = \pm 1$ *т. т. т. когда* $\xi_1 = \alpha + \beta\, \xi_2$ (п.н.),

то есть сл.в. ξ_1, ξ_2 линейно связаны, причём знак $\operatorname{sign}(\beta) = \operatorname{sign}(\rho)$.

(R_3) *Если сл.в. ξ_1, ξ_2 независимы, то $\rho = 0$.*

(R_4) $\mathbf{Corr}(v\,\xi + u, b\,\eta + a) = \operatorname{sign}(vb)\, \mathbf{Corr}(\xi, \eta)$

для $\forall\, a, b, u, v \in \mathbb{R}^1,\ vb \neq 0$, т.е. коэффициент корреляции не изменяется при *линейных преобразованиях сл.в., если угловые коэффициенты преобразований имеют одинаковый знак.*

$\Rrightarrow$ (R_1) следует из неравенства $\sigma_1^2(1-\rho^2) \geqslant 0$; (R_2) — из соотношений $|\rho| = 1 \;\Leftrightarrow\; 0 = \sigma_1^2(1-\rho^2) = \mathbf{E}\left[\xi-\mu_1-\beta(\xi_2-\mu_2)\right]^2$ и свойств интеграла Лебега; (R_3) следует из теоремы 48, стр. 56:

$$\mathbf{E}\left[(\xi_1-\mu_1)(\xi_2-\mu_2)\right] \;=\; \mathbf{E}(\xi_1-\mu_1)\,\mathbf{E}(\xi_2-\mu_2) \;=\; 0.$$

Свойство (R_4) легко проверить (?!) непосредственной подстановкой в формулу для коэффициента корреляции. $\Lleftarrow$

⚠ Свойство (R_3), вообще говоря, необратимо. Можно привести примеры (см. 99, стр. 100), в которых сл.в. зависимы, но не коррелируют. С другой стороны, для некоторых частных многомерных моделей, например нормальной (стр. 70) и пуассоновской (пример 127, стр. 122), понятия независимости и некоррелированности совпадают.

Отметим ещё одно свойство регрессии. Остаток прогноза $\xi_1-\alpha-\beta\xi_2$ не коррелирует (?!) с предикторной сл.в. ξ_2 :

$$\mathbf{Corr}\,(\xi_2,\xi_1-\alpha-\beta\xi_2) = 0. \tag{25}$$

Из (25) следует, что дисперсия $\sigma_1^2 = \sigma_{1\cdot2}^2+\sigma_{1/2}^2$, где $\sigma_{1\cdot2}^2 = \sigma_1^2(1-\rho^2)$ — остаточная дисперсия, $\sigma_{1/2}^2$ — дисперсия регрессии. Другими словами, изменчивость сл.в. ξ_1 на $\rho^2\cdot 100\%$ можно объяснить влиянием на неё ξ_2.

73| Упр. а) Покажите, что $\mathbf{D}(\xi_1+\xi_2) \;=\; \mathbf{D}\xi_1+\mathbf{D}\xi_2+2\rho\sqrt{\mathbf{D}\xi_1\mathbf{D}\xi_2}$. Усиливая известное свойство дисперсии, можно сказать, что дисперсия суммы двух сл.в. равна сумме их дисперсий т. т. т. когда сл.в. не коррелируют.

б) Опишите ситуации, когда $\mathbf{D}(\xi_1-\xi_2) \;=\; \mathbf{D}\xi_1-\mathbf{D}\xi_2$.

Определения. Квадратная матрица $\boldsymbol{\Sigma}$ (сигма), составленная из всех пар ковариаций сл.вектора $\vec{\xi}$:

$$\boldsymbol{\Sigma} \;=\; \mathbf{Cov}(\,\vec{\xi}\,) \;:=\; \Big(\mathbf{Cov}\,(\xi_i,\xi_j)\Big)_{i,j=1}^{k} \;=\; \Big(\sigma_{ij}\Big)_{i,j=1}^{k},$$

называется *матрицей ковариаций* или *дисперсионной матрицей*. Диагональ этой матрицы содержит дисперсии сл.в.: $\sigma_{jj} = \mathbf{Cov}\,(\xi_j,\xi_j) \;=\; \mathbf{D}\xi_j$.

Матрица, составленная из всех пар корреляций сл.вектора $\vec{\xi}$:

$$\mathcal{P} \;=\; \mathbf{Corr}(\,\vec{\xi}\,) \;:=\; \Big(\mathbf{Corr}\,(\xi_i,\xi_j)\Big)_{i,j=1}^{k} \;=\; \Big(\rho_{ij}\Big)_{i,j=1}^{k},$$

называется *матрицей корреляций* сл.вектора $\vec{\xi}$. Диагональные элементы этой матрицы $\rho_{jj} = 1$.

Определим для любой матрицы (Кси) $\Xi = \Xi^{(k\times m)} = \big(\xi_{ij}\big)_{i,j}$ размера $k\times m$, элементы которой суть сл.величины, математическое ожидание

$$\mathbf{E}\,\Xi \;=\; \big(\mathbf{E}\,\xi_{ij}\big)_{i,j}\,.$$

В частности, $\mathbf{E}\,\vec{\xi} = \vec{\mu} = (\mathbf{E}\,\xi_1,\ldots,\mathbf{E}\,\xi_k)^\flat$. Легко устанавливаются простейшие алгебраические свойства такой операции.

74| **Лемма.** (?!) *Если математические ожидания компонент матриц $\Xi^{(k\times m)}$, $\Xi_1^{(k\times m)}$ конечны, то для любых констант $a, b\in\mathbb{R}^1$ и детерминированных матриц $A = A^{(l\times k)}, B = B^{(m\times n)}$, $k, l, m, n \geqslant 1$:*

- *(линейность)* $\mathbf{E}[a\Xi + b\Xi_1] = a\,\mathbf{E}\,\Xi + b\,\mathbf{E}\,\Xi_1$;
- *(однородность)* $\mathbf{E}[A\Xi B] = A(\mathbf{E}\,\Xi)B$.

Следовательно, $\mathbf{E}\left[\vec{u}^\flat\,\vec{\xi}\,\right] = \vec{u}^\flat\,\vec{\mu} = \vec{\mu}^\flat\,\vec{u}$ *для* $\forall\,\vec{u}\in\mathbb{R}^k$, *если все компоненты вектора математических ожиданий* $\vec{\mu} = \mathbf{E}\,\vec{\xi}$ *конечны.*

75| **Лемма.** *Пусть $\vec{\xi}$ — сл.вектор, у которого все дисперсии конечны: $\mathbf{D}\xi_j < \infty$, $j = \overline{1,k}$, и вектор средних $\vec{\mu} = \mathbf{E}\,\vec{\xi}$. Тогда:*

a) *матрица ковариаций* $\boldsymbol{\Sigma} = \mathbf{Cov}(\,\vec{\xi}\,) \;=\; \mathbf{E}\left[\,\vec{\xi}\,\vec{\xi}^\flat\right] - \vec{\mu}\,\vec{\mu}^\flat$;

b) *матрица ковариаций неотрицательно определена:* $\boldsymbol{\Sigma}\geqslant 0$;

c) *для любой детерминированной матрицы* $A = A^{(m\times k)}$

$$\mathbf{Cov}(A\vec{\xi}) \;=\; A\,\mathbf{Cov}(\,\vec{\xi}\,)A^\flat \;=\; A\boldsymbol{\Sigma}A^\flat.$$

В частности, дисперсия $\mathbf{D}(\vec{u}^\flat\,\vec{\xi}) \;=\; \vec{u}^\flat\boldsymbol{\Sigma}\,\vec{u}$.

⤇ a) Произведение $\vec{\xi}\,\vec{\xi}^\flat$ столбца $\vec{\xi}$ на строку $\vec{\xi}^\flat$ есть матрица размера $k\times k$, составленная из всех пар произведений компонент вектора $\vec{\xi}$. Далее можно использовать свойство 69.

c) Из свойств математического ожидания $\mathbf{E}\,A\vec{\xi} = A\vec{\mu}$. В силу a)

$$\mathbf{Cov}(A\vec{\xi}) \;=\; \mathbf{E}\left[(A\vec{\xi})(A\vec{\xi})^\flat\right] - (A\vec{\mu})(A\vec{\mu})^\flat \;=\; A\big(\mathbf{E}\left[\vec{\xi}\,\vec{\xi}^\flat\right] - \vec{\mu}\vec{\mu}^\flat\big)A^\flat$$

по свойству операции транспонирования $((A\vec{\xi})^\flat = \vec{\xi}^\flat A^\flat,\ (A\vec{\mu})^\flat = \vec{\mu}^\flat A^\flat)$ и свойству математического ожидания. Дисперсия скалярной сл.в. совпадает с её (одномерной) матрицей ковариаций: $\mathbf{D}(\vec{u}^\flat\,\vec{\xi}) = \mathbf{Cov}(\vec{u}^\flat\,\vec{\xi})$.

b) В силу c) $\vec{u}^\flat\,\boldsymbol{\Sigma}\,\vec{u} = \mathbf{D}(\vec{u}^\flat\,\vec{\xi}) \geqslant 0$ для $\forall\,\vec{u}$. ⤆

Из доказательства последнего факта можно вывести

76| **Следствие.** *Матрица ковариаций вырождена, т.е.* $\left|\mathbf{Cov}(\vec{\xi})\right| = 0$ *т. т. т. когда компоненты вектора* $\vec{\xi}$ *линейно связаны:* $\vec{u}^{\flat}\,\vec{\xi} - b = 0$ (п.н.) *с некоторым детерминированным вектором* $\vec{u}$ *и скалярной константой* $b \in \mathbb{R}^1$.

Этим объясняется, в частности, введение следующего показателя, характеризующего изменчивость сл.вектора во всём k-мерном пространстве $\mathbb{R}^k$.

Определение. Определитель матрицы ковариаций $\left|\mathbf{\Sigma}\right|$ называется *обобщённой дисперсией* случайного вектора.

Обобщённая дисперсия двумерного сл.вектора $\left|\mathbf{\Sigma}\right| = \sigma_1^2\sigma_2^2(1-\rho^2)$ принимает малые значения не только при малых значениях дисперсий компонент, но и при коэффициенте их связности $\rho \approx \pm 1$.

Другая характеристика изменчивости сл.вектора, не учитывающая связей компонент, равна сумме дисперсий: $\sigma_1^2 + \ldots + \sigma_k^2 = \mathrm{tr}(\mathbf{\Sigma})$, т.е. равна следу матрицы ковариаций.

Утверждение b) предыдущей леммы может быть обращено.

77| **Следствие.** *Симметричная матрица* Σ *является матрицей ковариаций некоторого сл.вектора т. т. т. когда она неотрицательно определена.*

$\Rrightarrow$ Нам понадобится факт из теории матриц, по которому для любой матрицы $\mathbf{\Sigma} \geqslant 0$ найдётся матрица $A\ (=\sqrt{\mathbf{\Sigma}})$ такая, что $\mathbf{\Sigma} = A^{\flat}A$ (см. стр. 534). Пусть $\vec{\xi}$ — сл.вектор с независимыми компонентами и единичными дисперсиями. Тогда матрица ковариаций $\vec{\xi}$ совпадает с единичной матрицей и в силу утверждения c) леммы 75 ковариация $\mathbf{Cov}(A^{\flat}\,\vec{\xi}) = A^{\flat}\,\dot{\mathbb{I}}\,A = \mathbf{\Sigma}$. $\Lleftarrow$

✧ **Линейная регрессия сл.величины на вектор.** Построим теперь прогноз значений одной сл.в. с помощью линейной функции от вектора сл.в. Обозначим через ζ прогнозируемую сл.в., и пусть $\vec{\xi} = (\xi_1, \ldots, \xi_k)^{\flat}$ — наблюдаемый вектор, по значениям которого мы хотим предсказать ζ. Необходимо найти вектор $\vec{c}$ и константу b, для которых достигается минимум среднеквадратического расхождения $\mathbf{E}(\zeta - \vec{c}^{\,\flat}\,\vec{\xi} - b)^2$. Так как минимум квадратической ошибки прогноза какой-либо сл.в. с помощью константы доставляет её среднее, то поставленная задача эквивалентна поиску минимума функции

$$H(\vec{c}) \;=\; \mathbf{E}\big(\zeta - \mu_\zeta - \vec{c}^{\,\flat}\,(\vec{\xi} - \vec{\mu}_\xi)\big)^2 \;=\; \mathbf{E}\big(\zeta - \vec{c}^{\,\flat}\,\vec{\xi}\,\big)^2,$$

где $\mu_\zeta = \mathbf{E}\,\zeta$, $\vec{\mu}_\xi = \mathbf{E}\,\vec{\xi}$, и, сокращая запись (переобозначив переменные), мы положили $\mu_\zeta = 0$, $\vec{\mu}_\xi = \vec{0}$.

Рассмотрим сл.вектор $\vec{\xi}_*^\flat = (\zeta, \vec{\xi}^\flat)$ и вектор констант $\vec{c}_*^\flat = (1, -\vec{c}^\flat)$ размерности $(k+1)$. По лемме 75, a) для вектора $\vec{\xi}_*$ матрица ковариаций

$$\mathbf{Cov}(\vec{\xi}_*) \;=\; \mathbf{E}\left[\begin{pmatrix}\zeta\\ \vec{\xi}\end{pmatrix}(\zeta, \vec{\xi}^\flat)\right] \;=\; \begin{pmatrix}\mathbf{E}\,\zeta^2 & \mathbf{E}\,\zeta\vec{\xi}^\flat\\ \mathbf{E}\,\vec{\xi}\zeta & \mathbf{E}\,\vec{\xi}\vec{\xi}^\flat\end{pmatrix} \;=\; \begin{pmatrix}\sigma_\zeta^2 & \vec{\sigma}^\flat_{\zeta\xi}\\ \vec{\sigma}_{\zeta\xi} & \Sigma_\xi\end{pmatrix}, \tag{26}$$

где $\sigma_\zeta^2 = \mathbf{D}\zeta$ — дисперсия ζ, $\vec{\sigma}_{\zeta\xi}$ — вектор (k-мерный) ковариаций ζ со всеми компонентами $\vec{\xi}$, $\mathbf{\Sigma} = \mathbf{Cov}(\vec{\xi})$ — матрица ковариаций $\vec{\xi}$. В этих обозначениях, с учётом утверждения леммы 75, c) и правил обращения с блочными матрицами, функция $H(\vec{c}) = \mathbf{D}\,\vec{c}_*^\flat\vec{\xi}_* = \vec{c}_*^\flat\,\mathbf{Cov}(\vec{\xi}_*)\vec{c}_* = \sigma_\zeta^2 - 2\vec{c}^\flat\vec{\sigma}_{\zeta\xi} + \vec{c}^\flat\mathbf{\Sigma}\vec{c}$.

Заметим, что, если сл.в. ζ не коррелирует со всеми компонентами $\vec{\xi}$, т.е. в ситуации, когда $\vec{\sigma}_{\zeta\xi} = \vec{0}$, функция $H = \sigma_\zeta^2 + \vec{c}^\flat\mathbf{\Sigma}\vec{c} \geqslant \sigma_\zeta^2$ ввиду неотрицательной определённости $\mathbf{\Sigma}$; поэтому минимум H достигается при $\vec{c} = \vec{0}$.

Найдём такой вектор $\vec{e}$, что сл.в. $\eta = \zeta + \vec{\xi}^\flat\vec{e}$ не коррелирует со всеми компонентами $\vec{\xi}$. Вектор ковариаций η с $\vec{\xi}$ равен

$$\mathbf{E}\left[\vec{\xi}\eta\right] \;=\; \mathbf{E}\left[\vec{\xi}\zeta\right] + \mathbf{E}\left[\vec{\xi}\vec{\xi}^\flat\vec{e}\right] \;=\; \vec{\sigma}_{\zeta\xi} + \mathbf{\Sigma}\vec{e}.$$

Таким образом, если $\mathbf{\Sigma} > 0$, то этот вектор ковариаций будет иметь нулевые компоненты при $\vec{e} = -\mathbf{\Sigma}^{-1}\vec{\sigma}_{\zeta\xi}$. Запишем функцию H (с найденным η) в виде $H(\vec{c}) = \mathbf{E}\left(\eta - (\vec{c}+\vec{e})^\flat\vec{\xi}\,\right)^2$. Так как η не коррелирует с $\vec{\xi}$, то в силу предыдущих рассуждений минимум функции H достигается при $\vec{c} + \vec{e} = \vec{0}$. Итак, в представленных обозначениях справедлива

78 **Теорема.** I) *Если* $\mathbf{\Sigma} = \mathbf{Cov}(\vec{\xi}) > 0$, *то линейная функция*

$$Z(\vec{x}) \;=\; \mu_\zeta + \vec{\beta}^\flat\big(\vec{x} - \vec{\mu}_\xi\big), \tag{27}$$

где $\vec{\beta}^\flat = \vec{\sigma}^\flat_{\zeta\xi}\,\mathbf{\Sigma}^{-1}$, $\vec{\sigma}_{\zeta\xi} = \mathbf{Cov}(\zeta, \vec{\xi})$, *даёт наилучший в среднем квадратическом линейный прогноз значения сл.в.* ζ *по реализациям сл.вектора* $\vec{\xi} = \vec{x}$.

II) *Остаточная дисперсия прогноза (дисперсия ошибки)*

$$\sigma^2_{\zeta\cdot\xi} \;=\; \mathbf{E}\big(\zeta - Z(\vec{\xi})\big)^2 \;=\; \sigma_\zeta^2 - \vec{\sigma}^\flat_{\zeta\xi}\,\mathbf{\Sigma}^{-1}\vec{\sigma}_{\zeta\xi} \;=\; \mathbf{D}\zeta - \mathbf{D}\,Z(\vec{\xi}). \tag{28}$$

⚠ Применяя известные операции к блочным матрицам, замечаем, что остаточная дисперсия с точностью до множителя $\left|\mathbf{\Sigma}_\xi\right|$ совпадает с определителем матрицы ковариаций (26). Таким образом, остаточную дисперсию можно записать в виде $\left|\,\mathbf{Cov}(\vec{\xi}_*)\right| \,/\, \left|\mathbf{\Sigma}_\xi\right| = 1/\tilde{\sigma}_\zeta^2$, где $\tilde{\sigma}_\zeta^2$ — первый диагональный элемент обратной матрицы ковариаций составного вектора $\vec{\xi}_* = (\zeta, \vec{\xi}^\flat)$.

✧ **Множественный и частный коэффициенты корреляции.** При выводе теоремы 78 было отмечено, что остаток линейного прогноза $\zeta - Z(\vec{\xi})$ не коррелирует со всеми компонентами $\vec{\xi}$. Этот факт можно интерпретировать в том духе, что после вычитания Z не осталось влияния вектора $\vec{\xi}$ на сл.в. ζ или, в совсем вольной интерпретации, что сл.функция $Z(\vec{\xi})$ несёт в себе «всю» информацию о влиянии $\vec{\xi}$ на ζ. Это даёт основание ввести следующее

Определение. Пусть $Z(\vec{x})$ — линейная среднеквадратическая регрессия сл.в. ζ на сл.вектор $\vec{\xi}$. Коэффициент корреляции

$$\rho_{\zeta\boldsymbol{\cdot}\xi} = \mathbf{Corr}(\zeta, Z(\vec{\xi}))$$

называется *множественным (сводным) коэффициентом корреляции* сл.в. ζ и сл.вектора $\vec{\xi}$.

В качестве альтернативного определения множественного коэффициента корреляции может быть взято любое из свойств (Rm_2), (Rm_5) следующей теоремы, в которой сохранены предыдущие обозначения.

79| Теорема. *Пусть $\mathbf{\Sigma} = \mathbf{Cov}(\vec{\xi}) > 0$. Множественный коэффициент корреляции $\rho_{\zeta\boldsymbol{\cdot}\xi}$ удовлетворяет следующим свойствам:*

(Rm_1) $0 \leqslant \rho_{\zeta\boldsymbol{\cdot}\xi} \leqslant 1$;

(Rm_2) $$\rho_{\zeta\boldsymbol{\cdot}\xi} = \frac{\sqrt{\vec{\sigma}^{\flat}_{\zeta\xi}\,\mathbf{\Sigma}^{-1}\vec{\sigma}_{\zeta\xi}}}{\sigma_\zeta} = \left(1 - \frac{\sigma^2_{\zeta\boldsymbol{\cdot}\xi}}{\sigma^2_\zeta}\right)^{1/2};$$

(Rm_3) $\rho_{\zeta\boldsymbol{\cdot}\xi} = 0$ *т. т. т. когда сл.в. ζ не коррелирует со всеми сл.в. ξ_j, $j = \overline{1,k}$;*

(Rm_4) $\rho_{\zeta\boldsymbol{\cdot}\xi} = 1$ *т. т. т. когда сл.в. ζ линейно зависит от компонент вектора $\vec{\xi}$, т.е. почти наверное $\zeta = Z(\vec{\xi})$;*

(Rm_5) $\rho_{\zeta\boldsymbol{\cdot}\xi}$ *есть максимальный коэффициент корреляции между сл.в. ζ и линейными функциями компонент вектора $\vec{\xi}$:*

$$\rho_{\zeta\boldsymbol{\cdot}\xi} = \max_{\vec{c}} \mathbf{Corr}(\zeta, \vec{c}^{\flat}\vec{\xi}).$$

⤇ (Rm_5) Пусть $\eta = \vec{c}^{\flat}\vec{\xi}$ — некоторая линейная комбинация $\vec{\xi}$. Заметим сначала, что по свойству обычного коэффициента корреляции этот коэффициент не изменяется при линейных преобразованиях сл.величин. Поэтому, не умаляя общности, можно предположить, что $\mathbf{E}\zeta = 0$, $\mathbf{E}\vec{\xi} = \vec{0}, \mathbf{E}\eta = 0$ и дисперсии $\mathbf{D}\zeta = 1, \mathbf{D}\eta = \vec{c}^{\flat}\mathbf{\Sigma}\vec{c} = 1$. В этих предположениях искомый коэффициент корреляции равен коэффициенту ковариации:

$$\mathbf{Corr}(\zeta,\eta)=\mathbf{E}\left[\eta\zeta\right]=\mathbf{E}\left[\vec{c}^{\flat}\vec{\xi}\zeta\right]=\vec{c}^{\flat}\,\mathbf{E}\left[\vec{\xi}\zeta\right]=\vec{c}^{\flat}\vec{\sigma}_{\zeta\xi}. \tag{29}$$

Функцию линейной регрессии с этой же целью представим в виде $Z(\vec{\xi})= \vec{\beta}^{\flat}\vec{\xi}=\sigma_z\vec{\alpha}^{\flat}\vec{\xi}$ с коэффициентом $\sigma_z^2=\mathbf{D}Z=\vec{\sigma}^{\flat}_{\zeta\xi}\boldsymbol{\Sigma}^{-1}\vec{\sigma}_{\zeta\xi}$ (т.е. $\mathbf{D}\left[\vec{\alpha}^{\flat}\vec{\xi}\right]= \mathbf{E}\left(\vec{\alpha}^{\flat}\vec{\xi}\right)^2=1$). В предположении $\mathbf{D}\zeta=1$ имеем $\mathbf{Corr}(\zeta,Z)=\vec{\alpha}^{\flat}\vec{\sigma}_{\zeta\xi}$.

По определению регрессии $\mathbf{E}\left(\zeta-\sigma_z\vec{\alpha}^{\flat}\vec{\xi}\right)^2\leqslant\mathbf{E}\left(\zeta-\sigma_z\vec{c}^{\flat}\vec{\xi}\right)^2$. Откуда, учитывая предположение о дисперсиях линейных комбинаций, получаем

$$1+\sigma_z^2-2\sigma_z\vec{\alpha}^{\flat}\vec{\sigma}_{\zeta\xi}\leqslant 1+\sigma_z^2-2\sigma_z\vec{c}^{\flat}\vec{\sigma}_{\zeta\xi},$$

что в сочетании с (29) доказывает (Rm$_5$). ⋖

80| Упр. Докажите свойства (Rm$_1$)–(Rm$_4$).

81| Примеры. 1) Пусть $\vec{\xi}\backsim\mathcal{H}\mathrm{g}_k(n;N_1,\ldots,N_k)$. Как отмечалось в примере 22, стр. 42, вектор $(\xi_1,n-\xi_1)\backsim\mathcal{H}\mathrm{g}_2(n;N_1,N-N_1)$, т.е. ξ_1 имеет гипергеометрическое (одномерное) распределение. Поэтому математическое ожидание и дисперсия компонент многомерного гипергеометрического распределения равны соответственно

$$\mathbf{E}\,\xi_j=nq_j,\qquad \mathbf{D}\xi_j=nq_j(1-q_j)\,\frac{(N-n)}{(N-1)}\,,$$

где $q_j=N_j/N$ — относительная доля j-й категории $(j=\overline{1,k})$.

Для вычисления $\mathbf{Cov}(\xi_1,\xi_2)$ достаточно рассмотреть распределение $\mathcal{H}\mathrm{g}_3(n;N_1,N_2,N-N_1-N_2)$. Мы не будем приводить здесь подробных вычислений, скажем только, что коэффициент корреляции

$$\rho_{jm}\;=\;\mathbf{Corr}(\xi_j,\xi_m)\;=\;-\sqrt{\frac{q_jq_m}{(1-q_j)(1-q_m)}}\,.$$

Заметим, что эта формула справедлива и для случая с $N=N_1+N_2$, когда в эксперименте наблюдаются только две категории: $\rho_{12}=-1$.

2) Если $(\xi_1,\ldots,\xi_k)\backsim\mathcal{M}\mathtt{ult}_k(n;q_1,\ldots,q_k)$, то, как и в предыдущем примере, $(\xi_1,n-\xi_1)\backsim\mathcal{M}\mathtt{ult}_2(n;q_1,1-q_1)$, т.е. $\xi_1\backsim\mathcal{B}\mathtt{in}(n;q_1)$. Отсюда

$$\mathbf{E}\,\xi_j\;=\;nq_j,\quad \mathbf{D}\xi_j\;=\;nq_j(1-q_j)\ \text{для}\ \forall\, j=1,\ldots,k.$$

Для вычисления ковариации достаточно рассмотреть трёхмерный вектор

$(\xi_1, \xi_2, \xi_3) \backsim \mathtt{Mult}_3(n; q_1, q_2, 1 - q_1 - q_2)$:

$$\mathbf{E}\,\xi_1\xi_2 \;=\; \sum_{m=0}^{n}\sum_{j=0}^{n-m} mj\frac{n!}{m!j!(n-m-j)!}q_1^m q_2^j (1-q_1-q_2)^{n-m-j}.$$

Поскольку при $m = 0, j = 0$ и $m = n$ слагаемые равны нулю $(0 \leqslant j \leqslant n-m = 0)$, можно ограничиться областью суммирования $1 \leqslant m \leqslant (n-1), 1 \leqslant j \leqslant (n-m)$. Произведя соответствующие сокращения и замены $(m-1) \dashrightarrow x$, $(j-1) \dashrightarrow y$, $(n-2) \dashrightarrow n'$, получаем

$$\begin{aligned}\mathbf{E}\,\xi_1\xi_2 \;&=\; q_1q_2\sum_{x=0}^{n'}\sum_{y=0}^{n'-x}\frac{n(n-1)(n')!}{x!y!(n'-x-y)!}\,q_1^x q_2^y (1-q_1-q_2)^{n'-x-y} \;=\\ &=\; q_1q_2n(n-1),\end{aligned}$$

где последнее равенство справедливо ввиду равенства 1 суммы всех вероятностей в модели $\mathtt{Mult}_3(n'; q_1, q_2, 1 - q_1 - q_2)$. Таким образом, коэффициент ковариации $\mathbf{Cov}(\xi_1, \xi_2) = q_1q_2n(n-1) - nq_1\,nq_2 = -nq_1q_2$, а коэффициент корреляции

$$\rho_{12} \;=\; \mathbf{Corr}(\xi_1, \xi_2) \;=\; -\sqrt{\frac{q_1q_2}{(1-q_1)(1-q_2)}},$$

причём, как и ожидалось, $\rho_{12} = -1$, если $q_1 + q_2 = 1$. Матрицу ковариаций $\vec{\xi} \backsim \mathtt{Mult}_k(n; q_1, \ldots, q_k)$ можно записать в виде $\mathbf{Cov}(\,\vec{\xi}\,) = n\bigl(Q - \vec{q}\,\vec{q}^{\,\flat}\bigr)$, где вектор-столбец $\vec{q} = (q_1, \ldots, q_k)^\flat$, Q — диагональная матрица с элементами $\vec{q}$ на диагонали, $\vec{q}\,\vec{q}^{\,\flat}$ — матрица всех пáрных произведений q_jq_m.

Произведя несложные арифметические преобразования, получаем из предыдущего, что наилучший линейный прогноз значения одной компоненты, скажем ξ_1, по значению $\xi_2 = x_2$ другой компоненты мультиномиального (и многомерного гипергеометрического) сл.вектора $(\xi_1, \xi_2, \ldots, \xi_k)$ равен

$$x_1^* = \frac{q_1}{1-q_2}(n - x_2). \tag{30}$$

Здесь есть проблема применимости, связанная с дискретностью ξ_1. Рассмотрим конкретный пример. Предположим, что при одном выстреле снайпер с вероятностью 0.90 попадает в центральное «яблочко» и с вероятностью 0.01 промахивается по мишени, остальные 9% вероятности приходятся на кольцевые участки мишени. После серии в $n = 10$ выстрелов жюри сообщило, что ровно при одном выстреле снайпер промахнулся по мишени, а для остальных выстрелов требуется дополнительное исследование. Во время ожидания результатов этого исследования может возникнуть желание заключить пари на количество попаданий в «яблочко». Линейный прогноз даёт $x_1 = 8.18$. Такой ответ, естественно, никого не устроит. Чтобы дать приемлемый ответ в виде целого числа, не надо торопиться округлять, прогнозируя $x_1 = 8$. Здесь придётся пересмотреть саму постановку задачи о наилучшем прогнозе.

В данной ситуации штраф за ошибочный прогноз редко представляет собой квадрат разности $(d-r)^2$, где d — выдвинутый прогноз, а r — действительный результат. В простейшем случае штрафная функция равна 1 (потерял свою ставку), если $d \neq r$, и равна -1 (приобрел ставку оппонента), если $d = r$. При такой штрафной функции наилучший прогноз равен исходу с наибольшей условной вероятностью. Для нашего примера это $x_1^* = 9$. Условная вероятность такого исхода равна всего $0.424 < 0.5$.

з) Для сл.вектора $(\xi_1, \xi_2) \backsim \mathcal{D}\mathtt{ir}_2(a_1, a_2, a_3)$ (см. (22), стр. 71), используя определение бета-функции и формулу понижения $\mathrm{B}(x+1, y, z)/\mathrm{B}(x, y, z) = x/(x+y+z)$, получаем смешанный момент (при $m = 1, r = 1$)

$$\mathbf{E}\,\xi_1^m \xi_2^r \;=\; \int_0^1 dx_1 \int_0^{1-x_1} x_1^m x_2^r \frac{x_1^{a_1-1} x_2^{a_2-1}}{\mathrm{B}(a_1, a_2, a_3)} (1-(x_1+x_2))^{a_3-1}\, dx_2 \;=$$

$$= \left.\frac{\mathrm{B}(m+a_1, r+a_2, a_3)}{\mathrm{B}(a_1, a_2, a_3)}\right|_{\substack{m=1,\\ r=1}} \;=\; \frac{a_1 a_2}{(a_1+a_2+a_3+1)(a_1+a_2+a_3)}\,.$$

Пусть $\vec{\xi} \backsim \mathcal{D}\mathtt{ir}_k(a_1, \ldots, a_{k+1})$. Положим $\boldsymbol{a} = \sum_1^{k+1} a_j,\ q_j = a_j/\boldsymbol{a},\ j = \overline{1, k}$. Тогда смешанный момент $\mathbf{E}[\xi_1 \xi_2] = a_1 a_2/(\boldsymbol{a}(\boldsymbol{a}+1))$, математическое ожидание $\mathbf{E}\,\xi_1 = \mathrm{B}(1+a_1, a_2, a_3)/\mathrm{B}(a_1, a_2, a_3) = a_1/\boldsymbol{a} = q_1$, ковариация $\mathbf{E}[\xi_1\xi_2] - \mathbf{E}\,\xi_1\ \mathbf{E}\,\xi_2 = -a_1 a_2/(\boldsymbol{a}^2(\boldsymbol{a}+1)) = -q_1 q_2/(\boldsymbol{a}+1)$ и, наконец, коэффициент корреляции

$$\mathbf{Corr}\,(\xi_1, \xi_2) \;=\; -\sqrt{\frac{a_1 a_2}{(\boldsymbol{a}-a_1)(\boldsymbol{a}-a_2)}} \;=\; -\sqrt{\frac{q_1 q_2}{(1-q_1)(1-q_2)}}\,.$$

Сравните с гипергеометрической и мультиномиальной моделями. ◎

Свойство (Rm$_5$) может быть положено в основу определения коэффициента корреляции со сл.вектором $\vec{\xi}$ не только одной сл.в. ζ, но и вектора сл.в. $\vec{\zeta}$ (см. [1]). Этот коэффициент применяется в тех случаях, когда требуется небольшим набором характеристик описать степень связности двух групп случайных признаков, например, связь успеваемости по различным школьным предметам с рядом социально-экономических признаков.

Определение. Максимальный коэффициент корреляции

$$\max_{\vec{a},\vec{c}} \mathbf{Corr}(\vec{a}^{\flat}\vec{\xi}, \vec{c}^{\flat}\vec{\zeta})$$

между линейными комбинациями сл.векторов $\vec{\zeta}, \vec{\xi}$ называется *каноническим коэффициентом корреляции* между $\vec{\zeta}$ и $\vec{\xi}$.

82| Упр. Покажите, что каноническая корреляция не меньше нуля.

Функция $\mathbf{Corr}(\vec{a}^{\flat}\vec{\xi}, \vec{c}^{\flat}\vec{\zeta})$ имеет несколько неотрицательных локальных максимумов — все они называются каноническими коэффициентами корреляции и используются для анализа связей рассматриваемых векторов. При этом значения компонент векторов $\vec{a}, \vec{c}$, на которых достигается тот или иной максимум, а точнее, их относительная величина, показывают степень влияния соответствующих компонент сл.векторов $\vec{\zeta}, \vec{\xi}$, участвующих в формировании корреляционной связи.

Пусть k_1 — размерность вектора $\vec{\zeta}$, k_2 — размерность $\vec{\xi}$. Для определённости будем считать, что $k_1 \leqslant k_2$. Предположим, что матрицы корреляций $\mathbf{\Sigma}_{11} = \mathbf{Corr}(\vec{\zeta}), \mathbf{\Sigma}_{22} = \mathbf{Corr}(\vec{\xi})$ невырожденные, а матрица взаимных корреляций $\mathbf{\Sigma}_{12} = \mathbf{Corr}(\vec{\zeta}, \vec{\xi}) = \big(\mathbf{Corr}(\zeta_i, \xi_j)\big)_{i,j}$ имеет полный ранг k_1. Тогда квадраты канонических коэффициентов корреляции равны собственным числам матрицы $\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{12}^{\flat}$. Соответствующие собственные векторы этой матрицы дают коэффициенты $\vec{c}^{\flat}\vec{\zeta}$ линейной комбинации вектора $\vec{\zeta}$. Коэффициенты линейной комбинации $\vec{a}^{\flat}\vec{\xi}$ находятся по формуле $\vec{a} = \mathbf{\Sigma}_{12}^{\flat}\mathbf{\Sigma}_{11}^{-1}\vec{c}$.

Факт некоррелированности остатков регрессии от регрессоров может быть положен в основу определения ещё одной меры зависимости. Пусть ζ_1, ζ_2 — некоторые сл.в., $\vec{\xi}$ — сл.вектор.

Определение. *Частным коэффициентом корреляции* между ζ_1, ζ_2 за вычетом влияния $\vec{\xi}$ называется величина

$$\rho_{1,2\cdot\xi} = \mathbf{Corr}(\zeta_1, \zeta_2 \mid \vec{\xi}) := \mathbf{Corr}\,\big(\zeta_1 - Z_1(\vec{\xi}), \zeta_2 - Z_2(\vec{\xi})\big),$$

равная коэффициенту корреляции между остатками линейной среднеквадратической регрессии ζ_j, $j = 1, 2$, на сл.вектор $\vec{\xi}$.

⚠ Предложенное здесь обозначение для частного коэффициента корреляции не случайно. Для нормального вектора $(\zeta_1, \zeta_2, \vec{\xi})$ этот коэффициент равен коэффициенту корреляции условного распределения (ζ_1, ζ_2) при фиксированном значении $\vec{\xi} = \vec{x}$ (независимо от $\vec{x}$). Отсюда и практическая интерпретация этого коэффициента как условного коэффициента корреляции между двумя сл.в. при фиксированных значениях остальных компонент вектора.

Не умаляя общности, можно считать математические ожидания всех сл.в. равными нулю. В соответствии с теоремой 78 для $j = 1, 2$,

$$\zeta_j - Z_j(\vec{\xi}) \;=\; \zeta_j - \vec{\sigma}^{\flat}_{j\boldsymbol{\cdot}\xi}\, \mathbf{\Sigma}^{-1}\vec{\xi},$$

где $\mathbf{\Sigma} = \mathbf{Cov}(\vec{\xi})$, $\vec{\sigma}_{j\boldsymbol{\cdot}\xi} = \mathbf{Cov}(\zeta_j, \vec{\xi})$; дисперсии остатков $\mathbf{D}\big(\zeta_j - Z_j\big) = \mathbf{D}\zeta_j - - \vec{\sigma}^{\flat}_{j\boldsymbol{\cdot}\xi}\, \mathbf{\Sigma}^{-1}\vec{\sigma}_{j\boldsymbol{\cdot}\xi}$. Таким образом, в силу отмеченной некоррелированности остатков от сл.вектора $\vec{\xi}$ имеем

$$\mathbf{E}\left[\zeta_1 - Z_1(\vec{\xi}), \zeta_2 - Z_2(\vec{\xi})\right] \;=\; \mathbf{E}\left[\zeta_1\big(\zeta_2 - Z_2(\vec{\xi})\big)\right] \;=\; \rho_{12}\sigma_1\sigma_2 - \vec{\sigma}^{\flat}_{2\boldsymbol{\cdot}\xi}\, \mathbf{\Sigma}^{-1}\vec{\sigma}_{1\boldsymbol{\cdot}\xi},$$

где $\rho_{12} = \mathbf{Corr}(\zeta_1, \zeta_2)$, $\sigma_j^2 = \mathbf{D}\zeta_j$, $j = 1, 2$. Учитывая связь между коэффициентами ковариации и корреляции, приходим к следующему утверждению.

83] **Теорема.** *Пусть* $\rho_{12} = \mathbf{Corr}(\zeta_1, \zeta_2)$, $\vec{\rho}_{j\boldsymbol{\cdot}\xi} = \mathbf{Corr}(\zeta_j, \vec{\xi})$ *— вектор корреляций* ζ_j *с компонентами вектора* $\vec{\xi}$, $\mathcal{P} = \mathbf{Corr}(\vec{\xi})$ *— матрица коэффициентов корреляции* $\vec{\xi}$. *Тогда частный коэффициент корреляции*

$$\mathbf{Corr}(\zeta_1, \zeta_2 \,|\, \vec{\xi}) \;=\; \frac{\rho_{12} - \vec{\rho}^{\flat}_{1\boldsymbol{\cdot}\xi}\mathcal{P}^{-1}\vec{\rho}_{2\boldsymbol{\cdot}\xi}}{\sqrt{(1 - \vec{\rho}^{\flat}_{1\boldsymbol{\cdot}\xi}\mathcal{P}^{-1}\vec{\rho}_{1\boldsymbol{\cdot}\xi})(1 - \vec{\rho}^{\flat}_{2\boldsymbol{\cdot}\xi}\mathcal{P}^{-1}\vec{\rho}_{2\boldsymbol{\cdot}\xi})}}.$$

84] **Следствие.** *Пусть* $\rho_{jm} = \mathbf{Corr}(\zeta_j, \zeta_m)$, $j, m = 1, 2, 3$. *Тогда частный коэффициент корреляции между* ζ_1, ζ_2 *за вычетом влияния сл.в.* ζ_3 *равен*

$$\mathbf{Corr}(\zeta_1, \zeta_2 \,|\, \zeta_3) \;=\; \frac{\rho_{12} - \rho_{13}\rho_{23}}{\sqrt{(1 - \rho_{13}^2)(1 - \rho_{23}^2)}}.$$

85] Пример. В конце XIX в. одной страховой компанией была обнаружена высокая положительная корреляционная связь между ущербом от пожара (сл.в. ζ_1) и фактом приезда пожарной команды (сл.в. $\zeta_2 = 0, 1$, где 1 — участие пожарной команды в процессе тушения). Это было воспринято как сигнал к отказу от вызова пожарной команды и к необходимости расследования деятельности всех этих команд. Ларчик открывался просто: обе эти величины имеют высокую положительную корреляционную связь с третьей величиной ζ_3, равной стоимости объекта до пожара. В качестве иллюстрации (не обязательно соответствующей описываемому историческому примеру) предположим, что $\rho_{12} = 0.5$, $\rho_{13} = \rho_{23} = 0.8$. Тогда при одинаковой стоимости объекта до пожара корреляция между ущербом и фактом участия в тушении пожарной команды будет отрицательной: $\mathbf{Corr}(\zeta_1, \zeta_2 \mid \zeta_3) = -0.23$. ⊚

§ 7. Эллипсоид рассеяния

Наряду с числовыми характеристиками положения $\vec{\mu} = \mathbf{E}\vec{\xi}$ и изменчивости $\mathbf{\Sigma} = \mathbf{Cov}(\vec{\xi})$ сл.вектора $\vec{\xi}$ часто рассматривают «производную» от них геометрическую характеристику — так называемый эллипсоид рассеяния. Кроме области изменения сл.в. (в основном), эллипсоид рассеяния показывает характер зависимости сл.в. Из всего разнообразия геометрических фигур эллипсоид выбран по следующим причинам. Во-первых, эллипсоид имеет удобное аналитическое представление (сравните, например, с аналитическим представлением прямоугольника, у которого грани не параллельны координатным плоскостям). Во-вторых, реальные данные показывают, что основная масса реализаций сл.вектора визуально располагается внутри фигуры, подобной эллипсоиду. Наконец, у наиболее популярного многомерного нормального распределения линии постоянства ф.пл. образованы именно эллипсоидами.

Определение. Пусть $\vec{\mu} = \mathbf{E}\vec{\xi}$ — вектор математических ожиданий k-мерного сл.вектора $\vec{\xi}$, $\mathbf{\Sigma} = \mathbf{Cov}(\vec{\xi})$ — его матрица ковариаций, $\mathbf{\Sigma} > 0$. Эллипсоид

$$\{\, \vec{x} \in \mathbb{R}^k : \ (\vec{x} - \vec{\mu})^\flat \, \mathbf{\Sigma}^{-1} \, (\vec{x} - \vec{\mu}) \leqslant k + 2 \,\} \tag{31}$$

называется *эллипсоидом рассеяния* сл.вектора $\vec{\xi}$.

Причины выбора конкретных параметров эллипсоида объясняет

86| **Теорема.** *Пусть* $\mathbf{\Sigma} > 0$, *тогда эллипсоид* (31) *— единственный эллипсоид, равномерное распределение внутри которого имеет идентичные с* $\vec{\xi}$ *вектор математических ожиданий* $\vec{\mu}$ *и матрицу ковариаций* $\mathbf{\Sigma}$.

⇛ Пусть $\mathfrak{E} : (\vec{x} - \vec{a})^\flat \, B \, (\vec{x} - \vec{a}) \leqslant c^2$ — некоторый эллипсоид с положительно определённой матрицей $B > 0$. Плотность равномерного распределения в эллипсоиде $f(\vec{x}) = 1/V$ при $\vec{x} \in \mathfrak{E}$, где V — объём эллипсоида. Как известно, для матрицы $B > 0$ существует ортогональная матрица Q, приводящая $\mathfrak{E}$ к главным осям: $Q^\flat Q = QQ^\flat = \dot{\mathbb{I}}$, $QBQ^\flat = \Lambda^2$, где Λ^2 — диагональная матрица с положительными элементами $\lambda_1^2, \ldots, \lambda_k^2$, равными собственным числам B. Легко проверить, что $(\Lambda^2)^{-1} = (\Lambda^{-1})^2 = QB^{-1}Q^\flat$.

Преобразование $(\vec{x} - \vec{a}) = Q^\flat \Lambda^{-1} \vec{y}$ взаимно однозначно переводит эллипсоид $\mathfrak{E}$ в шар $\mathfrak{O} : \vec{y}^\flat \vec{y} \leqslant c^2$. Поскольку якобиан этого преобразования равен $\big|\big|Q^\flat \Lambda^{-1}\big|\big| = 1/\sqrt{|B|}$, то объём

$$V \;=\; \int_{\mathfrak{E}} d\vec{x} \;=\; \frac{1}{\sqrt{|B|}} \int_{\mathfrak{O}} d\vec{y}.$$

Точное значение последнего интеграла нам здесь не понадобится. Для особо любознательных скажем, что объём k-мерного шара радиуса c равен $c^k\sqrt{\pi^k}/\Gamma\big(1 + \frac{k}{2}\big)$. Найдём преобразование этого интеграла с помощью сферической замены переменных, т.е. для $r > 0$ и $(\theta, \alpha_1, \ldots, \alpha_{k-2}) \in \Theta$, где

$$\Theta := \{\theta \in [0; 2\pi),\, \alpha_j \in (0; \pi],\; j = 1, \ldots, k-2\},$$

положим

$$\begin{array}{rllll}
y_1 = & r\sin\theta & \cdot\sin\alpha_1 & \ldots & \cdot\sin\alpha_{k-2}, \\
y_2 = & r\cos\theta & \cdot\sin\alpha_1 & \ldots & \cdot\sin\alpha_{k-2}, \\
y_3 = & & r\cos\alpha_1 & \ldots & \cdot\sin\alpha_{k-2}, \\
 & & \ldots & & \\
y_k = & & & & r\cos\alpha_{k-2}.
\end{array}$$

Якобиан сферической замены равен $r^{k-1}J{:=}r^{k-1}\sin\alpha_1\sin^2\alpha_2\cdots\sin^{k-2}\alpha_{k-2}$. Легко понять, что $y_1^2 + \ldots + y_k^2 = r^2$. Поэтому, обозначив $d\vec{\alpha} = d\alpha_1 \ldots d\alpha_{k-2}$, получим

$$V \;=\; \frac{1}{\sqrt{|B|}} \int\!\!\cdots\!\!\int\limits_{\Theta} J\, d\theta d\vec{\alpha} \cdot \int_0^c r^{k-1}\, dr \;=\; \frac{c^k}{k\sqrt{|B|}} \int\!\!\cdots\!\!\int\limits_{\Theta} J\, d\theta d\vec{\alpha}.$$

Аналогично, $\int_{\mathfrak{E}}(\vec{x} - \vec{a})d\vec{x} \;=\; Q^\flat\Lambda^{-1}\int_{\mathfrak{O}}\vec{y}d\vec{y}/\sqrt{|B|} \;=\; \vec{0}$ в силу симметрии области интегрирования. То есть математическое ожидание вектора с равномерным распределением внутри эллипсоида $\mathfrak{E}$ равно $\vec{a}$.

Ковариация (в матричной форме записи интегралов) вычисляется как

$$\frac{1}{V}\int_{\mathfrak{E}}(\vec{x} - \vec{a})(\vec{x} - \vec{a})^\flat\, d\vec{x} \;=\; \frac{1}{V\sqrt{|B|}}\, Q^\flat\Lambda^{-1}\Big(\int_{\mathfrak{O}} \vec{y}\vec{y}^\flat d\vec{y}\Big)\Lambda^{-1}Q.$$

Легко понять, что в силу симметрии интегралы от смешанных произведений $y_j y_m$, $j \neq m$, здесь равны нулю. Также в силу симметрии

$$\int_{\mathfrak{O}} y_1^2 d\vec{y} \;=\; \frac{1}{k}\int_{\mathfrak{O}}(y_1^2 + \ldots + y_k^2)d\vec{y}.$$

Снова применяя сферическую замену, получаем

$$\int\limits_{\mathfrak{O}} y_1^2 d\vec{y} \;=\; \frac{1}{k}\int\limits_{\Theta}\!\!\cdots\!\!\int J\,d\theta d\vec{\alpha}\cdot\int\limits_0^c r^2 r^{k-1}\,dr \;=\; \frac{1}{k}\frac{k\sqrt{|B|}\,c^{k+2}}{c^k(k+2)}V.$$

Поэтому матрица ковариаций равномерного распределения совпадает с

$$\frac{c^2}{k+2}Q^\flat\Lambda^{-1}\dot{\mathbb{I}}\Lambda^{-1}Q \;=\; \frac{c^2}{k+2}B^{-1}.$$

Теорема доказана. $\Lleftarrow$

Из доказательства теоремы видно, что определитель матрицы ковариаций сл.вектора пропорционален квадрату объёма эллипсоида рассеяния. Отсюда

Определение. *Обобщённой дисперсией* сл.вектора $\vec{\xi}$ называется определитель матрицы ковариаций $\mathfrak{S}^2(\vec{\xi}) = \left|\,\mathbf{Cov}(\vec{\xi})\right|$.

В связи с этим см. также определение на стр. 77.

В одномерном случае обобщённая дисперсия $\mathfrak{S}^2(\xi) = \sigma^2$, т.е. равна дисперсии ξ; эллипсоид рассеяния совпадает с отрезком $[\mu - \sqrt{3}\sigma; \mu + \sqrt{3}\sigma]$.

В двумерном случае эллипсоид рассеяния можно представить в виде

$$Y^2 - 2\rho YX + X^2 \leqslant 4(1-\rho^2),$$

где ρ — коэффициент корреляции между сл.в. ξ_1, ξ_2, переменные $X = (x_1 - \mu_1)/\sigma_1$, $Y = (x_2 - \mu_2)/\sigma_2$. Обобщённая дисперсия $\mathfrak{S}^2(\xi_1, \xi_2) = \sigma_1^2\sigma_2^2(1-\rho^2)$. Ясно, что оси эллипса рассеяния параллельны осям координат т. т. т. когда коэффициент корреляции $\rho = 0$.

По виду двумерного эллипсоида рассеяния легко восстановить линии регрессии. Если неявное выражение $Y = Y(X)$, описывающее границу эллипсоида, продифференцировать по X, получим уравнение

$$2Y\,Y' - 2\rho X\,Y' - 2\rho Y + 2X = 0.$$

Для касательных к границе, параллельных оси OX, производная равна нулю. Следовательно, точки касания удовлетворяют уравнению $X = \rho Y$ или $(x_1 - \mu_1)/\sigma_1 = \rho(x_2 - \mu_2)/\sigma_2$, что совпадает с уравнением регрессии ξ_1 на ξ_2. Другими словами, если значения сл.в. ξ_1 откладываются на оси абсцисс (OX), а значения ξ_2 — на оси ординат (OY), то линия, соединяющая точки касания двух касательных к эллипсу, параллельных оси абсцисс, и будет линией регрессии ξ_1 на ξ_2 (X на Y). Аналогично для регрессии ξ_2 на ξ_1.

§ 8. Метод главных компонент

Во многих исследовательских работах число показателей, которые требуется обработать, слишком велико. В таких изысканиях интерес зачастую представляют не значения самих показателей, а их разброс между изучаемыми объектами. Так, например, антрополог-физиономист проводит измерения нескольких десятков характеристик, таких как длина уха, ширина уха, длина лица, ширина лица и т.п. Его может интересовать описание и анализ различий индивидов по подобного рода характеристикам (скажем, для дальнейшего изучения причин этих различий). Поэтому, естественно, вначале он должен будет отбросить характеристики или комбинации характеристик с малыми различиями, а затем выделить характеристики или комбинации характеристик, обобщающие разделительные свойства индивидов. Для этих целей и служит метод главных компонент.

Определение. Пусть $\vec{\xi}$ – k-мерный сл.вектор с вектором средних $\vec{\mu}$. Вектор $\vec{\zeta} = (\zeta_1, \ldots, \zeta_k)$ называется вектором *главных компонент* для $\vec{\xi}$, если:

(✓) все компоненты $\vec{\zeta}$ получены как нормированные линейные комбинации центрированного вектора $\vec{\xi}$:

$$\zeta_j = \vec{c}_j^{\,\flat}(\vec{\xi} - \vec{\mu}) = c_{j1}(\xi_1 - \mu_1) + \ldots + c_{jk}(\xi_k - \mu_k),$$
$$\vec{c}_j^{\,\flat}\,\vec{c}_j = c_{j1}^2 + \ldots + c_{jk}^2 = 1, \quad j = 1, \ldots, k;$$

(✓) дисперсия 1-й компоненты ζ_1 максимальна среди всех нормированных линейных комбинаций $\vec{\xi}$;

(✓) дисперсия j-й компоненты ζ_j максимальна среди всех нормированных линейных комбинаций вектора $\vec{\xi}$, не коррелирующих с компонентами $\zeta_1, \ldots, \zeta_{j-1}$, $j = 2, \ldots, k$.

Решение проблемы главных компонент основано на решении характеристического уравнения $\mathbf{\Sigma}\vec{c} = \gamma\vec{c}$. Напомним некоторые свойства этого решения. Для любой k-мерной неотрицательно определённой матрицы $\mathbf{\Sigma} \geqslant 0$ существует k собственных чисел $\gamma_1 \geqslant \ldots \geqslant \gamma_k \geqslant 0$, причём $\gamma_k = 0$ т. т. т. когда $|\mathbf{\Sigma}| = 0$. Пусть матрица $Q = (\vec{c}_1, \ldots, \vec{c}_k)$, т.е. её j-й столбец образован нормированным собственным вектором $\vec{c}_j$, отвечающим собственному числу γ_j,

тогда

$$\mathbf{\Sigma} Q = Q\Gamma$$

с диагональной матрицей $\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_k)$. Собственные векторы, отвечающие различным собственным числам, ортогональны. Поэтому в случае, когда все собственные числа различны, матрица Q будет ортогональной: $QQ^\flat = Q^\flat Q = \dot{\mathbb{I}}$.

При доказательстве следующей теоремы некоторое затруднение возникает в случае, когда часть собственных чисел совпадает. Обходя это неудобство, будем считать, что все собственные числа матрицы ковариаций $\mathbf{\Sigma}$ различны.

87| **Теорема.** *Пусть* $\mathbf{\Sigma}$ *— матрица ковариаций* $\vec{\xi}$. *Тогда векторы* $\vec{c}_j$, *определяющие главные компоненты, совпадают с нормированными собственными векторами матрицы* $\mathbf{\Sigma}$:

$$\mathbf{\Sigma}\vec{c}_j = \gamma_j \vec{c}_j\,,$$
$$\|\vec{c}_j\|^2 = \vec{c}_j^{\,\flat}\,\vec{c}_j = 1, \quad j = 1, \ldots, k,$$

причём собственные числа выбраны так, что $\gamma_1 > \cdots > \gamma_k > 0$.

⊫ Во-первых, как всегда, когда дело касается дисперсий сл.в., можно предположить, что математические ожидания равны нулю. Пусть столбцы матрицы Q образованы нормированными решениями характеристического уравнения с упорядоченными по убыванию собственными числами. Тогда для сл.вектора $\vec{\zeta} = Q^\flat \vec{\xi}$ матрица ковариаций

$$\mathbf{Cov}(\vec{\zeta}) \;=\; Q^\flat\,\mathbf{Cov}(\vec{\xi})Q \;=\; Q^\flat \mathbf{\Sigma} Q \;=\; Q^\flat Q\Gamma \;=\; \Gamma$$

по свойству ковариации линейного преобразования. Таким образом, все компоненты вектора $\vec{\zeta}$ не коррелируют, а дисперсии $\mathbf{D}\zeta_j = \gamma_j,\ j = 1, \ldots, k$.

Произвольный (нормированный) вектор $\vec{a}$ можно представить в виде линейной комбинации собственных векторов:

$$\vec{a} \;=\; b_1\vec{c}_1 + \ldots + b_k\vec{c}_k \;=\; Q\vec{b},$$

где вектор $\vec{b} = Q^\flat \vec{a}$; причём квадрат нормы $\|\vec{b}\|^2 = \vec{b}^\flat\,\vec{b} = \vec{a}^\flat\,QQ^\flat\vec{a} = \vec{a}^\flat\,\dot{\mathbb{I}}\vec{a} = 1$. Следовательно, дисперсия сл.в. $\vec{a}^\flat\,\vec{\xi}$

$$\mathbf{D}(\vec{a}^\flat\,\vec{\xi}) \;=\; \vec{a}^\flat\mathbf{\Sigma}\,\vec{a} \;=\; \vec{b}^\flat\,Q^\flat\mathbf{\Sigma} Q\vec{b} \;=\; \vec{b}^\flat\,\Gamma\vec{b} \;=\; b_1^2\gamma_1 + \ldots + b_k^2\gamma_k \;\leqslant\; \gamma_1\|\vec{b}\|^2,$$

поскольку все собственные числа $\gamma_2, \ldots, \gamma_k < \gamma_1$. Последнее неравенство будет строгим т. т. т. когда коэффициент $|b_1| \neq 1$ (< 1). Это доказывает, что сл.в. $\zeta_1 = \vec{c}_1^{\,\flat}\vec{\xi}$ есть первая главная компонента.

Пусть теперь вектор $\vec{a}$ выбран так, что сл.в. $\vec{a}^{\flat}\vec{\xi}$ не коррелирует с $\zeta_n = \vec{c}_n^{\,\flat}\vec{\xi},\ n \leqslant j-1$. Равенство нулю корреляции означает, что

$$0 \;=\; \mathbf{Cov}(\vec{a}^{\flat}\vec{\xi}, \zeta_n) \;=\; \mathbf{E}(\vec{a}^{\flat}\vec{\xi}\vec{\xi}^{\flat}\vec{c}_n) \;=\; \vec{a}^{\flat}\mathbf{\Sigma}\vec{c}_n \;=\; \gamma_n\vec{a}^{\flat}\vec{c}_n,$$

т.е. $\vec{a} \perp \vec{c}_n,\ n < j$. Поэтому $\vec{a} = b_j\vec{c}_j + \ldots + b_k\vec{c}_k$ и, как и выше,

$$\mathbf{D}(\vec{a}^{\flat}\vec{\xi}) \;=\; b_j^2\gamma_j + \ldots + b_k^2\gamma_k \;\leqslant\; \gamma_j,$$

т.е. дисперсия ζ_j максимальна среди всех нормированных комбинаций вектора $\vec{\xi}$, не коррелирующих с первыми $j-1$ главными компонентами. ⫤

► **Нормирование измеряемых характеристик.** Метод главных компонент весьма чувствителен к изменению единиц измерения характеристик исследуемого объекта. Поэтому применение этого метода будет наиболее плодотворным в ситуациях, когда показатели $\xi_1, \ldots, \xi_k$ имеют общую физическую природу и измеряются в одних и тех же единицах. В противном случае обычно прибегают к предварительной нормировке показателей и в качестве нормирующей константы выбирают стандартное отклонение: $\xi_j^* = (\xi_j - \mu_j)/(\sigma_j)$. В этом случае ковариационная матрица $\mathbf{\Sigma}$ совпадает с корреляционной матрицей, т.е. метод главных компонент опирается на решение характеристического уравнения относительно матрицы корреляций, а не ковариаций. Заметим, кстати, что тогда сумма дисперсий главных компонент равна k, так как на диагонали матрицы корреляций стоят единицы (см. далее свойство 5)).

► **Свойства главных компонент.** Пусть $\vec{\zeta} = Q^{\flat}\vec{\xi}$ — вектор главных компонент $\vec{\xi}$. Тогда имеют место следующие свойства.

1) Матрица Q ортогональна: $Q^{\flat}Q = QQ^{\flat} = \dot{\mathbb{I}}$.

2) Матрица ковариаций $\vec{\zeta}$ диагональна: $\mathbf{Cov}(\vec{\zeta}) = \mathrm{diag}(\gamma_1, \ldots, \gamma_k)$.

3) Дисперсия j-й главной компоненты равна j-му собственному числу матрицы $\mathbf{\Sigma}$: $\mathbf{D}\zeta_j = \gamma_j$, $j = 1, \ldots, k$.

4) Переход к главным компонентам сохраняет обобщённую дисперсию:

$$\mathfrak{S}^2(\vec{\xi}) \;=\; |\mathbf{\Sigma}| \;=\; \mathfrak{S}^2(\vec{\zeta}) \;=\; \textstyle\prod_1^k \gamma_j.$$

5) Переход к главным компонентам сохраняет сумму дисперсий:

$$\textstyle\sum_1^k \mathbf{D}\xi_j \;=\; \sum_1^k \mathbf{D}\zeta_j \;=\; \sum_1^k \gamma_j.$$

6) Переход к главным компонентам представляет собой поворот осей системы координат параллельно осям эллипсоида рассеяния $\vec{\xi}$.

Доказательства этих свойств сразу следуют либо из утверждения предыдущей теоремы, либо из её доказательства. Например,

$$\sum_1^k \mathbf{D}\xi_j = \mathbf{E}\left[\vec{\xi}^\flat \vec{\xi}\right] = \mathbf{E}\left[\vec{\zeta}^\flat Q^\flat Q \vec{\zeta}\right] = \mathbf{E}\left[\vec{\zeta}^\flat \vec{\zeta}\right] = \sum_1^k \gamma_j.$$

Свойство 6) есть простая констатация известного свойства, что переход к канонической форме записи эллипсоида определяется собственными векторами матрицы, задающей этот эллипсоид.

► **Понижение размерности наблюдаемого вектора**. Свойство 5) лежит в основе применения метода главных компонент на практике. Величина

$$\mathfrak{m}_q = \frac{\gamma_1 + \ldots + \gamma_q}{\gamma_1 + \ldots + \gamma_k} \cdot 100\%$$

показывает ту долю разброса наблюдений, которая может быть объяснена факторами, определяющими первые q главных компонент. Если $\mathfrak{m}_q$ достаточно велико, то оставшиеся $k-q$ главных компонент могут быть исключены из дальнейшего рассмотрения. Точное значение понятия «достаточно велико» конкретизируется в соответствии с рассматриваемой исследовательской задачей. Можно предложить следующую градацию для описания информативности главных компонент.

Если $\mathfrak{m}_q > 75\%$, первые q главных компонент содержат *основную массу* информации о разбросе наблюдаемых признаков, если $\mathfrak{m}_q > 90\%$, то *почти всю* информацию, и, наконец, если $\mathfrak{m}_q > 95\%$, то *всю* информацию. Приведённые здесь проценты, конечно, условны и не обязательны к исполнению.

Заметим, что, обладая информацией о значениях $\vec{\zeta} = \vec{z}$ всех главных компонент, можно восстановить исходные значения вектора $\vec{\xi}$ по формуле $\vec{x} = Q\vec{z}$. Восстановить значения $\vec{\xi}$ по части главных компонент уже не удастся. Можно только построить наилучшее приближение для этих значений.

88| **Лемма.** (?!) 1) *Наилучшая линейная аппроксимация вектора* $\vec{\xi}$ *по значениям первых* q *главных компонент* $\vec{\zeta}^{(q)}$ *достигается при линейном преобразовании* $\vec{\xi}^* = Q^{(q)}\vec{\zeta}^{(q)}$ *с матрицей* $Q^{(q)}$ *размера* $k \times q$, *построенной на первых* q *столбцах матрицы* Q.

2) *Относительная суммарная ошибка этой аппроксимации*

$$\frac{\sum_1^k \mathbf{E}(\xi_j - \xi_j^*)^2}{\sum_1^k \sigma_j^2} = 1 - \frac{1}{100}\mathfrak{m}_q.$$

§ 9. Условное распределение

Строгому определению условного математического ожидания и связанному с ним понятию условного распределения для общей вероятностной модели будет посвящён специальный раздел (глава III). Здесь же мы введём эти понятия для дискретных и абсолютно непрерывных моделей. В дальнейшем знаки $\mathbf{P}, \mathbf{E}$ будут использоваться для обозначения вероятности и математического ожидания относительно распределения на исходном вероятностном пространстве, знаки P, E будут обозначать соответствующие операции в пространстве значений сл.величин $(\mathbb{R}^k)$. Такое разграничение используется только для того, чтобы обратить внимание на способ вычисления соответствующей характеристики, например: для любой функции h от сл.в. ξ математическое ожидание

$$\mathbf{E}\big[h(\xi)\big] \;=\; \int_\Omega h(\xi(\omega))\;\mathbf{P}(d\omega) \;=\; \int_\mathbb{R} h(x)\;\mathrm{P}_\xi(dx) \;=\; \mathrm{E}\big[h(\xi)\big].$$

✧ **Дискретные модели**. Пусть $p(x,y)$, $x \in \mathscr{X}$, $y \in \mathscr{Y}$, — функция вероятностей дискретного сл.вектора $(\xi,\eta) \in \mathbb{R}^2$. Если $p_\eta(y) = \mathbf{P}\{\eta = y\} > 0$, то для каждого $x \in \mathbb{R}^1$ можно определить условную вероятность $\mathbf{P}\big\{\xi = x \,\big|\, \eta = y\big\} = p(x,y)/p_\eta(y)$. Легко понять из (8), стр. 42, что сумма этих вероятностей по всем $x \in \mathscr{X}$ равна 1. Таким образом, набор вероятностей

$$\mathrm{P}_{\xi|\eta}(x \,|\, y) := \frac{p(x,y)}{p_\eta(y)}, \quad x \in \mathscr{X}, \tag{32}$$

определяет дискретное распределение, которое называется *условным распределением ξ при фиксированном значении случайной величины $\eta = y$*. Носитель этого распределения, вообще говоря, зависит от y.

Хотя ясно, что сл.в. η не может принимать значения вне носителя $\mathscr{Y}$, с формальной точки зрения необходимо доопределить $\mathrm{P}_{\xi|\eta}(x \,|\, y)$ и при $y \notin \mathscr{Y}$. Поскольку вероятность совокупности всех точек $y \notin \mathscr{Y}$ (подчеркнём, что именно всей совокупности таких точек, а не каждой точки в отдельности) равна нулю, такое доопределение можно осуществить произвольным образом, учитывая, например, удобство записи, лишь бы только функция $\mathrm{P}_{\xi|\eta}(x \,|\, y)$ задавала некоторое вероятностное распределение. Например, можно положить $\mathrm{P}_{\xi|\eta}(x_0\,|\,y) = 1, \mathrm{P}_{\xi|\eta}(x\,|\,y) = 0,\ x \neq x_0$, с произвольным $x_0 \in \mathscr{X}$.

Относительно условного распределения можно определить любые числовые характеристики. Например, для измеримой функции $h(x, y)$ величину

$$\mathrm{E}\big[h(\xi,\eta)\,|\,y\big] := \sum_{x\in\mathscr{X}} h(x,y)\,\mathrm{P}_{\xi|\eta}(x\,|\,y) \tag{33}$$

естественно назвать *условным математическим ожиданием* $h(\xi, \eta)$ (кратко, усл.м.о.) при фиксированном $\eta = y$. В частности, для индикаторной функции $h(x, y) = \dot{\mathrm{I}}_A(x)$ получим классическую условную вероятность (усл.вер.) $\mathrm{P}_{\xi|\eta}\{A\,|\,y\} = \mathbf{P}\big\{\xi \in A\,|\,\eta = y\big\}$ события $\{\xi \in A\}$ при условии, что $\eta = y$.

Заметим, что при $y \notin \mathscr{Y}$ усл.м.о. может принимать произвольные значения. Другими словами, определённое выше усл.м.о. есть только один из вариантов усл.м.о. Важно, что любой другой вариант усл.м.о. будет отличаться от него только на множестве $\mathscr{Y}^c$, мера которого равна нулю.

Легко видеть, что усл.м.о. $\mathrm{E}\big[h(\xi,\eta)\,|\,y\big]$ есть измеримая функция y, поэтому суперпозиция $\mathrm{E}\big[h(\xi,\eta)\,|\,\eta\big]$ представляет собой сл.в. Её математическое ожидание

$$\begin{aligned}\mathrm{E}\Big[\,\mathrm{E}\big[h(\xi,\eta)\,|\,\eta\big]\,\Big] \;&=\; \sum_{y\in\mathscr{Y}}\Big[\sum_{x\in\mathscr{X}} h(x,y)\,\mathrm{P}_{\xi|\eta}(x\,|\,y)\Big]p_\eta(y) \;=\\ &=\; \sum_{(x,y)\in\mathscr{X}\times\mathscr{Y}} h(x,y)p(x,y) \;=\; \mathrm{E}[h(\xi,\eta)],\end{aligned} \tag{34}$$

при любом способе определения $\mathrm{P}_{\xi|\eta}$ в точках y с $p_\eta(y) = 0$. Другими словами, математическое ожидание $h(\xi, \eta)$ можно вычислить поэтапно, сначала при фиксированном значении η, затем усредняя по η.

Для индикаторной функции $h(x, y) = \dot{\mathrm{I}}_A(x)$ соотношение (34) даёт хорошо известную формулу полной вероятности:

$$\mathbf{P}\{\xi \in A\} \;=\; \mathrm{E}[\dot{\mathrm{I}}_A(\xi)] \;=\; \sum_{y\in\mathscr{Y}} \mathrm{P}_{\xi|\eta}\{A\,|\,y\}p_\eta(y).$$

✧ **Абсолютно непрерывные модели.** Пусть теперь сл.вектор (ξ, η) имеет абсолютно непрерывную ф.р. В этой ситуации мы не можем вычислить условную вероятность $\mathbf{P}\big\{\xi \in A\,|\,\eta = y\big\}$ стандартным способом, поскольку вероятность любого одноточечного события $\{\eta = y\}$ равна нулю. Приведём некоторые соображения, позволяющие и в этом случае ввести понятие условного распределения, удовлетворяющее свойству (34).

Первая идея исходит из принципа «по аналогии». Усл.вер. равна отношению совместной вероятности событий на вероятность условия. В непрерывном случае аналогом вероятности одноточечного события выступает функция плотности. Пусть $f(x, y)$ — совместная ф.пл. вектора (ξ, η), а $f_\eta(y) = \int_{\mathbb{R}} f(x, y)\,dx$ — частная ф.пл. η, тогда для любого y с $f_\eta(y) > 0$ можно определить функцию

$$f_{\xi|\eta}(x \mid y) = \frac{f(x, y)}{f_\eta(y)}, \qquad x \in \mathbb{R}^1. \tag{35}$$

89| Лемма. (?!) *При* $f_\eta(y) > 0$ *функция* (35) *есть плотность некоторого распределения, т.е.* $f_{\xi|\eta}(x \mid y) \geqslant 0$, $\forall\, x \in \mathbb{R}^1$, *и* $\int_{\mathbb{R}} f_{\xi|\eta}(x \mid y)\,dx = 1$.

К той же самой ф.пл. мы придём, если определим условную плотность в точке «естественным» образом — как предел условных вероятностей в «малой» окрестности условия:

$$\lim_{\varepsilon \to +0} \frac{1}{\varepsilon}\, \mathbf{P}\big\{\xi \in (x-\varepsilon; x+\varepsilon] \mid \eta \in (y-\varepsilon; y+\varepsilon]\big\} =$$
$$= \lim_{\varepsilon \to +0} \frac{\big(\Delta_{\vec{u}-\varepsilon, \vec{u}+\varepsilon} F\big)/\varepsilon^2}{\big(F_\eta(y+\varepsilon) - F_\eta(y-\varepsilon)\big)/\varepsilon} = \frac{f(x, y)}{f_\eta(y)},$$

где $\vec{u} \pm \varepsilon = (x \pm \varepsilon, y \pm \varepsilon)$ и последнее равенство справедливо для почти всех (по распределению F) точек (x, y) с $f_\eta(y) > 0$. Заметим, что такой способ вычисления условных характеристик (ф.р., ф.пл., математического ожидания) представляется единственно возможным, если дело касается одного условия $\eta = y$, поскольку позволяет дать этим характеристикам практически приемлемую интерпретацию. В качестве примера можно привести свойство броуновского моста как гауссовского процесса, распределение которого совпадает с условным распределением винеровского процесса (в духе предыдущих построений) при условии, что его значение в момент $t = 1$ равно нулю (см. 406, стр. 363).

Построенные здесь условные распределения и усл.м.о. обладают всеми свойствами общего усл.м.о. Ключевым является свойство (34) (или его более широкий вариант (37)), позволяющее вычислять любые характеристики сл.вектора (ξ, η) последовательно. Для осуществления этой возможности в данном случае необходимо доопределить условную плотность (35) для точек с $f_\eta(y) = 0$. Поскольку вероятность совокупности всех таких точек равна нулю (?!), можно определить $f_{\xi|\eta}(x \mid y)$ произвольным образом, лишь бы только эта функция задавала ф.пл. Например, положить $f_{\xi|\eta}(x \mid y) = f_0(x)$ с некоторой фиксированной ф.пл. $f_0(x)$ на $\mathbb{R}^1$ относительно меры Лебега (хотя последнее нужно только для соблюдения единообразия вводимых определений).

Математическое ожидание относительно условной плотности

$$\mathrm{E}\{h(\xi,\eta)\,|\,y\} \;=\; \mathrm{E}\{h(\xi,\eta)\,|\,\eta=y\} \;:=\; \int_{\mathbb{R}} h(x,y) f_{\xi|\eta}(x\,|\,y)\,dx \tag{36}$$

назовём усл.м.о. $h(\xi,\eta)$ при фиксированном $\eta=y$. Выбирая здесь $h(x,y) = \dot{\mathrm{I}}_A(x)$ — индикаторную функцию произвольного борелевского множества $A\subset\mathbb{R}^1$ или множества $A=\{x\in\mathbb{R}^1 : x\leqslant z\}$, $z\in\mathbb{R}^1$, придём к условному распределению или, соответственно, к условной ф.р. сл.в. ξ [‡] :

$$\mathrm{P}_{\xi|\eta}\{A\,|\,y\} := \int\limits_A f_{\xi|\eta}(x\,|\,y)\,dx, \qquad F_{\xi|\eta}(z\,|\,y) := \int\limits_{-\infty}^{z} f_{\xi|\eta}(x\,|\,y)\,dx.$$

Следующее соотношение берётся за основу при определении усл.м.о. для общих вероятностных моделей.

90| **Лемма.** (?!) *Для любых неотрицательных борелевских функций* $g(y):\mathbb{R}^1\mapsto\mathbb{R}^1_+$, $h(x,y):\mathbb{R}^2\mapsto\mathbb{R}^1_+$ *справедливо*

$$\begin{aligned}\mathbf{E}\big[g(\eta)h(\xi,\eta)\big] \;&=\; \mathbf{E}\Big(g(\eta)\,E\big[h(\xi,\eta)\,|\,\eta\big]\Big) \;= \\ &=\; \int_{\mathbb{R}} g(y)\Big[\int_{\mathbb{R}} h(x,y) f_{\xi|\eta}(x\,|\,y)\,dx\Big] f_\eta(y)\,dy.\end{aligned} \tag{37}$$

Выбирая g и h как индикаторные функции $g(y)=\dot{\mathrm{I}}(y; y\leqslant z)$ и $h(x,y) = \dot{\mathrm{I}}(x,y;\, x\leqslant u)$, $z,u\in\mathbb{R}^1$, получаем

91| **Следствие.** *Пусть* F_η, f_η *— соответственно ф.р. и ф.пл. сл.в.* η*. Тогда совместная ф.р. сл.вектора* (ξ,η) *может быть вычислена как:*

$$\begin{aligned} F(u,z) \;&\overset{\mathrm{df}}{=}\; \mathbf{P}\{\xi\leqslant u, \eta\leqslant z\} \;=\; \int_{-\infty}^{z} F_{\xi|\eta}(u\,|\,y) f_\eta(y)\,dy \;= \\ &=\; \int_{-\infty}^{z} F_{\xi|\eta}(u\,|\,y)\,dF_\eta(y).\end{aligned} \tag{38}$$

Если в (37) выбрать $g\equiv 1$, $h(x,y)=\dot{\mathrm{I}}_A(x)$ с борелевским множеством $A\subset\mathbb{R}^1$, то получим непрерывный аналог формулы полной вероятности:

$$\mathbf{P}\{\xi\in A\} \;=\; \mathbf{E}\big[\mathbf{P}\big\{\xi\in A\,|\,\eta\big\}\big] \;=\; \int_{\mathbb{R}} \mathrm{P}_{\xi|\eta}\{A\,|\,y\} f_\eta(y)\,dy.$$

[‡] По свойству ф.р. обязательно $F_j(x)=0$ при $x\leqslant 0$ и $F_j(x)=1$ при $x\geqslant 1$.

⚠ Как и в дискретном случае, здесь также для точек из некоторого множества вероятности нуль усл.м.о. определяется произвольным образом. Для абсолютно непрерывной модели это связано не только с необходимостью доопределения условного распределения при $f_\eta(y) = 0$. Напомним, что плотность есть производная Радона–Никодима, представляющая собой целое семейство функций, совпадающих только почти наверное. Даже если ограничиться вариантом плотности, вычисляемым как обычная производная, иногда приходится доопределять её некоторым (опять произвольным) способом в точках, где эта производная не существует.

Общая концепция условного распределения уже не может опираться на соображения, использованные нами в дискретной и абсолютно непрерывной моделях. Оказывается, можно ввести определение усл.м.о. $\mathbf{E}\big[\,h(\xi,\eta)\,|\,\eta\,\big]$, а вслед за ним и условного распределения, опираясь только на одно ключевое соотношение, аналогичное равенствам (34), (37):

$$\mathbf{E}\left[\,\dot{\mathrm{I}}_B(\eta)h(\xi,\eta)\right] \;=\; \mathbf{E}\left[\,\dot{\mathrm{I}}_B(\eta)\,\mathbf{E}\big[\,h(\xi,\eta)\,|\,\eta\,\big]\right], \qquad \forall\, B\in\mathscr{B}.$$

В соответствии с этим определением, усл.ф.р. ξ при фиксированном $\eta = y$ есть измеримая (по y) функция распределения (по x) $F_{\xi|\eta}(x\,|\,y)$, удовлетворяющая (38). Конструкции условных распределений, приведённые выше, представляют собой только частные варианты общей конструкции, в том смысле, что любой другой вариант будет отличаться от них лишь почти наверное.

⚠ Отметим одно важное обстоятельство, связанное с общей конструкцией усл.м.о. Исходя из (33) и (36), можно заметить, что при всех y

$$\mathbf{E}\big[\,h(\xi,\eta)\,|\,\eta=y\,\big] \;=\; \mathbf{E}\big[\,h(\xi,y)\,|\,\eta=y\,\big]. \tag{39}$$

В общем случае такое равенство можно обосновать либо при дополнительных условиях на вектор (ξ,η), например когда сл.в. ξ,η независимы, либо при конкретизации способа вычисления соответствующих усл.м.о. (например, через ф.плотности или ф.вероятности, как это сделано выше, или условное распределение, если указан способ его построения).

Если распределение η имеет плотность f_η, то из соотношения (38) следует, что совместная ф.р. $F(x,y) = \int_{-\infty}^{y} F(x\,|\,t) f_\eta(t)\,dt$. Таким образом, при фиксированном x функция $F(x,y)$ абсолютно непрерывна по y и, следовательно, условную ф.р. ξ при почти всех (по мере Лебега) $y\in\mathbb{R}^1$, $f_\eta(y) > 0$, можно вычислять по формуле

$$F(x\,|\,y) = \frac{1}{f_\eta(y)}\frac{\partial F(x,y)}{\partial y}\ .$$

92 Пример. Пусть $F(x,y) = G(x)G(y)\,\dot{\mathrm{I}}\,(x \leqslant y) + G^2(y)\,\dot{\mathrm{I}}\,(x \geqslant y)$ с некоторой непрерывной ф.р. G (см. пример 62, стр. 66). Маргинальная ф.р. второй компоненты $F_\eta(y) = F(+\infty, y) = G^2(y)$. Если ф.р. G имеет плотность g, то плотность $f_\eta(y) = 2g(y)G(y)$. В этом случае по предыдущей формуле

$$F(x\,|\,y) = \frac{1}{2}\,\frac{G(x)}{G(y)}\;\dot{\mathrm{I}}\,(x < y) + 1\cdot\dot{\mathrm{I}}\,(x > y).$$

При $x = y$ эту усл.ф.р. мы вольны выбрать произвольно, дабы сохранить свойство непрерывности справа, положим $F(y\,|\,y) = 1$. Выражение для $F(x\,|\,y)$ не содержит плотность g. Можно предположить, что этот результат имеет место и при отсутствии плотности. Проверим справедливость (38).

В случае $u < z$ ф.р. $F(u,z) = G(u)G(z)$. Для вычисления интеграла в (38) его необходимо разбить на две части $y \leqslant u$ и $u < y \leqslant z$:

$$\int_{-\infty}^{u} + \int_{u}^{z} F(u\,|\,y)\,dG^2(y) = \int_{-\infty}^{u} 1\,dG^2(y) + \frac{1}{2}\,G(u)\int_{u}^{z}\frac{1}{G(y)}\,dG^2(y).$$

Первый интеграл, очевидно, равен $G^2(u)$. Во втором интеграле можно сначала осуществить замену $G(y) \dashrightarrow x,\ G(u) \leqslant x \leqslant G(z)$, см. 40, стр. 53. В результате после простых вычислений получим $G(u)G(z)$. В случае $u > z$ в предыдущем уравнении останется только первый интеграл (здесь уже до z), т.е. снова правая часть (38) совпадёт с ф.р. $F(u,z) = G^2(z)$. ◎

✧ Совместное распределение через условное. Равенство (38) можно использовать и в обратном направлении. Часто ф.р. наблюдаемой в эксперименте сл.в. ξ зависит от некоторых параметров y : $F_\xi(x) = F_\xi(x\,|\,y)$. Если, в свою очередь, значения y представляют собой реализации некоторого сл.вектора η, то (38) даёт способ определения совместного распределения сл.вектора (ξ, η) (аналог формулы Фубини через переходную ф.р.).

93 **Теорема.** *Пусть функция $G(x\,|\,y)$ при каждом $y \in \mathbb{R}^1$ есть ф.р., а при каждом фиксированном x она измерима как функция переменной y; F_0 — некоторая ф.р. Тогда:*

i) *функция*

$$F(x,y) = \int_{-\infty}^{y} G(x\,|\,u)\,dF_0(u)$$

задаёт ф.р. двумерного сл.вектора (ξ,η), *где* $\eta \backsim F_0$;

ii) *для любой неотрицательной измеримой функции* $h(x,y)$ [‡]

$$\mathbf{E}\left[h(\xi,\eta)\right] = \int\limits_{y\in\mathbb{R}^1}\left[\int\limits_{x\in\mathbb{R}^1} h(x,y)\,dG(x\,|\,y)\right]dF_0(y), \tag{40}$$

в частности при $\forall\, A, B \in \mathscr{B}$

$$\mathbf{P}\{\xi\in A, \eta\in B\} = \int\limits_{y\in B}\left[\int\limits_{x\in A} dG(x\,|\,y)\right]dF_0(y).$$

$\Rrightarrow$ i) Свойства нормированности и непрерывности ф.р. легко проверяются с привлечением теоремы Лебега о мажорируемой сходимости. Свойство монотонности следует из представления

$$\Delta_{\vec{a},\vec{c}}F(x,y) \;=\; \int_{a_1}^{c_1}\big(G(c_2\,|\,t)-G(a_2\,|\,t)\big)\,dF_0(t) \;\geqslant\; 0.$$

ii) Пусть $\mathcal{H}$ — класс неотрицательных борелевских функций из $\mathbb{R}^2$ в $\mathbb{R}^1$, для которых справедливо доказываемое равенство. В частности, для этих функций интеграл $\int_{x\in\mathbb{R}^1} h(x,y)\,dG(x\,|\,y)$ измерим по Борелю как функция $y\in\in\mathbb{R}^1$. По построению в этот класс входят индикаторные функции $\dot{\mathrm{I}}_U$ всех угловых множеств $U = (-\infty;x]\times(-\infty;y]$, а также индикаторная функция $\mathbb{R}^2$. Класс $\mathcal{H}$ замкнут относительно неотрицательных линейных комбинаций. Пусть $h_n\in\mathcal{H}$ и $h_n\nearrow h_0,\ n\to\infty$. По теореме Б. Леви о монотонной сходимости $\int_{\mathbb{R}} h_n(x,y)\,dG(x\,|\,y)\nearrow\int_{\mathbb{R}} h_0(x,y)\,dG(x\,|\,y)$ при $\forall\, y\in\mathbb{R}^1$ и

$$\int_{\mathbb{R}}\int_{\mathbb{R}} h_n(x,y)\,dG(x\,|\,y)\,dF_0(y) \;\nearrow\; \int_{\mathbb{R}}\int_{\mathbb{R}} h_0(x,y)\,dG(x\,|\,y)\,dF_0(y).$$

Таким образом, класс $\mathcal{H}$ замкнут относительно монотонных пределов, следовательно, по лемме 515, стр. 468, этот класс содержит все неотрицательные функции, измеримые относительно σ-алгебры, порождённой угловыми множествами U, т.е. измеримые относительно борелевской σ-алгебры. $\Lleftarrow$

94| **Следствие.** *Пусть при каждом* $y\in\mathbb{R}^1$ *функция* $f(x\,|\,y)$ *есть ф.плотности относительно некоторой меры* μ_1 *на* $(\mathbb{R}^1,\mathscr{B}^1)$ *и, кроме того, при каждом* x *эта функция измерима по* y. *Пусть также* $f_\eta(y),\ y\in\mathbb{R}^1$, *есть ф.плотности относительно какой-либо меры* μ_2 *на* $(\mathbb{R}^1,\mathscr{B}^1)$. *Тогда функция* $f(x,y) = f(x\,|\,y)f_\eta(y)$ *есть ф.плотности распределения некоторого сл.вектора*

[‡] Запись вида $\int_{u\in A}$ используется для указания на u как переменную интегрирования.

(ξ,η) *относительно произведения мер* $\mu_1\times\mu_2$ *на* $(\mathbb{R}^2,\mathscr{B}^2)$. *При этом*

$$\mathbf{E}\left[h(\xi,\eta)\right] \;=\; \int\limits_{y\in\mathbb{R}} \Big[\int\limits_{x\in\mathbb{R}} h(x,y)f(x\,|\,y)\,\mu_1(dx)\Big] f_\eta(y)\,\mu_2(dy).$$

95| Упр. Докажите это следствие с помощью теоремы Фубини.

⚠ Предыдущие построения легко переносятся на случай многомерных $\vec{\xi}$ и $\vec{\eta}$. Например, если $f(x_1,x_2,y_1,y_2,y_3)$ — совместная ф.пл. вектора $(\xi_1,\xi_2,\eta_1,\eta_2,\eta_3)$, то $f_{\vec{\eta}}(y_1,y_2,y_3)=\int_{\mathbb{R}^2} f(x_1,x_2,y_1,y_2,y_3)\,dx_1dx_2$ — частная плотность $\vec{\eta}=(\eta_1,\eta_2,\eta_3)$ и усл.м.о. при фиксированном значении $\vec{\eta}=\vec{y}$ любой неотрицательной функции $h(x_1,x_2,y_1,y_2,y_3)$ определяется как

$$\mathbf{E}\big[\,h(\vec{\xi},\vec{\eta})\mid\vec{y}\,\big]=\frac{1}{f_{\vec{\eta}}(\vec{y})}\int_{\mathbb{R}^2} h(\vec{x},\vec{y})f(\vec{x},\vec{y})\,d\vec{x}.$$

96| П р и м е р. Статистический эксперимент состоит в наблюдении последовательности независимых сл.в. $\xi_1,\ldots,\xi_n$ с общей ф.р. $F(x\,|\,\theta)$, зависящей от параметра θ; цель исследования — принятие решения о значении θ. При байесовском подходе к принятию решения, когда одна и та же статистическая задача возникает многократно, иногда можно предположить, что параметр θ в каждом из экспериментов есть реализация некоторой сл.в. ϑ с ф.р. $G(\theta)$, называемой *априорной* ф.р. Тогда ф.р.

$$F(\vec{x},\theta)=\int_{-\infty}^{\theta}\prod\nolimits_1^n F(x_i\,|\,u)\,dG(u)$$

описывает совместное распределение вектора $\vec{\xi}=(\xi_1,\ldots,\xi_n)$ и параметра ϑ. Оптимальное в байесовском подходе решение всегда основывается на *апостериорном распределении* — условном распределении ϑ при фиксированном значении наблюдаемого вектора $\vec{\xi}=\vec{x}$. ◎

97| П р и м е р. С помощью формулы (40) легко доказывается равенство (68), стр. 72. Пусть $F_{\eta|\xi}(y\,|\,x)$ — условная ф.р. η при фиксированном значении $\xi=x$. Тогда для любой неотрицательной измеримой функции $h(x)$ и её естественного продолжения $\widetilde{h}(x,y)=h(x)$ на $\mathbb{R}^2$ имеем

$$\int\limits_{\mathbb{R}^2}\widetilde{h}(x,y)\,dF(x,y) \;=\; \int\limits_{z\in\mathbb{R}^1} h(x)\Big[\int\limits_{y\in\mathbb{R}^1} dF_{\eta|\xi}(y\,|\,x)\Big]dF_\xi(x) \;=\; \int\limits_{x\in\mathbb{R}^1} h(x)\,dF_\xi(x).$$

◎

✧ **Наилучший нелинейный прогноз.** Теория условных распределений помогает отыскать наилучший прогноз значений одной сл.в. по значениям любого сл.вектора. Рассмотрим две сл.в. ξ, η и найдём функцию h, для которой значения $h(\eta)$ наилучшим образом прогнозируют значения ξ в смысле достижения минимума среднеквадратической ошибки прогноза $\mathbf{E}\big(\xi - h(\eta)\big)^2$. Пусть $F_{\xi|\eta}(x \mid y)$ — какой-либо из вариантов усл.ф.р. ξ при фиксированном η. Вычислим усл.м.о. относительно этой усл.ф.р.:

$$\mu^*(y) \;=\; \mathrm{E}\big[\xi \mid \eta = y\big] \;=\; \int\limits_{x\in\mathbb{R}^1} x\, dF_{\xi|\eta}(x \mid y).$$

98| **Теорема.** *Если* $\mathbf{E}\xi^2 < \infty$, *то функция* $E\big[\xi \mid \eta = y\big]$ *минимизирует среднеквадратическую ошибку прогноза значений сл.в.* ξ *по значениям* η.

⇛ Ввиду (37) $\mathrm{E}\big[\xi^2 \mid \eta = y\big] < \infty$, а значит, и $|\mu^*(y)| < \infty$ для почти всех $y \in \mathbb{R}^1$. Для любой функции h справедливо простое тождество

$$\mathbf{E}(\xi - h(\eta))^2 \;=\; \mathbf{E}(\xi - \mu^*(\eta))^2 + \mathbf{E}(\mu^*(\eta) - h(\eta))^2 + \\ + 2\,\mathbf{E}\left[(\xi - \mu^*(\eta))(\mu^*(\eta) - h(\eta))\right].$$

В силу равенства, подобного (37), третье слагаемое равно нулю:

$$\mathbf{E}\left[(\xi - \mu^*(\eta))(\mu^*(\eta) - h(\eta))\right] \;=\; \mathbf{E}\Big[(\mu^*(\eta) - h(\eta))\,\mathbf{E}\big[\,(\xi - \mu^*(\eta)) \mid \eta\,\big]\Big] \;=\; 0,$$

т.к. $\mathbf{E}\big[\,(\xi - \mu^*(\eta)) \mid \eta = y\,\big] = 0$ при всех y. Отсюда ясно, что минимум среднеквадратического расхождения $\mathbf{E}(\xi - h(\eta))^2$ достигается, когда $\mathbf{E}(\mu^*(\eta) - h(\eta))^2$ равно нулю, т.е. когда $h = \mu^*$. ⇚

Определение. Условное математическое ожидание $\mathrm{E}\big[\xi \mid \eta = y\big]$ называется *(среднеквадратической) регрессией* ξ *на* η.

⚠ Теорема 98 даёт ещё один способ определения усл.м.о. случайной величины ξ при фиксированном значении $\eta = y$ как функции $h^*(y)$, для которой достигается минимум математического ожидания $\mathbf{E}\big(\xi - h(\eta)\big)^2$ по всем измеримым функциям h. Существование минимума легко устанавливается с привлечением теории гильбертовых пространств. Единственное неудобство — необходимость предположить конечность второго момента у ξ, что, конечно, несущественно для индикаторной функции $\xi = \dot{\mathrm{I}}_B$.

99| Примеры. 1) Так как для независимых сл.в. совместная плотность равна произведению частных плотностей, то условная плотность ξ_1 при любом фиксированном $\xi_2 = x_2$ равна частной (безусловной) плотности ξ_1; условное среднее $\mathrm{E}\big[\xi_1 \,|\, \xi_2\big]$ совпадает с математическим ожиданием $\mathbf{E}\,\xi_1$.

2) Плотность двумерной нормальной модели (21), стр. 70 (при $|\rho| < 1$), после элементарных преобразований можно записать в виде

$$f(y,x) \;=\; \frac{1}{2\pi\sqrt{(1-\rho^2)}}\exp\left\{-\frac{(y-\rho x)^2}{2(1-\rho^2)} - \frac{x^2}{2}\right\} \;=\; \;=\; \phi\big(x \,|\, 0,1\big)\;\phi\big(y \,|\, \rho x, (1-\rho^2)\big),$$

где $\phi(x \,|\, \mu,\sigma^2) = \exp\{-(x-\mu)^2/2\sigma^2\}/(\sqrt{2\pi}\sigma)$ — плотность одномерного нормального распределения со средним μ и дисперсией σ^2.

Отсюда ясно, что, т.к. $\int_{\mathbb{R}^1}\phi\big(y \,|\, \rho x,(1-\rho^2)\big)\,dy = 1$, частная плотность $f_\xi(x) = \phi(x \,|\, 0,1)$. Следовательно, при фиксированном $\xi = x$ условное распределение $\eta\big|_{\xi=x} \backsim \phi\big(\,\cdot\, \,|\, \rho x,(1-\rho^2)\big)$, т.е. нормально с соответствующими параметрами. Таким образом, регрессия $\mu_{\eta|\xi}(x) = \rho x$ линейна.

3) Пусть сл.вектор (ξ_1,ξ_2) имеет равномерное распределение внутри треугольника с вершинами $(-1,0),(0,1),(1,0)$. Легко понять, что при фиксированном $\xi_2 = x_2 \in [0;1]$ условная плотность ξ_1 также равномерна с носителем $(x_2-1;1-x_2)$ (от левой стороны треугольника до правой стороны). Поэтому условное среднее равно 0. Снова мы получили регрессию, совпадающую с линейной регрессией: $\mu_{1|2}(x_2) = 0 + 0x_2$. Из общего вида линии регрессии (теорема 71, стр. 74) отсюда следует, что коэффициент корреляции $\rho = 0$. При этом сл.в. ξ_1,ξ_2 зависимы, поскольку условная плотность одной сл.в. зависит от значения другой сл.в. Интересно, что регрессия ξ_2 на $\xi_1 = x_1$ уже не линейна: $\mu_{2|1}(x_1) = (1-|x_1|)/2$ при $x_1 \in [-1;1]$.

4) Из конструкции многомерного гипергеометрического распределения $\mathcal{H}\mathrm{g}_k(n;N_1,\ldots,N_k)$ (стр. 32) ясно, что если в выборке зафиксировать количество элементов какой-либо категории (или группы категорий), то распределение оставшихся категорий снова будет подчинено многомерному гипергеометрическому закону. Например, условное распределение $(\xi_1,\ldots,\xi_{k-1})$ при фиксированном $\xi_k = n_k$ совпадает с $\mathcal{H}\mathrm{g}_{(k-1)}(n-n_k;N_1,\ldots,N_{k-1})$. Впрочем, этот результат может быть получен и чисто формальными рассуждениями, если вспомнить, что здесь любой маргинальный закон принадлежит тому же типу. Следовательно, условное среднее ξ_1 можно записать в виде

$$\mathrm{E}\left[\xi_1 \,|\, \xi_2 = n_2,\ldots,\xi_J = n_J\right] \;=\; \frac{N_1\big(n-(n_2+\ldots+n_J)\big)}{N-(N_2+\ldots+N_J)}\,.$$

Таким образом, регрессия зависит линейно от условий. Полученное здесь выражение совпадает с линейной регрессией (30), стр. 81, с $q_r = N_r/N$.

Аналогичное верно и для мультиномиальной модели (стр. 33).

5) Если у вектора $\vec{\xi} \backsim \mathcal{D}\mathtt{ir}_k(a_1, \ldots, a_{k+1})$, имеющего распределение Дирихле (22), стр. 71, зафиксирована одна компонента $\xi_k = p_k$ $(\in (0;1))$, то область возможных значений остальных компонент сужается: $\xi_1 + \ldots + \xi_{k-1} \leqslant \leqslant 1 - p_k$. Чтобы привести условное распределение к стандартному виду, следует рассмотреть нормированный вектор $(\xi_1, \ldots, \xi_{k-1})/(1-p_k)$. Так как маргинальное распределение $\xi_k \backsim \mathcal{B}\mathtt{et}(a_k, \boldsymbol{a} - a_k)$ $(\boldsymbol{a} = a_1 + \ldots + a_{k+1})$, то легко показывается, что условное распределение

$$\frac{1}{(1-p_k)}(\xi_1, \ldots, \xi_{k-1})\Big|_{\xi_k = p_k} \backsim \ \mathcal{D}\mathtt{ir}_{(k-1)}(a_1, \ldots, a_{k-1}, a_{k+1}).$$

Таким образом, условное среднее $\mathtt{E}\big[\xi_1 \,|\, \xi_k\big] = (1-\xi_k)a_1/(\boldsymbol{a} - a_k)$.

6) При аттестации партии кофе после прохождения контроля в паспорте указывается доля $\theta \cdot 100\%$ зёрен сорта «Арабика». Контроль состоит в случайном отборе из партии n зёрен и в последующем их анализе. Поскольку количество отобранных зёрен обычно значительно меньше общего объёма партии, можно считать, что количество зёрен сорта «Арабика» в контрольной выборке есть реализация биномиальной сл.в. $\xi \backsim \mathcal{B}\mathtt{in}(n, \theta)$ с вероятностью успеха θ. Заметим, что в данной ситуации, когда требуется произвести оценку θ не для одной партии, а для ряда последовательно поступающих партий в отдельности, необходимо учитывать информацию о характере изменчивости величины θ от партии к партии. Пусть априори (до опыта) доля θ в каждой партии представляет собой реализацию бета-сл.в. $\vartheta \backsim \mathcal{B}\mathtt{et}(a_1, a_2)$ с плотностью $f(\theta \,|\, a_1, a_2)$, $0 \leqslant \theta \leqslant 1$, с некоторыми параметрами $a_1, a_2 > 0$ (см. стр. 107). Совместное распределение вектора (ξ, ϑ) (одна компонента дискретна, вторая непрерывна) удобнее описать с помощью функции плотности относительно произведения считающей меры и меры Лебега на прямой:

$$p(k, \theta) \ := \ p_\xi(k \,|\, \theta)\, f(\theta \,|\, a_1, a_2) \ = \ \mathrm{C}_n^k \frac{1}{\mathrm{B}(a_1, a_2)} \theta^{k + a_1 - 1}(1-\theta)^{n-k+a_2-1}.$$

Понятно, что условное распределение $\vartheta\big|_{\xi = k} \backsim \mathcal{B}\mathtt{et}(k + a_1, n - k + a_2)$. Поэтому регрессия ϑ на ξ, доставляющая по теореме 98 наилучшую оценку θ по результатам контрольных испытаний, вычисляется по формуле

$$\hat{\theta} \ = \ \mathtt{E}\big[\vartheta \,|\, \xi = k\big] \ = \ \frac{k + a_1}{n + a_1 + a_2}. \qquad \circledcirc$$

А. Доказательства

⊫ [*Теорема* 51, стр. 58] Пусть $\mathscr{X} = \{x_k,\ k = 1, \ldots, N\}$, $N \leqslant \leqslant \infty$, — не более чем счётное множество точек разрыва функции F (см. лемму 23, стр. 43), $p_k = F(x_k) - F(x_k - 0)$ — величина скачка (> 0) в точке x_k, $k = 1, \ldots, N$. Определим «дискретную» функцию

$$F_d(y) = \sum_{x_k \leqslant y} p_k, \qquad y \in \mathbb{R}^1.$$

В силу 50 эта функция непрерывна справа. Кроме того, т.к., очевидно, $F_d(y) < \infty$ для $\forall\, y$, то $F_d(-\infty) = 0$. Осталось показать, что разность $F_c(x) = F(x) - F_d(x)$ есть непрерывная и монотонно неубывающая функция.

Очевидно, $F_c(x+0) - F_c(x-0) = \big[F(x) - F(x-0)\big] - \big[F_d(x) - F_d(x-0)\big]$, поэтому согласно 50 $F_c(x+0) - F_c(x-0) = 0$, если $x \notin \mathscr{X}$, и $F_c(x+0) - F_c(x-0) = p_k - p_k = 0$, если $x = x_k \in \mathscr{X}$.

Пусть $Z = \mathscr{X} \cap (a; b]$ — совокупность точек на фиксированном конечном интервале $(a;\, b]$, в которых функция F терпит разрыв. Выберем произвольным образом возрастающую последовательность конечных наборов $A_n = \big\langle\, y_{n1} < \ldots < y_{nn} \,\big\rangle \subseteq Z$ так, чтобы $A_n \nearrow Z$ (или $A_n = Z$ для $\forall\, n$, если Z конечно). Тогда

$$F_c(b) - F_c(a) \;=\; F(b) - F(a) - \sum_{x_k \in Z} p_k \;=\; \lim_n \Big[F(b) - F(a) - \sum_{x_k \in A_n} p_k\Big].$$

Выражение в квадратных скобках здесь не меньше нуля при любом n. Действительно, в силу монотонности F разность функции $F(u) - F(v) \geqslant F(u) - F(u-0)$ при любых $v < u$. Поэтому

$$\begin{aligned} F(b) - F(a) &= \big[F(b) - F(y_{nn})\big] + \sum_{k=2}^{n} \big[F(y_{nk}) - F(y_{n(k-1)})\big] + \big[F(y_{n1}) - F(a)\big] \geqslant \\ &\geqslant \big[F(b) - F(y_{nn})\big] + \sum_{k=2}^{n} \big[F(y_{nk}) - F(y_{nk} - 0)\big] + \big[F(y_{n1}) - F(y_{n1} - 0)\big] \geqslant \\ &\geqslant \sum_{k=1}^{n} \big[F(y_{nk}) - F(y_{nk} - 0)\big] \;=\; \sum_{x_k \in A_n} p_k \;\geqslant\; 0, \end{aligned}$$

что доказывает монотонность F_c.

В силу непрерывности F_c, точки разрыва функции F_d совпадают с точками разрыва функции F с идентичными скачками. Поэтому в представлении $F = F_c + F_d$ дискретная функция определяется однозначно (при условии $F_d(-\infty) = 0$), стало быть, однозначно определяется и функция F_c. $\Lleftarrow$

$\Rrightarrow$ [*Лемма* 68, стр. 72.] Достаточно рассмотреть неотрицательные функции. Воспользуемся леммой 515, стр. 468. Пусть $\mathcal{H}$ — класс неотрицательных борелевских функций, для которых справедливо (23), стр. 72. Этот класс содержит все индикаторные функции $h(x) = \dot{\mathrm{I}}_B(x)$, $x \in \mathbb{R}^1$, $B \in \mathscr{B}^1$. Действительно, для таких функций естественное продолжение $\widetilde{h}(x,y) = \dot{\mathrm{I}}_{\widetilde{B}}(x,y)$ с 2-мерным «цилиндром» $\widetilde{B} = B \times \mathbb{R}^1$ и по определению интеграла

$$\mathbf{E}\,\widetilde{h}(\xi,\eta) \;=\; \mathbf{P}\big\{\,\xi \in B, \eta \in \mathbb{R}^1\,\big\} \;=\; \mathbf{P}\big\{\,\xi \in B\,\big\} \;=\; \mathbf{E}\,h(\xi).$$

Кроме того, очевидно, класс $\mathcal{H}$ замкнут относительно линейных комбинаций с неотрицательными коэффициентами. Пусть теперь $\langle\, h_n \,\rangle_1^\infty \subset \mathcal{H}$ — последовательность функций из $\mathcal{H}$, монотонно сходящаяся к некоторой функции h_0. Тогда $\widetilde{h}_n(x,y) \nearrow \widetilde{h}_0(x,y)$, и по теореме о монотонной сходимости Б. Леви

$$\mathbf{E}\,\widetilde{h}_0(\xi,\eta) \;=\; \lim_{n}{}_{\uparrow}\, \mathbf{E}\,\widetilde{h}_n(\xi,\eta) \;=\; \lim_{n}{}_{\uparrow}\, \mathbf{E}\,h_n(\xi) \;=\; \mathbf{E}\,h_0(\xi).$$

Следовательно, $h_0 \in \mathcal{H}$. В силу 515, стр. 468, класс $\mathcal{H}$ содержит все неотрицательные борелевские функции. $\Lleftarrow$

$\Rrightarrow$ [*Формула Гёфдинга* 24, стр. 74.] Нам понадобится следующее легко устанавливаемое равенство:

$$(*) \qquad \sum\nolimits_{j=0}^{n-1} j\big(g_{j+1} - g_j\big) \;=\; n g_n - \sum\nolimits_{j=0}^{n-1} g_{j+1},$$

справедливое для любой числовой последовательности $g_j, j \geqslant 0$. Рассмотрим интеграл от функции xy по области $\{(x,y) \in (-A;0] \times (0;A]\}$ с некоторым $A > 0$. Разобьём эту область на n^2 «равных» частей $B_{kj} = \{(x,y) : -\frac{A(k+1)}{n} < x \leqslant -\frac{Ak}{n}, \frac{Aj}{n} < y \leqslant \frac{A(j+1)}{n}\}$, $k,j = 0,\ldots,n-1$. Очевидно, последовательность простых функций

$$h_n(x,y) \;=\; \sum_{k=0}^{n-1}\sum_{j=0}^{n-1} \frac{Ak}{n}\,\frac{Aj}{n}\;\dot{\mathrm{I}}\,\big((x,y) \in B_{kj}\big) \;\nearrow\; -xy$$

при $n \to \infty$ для $\forall\,(x,y) \in (-A;0] \times (0;A]$. По определению интеграла Лебега

$$\iint\limits_{(-A;0]\times(0;A]} -xy\,dF(x,y) = \; A^2 \lim_n \frac{1}{n^2} \sum_{k=0}^{n-1}\sum_{j=0}^{n-1} kj\,\mu_F\big((x,y) \in B_{kj}\big) \;=$$

$$= A^2 \lim_n \frac{1}{n^2} \sum_{k=0}^{n-1} k \sum_{j=0}^{n-1} j \Big(\Big[F\Big(-\frac{Ak}{n}, \frac{A(j+1)}{n} \Big) - F\Big(-\frac{Ak}{n}, \frac{Aj}{n} \Big) \Big] -$$

$$- \Big[F\Big(-\frac{A(k+1)}{n}, \frac{A(j+1)}{n} \Big) - F\Big(-\frac{A(k+1)}{n}, \frac{Aj}{n} \Big) \Big] \Big).$$

В силу равенства $(*)$ двойная сумма здесь преобразуется к виду

$$\sum_{k=0}^{n-1} k \Big[nF\Big(-\frac{Ak}{n}, A \Big) - \sum_{j=0}^{n-1} F\Big(-\frac{Ak}{n}, \frac{A(j+1)}{n} \Big) -$$

$$- nF(-\frac{A(k+1)}{n}, A) + \sum_{j=0}^{n-1} F\Big(-\frac{A(k+1)}{n}, \frac{A(j+1)}{n} \Big) \Big] =$$

$$= -n^2 F(-A, A) + n \sum_{k=0}^{n-1} F\Big(-\frac{A(k+1)}{n}, A \Big) + n \sum_{j=0}^{n-1} F\Big(-A, \frac{A(j+1)}{n} \Big) -$$

$$- \sum_{k=0}^{n-1} \sum_{j=0}^{n-1} F\Big(-\frac{A(k+1)}{n}, \frac{A(j+1)}{n} \Big).$$

Переходя к пределу при $n \to \infty$, получаем

$$\int_{-A}^{0} \int_{0}^{A} (-xy)\, dF(x,y) = \int_{-A}^{0} \int_{0}^{A} \big[F(x,A) + F(-A,y) - F(x,y) - F(-A,A) \big]\, dxdy =$$

$$= \int_{-A}^{0} \int_{0}^{A} \mathbf{P} \{ (\xi, \eta) \in (-A; x] \times (y; A] \}\, dxdy.$$

Область интегрирования и подынтегральная функция возрастают по $A > 0$. Стало быть, по теореме о монотонной сходимости 534, стр. 480,

$$\int_{-\infty}^{0} \int_{0}^{\infty} xy\, dF(x,y) = \int_{-\infty}^{0} \int_{0}^{\infty} \big[F(x,y) - F_1(x) \big]\, dxdy.$$

Поскольку доказанное равенство справедливо для любой ф.р. $F(x,y)$, в частности для $F_1(x)F_2(y)$, получаем

$$\int_{-\infty}^{0} \int_{0}^{\infty} xy\, dF(x,y) - \int_{-\infty}^{0} x\, dF_1(x) \int_{0}^{\infty} y\, dF_2(y) = \int_{-\infty}^{0} \int_{0}^{\infty} \big[F(x,y) - F_1(x)F_2(y) \big]\, dxdy,$$

если слагаемые конечны, что обеспечивается условиями теоремы. Аналогичным образом исследуются три других квадранта декартовой системы. $\Lleftarrow$

В. Упражнения

Упр. В.1. Пусть сл.в. ξ, η таковы, что $\mathbf{E}\xi = \mathbf{E}\eta = 0, \mathbf{D}\xi = \mathbf{D}\eta = 1, \rho = \mathbf{Corr}(\xi, \eta)$. Доказать следующие неравенства:

a) $\mathbf{E}\max(\xi^2, \eta^2) \leqslant 1 + \sqrt{1-\rho^2}$;

b) $\mathbf{P}\big\{\, |\xi| \geqslant \varepsilon \text{ или } |\eta| \geqslant \varepsilon \,\big\} \leqslant \frac{1}{\varepsilon^2}\big(1 + \sqrt{1-\rho^2}\,\big)$.

Упр. В.2. Пусть ξ, η, ζ — попарно некоррелированные сл.в. Будут ли некоррелированными сл.в. a) ξ и $\eta + \zeta$; b) ξ и $\eta\zeta$?

Упр. В.3. Пусть $\langle\, \xi_k \,\rangle_1^n$ — независимые одинаково распределённые сл.в. с коэффициентом асимметрии $\mathbf{E}(\xi_1 - \mathbf{E}\,\xi_1)^3 = 0$. Показать некоррелированность среднего $\overline{\xi} = \frac{1}{n}(\xi_1 + \ldots + \xi_n)$ и дисперсии $S^2 = \frac{1}{n}\sum_1^n(\xi_k - \overline{\xi}\,)^2$.

Упр. В.4. Привести пример разрывной двумерной плотности, у которой обе маргинальные плотности непрерывны.

Упр. В.5. Сл.вектор $\vec{\xi}$ имеет *сферически симметричное* распределение, если $Q\vec{\xi} \backsim \vec{\xi}$ для любой ортогональной матрицы Q, т.е. распределение вектора инвариантно относительно поворотов вокруг нуля. Показать, что в этом случае компоненты вектора центрированы и попарно некоррелированы.

Упр. В.6. Пусть $\vec{\xi} \in \mathbb{R}^k$ — сл.вектор с одинаковыми коэффициентами корреляции $\rho = \mathbf{Corr}(\xi_i, \xi_j)$, $i, j = \overline{1, k}$. Доказать, что $\rho \geqslant -\frac{1}{(k-1)}$.

Упр. В.7. Пусть F, G — две одномерные ф.р. Показать, что функции $H_1(x, y) = \min\big\{F(x), F(y)\big\}$ и $H_2(x, y) = F(x)G(y)\big(1 + c(1 - F(x))(1 - G(y))\big)$, где $|c| < 1$, являются ф.р. и их маргинальные ф.р. совпадают с F и G.

Упр. В.8. Пусть $\mathbf{D}\big[\xi \mid \eta\big]$ — условная дисперсия сл.в. ξ при фиксированной сл.в. η. Покажите, что $\mathbf{D}\xi = \mathbf{E}\big(\mathbf{D}\big[\xi \mid \eta\big]\big) + \mathbf{D}\big(\,\mathbf{E}\big[\,\xi \mid \eta\,\big]\big)$.

Упр. В.9. Пусть $\xi_1 \backsim \xi_2$. Верно ли, что $\eta\xi_1 \backsim \eta\xi_2$ для любой сл.в. η?

Упр. В.10. Описать класс сл.в., не зависящих от самих себя.

Упр. В.11. А) Пусть m — медиана сл.в. ξ, т.е. $\mathbf{P}\{\xi \leqslant m\} \geqslant \frac{1}{2} \leqslant \mathbf{P}\{\xi \geqslant m\}$. Показать, что $\mathbf{P}\{\xi < m\} \leqslant \frac{1}{2} \geqslant \mathbf{P}\{\xi > m\}$ и при $\mathbf{E}\big|\xi\big| < \infty$

$$\inf_{c\,\in\mathbb{R}} \mathbf{E}\big|\xi - c\big| = \mathbf{E}\big|\xi - m\big|.$$

Б) Пусть $\mathbf{D}\xi < \infty$, $\mu = \mathbf{E}\,\xi$. Показать, что $\inf\limits_{c\,\in\mathbb{R}} \mathbf{E}\left(\xi - c\right)^2 = \mathbf{E}\left(\xi - \mu\right)^2$.

C. Основные вероятностные законы

Дискретные распределения

✦ Биномиальное — $\mathcal{B}\mathtt{in}(n,p)$, $p \in (0;1), n \in \mathbb{N}$:

$$\mathbf{P}\{\xi = k\} = \mathrm{C}_n^k\, p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n, \qquad \mathbf{E}\xi = np, \quad \mathbf{D}\xi = np(1-p).$$

✦ Гипергеометрическое — $\mathcal{H}\mathtt{g}(n, N_1, N_2)$, $1 \leqslant n \leqslant N = N_1 + N_2$:

$$\mathbf{P}\{\xi = k\} = \frac{\mathrm{C}_{N_1}^{k}\, \mathrm{C}_{N_2}^{n-k}}{\mathrm{C}_N^n}, \quad \max\{0, n - N_2\} \leqslant k \leqslant \min\{N_1, n\}, \quad p = N_1/N,$$

$$\mathbf{E}\xi = np, \quad \mathbf{D}\xi = np(1-p)\frac{N-n}{N-1}.$$

✦ Геометрическое — $\mathcal{G}\mathtt{eo}(p)$, $p \in (0;1)$:

$$\mathbf{P}\{\xi = k\} = p(1-p)^{k-1}, \quad k = 1, 2, \ldots, n, \qquad \mathbf{E}\xi = \frac{1}{p}, \quad \mathbf{D}\xi = \frac{(1-p)}{p^2}.$$

✦ Пуассона — $\mathcal{P}\mathtt{ois}(\lambda)$, $\lambda > 0$:

$$\mathbf{P}\{\xi = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k \in \mathbb{N}, \qquad \mathbf{E}\xi = \lambda, \quad \mathbf{D}\xi = \lambda.$$

Абсолютно непрерывные распределения

✦ Равномерное — $\mathcal{U}\mathtt{n}(a,b)$, $a < b,\ a, b \in \mathbb{R}^1$:

$$f(x) = \frac{1}{b-a}, \quad x \in [a;b], \qquad \mathbf{E}\xi = \frac{b+a}{2}, \quad \mathbf{D}\xi = \frac{(b-a)^2}{12}.$$

✦ Бета — $\mathcal{B}\mathtt{et}(\alpha_1, \alpha_2)$, $\alpha_1, \alpha_2 > 0$: $(x \in (0;1))$

$$f(x) = \frac{x^{\alpha_1 - 1}(1-x)^{\alpha_2 - 1}}{\mathrm{B}(\alpha_1, \alpha_2)}, \qquad \mathbf{E}\xi = \frac{\alpha_1}{\alpha_1 + \alpha_2}, \ \mathbf{D}\xi = \frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)^2(\alpha_1 + \alpha_2 + 1)}.$$

✦ Показательное — $\mathcal{E}\mathtt{x}(\lambda)$, $\lambda > 0$:

$$f(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}}, \quad x \geqslant 0, \qquad \mathbf{E}\xi = \lambda, \quad \mathbf{D}\xi = \lambda^2.$$

✦ Лапласа (двустороннее показательное) — $\mathcal{L}\mathtt{ap}(\lambda)$, $\lambda > 0$:

$$f(x) = \frac{1}{2\lambda} e^{-\frac{|x|}{\lambda}}, \quad x \in \mathbb{R}^1, \qquad \mathbf{E}\xi = 0, \quad \mathbf{D}\xi = 2\lambda^2.$$

✦ Гамма — $\mathcal{G}\mathtt{am}(a, \lambda)$, $a, \lambda > 0$:

$$f(x) = \frac{1}{\Gamma(a)\lambda^a} x^{a-1} e^{-\frac{x}{\lambda}}, \quad x > 0, \qquad \mathbf{E}\xi = a\lambda, \quad \mathbf{D}\xi = a\lambda^2.$$

✦ Нормальное (гауссовское) — $\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}^1, \sigma^2 > 0$:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big\{ -\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\Big\}, \quad x \in \mathbb{R}^1, \qquad \mathbf{E}\xi = \mu, \quad \mathbf{D}\xi = \sigma^2.$$

✦ Коши — $\mathcal{C}\mathtt{auch}(\mu, \sigma)$, $\mu \in \mathbb{R}^1, \sigma > 0$:

$$f(x) = \frac{1}{\pi\sigma}\,\frac{1}{1 + \big(\frac{x-\mu}{\sigma}\big)^2}, \quad x \in \mathbb{R}^1, \qquad \mathbf{E}\xi^{\pm} = \infty.$$

✦ Стьюдента (стр. 124) — $\mathcal{S}\mathtt{tud}(k)$, $k = 1, 2, \ldots,$ $x \in \mathbb{R}^1$,

$$f(x) = \frac{\Gamma((k+1)/2)}{\sqrt{\pi k}\,\Gamma(k/2)} \cdot \Big(1 + \frac{x^2}{k}\Big)^{-\frac{1}{2}(k+1)}, \quad \mathbf{E}\xi = 0 \ (k > 1), \ \mathbf{D}\xi = \frac{k}{k-2} \ (k > 2).$$

✦ Хи-квадрат — $\mathcal{K}\mathtt{hi}(k)$, $k = 1, 2, \ldots,$

$$f(x) \backsim \mathcal{G}\mathtt{am}\Big(\frac{k}{2}, 2\Big), \qquad \mathbf{E}\xi = k, \quad \mathbf{D}\xi = 2k.$$

Многомерные распределения

✦ Гипергеометрическое — $\mathcal{H}\mathtt{g}_k(n; N_1, \ldots, N_k)$ — стр. 32, 42, 80.

✦ Дирихле (многомерное бета) — $\mathcal{D}\mathtt{ir}_k(\alpha_1, \ldots, \alpha_{k+1})$ — стр. 70.

✦ Мультиномиальное — $\mathcal{M}\mathtt{ult}_k(n, q_1, \ldots, q_k)$ — стр. 33, 42, 80.

✦ Нормальное — $\mathcal{N}_2(\vec{\mu}, \mathbf{\Sigma})$ — стр. 181.

✦ Пуассона — $\mathcal{P}\mathtt{ois}_k(\lambda_0, \lambda_1, \ldots, \lambda_k)$ — стр. 122.

✦ Равномерное — $\mathcal{U}\mathtt{n}(\mathscr{X})$ — стр. 70.

D. Указания к решению задач

✦ 8, стр. 30. Функция $\xi : (\Omega, \mathscr{F}) \mapsto (X, \mathscr{X})$ измерима относительно σ-алгебры $\mathscr{X} = \sigma(\mathcal{U})$, порождённой некоторым классом подмножеств $\mathcal{U}$, т. т. т. когда для $\forall B \in \mathcal{U}$ прообраз $\xi^{-1}(B) \in \mathscr{F}$. Для каждого из пунктов леммы воспользоваться указанными ссылками со специфическим классом $\mathcal{U}$.

✦ 10, стр. 33. а) Гипергеометрическая модель: общее число исходов — C_N^n (из N объектов выбирается n), в «благоприятные» попадают любые n_1 из первой группы ($C_{N_1}^{n_1}$ вариантов), ..., n_k из k-й группы ($C_{N_k}^{n_k}$ вариантов). Мультиномиальная модель: вероятность конкретной последовательности исходов при n испытаниях, скажем, $(R_1, R_3, R_1, \ldots, R_2)$, равна $q_1 q_3 q_1 \cdots q_2 = q_1^{n_1} \cdots q_k^{n_k}$. Учесть, что среди $n!$ перестановок исходов с тем же составом будут совпадающие, т.к. перестановки внутри группы с одним исходом неотличимы между собой.

б) Доказать, что $M^{-m}M!/(M-m)! \to 1$ при $M \to \infty$ и фиксированном $m > 0$.

✦ 15, стр. 37. $\mathbf{P}\{a < \xi < c\} = \lim_n \mathbf{P}\{a < \xi \leqslant c - 1/n\} = \lim_n F(c-1/n) - F(a)$.

✦ 17, стр. 38. Найти прямоугольник $(\vec{a};\vec{c}]$, для которого $\Delta_{\vec{a};\vec{c}}G < 0$.

✦ 19, стр. 40. Воспользоваться (F_4). (F_2) справедливо для любой ф.р.

✦ 21, стр. 41. $\mathbf{P}\{\,\xi_1 \leqslant x_1\,\} = \mathbf{P}\{\,\xi_1 \leqslant x_1, \xi_2 \in \mathbb{R}^1\,\}$.

✦ (8), стр. 42. $\mathbf{P}\{\xi_1 = n\} = \sum_{m \in \mathscr{X}_2} \mathbf{P}\{\xi_1 = n, \xi_2 = m\}$.

✦ 23, стр. 43. Оценить количество точек $\bigcup_n \{x \in \mathbb{R}^1 : \mathbf{P}\{\,\xi = x\,\} \geqslant 1/n\}$.

✦ 25, стр. 44. i) Границу углового множества представить в виде монотонного предела $\lim\limits_{z\downarrow 0}{}_{\downarrow} \{\vec{x}:\ \vec{x} \leqslant \vec{x}_0\} \setminus \{\vec{x}:\ \vec{x} \leqslant \vec{x}_0 - z\}$.

ii) $|F(x_1,y_1) - F(x_2,y_2)| \leqslant F(x_{(2)},y_{(2)}) - F(x_{(1)},y_{(1)}) \leqslant [F_\xi(x_{(2)}) - F_\xi(x_{(1)})] + [F_\eta(y_{(2)}) - F_\eta(y_{(1)})]$, где $a_{(2)} = \max\{a_1,a_2\}, a_{(1)} = \min\{a_1,a_2\}$.

iii) Если $\mathbf{P}\{\xi = a\} = q > 0 \Rightarrow F(x,y) = \mathbf{P}\{\xi \leqslant x, \eta \leqslant y\} < \mathbf{P}\{\xi \leqslant a\} - q,\ \forall x < a, \forall y,$ и $\exists b:\ F(a,b) = \mathbf{P}\{\xi \leqslant a, \eta \leqslant b\} > \mathbf{P}\{\xi \leqslant a\} - q/2$.

✦ (11), стр. 47. Проверить свойства вероятности.

✦ 31, стр. 48. Так как g_j измеримо, то $g_j^{-1}(B_j) \in \mathscr{B}^{k_j},\ \forall B_j \in \mathscr{B}^{m_j}$; далее применить условие независимости векторов $\vec{\eta}_1, \ldots, \vec{\eta}_n$.

✦ 34, стр. 49. $\forall A_j \in \sigma(\vec{\xi}_j)\ \exists B_j \in \mathscr{B}^{m_j}:\ A_j = \vec{\xi}_j^{\,-1}(B_j)$.

✦ (?!), стр. 50. Показать: $\mathbf{P}\{(A_1 \uplus A_2)B\} = \mathbf{P}\{A_1B\} + \mathbf{P}\{A_2B\} =$ (независимость) $= \mathbf{P}\{A_1\}\,\mathbf{P}\{B\} + \mathbf{P}\{A_2\}\,\mathbf{P}\{B\} = \mathbf{P}\{A_1 \uplus A_2\}\,\mathbf{P}\{B\}$.

✦ 44, стр. 55. Заметить, что функции $\mathbf{P}\{\xi > x\}$ и $\mathbf{P}\{\xi \geqslant x\}$ отличаются на не более чем счётном множестве $x \in \mathbb{R}^1$.

✦ 46, стр. 55. Применить формулу интегрирования по частям к $\int_a^b x^q\,dF$ при достаточно большом $a > 0$; устремить $b \to \infty$.

✦ 50, стр. 58. Пусть $c,\ c \notin \mathscr{X}$, и $x_{i_j} \in (c-1;c) \cap \mathscr{X}$. Если $a \nearrow c$, то $\exists j_m \to \infty$: $F(c) - F(a) = \sum\limits_{j \geqslant j_m} p_{i_j} \to 0$ по свойству сходящихся рядов. То же самое верно для $\forall\, a$ и $c \searrow a$. Для $\forall\, c = x_j \in \mathscr{X}$ разность $F(c) - F(a) = p_j + \sum\limits_{a < x_k < c} p_k$.

✦ 61, стр. 66. Воспользоваться доказательством свойства (F_3) для ф.р.

✦ (18), стр. 67. (1) $\mathbf{P}\{\xi = \eta\} = 0$; (2) $(\xi,\eta) \backsim (\eta,\xi) \Rightarrow \mathbf{P}\{\xi < \eta\} = \mathbf{P}\{\xi > \eta\}$.

✦ 69, стр. 74. Применить свойство линейности математического ожидания.

✦ 73, стр. 75. б) $\mathbf{D}(\xi_2 + (\xi_1 - \xi_2)) = \mathbf{D}\xi_2 + \mathbf{D}(\xi_1 - \xi_2)$.

✦ (25), стр. 75. $\mathbf{Cov}(\xi_2, \xi_1 - \rho\sigma_1\xi_2/\sigma_2) = \sigma_{12} - \rho\sigma_1\sigma_2 = 0$.

✦ 74, стр. 76. В матрице $\mathbf{E}\,[A\Xi]$ элемент на i-й строке и j-м столбце равен $\mathbf{E}[a_{i1}\xi_{1j} + \ldots + a_{ik}\xi_{kj}] = a_{i1}\,\mathbf{E}\,\xi_{1j} + \ldots + a_{ik}\,\mathbf{E}\,\xi_{kj}$, т.е. равен (i,j) элементу матрицы $A\,\mathbf{E}\,[\Xi]$.

✦ 80, стр. 80. (R_2) Для вычисления коэффициента корреляции надо найти дисперсии компонент и их ковариацию. Из некоррелированности остатка $\zeta - Z(\vec{\xi})$ и компонент $\vec{\xi}$ вывести, что ковариация $\mathbf{Cov}(\zeta, Z(\vec{\xi})) = \mathbf{D}Z(\vec{\xi})$. Найти вид этой дисперсии из предшествующей теоремы. Свойство (R_4) и неравенство $\rho_{\zeta\cdot\xi} \leqslant 1$ следуют из свойств классического коэффициента корреляции. Свойство (R_3) следует из (R_2) (квадратичная форма с невырожденной матрицей может равняться нулю только при нулевом векторе ковариаций).

✦ 82, стр. 83. Использовать то, что $\mathbf{Corr}\left(\vec{a}^\flat\vec{\xi}, \vec{c}^\flat\vec{\zeta}\right) = -\mathbf{Corr}\left(\vec{a}^\flat\vec{\xi}, -\vec{c}^\flat\vec{\zeta}\right)$.

✦ (37), стр. 95. В записи $\mathbf{E}\left[g(\eta)h(\xi,\eta)\right]$ через совместную плотность воспользоваться условной плотностью; применить теорему Фубини–Тонелли.

✦ В.1. a) Применить тождество $\max(a,b) = \frac{1}{2}\,(a+b+|a-b|)$ и неравенство Коши–Буняковского; b) следует из a) и неравенства Маркова.

✦ В.2. a) Да; b) Не обязательно, например, $\xi = \alpha_1\alpha_2, \eta = \alpha_1\alpha_3, \zeta = \alpha_2\alpha_3$, где сл.в. $\alpha_1, \alpha_2, \alpha_3$ независимы и имеют нулевые математические ожидания.

✦ В.3. Можно считать $\mathbf{E}\xi_1 = 0$; показать, что в $n^2\overline{\xi}S^2 = \sum_j \xi_j\left(\sum_k \xi_k^2 - n\sum_i \xi_i \sum_m \xi_m\right)$ все слагаемые имеют нулевое математическое ожидание.

✦ В.4. Равномерное распределение в прямоугольнике $(-1,0)$, $(0,-1)$, $(1,0)$, $(0,1)$.

✦ В.5. По условию $\mathbf{E}\left[h(\vec{\xi})\right] = \mathbf{E}\left[h(Q\vec{\xi})\right]$ для любой интегрируемой функции h. Выбрать здесь Q (при соответствующих h) так, что $x_1 \dashrightarrow -x_1$.

✦ В.6. Определитель матрицы корреляций равен $(1-\rho)^{k-1}(1+(k-1)\rho)$. Другой способ: $0 \leqslant \mathbf{E}\left[\sum_1^k \xi_j\right]^2 = \sum_1^k \mathbf{D}\xi_j + \sum_{i\neq j}\mathbf{E}\left[\xi_i\xi_j\right] = k + k(k-1)\rho$, если $\mathbf{E}\xi_j = 0$, $\mathbf{D}\xi_j = 1$.

✦ В.7. $F(x) = H(x, +\infty)$.

✦ В.8. Применить формулу для маргинальной плотности.

✦ В.9. Не обязательно: $\eta = \xi_1$ и ξ_1, ξ_2 — независимы.

✦ В.10. $\mathbf{P}\{\xi = const\} = 1$.

✦ В.11. А) $\mathbf{P}\{\xi < m\} = 1 - \mathbf{P}\{\xi \geqslant m\} \geqslant 1 - \frac{1}{2}$. Пусть $c > m$, тогда $|x-c| - |x-m| = (c-m)\,\dot{\mathrm{I}}(x \leqslant m) + (c+m-2x)\,\dot{\mathrm{I}}(m < x \leqslant c) - (c-m)\,\dot{\mathrm{I}}(x > c) \geqslant (c-m)\left(\dot{\mathrm{I}}(x \leqslant m) - \dot{\mathrm{I}}(x > m)\right)$. Показать отсюда, что $\mathbf{E}\left[|\xi - c| - |\xi - m|\right] \geqslant 0$.

Б) $\mathbf{E}(\xi - c)^2 = \mathbf{E}(\xi - \mu)^2 + (\mu - c)^2$.

II Преобразования случайных величин

В некоторых классических учебниках теория вероятностей понимается как математическая наука, «позволяющая по вероятностям одних случайных событий находить вероятности других случайных событий, связанных каким-либо способом с первыми» (С. Н. Бернштейн). В этой главе мы рассмотрим некоторые методы отыскания распределений функций от случайных величин с известным распределением.

§ 1. Индуцированное распределение

Пусть $\vec{\xi} = (\xi_1, \ldots, \xi_k)$ — случайный вектор, для которого известен закон распределения, т.е. известны вероятности $\mathsf{P}_{\vec{\xi}}\{A\} \stackrel{\mathrm{df}}{=} \mathbf{P}\{\vec{\xi} \in A\}$ для всех борелевских подмножеств $A \subset \mathbb{R}^k$. Рассмотрим некоторую борелевскую функцию $h : (\mathbb{R}^k, \mathscr{B}^k) \mapsto (\mathbb{R}^m, \mathscr{B}^m)$ и определим m-мерный сл.вектор

$$\vec{\eta} = (\eta_1, \ldots, \eta_m) = h(\vec{\xi}) = \big(h_1(\vec{\xi}), \ldots, h_m(\vec{\xi}) \big).$$

Отображение h порождает (индуцирует) вероятностный закон (?!) на $(\mathbb{R}^m, \mathscr{B}^m)$ по формуле

$$\mathsf{P}_{\vec{\eta}}\{B\} = \mathsf{P}_{\vec{\xi}}\{h^{-1}(B)\}, \qquad B \in \mathscr{B}^m.$$

Этот закон и есть распределение сл.вектора $\vec{\eta}$. Поскольку привычный способ представления любого закона распределения подразумевает описание его функции распределения (ф.р.) или функции плотности (ф.пл.), естественно, хотелось бы описать распределение $\vec{\eta}$ аналогичным образом.

Анализ способов построения ф.р. начнём с простейшего преобразования. Интегралы, возникающие в процессе изложения, будем понимать как лебеговские, подразумевая возможность их вычисления, если понадобится, по Риману (см. 566, стр. 497).

100 Пример. Пусть $F_\xi(x)$ — ф.р. сл.величины ξ и $h(x) = bx + a$ с некоторыми $a \in \mathbb{R}^1$ и $b > 0$. Заметим, что обратная функция $h^{-1}(y) = \frac{1}{b}(y-a)$, поэтому неравенство $h(x) \leqslant y$ эквивалентно $x \leqslant \frac{1}{b}(y-a)$. Следовательно, ф.р. сл.величины $\eta = b\xi + a$

$$F_\eta(y) \overset{\text{df}}{=} \mathbf{P}\{\eta \leqslant y\} = \mathbf{P}\big\{\xi \leqslant \tfrac{1}{b}(y-a)\big\} = F_\xi\big(\tfrac{1}{b}(y-a)\big), \quad y \in \mathbb{R}^1.$$

Если распределение ξ имеет плотность $f_\xi(x)$ с носителем $\mathscr{X}_\xi = (L; R)$, с конечными или бесконечными границами, то

$$F_\eta(y) = F_\xi\big(\tfrac{1}{b}(y-a)\big) = \int_{-\infty}^{\frac{1}{b}(y-a)} f_\xi(x)\,\dot{\mathrm{I}}_{(L;R)}(x)\,dx.$$

Замена переменной $x = h^{-1}(t) = \frac{1}{b}(t-a)$, $-\infty < t < y$, в этом интеграле переводит индикаторную функцию $\dot{\mathrm{I}}_{(L;R)}(x)$ в индикаторную функцию $\dot{\mathrm{I}}_{(U;V)}(x)$, где $U = bL + a$ $(= \lim_{x\to L}(bx+a))$ и $V = bR + a$. В силу возможности линейной замены в интеграле Лебега (см. 543, стр. 484) получаем

$$F_\eta(y) = \int\limits_{-\infty}^{y} \underbrace{\frac{1}{b}\, f_\xi\big(\tfrac{1}{b}(t-a)\big)\,\dot{\mathrm{I}}_{(U;V)}(t)}_{f_\eta(t)}\,dt.$$

Следовательно, плотность $\eta = b\xi + a$ равна $f_\eta(t) = \frac{1}{b} f_\xi\big(\frac{1}{b}(t-a)\big)$, $t \in (U; V)$.

Так, если $\xi \backsim \mathcal{N}(0,1)$, то $\sigma\xi + \mu \backsim \mathcal{N}(\mu, \sigma^2)$. ◎

101 Упр. Как изменятся ф.р. и ф.пл. $b\xi + a$, если $b < 0$?

✧ **Плотность монотонного преобразования.** $\eta = h(\xi)$ можно найти по аналогичной схеме — сначала ф.р. η (с возрастающей функцией h) представляется через интеграл от плотности: $F_\eta(y) = \mathbf{P}\{\eta \leqslant y\} = \mathbf{P}\big\{\xi \leqslant h^{-1}(y)\big\} = \int_{-\infty}^{h^{-1}(y)} f_\xi(t)\,dt$; затем производится замена $t = \widetilde{h}(u) = h^{-1}(u)$, $u \leqslant y$, каковая приводит ф.р. к виду $F_\eta(y) = \int_{-\infty}^{y} f_\xi(\widetilde{h}(u))\widetilde{h}'(u)\,du$ (с учётом якобиана $dt = \widetilde{h}'(u)du$). Подынтегральное выражение и даст плотность η.

Следующая теорема содержит формальное обоснование этой схемы. Чтобы избежать излишних подробностей, под носителем распределения будем понимать любой открытый интервал, вне которого плотность равна нулю.

102| **Теорема.** *Пусть* $f_\xi(x)$ *— плотность распределения сл.в.* ξ *с носителем* $\mathscr{X}_\xi = (A;B)$, $h:(A;B)\mapsto(C;D)$ *— строго монотонная непрерывная функция с обратной функцией* $\widetilde{h}(y) = h^{-1}(y)$, *определённой при* $\forall\, y\in(C;D)$. *Если* $\widetilde{h}$ *всюду непрерывно дифференцируема, то сл.в.* $\eta = h(\xi)$ *имеет распределение абсолютно непрерывного типа с ф.пл.*

$$f_\eta(y) = \bigl|(h^{-1}(y))'\bigr|\, f_\xi\bigl(h^{-1}(y)\bigr), \qquad y\in(C;D).$$

⊫ Пусть h строго возрастает. Для любого конечного интервала $(u;v]\subset\subset(C;D)$ в силу строгой монотонности и непрерывности (а значит, и взаимной однозначности) обратной функции $\widetilde{h} = h^{-1}$ имеем

$$\mathbf{P}\{u<\eta\leqslant v\} \;=\; \mathbf{P}\bigl\{\,\widetilde{h}(u)<\xi\leqslant\widetilde{h}(v)\,\bigr\} \;=\; \int\limits_{\widetilde{h}(u)}^{\widetilde{h}(v)} f_\xi(x)\,dx \;=\; \int\limits_{\widetilde{h}(u)}^{\widetilde{h}(v)} f_\xi\bigl(\widetilde{h}(h(x))\bigr)\,dx.$$

Замена $t = h(x)$ по формуле замены переменных в интеграле Лебега (542, стр. 483) переводит этот интеграл относительно меры Лебега λ в интеграл относительно индуцированной меры μ_h, которая по определению на любом конечном интервале $(k;m]$ вычисляется как

$$\mu_h(k;m] \;=\; \lambda\bigl(h^{-1}(k);h^{-1}(m)\bigr] \;=\; \widetilde{h}(m)-\widetilde{h}(k).$$

Другими словами, мера μ_h есть мера Лебега–Стилтьеса с генерирующей функцией $\widetilde{h}$. Так как функция $\widetilde{h}$ непрерывно дифференцируема, то на конечном отрезке $(u;v]$ она абсолютно непрерывна. Поэтому мера μ_h абсолютно непрерывна относительно меры Лебега с плотностью $\widetilde{h}'$. Следовательно,

$$\mathbf{P}\{u<\eta\leqslant v\} \;=\; \int_u^v f_\xi\bigl(\widetilde{h}(t)\bigr)\,\mu_h(dt) \;=\; \int_u^v f_\xi\bigl(\widetilde{h}(t)\bigr)\,\widetilde{h}'(t)\,dt.$$

Такое представление ф.р. в виде интеграла эквивалентно абсолютной непрерывности распределения η с заявленной в теореме плотностью. Случай убывающей функции h аналогичен рассмотренному. ⫤

103| Пример. Пусть $\xi \backsim \mathfrak{U}\mathrm{n}(-1,1)$ с плотностью $f_\xi(x) = {}^1\!/_2$ при $x\in\mathscr{X}_\xi = (-1;1)$. Функция $h(x) = \sin(\pi x/2)$ строго возрастает на носителе $\mathscr{X}_\xi$; при этом $\mathscr{Y} = h(\mathscr{X}_\xi) = (-1;1)$. Обратная функция $\widetilde{h}(y) = h^{-1}(y) = 2\arcsin(y)/\pi$, а производная $\widetilde{h}'(y) = 2/(\pi\sqrt{1-y^2})$ при $y\in(-1;1)$. Поэтому функция плотности сл.в. $\eta = \sin(\pi\xi/2)$ равна $f_\eta(y) = 2\widetilde{h}'(y)$, $-1<y<1$.

Напомним, что в точках, которые не были указаны при описании плотности, её значение предполагается равным нулю. ◎

Метод доказательства предыдущей теоремы может быть применён и к кусочно-монотонным функциям.

104| Пример. Пусть $\xi \backsim \mathfrak{Un}(-1,1)$ с ф.плотности $f(x) = {}^1\!/_2, x \in$ $\in [-1;1]$. Функция $h(x) = x^2$ переводит отрезок $[-1;1]$ в $[0;1]$. Для точек $u, v \in (0;1)$, $u \leqslant v$, вероятность

$$\mathbf{P}\{u \leqslant \xi^2 \leqslant v\} = \mathbf{P}\{-\sqrt{v} \leqslant \xi \leqslant -\sqrt{u}\} + \mathbf{P}\{\sqrt{u} \leqslant \xi \leqslant \sqrt{v}\} =$$

$$= \frac{1}{2}\int_{-\sqrt{v}}^{-\sqrt{u}} dx + \frac{1}{2}\int_{\sqrt{u}}^{\sqrt{v}} dx = \int_{\sqrt{u}}^{\sqrt{v}} dx.$$

Замена $x = \sqrt{t}$, $t \in [u;v]$, в силу её непрерывной дифференцируемости на интервалах, не содержащих ноль, даёт

$$\mathbf{P}\{u \leqslant \xi^2 \leqslant v\} = \int_u^v \frac{1}{2\sqrt{t}}\, dt.$$

Таким образом, плотность $\eta = \xi^2$ равна $1/(2\sqrt{y})$, $0 < y < 1$. ◎

Рассмотренный в этом примере приём формализует следующая

105| **Лемма.** *Пусть $f_\xi(x)$ — плотность распределения сл.в. ξ с носителем $\mathscr{X}_\xi = (A;B)$ и $h(x) = \sum_{k=1}^n h_k(x)\,\dot{\mathrm{I}}(x\,;\,J_k)$ — кусочно строго монотонная функция, где непересекающиеся интервалы $J_k \subset (A;B)$, $k = \overline{1,n}$, и $\bigcup_1^n [J_k] = [A;B]$, функции h_k строго монотонны и непрерывно дифференцируемы с обратными $h_k^{-1}(y)$, определёнными при $\forall\, y \in h_k(J_k)$. Тогда плотность f_η распределения сл.в. $\eta = h(\xi)$ можно вычислить как*

$$f_\eta(y) = \sum_{k=1}^n \left|\big(h_k^{-1}(y)\big)'\right| f_\xi\big(h_k^{-1}(y)\big)\ \dot{\mathrm{I}}\ \big(y\,;\,h_k(J_k)\big), \qquad y \in h\big((A;B)\big).$$

Другой способ нахождения плотности использует производную ф.р. Предположим, что каким-то образом нам удалось представить ф.р. сл.в. $\eta = h(\xi)$ в виде $F_\eta(y) = F_0(\widetilde{h}(y))$ с некоторой функцией $\widetilde{h}$ и абсолютно непрерывной функцией F_0, производная которой равна f_0. Тогда, по правилу дифференцирования сложной функции, производная

$$\frac{d}{dy} F_\eta(y) = f_0(\widetilde{h}(y))\widetilde{h}'(y).$$

Полученное выражение будет плотностью распределения η, если функция $\widetilde{h}$ непрерывно дифференцируема. В случае кусочной непрерывности f и (или) $\widetilde{h}'$ абсолютную непрерывность F_η можно доказать, проверив равенство $\int_{\mathbb{R}} \left(\frac{d}{dy} F_\eta(y)\right) dy = 1$ (см. 60, стр. 63).

106| Пример. Пусть $\xi \backsim \mathcal{N}(0,1)$, найдём распределение $\chi_1^2 = \xi^2$. Искомая ф.р. (при $y > 0$)

$$F(y) \ \stackrel{\mathrm{df}}{=} \ \mathbf{P}\left\{\xi^2 \leqslant y\right\} \ = \ \Phi(\sqrt{y}\,) - \Phi(-\sqrt{y}\,) \ = \ 2\Phi(\sqrt{y}\,) - 1.$$

Ясно, что $F(y) = 0$ при $y \leqslant 0$. Функция F, очевидно, всюду непрерывна, и её производная (при $y > 0$) равна

$$F'(y) \ = \ \frac{1}{\sqrt{2\pi}} y^{\frac{1}{2}-1} \exp\left(-\frac{1}{2}y\right) \ = \ \frac{(1/2)^{1/2}}{\Gamma(1/2)} y^{\frac{1}{2}-1} \exp\left(-\frac{1}{2}y\right).$$

Необходимость в проверке равенства $\int_{\mathbb{R}} F'(y)\,dy = 1$ здесь отпадает, т.к. F' совпадает с плотностью хорошо известного гамма-распределения $\mathcal{G}\mathtt{am}\left(\frac{1}{2}, 2\right)$ с параметром формы $\frac{1}{2}$ и параметром масштаба 2 (см. стр. 107). Такое распределение называется хи-квадрат распределением с одной степенью свободы (об общем хи-квадрат распределении см. стр. 167). ⊚

✧ **Плотность функции вектора случайных величин.** Аналогичные схемы используются и при многомерных преобразованиях. Различные варианты (достаточно сложного) доказательства следующей теоремы можно найти в [11], [12], [24]. Случай линейного преобразования рассмотрен на стр. 130.

107| **Теорема.** *Пусть отображение $h : \mathbb{R}^k \mapsto \mathbb{R}^k$ взаимно однозначным образом переводит носитель сл.вектора $\vec{\xi}$ в множество $\mathscr{Y} \subset \mathbb{R}^k$; при этом каждая компонента обратной функции $h^{-1}(\vec{y}) = (\widetilde{h}_1(\vec{y}), \ldots, \widetilde{h}_k(\vec{y}))$ непрерывно дифференцируема во внутренней части $\mathscr{Y}^o$ и якобиан функции h^{-1} строго положителен:*

$$J_{h^{-1}}(\vec{y}) \ = \ \left\|\left(\frac{\partial \widetilde{h}_j}{\partial y_l}\right)_{j,l}\right\| > 0 \qquad \text{для} \ \ \forall\, \vec{y} \in \mathscr{Y}^o.$$

Если $f_{\vec{\xi}}(\vec{x})$ — плотность распределения вектора $\vec{\xi} = (\xi_1, \ldots, \xi_k)$, то плотность распределения сл.вектора $\vec{\eta} = h(\vec{\xi})$ равна

$$f_{\vec{\eta}}\,(\vec{y}) \ = \ J_{h^{-1}}(\vec{y})\, f_{\vec{\xi}}\big(h^{-1}(\vec{y})\big), \qquad \vec{y} \in \mathscr{Y}^o.$$

108| Примеры. 1) Пусть $\xi_1, \xi_2 \backsim \mathcal{E}\mathrm{x}(1)$ — независимые показательные сл.в. Найдём совместное распределение сл.в. $\eta_1 = \xi_1 + \xi_2$, $\eta_2 = \xi_1/(\xi_1 + \xi_2)$. Преобразование $y_1 = x_1 + x_2$, $y_2 = x_1/(x_1 + x_2)$ переводит носитель $\mathscr{X} = \mathbb{R}^1_+ \times \mathbb{R}^1_+$ вектора (ξ_1, ξ_2) в область $\mathscr{Y} = \mathbb{R}^1_+ \times (0;1)$. Якобиан обратного преобразования $x_1 = y_1 y_2$, $x_2 = y_1 - y_1 y_2$, $0 < y_1 < \infty$, $0 < y_2 < 1$,

$$J_{H^{-1}} = \left\| \begin{vmatrix} y_2 & y_1 \\ 1 - y_2 & -y_1 \end{vmatrix} \right\| = y_1 > 0.$$

Так как плотность $\vec{\xi} = (\xi_1, \xi_2)$ равна $f_{\vec{\xi}}(x_1, x_2) = e^{-(x_1+x_2)}$, то

$$f_{\vec{\eta}}(y_1, y_2) = J_{H^{-1}} \, e^{-(x_1+x_2)} = y_1 e^{-y_1} \cdot 1, \qquad y_1 > 0, \ 0 < y_2 < 1.$$

Другими словами, сл.в. η_1, η_2 независимы, $\eta_1 \backsim \mathcal{G}\mathrm{am}(2,1)$ с плотностью $f_{\eta_1}(y_1) = y_1 e^{-y_1}$, а $\eta_2 \backsim \mathcal{U}\mathrm{n}(0,1)$ с плотностью $f_{\eta_2}(y_2) = 1$.

2) [*Генерирование нормальных сл.в.*] Пусть $\xi_1, \xi_2 \backsim \mathcal{U}\mathrm{n}(0,1)$ — независимые равномерные сл.в. Покажем, что сл.в.

$$\eta_1 = \sqrt{-2 \ln \xi_1} \cos(2\pi\xi_2), \qquad \eta_2 = \sqrt{-2 \ln \xi_1} \sin(2\pi\xi_2)$$

также независимы и имеют стандартное нормальное $\mathcal{N}(0,1)$ распределение. Можно считать, что $\xi_1, \xi_2 \in (0;1)$, тогда обратное преобразование:

$$x_1 = \exp\left(-\frac{1}{2}(y_1^2 + y_2^2)\right), \qquad x_2 = \frac{1}{2\pi}\mathrm{Atan}(y_2, y_1), \qquad y_1, y_2 \neq 0,$$

где функция $\mathrm{Atan}(y_2, y_1) = \arctan(y_2/y_1)$, если $y_2, y_1 > 0$, $\mathrm{Atan}(y_2, y_1) = \arctan(y_2/y_1) + 2\pi$, если $y_2 > 0, y_1 < 0$, $\mathrm{Atan}(y_2, y_1) = \arctan(y_2/y_1) + \pi$, если $y_2 < 0$. Условия теоремы 107 выполнены, т.к. якобиан обратного преобразования не обращается в ноль в области $(y_1, y_2) \in \mathbb{R}^2 \setminus (0,0)$, ибо

$$J = \left\| \begin{vmatrix} -y_1 e^{-\frac{1}{2}(y_1^2+y_2^2)} & -y_2 e^{-\frac{1}{2}(y_1^2+y_2^2)} \\ -\frac{y_2}{2\pi(y_1^2+y_2^2)} & \frac{y_1}{2\pi(y_1^2+y_2^2)} \end{vmatrix} \right\| = \frac{1}{2\pi} e^{-\frac{1}{2}(y_1^2+y_2^2)} > 0.$$

Плотность вектора (ξ_1, ξ_2) равна 1 при $0 < x_1, x_2 < 1$, поэтому плотность (η_1, η_2) равна $\frac{1}{2\pi} e^{-\frac{1}{2}(y_1^2+y_2^2)}$, что и требовалось. ◎

Особенно просто находится плотность линейного преобразования.

109| **Лемма.** (?!) *Пусть Q — невырожденная $(k \times k)$-матрица и вектор $\vec{a} \in \mathbb{R}^k$. Если распределение сл.вектора $\vec{\xi}$ имеет плотность $f_{\vec{\xi}}(\vec{x})$, то ф.пл. случайного вектора $\vec{\eta} = Q\vec{\xi} + \vec{a}$ равна* $f_{\vec{\eta}}(\vec{y}) = \frac{1}{\|Q\|} f_{\vec{\xi}}(Q^{-1}(\vec{y} - \vec{a}))$.

Некоторые соображения к доказательству леммы, не связанные с теоремой 107, приведены в дополнении к этой главе (стр. 130).

110| Упр. Докажите, что если ξ_1, ξ_2 — независимые нормальные $\mathcal{N}(0,1)$

сл.в., то сл.в. $\eta_1 = \xi_1 + \xi_2$ и $\eta_2 = \xi_1 - \xi_2$ также независимы и имеют нормальное распределение.

111| Пример. Пусть $\vec{\xi} \backsim \mathcal{D}\mathtt{ir}_2(a_1, a_2, a_3)$ с плотностью (22), стр. 71. Вектор $(\xi_1 + \xi_2, \xi_2)$, получаемый при линейном преобразовании $y_1 = x_1 + x_2, y_2 = x_2$ $(x_1 = y_1 - y_2, x_2 = y_2)$, имеет матрицу преобразований, для которой определитель $|Q| = 1$. При этом носитель $\mathscr{X} = \{0 < x_1, x_2 < 1, x_1 + x_2 < 1\}$ переходит в $\mathscr{Y} = \{0 < y_1 < 1, 0 < y_2 < y_1\}$. Следовательно, плотность $(\xi_1 + \xi_2, \xi_2)$ при $(y_1, y_2) \in \mathscr{Y}$ имеет вид

$$f(y_1, y_2) = \frac{\Gamma(a_1 + a_2 + a_3)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)} (y_1 - y_2)^{a_1 - 1} y_2^{a_2 - 1} (1 - y_1)^{a_3 - 1}.$$

Отсюда легко найти плотность $\xi_1 + \xi_2$, если при каждом фиксированном y_1 проинтегрировать последнее выражение по $y_2 \in [0; y_1]$ — здесь y_2 пробегает все действительные значения, при которых вектор $(y_1, y_2) \in \mathscr{Y}$. После несложных преобразований получаем, что $\xi_1 + \xi_2 \backsim \mathcal{B}\mathtt{et}(a_1 + a_2, a_3)$. Ясно, что справедливо и более общее свойство:

$$(\xi_1, \ldots, \xi_k) \backsim \mathcal{D}\mathtt{ir}_k(a_1, \ldots, a_{k+1}) \quad \Rightarrow$$
$$\Big(\sum_{i=1}^{r_1} \xi_i, \ldots, \sum_{i=1}^{r_m} \xi_i\Big) \backsim \mathcal{D}\mathtt{ir}_m\Big(\sum_{i=1}^{r_1} a_i, \ldots, \sum_{i=1}^{r_m} a_i, a_{k+1}\Big)$$

для любых натуральных чисел $r_1 + \ldots + r_m = k, \quad 1 \leqslant m \leqslant k$. ⊚

§ 2. Распределение суммы случайных величин. Свёртка

Чтобы найти распределение суммы сл.в., можно воспользоваться приёмом из примера 111, т.е. вложить отображение $(x_1, \ldots, x_k) \mapsto \sum_1^k x_j$ в некоторое взаимно однозначное отображение большей размерности, после чего применить стандартный способ выделения маргинальных распределений.

Рассмотрим преобразование $y_1 = \sum_1^k x_j, y_i = x_i, \ i = \overline{2, k}$, с обратным $x_1 = y_1 - \sum_2^k y_j, x_i = y_i, \ i = \overline{2, k}$. Определитель матрицы преобразования равен 1. Поэтому в силу леммы 109 ф.пл. $\vec{\eta} = (\xi_1 + \ldots + \xi_k, \xi_2, \ldots, \xi_k)$ связана с плотностью $\vec{\xi}$ уравнением $f_{\vec{\eta}}(y_1, \ldots, y_k) = f_{\vec{\xi}}(y_1 - y_2 - \ldots - y_k, y_2, \ldots, y_k)$. В соответствии с теоремой о частной плотности 63, стр. 68, справедлива

112| **Теорема.** *Если* $f(x_1, \ldots, x_k)$ *— ф.пл. сл.вектора* $(\xi_1, \ldots, \xi_k)$, *то* *плотность распределения суммы* $S = \xi_1 + \ldots + \xi_k$ *равна*

$$f_S(t) = \int\limits_{\mathbb{R}^{k-1}} f(t - y_2 - \ldots - y_k,\, y_2, \ldots, y_k)\, dy_2 \cdots dy_k.$$

113| Упр. Докажите, что

$$f_S(t) = \int\limits_{\mathbb{R}^{k-1}} f(y_1, \ldots, y_{k-1},\, t - y_1 - \ldots - y_{k-1})\, dy_1 \cdots dy_{k-1}.$$

Наиболее востребован случай с независимыми слагаемыми.

Определение. Ф.р. (плотность) суммы независимых сл.в. $\xi_1, \ldots, \xi_n$ называется *свёрткой* распределений (свёрткой плотностей) и обозначается $F_{\xi_1} * \ldots * F_{\xi_n}$ или $\mathop{*}\limits_{j=1}^{n} F_{\xi_j}$ (соответственно $f_{\xi_1} * \ldots * f_{\xi_n}$ или $\mathop{*}\limits_{j=1}^{n} f_{\xi_j}$).

Мы приведём здесь выражения для свёртки двух распределений в многомерном случае. Представим ф.р. $F(\vec{u})$ в виде

$$F(\vec{u}) \;=\; \int\limits_{\mathbb{R}^k} \dot{\mathrm{I}}_{\,\vec{u}}(\vec{y})\, dF(\vec{y}) \stackrel{(a)}{=} \int\limits_{\mathbb{R}^k} \dot{\mathrm{I}}_{\,\vec{u}}(\vec{y})\, f(\vec{y})\, d\vec{y}, \tag{1}$$

где $\dot{\mathrm{I}}_{\,\vec{u}}(\vec{y})$ — индикаторная функция множества $\{\vec{y}:\ \vec{y} \leqslant \vec{u}\}$ и равенство $\stackrel{(a)}{=}$ верно для абсолютно непрерывных распределений с плотностью f.

114| **Теорема.** [*О свёртке.*] I) *Свёртка двух ф.р.* F_1, F_2

$$F_1 * F_2\,(\vec{u}) \;=\; \int\limits_{\mathbb{R}^k} F_1(\vec{u} - \vec{y})\, dF_2(\vec{y}) \;=\; \int\limits_{\mathbb{R}^k} F_2(\vec{u} - \vec{x})\, dF_1(\vec{x}). \tag{2}$$

II) *Если* f_1 *— плотность* F_1, *то свёртка* $F_1 * F_2$ *также имеет плотность*

$$f_1 * F_2\,(\vec{u}) \;=\; \int\limits_{\mathbb{R}^k} f_1(\vec{u} - \vec{y})\, dF_2(\vec{y}). \tag{3}$$

III) *Если обе ф.р.* $F_1(\vec{x}), F_2(\vec{x})$ *имеют плотности* $f_1(\vec{x})$ *и* $f_2(\vec{x})$ *соответственно, то плотность их свёртки имеет вид*

$$f_1 * f_2\,(\vec{u}) \;=\; \int\limits_{\mathbb{R}^k} f_1(\vec{u} - \vec{y})\, f_2(\vec{y})\, d\vec{y} \;=\; \int\limits_{\mathbb{R}^k} f_2(\vec{u} - \vec{y})\, f_1(\vec{y})\, d\vec{y}. \tag{4}$$

⤃ I) Пусть $\vec{\xi} \backsim F_1(\vec{x})$, $\vec{\eta} \backsim F_2(\vec{y})$. По условию теоремы совместная ф.р. $(\vec{\xi}, \vec{\eta})$ равна $G(\vec{x}, \vec{y}) = F_1(\vec{x})F_2(\vec{y})$. В силу теоремы Фубини для неотрицательных функций

$$F_1 * F_2\,(\vec u) \;=\; \mathbf{P}\big\{\,\vec\xi + \vec\eta \leqslant \vec u\,\big\} \;=\; \int_{\mathbb{R}^{2k}} \dot{\mathrm{I}}_{\,\vec u}(\vec x + \vec y)\, dG(\vec x, \vec y) \;= \tag{5}$$
$$=\; \int_{\mathbb{R}^k}\bigg(\int_{\mathbb{R}^k} \dot{\mathrm{I}}_{\,\vec u}(\vec x + \vec y)\, dF_1(\vec x)\bigg) dF_2(\vec y) \;=\; \int_{\mathbb{R}^k} F_1(\vec u - \vec y)\, dF_2(\vec y).$$

Последнее равенство следует из (1) ввиду того, что при фиксированном $\vec y$ индикаторная функция $\dot{\mathrm{I}}_{\,\vec u}(\vec x + \vec y) = \dot{\mathrm{I}}_{\,(\vec u - \vec y)}(\vec x),\ \ \vec x \in \mathbb{R}^k$.

II) Запишем ф.р. $F_1(\vec u - \vec y)$ при фиксированном $\vec y$ через функцию плотности и произведём замену переменных $\vec x \dashrightarrow (\vec v - \vec y),\ \ \vec v \in \mathbb{R}^k$:

$$F_1(\vec u - \vec y) \;=\; \int_{\mathbb{R}^k} \dot{\mathrm{I}}_{\,\vec u}(\vec x + \vec y)\, f_1(\vec x)\, d\vec x \;=\; \int_{\mathbb{R}^k} \dot{\mathrm{I}}_{\,\vec u}(\vec v)\, f_1(\vec v - \vec y)\, d\vec v,$$

где индикатор $\dot{\mathrm{I}}_{\,\vec u}(\vec v)$ не зависит от $\vec y$. Таким образом, по теореме Фубини

$$\mathbf{P}\big\{\,\vec\xi + \vec\eta \leqslant \vec u\,\big\} \;=\; \int_{\mathbb{R}^k}\bigg(\int_{\mathbb{R}^k} \dot{\mathrm{I}}_{\,\vec u}(\vec v)\, f_1(\vec v - \vec y)\, d\,\vec v\bigg) dF_2(\vec y) \;=$$
$$=\; \int_{\mathbb{R}^k} \dot{\mathrm{I}}_{\,\vec u}(\vec v)\bigg(\int_{\mathbb{R}^k} f_1(\vec v - \vec y)\, dF_2(\vec y)\bigg)\, d\,\vec v,$$

что эквивалентно утверждению теоремы с заявленной плотностью.

Утверждение III) следует из II) и теоремы Радона–Никодима. ⤆

115| Упр. Докажите пункт III) теоремы в одномерном случае, воспользовавшись утверждением теоремы <u>112</u>.

⚠ Под свёрткой часто понимают только операцию над функциями типа (2), (4). Терминология, связанная со сл.в., представляется нам более наглядной. Так, пункт I) можно доказать, воспользовавшись свойствами условного математического ожидания. Заметим, что если сл.в. ξ_1, ξ_2 независимы, то справедливо равенство <u>39</u>, стр. 96, и потому при почти всех x_2

$$\mathbf{E}\big[\,\dot{\mathrm{I}}\,(\xi_1 + \xi_2 \leqslant u)\,\big|\,\xi_2 = x_2\big] \;=\; \mathbf{E}[\dot{\mathrm{I}}\,(\xi_1 \leqslant u - x_2)] \;=\; F_1(u - x_2).$$

Следовательно, по формуле полного математического ожидания

$$\mathbf{P}\{\xi_1 + \xi_2 \leqslant u\} \;=\; \mathbf{E}\big[\,\mathbf{E}\big[\,\dot{\mathrm{I}}\,(\xi_1 + \xi_2 \leqslant u)\,\big|\,\xi_2\big]\big] \;=$$
$$=\; \int_{\mathbb{R}} \mathbf{E}\big[\,\dot{\mathrm{I}}\,(\xi_1 + \xi_2 \leqslant u)\,\big|\,\xi_2 = x_2\big]\, dF_2(x_2) \;=\; \int_{\mathbb{R}} F_1(u - x_2)\, dF_2(x_2).$$

116| Упр. Докажите, что при $\forall\, a > 0,\ b \in \mathbb{R}^1$ справедливо равенство

$$\mathbf{P}\big\{\,a\vec\xi_1 + b\vec\xi_2 \leqslant \vec u\,\big\} \;=\; \int_{\mathbb{R}^k} F_1\big(\tfrac{1}{a}(\vec u - b\vec y)\big)\, dF_2(\vec y).$$

117| Упр. Воспользовавшись частью II) теоремы о свёртке, докажите, что для любой ф.р. F функция $F(x) - F(x-1)$, $x \in \mathbb{R}^1$, есть функция плотности, в частности $\int_{-\infty}^{\infty} \big(F(x) - F(x-1)\big)\, dx = 1$. Чему равен интеграл $\int_{-\infty}^{\infty} \big(F(x) - F(x-c)\big)\, dx$ при $c \in \mathbb{R}^1$?

Если сл.в. ξ_1, ξ_2 неотрицательны, то вид свёртки их плотностей немного упрощается.

118| **Лемма.** (?!) *Пусть независимые неотрицательные сл.величины ξ_1, ξ_2 имеют плотности f_1, f_2 соответственно. Тогда плотности распределений их суммы и их разности равны соответственно*

$$\begin{aligned} f_{\xi_1+\xi_2}(u) &= \int_0^u f_2(u-y) f_1(y)\, dy, \quad u > 0, \\ f_{\xi_1-\xi_2}(u) &= \int\limits_{\max\{0,-u\}}^{\infty} f_1(u+y) f_2(y)\, dy, \quad u \in \mathbb{R}^1. \end{aligned} \tag{6}$$

119| Пример. Пусть сл.в. $\xi \backsim \mathcal{G}\mathtt{am}(p,1)$, а сл.в. $\eta \backsim \mathcal{G}\mathtt{am}(q,1)$. Носители их распределений совпадают с $\mathbb{R}^1_+$, поэтому носитель суммы $\zeta = \xi + \eta$ также совпадает с $\mathbb{R}^1_+$. По предыдущей лемме свёртка (при $u > 0$)

$$f_\xi * f_\eta(u) = \int_0^u \frac{(u-y)^{p-1} y^{q-1}}{\Gamma(p)\Gamma(q)} e^{-(u-y)} e^{-y}\, dy = \frac{e^{-u} u^{p-1}}{\Gamma(p)\Gamma(q)} \int_0^u \Big(1 - \frac{y}{u}\Big)^{p-1} y^{q-1} dy.$$

После замены $y \dashrightarrow tu$, $0 \leqslant t \leqslant 1$, последний интеграл переходит в

$$u^q \int_0^1 (1-t)^{p-1} t^{q-1} dt = u^q\, \mathrm{B}(p,q) = u^q \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

Следовательно, $f_\xi * f_\eta(u) = u^{p+q-1} e^{-u} / \Gamma(p+q)$. В неформальном виде это можно записать как $\mathcal{G}\mathtt{am}(p,1) * \mathcal{G}\mathtt{am}(q,1) \backsim \mathcal{G}\mathtt{am}(p+q,1)$. ◎

120| Упр. Покажите, что

(а) $\mathcal{G}\mathtt{am}(p,\lambda) * \mathcal{G}\mathtt{am}(q,\lambda) \backsim \mathcal{G}\mathtt{am}(p+q,\lambda)$ при любых $p, q, \lambda > 0$;

(б) если $\xi_1, \xi_2 \backsim \mathcal{E}\mathtt{x}(1)$ — независимые показательные сл.в., то их разность имеет распределение Лапласа: $\xi_1 - \xi_2 \backsim \mathcal{L}\mathtt{ap}(0,1)$;

(в) если $\xi_1, \xi_2 \backsim \mathcal{U}\mathtt{n}(0,1)$ — независимые сл.в. с равномерным распределением в интервале $[0;1]$, то $\xi_1 - \xi_2 \backsim \xi_1 + \xi_2 - 1$.

► Во второй части теоремы о свёртке утверждается, что сумма независимых сл.в. имеет абсолютно непрерывный тип распределения, даже если только одна из них абсолютно непрерывного типа. В частности, справедлива следующая лемма.

121| **Лемма.** (?!) *Если сл.в.* η *имеет дискретное распределение, сосредоточенное на* $\mathcal{Y}$ *с вероятностями* $p_j = \mathbf{P}\{\eta = y_j\}$, $y_j \in \mathcal{Y}$, $1 \leqslant j \leqslant N$, $N \leqslant \infty$, *а сл.в.* ξ *имеет абсолютно непрерывное распределение с плотностью* f_ξ, *то плотность свёртки (плотность* $\xi + \eta$*) равна*

$$f_{\xi+\eta}(x) = \sum\nolimits_1^N f_\xi(x - y_j)p_j.$$

122| Пример. Пусть сл.в. $\xi \backsim \mathcal{U}\mathtt{n}(0,1)$, а сл.в. η принимает значения 0 и 1 с равными вероятностями. Тогда $f_{\xi+\eta}(x) = \frac{1}{2} f_\xi(x) + \frac{1}{2} f_\xi(x-1)$. Первое слагаемое здесь отлично от нуля (и $f_\xi(x) = 1$) только при $x \in [0;1]$. Второе слагаемое $f_\xi(x-1) = 1$ при $x \in [1;2]$. Таким образом, $f_{\xi+\eta}(x) = {}^1\!/_2$ при $0 \leqslant x \leqslant 2$, т.е. свёртка $\mathcal{U}\mathtt{n}(0,1) * \mathcal{B}\mathtt{ern}({}^1\!/_2) \backsim \mathcal{U}\mathtt{n}(0;2)$. ⊚

123| Упр. Докажите, что если одна из функций распределения непрерывна, то свёртка также будет иметь всюду непрерывную ф.р.

124| **Лемма.** (?!) *Пусть* ξ, η *— независимые сл.в., сосредоточенные на носителях* $\mathcal{X}_\xi = \langle\, x_k \,\rangle_{k=1}^N$, $\mathcal{X}_\eta = \langle\, y_j \,\rangle_{j=1}^M$, $1 \leqslant N, M \leqslant \infty$. *Тогда* $\mathcal{X}_S = \langle\, z = x_k + y_j : \ k = \overline{1,N},\ j = \overline{1,M} \,\rangle$ *— носитель* $\xi + \eta$ *и для* $\forall\, z \in \mathcal{X}_S$

$$\mathbf{P}\{\xi + \eta = z\} \ = \ \sum\nolimits_1^M \mathbf{P}\{\xi = z - y_j\}\, \mathbf{P}\{\eta = y_j\}\ .$$

125| Пример. Пусть независимые сл.в. $\xi, \eta \backsim \mathcal{P}\mathtt{ois}(\lambda)$. Обе сл.в. принимают значения на множестве $\mathbb{N}_0$ всех неотрицательных целых чисел, поэтому их сумма также будет принадлежать $\mathbb{N}_0$. Прежде найдём те значения $y_j = j \in \mathbb{N}_0$, для которых слагаемые в формуле свёртки отличны от нуля. При целых $z, j \geqslant 0$ произведение вероятностей $\mathbf{P}\{\xi = z - j\}\,\mathbf{P}\{\eta = j\} \neq 0$, если $0 \leqslant j \leqslant z$ (сравните с областью интегрирования в формуле для плотности свёртки неотрицательных сл.в.), поэтому вероятность $\mathbf{P}\{\xi + \eta = z\}$ равна

$$\sum_{j=0}^{\infty} \mathbf{P}\{\xi = z - j\}\ \mathbf{P}\{\eta = j\} \ = \ \sum_{j=0}^{z} \frac{\lambda^{z-j}}{(z-j)!} e^{-\lambda}\, \frac{\lambda^j}{j!}\, e^{-\lambda} \ =$$

$$= \ \lambda^z e^{-2\lambda} \sum_{j=0}^{z} \frac{1}{(z-j)!\, j!} \ = \ \frac{\lambda^z}{z!}\, e^{-2\lambda} \underbrace{\sum_{j=0}^{z} \frac{z!}{(z-j)!\, j!}}_{2^z} \ = \ \frac{(2\lambda)^z}{z!} e^{-2\lambda},$$

где мы воспользовались формулой бинома Ньютона. Следовательно, $\xi+\eta \backsim \backsim \mathcal{P}\mathtt{ois}(2\lambda)$ или в неформальном виде $\mathcal{P}\mathtt{ois}(\lambda) * \mathcal{P}\mathtt{ois}(\lambda) \backsim \mathcal{P}\mathtt{ois}(2\lambda)$. ◎

✧ **Таблица распределений сумм случайных величин**. Как видно из следующей теоремы, основные вероятностные модели (см. стр. 107) обладают свойством замкнутости относительно свёртки — свёртка принадлежит тому же типу $\mathcal{P}$, что и слагаемые: $(F_1,\ldots,F_n)\subset\mathcal{P} \;\Rightarrow\; \mathop{*}\limits_{j=1}^{n} F_j \in \mathcal{P}$.

126| **Теорема.** *Пусть* $\langle\,\xi_j\,\rangle_1^n$ *— независимые сл.в.,* $S=\sum_1^n \xi_j$.

Распределение слагаемых	*Распределение суммы*	*Параметры распределения* S
$\xi_j \backsim \mathcal{B}\mathtt{ern}(p)$	$S \backsim \mathcal{B}\mathtt{in}(n,p)$	$n,\;\; p$
$\xi_j \backsim \mathcal{B}\mathtt{in}(k_j,p)$	$S \backsim \mathcal{B}\mathtt{in}(K,p)$	$K=\sum_1^n k_j,\;\; p$
$\xi_j \backsim \mathcal{P}\mathtt{sc}(k_j,p)$	$S \backsim \mathcal{P}\mathtt{sc}(K,p)$	$K=\sum_1^n k_j,\;\; p$
$\xi_j \backsim \mathcal{P}\mathtt{ois}(\lambda_j)$	$S \backsim \mathcal{P}\mathtt{ois}(\Lambda)$	$\Lambda=\sum_1^n \lambda_j$
$\xi_j \backsim \mathcal{N}(\mu_j,\sigma_j^2)$	$S \backsim \mathcal{N}(\mu,\mathfrak{S}^2)$	$\mu=\sum_1^n \mu_j,\;\; \mathfrak{S}^2=\sum_1^n \sigma_j^2$
$\xi_j \backsim \mathcal{E}\mathtt{x}(\lambda)$	$S \backsim \mathcal{G}\mathtt{am}(n,\lambda)$	$\lambda,\;\; n$
$\xi_j \backsim \mathcal{G}\mathtt{am}(r_j,\lambda)$	$S \backsim \mathcal{G}\mathtt{am}(R,\lambda)$	$\lambda,\;\; R=\sum_1^n r_j$
$\xi_j \backsim \mathcal{C}\mathtt{auch}(\mu_j,\sigma_j)$	$S \backsim \mathcal{C}\mathtt{auch}(\mu,\mathfrak{S})$	$\mu=\sum_1^n \mu_j,\;\; \mathfrak{S}=\sum_1^n \sigma_j$

Доказательство утверждений мы оставляем читателю в качестве упражнения. Заметим, что распределение свёртки дискретных сл.в. может быть найдено непосредственно из определения соответствующей вероятностной модели.

127| Пример. Двумерное распределение Пуассона с параметрами $\lambda_1,\lambda_2>0,\; \lambda_0\geqslant 0$ определяется ф.вер.

$$p(x_1,x_2) \;=\; e^{-\lambda_0-\lambda_1-\lambda_2} \sum_{j=0}^{\min(x_1,x_2)} \frac{\lambda_0^j\,\lambda_1^{x_1-j}\,\lambda_2^{x_2-j}}{j!\,(x_1-j)!\,(x_2-j)!} \tag{7}$$

для всех пар неотрицательных целых чисел (x_1,x_2). Заметим, что при $\lambda_0= =0$[‡] эта ф.вер. равна произведению одномерных пуассоновских вероятностей. Вместо прямого вычисления характеристик распределения заметим, что эта модель допускает следующую интерпретацию. Пусть $\eta_j\backsim\mathcal{P}\mathtt{ois}(\lambda_j)$, $j= =0,1,2,$ — независимые (одномерные) пуассоновские сл.в. с соответствующими параметрами. Тогда сл.вектор $(\xi_1,\xi_2)=(\eta_0+\eta_1,\eta_0+\eta_2)$ имеет (?!)

[‡] По договорённости $0^0=1$.

распределение (7). По предыдущей теореме $\xi_j \backsim \mathcal{P}\mathtt{ois}(\lambda_0 + \lambda_j)$, $j = 1, 2$. Как известно (см. стр. 107), $\mathbf{E}\,\xi_j = \mathbf{D}\xi_j = \lambda_0 + \lambda_j$, $j = 1, 2$. Поэтому ковариация

$$\mathbf{Cov}(\xi_1, \xi_2) \;=\; \mathbf{E}\left[\big((\eta_0 - \lambda_0) + (\eta_1 - \lambda_1)\big)\big((\eta_0 - \lambda_0) + (\eta_2 - \lambda_2)\big)\right] \;=\; \mathbf{D}\eta_0 \;=\; \lambda_0$$

в силу независимости η_0, η_1, η_2. Коэффициент корреляции $\mathbf{Corr}(\xi_1, \xi_2) = \lambda_0 / \sqrt{(\lambda_0 + \lambda_1)(\lambda_0 + \lambda_2)}$ всегда неотрицателен, и, кроме того, компоненты пуассоновского вектора независимы т. т. т. когда они некоррелированы. ◎

128| Упр. Воспользовавшись только видом совместной вероятности (7), покажите, что частное распределение первой компоненты $\xi_1 \backsim \mathcal{P}\mathtt{ois}(\lambda_0 + \lambda_1)$.

§ 3. Распределение немонотонных преобразований

✧ Распределение Стьюдента. В статистике особую роль играет распределение Стьюдента, описывающее поведение отношения $\zeta = \sqrt{2p}\,\xi\,/\sqrt{\eta}$ с независимыми сл.в. $\xi \backsim \mathcal{N}(0, 1)$ и $\eta \backsim \mathcal{G}\mathtt{am}(p, 2)$; $p > 0$ — фиксированный параметр. Найдём сначала плотность сл.вектора (ϑ, η), где $\vartheta = \xi\,/\sqrt{\eta}$.

Рассмотрим преобразование $H : \mathbb{R}^1 \times \mathbb{R}^1_+ \mapsto \mathbb{R}^1 \times \mathbb{R}^1_+$ и обратное преобразование H^{-1}, определяемые соотношениями

$$\begin{aligned}
(\theta, v) &= H(x, y) = \left(\frac{x}{\sqrt{y}},\ y\right), \qquad x \in \mathbb{R}^1, y > 0,\\
(x, y) &= H^{-1}(\theta, v) = \big(\theta\sqrt{v},\ v\big), \qquad \theta \in \mathbb{R}^1, v > 0.
\end{aligned}$$

Якобиан $J_{H^{-1}} = \sqrt{v}$. Поэтому вектор $(\vartheta, V) = \big(\xi\,/\sqrt{\eta},\ \eta\big)$ имеет плотность

$$f(\theta, v) \;=\; J_{H^{-1}} \cdot f_\xi(x) f_\eta(y)\Big|_{(x,y) = H^{-1}(\theta, v)} \;=\; \sqrt{v}\; f_\xi\big(\theta\sqrt{v}\big)\, f_\eta(v), \quad \theta \in \mathbb{R}^1,\ v > 0.$$

Чтобы найти частную плотность ϑ, подставим сюда плотности нормального $\mathcal{N}(0, 1)$ и гамма $\mathcal{G}\mathtt{am}(p, 2)$ законов и проинтегрируем по $v > 0$:

$$\begin{aligned}
f_\vartheta(\theta) \;&=\; \int_0^\infty \sqrt{v}\frac{1}{\sqrt{2\pi}}\exp\Big\{-\tfrac{1}{2}\theta^2 v\Big\}\frac{1}{2^p\Gamma(p)} v^{p-1} e^{-\frac{1}{2}v}\, dv \;= \\
&= \frac{1}{\sqrt{2\pi}\, 2^p\, \Gamma(p)} \int_0^\infty v^{p + \frac{1}{2} - 1}\, \exp\Big\{-\Big(\tfrac{1}{2} + \tfrac{\theta^2}{2}\Big)\, v\Big\}\, dv.
\end{aligned}$$

Последний интеграл равен $\Gamma\big(p+\frac{1}{2}\big)\big(\frac{1}{2}+\frac{1}{2}\,\theta^2\big)^{-p-\frac{1}{2}}$. Итак,

$$f_\vartheta(\theta) \;=\; \frac{\Gamma(p+{}^1\!/_2)}{\sqrt{2\pi}\,2^p\,\Gamma(p)}\Big(\frac{1}{2}+\frac{\theta^2}{2}\Big)^{-p-\frac{1}{2}} \;=\; \frac{\Gamma(p+{}^1\!/_2)}{\sqrt{\pi}\,\Gamma(p)}\big(1+\theta^2\big)^{-\frac{(2p+1)}{2}}.$$

У классического распределения Стьюдента $p=k/2$, где k — целое число, называемое числом степеней свободы распределения Стьюдента. В этом случае плотность сл.в. $\zeta=\sqrt{k}\,\vartheta$ равна

$$f_\zeta(z) \;=\; \frac{\Gamma\big((k+1)/2\big)}{\sqrt{\pi k}\,\Gamma(k/2)}\cdot\Big(1+\frac{z^2}{k}\Big)^{-\frac{1}{2}(k+1)}, \qquad z\in\mathbb{R}^1. \tag{8}$$

Для пояснения терминологии заметим, что гамма-распределение $\mathcal{G}\mathtt{am}(k/2\,,2)$ совпадает с хи-квадрат распределением $\mathcal{K}\mathtt{hi}(k)$ с k степенями свободы.

Нормировка на $\sqrt{k}$ здесь полезна, поскольку, как легко видеть с помощью «второго замечательного предела» и формулы Стирлинга при больших k, плотность (8) близка к нормальной $\mathcal{N}(0,1)$ плотности.

✧ Плотность отношения случайных величин. Вывод распределения Стьюдента может быть осуществлён без обращения к теореме 107 (напомним, что эта теорема осталась у нас без доказательства). Идея вывода здесь аналогична идее доказательства леммы 109 (стр. 130): функция $x/\sqrt{y}$ линейна по переменной x при каждом y. Продемонстрируем схему применения этого свойства на следующем примере, где заодно продемонстрируем способ нивелирования возможной неоднозначности рассматриваемого преобразования, который заключается, конечно, в разбиении пространства значений сл.векторов на подходящие подмножества.

129| **Теорема.** *Пусть $f(x,y)$ — ф.пл. сл.вектора (ξ,η). Тогда плотность распределения отношения $\zeta=\xi/\eta$ равна*

$$f_\zeta(z)=\int_{\mathbb{R}}|y|\,f(zy,y)\,dy.$$

⊨⇒ По теореме Фубини находим ф.р. ζ:

$$F_\zeta(z) \;=\; \mathbf{P}\Big\{\,\frac{\xi}{\eta}\leqslant z\,\Big\} \;=\; \int_{\mathbb{R}^2}\dot{\mathrm{I}}\Big(\,\frac{x}{y}\;\leqslant\; z\Big)f(x,y)\,dxdy \;=$$

$$=\;\int\limits_{y<0}\Big[\int\limits_{x/y\leqslant z}f(x,y)\,dx\Big]\,dy+\int\limits_{y>0}\Big[\int\limits_{x/y\leqslant z}f(x,y)\,dx\Big]\,dy,$$

где область $\{(x,y)\;:\;\; x\in\mathbb{R}^1,\, y=0\}$ можно исключить из области интегрирования, т.к. она имеет нулевую лебегову меру. В каждом внутреннем интеграле (в квадратных скобках) по теореме 543, стр. 484, мы вправе при фиксированном $y \neq 0$ осуществить подстановку $x = yu$, т.е. $\frac{x}{y} = u$, $u \leqslant z$. Итак,

$$F_\zeta(z) = \int\limits_{y<0} \Big[|y| \int\limits_{u \leqslant z} f(uy, y)\, du \Big]\, dy + \int\limits_{y>0} \Big[|y| \int\limits_{u \leqslant z} f(uy, y)\, du \Big]\, dy,$$

что после применения теоремы Фубини для положительных функций даёт

$$F_\zeta(z) = \int\limits_{u \leqslant z} \Big[\int\limits_{y \in \mathbb{R}^1} |y| f(uy, y)\, dy \Big]\, du.$$

Такая запись ф.р. эквивалентна утверждению теоремы. ⤆

130 Упр. Покажите, что отношение независимых стандартных нормальных сл.в. имеет стандартное распределение Коши, т.е.

$$\text{если } \xi, \eta \backsim \mathcal{N}(0,1), \quad \text{то } \frac{\xi}{\eta} \backsim \mathfrak{Cauch}(0,1).$$

§ 4. Распределение порядковых статистик

В математической статистике часто возникает следующая задача. Имеется выборка $(x_1, \ldots, x_n)$, полученная из реализаций n независимых одинаково распределённых сл.в. $\xi_1, \ldots, \xi_n$ с общей ф.р. F. Эта выборка упорядочивается по возрастанию $x_{(1)} \leqslant x_{(2)} \leqslant \ldots \leqslant x_{(n)}$, т.е. строится так называемый вариационный ряд. Сл.величину, описываюшую k-е по порядку значение, обозначим через $\xi_{(k)}$. Необходимо найти распределение фиксированной части вариационного ряда, например, первого члена (минимального значения исходной выборки) $\xi_{(1)}$ и (или) последнего члена (максимального значения исходной выборки) $\xi_{(n)}$.

Распределения крайних членов вариационного ряда находятся просто. Например, событие $\{\xi_{(n)} \leqslant x\}$ происходит т. т. т. когда значения выборки $\xi_k \leqslant x$ для $\forall\, k = \overline{1,n}$, поэтому в силу независимости и одинаковой распределённости

$$F_{(n)}(x) := \mathbf{P}\left\{\xi_{(n)} \leqslant x\right\} = \prod_1^n \mathbf{P}\left\{\xi_k \leqslant x\right\} = F^n(x).$$

Аналогично, $F_{(1)}(x) := \mathbf{P}\left\{\xi_{(1)} \leqslant x\right\} = 1 - (1 - F(x))^n$. (?!)

Если ф.р. F имеет плотность f, то формальным дифференцированием отсюда можно получить ф.пл. интересующего нас распределения. Например, производная ф.р. максимума выборки удовлетворяет соотношению

$$f_{(n)}(x) \;=\; F'_{(n)}(x) \;=\; nF^{n-1}(x)f(x). \tag{9}$$

Осталось только доказать, что такой формализм здесь допустим, т.е. доказать, что любая степень абсолютно непрерывной ф.р. снова абсолютно непрерывна. Вместо прямой проверки этого утверждения воспользуемся утверждением 40, стр. 53. Как уже отмечалось, производная ф.р. является плотностью т. т. т. когда эта производная удовлетворяет свойствам ф.плотности. Для производной ф.р. $F_{(n)}$ имеем

$$\int_{\mathbb{R}} f_{(n)}(x)\,dx \;=\; \int_{\mathbb{R}} nF^{n-1}(x)f(x)\,dx \;=\; n\int_{\mathbb{R}} F^{n-1}(x)\,dF(x) \;=\; \\ =\; n\int_0^1 t^{n-1}\,dt \;=\; 1,$$

т.е. (9) действительно есть ф.пл. $\xi_{(n)}$.

Найдём теперь совместное распределение $(\xi_{(1)},\xi_{(n)})$. Ясно, что ф.р. $F_{(1,n)}(u_1,u_2)$ при $u_1 \geqslant u_2$ равна

$$F_{(1,n)}(u_1,u_2) \;=\; \mathbf{P}\left\{\xi_{(1)} \leqslant u_1, \xi_{(n)} \leqslant u_2\right\} \;=\; \mathbf{P}\left\{\xi_{(n)} \leqslant u_2\right\} \;=\; F^n(u_2).$$

При $u_1 < u_2$ необходимо найти вероятность того, что все элементы выборки будут меньше u_2 (событие B_2), при этом хотя бы один элемент попадёт в интервал $(-\infty;u_1]$, иначе говоря, найти вероятность события B_2 за вычетом события, что вся выборка лежит в интервале $(u_1;u_2]$:

$$F_{(1,n)}(u_1,u_2) \;=\; F^n(u_2) - \big(F(u_2)-F(u_1)\big)^n.$$

Окончательно получаем ф.р.

$$F_{(1,n)}(u_1,u_2) = \begin{cases} F^n(u_2) - \big(F(u_2)-F(u_1)\big)^n, & \text{если } u_1 < u_2, \\ F^n(u_2), & \text{если } u_1 \geqslant u_2, \end{cases}$$

и ф.пл. (при $u_1 < u_2$)

$$f_{(1,n)}(u_1,u_2) \;=\; n(n-1)\big(F(u_2)-F(u_1)\big)^{n-2} f(u_1)f(u_2).$$

131| ⚠ Вывод совместной ф.р. $(\xi_{(1)},\xi_{(n)})$ удобнее начинать с отыскания вероятности $G(u_1,u_2) := \mathbf{P}\left\{\xi_{(1)} > u_1, \xi_{(n)} \leqslant u_2\right\} = \big(F(u_2)-F(u_1)\big)^n$, $u_1 \leqslant u_2$, с которой искомая ф.р. связана простым соотношением $F_{(1,n)}(u_1,u_2) = G(-\infty,u_2) - G(u_1,u_2)$.

Найдём плотность распределения «размаха» выборки $\rho = \xi_{(n)} - \xi_{(1)}$. Определитель матрицы линейного преобразования $(y_1, y_2) = h(u_1, u_2) = (u_1, u_2 - u_1)$ с обратным $h^{-1}(y_1, y_2) = (y_1, y_2 + y_1)$ равен 1. Поэтому плотность вектора $(\xi_{(1)}, \rho)$ равна $f_{(1,n)}(y_1, y_2 + y_1)$. Ясно, что эта плотность отлична от нуля только при $y_2 > 0$. Итак, плотность размаха выборки при $y > 0$ равна

$$f_\rho(y) \ = \ n(n-1)\int_{-\infty}^{\infty}(F(y+y_1) - F(y_1))^{n-2} f(y_1) f(y+y_1)\, dy_1. \qquad (10)$$

132| Пример. Пусть $\xi_k \backsim \mathcal{E}\text{x}(1)$, т.е. $F(x) = 1 - e^{-x}$, $x \geqslant 0$. Произведение плотностей $f(y_1)f(y+y_1)$ отлично от нуля только при $y_1 > 0$ (напомним, что $y > 0$). Поэтому для ф.пл. размаха получаем

$$f_\rho(y) \ = \ n(n-1)\int_0^{\infty}(e^{-y_1} - e^{-y-y_1})^{n-2} e^{-y_1} e^{-y-y_1}\, dy_1 \ =$$
$$= \ n(n-1)(1-e^{-y})^{n-2} e^{-y}\int_0^{\infty} e^{-y_1(n-2)} e^{-2y_1}\, dy_1 \ = \ (n-1)(1-e^{-y})^{n-2} e^{-y}.$$

Ф.р. размаха, очевидно, равна $F_\rho(y) = (1 - e^{-y})^{n-1}$, $y > 0$. ◎

Перейдём к распределению некоторых произвольных членов вариационного ряда, например $(\xi_{(k)}, \xi_{(m)})$, $k < m$.

133| Упр. Покажите, что при $u_1 \leqslant u_2$ ф.р.

$$\mathbf{P}\{\, \xi_{(k)} \leqslant u_1, \xi_{(m)} \leqslant u_2 \,\} =$$
$$= \sum_{i=k}^{n} \ \sum_{j=max\{0,m-i\}}^{n-i} \frac{n!}{i!j!(n-i-j)!} F^i(u_1)\big(F(u_2) - F(u_1)\big)^j \big(1 - F(u_2)\big)^{n-i-j}.$$

В данной ситуации удобнее оперировать с ф.пл. При $u_1 < u_2$ и малом $\varepsilon > 0$ событие $\{\xi_{(k)} \in (u_1; u_1 + \varepsilon], \xi_{(m)} \in (u_2; u_2 + \varepsilon]\}$ заведомо выполняется, если количество элементов выборки, попавших в тот или иной интервал, удовлетворяет следующей схеме:

$(-\infty; u_1]$	$(u_1; u_1+\varepsilon]$	$(u_1+\varepsilon; u_2]$	$(u_2; u_2+\varepsilon]$	$(u_2+\varepsilon; \infty)$
$k-1$	1	$m-k-1$	1	$n-m$

.

Обозначим полиномиальный коэффициент

$$\mathrm{C}_n^{k,m} = \frac{n!}{(k-1)!(m-k-1)!(n-m)!},$$

тогда в соответствии с полиномиальной моделью имеем

$$\mathbf{P}\big\{\,\xi_{(k)}\in(u_1;u_1+\varepsilon],\ \xi_{(m)}\in(u_2;u_2+\varepsilon]\,\big\}\ \geqslant$$
$$\geqslant\ \mathrm{C}_n^{k,\,m}\,F^{k-1}(u_1)\big(F(u_1+\varepsilon)-F(u_1)\big)\times$$
$$\times\big(F(u_2)-F(u_1+\varepsilon)\,\big)^{m-k-1}\big(F(u_2+\varepsilon)-F(u_2)\big)\big(1-F(u_2)\big)^{n-m}.$$

Следовательно, правая «производная» совместной ф.р.

$$\varliminf_{\varepsilon\searrow 0}\frac{1}{\varepsilon^2}\,\mathbf{P}\big\{\,\xi_{(k)}\in(u_1;u_1+\varepsilon],\xi_{(m)}\in(u_2;u_2+\varepsilon]\,\big\}\ \geqslant f_{(k,m)}(u_1,u_2),$$

где

$$f_{(k,m)}(u_1,u_2)\ := \tag{11}$$
$$=\ \mathrm{C}_n^{k,\,m}\ f(u_1)\,f(u_2)\ F^{k-1}(u_1)\big(F(u_2)-F(u_1)\big)^{m-k-1}\big(1-F(u_2)\big)^{n-m}.$$

Покажем, что интеграл $\int\int_{\mathbb{R}^2} f_{(k,m)}(u_1,u_2)\,du_1du_2=1$, где из очевидных соображений при $u_1>u_2$ полагается $f_{(k,m)}(u_1,u_2)=0$.

Применяя теорему Фубини и (два раза) лемму 40, стр. 53, получаем после замен $F(u_2)\dashrightarrow t_2\dashrightarrow x_2+x_1,\ \ F(u_1)\dashrightarrow t_1\dashrightarrow x_1$:

$$\int_{\mathbb{R}^2} f_{(k,m)}(u_1,u_2)\,du_1du_2=$$
$$=\mathrm{C}_n^{k,\,m}\int_0^1 dt_1\,\int_{t_1}^1\Big[t_1^{k-1}(t_2-t_1)^{m-k-1}(1-t_2)^{n-m}\Big]\,dt_2=$$
$$=\mathrm{C}_n^{k,\,m}\int_0^1 dx_1\,\int_0^{1-x_1}\Big[x_1^{k-1}x_2^{m-k-1}(1-x_2-x_1)^{n-m}\Big]\,dx_2=1,$$

т.к. подынтегральное выражение (вместе с коэффициентом $\mathrm{C}_n^{k,\,m}$) есть ф.пл. распределения Дирихле (стр. 70). В силу теоремы 60, стр. 63, отсюда следует, что найденное в (11) выражение для $f_{(k,m)}(u_1,u_2)$ и есть ф.пл. вектора $(\xi_{(k)},\xi_{(m)})$ (при $u_1<u_2$; в противном случае плотность равна нулю).

Аналогичным образом можно найти плотность любого набора элементов вариационного ряда, в частности плотность всего вектора вариационного ряда $(\xi_{(1)},\dots,\xi_{(n)})$ равна

$$f_{(1,\dots,n)}(u_1,\dots,u_n)=n!\prod_1^n f(u_k),\qquad u_1<u_2<\ldots<u_n.$$

134| Примеры. 1) Пусть $\xi_k\backsim\mathfrak{Un}(0,T)$, т.е. имеет равномерное распределение в отрезке $[0;T]$ с плотностью $f(x)=1/T,\ 0\leqslant x\leqslant T$. Функция плотности вектора $(\xi_{(1)},\dots,\xi_{(n)})$ равна $n!/T^n$ в области $0\leqslant u_1\leqslant\ldots\leqslant u_n\leqslant$

$\leqslant T$. В этом случае из выражения (11) для ф.пл. легко найти все числовые характеристики членов вариационного ряда. Так, при $k \leqslant m$

$$\mathbf{E}\,\xi_{(k)} = \frac{k}{n+1}, \quad \mathbf{D}\xi_{(k)} = \frac{k(n-k+1)}{(n+1)^2(n+2)}, \quad \mathbf{Cov}\,(\xi_{(k)}, \xi_{(m)}) = \frac{k(n-m+1)}{(n+1)^2(n+2)}.$$

Здесь можно вычислить и ф.р.:

$$\begin{gathered}\mathbf{P}\left\{\xi_{(1)} \leqslant u_1, \ldots, \xi_{(n)} \leqslant u_n\right\} \;=\; \frac{n!}{T^n}\int_0^{u_1} dt_1 \int_{t_1}^{u_2} dt_2 \cdots \int_{t_{n-1}}^{u_n} dt_n, \\ 0 \leqslant u_1 \leqslant \ldots \leqslant u_n \leqslant T.\end{gathered} \tag{12}$$

Поскольку неравенство $\xi_{(k)} \leqslant u$ влечёт справедливость неравенств $\xi_{(j)} \leqslant u$ при любых $j < k$, то для произвольных $u_k \geqslant 0$, $k = \overline{1, n}$, ф.р. вычисляется по той же формуле с заменой верхних пределов интегрирования $u_k \dashrightarrow u'_k = \min\{u_k, \ldots, u_n, T\}$, $k = \overline{1, n}$.

2) Пусть объём выборки $n = 2k+1$, найдём распределение $(k+1)$-го члена вариационного ряда $\xi_{(k+1)}$, т.е. медианы выборки. По схеме, разработанной выше, если f — плотность наблюдений, то плотность $\xi_{(k+1)}$ можно представить как

$$f_{(k+1)}(x) \;=\; \frac{(2k+1)!}{k!k!}\,F^k(x)(1-F(x))^k f(x).$$

Пусть m — медиана ф.р. F, т.е. $F(m) = \int_{-\infty}^{m} f(x)\,dx = \frac{1}{2}$. Тогда

$$\begin{aligned}\mathbf{P}\left\{\xi_{(k+1)} \leqslant m\right\} \;&=\; \frac{(2k+1)!}{k!k!}\int_{-\infty}^{m} F^k(x)(1-F(x))^k f(x)\,dx \;= \\ &= \frac{(2k+1)!}{k!k!}\int_0^{1/2} t^k(1-t)^k\,dt \;=\; \frac{1}{2},\end{aligned}$$

ввиду очевидной симметрии бета-распределения $\mathcal{B}\mathtt{et}(k+1, k+1)$ относительно точки $\frac{1}{2}$. Другими словами, теоретическая медиана F служит медианой распределения выборочной медианы. В теории оценивания такое свойство называется несмещённостью по медиане. Этот результат справедлив и когда у исходной ф.р. F нет плотности. Действительно, полагая $t = F(x)$, имеем

$$\mathbf{P}\left\{\xi_{(k+1)} \leqslant x\right\} \;=\; \sum_{j=k+1}^{n} \mathrm{C}_n^j F^j(x)(1-F(x))^{n-j} \;=\; \sum_{j=k+1}^{n} \mathrm{C}_n^j t^j(1-t)^{n-j}.$$

Продифференцировав это выражение по t, получим

$$\frac{d}{dt}\,\mathbf{P}\left\{\xi_{(k+1)} \leqslant x\right\} \;=$$

$$= \Big[\frac{n!}{k!(n-k-1)!}t^k(1-t)^{n-k-1} - \frac{n!}{(k+1)!(n-k-2)!}t^{k+1}(1-t)^{n-k-2} +$$
$$+ \frac{n!}{(k+1)!(n-k-2)!}t^{k+1}(1-t)^{n-k-2} - \frac{n!}{(k+2)!(n-k-3)!}t^{k+2}(1-t)^{n-k-3} +$$
$$+ \cdots + \frac{n!}{(n-2)!1!}t^{n-2}(1-t)^1 - \frac{n!}{(n-1)!0!}t^{n-1}(1-t)^0 + \frac{n!}{(n-1)!0!}t^{n-1}(1-t)^0\Big] =$$
$$= \frac{n!}{k!(n-k-1)!}t^k(1-t)^{(n-k-1)} = \frac{(2k+1)!}{k!k!}t^k(1-t)^k.$$

Последнее выражение хорошо известно и представляет собой ф.пл. бета-распределения $\mathcal{B}\mathtt{et}(k+1,k+1)$. Пусть $\mathcal{B}(t\,|\,k+1,k+1)$ — ф.р. соответствующего бета-закона, тогда ф.р. выборочной медианы задаётся так:

$$\mathbf{P}\left\{\xi_{(k+1)} \leqslant x\right\} = \mathcal{B}\big(F(x)\,|\,k+1,k+1\big).$$

Число $\mathfrak{m}$ есть медиана (непрерывной справа) ф.р. F т. т. т. когда $F(\mathfrak{m}) \geqslant {}^1\!/_2$ и $\lim_{x\to\mathfrak{m}-0} F(x) \leqslant {}^1\!/_2$. Из предыдущего представления вполне очевидно, что тогда $\mathfrak{m}$ — медиана ф.р. $\xi_{(k+1)}$.

Если $F(x)=x$, то при $n=2k+1$ математическое ожидание выборочной медианы $\mathbf{E}\xi_{(k+1)} = {}^1\!/_2$, а дисперсия $\mathbf{D}\xi_{(k+1)} = 1/(12+8k) \asymp (2\sqrt{2k})^{-2} \asymp (2\sqrt{n})^{-2}$. После элементарных преобразований получаем, что плотность нормированной выборочной медианы $\big(\xi_{(k+1)} - \frac{1}{2}\big)2\sqrt{2k}$ равна

$$\frac{(2k+1)!}{k!\,k!\,\sqrt{8k}\,2^{2k}}\Big(1-\frac{y^2}{2k}\Big)^k \underset{k\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}y^2}$$

в силу формулы Стирлинга. Таким образом, в соответствии с леммой Шеффе 263, стр. 226, распределение выборочной медианы при выборе из равномерного закона асимптотически нормально $\mathcal{N}\big(\frac{1}{2},\frac{1}{4n}\big)$. ◎

135| Упр. Пусть $\vec{\zeta}_n = (\xi_{(1)},\ldots,\xi_{(n)})$ — вариационный ряд, построенный по выборке объёма n из равномерного распределения. Докажите, что сл.вектор $\big(\frac{\xi_{(1)}}{\xi_{(n)}},\ldots,\frac{\xi_{(n-1)}}{\xi_{(n)}}\big) \backsim \vec{\zeta}_{n-1}$ и не зависит от $\xi_{(n)}$.

A. Доказательство леммы о линейном преобразовании

⊨⇒ 109, стр. 116. Предположим сначала, что в матрице Q все главные миноры от нижнего правого угла невырождены. Рассмотрим итерационный процесс приведения матрицы Q к единичной матрице вида $Q_i = A_iQ_{i-1}$, $i = 1,\ldots,k$, $Q_0 = Q$, где строки матрицы A_i, кроме i-й, совпадают со строками единичной матрицы, а i-я строка равна i-й строке обратной матрицы Q_{i-1}^{-1}. Легко понять, что матрица Q_i будет отличаться от матрицы Q_{i-1} только i-й строкой, в которой все элементы равны нулю, кроме i-го элемента, равного единице. В силу предположения относительно матрицы Q на каждом шаге такого процесса $|Q_i| \neq 0$ и $|A_i| \neq 0$. Поскольку по построению $A_k A_{k-1} \cdots A_1 Q = \mathbb{I}$, то имеем представление для матрицы $Q = A_1^{-1} \cdots A_k^{-1}$.

Структура матрицы A_i^{-1} аналогична A_i, в частности её i-й диагональный элемент (равный определителю матрицы) отличен от нуля.

Пусть $f(\vec{x})$ — плотность сл.вектора $\vec{\xi}$. Найдём ф.р. сл.вектора $\vec{\eta} = A_1^{-1}\vec{\xi} = (\vec{a}^\flat \vec{\xi}, \xi_2, \ldots, \xi_k)^\flat$, где $\vec{a}^\flat = (a_1, \ldots, a_k)$ — первая строка A_1^{-1}, $a_1 = |A_1^{-1}|$. Первая строка A_1 равна $\vec{b}^\flat = (1, -a_2, \ldots, -a_k)/a_1$. Отсюда

$$\begin{aligned} F_{\vec{\eta}}(\vec{y}) &= \mathbf{P}\left\{\eta_1 \leqslant y_1,\, \eta_2 \leqslant y_2, \ldots, \eta_k \leqslant y_k\right\} = \\ &= \mathbf{P}\left\{\, a_1\xi_1 + \ldots + a_k\xi_k \leqslant y_1,\, \xi_2 \leqslant y_2, \ldots, \xi_k \leqslant y_k \,\right\} = \\ = \int\limits_{\substack{x_i \leqslant y_i,\\ i=\overline{2,k}}} \ldots \int & \Big[\int_{\mathbb{R}} \dot{\mathbb{I}}\, \big(a_1 x_1 + \ldots + a_k x_k \leqslant y_1\big) f(\vec{x})\, dx_1 \Big]\, dx_2 \cdots dx_k. \end{aligned}$$

Во внутреннем интеграле (в квадратных скобках) при фиксированных $(x_2, \ldots, x_k)$ осуществим линейную замену переменной интегрирования x_1 : $a_1x_1 + a_2x_2 + \ldots + a_kx_k \dashrightarrow u_1$, что можно записать в виде $x_1 = \vec{b}^\flat \vec{u}$ с вектор-столбцом переменных $\vec{u} = (u_1, x_2, \ldots, x_k)^\flat$. Таким образом, согласно теореме 543, стр. 484 о замене переменных в интеграле Лебега, имеем

$$F_{\vec{\eta}}(\vec{y}) = \frac{1}{|\,a_1|} \int\limits_{\substack{x_i \leqslant y_i,\\ i=\overline{2,k}}} \ldots \int \Big[\int\limits_{u_1 \leqslant y_1} f(\vec{b}^\flat \vec{u}, x_2, \ldots, x_k)\, du_1 \Big]\, dx_2 \cdots dx_k.$$

Следовательно, функция плотности $\vec{\eta} = A_1^{-1}\vec{\xi}$ равна

$$f_{\vec{\eta}}(\vec{y}) = \frac{1}{|\,a_1|} f(b^\flat \vec{y}, y_2, \ldots, y_k) = \big\|A_1\big\|\, f\big(A_1\vec{y}\big).$$

Аналогично для $A_j, j > 1$.

Применим полученный результат последовательно к матрицам $A_k^{-1}, \ldots, A_1^{-1}$ из представления матрицы Q:

ф.пл.	$\vec{\eta}^{(1)} = \big|A_k^{-1}\big|\,\vec{\xi}$	равна	$\big\|A_k\big\|\; f(A_k\vec{y})$,
ф.пл.	$\vec{\eta}^{(2)} = A_{k-1}^{-1}A_k^{-1}\vec{\xi} = A_{k-1}^{-1}\vec{\eta}^{(1)}$	равна	$\big\|A_{k-1}\big|\,\big|A_k\big\|\; f(A_kA_{k-1}\vec{y})$, $\ldots$,
ф.пл.	$A_1^{-1} \cdots A_k^{-1}\vec{\xi}$	равна	$\big\|A_1\big|\cdots\big|A_k\big\|\; f(A_k \cdots A_1\vec{y})$.

Поскольку, как отмечено выше, $A_k \cdots A_1 = Q^{-1}$, что и доказывает лемму.

Воспользовавшись разложением определителя матрицы по первому столбцу, легко понять, что всегда можно так переставить строки невырожденной матрицы, чтобы все миноры от нижнего правого угла матрицы имели ненулевой определитель. Любая такая перестановка эквивалентна умножению матрицы Q слева на матрицу J, получающуюся из единичной матрицы соответствующей перестановкой строк. Для матрицы J и её обратной $J^{-1} = J^{\flat}$ утверждение леммы доказывается (?!) простыми рассуждениями. Так как $\|J\| = 1$, то вместе с предыдущим имеем

$$JQ\vec{\xi} \backsim \frac{1}{\|Q\|} f(Q^{-1}J^{-1}\vec{y}) \Rightarrow$$
$$\Rightarrow Q\vec{\xi} = J^{-1}(EQ\vec{\xi}) \backsim \frac{1}{\|Q\|} f(Q^{-1}J^{-1}J\vec{y}) = \frac{1}{\|Q\|} f(Q^{-1}\vec{y}),$$

что и требовалось. ⋠

B. Упражнения

Вид распределений сл.в. для следующих задач приведён на стр. 107.

Упр. B.1. Найти функцию вероятностей:

(a) $[\![\xi]\!] + 1$, если $\xi \backsim \mathcal{E}\mathrm{x}(1)$;

(b) ξ, если $\xi \backsim \mathcal{P}\mathrm{ois}(\vartheta)$ со случайным параметром $\vartheta \backsim \mathcal{E}\mathrm{x}(1)$;

(c) $\xi_1 - \xi_2$, если $\xi_j \backsim \mathcal{P}\mathrm{ois}(1)$, $j = 1, 2$, и независимы.

Упр. B.2. Найти плотность распределения:

(a) ξ^{-1}, если $\xi \backsim \mathcal{C}\mathrm{auch}(0, 1)$;

(b) $|\xi|$, если $\xi \backsim \mathcal{L}\mathrm{ap}(0, 1)$;

(c) ξ, если $\ln \xi \backsim \mathcal{N}(0, 1)$ (логнормальное распределение);

(d) ξ, если $\xi \backsim \mathcal{E}\mathrm{x}(1/\vartheta)$ со случайным параметром $\vartheta \backsim \mathcal{E}\mathrm{x}(1)$;

(e) (ρ, ϑ), где $\rho = \xi^2 + \eta^2$, $\vartheta = \xi/\eta$, если $(\xi, \eta) \backsim \mathcal{U}\mathrm{n}(\Omega)$, где Ω — единичный круг с центром в начале координат;

(f) $\min(\xi_1, \ldots, \xi_k)$, если $\xi_j \backsim \mathcal{E}\mathrm{x}(1)$, $j = \overline{1, k}$, и независимы;

(g) $\xi_1 - \xi_2$, если $\xi_j \backsim \mathcal{E}\mathrm{x}(\lambda_j)$, $j = 1, 2$, и независимы;

(h) $\xi_1/(\xi_1 + \xi_2)$, если $\xi_j \backsim \mathcal{E}\mathrm{x}(1)$, $j = 1, 2$, и независимы.

(i) $\vec{\eta} = (\xi_1/\xi_k, \ldots, \xi_{k-1}/\xi_k)$, если $\xi_j \backsim \mathcal{E}\mathrm{x}(1)$, $j = \overline{1, k}$, и независимы;

Упр. B.3. Найти $h(x)$ такое, что $h(\xi) \backsim \mathcal{C}\mathrm{auch}(0, 1)$, если $\xi \backsim \mathcal{E}\mathrm{x}(1)$.

Упр. B.4. Пусть $\left\langle \xi_k \backsim F \right\rangle_1^n$ — независимые сл.в. с общей абсолютно-непрерывной ф.р. F. Показать, что при $\forall\, k \leqslant n$ справедливы соотношения:

$$\text{(а)} \qquad \mathbf{P}\left\{ \xi_k \leqslant x \mid \xi_{(n)} = t \right\} \;=\; \begin{cases} \dfrac{n-1}{n}\dfrac{F(x)}{F(t)} & \text{при} \quad x < t, \\ \quad 1 & \text{при} \quad x \geqslant t; \end{cases}$$

$$\text{(б)} \qquad \mathbf{E}\left[\xi_k \mid \xi_{(n)} = t \right] \;=\; \frac{n-1}{n}\frac{1}{F(t)}\int_{-\infty}^{t} x\,dF(x) + \frac{t}{n}\,;$$

$$\text{(в)} \qquad \mathbf{P}\left\{ \xi_{(1)} > x \mid \xi_{(n)} = t \right\} \;=\; \left(1 - \frac{F(x)}{F(t)}\right)^{n-1}, \quad x < t.$$

Упр. B.5. Пусть ξ_1, ξ_2 — независимые сл.в. Доказать, что

а) при $\forall\, x, y \in \mathbb{R}^1$ справедливы неравенства:

$$\mathbf{P}\{\xi_1 + \xi_2 \leqslant x\} \;\geqslant\; \mathbf{P}\{\xi_1 \leqslant x - y\}\,\mathbf{P}\{\xi_2 \leqslant y\}\,,$$
$$\mathbf{P}\{\xi_1 + \xi_2 > x\} \;\geqslant\; \mathbf{P}\{\xi_1 > x - y\}\,\mathbf{P}\{\xi_2 > y\}\,;$$

б) если ф.р. ξ_1 непрерывна всюду, то $\mathbf{P}\{\xi_1 = \xi_2\} = 0$;

в) если $\xi_1 \backsim \xi_2$, то сл.в. $\eta = \xi_1 - \xi_2$ имеет симметричное распределение;

г) если $\xi_1, \xi_2 \backsim \mathcal{E}\mathrm{x}(\lambda)$, то сл.в. ξ_1/ξ_2 и $\xi_1 + \xi_2$ также независимы.

Упр. B.6. Доказать утверждение леммы 66, стр. 72.

Упр. B.7. Какое распределение имеет отношение ξ/η, если независимые одинаково распределённые сл.в. ξ, η имеют плотность $\dfrac{\sqrt{2}}{\pi(1+x^4)}, x \in \mathbb{R}^1$?

C. Указания к решению задач

✦ (?!), стр. 111. Проверить свойства вероятностной меры. Например,
$\mathsf{P}_\eta\{B_1 \uplus B_2\} \overset{\mathrm{df}}{=} \mathsf{P}_\xi\{h^{-1}(B_1 \uplus B_2)\} = \mathsf{P}_\xi\{h^{-1}(B_1)\} + \mathsf{P}_\xi\{h^{-1}(B_2)\} = \mathsf{P}_\eta\{B_1\} + \mathsf{P}_\eta\{B_2\}$.

✦ 101, стр. 112. Неравенство $\{\eta \leqslant y\} \Leftrightarrow \{\xi < (y-a)/b\}^c$. (!) Знак «$<$».

✦ 109, стр. 116. Применить теорему 107, стр. 115.

✦ 110, стр. 116. Найти обратное преобразование.

✦ 115, стр. 119. Воспользоваться теоремой 112.

✦ 118, стр. 120. Подынтегральное выражение в формуле для свёртки отлично от нуля в первом случае при $y \geqslant 0, u - y \geqslant 0$, во втором случае при $y \geqslant 0, u + y \geqslant 0$.

✦ 120, стр. 120. Применить (6), стр. 120. При отрицательном u произвести замену переменной интегрирования. (в) Первый способ: показать, что обе части имеют треугольное распределение на отрезке $[-1;1]$. Другой способ: заметить, что $1 - \xi_2 \backsim \mathcal{U}\mathrm{n}(0,1)$.

✦ 121, стр. 121. В формуле (3), стр. 118, записать математическое ожидание дискретной сл.в. в виде ряда.

✦ 123, стр. 121. Применить теорему Лебега к (2), стр. 118.

✦ 124, стр. 121. Применить тождество $\mathbf{P}\{\,\xi+\eta=u\,\}=\sum_j\mathbf{P}\{\,\xi+\eta=u,\eta=y_j\,\}$.

✦ (?!), стр. 123. Воспользоваться представлением события
$\{(\eta_0+\eta_1,\eta_0+\eta_2)=(x_1,x_2)\}=\bigcup_{x_0\leqslant\min(x_1,x_2)}\{\eta_0=x_0,\eta_1=x_1-x_0,\eta_2=x_2-x_0\}$.

✦ 128, стр. 123. В выражении для частной вероятности $\mathbf{P}\{\xi_1=x_1\}=\sum_{x_2=0}^{\infty}p(x_1,x_2)$ при фиксированном x_1 произвести замену порядка суммирования; бесконечный ряд даст значение e^{λ_2}, а оставшийся конечный ряд после соответствующей нормировки совпадёт с разложением $(\lambda_0+\lambda_1)^{x_1}$ по формуле бинома Ньютона.

✦ 130, стр. 125. Подынтегральное выражение плотности в теореме 129, стр. 124, равно $|y|\exp\big(-\frac{1}{2}y^2(z^2+1)\big)dy=\exp\big(-\frac{1}{2}y^2(z^2+1)\big)d\,\frac{1}{2}y^2$.

✦ (?!), стр. 125. $\{\xi_{(1)}>x\}$ т. т. т. когда все сл.в. $\xi_k>x$, $k=\overline{1,n}$.

✦ 133, стр. 127. С помощью мультиномиальной модели найти вероятность попадания ровно i сл.в. в интервал $(-\infty;u_1]$, ровно j сл.в. в интервал $(u_1;u_2]$ и ровно $n-i-j$ сл.в. в интервал $(u_2;\infty)$.

✦ 135, стр. 130. Якобиан преобразования $x_1=u_nu_1,\ldots,x_{n-1}=u_nu_{n-1},x_n=u_n$ равен $J=u_n^{n-1}$, поэтому совместная плотность искомого вектора и сл.в. $\xi_{(n)}$ равна $(n-1)!\,n\,u_n^{n-1}$, $0\leqslant u_1\leqslant\ldots\leqslant u_{n-1}\leqslant1$, $0\leqslant u_n\leqslant1$.

✦ B.1. a) $\xi\backsim\mathcal{G}\mathrm{eo}(e^{-1})$; b) $\mathbf{P}\{\xi=k\}=2^{-(k+1)}$, применить 93, стр. 97;

c) $\mathbf{P}\{\xi=k\}=e^{-2}\sum_{j=\max(0,-k)}^{\infty}1/j!(j+k)!$ (функция Бесселя).

✦ B.2. a) $\mathcal{C}\mathrm{auch}(0,1)$;

b) $\mathcal{E}\mathrm{x}(1)$, найти ф.р.; d) найти частную плотность из 94, стр. 98;

e) независимы, $\rho\backsim\mathcal{U}\mathrm{n}(0,1)$, $\vartheta\backsim\mathcal{C}\mathrm{auch}(0,1)$, найти площадь части $x^2+y^2<r,x/y<\theta$ единичного круга; f) $\mathcal{E}\mathrm{x}(1/k)$, см. указание к (?!), стр. 125, выше;

h) $\mathcal{U}\mathrm{n}(0,1)$, найти частную плотность вектора $(\xi_1/(\xi_1+\xi_2),\xi_2)$;

i) $f_{\vec{\eta}}(y_1,\ldots,y_{k-1})=(k-1)!(y_1+\ldots+y_{k-1}+1)^{-k}$, $\vec{y}\geqslant0$.

✦ B.3. Проверить и применить следующие утверждения: если $\zeta\backsim F$ с непрерывной, строго монотонной ф.р. F и $\omega\sim\mathcal{U}\mathrm{n}(0,1)$, то $F(\zeta)\backsim\omega$, $F^{-1}(\omega)\backsim\zeta$.

✦ B.4. По аналогии с 92, стр. 97, проверить (38), стр. 95.

✦ B.5. a) $\int_{\mathbb{R}}F_1(x-t)dF_2(t)\geqslant\int_{-\infty}^{y}F_1(x-t)dF_2(t)\geqslant F_1(x-y)F_2(y)$. Ещё проще так: $\{\xi_1<_{(>)}a\}\cap\{\xi_2<_{(>)}b\}\Rightarrow\{\xi_1+\xi_2<_{(>)}a+b\}$; б) воспользоваться 123, стр. 121;

в) т.к. $(\xi_1,\xi_2)\backsim(\xi_2,\xi_1)$, то $\mathsf{P}\{\xi_1-\xi_2\leqslant y\}=\mathsf{P}\{\xi_2-\xi_1\leqslant y\}=\mathsf{P}\{-(\xi_1-\xi_2)\leqslant y\}$;

г) преобразование $u=x/y,v=x+y$ имеет обратное $x=uv/(1+u),y=v/(1+u)$. Найти якобиан; применив теорему 107, стр. 115, представить плотность $(\xi/\eta,\xi+\eta)$ в виде произведения плотностей относительно переменных u и v.

✦ B.6. Рассмотреть преобразование $x_j=y_j/(y_1+\ldots+y_{k+1})$, $j=\overline{1,k}$, $x_{k+1}=$ $=(y_1+\ldots+y_{k+1})$; показать, что якобиан обратного преобразования равен x_{k+1}^k (см. пример 108, стр. 115); искомую плотность вычислить как частную плотность совместного распределения $(\xi_1,\ldots,\xi_{k+1})$.

✦ B.7. $\mathcal{C}\mathrm{auch}(0,1)$.

III Условное математическое ожидание

Условное математическое ожидание представляет собой один из самых полезных инструментов в арсенале методов теории вероятностей. В этой главе введём понятия условного математического ожидания и условного распределения для общих вероятностных моделей. Заметим, что значение условной вероятности допускает привычную частотную интерпретацию только для дискретных моделей с положительной вероятностью условия. Представляется невозможным эксперимент, в котором некое событие нулевой вероятности может повториться хотя бы дважды. В таких ситуациях на условное математическое ожидание следует смотреть как на функцию случайного значения, полученного в эксперименте, и все свойства этого ожидания воспринимать только в усреднённом смысле. Так, если попытаться спрогнозировать значение одной случайной величины по реализациям другой случайной величины с непрерывным распределением, то для каждой конкретной реализации, ввиду «ничтожности» произошедшего, прогноз может обладать непредсказуемыми свойствами. С другой стороны, можно построить прогноз, у которого усреднённая ошибка (по всем возможным реализациям) будет минимальна. Классическая формула полной вероятности есть не что иное, как способ вычисления вероятности события посредством усреднения значений условной вероятности по всем возможным реализациям из полной группы событий. Для дискретных или абсолютно непрерывных вероятностных моделей условное ожидание может быть определено через соответствующее распределение (см. стр. 92). Однако стремление к общности, обусловленное, например, потребностями теории случайных процессов, диктует необходимость введения такого определения, которое было бы связано с вероятностной мерой основного вероятностного пространства $(\Omega, \mathscr{F}, \mathbf{P})$ и не зависело от способа описания распределения сл.в. в $\mathbb{R}^k$. В качестве первого шага в этом направлении заметим, что по теореме 516, стр. 469, возможность представления функции h на Ω в виде $h(\omega) = h(g(\omega))$ с некоторой фиксированной функцией g эквивалентна измеримости h относительно σ-алгебры на Ω, порождённой g.

§ 1. Условное ожидание относительно σ-алгебры

Рассмотрим сначала класс неотрицательных сл.в., определённых на одном вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$. Кроме того, будем понимать эти сл.в. в расширенном смысле, т.е. для них допускается значение $+\infty$ (возможно, с положительной вероятностью). Напомним, что математическое ожидание (интеграл Лебега) неотрицательной сл.в., принимающей с положительной вероятностью бесконечное значение, полагается равным $+\infty$. Измеримость функции $\xi : \Omega \mapsto (\overline{\mathbb{R}}, \mathscr{B}(\overline{\mathbb{R}}))$ относительно некоторой σ-алгебры $\mathscr{L}$ подмножеств Ω ($\mathscr{L}$-измеримость) будем обозначать $\xi \Subset \mathscr{L}$.

Определения. Пусть ξ — неотрицательная сл.в., $\mathscr{L}$ — σ-подалгебра $\mathscr{F}$ ($\mathscr{L} \subset \mathscr{F}$). *Условным математическим ожиданием (усл.м.о.)* ξ относительно $\mathscr{L}$ называется $\mathscr{L}$-измеримая сл.в. $\eta \geqslant 0$ такая, что

$$\mathbf{E}\left[\eta\, \dot{\mathrm{I}}_A\right] = \mathbf{E}\left[\xi\, \dot{\mathrm{I}}_A\right], \quad \forall A \in \mathscr{L}. \tag{1}$$

Обозначается $\eta = \mathbf{E}\left[\xi \,|\, \mathscr{L}\right]$. Усл.м.о. $\mathbf{E}\left[\, \dot{\mathrm{I}}_B \,|\, \mathscr{L}\,\right]$ называется *условной вероятностью (усл.вер.)* $B \in \mathscr{F}$ относительно $\mathscr{L}$ и обозначается $\mathbf{P}\left\{B \,|\, \mathscr{L}\right\}$.

136| Теорема. I. *Для любой неотрицательной сл.в.* ξ *и любой сигма-подалгебры* $\mathscr{L} \subset \mathscr{F}$ *существует неотрицательная* $\mathscr{L}$*-измеримая сл.в.* η, *удовлетворяющая* (1)*; любая сл.в.* $\widetilde{\eta} \Subset \mathscr{L}$, *отличающаяся от* η *на множестве* $\mathbf{P}$*-меры нуль, также представляет вариант усл.м.о.* $\mathbf{E}\left[\xi \,|\, \mathscr{L}\right]$.

II. *Если сл.в.* ξ *интегрируема, то существует вариант усл.м.о.* $\mathbf{E}\left[\xi \,|\, \mathscr{L}\right]$, *принимающий конечные значения для* $\forall\, \omega \in \Omega$.

III. *Для любого события* $B \in \mathscr{F}$ *существует вариант усл.вер.* $\mathbf{P}\left\{B \,|\, \mathscr{L}\right\} \in [0; 1]$ *при* $\forall\, \omega \in \Omega$.

⊫ I. Введём вспомогательное вероятностное пространство $(\Omega, \mathscr{L}, \mathbf{P})$ с той же мерой $\mathbf{P}$, суженной до σ-алгебры $\mathscr{L} \subset \mathscr{F}$. Определим меру на пространстве $(\Omega, \mathscr{L})$:

$$\mu(A) = \int_A \xi \, d\mathbf{P}, \qquad A \in \mathscr{L}.$$

По свойству интеграла Лебега эта мера абсолютно непрерывна относительно меры $\mathbf{P}$, т.е. $\mu(A) = 0$, если $\mathbf{P}\{A\} = 0$. В силу теоремы Радона–Никодима (см. 558, стр. 491, или вариант из замечания 560, стр. 492) существует почти наверное единственная $\mathscr{L}$-измеримая сл.в. $\eta = \dfrac{d\mu}{d\mathbf{P}}$ такая, что $\mu(A) = \int_A \eta \, d\mathbf{P}$, т.е. η удовлетворяет определению усл.м.о.

Второе утверждение следует из свойств математического ожидания, поскольку, очевидно, $\eta \dot{\mathrm{I}}_A = \widetilde{\eta} \, \dot{\mathrm{I}}_A$ (п.н.), стало быть, $\mathbf{E}\big[\eta \dot{\mathrm{I}}_A\big] = \mathbf{E}\big[\widetilde{\eta} \, \dot{\mathrm{I}}_A\big]$.

II. Если ξ интегрируема, то мера μ конечна:

$$\mu(A) \; = \; \int_A \eta \, d\mathbf{P} \; = \; \int_A \xi \, d\mathbf{P} \; \leqslant \; \int_\Omega \xi \, d\mathbf{P} \; < \; \infty.$$

Поэтому вероятность $\mathbf{P}\{\eta = \infty\} = 0$. Если $\eta = \mathbf{E}\big[\xi \,\big|\, \mathscr{L}\big]$, то в силу предыдущего $\eta \, \dot{\mathrm{I}}(\eta < \infty)$ также есть вариант усл.м.о.

III. Пусть $\eta = \mathbf{P}\big\{B \,\big|\, \mathscr{L}\big\}$ — некоторый вариант усл.вер. (существующий по первому пункту данной теоремы). Множество $A = \{\omega : \eta(\omega) > 1\} \in \mathscr{L}$, поэтому в силу определения η

$$\mathbf{P}\{A\} \; \leqslant \; \mathbf{E}\big[\eta \dot{\mathrm{I}}_A\big] \; = \; \mathbf{E}\big[\dot{\mathrm{I}}_B \dot{\mathrm{I}}_A\big] \; \leqslant \; \mathbf{P}\{A\},$$

где первое неравенство будет строгим, если $\mathbf{P}\{A\} > 0$. Итак, $\mathbf{P}\{\eta > 1\} = 0$, поэтому сл.в. $\widetilde{\eta} = \eta \, \dot{\mathrm{I}}(\eta \leqslant 1) = \mathbf{P}\big\{B \,\big|\, \mathscr{L}\big\}$ и $0 \leqslant \widetilde{\eta} \leqslant 1$. ⫤

137 Упр. Если для интегрируемых $\mathscr{L}$-измеримых сл.величин η_1, η_2 равенство $\mathbf{E}\big[\eta_1 \dot{\mathrm{I}}_A\big] = \mathbf{E}\big[\eta_2 \dot{\mathrm{I}}_A\big]$ имеет место при $\forall \, A \in \mathscr{L}$, то $\mathbf{P}\{\eta_1 \neq \eta_2\} = 0$, т.е. $\eta_1 = \eta_2$ (п.н.) .

Подсказка. $|\eta_1 - \eta_2| = (\eta_1 - \eta_2) \, \dot{\mathrm{I}}_A - (\eta_1 - \eta_2) \, \dot{\mathrm{I}}_{A^c}$ при $A = \{\eta_1 > \eta_2\}$.

⚠ ✓ Условие (1) может выполняться для нескольких сл.в. Поэтому в действительности усл.м.о. $\mathbf{E}\big[\xi \,\big|\, \mathscr{L}\big]$ есть семейство сл.в., и любое соотношение типа $\eta = \mathbf{E}\big[\xi \,\big|\, \mathscr{L}\big]$ должно быть записано в виде $\eta \in \mathbf{E}\big[\xi \,\big|\, \mathscr{L}\big]$. Поскольку в силу 137 (или первого пункта 136) любые два варианта из семейства усл.м.о. отличаются лишь на множестве нулевой вероятности, то вместо этого принято писать $\eta = \mathbf{E}\big[\xi \,\big|\, \mathscr{L}\big]$ (п.н.). Заметим, что если множество $B \in \mathscr{L}$ и $\mathbf{P}\{B\} > 0$, причём ни одно $\mathbf{P}$-нулевое подмножество $A \subset B$ не входит в $\mathscr{L}$, то все варианты усл.м.о. имеют одно и то же значение при всех $\omega \in B$. Если σ-алгебра $\mathscr{L}$ состоит только из таких множеств, то семейство $\mathbf{E}\big[\,\xi \mid \mathscr{L}\,\big]$ содержит только одну сл.в. и уточняющая запись (п.н.) не нужна.

В дальнейшем мы будем добавлять это уточнение только в тех случаях, когда могут возникнуть проблемы с толкованием выражения, содержащего усл.м.о., например в соотношениях, подобных 139, a. Здесь это означает, что любой вариант усл.м.о. левой части неравенства почти наверное не превосходит любого из вариантов усл.м.о. правой части.

✓ Усл.м.о. есть сл.величина, и её значение зависит от реализаций ω вероятностного эксперимента. Чтобы подчеркнуть этот факт, иногда пишут $\mathbf{E}\big[\,\xi \mid \mathscr{L}\,\big](\omega)$, особенно если требуется выделить какой-либо вариант усл.м.о.

✓ Если сл.в. ξ $\mathscr{L}$-измерима, то, очевидно, она может быть взята в качестве варианта усл.м.о. $\mathbf{E}\big[\,\xi \mid \mathscr{L}\,\big]$.

✓ Если $\eta = \mathbf{E}\big[\,\xi \mid \mathscr{L}\,\big]$ и $\widetilde{\eta} = \eta$ (п.н.), но $\widetilde{\eta} \notin \mathscr{L}$, то равенство $\mathbf{E}\big[\widetilde{\eta}\,\dot{\mathrm{I}}_A\big] = \mathbf{E}\big[\xi\,\dot{\mathrm{I}}_A\big]$ также справедливо для $\forall A \in \mathscr{L}$, однако $\widetilde{\eta}$ не может считаться усл.м.о. $\mathbf{E}\big[\,\xi \mid \mathscr{L}\,\big]$. В целях исключения подобных ситуаций часто рассматривают только полные σ-алгебры, содержащие все подмножества всех нулевых событий относительно меры $\mathbf{P}$.

138| Примеры. 1) Относительно σ-алгебры $\mathscr{L} = \big\langle \varnothing, \Omega \big\rangle$ измеримы только постоянные «случайные» величины $\eta(\omega) \equiv b$. Из определения усл.м.о. сразу следует, что в этом случае $\mathbf{E}\big[\,\xi \mid \mathscr{L}\,\big](\omega) = \mathbf{E}\,\xi$ при всех $\omega \in \Omega$.

2) Если σ-подалгебра $\mathscr{L}$ порождена множеством $U \in \mathscr{F}$, т.е. $\mathscr{L} = \sigma\big(U\big) = \big\langle \varnothing, \Omega, U, U^c \big\rangle$, то любая $\mathscr{L}$-измеримая сл.в. η обязана принимать постоянные значения на множествах U и U^c, т.е. $\mathbf{E}\big[\,\xi \mid \mathscr{L}\,\big](\omega) = b_1\,\dot{\mathrm{I}}_U(\omega) + b_2\,\dot{\mathrm{I}}_{U^c}(\omega)$. Выбирая в (1) множество $A = U$, получаем $b_1\,\mathbf{E}[\dot{\mathrm{I}}_U] = \mathbf{E}[\xi\,\dot{\mathrm{I}}_U]$; для $A = U^c$ имеем $b_2\,\mathbf{E}[\dot{\mathrm{I}}_{U^c}] = \mathbf{E}[\xi\,\dot{\mathrm{I}}_{U^c}]$.

Таким образом, если $0 < \mathbf{P}(U) = \mathbf{E}[\dot{\mathrm{I}}_U] < 1$, то для $\forall\,\omega \in \Omega$

$$\mathbf{E}\big[\,\xi \mid \sigma\big(U\big)\,\big](\omega) \;=\; \frac{\mathbf{E}[\xi\,\dot{\mathrm{I}}_U]}{\mathbf{P}\{U\}}\,\dot{\mathrm{I}}_U(\omega) + \frac{\mathbf{E}[\xi\,\dot{\mathrm{I}}_{U^c}]}{\mathbf{P}\{U^c\}}\,\dot{\mathrm{I}}_{U^c}(\omega).$$

В частности, полагая здесь $\xi = \dot{\mathrm{I}}_B$, получаем, что, если происходит U, т.е.

$\omega \in U$, тогда усл.вер. $\mathbf{P}\{B \,|\, \sigma(U)\}(\omega)$, как и следовало ожидать, равна классической усл.вер. $\mathbf{P}\{B \,|\, U\} = \mathbf{P}\{BU\}/\mathbf{P}\{U\}$.

3) Если $\mathscr{L} = \sigma(U_j, j = \overline{1,N})$ — σ-алгебра, порождённая разбиением (конечным или счётным) $\Omega = \biguplus_1^N U_j$, $N \leqslant \infty$, причём $\mathbf{P}\{U_j\} > 0$ при $\forall j$, то

$$\mathbf{P}\{B \,|\, \mathscr{L}\}(\omega) \;=\; \sum\nolimits_1^N \mathbf{P}\{B \,|\, U_j\}\, \dot{\mathrm{I}}_{U_j}(\omega), \qquad \forall\, \omega \in \Omega.$$

В случае, когда $\mathbf{P}\{U_j\} = 0$ при некоторых $1 \leqslant j \leqslant N$, усл.вер. $\mathbf{P}\{B \,|\, \mathscr{L}\}(\omega) = \sum_1^N b_j\, \dot{\mathrm{I}}_{U_j}(\omega)$ с $b_j = \mathbf{P}\{B \,|\, U_j\}$, если $\mathbf{P}\{U_j\} > 0$, и произвольным b_j, если $\mathbf{P}\{U_j\} = 0$.

Если в определении усл.вер. для $B \in \mathscr{F}$ относительной этой σ-алгебры $\mathscr{L}$ выбрать $A = \Omega$, то получим формулу полной вероятности:

$$\mathbf{P}\{B\} \;=\; \mathbf{E}\big[\mathbf{P}\{B \,|\, \mathscr{L}\}\big] \;=\; \sum\nolimits_1^N \mathbf{P}\{B \,|\, U_j\}\, \mathbf{P}\{U_j\}\,.$$

Следует заметить, что эта формула остаётся справедливой при любом определении усл.вер. $\mathbf{P}\{B \,|\, U_j\}$ $(\in [0;1])$, когда $\mathbf{P}\{U_j\} = 0$.

4) Пример 147, стр. 149, показывает, что в некоторых случаях и для неинтегрируемых сл.в. ξ существует всюду конечный вариант усл.м.о. ◎

Прежде чем обратиться к свойствам усл.м.о., приведём вспомогательные утверждения, обобщающие известные теоремы Б. Леви и Фату.

139| Лемма. *Пусть* $\boldsymbol{\xi} = \langle \xi_n \geqslant 0 \rangle_1^\infty$ *— последовательность неотрицательных сл.в.,* $\langle \mathbf{E}[\xi_n \,|\, \mathscr{L}] \rangle_1^\infty$ *— соответствующая последовательность (произвольных вариантов) усл.м.о. Тогда:*

(a) *если* $\xi_1 \leqslant \xi_2$, *то* $\mathbf{E}[\xi_1 \,|\, \mathscr{L}] \leqslant \mathbf{E}[\xi_2 \,|\, \mathscr{L}]$ (п.н.);

(b) *если* $\boldsymbol{\xi}$ *неубывающая последовательность, т.е.* $\xi_n \leqslant \xi_{n+1}$ (п.н.), $n \geqslant 1$, *и* $\xi_n \xrightarrow{\text{п.н.}} \xi_0$, *то п.н. существует предел*

$$\lim_n \mathbf{E}[\xi_n \,|\, \mathscr{L}] = \mathbf{E}[\xi_0 \,|\, \mathscr{L}];$$

(c) $\varliminf\limits_n \mathbf{E}[\xi_n \,|\, \mathscr{L}] \geqslant \mathbf{E}[\varliminf\limits_n \xi_n \,|\, \mathscr{L}]$ (п.н.).

⊨ (a) Множество $A = \{\omega : \mathbf{E}[\xi_1 \,|\, \mathscr{L}] > \mathbf{E}[\xi_2 \,|\, \mathscr{L}]\} \in \mathscr{L}$, поэтому

$$\mathbf{E}[\xi_1\, \dot{\mathrm{I}}_A] \;=\; \mathbf{E}\big[\mathbf{E}[\xi_1 \,|\, \mathscr{L}]\, \dot{\mathrm{I}}_A\big] \;\geqslant\; \mathbf{E}\big[\mathbf{E}[\xi_2 \,|\, \mathscr{L}]\, \dot{\mathrm{I}}_A\big] \;=\; \mathbf{E}[\xi_2\, \dot{\mathrm{I}}_A]$$

со строгим неравенством, если $\mathbf{P}\{A\} \neq 0$, что невозможно.

(b) Пусть $\eta_n = \mathbf{E}[\xi_n \,|\, \mathscr{L}]$, $n \geqslant 1$. В силу (a) множество U^c тех $\omega \in \Omega$, где нарушается монотонность последовательности η_n, имеет нулевую вероятность. Поэтому последовательность $\eta_n \dot{\mathrm{I}}_U$ не убывает при всех $\omega \in \Omega$. Пусть $\eta_0 = \lim_n \eta_n$; покажем, что $\eta_0 = \mathbf{E}\big[\xi_0 \,|\, \mathscr{L}\big]$. Выберем произвольное $A \in \mathscr{L}$, тогда в силу теоремы о монотонной сходимости Б. Леви 534, стр. 480,

$$\mathbf{E}\left[\eta_0 \dot{\mathrm{I}}_A\right] = \lim_{n}{}_{\uparrow} \mathbf{E}\left[\eta_n \dot{\mathrm{I}}_A\right] = \lim_{n}{}_{\uparrow} \mathbf{E}\left[\xi_n \dot{\mathrm{I}}_A\right] = \mathbf{E}\left[\xi_0 \dot{\mathrm{I}}_A\right].$$

(c) Определим $\gamma_n = \inf\limits_{k \geqslant n} \xi_k$, тогда $\gamma_{n-1} \leqslant \gamma_n \underset{n}{\rightarrow} \varliminf\limits_{n} \xi_n$. Поэтому в силу (b)

$$\mathbf{E}\big[\varliminf_{n} \xi_n \,|\, \mathscr{L}\big] = \lim_{n} \mathbf{E}\big[\gamma_n \,|\, \mathscr{L}\big].$$

Так как, очевидно, $\gamma_n \leqslant \xi_n$, то, применяя к правой части этого равенства свойство (a), получим доказательство аналога леммы Фату для усл.м.о. ⇚

140| **Теорема.** *Пусть ξ, ζ — неотрицательные сл.в. Тогда справедливы следующие утверждения:*

Ср1) (a) $\mathbf{E}\big[b\xi + c\zeta \,|\, \mathscr{L}\big] = b\,\mathbf{E}\big[\xi \,|\, \mathscr{L}\big] + c\,\mathbf{E}\big[\zeta \,|\, \mathscr{L}\big], \qquad \forall\, b, c \in \mathbb{R}^1_+;$

(b) *если $\xi \geqslant \zeta \geqslant 0$, то на множестве $Q = \{\omega : \mathbf{E}\big[\zeta \,|\, \mathscr{L}\big](\omega) < \infty\}$*

$$\mathbf{E}\big[\xi - \zeta \,|\, \mathscr{L}\big] = \mathbf{E}\big[\xi \,|\, \mathscr{L}\big] - \mathbf{E}\big[\zeta \,|\, \mathscr{L}\big] \quad (\text{п.н.});$$

Ср2) *если $\xi \Subset \mathscr{L}$, то $\mathbf{E}\big[\xi\zeta \,|\, \mathscr{L}\big] = \xi\,\mathbf{E}\big[\zeta \,|\, \mathscr{L}\big]$, в частности усл.м.о. $\mathbf{E}\big[\xi \,|\, \mathscr{L}\big] = \xi$;*

Ср3) *если $\xi \leqslant \zeta$ (п.н.), то $\mathbf{E}\big[\xi \,|\, \mathscr{L}\big] \leqslant \mathbf{E}\big[\zeta \,|\, \mathscr{L}\big]$ (п.н.);*

Ср4) *если ξ не зависит от σ-подалгебры $\mathscr{L}$, т.е. для $\forall\, U \in \mathscr{L}$ сл.в. ξ и $\dot{\mathrm{I}}_U$ независимы, то $\mathbf{E}\big[\xi \,|\, \mathscr{L}\big] = \mathbf{E}\xi$;*

Ср5) *(телескопическое свойство) если σ-подалгебры $\mathscr{L} \subset \mathscr{M} \subset \mathscr{F}$, то*

$$\mathbf{E}\big[\,\mathbf{E}\big[\xi \,|\, \mathscr{L}\big] \,|\, \mathscr{M}\big] = \mathbf{E}\big[\,\mathbf{E}\big[\xi \,|\, \mathscr{M}\big] \,|\, \mathscr{L}\big] = \mathbf{E}\big[\xi \,|\, \mathscr{L}\big].$$

⇛ Ср1) (a) В силу линейности интеграла Лебега при $\forall\, A \in \mathscr{L}$

$$\begin{aligned}\mathbf{E}\big(\big(b\,\mathbf{E}\big[\xi \,|\, \mathscr{L}\big] + c\,\mathbf{E}\big[\zeta \,|\, \mathscr{L}\big]\big)\dot{\mathrm{I}}_A\big) &= \\ &= b\,\mathbf{E}\big(\mathbf{E}\big[\xi \,|\, \mathscr{L}\big]\dot{\mathrm{I}}_A\big) + c\,\mathbf{E}\big(\mathbf{E}\big[\zeta \,|\, \mathscr{L}\big]\dot{\mathrm{I}}_A\big) = \\ &= b\,\mathbf{E}\big(\xi\dot{\mathrm{I}}_A\big) + c\,\mathbf{E}\big(\zeta\dot{\mathrm{I}}_A\big) = \mathbf{E}\big((b\xi + c\zeta)\dot{\mathrm{I}}_A\big).\end{aligned}$$

(b) По свойству (a)

$$\mathbf{E}\big[\xi\dot{\mathrm{I}}_Q \,|\, \mathscr{L}\big] = \mathbf{E}\big[(\xi-\zeta)\dot{\mathrm{I}}_Q + \zeta\dot{\mathrm{I}}_Q \,|\, \mathscr{L}\big] = \mathbf{E}\big[(\xi-\zeta)\dot{\mathrm{I}}_Q \,|\, \mathscr{L}\big] + \mathbf{E}\big[\zeta\dot{\mathrm{I}}_Q \,|\, \mathscr{L}\big].$$

Поскольку по свойству Ср2) (в доказательстве которого используется только свойство (a)) $\mathbf{E}\big[\zeta\dot{\mathrm{I}}_Q \,|\, \mathscr{L}\big] = \mathbf{E}\big[\zeta \,|\, \mathscr{L}\big]\dot{\mathrm{I}}_Q < \infty$, то из этого равенства (опять же в силу Ср2) имеем

$$\mathbf{E}\big[(\xi-\zeta)\,|\,\mathscr{L}\,\big]\,\dot{\mathrm{I}}_Q \;=\; \mathbf{E}\big[\xi\,|\,\mathscr{L}\,\big]\,\dot{\mathrm{I}}_Q-\mathbf{E}\big[\zeta\,|\,\mathscr{L}\,\big]\,\dot{\mathrm{I}}_Q.$$

Ср2) Пусть сначала $\xi=\dot{\mathrm{I}}_B$ с некоторым множеством $B\in\mathscr{L}$. По определению усл.м.о. для $\forall\,A\in\mathscr{L}$

$$\mathbf{E}\Big[\big(\xi\;\mathbf{E}\big[\zeta\,|\,\mathscr{L}\,\big]\big)\,\dot{\mathrm{I}}_A\Big] \;=\; \mathbf{E}\Big[\mathbf{E}\big[\zeta\,|\,\mathscr{L}\,\big]\;\dot{\mathrm{I}}_{AB}\Big] \;=\; \mathbf{E}\Big[\zeta\;\dot{\mathrm{I}}_{AB}\Big] \;=\; \mathbf{E}\Big[\big(\xi\zeta\big)\;\dot{\mathrm{I}}_A\Big],$$

т.е. Ср2) справедливо для индикаторных функций. Из Ср1,(a) следует, что это свойство будет справедливо и для любых конечных линейных комбинаций индикаторных функций с неотрицательными коэффициентами, т.е. для неотрицательных простых функций. Любую неотрицательную сл.в. ξ можно приблизить монотонной последовательностью неотрицательных простых функций: $\eta_n\nearrow\xi$. Таким образом, $\zeta\,\eta_n\leqslant\zeta\,\eta_{n+1}$ и в силу леммы 139, b) для любых вариантов соответствующих усл.м.о. почти наверное

$$\mathbf{E}\big[\xi\zeta\,|\,\mathscr{L}\,\big] \;=\; \lim_n\mathbf{E}\big[\eta_n\zeta\,|\,\mathscr{L}\,\big] \;=\; \lim_n\eta_n\,\mathbf{E}\big[\zeta\,|\,\mathscr{L}\,\big] \;=\; \xi\,\mathbf{E}\big[\zeta\,|\,\mathscr{L}\,\big].$$

Ср3) Повторяет 139, (a).

Ср4) Пусть $\xi=\dot{\mathrm{I}}_B$, $B\in\mathscr{F}$. Независимость ξ от $\mathscr{L}$ означает, что математическое ожидание $\mathbf{E}\big[\,\dot{\mathrm{I}}_A\,\dot{\mathrm{I}}_B\,\big]=\mathbf{E}\big[\,\dot{\mathrm{I}}_A\,\big]\,\mathbf{E}\big[\,\dot{\mathrm{I}}_B\,\big]$ для $\forall\,A\in\mathscr{L}$. Поэтому

$$\mathbf{E}\big[\,\dot{\mathrm{I}}_A\,\xi\big] \;=\; \mathbf{E}\big[\,\dot{\mathrm{I}}_A\,\big]\,\mathbf{E}\xi \;=\; \mathbf{E}\big[\,\dot{\mathrm{I}}_A\,\mathbf{E}\xi\big],$$

т.е. свойство Ср4) справедливо для индикаторных функций и, следовательно, для неотрицательных простых функций. Пусть теперь монотонная последовательность неотрицательных простых функций $\eta_n\nearrow\xi$. Тогда по лемме 139

$$\mathbf{E}\big[\xi\,|\,\mathscr{L}\,\big] \;=\; \lim_n\mathbf{E}\big[\eta_n\,|\,\mathscr{L}\,\big] \;=\; \lim_n\mathbf{E}\,\eta_n \;=\; \mathbf{E}\,\xi$$

в силу теоремы о монотонной сходимости Б. Леви 534, стр. 480.

Ср5) Ввиду Ср2) необходимо доказать только второе равенство. Пусть $A\in\mathscr{L}$, тогда $A\in\mathscr{M}$, поэтому в соответствии с определением усл.м.о.

$$\mathbf{E}\Big[\,\dot{\mathrm{I}}_A\big(\mathbf{E}\big[\xi\,|\,\mathscr{M}\,\big]\big)\Big] \;=\; \mathbf{E}\big[\,\dot{\mathrm{I}}_A\,\xi\big] \;=\; \mathbf{E}\Big[\,\dot{\mathrm{I}}_A\,\mathbf{E}\big[\xi\,|\,\mathscr{L}\,\big]\Big].$$

Так как $\eta=\mathbf{E}\big[\xi\,|\,\mathscr{L}\,\big]\Subset\mathscr{L}$, то по определению $\eta=\mathbf{E}\big[\;\mathbf{E}\big[\xi\,|\,\mathscr{M}\,\big]\,|\,\mathscr{L}\,\big]$. ⊲

► Прежде чем вводить усл.м.о. для произвольных сл.в., необходимо решить ещё одну проблему корректности введённого определения, связанную с возможностью задания сл.в. в виде $\xi(T(\omega))$, $\omega \in \Omega$, т.е. как суперпозиции $\xi \circ T$ измеримых функций T и ξ. Точнее, пусть $\xi : (\Omega, \mathscr{F}, \mathbf{P}) \mapsto (\mathbb{R}^1_+, \mathscr{B})$ — неотрицательная сл.в., $T : \Omega' \mapsto \Omega$ — некоторое сюръективное отображение, т.е. $T(\Omega') = \Omega$. Определим σ-алгебру $\mathscr{F}' = T^{-1}(\mathscr{F})$ и вероятностную меру на ней $\mathbf{P}'\{T^{-1}(F)\} = \mathbf{P}\{F\}$, $F \in \mathscr{F}$. Таким образом, сл.в. ξ может рассматриваться как функция на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$ или как функция $\xi(T(\omega'))$ на $(\Omega', \mathscr{F}', \mathbf{P}')$. Напомним, что по теореме о замене переменных в интеграле Лебега $\int_{\Omega'} \xi(T(\omega'))\, d\mathbf{P}' = \int_\Omega \xi(\omega)\, d\mathbf{P}$. Естественно ожидать, что и усл.м.о. ξ относительно какой-либо σ-подалгебры $\mathscr{L} \subset \mathscr{F}$ будет совпадать с усл.м.о. $\xi \circ T$ относительно σ-подалгебры $\mathscr{L}' = T^{-1}(\mathscr{L}) \subset \mathscr{F}'$.

Пусть сл.в. $\zeta' = \mathbf{E}\big[\, \xi \circ T \,\big|\, \mathscr{L}' \big]$, т.е. она измерима относительно $\mathscr{L}'$ и удовлетворяет равенству

$$\int_{\Omega'} \zeta' \,\dot{\mathrm{I}}_{A'}\, d\mathbf{P}' = \int_{\Omega'} (\xi \circ T)\, \dot{\mathrm{I}}_{A'}\, d\mathbf{P}'$$

для любых $A' \in \mathscr{L}'$. Так как σ-алгебра $\mathscr{L}'$ индуцирована функцией T, то подмножество $A' \in \mathscr{L}'$ т. т. т. когда $A' = T^{-1}(A)$ с некоторым $A \in \mathscr{L}$. Кроме того, по теореме 516, стр. 469, $\zeta'(\omega') = \zeta(T(\omega'))$ с некоторой $\mathscr{L}$-измеримой функцией $\zeta : \Omega \mapsto \mathbb{R}^1$. Таким образом, по теореме о замене переменных в интеграле Лебега получаем $\int_\Omega \zeta \,\dot{\mathrm{I}}_A\, d\mathbf{P} = \int_\Omega \xi\, \dot{\mathrm{I}}_A\, d\mathbf{P}$ для любых $A \in \mathscr{L}$. Это означает (по определению усл.м.о.), что сл.в. ζ есть усл.м.о. $\mathbf{E}\big[\,\xi \,\big|\, \mathscr{L}\,\big]$.

По тем же соображениям справедливо и обратное утверждение: если $\zeta = \mathbf{E}\big[\,\xi \,\big|\, \mathscr{L}\,\big]$, то $\mathscr{L}'$-измеримая сл.в. $\zeta' = \zeta \circ T$ есть усл.м.о. $\mathbf{E}\big[\,\xi \circ T \,\big|\, \mathscr{L}'\big]$.

► Определим теперь усл.м.о. от произвольной сл.в.

Определение. Пусть $\xi : (\Omega, \mathscr{F}, \mathbf{P}) \mapsto (\mathbb{R}^1, \mathscr{B})$ — сл.в., $\mathscr{L}$ — σ-подалгебра $\mathscr{F}$. Если для каких-либо (следовательно, для всех) вариантов усл.м.о. $\mathbf{E}\big[\,\xi^+ \,\big|\, \mathscr{L}\,\big]$ и $\mathbf{E}\big[\,\xi^- \,\big|\, \mathscr{L}\,\big]$ множество $B = \big\{\omega : \min\big(\mathbf{E}\big[\,\xi^+ \,\big|\, \mathscr{L}\,\big],\ \mathbf{E}\big[\,\xi^- \,\big|\, \mathscr{L}\,\big]\big) < +\infty\big\}$ имеет вероятность $\mathbf{P}\{B\} = 1$, то *усл.м.о.* сл.в. ξ относительно $\mathscr{L}$ называется любая сл.в. η такая, что для $\omega \in B$

$$\eta = \mathbf{E}\big[\,\xi^+ \,\big|\, \mathscr{L}\,\big] - \mathbf{E}\big[\,\xi^- \,\big|\, \mathscr{L}\,\big] \quad (\text{п.н.});$$

класс таких сл.в. η обозначается $\mathbf{E}\big[\,\xi \,\big|\, \mathscr{L}\,\big]$. Если $\mathbf{P}\{B\} < 1$, усл.м.о. не определено.

141| Примеры. 1) Если существует $\mathbf{E}\xi$, то, скажем, $\mathbf{E}\xi^+ < \infty$, и по теореме 136, II существует вариант усл.м.о. $\mathbf{E}\big[\,\xi^+ \,\big|\, \mathscr{L}\,\big] < \infty$ при любых

$\omega \in \Omega$. Таким образом, в этом случае усл.м.о. определено. В частности, если $\mathbf{E}|\xi| < \infty$, то найдётся вариант усл.м.о. $\bigl|\mathbf{E}\bigl[\xi \,\big|\, \mathscr{L}\bigr](\omega)\bigr| < \infty$ для $\forall\, \omega \in \Omega$.

В примере 155, стр. 156, приводится ситуация, когда усл.м.о. определено и всюду конечно, однако математическое ожидание $\mathbf{E}\xi$ не существует.

2) Если сл.в. $\xi \Subset \mathscr{L}$, то $\xi^{\pm} \Subset \mathscr{L}$, и в силу свойства Ср2) усл.м.о. $\mathbf{E}\bigl[\xi \,\big|\, \mathscr{L}\bigr] \stackrel{\mathrm{df}}{=} \mathbf{E}\bigl[\xi^+ \,\big|\, \mathscr{L}\bigr] - \mathbf{E}\bigl[\xi^- \,\big|\, \mathscr{L}\bigr] = \xi^+ - \xi^- = \xi$. ◎

Существование математического ожидания $\mathbf{E}\xi$ является достаточным условием для применения соотношения (1) при проверке свойств усл.м.о.

142| **Лемма.** *Если для $\mathscr{L}$-измеримых сл.в. η_1, η_2, принимающих значения в расширенной прямой $\overline{\mathbb{R}}$, существуют математические ожидания $\mathbf{E}\eta_1, \mathbf{E}\eta_2$, то $\eta_1 \leqslant \eta_2$ (или $\eta_1 = \eta_2$)* (п.н.) *т. т. т. когда для $\forall\, A \in \mathscr{L}$*

$$\mathbf{E}\bigl[\eta_1 \, \dot{\mathrm{I}}_A\bigr] \underset{(=)}{\leqslant} \mathbf{E}\bigl[\eta_2 \, \dot{\mathrm{I}}_A\bigr]. \tag{2}$$

⇛ Очевидно, достаточно доказать только утверждение относительно неравенства «$\leqslant$». В одну сторону («тогда») это утверждение следует из свойств интеграла Лебега. Пусть, обратно, выполнено (2) для $\forall\, A \in \mathscr{L}$. Рассмотрим множество $A_N = \{\omega :\ |\eta_2| < N, \eta_1 > \eta_2\} \in \mathscr{L}$. Если $\mathbf{P}\{A_N\} > 0$, то ввиду конечности $\mathbf{E}\bigl[\eta_2 \, \dot{\mathrm{I}}_{A_N}\bigr]$ имеем строгое неравенство

$$\mathbf{E}\bigl[\eta_1 \, \dot{\mathrm{I}}_{A_N}\bigr] > \mathbf{E}\bigl[\eta_2 \, \dot{\mathrm{I}}_{A_N}\bigr].$$

Полученное противоречие показывает (после $N \to \infty$), что событие $\{\omega :\ |\eta_2| < +\infty, \eta_1 > \eta_2\}$ имеет нулевую вероятность. Аналогично для событий $\{\omega :\ |\eta_1| < +\infty, \eta_1 > \eta_2\}$ и $\{\omega :\ \eta_1 = +\infty, \eta_2 = -\infty\}$. ⇚

143| **Лемма.** *Если существует $\mathbf{E}\xi$, то $\mathbf{E}\bigl[\xi \,\big|\, \mathscr{L}\bigr]$ определено и $\mathscr{L}$-измеримая сл.в. η есть вариант усл.м.о. $\mathbf{E}\bigl[\xi \,\big|\, \mathscr{L}\bigr]$ т. т. т. когда для $\forall\, A \in \mathscr{L}$*

$$\mathbf{E}\bigl[\eta \, \dot{\mathrm{I}}_A\bigr] = \mathbf{E}\bigl[\xi \, \dot{\mathrm{I}}_A\bigr].$$

⇛ Пусть, к примеру, $\mathbf{E}\xi^- < \infty$, тогда $\mathbf{E}\bigl[\mathbf{E}\bigl[\xi^- \,\big|\, \mathscr{L}\bigr] \dot{\mathrm{I}}_A\bigr] = \mathbf{E}\bigl[\xi^- \, \dot{\mathrm{I}}_A\bigr] < < \infty$ для $\forall\, A \in \mathscr{L}$. Поэтому в силу свойств интеграла Лебега

$$\mathbf{E}\bigl[\xi \, \dot{\mathrm{I}}_A\bigr] = \mathbf{E}\bigl[\xi^+ \, \dot{\mathrm{I}}_A\bigr] - \mathbf{E}\bigl[\xi^- \, \dot{\mathrm{I}}_A\bigr] =$$

$$= \mathbf{E}\bigl[\mathbf{E}\bigl[\xi^+ \,\big|\, \mathscr{L}\bigr] \dot{\mathrm{I}}_A\bigr] - \mathbf{E}\bigl[\mathbf{E}\bigl[\xi^- \,\big|\, \mathscr{L}\bigr] \dot{\mathrm{I}}_A\bigr] =$$

$$= \mathbf{E}\big[\big(\mathbf{E}[\xi^+ \,|\, \mathscr{L}\,] - \mathbf{E}[\xi^- \,|\, \mathscr{L}\,]\big)\,\dot{\mathrm{I}}_A\big] \;=\; \mathbf{E}\big[\,\mathbf{E}[\xi \,|\, \mathscr{L}\,]\,\dot{\mathrm{I}}_A\big] = \mathbf{E}\big[\eta\,\dot{\mathrm{I}}_A\big],$$

что с учётом предыдущей леммы доказывает требуемое. ⫷

§ 2. Свойства условного ожидания

Напомним, что любое соотношение между двумя и более усл.м.о. означает, как правило, что это соотношение выполняется почти наверное (п.н.) для любых вариантов выбора соответствующих усл.м.о.

144| **Теорема.** *Пусть* $\xi, \zeta \,:\, (\Omega, \mathscr{F}, \mathbf{P}) \mapsto \big(\overline{\mathbb{R}}^1, \mathscr{B}(\overline{\mathbb{R}}^1)\big)$ *— сл.в.;* $\mathscr{L}, \mathscr{M}$ *— некоторые* σ*-подалгебры* $\mathscr{F}$.

C1) *Если* $\xi = \zeta$ (п.н.) *, то усл.м.о.* $\mathbf{E}[\xi \,|\, \mathscr{L}\,]$, $\mathbf{E}[\zeta \,|\, \mathscr{L}\,]$ *определены (и в этом случае они совпадают) или не определены одновременно.*

C2) $\mathbf{E}[\xi + \zeta \,|\, \mathscr{L}\,] = \mathbf{E}[\xi \,|\, \mathscr{L}\,] + \mathbf{E}[\zeta \,|\, \mathscr{L}\,]$,
если существуют варианты усл.м.о, для которых выражение в правой части определено для $\forall\, \omega \in \Omega$.

C3) *Если* $\xi \geqslant \zeta$ (п.н.) *, то* $\mathbf{E}[\xi \,|\, \mathscr{L}\,] \geqslant \mathbf{E}[\zeta \,|\, \mathscr{L}\,]$, *когда определены оба усл.м.о.*

C4) *(Формула полного математического ожидания.) Если существует математическое ожидание* $\mathbf{E}\xi$*, то*

$$\mathbf{E}\big(\mathbf{E}[\xi \,|\, \mathscr{L}\,]\big) \;=\; \mathbf{E}\xi.$$

В частности, для $\forall\, B \in \mathscr{F}$ *имеет место формула полной вероятности:*

$$\mathbf{E}\big(\mathbf{P}\{B \,|\, \mathscr{L}\,\}\big) \;=\; \mathbf{P}\{B\}\,.$$

C5) *Пусть* ξ *— конечная* $\mathscr{L}$*-измеримая сл.в. и определено усл.м.о.* $\mathbf{E}[\zeta \,|\, \mathscr{L}\,]$, *тогда* $\mathbf{E}[\xi\zeta \,|\, \mathscr{L}\,] = \xi\,\mathbf{E}[\zeta \,|\, \mathscr{L}\,]$.

C6) *(Телескопическое свойство.) Если* $\mathscr{L} \subset \mathscr{M}$ *и существует* $\mathbf{E}\xi$*, то*

$$\mathbf{E}\big[\,\mathbf{E}[\xi \,|\, \mathscr{L}\,] \,|\, \mathscr{M}\,\big] \;=\; \mathbf{E}[\xi \,|\, \mathscr{L}\,] \;=\; \mathbf{E}\big[\,\mathbf{E}[\xi \,|\, \mathscr{M}\,] \,|\, \mathscr{L}\,\big].$$

C7) *(Неравенство «треугольника».)* $\big|\mathbf{E}[\xi \,|\, \mathscr{L}\,]\big| \leqslant \mathbf{E}\big[|\xi| \,|\, \mathscr{L}\,\big]$.

C8) *Если существует* $\mathbf{E}\xi$ *и сл.в.* $\xi, \dot{\mathrm{I}}_U$ *независимы для* $\forall\, U \in \mathscr{L}$ *, т.е. сл.в.* ξ *не зависит от* σ*-алгебры* $\mathscr{L}$ *, то* $\mathbf{E}[\xi \,|\, \mathscr{L}\,] = \mathbf{E}\xi$.

C9) *Если* $|\xi_n| \leqslant \zeta$ (п.н.)*,* $\mathbf{E}\zeta < \infty$*, и* $\xi_0 = \lim_n \xi_n$*, то*

$$\lim_n \mathbf{E}[\xi_n \,|\, \mathscr{L}\,] = \mathbf{E}[\xi_0 \,|\, \mathscr{L}\,].$$

т.е. для любой последовательности вариантов усл.м.о. $\eta_n = \mathbf{E}\big[\,\xi_n \,|\, \mathscr{L}\,\big]$, $n \geqslant 1$, *существует* (п.н.) *предел* $\eta_0 = \lim_n \eta_n$ *и* $\eta_0 = \mathbf{E}\big[\,\xi_0 \,|\, \mathscr{L}\,\big]$.

$\Rrightarrow$ C1) Достаточно заметить, что по теореме 136, I класс сл.в. $\mathbf{E}\big[\,\xi^+ \,|\, \mathscr{L}\,\big] = \mathbf{E}\big[\,\zeta^+ \,|\, \mathscr{L}\,\big]$ (аналогично для сл.в. ξ^-, ζ^-).

C2) Воспользуемся очевидным соотношением $(\xi+\zeta)^- + \xi^+ + \zeta^+ = \xi^- + \zeta^- + (\xi+\zeta)^+$. Так как обе части этого равенства содержат неотрицательные сл.в., то в силу свойства линейности Cp1) для неотрицательных сл.в.

$$\mathbf{E}\big[\,(\xi+\zeta)^- \,|\, \mathscr{L}\,\big] + \mathbf{E}\big[\,\xi^+ \,|\, \mathscr{L}\,\big] + \mathbf{E}\big[\,\zeta^+ \,|\, \mathscr{L}\,\big] \;=$$
$$=\; \mathbf{E}\big[\,(\xi+\zeta)^+ \,|\, \mathscr{L}\,\big] + \mathbf{E}\big[\,\xi^- \,|\, \mathscr{L}\,\big] + \mathbf{E}\big[\,\zeta^- \,|\, \mathscr{L}\,\big],$$

где варианты всех выписанных здесь усл.м.о. выбраны так, чтобы правая часть Cp2) была определена при всех ω и, кроме того, указанное равенство имело место также при всех ω. Заметим, что в любом равенстве перенос различных частей слева направо и наоборот возможен только для конечных слагаемых.

Правая часть C2) будет определена, если, например, $\mathbf{E}\big[\,\xi^+ \,|\, \mathscr{L}\,\big] < \infty$, $\mathbf{E}\big[\,\zeta^+ \,|\, \mathscr{L}\,\big] < \infty$. Так как $(\xi+\zeta)^+ \leqslant \xi^+ + \zeta^+$, то в этом случае $\mathbf{E}\big[\,(\xi+\zeta)^+ \,|\, \mathscr{L}\,\big] < \infty$ по свойству монотонности усл.м.о. от неотрицательных сл.в. В силу конечности этих усл.м.о. из предыдущего равенства имеем

$$-\mathbf{E}\big[\,(\xi+\zeta) \,|\, \mathscr{L}\,\big] \;=\; \mathbf{E}\big[\,(\xi+\zeta)^- \,|\, \mathscr{L}\,\big] - \mathbf{E}\big[\,(\xi+\zeta)^+ \,|\, \mathscr{L}\,\big] \;=$$
$$=\; \mathbf{E}\big[\,\xi^- \,|\, \mathscr{L}\,\big] + \mathbf{E}\big[\,\zeta^- \,|\, \mathscr{L}\,\big] - \mathbf{E}\big[\,\xi^+ \,|\, \mathscr{L}\,\big] - \mathbf{E}\big[\,\zeta^+ \,|\, \mathscr{L}\,\big] \;=$$
$$=\; -(\mathbf{E}\big[\,\xi \,|\, \mathscr{L}\,\big] + \mathbf{E}\big[\,\zeta \,|\, \mathscr{L}\,\big]),$$

что и доказывает C2). Если же $\mathbf{E}\big[\,\xi^+ \,|\, \mathscr{L}\,\big] = \infty$, то для существования $\mathbf{E}\big[\,\xi \,|\, \mathscr{L}\,\big]$ и определённости правой части C2) необходимо, чтобы одновременно $\mathbf{E}\big[\,\xi^- \,|\, \mathscr{L}\,\big] < \infty, \mathbf{E}\big[\,\zeta^- \,|\, \mathscr{L}\,\big] < \infty$. Отсюда, в силу неравенства $(\xi+\zeta)^- \leqslant \xi^- + \zeta^-$, по аналогичным соображениям получаем доказательство C2).

C3) По условию $\xi^+ + \zeta^- \geqslant \zeta^+ + \xi^-$ (п.н.), поэтому, в силу свойства аддитивности усл.м.о. от неотрицательных сл.в., найдутся варианты соответствующих усл.м.о. такие, что

$$\mathbf{E}\big[\,\xi^+ \,|\, \mathscr{L}\,\big] + \mathbf{E}\big[\,\zeta^- \,|\, \mathscr{L}\,\big] \;\geqslant\; \mathbf{E}\big[\,\zeta^+ \,|\, \mathscr{L}\,\big] + \mathbf{E}\big[\,\xi^- \,|\, \mathscr{L}\,\big]$$

для всех $\omega \in \Omega$. Если $\mathbf{E}\big[\,\zeta^- \,|\, \mathscr{L}\,\big] < \infty$, то (?!) и $\mathbf{E}\big[\,\xi^- \,|\, \mathscr{L}\,\big] < \infty$. В этом случае C3) легко получается преобразованием предыдущего неравенства. При $\mathbf{E}\big[\,\zeta^- \,|\, \mathscr{L}\,\big] = \infty$, т.е. при $\mathbf{E}\big[\,\zeta \,|\, \mathscr{L}\,\big] = -\infty$, неравенство C3) очевидно.

C4) Если в утверждении леммы 143 выбрать множество $A = \Omega$, получим требуемое соотношение.

C5) По свойству Cp2) для неотрицательных сл.в. имеем $\mathbf{E}\left[\xi^{\pm}\zeta^{\pm} \mid \mathscr{L}\right] = \xi^{\pm}\,\mathbf{E}\left[\zeta^{\pm} \mid \mathscr{L}\right]$, причём все варианты усл.м.о. можно выбрать так, чтобы равенства при всех сочетаниях знаков $\pm$ выполнялись для $\forall\omega \in \Omega$. Пусть ω таково, что $\mathbf{E}\left[\zeta^{\pm} \mid \mathscr{L}\right](\omega) < \infty$. Тогда конечно усл.м.о.

$$\begin{aligned}\mathbf{E}\left[(\xi\zeta)^{+} \mid \mathscr{L}\right] &= \mathbf{E}\left[\xi^{+}\zeta^{+} \mid \mathscr{L}\right] + \mathbf{E}\left[\xi^{-}\zeta^{-} \mid \mathscr{L}\right] = \\ &= \xi^{+}\,\mathbf{E}\left[\zeta^{+} \mid \mathscr{L}\right] + \xi^{-}\,\mathbf{E}\left[\zeta^{-} \mid \mathscr{L}\right],\end{aligned}$$

т.е. для такого ω определено усл.м.о. $\mathbf{E}\left[\xi\zeta \mid \mathscr{L}\right]$.

Если, например, $\mathbf{E}\left[\zeta^{+} \mid \mathscr{L}\right](\omega) = \infty$, $\mathbf{E}\left[\zeta^{-} \mid \mathscr{L}\right](\omega) < \infty$ и $\xi(\omega) < 0$, то $\mathbf{E}\left[(\xi\zeta)^{+} \mid \mathscr{L}\right](\omega) = \xi^{-}(\omega)\,\mathbf{E}\left[\zeta^{-} \mid \mathscr{L}\right](\omega) < \infty$. При $\xi(\omega) \geqslant 0$ имеем $\mathbf{E}\left[(\xi\zeta)^{-} \mid \mathscr{L}\right](\omega) = \xi^{+}(\omega)\,\mathbf{E}\left[\zeta^{-} \mid \mathscr{L}\right](\omega) < \infty$, т.е. в любом случае усл.м.о. $\mathbf{E}\left[\xi\zeta \mid \mathscr{L}\right]$ определено. Остальные ситуации разбираются аналогично.

C6) Как и для неотрицательных сл.в., достаточно доказать второе равенство (см. пример 141). Пусть $A \in \mathscr{L}$, тогда $A \in \mathscr{M}$ и, ввиду конечности $\mathbf{E}\xi^{+}\,\dot{\mathrm{I}}_A$ (или $\mathbf{E}\xi^{-}\,\dot{\mathrm{I}}_A$),

$$\begin{aligned}\mathbf{E}\left[\mathbf{E}\left[\xi \mid \mathscr{M}\right]\dot{\mathrm{I}}_A\right] &= \mathbf{E}\left[(\mathbf{E}\left[\xi^{+} \mid \mathscr{M}\right] - \mathbf{E}\left[\xi^{-} \mid \mathscr{M}\right])\,\dot{\mathrm{I}}_A\right] = \\ &= \mathbf{E}\left[\xi^{+}\,\dot{\mathrm{I}}_A\right] - \mathbf{E}\left[\xi^{-}\,\dot{\mathrm{I}}_A\right] = \\ &= \mathbf{E}\left[(\mathbf{E}\left[\xi^{+} \mid \mathscr{L}\right] - \mathbf{E}\left[\xi^{-} \mid \mathscr{L}\right])\,\dot{\mathrm{I}}_A\right] = \mathbf{E}\left[\mathbf{E}\left[\xi \mid \mathscr{L}\right]\dot{\mathrm{I}}_A\right].\end{aligned}$$

Так как $\mathbf{E}\left[\xi \mid \mathscr{L}\right] \Subset \mathscr{L}$, то по лемме 143 $\mathbf{E}\left[\xi \mid \mathscr{L}\right] = \mathbf{E}\left[\,\mathbf{E}\left[\xi \mid \mathscr{M}\right] \mid \mathscr{L}\right]$.

C7) По определению усл.м.о.

$$\begin{aligned}|\,\mathbf{E}\left[\xi \mid \mathscr{L}\right]| &= |\,\mathbf{E}\left[\xi^{+} \mid \mathscr{L}\right] - \mathbf{E}\left[\xi^{-} \mid \mathscr{L}\right]| \leqslant \\ &\leqslant \mathbf{E}\left[\xi^{+} \mid \mathscr{L}\right] + \mathbf{E}\left[\xi^{-} \mid \mathscr{L}\right] = \mathbf{E}\left[|\xi| \mid \mathscr{L}\right].\end{aligned}$$

C8) Очевидно, $\xi^{+} = \xi\,\dot{\mathrm{I}}(\xi > 0)$ и $\xi^{-} = -\xi\,\dot{\mathrm{I}}(\xi < 0)$ не зависят от $\mathscr{L}$, т.е. $\mathbf{E}\left[\xi^{\pm} \mid \mathscr{L}\right] = \mathbf{E}\left[\xi^{\pm}\right]$ по свойству Cp4). Так как по условию или $\mathbf{E}\left[\xi^{+}\right] < \infty$, или $\mathbf{E}\left[\xi^{-}\right] < \infty$, то усл.м.о. $\mathbf{E}\left[\xi \mid \mathscr{L}\right] = \mathbf{E}\left[\xi^{+}\right] - \mathbf{E}\left[\xi^{-}\right]$ определено.

C9) В силу теоремы Лебега для $\forall A \in \mathscr{L}$ математическое ожидание $\mathbf{E}\left[|\xi_0|\,\dot{\mathrm{I}}_A\right] < \infty$, поэтому ввиду леммы 143

$$\mathbf{E}\left[\xi_0\,\dot{\mathrm{I}}_A\right] = \lim_n \mathbf{E}\left[\xi_n\,\dot{\mathrm{I}}_A\right] = \lim_n \mathbf{E}\left[\mathbf{E}\left[\xi_n \mid \mathscr{L}\right]\dot{\mathrm{I}}_A\right] = \lim_n \mathbf{E}\left[\eta_n\,\dot{\mathrm{I}}_A\right].$$

Выберем $A = \bigcap_j \{\omega : \ \eta_{n_j} > \varepsilon + \mathbf{E}\big[\,\xi_0 \,|\, \mathscr{L}\,\big]\}$ $(\in \mathscr{L})$ для произвольной подпоследовательности $n_j \to \infty$ и $\varepsilon > 0$. Тогда

$$\mathbf{E}\left[\xi_0\, \dot{\mathrm{I}}_A\right] \;=\; \lim_j \mathbf{E}\left[\eta_{n_j}\, \dot{\mathrm{I}}_A\right] \;\geqslant\; \lim_j \mathbf{E}\left[\,\mathbf{E}\big[\,\xi_0 \,|\, \mathscr{L}\,\big]\, \dot{\mathrm{I}}_A\right] + \varepsilon\, \mathbf{P}\{A\}\,,$$

что возможно, только если $\mathbf{P}\{A\} = 0$ (все математические ожидания конечны). Другими словами, $\overline{\lim\limits_n}\, \eta_n \leqslant \mathbf{E}\big[\,\xi_0 \,|\, \mathscr{L}\,\big]$ (п.н.) . Аналогично, $\underline{\lim\limits_n}\, \eta_n \geqslant$ $\geqslant \mathbf{E}\big[\,\xi_0 \,|\, \mathscr{L}\,\big]$ (п.н.) , поэтому $\lim_n \eta_n = \mathbf{E}\big[\,\xi_0 \,|\, \mathscr{L}\,\big]$ (п.н.) . ⫤

§ 3. Условное ожидание относительно случайной величины

Особенно интересен случай, когда $\mathscr{L}$ совпадает с минимальной σ-алгеброй, относительно которой измерим заданный сл.вектор $\vec{\zeta}$: $\mathscr{L} = \sigma\big(\vec{\zeta}\big) = \vec{\zeta}^{\,-1}(\mathscr{B}^k)$. Как следует из теоремы 516, стр. 469, любая $\sigma\big(\vec{\zeta}\big)$-измеримая сл.в. принимает постоянные значения на тех ω, на которых постоянен (фиксирован) сл.вектор $\vec{\zeta}$. Этим объясняется введение следующего определения.

Определение. Если σ-алгебра $\mathscr{L} = \sigma\big(\vec{\zeta}\big)$, т.е. порождена сл.вектором $\vec{\zeta} : (\Omega, \mathscr{F}, \mathbf{P}) \mapsto (\mathbb{R}^k, \mathscr{B}^k)$, то усл.м.о. $\mathbf{E}\big[\,\xi \,|\, \mathscr{L}\,\big] =: \mathbf{E}\big[\,\xi \,|\, \vec{\zeta}\,\big]$ называется *усл.м.о. ξ при фиксированном значении $\vec{\zeta}$.*

⚠ Чтобы подчеркнуть зависимость усл.м.о. от значений $\vec{\zeta}$, будем писать $\mathbf{E}\big[\,\xi \,|\, \vec{\zeta} = \vec{y}\,\big]$ вместо $\mathbf{E}\big[\,\xi \,|\, \vec{\zeta}\,\big](\omega)$, если $\vec{\zeta}(\omega) = \vec{y}$.

145 Примеры. 1) Любая детерминированная константа $\zeta(\omega) \equiv c$ порождает σ-алгебру $\mathscr{L} = \big\langle\, \varnothing, \Omega\,\big\rangle$. Поэтому для сл.в. ξ, у которой существует $\mathbf{E}\xi$, усл.м.о. $\mathbf{E}\big[\,\xi \,|\, c\,\big] = \mathbf{E}\xi$.

2) Близкую σ-алгебру порождает сл.в., принимающая с вероятностью 1 какое-либо одно значение, скажем x, и с вероятностью 0 остальные значения. Для такой сл.в. ζ подмножества σ-алгебры $\sigma\big(\zeta\big)$ можно разделить на два типа: либо $A \supset X = \big\{\omega : \zeta(\omega) = x\big\}$, и в этом случае $\mathbf{P}\{A\} = 1$, либо $A \cap X = \varnothing$, и тогда $\mathbf{P}\{A\} = 0$. Поэтому усл.м.о. $\mathbf{E}\big[\,\xi \,|\, \zeta\,\big](\omega) = \mathbf{E}\xi$ для ω таких, что $\zeta(\omega) = x$. Для остальных ω усл.м.о. можно определить произвольным образом, лишь бы сл.в. $\mathbf{E}\big[\,\xi \,|\, \zeta\,\big] \Subset \sigma\big(\zeta\big)$, например, положить равным $\mathbf{E}\xi$. Более простой путь — записать $\mathbf{E}\big[\,\xi \,|\, \zeta\,\big] = \mathbf{E}\xi$ (п.н.) . ◎

Представляется удобным при определении усл.м.о. одной сл.в. при фиксированном значении другой сл.в. исходить только из распределений этих сл.в., не обращаясь к исходному вероятностному пространству. Для упрощения рассмотрим одномерную сл.в. ζ. В этом случае вектор (ξ,ζ) как измеримое отображение $(\Omega,\mathscr{F})$ в $(\mathbb{R}^2,\mathscr{B}^2)$ порождает на нём вероятностную меру $\mathsf{P}_{(\xi,\zeta)}$. Пусть сл.величины X,Z определены на $(\mathbb{R}^2,\mathscr{B}^2,\mathsf{P}_{(\xi,\zeta)})$ по формулам $X(x_1,x_2)=x_1, Z(x_1,x_2)=x_2$. По построению $(\xi,\zeta)\backsim(X,Z)$ и в силу тех же соображений, что были приведены на стр. 141 при обосновании единственности усл.м.о., для любого варианта усл.м.о. $\mathbf{E}\big[\,\xi\mid\zeta\,\big]$ найдётся такой вариант усл.м.о. $\mathbf{E}\big[\,X\mid Z\,\big]$ (и наоборот), что $\mathbf{E}\big[\,\xi\mid\zeta=y\,\big]=\mathbf{E}\big[\,X\mid Z=y\,\big]$ для $\forall\,y\in\mathbb{R}^1$. Заметим, что прообраз любого множества $Z^{-1}(A)=\mathbb{R}^1\times A$, т.е. σ-алгебра $\sigma(Z)=\mathbb{R}^1\times\mathscr{B}$. Поэтому функция $g(x,y)$ измерима относительно σ-алгебры $\sigma(Z)$ т. т. т. когда $g(x,y)=\eta(y)$ с некоторой борелевской функцией η. Следовательно, усл.м.о. $\eta(y)=\mathbf{E}\big[\,h(\xi)\mid\zeta=y\,\big]$ от любой неотрицательной функции $h(x)$ можно определить как такую измеримую по Борелю неотрицательную функцию $\eta:\mathbb{R}^1\mapsto\mathbb{R}^1_+$, для которой при $\forall\,A\in\mathscr{B}$

$$\int_{\mathbb{R}}\eta(y)\,\dot{\mathrm{I}}_A(y)\,dF_\zeta(y)=\iint_{\mathbb{R}^2}h(x)\,\dot{\mathrm{I}}_A(y)\,dF(x,y), \tag{3}$$

где F — совместная ф.р. вектора (ξ,ζ), F_ζ — маргинальная ф.р. ζ. Для произвольной функции $h(x)$ усл.м.о. равно разности $\mathbf{E}\big[\,h(\xi)\,\dot{\mathrm{I}}\,\big(h(\xi)>0\big)\mid\zeta=$ $=y\,\big]-\mathbf{E}\big[-h(\xi)\,\dot{\mathrm{I}}\,\big(h(\xi)<0\big)\mid\zeta=y\,\big]$, если хотя бы одно из указанных усл.м.о. конечно. Любое утверждение относительно усл.м.о. $\mathbf{E}\big[\,h(\xi)\mid\zeta=y\,\big]$ будет выполняться с точностью до множества $y\in Q\subset\mathbb{R}^1$, для которого маргинальная вероятность $\mathsf{P}_\zeta\{Q\}=\mathbf{P}\{\zeta\in Q\}=0$, т.е. P_ζ-почти наверное.

146| Примеры. 1) Пусть распределение (ξ,ζ) имеет плотность $f(x,y)$ (производную Радона–Никодима $d\,\mathsf{P}_{(\xi,\zeta)}/d\mu$) относительно произведения некоторых мер $\mu=\mu_1\times\mu_2$ на $(\mathbb{R}^2,\mathscr{B}^2)$, например мер Лебега или считающих мер; $f_\zeta(y)=\int_{\mathbb{R}}f(x,y)\mu_1(dx)$ — плотность (относительно μ_2) распределения сл.в. ζ. По теореме Фубини в правой часть соотношения (3)

$$\int\limits_{x\in\mathbb{R}^1,\,y\in A}x\,f(x,y)\,d(\mu_1\times\mu_2)=\int\limits_{y\in A}\underbrace{\left(\int\limits_{\mathbb{R}}x\,\frac{f(x,y)}{f_\zeta(y)}\,d\mu_1\right)}_{\mathbf{E}[\xi\,|\,\eta=y]}f_\zeta(y)\,d\mu_2. \tag{4}$$

Итак, усл.м.о., вычисляемое через совместную плотность распределения (совместную вероятность), даёт вариант общей конструкции усл.м.о.

2) Пусть сл.в. $\zeta(\omega)=1$ при $\forall\,\omega$, а ξ — сл.в., для которой $\mathbf{E}\,\xi=0$. Тогда, как уже отмечалось, усл.м.о. $\mathbf{E}\big[\,\xi\mid\sigma\big(\zeta\big)\,\big](\omega)=0$ при $\forall\,\omega$. С другой стороны,

т.к. распределение сл.в. ζ сосредоточено в точке $y = 1$, то требованию (3) для функции $h(x) = x$ удовлетворяет любая функция $\eta(y)$ с условием $\eta(1) = 0$. Таким образом, можно выбрать усл.м.о. $\mathbf{E}\big[\xi \mid \zeta = y\big] = \sin \pi y$ или $\mathbf{E}\big[\xi \mid \zeta = y\big] = \ln y$. Понятно, что из всех таких «странных» вариантов самое естественное — положить $\mathbf{E}\big[\xi \mid \zeta = y\big] = 0$ $(= \mathbf{E}\xi)$ при $\forall y$. ◎

Следующий пример обращает внимание на необходимость отдельного рассмотрения положительной и отрицательной частей сл.в.

147| Пример. Пусть $f(x,y) = \frac{1}{2} y\, e^{-y(|x|+1)}\, \dot{\mathrm{I}}_{\{y>0\}}(y)$ — совместная плотность (ξ,ζ) относительно меры Лебега. Легко понять, что частная плотность ξ равна $f_\xi(x) = \frac{1}{2}(|x|+1)^{-2}$. Её математическое ожидание $\mathbf{E}\xi$, очевидно, не существует. С другой стороны, для положительной части $x\,\dot{\mathrm{I}}_{\{x>0\}}$ функции x по теореме Фубини правая часть (3) принимает вид

$$\int_{\mathbb{R}^2} x\,\dot{\mathrm{I}}_{\{x>0\}}\,\dot{\mathrm{I}}_A(y) f(x,y)\,dxdy = \int_0^\infty \frac{y}{2}\,\dot{\mathrm{I}}_A(y)\Big(\int_0^\infty x\,e^{-yx}\,dx\Big)e^{-y}\,dy = $$
$$= \int_0^\infty \frac{1}{2y}\,\dot{\mathrm{I}}_A(y)e^{-y}\,dy,$$

даже если интеграл равен $+\infty$. Равенство в (3) достигается, если выбрать $\eta(y) = \frac{1}{2y}$. Аналогично, $\mathbf{E}\big[-\xi\,\dot{\mathrm{I}}_{\{\xi<0\}} \mid \zeta = y\big] = \frac{1}{2y}$. В соответствии с определением $\mathbf{E}\big[\xi \mid \zeta = y\big] = \frac{1}{2y} - \frac{1}{2y} = 0$ для почти всех y. Тот же результат получается, если обратиться к понятию условной плотности (в смысле (4) или (35), стр. 94). В данном случае плотность ζ есть плотность показательного закона: $f_\zeta(y) = e^{-y}\,\dot{\mathrm{I}}_{\{y>0\}}(y)$, а условная плотность ξ при фиксированном $\zeta = y$ (> 0) есть плотность распределения Лапласа: $f_{\xi|\zeta}(x \mid y) = \frac{1}{2} y e^{-y|x|}$. При $y < 0$ (вероятность чего равна нулю), ради соблюдения единообразия, можно положить $f_{\xi|\zeta}(x \mid y) = \frac{1}{2}|y|e^{-|y||x|}$, а при $y = 0$ $f_{\xi|\zeta}(x \mid 0) = \frac{1}{2}e^{-|x|}$. Таким образом, усл.м.о. $\mathbf{E}\big[\xi \mid \zeta = y\big] = 0$ при всех y.

Заметим, что в данном случае $\mathbf{E}|\xi| = \infty$, поэтому формула полного математического ожидания $\mathbf{E}\xi = \mathbf{E}\big[\mathbf{E}\big[\xi \mid \eta\big]\big] = 0$ здесь неприменима. ◎

Некоторые свойства усл.м.о. в теореме 144 применительно к $\mathbf{E}\big[\xi \mid \zeta\big]$ могут быть сформулированы в более удобном виде.

148| **Следствие.** *Для почти всех $y \in \mathbb{R}^1$ относительно распределению сл.в. ζ справедливы свойства:*

C5) $\mathbf{E}\big[\xi\zeta \mid \zeta = y\big] = y\,\mathbf{E}\big[\xi \mid \zeta = y\big]$, *если существует $\mathbf{E}\big[\xi \mid \zeta\big]$;*

C8) $\mathbf{E}\big[\xi \mid \zeta = y\big] = \mathbf{E}\xi$, *если сл.в. ξ, ζ независимы и $\exists\ \mathbf{E}\xi$.*

Из свойства C8) следует, что для (стохастически) независимых сл.в. условное ожидание одной сл.в. не зависит (функционально) от значений другой сл.в. Свойство C5) можно проинтерпретировать в следующем духе: если зафиксировано значение $\zeta = y$, то в процессе вычисления усл.м.о. это значение не изменяется и y, как константу, можно вынести за знак усл.м.о., лишь бы это усл.м.о. было определено.

Подобная интерпретация порождает надежду на возможность равенства

$$\mathbf{E}\big[\, h(\xi,\zeta) \,|\, \zeta = y \,\big] = \mathbf{E}\big[\, h(\xi,y) \,|\, \zeta = y \,\big]. \tag{5}$$

Вообще говоря, такое равенство не вполне согласовано, поскольку левая и правая части здесь формируются различными способами. В левой части сначала из класса почти наверное совпадающих сл.в. $\mathbf{E}\big[\, h(\xi,\zeta) \,|\, \zeta \,\big]$ выбирается представитель, для которого затем находится значение при $\zeta = y$. В правой части сначала строится семейство (по y) классов сл.в. $g(y,\zeta) = \mathbf{E}\big[\, h(\xi,y) \,|\, \zeta \,\big]$, затем в некотором классе $g(y,\zeta)$ выбирается представитель, и для него находится значение при $\zeta = y$. Ввиду возможной несчётности y нельзя утверждать, что при <u>любом способе</u> выбора этих представителей требуемое равенство будет выполняться сразу для почти всех y. Свойство C5) справедливо, так как здесь выбор вариантов усл.м.о. в правой части из семейства $g(y,\zeta) = y\,\mathbf{E}\big[\, \xi \,|\, \zeta \,\big]$ происходит, по существу, независимо от y. В общем случае можно предложить только конкретный способ выбора представителей классов (5), при котором это равенство будет справедливо почти наверное. Такой способ может быть основан на усл.ф.р. $F_\xi(x \,|\, \zeta = y)$ сл.в. ξ при фиксированном значении $\zeta = y$ (см. (7) ниже). В следующей теореме, сформулированной для одномерных сл.в. (общий случай $\vec{\xi} \in \mathbb{R}^k, \vec{\zeta} \in \mathbb{R}^m$ получается простой заменой), приводятся ещё две ситуации, при которых равенство (5) справедливо при почти всех реализациях сл.в. ζ для любых вариантов усл.м.о.

<u>149</u>| **Теорема.** *Пусть ξ, ζ — случайные величины и борелевская функция $h : \mathbb{R}^2 \mapsto \mathbb{R}^1$ такова, что $\mathbf{E}\,|h(\xi,\zeta)| < \infty$.*

A) *Если распределение (ξ,ζ) имеет плотность $f(x,y)$ относительно произведения $\mu_1 \times \mu_2$ некоторых мер на $(\mathbb{R}^2, \mathscr{B}^2)$ и $f_\zeta(y) = \int_{\mathbb{R}} f(x,y)\,\mu_1(dx)$ — плотность ζ, то для P_ζ-почти всех $y \in \mathbb{R}^1$ с $f_\zeta(y) > 0$ имеет место*

$$\mathbf{E}\big[\, h(\xi,\zeta) \,|\, \zeta = y \,\big] \;=\; \mathbf{E}\big[\, h(\xi,y) \,|\, \zeta = y \,\big] \;=\; \int_{\mathbb{R}} h(x,y) \frac{f(x,y)}{f_\zeta(y)}\, \mu_1(dx),$$

где интеграл конечен для P_ζ *-почти всех* $y \in \mathbb{R}^1$.

в) *Если сл.в.* ξ, ζ *независимы, то для* P_ζ *-почти всех* $y \in \mathbb{R}^1$

$$\mathbf{E}\big[\, h(\xi,\zeta) \mid \zeta = y \,\big] = \mathbf{E}\, h(\xi, y),$$

где математическое ожидание $\mathbf{E}\, h(\xi, y) = \int_{x\in\mathbb{R}^1} h(x,y)\ \mathrm{P}_\xi(dx)$ *вычисляется относительно частного распределения* P_ξ *сл.в.* ξ .

$\Rrightarrow$ Пункт а) следует из теоремы Фубини и представления, аналогичного (4). Пункт в) есть следствие а), поскольку для независимых ξ, ζ в качестве μ_1, μ_2 можно выбрать вероятностные меры $\mathrm{P}_\xi, \mathrm{P}_\zeta$, описывающие распределения ξ и ζ соответственно. Тогда плотности $f(x,y) \equiv 1$ и $f_\zeta(y) \equiv 1$ и $\mathbf{E}\big[\, h(\xi,y) \mid \zeta = y \,\big] = \int_{\mathbb{R}} h(x,y)\, \mu_1(dx) = \mathbf{E}\, h(\xi, y)$. $\Lleftarrow$

150| Упр. Докажите утверждение в) теоремы 149 для независимых случайных элементов, принимающих значения в метрическом пространстве.

151| Упр. Докажите утверждение о том, что если $\mathbf{E}\,\xi^2 < \infty$, то функция $h^*(y) = \mathbf{E}\big[\,\xi \mid \zeta = y\,\big]$ доставляет минимум среднеквадратического отклонения $\mathbf{E}(\xi - h(\zeta))^2$ среди всех борелевских функций h.

Подсказка. Применить формулу сокращённого умножения к $\mathbf{E}\left[(\xi - h^*) + (h^* - h)\right]^2$; воспользоваться формулой полного математического ожидания и свойством C5).

§ 4. Регулярные условные распределения

Для любого сл.вектора $\vec{\xi}$ можно определить функцию

$$F(\vec{x} \mid \mathscr{L}) \ := \ \mathbf{P}\big\{\, \vec{\xi} \leqslant \vec{x} \mid \mathscr{L} \,\big\} \ \stackrel{\mathrm{df}}{=} \ \mathbf{E}\big[\ \dot{\mathrm{I}}_{\{\vec{\xi} \leqslant \vec{x}\}} \mid \mathscr{L} \,\big], \qquad \vec{x} \in \mathbb{R}^k,$$

которую хотелось бы назвать условной функцией распределения (усл.ф.р.) $\vec{\xi}$ относительно σ-алгебры $\mathscr{L}$. Проблема заключается в том, что при каждом $\vec{x}$ усл.м.о. выбирается не единственным образом: при некоторых ω (зависящих от $\vec{x}$) усл.м.о. может принимать произвольные значения. Далее будет показано, что для сл.векторов $\vec{\xi} \in \mathbb{R}^k$ можно так подобрать варианты усл.ф.р., что при всех $\omega \in \Omega$ функция $F(\vec{x} \mid \mathscr{L})(\omega)$ будет удовлетворять свойствам, присущим ф.р. (монотонность, непрерывность справа и нормированность).

Определение. Функция $F(\vec{x} \mid \omega)$, $\vec{x} \in \mathbb{R}^k, \omega \in \Omega$, называется *регулярной усл.ф.р.* $\vec{\xi}$ относительно σ-алгебры $\mathscr{L}$, если

а) при $\forall\, \vec{x} \in \mathbb{R}^k$ сл.в. $F(\vec{x} \mid \omega)$ есть вариант усл.вер. $\mathbf{P}\big\{\, \vec{\xi} \leqslant \vec{x} \mid \mathscr{L} \,\big\}$,

b) для $\forall\,\omega$ функция $F(\vec{x}\,|\,\omega)$, $\vec{x}\in\mathbb{R}^k$, есть ф.р.

Рассмотрим вектор $\vec{\xi}=(\xi_1,\xi_2)$, и пусть $F(x_1,x_2\,|\,\omega)$ — регулярная усл.ф.р. $\vec{\xi}$ относительно σ-алгебры $\mathscr{L}=\sigma(\xi_2)$, порождённой сл.в. ξ_2. В силу a) и свойства C5) при почти всех ω эта ф.р.

$$F(x_1,x_2\,|\,\omega)=\mathbf{E}\big[\,\dot{\mathrm{I}}_{\{\xi_1\leqslant x_1\}}\,\dot{\mathrm{I}}_{\{\xi_2\leqslant x_2\}}\mid\mathscr{L}\,\big](\omega)=\dot{\mathrm{I}}_{\{\xi_2\leqslant x_2\}}(\omega)\,\mathbf{E}\big[\,\dot{\mathrm{I}}_{\{\xi_1\leqslant x_1\}}\mid\mathscr{L}\,\big](\omega).$$

Если ω таково, что $\xi_2(\omega)=y$, то по переменной x_2 эта функция определяет распределение, сосредоточенное в точке $x_2=y$. Такое распределение удобнее описывать как распределение только одной сл.в. Кроме того, здесь удобнее обращаться непосредственно к распределению самих сл.величин, а не оперировать в терминах исходного вероятностного пространства.

Определение. Функция $F(\vec{x}\,|\,\vec{y})$, $\vec{x}\in\mathbb{R}^k,\vec{y}\in\mathbb{R}^m$, называется *регулярной усл.ф.р.* вектора $\vec{\xi}$ при фиксированном значении $\vec{\zeta}=\vec{y}$, если

á) при $\forall\,\vec{x}$ функция $F(\vec{x}\,|\,\vec{y})$, $\vec{y}\in\mathbb{R}^m$, есть вариант усл.вер.

$$\mathbf{P}\big\{\,\vec{\xi}\leqslant\vec{x}\mid\vec{\zeta}=\vec{y}\,\big\}\,;$$

b́) при $\forall\,\vec{y}$ функция $F(\vec{x}\,|\,\vec{y})$, $\vec{x}\in\mathbb{R}^k$, есть ф.р.

⚠ Условие á) означает (в одномерном случае), что для $\forall\,B\in\mathscr{B},x\in\mathbb{R}^1$

$$\mathbf{E}\left[\,\dot{\mathrm{I}}_B(\zeta)\,\dot{\mathrm{I}}_{\{\xi\leqslant x\}}(\xi)\right]=\mathbf{E}\left[\,\dot{\mathrm{I}}_B(\zeta)F(x\,|\,\zeta)\right]=\int_{y\in B}F(x\,|\,y)\,dF_\zeta(y).$$

Выбирая $B=\{\zeta\leqslant z\}$, отсюда получаем, что совместная ф.р. вектора (ξ,ζ) может быть вычислена через усл.ф.р.:

$$F(x,z)=\int_{-\infty}^{z}F(x\,|\,y)\,dF_\zeta(y).$$

Ясно, что и обратно, если совместная ф.р. имеет подобное представление через некоторую функцию $F(x\,|\,y)$, представляющую собой ф.р. по переменной x, то $F(x\,|\,y)$ есть регулярная усл.ф.р. Из теоремы 152 и формулы полного математического ожидания C4) следует, что в этом случае математическое ожидание любой функции $h(\xi,\zeta)$ может быть вычислено как

$$\mathbf{E}\,h(\xi,\zeta)=\int\limits_{y\in\mathbb{R}^1}\Bigg(\int\limits_{x\in\mathbb{R}^1}h(x,y)\,dF(x\,|\,y)\Bigg)\,dF_\zeta(y).\tag{6}$$

Математическое ожидание относительно регулярной усл.ф.р. даёт один из вариантов усл.м.о.

152 **Теорема.** I. *Пусть $F(\vec{x}\,|\,\omega)$ — регулярная усл.ф.р. $\vec{\xi}$ относительно σ-алгебры $\mathscr{L}$. Тогда*

i) *при $\forall\,\omega$ функция множеств*

$$\widetilde{\mathsf{P}}\{B\,|\,\omega\} \;=\; \int_B dF(\vec{x}\,|\,\omega),\ B\in\mathscr{B}^k,$$

определяет вероятностную меру на $(\mathbb{R}^k,\mathscr{B}^k)$;

ii) *при $\forall\,B\in\mathscr{B}^k$ сл.в. $\widetilde{\mathsf{P}}\{B\,|\,\omega\}$ есть вариант усл.вер. $\mathbf{P}\big\{\vec{\xi}\in B\,\big|\,\mathscr{L}\big\}$;*

iii) *для любой неотрицательной измеримой функции $h(\vec{x})$ (или измеримой функции с конечным математическим ожиданием $\mathbf{E}\big|h(\vec{\xi})\big|<\infty$)*

$$\mathbf{E}\big[\,h(\vec{\xi})\,|\,\mathscr{L}\,\big](\omega) \;=\; \int\limits_{\vec{x}\in\mathbb{R}^k} h(\vec{x})\,dF(\vec{x}\,|\,\omega)\ \text{(п.н.)}.$$

II. *Пусть $F(\vec{x}\,|\,\vec{y})$ — регулярная усл.ф.р. $\vec{\xi}$ при фиксированном значении $\vec{\zeta}=\vec{y}$. Тогда для любой неотрицательной измеримой функции $h(\vec{x},\vec{y})$ (или измеримой функции с конечным математическим ожиданием $\mathbf{E}\big|h(\vec{\xi},\vec{\zeta})\big|<\infty$)*

$$\mathbf{E}\big[\,h(\vec{\xi},\vec{\zeta})\,|\,\vec{\zeta}=\vec{y}\,\big] \;=\; \int\limits_{\vec{x}\in\mathbb{R}^k} h(\vec{x},\vec{y})\,dF(\vec{x}\,|\,\vec{y}) \qquad (7)$$

для почти всех $\vec{y}\in\mathbb{R}^m$.

⊨ Рассмотрим только одномерные сл.векторы $\vec{\xi},\vec{\zeta}$ — многомерный случай доказывается аналогично.

I. Пункт i) есть простая констатация факта, что ф.р. порождает некоторую вероятностную меру. ii) Пусть $\mathcal{W}$ — класс борелевских подмножеств B, для которых $\widetilde{\mathsf{P}}\{B\,|\,\omega\}=\mathbf{P}\big\{\xi\in B\,\big|\,\mathscr{L}\big\}$ (п.н.). В силу свойства C9) теоремы 144, стр. 144, $\mathcal{W}$ есть монотонный класс. Кроме того, этот класс содержит все полубесконечные интервалы вида $(-\infty;c], c\in\mathbb{R}^1$. По свойству C3) (свойству аддитивности усл.м.о.) отсюда следует, что класс $\mathcal{W}$ содержит алгебру $\mathcal{A}$, состоящую из всех конечных объединений непересекающихся интервалов вида $(a;c],(-\infty;c],(a;\infty),-\infty\leqslant a<c<\infty$. Напомним, что минимальная σ-алгебра $\sigma\big(\mathcal{A}\big)$, порождённая алгеброй $\mathcal{A}$, совпадает с борелевской σ-алгеброй $\mathscr{B}$ и с минимальным монотонным классом $\mathfrak{M}(\mathcal{A})$, порождённым $\mathcal{A}$. Поэтому

$$\mathscr{B}=\sigma\big(\mathcal{A}\big) \;=\; \mathfrak{M}(\mathcal{A}) \;\subset\; \mathfrak{M}(\mathcal{W}) \;=\; \mathcal{W} \;\subset\; \mathscr{B}.$$

iii) Пусть $h(z)=\dot{\mathrm{I}}_B(z)$ с некоторым измеримым множеством $B\in\mathscr{B}$. По определению интеграла Лебега–Стилтьеса правая часть iii) равна $\widetilde{\mathrm{P}}\{B \mid \omega\} = \mathbf{P}\big\{\xi \in B \,\big|\, \mathscr{L}\big\}$ (п.н.) (в силу предыдущего). Отсюда, ввиду аддитивности интеграла Лебега и усл.м.о., требуемое равенство верно для любых неотрицательных простых функций. Далее, по построению, это равенство будет верно для неотрицательных и для любых интегрируемых функций.

II. Рассмотрим функцию $(x, y \in \mathbb{R}^1, \omega \in \Omega)$

$$\widetilde{F}(x, y \mid \omega) = F(x \mid \zeta(\omega)), \quad \text{если } \zeta(\omega) \leqslant y;$$

при $\zeta(\omega) > y$ функция $\widetilde{F} = 0$. Очевидно, для каждых (x, y) эта функция измерима относительно σ-алгебры $\sigma\big(\zeta\big)$, порождённой сл.в. ζ. Легко понять (см. замечание перед формулировкой теоремы), что $\widetilde{F}$ есть регулярная усл.ф.р. вектора (ξ, ζ) относительно $\sigma\big(\zeta\big)$. Причём если $\zeta(\omega) = y_0$, то ф.р. $\widetilde{F}(x, y \mid \omega) = F(x \mid y_0)$ для $\forall\, y \geqslant y_0$. Поэтому для почти всех ω имеем

$$\mathbf{E}\big[\, h(\xi, \zeta) \mid \zeta = y_0 \big] \;=\; \int\limits_{(x,y)\in\mathbb{R}^2} h(x, y)\, d\widetilde{F}(x, y \mid \omega), \quad \text{т.е.}$$

$$\mathbf{E}\big[\, h(\xi, \zeta) \mid \zeta = y_0 \big] \;=\; \int\limits_{x\in\mathbb{R}^1} h(x, y_0)\, dF(x \mid y_0),$$

где $y_0 = \zeta(\omega)$. ⫤

<u>**153**</u>| **Теорема.** *Для любого сл.вектора $\vec{\xi}$ существует регулярная усл.ф.р. относительно любой σ-подалгебры.*

⊨ Ввиду отсутствия принципиальных отличий рассмотрим только одномерный случай. Для каждого рационального числа q выберем по одному варианту усл.вер. $F(q \mid \omega) = \mathbf{P}\big\{\xi \leqslant q \,\big|\, \mathscr{L}\big\}$ относительно заданной σ-подалгебры $\mathscr{L}$. Обозначим через $B_{rq},\ r < q$, множество тех ω, на которых нарушается свойство монотонности ф.р.: $F(r \mid \omega) > F(q \mid \omega)$. По свойству монотонности усл.м.о. такое множество имеет меру нуль для любой пары чисел $r < q$. Таким образом, множество $B = \bigcup_{r<q} B_{rq}$, где объединение берётся по всем рациональным числам $r < q$, имеет меру нуль.

По свойству C9) теоремы <u>144</u>, стр. 144, п.н. предел $\lim\limits_{n\to\infty} F(n \mid \omega) = \lim\limits_{n\to\infty} \mathbf{E}\big[\, \dot{\mathrm{I}}_{\xi\leqslant n} \mid \mathscr{L}\, \big] = 1$ и, аналогично, $\lim\limits_{n\to\infty} F(-n \mid \omega) = 0$ (п.н.). Множество N тех ω, для которых нарушаются эти равенства, имеет $\mathbf{P}$-меру нуль.

Пусть Q — множество рациональных чисел, $G(x)$ — произвольная ф.р.

Определим новую функцию

$$\widetilde{F}(x \mid \omega) \ := \inf_{q:\ x<q\in Q} F(q \mid \omega)\, \dot{\mathrm{I}}_{B^c N^c}(\omega) + G(x)\, \dot{\mathrm{I}}_{B\cup N}(\omega).$$

При каждом ω эта функция удовлетворяет всем свойствам, присущим ф.р. Действительно, если $\omega \notin B \cup N$, то для любой последовательности $\langle t_n \rangle$ такой, что $t_n > x$ и $t_n \underset{n}{\to} x$, всегда найдется такая последовательность рациональных чисел $q_n \geqslant t_n$, что $q_n \underset{n}{\to} x$. В силу монотонности $\widetilde{F}$ (доказывается аналогичными рассуждениями) и по определению $\widetilde{F}$ справедливо:

$$\widetilde{F}(x \mid \omega) \ \leqslant \ \widetilde{F}(t_n \mid \omega) \ \leqslant \ F(q_n \mid \omega) \ \underset{n}{\to} \ \widetilde{F}(x \mid \omega),$$

где последнее соотношение следует из определения инфимума.

Заметим, что так как множества B, N измеримы относительно σ-алгебры $\mathscr{L}$ и предел измеримых функций измерим, то функция $\widetilde{F}(x \mid \omega)$ $\mathscr{L}$-измерима при каждом фиксированном x. Пусть монотонно убывающая последовательность рациональных чисел $q_n \underset{n}{\to} x$, тогда $\dot{\mathrm{I}}_{\{\xi\leqslant x\}} = \lim_{n}{}_{\downarrow} \dot{\mathrm{I}}_{\{\xi\leqslant q_n\}}$ и в силу свойства C9) и равенства нулю вероятности множества $B \cup N$ имеем (п.н.)

$$\widetilde{F}(x \mid \omega) \ \overset{\mathrm{df}}{=} \ \lim_{n}{}_{\downarrow} \mathbf{E}\big[\, \dot{\mathrm{I}}_{\{\xi\leqslant q_n\}} \mid \mathscr{L} \,\big](\omega) \ = \ \mathbf{E}\big[\, \dot{\mathrm{I}}_{\{\xi\leqslant x\}} \mid \mathscr{L} \,\big]\,(\omega). \qquad \Lleftarrow$$

154| П р и м е р ы. 1) Если сл.вектор (ξ_1, ξ_2) имеет дискретное распределение на $\mathscr{X} = \big\langle\, (x_i, y_j) \,\big\rangle = \mathscr{X}_1 \times \mathscr{X}_2$, где все точки носителя $\mathscr{X}_2 = \big\langle\, y_j \,\big\rangle_1^{n_2}$ имеют ненулевую вероятность, то условное распределение ξ_1 относительно ξ_2 можно задать с помощью условной функции вероятностей (усл.ф.вер.):

$$\mathbf{P}\big\{\, \xi_1 = x_i \mid \xi_2 = y_j \,\big\} = \frac{p(x_i, y_j)}{p_2(y_j)}, \qquad x_i \in \mathscr{X}_1,$$

где $p(x_i, y_j) = \mathbf{P}\{\, \xi_1 = x_i, \xi_2 = y_j \,\}$, $p_2(y_j) = \sum_{l=1}^{n_1} p(x_l, y_j)$. Условное м.о. равно среднему относительно усл.ф.вер.: для $\forall\, y_j \in \mathscr{X}_2$

$$\mathbf{E}\big[\, h(\xi_1, \xi_2) \mid \xi_2 = y_j \,\big] \ = \ \sum_{l=1}^{n_1} h(x_l, y_j)\, \mathbf{P}\big\{\, \xi_1 = x_l \mid \xi_2 = y_j \,\big\}.$$

2) Утверждение A) теоремы 149 показывает, что если совместное распределение (ξ_1, ξ_2) абсолютно непрерывно, то функция $F_{1|2}(x \mid x_2) = \int_{-\infty}^{x} f(t_1, x_2)\, dt_1 \,/ f_2(x_2)$, $x \in \mathbb{R}^1$, есть один из вариантов усл.ф.р., а функция $f_{1|2}(x_1 \mid x_2) = f(x_1, x_2)/f_2(x_2)$ — естественный кандидат на звание условной плотности ξ_1 при $\xi_2 = x_2$. Условное математическое ожидание (при

$f_2(x_2) > 0)$ от любой измеримой функции можно вычислять как

$$\mathbf{E}\big[\, h(\xi_1, \xi_2) \,|\, \xi_2 = x_2 \,\big] \;=\; \frac{1}{f_2(x_2)} \int_{\mathbb{R}} h(x_1, x_2)\, f(x_1, x_2)\, dx_1.$$

При $f_2(x_2) = 0$ усл.м.о. можно присвоить любое значение. ◎

Следующий пример ещё раз демонстрирует необходимость введения усл.м.о. сначала для неотрицательных, а затем для произвольных сл.в.

155| Пример. Пусть ξ_1, ξ_2 — независимые одинаково распределённые сл.в. Найдём усл.м.о. $\eta = \mathbf{E}\big[\, \xi_1 \,|\, \xi_1 + \xi_2 = y \,\big]$. Есть очень простой способ догадаться, чему оно равно. Из условий задачи ясно, что $(\xi_1, \xi_2) \backsim (\xi_2, \xi_1)$. Поэтому $\mathbf{E}\big[\, \xi_1 \,|\, \xi_1 + \xi_2 = y \,\big] = \mathbf{E}\big[\, \xi_2 \,|\, \xi_1 + \xi_2 = y \,\big]$. С другой стороны, в силу линейности усл.м.о.

$$2\eta = \mathbf{E}\big[\, \xi_1 \,|\, \xi_1 + \xi_2 = y \,\big] + \mathbf{E}\big[\, \xi_2 \,|\, \xi_1 + \xi_2 = y \,\big] = \mathbf{E}\big[\, (\xi_1 + \xi_2) \,|\, \xi_1 + \xi_2 = y \,\big] = y.$$

Таким образом, $\mathbf{E}\big[\, \xi_1 \,|\, \xi_1 + \xi_2 = y \,\big] = \frac{1}{2} y$. Здесь имеется некоторая проблема: свойство линейности усл.м.о. можно применять, только когда все слагаемые существуют. Из свойств усл.м.о. следует, что для этого достаточно, например, чтобы сл.в. были неотрицательны или чтобы существовало математическое ожидание $\mathbf{E}\,\xi_1 = \mathbf{E}\,\xi_2$. Покажем, что эти условия необязательны.

Пусть $f(x)$ — ф.пл. ξ_1, ξ_2, тогда $f(x_1) f(x_2)$ — плотность распределения вектора (ξ_1, ξ_2), а $f_1(t, y) = f(t) f(y - t)$ — плотность совместного распределения вектора $(\theta, \zeta) = (\xi_1, \xi_1 + \xi_2)$; плотность $\zeta = \xi_1 + \xi_2$ равна $f_\zeta(y) = \int_{\mathbb{R}} f(t) f(y - t)\, dt$. Таким образом, условная плотность θ при фиксированном $\xi_1 + \xi_2 = y$ равна $f_{\theta|\zeta}(t \,|\, y) = f(t) f(y - t) / f_\zeta(y)$.

Эта плотность симметрична относительно точки $t = \frac{1}{2} y$, ибо после сдвига на $\frac{1}{2} y$ (подстановки $t = u + \frac{1}{2} y$) функция $f(u + \frac{1}{2} y) f(-u + \frac{1}{2} y) / f_\zeta(y)$ становится чётной функцией u. Если математическое ожидание относительно $f_{\theta|\zeta}(t \,|\, y)$ существует, то оно обязано равняться центру симметрии, т.е. $\frac{1}{2} y$.

Пусть $\xi_1 \backsim \xi_2 \backsim \mathcal{C}\mathrm{auch}(0, 1)$ — сл.в. Коши с плотностью $f(x) = 1/(\pi(1 + x^2))$. Функция $f(t) f(y - t) = \mathcal{O}(1/t^4)$ при $t \to \pm\infty$, т.е. математическое ожидание относительно условной плотности существует. Таким образом, несмотря на неопределённость $\mathbf{E}\,\xi_1$, усл.м.о. $\mathbf{E}\big[\, \xi_1 \,|\, \xi_1 + \xi_2 = y \,\big] = \frac{1}{2} y$. Можно показать, что для положительной и отрицательной частей ξ_1 усл.м.о.

$$\mathbf{E}\,\big[\, \pm\, \xi_1\, \dot{\mathrm{I}}\, (\pm \xi_1 > 0) \,|\, \xi_1 + \xi_2 = y \big] \;=\; \pm\, \frac{y}{4} + \frac{(2 + y^2) \operatorname{arctg} y}{2\pi y} + \frac{\ln(1 + y^2)}{4\pi}.$$

Кстати, условная дисперсия $\mathbf{D}\{\xi_1 \,|\, \xi_1 + \xi_2 = y\} = 1 + \frac{1}{4} y^2$. ◎

IV Характеристические функции

Метод характеристических функций является наиболее действенным методом анализа распределений случайных величин, особенно если этот анализ касается асимптотического (в каком-либо смысле) поведения распределений. Эта сторона метода будет подробно освещена в главе VI. Здесь же остановимся на основных свойствах характеристических функций одномерных и многомерных случайных величин.

§ 1. Характеристическая функция одномерного распределения

Определим комплекснозначную сл.в. $\zeta = \alpha + \mathtt{i}\vartheta$, где $\mathtt{i}$ — мнимая единица, α, ϑ — действительные сл.в. Положим по определению $\mathbf{E}\,\zeta := \mathbf{E}\,\alpha + \mathtt{i}\,\mathbf{E}\,\vartheta$, если указанные математические ожидания конечны. Напомним знаменитую формулу Эйлера: $e^{\mathtt{i}a} = \cos a + \mathtt{i}\sin a$ для действительных a.

Операция взятия математического ожидания от комплексных сл.в. обладает свойствами обычного математического ожидания. В частности, справедливо неравенство «треугольника», а также свойство математического ожидания от произведения независимых сл.в.

156 **Лемма.** (a) *Если* $\mathbf{E}\,|\zeta| < \infty$, *то* $\mathbf{E}\,\zeta$ *существует и* $\bigl|\mathbf{E}\,\zeta\bigr| \leqslant \mathbf{E}\,|\zeta|$.

(b) *Если для сл.в.* $\zeta_j = \alpha_j + \mathtt{i}\vartheta_j$, $j = 1, 2$, *векторы* (α_1, ϑ_1) *и* (α_2, ϑ_2) *независимы и* $\mathbf{E}\,|\zeta_1|$, $\mathbf{E}\,|\zeta_2| < \infty$, *то* $\mathbf{E}\,[\zeta_1\zeta_2] = \mathbf{E}\,[\zeta_1] \cdot \mathbf{E}\,[\zeta_2]$.

⇛ (a) Пусть $\zeta = \alpha + \mathrm{i}\vartheta$. Поскольку $|\zeta| = \sqrt{\alpha^2 + \vartheta^2} \geqslant |\alpha|$ и $|\zeta| \geqslant |\vartheta|$, то из условия леммы следует конечность математических ожиданий $\mathbf{E}\,\alpha$ и $\mathbf{E}\,\vartheta$.

Далее, легко проверить напрямую (?!), или воспользовавшись свойствами выпуклых функций (см. пример на стр. 533), или сославшись на неравенство Коши–Буняковского, что для $\forall\, a, a_0, b, b_0 \in \mathbb{R}^1$

$$\sqrt{a^2 + b^2} \;\geqslant\; W(a,b) := (aa_0 + bb_0)/\sqrt{a_0^2 + b_0^2},$$

где $W \equiv 0$, если $a_0 = b_0 = 0$. Заметим, что $W(a_0, b_0) = (a_0^2 + b_0^2)^{1/2}$.

Если $a_0 = \mathbf{E}\,\alpha$, $b_0 = \mathbf{E}\,\vartheta$, то среднее $\mathbf{E}\,W(\alpha, \vartheta) = W(a_0, b_0) = |\,\mathbf{E}\,\zeta|$ и

$$\mathbf{E}\,|\zeta| \;=\; \mathbf{E}\,\sqrt{\alpha^2 + \vartheta^2} \;\;\geqslant\;\; \mathbf{E}\,W(\alpha, \vartheta) \;=\; |\,\mathbf{E}\,\zeta|.$$ ⇚

157| Упр. Докажите (b).

Определение. *Характеристической функцией* (коротко хар.ф.) сл.в. ξ называется функция

$$\varphi_\xi(t) \;=\; \mathbf{E}\,e^{\mathrm{i}t\xi} = \mathbf{E}\cos(t\xi) + \mathrm{i}\,\mathbf{E}\sin(t\xi), \qquad t \in \mathbb{R}^1.$$

⚠ Значение математического ожидания полностью определяется ф.р. F (ф.пл. f) сл.в. ξ. Так как

$$\varphi_\xi(t) = \varphi_F(t) = \int_{\mathbb{R}} e^{\mathrm{i}tx}\, dF(x) \quad \Big(= \varphi_f(t) = \int_{\mathbb{R}} e^{\mathrm{i}tx}\, f(x)\, dx \Big),$$

то хар.ф. связывают именно с распределением сл.в. В особенности это относится к абсолютно непрерывным законам, поскольку для них хар.ф. — это преобразование Фурье для ф.пл. f.

Если из контекста ясно, с какой сл.в. (ф.р.) связана рассматриваемая хар.ф., то нижний индекс в её записи мы будем опускать.

► Приведём основные свойства характеристических функций.

158| **Теорема.** *Характеристическая функция $\varphi_\xi(t)$ существует для любой сл.в. ξ и удовлетворяет следующим свойствам.*

(X_1) $\big|\varphi_\xi(t)\big| \leqslant 1$.

(X_2) $\varphi_\xi(0) = 1$.

(X_3) *Хар.ф. равномерно непрерывна на всей числовой прямой.*

(X_4) *Если $\eta = a\xi + b$ с действительными a, b, то $\varphi_\eta(t) = e^{\mathrm{i}tb}\varphi_\xi(at)$.*

(X_5) $\varphi_\xi(-t) = \overline{\varphi_\xi(t)}$ *(т.е. комплексно сопряжены).*

(X$_6$) *Хар.ф.* $\varphi_\xi(t)$ *вещественна т. т. т. когда распределение сл.в.* ξ *симметрично:* $\xi \backsim -\xi$; *при этом* $\varphi_\xi(t)$ *чётна.*

(X$_7$) *Если ожидание* $\mathbf{E}\left|\xi^b\right| < \infty$ *при некотором* $b \geqslant 1$, *то для* $\forall\, k \leqslant \llbracket b \rrbracket$ *существует производная хар.ф.* $\varphi_\xi^{(k)}(t) = \mathtt{i}^k\, \mathbf{E}\left[\xi^k e^{\mathtt{i}t\xi}\right]$ *для* $\forall\, t \in \mathbb{R}^1$ *и*

$$\mathbf{E}\,\xi^k = \frac{1}{\mathtt{i}^k}\,\varphi_\xi^{(k)}(0).$$

Обратно, если существует производная в нуле чётного порядка $\varphi_\xi^{(2k)}(0)$, $k = 1, 2, \ldots$, *то моменты* $\mathbf{E}\,\xi^m$ *конечны для* $\forall\, m \leqslant 2k$.(‡)

(X$_8$) *Если для некоторого целого* $m \geqslant 1$ *существует конечный момент* $\alpha_m = \mathbf{E}\,\xi^m$, *то хар.ф.* ξ *допускает представление*

$$\varphi_\xi(t) \;=\; 1 + \mathtt{i}\alpha_1 t - \frac{\alpha_2}{2}\,t^2 + \ldots + \mathtt{i}^m \frac{\alpha_m}{m!}\,t^m + \frac{t^m}{m!}\,\varepsilon_m(t)\,,$$

где $\varepsilon_m(t) \leqslant 2\,\mathbf{E}\,|\xi|^m$ *и* $\varepsilon_m(t) \to 0$ *при* $t \to 0$. *Если, кроме того,* $\mathbf{E}\,|\xi|^{m+1} < \infty$, *то* $\varepsilon_m(t) \leqslant \frac{|t|}{m+1}\,\mathbf{E}\,|\xi|^{m+1}$.

(X$_9$) *Если* $\left\langle\, \xi_j \,\right\rangle_1^n$ — *независимые сл.в. и* $S = \sum_1^n \xi_j$, *то хар.ф.*

$$\varphi_S(t) = \prod\nolimits_1^n \varphi_{\xi_j}(t).$$

В частности, если все слагаемые одинаково распределены и, следовательно, имеют одинаковую хар.ф. $\varphi(t)$, *то* $\varphi_S(t) = \big(\varphi(t)\big)^n$.

$\Rrightarrow$ Существование хар.ф. следует из ограниченности сл.в. $\cos(t\xi)$ и $\sin(t\xi)$ при любых $t \in \mathbb{R}^1$.

(X$_1$) Так как $|e^{\mathtt{i}u}| = 1$ для $\forall\, u \in \mathbb{R}^1$, то $|\varphi_\xi(t)| = |\,\mathbf{E}\,e^{\mathtt{i}t\xi}| \leqslant \mathbf{E}\,|e^{\mathtt{i}t\xi}| = 1$.

(X$_3$) Поскольку $|e^{\mathtt{i}(t+\Delta)\xi} - e^{\mathtt{i}t\xi}| = |e^{\mathtt{i}t\xi}|\,|e^{\mathtt{i}\Delta\xi} - 1| = |e^{\mathtt{i}\Delta\xi} - 1|$, то

$$\sup_t |\varphi(t+\Delta) - \varphi(t)| \;=\; \sup_t |\,\mathbf{E}\left[e^{\mathtt{i}(t+\Delta)\xi} - e^{\mathtt{i}t\xi}\right]| \;\leqslant\; \mathbf{E}\,|e^{\mathtt{i}\Delta\xi} - 1| \;\underset{\Delta\to 0}{\longrightarrow}\; 0\,,$$

где возможность предельного перехода следует из теоремы Лебега о мажорируемой сходимости 540, стр. 482, ибо $|e^{\mathtt{i}\Delta\xi} - 1| \leqslant 2$.

(X$_6$) По определению симметричной сл.в. ф.р. ξ и $-\xi$ совпадают, поэтому совпадают их хар.ф., т.е. для $\forall\, t \in \mathbb{R}^1$ справедливы равенства

$$\varphi_\xi(t) = \varphi_{-\xi}(t) = \varphi_\xi(-t) = \overline{\varphi_\xi(t)}$$

в силу свойств (X$_4$), (X$_5$). Комплексное число равно своему сопряжённому,

(‡) Можно привести пример сл.в. (см. стр. 173), у которой хар.ф. имеет конечную производную в нуле, однако её первый момент не существует.

только если его мнимая часть равна нулю. Чётность хар.ф. следует из второго равенства последней цепочки.

Обратное утверждение доказывается аналогичными рассуждениями, если принять во внимание приводимый ниже в теореме 162 факт взаимно однозначного соответствия между хар.ф. и функциями распределения.

(X_7) Продифференцировав k раз равенство $\varphi(t) = \mathbf{E}\, e^{\mathrm{i}t\xi}$ по t, неформально занося операцию дифференцирования под знак математического ожидания и полагая затем $t = 0$, придём к требуемым соотношениям:

$$\varphi^{(k)}(t)\Big|_{t=0} \;=\; \left(\mathbf{E}\left[e^{\mathrm{i}t\xi}\right]\right)^{(k)}\Big|_{t=0} \;=\; \mathrm{i}^k\,\mathbf{E}\Big[\xi^k e^{\mathrm{i}t\xi}\Big]_{t=0} \;=\; \mathrm{i}^k\,\mathbf{E}\,\xi^k.$$

Докажем корректность такого дифференцирования. Действительно, эти равенства очевидным образом справедливы для $k = 0$. Если они справедливы для какого-либо $k \geqslant 0$ и $\mathbf{E}\left|\xi^{k+1}\right| < \infty$, то

$$\begin{aligned}\frac{\varphi^{(k)}(t+\varepsilon) - \varphi^{(k)}(t)}{\varepsilon} &= \mathrm{i}^k\;\mathbf{E}\left[\xi^k\,\frac{e^{\mathrm{i}(t+\varepsilon)\xi} - e^{\mathrm{i}t\xi}}{\varepsilon}\right] = \mathrm{i}^k\;\mathbf{E}\left[\xi^k\,\frac{e^{\mathrm{i}\frac{1}{2}\varepsilon\xi} - e^{-\mathrm{i}\frac{1}{2}\varepsilon\xi}}{\varepsilon}\,e^{\mathrm{i}(t+\varepsilon/2)\xi}\right] \\ &= \mathrm{i}^{k+1}\;\mathbf{E}\left[\xi^{k+1}\,e^{\mathrm{i}t\xi}\;\frac{\sin(\varepsilon\xi/2)}{\varepsilon\xi/2}\,e^{\mathrm{i}\varepsilon\xi/2}\right].\end{aligned}$$

Выражение под знаком математического ожидания при $\varepsilon \to 0$ сходится к $\xi^{k+1}\,e^{\mathrm{i}t\xi}$ и ограничено:

$$\left|\xi^{k+1}\,\frac{\sin(\varepsilon\xi/2)}{(\varepsilon\xi/2)}\,e^{\mathrm{i}(t+\varepsilon/2)\xi}\right| \;\leqslant\; \left|\xi^{k+1}\right|.$$

Так как по условию $\mathbf{E}\left|\xi^{k+1}\right| < \infty$, то требуемое выражение для $(k+1)$-й производной $\varphi^{(k+1)}(t)$ следует теперь из теоремы Лебега.

Для доказательства обратного утверждения воспользуемся тем, что производную m-го порядка, если она существует, можно вычислить с помощью оператора конечных разностей m-го порядка (см. [23], § 12, стр. 244–245):

$$\varphi^{(m)}(t) \;=\; \lim_{\varepsilon\to 0}\frac{\Delta_\varepsilon^{(m)}\varphi(t)}{(2\varepsilon)^m},$$

где $\Delta_\varepsilon^{(1)}\varphi(t) = \varphi(t+\varepsilon) - \varphi(t-\varepsilon)$, $\Delta_\varepsilon^{(m)}\varphi(t) = \Delta_\varepsilon^{(1)}\left(\Delta_\varepsilon^{(m-1)}\varphi(t)\right)$. Применительно к функции $e^{\mathrm{i}tx}$ (от переменной t при фиксированном x) этот оператор даёт:

$$\begin{aligned}\Delta_\varepsilon^{(1)}e^{\mathrm{i}tx} \;&=\; e^{\mathrm{i}tx}(e^{\mathrm{i}\varepsilon x} - e^{-\mathrm{i}\varepsilon x}) \;=\; e^{\mathrm{i}tx}\,2\mathrm{i}\sin(\varepsilon x),\\ \Delta_\varepsilon^{(m)}e^{\mathrm{i}tx} \;&=\; \left(\Delta_\varepsilon^{(1)}e^{\mathrm{i}tx}\right)\left(2\mathrm{i}\sin(\varepsilon x)\right)^{m-1} \;=\; e^{\mathrm{i}tx}\left(2\mathrm{i}\sin(\varepsilon x)\right)^m.\end{aligned}$$

Таким образом, производная порядка $2k$ в точке $t = 0$

$$\varphi^{(2k)}(t)\Big|_{t=0} = \lim_{\varepsilon\to 0} \frac{\Delta_\varepsilon^{(2k)}\varphi(t)}{(2\varepsilon)^{2k}}\Big|_{t=0} = \lim_{\varepsilon\to 0} \mathbf{E}\left[\frac{\Delta_\varepsilon^{(2k)}e^{\mathrm{i}t\xi}}{(2\varepsilon)^{2k}}\Big|_{t=0}\right] =$$

$$= \lim_{\varepsilon\to 0} \mathbf{E}\left[\left(\frac{\mathrm{i}\sin(\varepsilon\xi)}{\varepsilon}\right)^{2k}\right] = (-1)^k \lim_{\varepsilon\to 0} \mathbf{E}\left[\xi^{2k}\left(\frac{\sin(\varepsilon\xi)}{\varepsilon\xi}\right)^{2k}\right].$$

Так как здесь под знаком математического ожидания стоит неотрицательная функция, то можно применить лемму Фату 537, стр. 481:

$$\mathbf{E}\,\xi^{2k} = \mathbf{E}\left[\lim_{\varepsilon\to 0} \xi^{2k}\left(\frac{\sin(\varepsilon\xi)}{\varepsilon\xi}\right)^{2k}\right] \leqslant \lim_{\varepsilon\to 0} \mathbf{E}\left[\xi^{2k}\left(\frac{\sin(\varepsilon\xi)}{\varepsilon\xi}\right)^{2k}\right] = \left|\varphi^{(2k)}(0)\right| < \infty.$$

Обращаем внимание, что для нечётных степеней лемму Фату применять уже нельзя. Существование моментов порядка $m < 2k$ следует из хорошо известного неравенства Ляпунова 549, стр. 487.

(X_8) Воспользуемся представлением (10), стр. 178:

$$e^{\mathrm{i}tx} = \sum\nolimits_{k=0}^{m} \frac{\mathrm{i}^k t^k}{k!}\, x^k + \frac{t^m}{m!}\left[x^m \delta_m(tx)\right],$$

где $|\delta_m(tx)| \leqslant \min\{2, tx/(m+1)\}$. Усреднив (по распределению ξ) обе части этого представления, получаем

$$\varphi(t) = \mathbf{E}\,e^{\mathrm{i}t\xi} = \sum\nolimits_{k=0}^{m} \frac{\mathrm{i}^k t^k}{k!}\,\alpha_k + \frac{t^m}{m!}\,\mathbf{E}\left[\xi^m \delta_m(t\xi)\right].$$

При $t \to 0$ по теореме Лебега $\varepsilon_m(t) = \mathbf{E}\left[\xi^m\delta_m(t\xi)\right] \to 0$, т.к. $\delta_m(t\xi) \overset{\text{п.н.}}{\to} 0$ и $|\xi^m\delta_m(t\xi)| \leqslant 2|\xi^m|$, причём по условию $\mathbf{E}\left[|\xi^m|\right] < \infty$.

Из неравенства $|\delta_m(tx)| \leqslant \dfrac{|tx|}{m+1}$ имеем $|\varepsilon_m| \leqslant \dfrac{|t|}{m+1}\,\mathbf{E}\,|\xi^{m+1}|$, что доказывает свойство (X_8). ⊲

159| Упр. Докажите свойства (X_2), (X_4), (X_5), (X_9).

✧ **Таблица характеристических функций основных распределений.** Применение метода хар.ф. особенно удобно, когда известен вид этих функций для исследуемых вероятностных законов. В следующей лемме приведена таблица хар.ф. основных вероятностных законов (стр. 107).

160| **Лемма.** *Таблица характеристических функций.*

	Распределение		*Хар.функция*
a)	*Константа*	$\xi \overset{\text{п.н.}}{=\!=} C$	$\varphi(t) = e^{\mathrm{i}tC}$
b)	*Биномиальное*	$\xi \backsim \mathrm{Bin}(n,p)$	$\varphi(t) = (1 - p + pe^{\mathrm{i}t})^n$

c)	*Геометрическое*	$\xi \backsim \mathcal{G}\mathtt{eo}(p)$	$\varphi(t) = \dfrac{p}{p-1+e^{-\mathtt{i}t}}$
d)	*Пуассона*	$\xi \backsim \mathcal{P}\mathtt{ois}(\lambda)$	$\varphi(t) = \exp\{\lambda(e^{\mathtt{i}t}-1)\}$
e)	*Нормальное*	$\xi \backsim \mathcal{N}(\mu, \sigma^2)$	$\varphi(t) = e^{\mathtt{i}t\mu}\exp\{-\frac{1}{2}\,\sigma^2 t^2\}$
f)	*Равномерное*	$\xi \backsim \mathcal{U}\mathtt{n}(-1, 1)$	$\varphi(t) = \dfrac{\sin t}{t}$
g)	*Показательное*	$\xi \backsim \mathcal{E}\mathtt{x}(\lambda)$	$\varphi(t) = \dfrac{1}{1-\mathtt{i}\,t\,\lambda}$
h)	*Гамма*	$\xi \backsim \mathcal{G}\mathtt{am}(p, \lambda)$	$\varphi(t) = \dfrac{1}{(1-\mathtt{i}\,t\,\lambda)^p}$
j)	*Коши*	$\xi \backsim \mathcal{C}\mathtt{auch}(0, 1)$	$\varphi(t) = e^{-\lvert t\rvert}$

⊨ c) $\mathbf{P}\{\xi = k\} = p(1-p)^{k-1}, \quad k \geqslant 1$. Так как $\left|e^{\mathtt{i}t}(1-p)\right| < 1$, то по формуле для суммы бесконечной геометрической прогрессии получаем

$$\varphi(t) \;=\; p\sum\nolimits_1^\infty e^{\mathtt{i}tk}(1-p)^{k-1} \;=\; p\cdot\frac{e^{\mathtt{i}t}}{1-e^{\mathtt{i}t}(1-p)}.$$

e) Найдём хар.ф. сл.в. $\eta \backsim \mathcal{N}(0, 1)$. По определению

$$\sqrt{2\pi}\varphi(t) \;=\; \lim_{A\to\infty}\int_{-A}^{A} e^{\mathtt{i}tx}e^{-\frac{1}{2}\,x^2}\,dx = e^{-\frac{1}{2}\,t^2}\lim_{A\to\infty}\int_{-A}^{A} e^{-\frac{1}{2}\,(x-\mathtt{i}t)^2}\,dx.$$

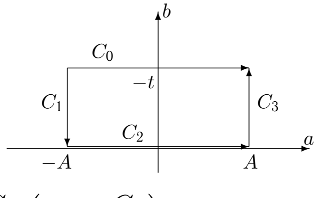
Так как $e^{-\frac{1}{2}\,z^2}$ — аналитическая функция комплексной переменной z, то последний интеграл можно рассматривать как интеграл от функции комплексной переменной по кривой $C = \{z = x + \mathtt{i}y : -A \leqslant x \leqslant A, y = 0\}$. Произведя в этом интеграле замену переменной $z = w + \mathtt{i}t$, получим

$$\int_{-A}^{A} e^{-\frac{1}{2}\,(x-\mathtt{i}t)^2}\,dx = \int_{C_0} e^{-\frac{1}{2}\,w^2}\,dw,$$

где линия $C_0 = \{w = a + \mathtt{i}b : b = -t, -A \leqslant a \leqslant A\}$. По теореме Коши интеграл от аналитической функции зависит только от крайних точек этой линии и не зависит от контура, соединяющего их.

Рассмотрим контур, состоящий из трех линий:

$$\begin{aligned} C_1 =&\{w = a + \mathtt{i}b : a = -A,\ b \in [-t; 0]\},\\ C_2 =&\{w = a + \mathtt{i}b : -A \leqslant a \leqslant A,\ b = 0\},\\ C_3 =&\{w = a + \mathtt{i}b : a = A,\ b \in [0; -t]\} \end{aligned}$$



(на рисунке $t < 0$). Криволинейный интеграл по C_1 (и по C_3)

$$\int_{C_{1(3)}} e^{-\frac{1}{2}\,w^2}\,dw = \mp e^{-\frac{1}{2}\,A^2}\int_0^{-t} e^{\pm\mathtt{i}Ab}e^{\frac{1}{2}\,b^2}\,db \;\underset{A\to\infty}{\longrightarrow}\; 0,$$

поскольку $\left|\int_0^{-t} e^{\mathtt{i}Ab} e^{-\frac{1}{2}b^2}\,db\right| \leqslant \int_0^{-t} e^{-\frac{1}{2}b^2}\,db$ и последнее выражение не зависит от A. Наконец, интеграл по C_2

$$\int_{C_2} e^{-\frac{1}{2}w^2}\,dw = \int\limits_{-A}^{A} e^{-\frac{1}{2}a^2}\,da \;\to\; \sqrt{2\pi}.$$

Можно было воспользоваться приёмом доказательства следующего утверждения. Здесь производная φ' удовлетворяет уравнению $\varphi'(t) = -t\varphi(t)$, решение которого, очевидно, совпадает с заявленной хар.ф.

g-h) Рассмотрим только гамма-распределение $\mathcal{G}\mathtt{am}(p,1)$ с плотностью $f(x) = x^{p-1}e^{-x}/\Gamma(p), \quad x>0$. Определим $h(t) = \int_0^\infty e^{\mathtt{i}tx}x^{p-1}e^{-x}\,dx,\ t\in\mathbb{R}^1$. Так как гамма-распределение имеет все моменты, то по свойству (X_7) производная

$$h'(t) \;=\; \mathtt{i}\int\limits_0^\infty e^{\mathtt{i}tx}x^pe^{-x}\,dx \;=\; \frac{\mathtt{i}}{\mathtt{i}t-1}\int\limits_0^\infty x^p\,de^{(\mathtt{i}t-1)x} \;=$$

$$= \;\frac{\mathtt{i}}{\mathtt{i}t-1}\Big(\int\limits_0^\infty x^p\,de^{-x}\cos(tx) + \mathtt{i}\int\limits_0^\infty x^p\,de^{-x}\sin(tx)\Big) \;=\; \frac{-\mathtt{i}p}{\mathtt{i}t-1}h(t)$$

после интегрирования по частям. Итак, функция h удовлетворяет дифференциальному уравнению $h'(t) = \mathtt{i}p(1-\mathtt{i}t)^{-1}h(t)$, решение которого хорошо известно: $h(t) = C(1-\mathtt{i}t)^{-p}$. Константа $C = h(0) = \Gamma(p)$.

j) Пусть $t\neq 0$; тогда

$$\varphi(t) \;=\; \frac{1}{\pi}\int_{\mathbb{R}} \frac{e^{\mathtt{i}tx}}{1+x^2}\,dx \;=\; \frac{|t|}{\pi}\int_{\mathbb{R}} \frac{e^{\mathtt{i}x}}{t^2+x^2}\,dx.$$

Интеграл такого сорта входит в обычный набор задач курса ТФКП. Наиболее удобный способ его вычисления — с помощью вычетов (res) подынтегральной функции в особых точках $(z=z_k)$ верхней полуплоскости:

$$\varphi(t) \;=\; \frac{|t|}{\pi}\,2\pi\mathtt{i}\sum_k \operatorname*{res}_{z=z_k}\frac{e^{\mathtt{i}z}}{(t^2+z^2)} \;=\; \frac{|t|}{\pi}\,2\pi\mathtt{i}\operatorname*{res}_{z=\mathtt{i}|t|}\frac{e^{\mathtt{i}z}}{(t^2+z^2)} = e^{-|t|},$$

т.к. единственная особая точка $z=\mathtt{i}|t|$ есть полюс 1-го порядка и вычет в этой точке равен

$$\lim_{z\to\mathtt{i}|t|}\frac{(z-\mathtt{i}|t|)e^{\mathtt{i}z}}{(t^2+z^2)} \;=\; \frac{e^{-|t|}}{2\mathtt{i}|t|}. \qquad \Lleftarrow$$

161| Упр. Докажите лемму для распределений b), d), f). Воспользовавшись (X_4), подтвердите общие формулы для распределений $\mathcal{N}(\mu,\sigma^2)$ и $\mathcal{G}\mathtt{am}(p,\lambda)$. Найдите хар.ф. ξ, если (i) $\xi \backsim \mathcal{U}\mathtt{n}[0;1]$, (ii) $\xi \backsim \mathcal{P}\mathtt{sc}(n,p)$.

§ 2. Теоремы единственности

Теперь, как и было обещано, установим взаимно однозначное соответствие между хар.ф. и функциями распределения.

162| **Теорема.** [*Единственности для хар.ф.*] *Сл.в. ξ, η имеют одинаковое распределение т. т. т. когда совпадают их характеристические функции:*

$$\Big[F_\xi(x) = F_\eta(x), \quad \forall\, x \in \mathbb{R}^1\Big] \quad \Leftrightarrow \quad \Big[\varphi_\xi(t) = \varphi_\eta(t), \quad \forall\, t \in \mathbb{R}^1\Big].$$

$\models\!\!\!\Rightarrow$ В одну сторону ($\Rightarrow$) утверждение очевидно. Идея следующего доказательства обратного утверждения состоит в установлении равенства $F_{\xi+s\theta} \equiv F_{\eta+s\theta}$ между ф.р. двух сумм независимых сл.в. $\xi + s\theta$ и $\eta + s\theta$, где θ — нормальная $\mathcal{N}(0,1)$ сл.в., параметр $s > 0$. По лемме 163 отсюда будет следовать, что $F_\xi(x) = F_\eta(x)$ сначала во всех точках непрерывности обеих ф.р., а затем, по теореме 27, стр. 45, и во всех остальных точках $\mathbb{R}^1$.

Пусть $\theta \backsim \mathcal{N}(0,1)$ не зависит от ξ (и от η). Тогда $s\theta \backsim \mathcal{N}(0,s^2)$, и по формуле для свёртки находим плотность

$$f_{\xi+s\theta}(z) = \frac{1}{\sqrt{2\pi}s}\int_{\mathbb{R}} e^{-(z-x)^2/2s^2}\,dF_\xi(x).$$

Замечательно, что под интегралом здесь находится хар.ф. нормального $\mathcal{N}(0,s^{-2})$ закона в точке $(x-z)$ (см. стр. 161). Заменим эту хар.ф. её выражением через интеграл от нормальной $\mathcal{N}(0,s^{-2})$ ф.пл.:

$$f_{\xi+s\theta}(z) = \frac{1}{\sqrt{2\pi}s}\int_{\mathbb{R}}\left(\int_{\mathbb{R}} e^{\mathrm{i}(x-z)y}\,\frac{s}{\sqrt{2\pi}}\,e^{-\frac{1}{2}s^2y^2}\,dy\right)dF_\xi(x).$$

Поскольку подынтегральная функция $\left|e^{\mathrm{i}(x-z)y}\right| \leqslant 1$, то она интегрируема сразу по обеим переменным x и y (по вероятностной мере). Следовательно, по теореме Фубини здесь можно поменять порядок интегрирования. Заметим, что $\int e^{\mathrm{i}(x-z)y}dF_\xi(x) = e^{-\mathrm{i}zy}\varphi_\xi(y)$, поэтому

$$f_{\xi+s\theta}(z) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-\mathrm{i}zy}\varphi_\xi(y)e^{-\frac{1}{2}s^2y^2}\,dy. \tag{1}$$

Так как по условию $\varphi_\xi = \varphi_\eta$, то $f_{\xi+s\theta} = f_{\eta+s\theta}$, а значит, и $F_{\xi+s\theta} = F_{\eta+s\theta}$. $\Leftarrow\!\!\!\!\dashv$

163| **Лемма.** *Пусть $\vec{\xi}, \vec{\theta}$ — независимые сл.векторы, s — действительный параметр. Тогда для ф.р. $F_{\vec{\xi}+s\vec{\theta}}$ суммы $\vec{\xi}+s\vec{\theta}$ справедливо*

$$\lim_{s\to 0} F_{\vec{\xi}+s\vec{\theta}}(\vec{x}) = F_{\vec{\xi}}(\vec{x})$$

в любой точке $\vec{x}$ *непрерывности ф.р.* $F_{\vec{\xi}}$ *сл.вектора* $\vec{\xi}$.

⫤ Как замечено в 116, стр. 119,

$$F_{\vec{\xi}+s\vec{\theta}}(\vec{x}) \;=\; \int_{\mathbb{R}^k} F_{\vec{\xi}}(\vec{x}-s\vec{u})\,dF_{\vec{\theta}}(\vec{u}).$$

Если $\vec{x}$ — точка непрерывности $F_{\vec{\xi}}$, то для $\forall\,\vec{u}$ при $s\to 0$ ф.р. $F_{\vec{\xi}}(\vec{x}-s\vec{u})\to$ $\to F_{\vec{\xi}}(\vec{x})$. По теореме Лебега об ограниченной сходимости можно перейти к пределу под знаком интеграла:

$$\lim_{s\to 0} F_{\vec{\xi}+s\vec{\theta}}(\vec{x}) \;=\; \int_{\mathbb{R}^k} F_{\vec{\xi}}(\vec{x})\,dF_{\vec{\theta}}(\vec{u}) \;=\; F_{\vec{\xi}}(\vec{x})\int_{\mathbb{R}^k} dF_{\vec{\theta}}(\vec{u}) \;=\; F_{\vec{\xi}}(\vec{x}). \qquad ⫣$$

164| Упр. Распределение Лапласа можно определить либо как а) распределение сл.в. $\xi_1-\xi_2$, где $\xi_1,\xi_2 \backsim \mathcal{E}\mathrm{x}(1)$ — независимые показательные сл.в., либо как б) абсолютно непрерывное распределение с плотностью $f(x)=$ $=\frac{1}{2}e^{-|x|},\ x\in\mathbb{R}^1$. Покажите, что эти определения эквивалентны.

✧ **Формулы обращения**. Другие способы доказательства теоремы единственности основаны на прямых соотношениях между хар.ф. и законами распределения сл.в. Три варианта таких соотношений приведены ниже.

165| **Теорема.** (?!) *Пусть* ξ *— целочисленная сл.в.,* $\sum_{-\infty}^{\infty}\mathbf{P}\{\xi=k\}=$ $=1$, *а* $\varphi(t)$ *— её хар.функция. Тогда для* $\forall\,m=0,\pm1,\pm2,\dots$,

$$\mathbf{P}\{\xi=m\} \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-imt}\,\varphi(t)\,dt.$$

⚠ Интеграл в правой части этого равенства (как и в подобных равенствах далее) может пониматься как интеграл Римана, поскольку подынтегральная функция здесь непрерывна в области интегрирования.

Положив $s=0$ в (1), получим формулу обратного преобразования Фурье. В следующей теореме содержится формальное обоснование такого действия. В части I теоремы приводится так называемая формула обращения, из которой также следует теорема единственности для хар.ф. Нам понадобятся свойства функции интегрального синуса (интеграла Маскерони–Дирихле, см. [23])

$$\mathrm{Si}(x) \;:=\; \int_0^x \frac{\sin t}{t}\,dt \;=\; \int_0^1 \frac{\sin(x\,t)}{t}\,dt, \qquad x\in\mathbb{R}^1.$$

(Si_1) Функция Si непрерывна и ограничена: $\left|\mathrm{Si}(x)\right|<2,\ \forall\,x\in\mathbb{R}^1$;

(Si_2) предел $\lim\limits_{x\to+\infty}\mathrm{Si}(\pm x)=\pm\frac{\pi}{2}$.

166| **Теорема.** *Пусть $F(x)$ — ф.р. сл.в. ξ, $\varphi(t)$ — её хар.ф.*

I. *(Формула обращения.) Для $\forall\, a, b \in \mathbb{R}^1$, где функция F непрерывна,*

$$F(b) - F(a) \;=\; \lim_{M\to\infty} \frac{1}{2\pi} \int_{-M}^{M} \frac{e^{-ita} - e^{-itb}}{it}\, \varphi(t)\, dt. \tag{2}$$

II. *(Обратное преобразование Фурье.) Если $\int_{\mathbb{R}} \bigl|\varphi(t)\bigr|\, dt < \infty$, то*

$$F(b) - F(a) \;=\; \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ita} - e^{-itb}}{it}\, \varphi(t)\, dt \tag{3}$$

для $\forall\, a, b \in \mathbb{R}^1$, и, кроме того, ф.р. F абсолютно непрерывна с равномерно непрерывной плотностью

$$f(x) \;=\; \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t)\, dt. \tag{4}$$

$\Rrightarrow$ I. Для $\forall\, M > 0$ определим

$$J(M) \;=\; \frac{1}{2\pi} \int_{-M}^{M} \frac{e^{-ita} - e^{-itb}}{it}\, \varphi(t)\, dt \;=\; \frac{1}{2\pi} \int_{-M}^{M} \frac{e^{-ita} - e^{-itb}}{it} \bigl(\mathbf{E}\, e^{it\xi}\bigr)\, dt.$$

Так как $\bigl|(e^{-ita} - e^{-itb}) e^{itx}/it\bigr| \leqslant 1$ при $t \in (-M; M)$, $x \in \mathbb{R}^1$, то по теореме Фубини можно поменять порядок интегрирования (по мере $dt \times dF(x)$) :

$$J(M) \;=\; \frac{1}{2\pi} \int_{\mathbb{R}} \Bigl[\int_{-M}^{M} \frac{e^{-ita} - e^{-itb}}{it}\, e^{itx}\, dt \Bigr]\, dF(x).$$

Заметим, что $\cos(ct)/t$ есть нечётная, а $\sin(ct)/t$ — чётная функции t. Поэтому, применяя формулу Эйлера и функцию Si, имеем

$$\begin{aligned} J(M) \;&=\; \frac{1}{\pi} \int_{\mathbb{R}} \Bigl[\int_{0}^{M} \frac{\sin(t(x-a)) - \sin(t(x-b))}{t}\, dt \Bigr]\, dF(x) \;= \\ &=\; \frac{1}{\pi}\, \mathbf{E} \bigl[\mathrm{Si}(M(\xi - a)) - \mathrm{Si}(M(\xi - b)) \bigr]. \end{aligned}$$

Согласно свойству (Si_2), при $M \to +\infty$ получаем

$$\frac{1}{\pi} \bigl[\mathrm{Si}(M(x-a)) - \mathrm{Si}(M(x-b)) \bigr] \;\to\; \begin{cases} 0, & x < a < b \ \text{ или } \ a < b < x, \\ \frac{1}{2}, & x = a < b \ \text{ или } \ a < b = x, \\ 1, & a < x < b. \end{cases} \tag{5}$$

По условию вероятность $\mathbf{P}\{ (\xi = a) \cup (\xi = b) \} = 0$. Поэтому, ввиду свойства

(Si$_1$) и согласно теореме Лебега об ограниченной сходимости,

$$\lim_{M\to\infty} J(M) \;=\; \mathbf{E}\left[\,\dot{\mathrm{I}}\,(a<\xi<b)\right] \;=\; F(b-0)-F(a) \;=\; F(b)-F(a).$$

Первая часть II следует из I и теоремы Лебега об ограниченной сходимости. Во второй части интеграл, определяющий функцию $f(x)$, существует в силу условий теоремы. Кроме того, эта функция неотрицательна как предел при $s\to 0$ правой части соотношения (1): $f(x)=\lim_s f_{\xi+s\theta}(x)\;\geqslant 0$, где переход к пределу под интегралом разрешён опять же по теореме Лебега.

По условию теоремы в интеграле $\int_a^b f(x)\,dx$ при любых $a,b\in\mathbb{R}^1$ можно произвести замену порядка интегрирования — в результате получим правую часть (3). Итак, $F(b)-F(a)=\int_a^b f(x)\,dx$, т.е. f есть плотность ф.р. F.

Равномерная непрерывность f доказывается аналогично (X$_3$). ⟽

⚠ Если вдобавок хар.ф. φ чётна (соответствующее распределение симметрично), то формулы (3) и (4) можно переписать в виде

$$F(x)=\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \frac{\sin(x\,t)}{t}\,\varphi(t)\,dt, \qquad f(x)=\frac{1}{\pi}\int_0^\infty \cos(x\,t)\,\varphi(t)\,dt.$$

167| Упр. Найдите с помощью (4) плотность по хар.ф.:
a) $\varphi(t)=(1-|t|)\,\dot{\mathrm{I}}\,(|t|\leqslant 1)$, b) $\varphi(t)=e^{-|t|}$ (Коши).

168| Упр. Докажите справедливость равенства (3) при более слабом условии $\int_{-\infty}^{\infty}|\varphi(t)|(1+|t|)^{-1}\,dt<\infty$.

169| Упр. Воспользовавшись (5), покажите, что в общем случае правая часть (2) равна $\frac{1}{2}\big(F(b)+F(b-)\big)-\frac{1}{2}\big(F(a)+F(a-)\big)$.

170| Упр. Проинтегрировав тождество (1) по $z\in[a;b]$, произведя замену $s^2 \dashrightarrow 2s^2$ и перейдя к пределу при $s\to 0$, получите ещё один вариант формулы обращения

$$F_\xi(b)-F_\xi(a) \;=\; \lim_{s\to 0}\frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{-\mathrm{i}ta}-e^{-\mathrm{i}tb}}{\mathrm{i}t}\varphi_\xi(t)e^{-s^2t^2}\,dt.$$

✧ **Распределение суммы случайных величин.** С помощью хар.ф. легко находится распределение суммы независимых сл.в.

171| Пример. В примере 106, стр. 115, было установлено, что сл.в. хи-квадрат $\chi_1^2 \backsim \mathcal{G}\mathrm{am}(\frac{1}{2},2)$, т.е. её хар.ф. равна $\varphi_{\chi_1^2}(t)=(1-2\mathrm{i}t)^{-1/2}$. Сл.в. хи-квадрат с k степенями свободы χ_k^2 по определению равна сумме квадратов k независимых стандартных нормальных сл.в.:

$$\chi_k^2 = \sum\nolimits_1^k \xi_j^2, \qquad \xi_j \backsim \mathcal{N}(0,1).$$

По свойству (X_9) хар.ф. χ_k^2 равна $(1-2\mathtt{i}t)^{-k/2}$, т.е. $\chi_k^2 \backsim \mathcal{G}\mathtt{am}(k/2\,,2)$.

В приложении В (стр. 176) показано, что в окрестности точки $t=0$ ($|t| < \frac{1}{2}\min\{1, \mathrm{tg}(2\pi/(k+4))\}$) логарифм этой хар.ф. допускает представление

$$\ln(\varphi_{\chi_k^2}(t)) \;=\; -\frac{k}{2}\sum_{j=1}^{\infty} \mathtt{i}^j(-1)\frac{2^j}{j}\,t^j \;=\; \sum_{j=1}^{\infty}\frac{\mathtt{i}^j}{j!}\,t^j\,\underbrace{k(j-1)!2^{j-1}}_{\gamma_j},$$

где выделенные здесь коэффициенты γ_j суть так называемые семиинварианты распределения сл.в. В том же приложении показано, что математическое ожидание сл.в. равно первому семиинварианту, а дисперсия — второму семиинварианту, т.е. $\mathbf{E}\chi_k^2 = \gamma_1 = k$, $\mathbf{D}\chi_k^2 = \gamma_2 = 2k$.

Интересно, что в случае $k=2$ имеем показательное распределение:

$$\xi_1^2+\xi_2^2 \backsim \mathcal{E}\mathrm{x}(2), \qquad \xi_1,\xi_2 \backsim \mathcal{N}(0,1) \text{ и независимы.} \qquad \circledcirc$$

172| Упр. Выведите с помощью хар.ф. утверждения теоремы 126, стр. 122, о распределениях сумм независимых сл.в.

⚠ Доказательство того, что распределение суммы независимых нормальных сл.в. нормально, почти тривиально:

$$\varphi_{\xi_1+\xi_2}(t) \;=\; \varphi_{\xi_1}(t)\varphi_{\xi_2}(t) \;=\; e^{\mathtt{i}\mu_1 t - \frac{1}{2}\sigma_1^2t^2}e^{\mathtt{i}\mu_2 t - \frac{1}{2}\sigma_2^2t^2} \;=\; e^{\mathtt{i}(\mu_1+\mu_2)t - \frac{1}{2}(\sigma_1^2+\sigma_2^2)t^2}.$$

Оказывается, справедлив и обратный результат (уже не такой тривиальный). Так называемая теорема Крамера утверждает, что если $\varphi_1(t)\varphi_2(t) = \exp\big(\mathtt{i}\mu t - \frac{1}{2}\sigma^2t^2\big)$, $t\in\mathbb{R}^1$, то φ_1,φ_2 суть хар.ф. некоторых нормальных законов.

Аналогичное утверждение имеет место и для пуассоновского закона (Д.А. Райков), и для биномиального закона.

§3. Характеристическая функция многомерного распределения

Пусть $\vec{\xi} = (\xi_1,\ldots,\xi_k)$ — случайный вектор, $\vec{t} = (t_1,\ldots,t_k)\in\mathbb{R}^k$.

Определение. Функция $\varphi_{\vec{\xi}}(t_1,\ldots,t_k) = \mathbf{E}\,e^{\mathtt{i}\vec{t}^{\,\flat}\vec{\xi}} = \mathbf{E}\,e^{\mathtt{i}(t_1\xi_1+\ldots+t_k\xi_k)}$, $\vec{t}\in\mathbb{R}^k$, называется *характеристической функцией* случайного вектора $\vec{\xi}$.

Свойства многомерных характеристических функций в основном идентичны свойствам хар.ф. одномерных сл.в.

173 **Теорема.** *Характеристическая функция $\varphi_{\vec{\xi}}(\vec{t}\,)$ существует для любого сл.вектора $\vec{\xi}=(\xi_1,\dots,\xi_k)$ и удовлетворяет следующим свойствам.*

($\mathrm{X_{k1}}$) $\left|\varphi_{\vec{\xi}}(\vec{t}\,)\right|\leqslant 1;\qquad \varphi_{\vec{\xi}}(\vec{0})=1;\qquad \varphi_{\vec{\xi}}(\vec{t}\,)$ *равномерно непрерывна на* $\mathbb{R}^k$.

($\mathrm{X_{k2}}$) *Функция* $\varphi_{\vec{\xi}}(t_1,\dots,t_m,0,\dots,0)$, $(t_1,\dots,t_m)\in\mathbb{R}^m$, $m<k$, *совпадает с хар.ф. сл.вектора* $(\xi_1,\dots,\xi_m)$.

($\mathrm{X_{k3}}$) *Пусть Q — действительная $(m\times k)$-матрица и $\vec{b}\in\mathbb{R}^m$, тогда хар.ф. $\vec{\eta}=Q\vec{\xi}+\vec{b}$ связана с хар.ф. $\vec{\xi}$ равенством*

$$\varphi_{\vec{\eta}}(\vec{u}\,)=e^{\mathrm{i}\vec{u}^{\flat}\vec{b}}\,\varphi_{\vec{\xi}}(Q^{\flat}\vec{u}\,),\qquad \vec{u}\in\mathbb{R}^m.$$

($\mathrm{X_{k4}}$) *Если для натуральных $r_1,\dots,r_k$ момент $\mathbf{E}\left|\xi_1^{r_1}\cdots\xi_k^{r_k}\right|<\infty$, то хар.ф. $\varphi_{\vec{\xi}}(\vec{t}\,)$ имеет производную в точке $\vec{t}=\vec{0}$ порядка $r=r_1+\ldots+r_k$ и*

$$\mathbf{E}\left[\xi_1^{r_1}\cdots\xi_k^{r_k}\right]\;=\;\frac{1}{\mathrm{i}^r}\;\frac{\partial^r\varphi_{\vec{\xi}}(\vec{t}\,)}{\partial t_1^{r_1}\cdots\partial t_k^{r_k}}\Bigg|_{\vec{t}=\vec{0}}.$$

($\mathrm{X_{k5}}$) *Если существуют (конечные) вектор средних $\vec{\mu}=\mathbf{E}\vec{\xi}$ и матрица ковариаций $\mathbf{\Sigma}=\mathbf{Cov}(\vec{\xi}\,)$, то хар.ф. допускает представление*

$$\varphi_{\vec{\xi}}(\vec{t}\,)\;=\;1+\mathrm{i}\vec{t}^{\,\flat}\vec{\mu}-\frac{1}{2}\vec{t}^{\,\flat}\mathbf{\Sigma}\vec{t}+o(\|\vec{t}\|^2),\qquad \vec{t}\to 0.$$

($\mathrm{X_{k6}}$) *Компоненты вектора $\vec{\xi}=(\xi_1,\dots,\xi_k)$ независимы в совокупности т. т. т. когда хар.ф. $\vec{\xi}$ равна произведению хар.ф. компонент:*

$$\varphi_{\vec{\xi}}(\vec{t})=\prod_{j=1}^{k}\varphi_{\xi_j}(t_j). \tag{6}$$

⇛ ($\mathrm{X_{k6}}$) Если компоненты $\vec{\xi}=(\xi_1,\dots,\xi_k)$ независимы, то

$$\varphi_{\vec{\xi}}(t_1,\dots,t_k)\;=\;\mathbf{E}\exp\Big\{\mathrm{i}\sum_j t_j\,\xi_j\Big\}\;=\;\prod_{j=1}^{k}\mathbf{E}\,e^{\mathrm{i}t_j\,\xi_j}\;=\;\prod_{j=1}^{k}\varphi_{\xi_j}(t_j)$$

по лемме 156, стр. 157, о среднем значении произведения независимых сл.в.

Обратно, пусть равенство (6) верно для $\forall\,\vec{t}$. Как только что было замечено, правая часть (6) есть хар.ф. сл.вектора с независимыми компонентами. Применяя теперь приводимую ниже теорему 176, заключаем, что ф.р. вектора $\vec{\xi}$ совпадает с ф.р. вектора с независимыми компонентами. ⇚

174 Упр. Докажите свойства ($\mathrm{X_{k1}}$) – ($\mathrm{X_{k5}}$).

Здесь также справедлива теорема единственности, доказательство которой можно провести аналогично одномерному случаю. Нам понадобится

175| Лемма. (?!) *Для вектора* $\vec{\theta}$, *составленного из независимых нормальных* $\mathcal{N}(0,1)$ *сл.в., хар.ф.* $a\vec{\theta}$, $a \in \mathbb{R}^1$, *равна* $\exp\left\{-a^2 \frac{1}{2} \vec{t}^{\flat} \vec{t}\right\}$, $\vec{t} \in \mathbb{R}^k$.

176| Теорема. *Хар.ф. сл.векторов* $\vec{\xi}, \vec{\eta}$ *совпадают т. т. т. когда совпадают их ф.р.:*

$$\left[F_{\vec{\xi}}(\vec{x}) = F_{\vec{\eta}}(\vec{x}),\ \forall\, \vec{x} \in \mathbb{R}^k\right] \quad \Leftrightarrow \quad \left[\varphi_{\vec{\xi}}(\vec{t}) = \varphi_{\vec{\eta}}(\vec{t}),\ \forall\, \vec{t} \in \mathbb{R}^k\right].$$

$\Rrightarrow$ Пусть $\vec{\theta}$ — k-мерный вектор независимых нормальных $\mathcal{N}(0,1)$ сл.в., независимый от $\vec{\xi}$ и $\vec{\eta}$. Очевидно, при $s > 0$ плотность вектора $s\vec{\theta}$ равна $\exp\{-\frac{1}{2s^2}\vec{z}^{\flat}\vec{z}\}/\left(\sqrt{2\pi}s\right)^k$, $\vec{z} \in \mathbb{R}^k$. Функция плотности суммы $\vec{\xi} + s\vec{\theta}$ по теореме о свёртке имеет вид

$$f_{\vec{\xi}+s\vec{\theta}}(\vec{z}) \;=\; \frac{1}{(\sqrt{2\pi}s)^k} \int\limits_{\mathbb{R}^k} \exp\left\{-\frac{1}{2s^2}(\vec{z}-\vec{y})^{\flat}(\vec{z}-\vec{y})\right\} dF_{\vec{\xi}}(\vec{y}).$$

По предыдущей лемме подынтегральная функция здесь равна хар.ф. вектора $\frac{1}{s}\vec{\theta}$, вычисленной в точке $(\vec{y}-\vec{z})$. Дальнейшие выкладки идентичны доказательству одномерного варианта теоремы (см. стр. 164). $\Lleftarrow$

Чтобы сформулировать k-мерный аналог формулы обращения, рассмотрим параллелепипед $(\vec{a};\vec{c}] = (a_1;c_1] \times \cdots \times (a_k;c_k]$. Говорят, что $(\vec{a};\vec{c}]$ является *интервалом непрерывности* F, если для $\forall\, m = \overline{1,k}$ точки a_m и c_m являются точками непрерывности маргинальной ф.р. $F_{\xi_m}(x)$, что обеспечивает непрерывность функции F во всех вершинах параллелепипеда $(\vec{a},\vec{c}]$ (см. 25, стр. 44). Кроме того, определим k-мерный куб $M^{(k)} = [-m;m]^k$.

177| Теорема. [*Многомерная формула обращения.*] *Если* $\varphi(\vec{t})$ — *характеристическая функция вектора* $\vec{\xi}$, *то для любого интервала непрерывности* $(\vec{a};\vec{c}]$ *функции распределения* $F(\vec{x})$ *справедливо равенство*

$$\mathbf{P}\left\{\, \vec{\xi} \in (\vec{a};\vec{c}]\,\right\} \;\stackrel{\mathrm{df}}{=}\; \Delta_{\vec{a},\vec{c}}F \;=\; \lim_{m\to\infty} \int\limits_{M^{(k)}}\!\!\ldots\!\int \prod_{j=1}^{k} \frac{e^{-\mathtt{i}t_j a_j} - e^{-\mathtt{i}t_j c_j}}{(2\pi)^k\, \mathtt{i}t_j}\, \varphi(\vec{t}\,)\, d\vec{t}.$$

В следующей лемме устанавливается простой способ идентификации распределений сл.векторов.

178| Лемма. *Сл.векторы* $\vec{\xi}$, $\vec{\eta}$ *имеют одинаковое распределение:* $\vec{\xi} \backsim \vec{\eta}$, *т. т. т. когда* $\vec{c}^{\flat}\vec{\xi} \backsim \vec{c}^{\flat}\vec{\eta}$ *для* $\forall\, \vec{c} \in \mathbb{R}^k$.

$\models$ По теореме 176 достаточно установить равенство соответствующих хар.ф. Если $\vec{\xi} \backsim \vec{\eta}$, то по свойству ($\text{X}_{\text{k3}}$) теоремы 173 для $\forall\, t \in \mathbb{R}^1$

$$\varphi_{\vec{c}^\flat \vec{\xi}}(t) \;=\; \varphi_{\vec{\xi}}(t\vec{c}) \;=\; \varphi_{\vec{\eta}}(t\vec{c}) \;=\; \varphi_{\vec{c}^\flat \vec{\eta}}(t).$$

Обратно, если $\vec{t}^\flat \vec{\xi} \;\backsim\; \vec{t}^\flat \vec{\eta}$ при $\forall\, \vec{t} \in \mathbb{R}^k$, то

$$\varphi_{\vec{\xi}}(\vec{t}\,) \;=\; \varphi_{\vec{t}^\flat \vec{\xi}}(1) \;=\; \varphi_{\vec{t}^\flat \vec{\eta}}(1) \;=\; \varphi_{\vec{\eta}}(\vec{t}\,). \qquad \Leftarrow\!\!|$$

A. Дополнения

► Если сл.в. ξ имеет все (конечные) моменты $\alpha_k = \mathbf{E}\,\xi^k$, $k \in \mathbb{N}$, то её хар.ф. бесконечно дифференцируема при всех t. В этом случае формальный ряд Маклорена этой хар.ф. полностью определяется указанными моментами. Проанализировав поведение остаточного члена ряда Маклорена, можно найти условия, при которых этот ряд будет сходиться к хар.ф. Поскольку хар.ф. единственным образом описывает распределение, то таким путём можно найти частное решение так называемой проблемы моментов — идентификации распределения сл.в. по её моментам.

[*Проблема моментов.*] *Набор моментов $\alpha_k = \mathbf{E}\,\xi^k$, $k \in \mathbb{N}$, сл.в. ξ единственным образом определяет распределение, если*

а) $\overline{\lim\limits_{k}}\, \dfrac{(\alpha_{2k})^{1/2k}}{k} < \infty$ *или* б) (*Карлеман*) $\sum_1^\infty \dfrac{1}{(\alpha_{2k})^{1/2k}} = \infty$.

(*Условие Крейна.*) *Если плотность распределения $f(x) > 0$, $\forall\, x \in \mathbb{R}^1$, и*

$$\int_{\mathbb{R}} \frac{-\ln f(x)}{1 + x^2}\, dx < \infty,$$

то распределение сл.в. не определяется однозначно своими моментами.

179| Примеры. [20] 1) Пусть $\zeta \backsim \mathcal{N}(0,1)$, тогда сл.в. $\xi = \zeta^3$ удовлетворяет условию Крейна. Для моментов этой сл.в. $\alpha_{2k} = (6k-1)!!$ не выполняются условия а) и б), т.к. по формуле Стирлинга $(\alpha_{2k})^{1/2k} = \mathcal{O}(k^{3/2})$. Если f — плотность ξ, то распределения с плотностью $f_\tau(x) = f(x)\big(1 + \tau \sin\big(\frac{\pi}{6} - \sqrt{3}|x|^{2/3}\big)\big)$, $|\tau| < 1$, обладают теми же моментами, что и ξ.

2) Распределение с плотностью $f(x) = \frac{1}{\sqrt{2\pi}x} \exp\{-\frac{1}{2}\ln^2(x)\}$, $x > 0$, (лог-нормальная модель) также не определяется своими моментами. Легко проверить, что распределения с плотностью $f_\tau(x) = f(x)\big(1 + \tau \sin(2\pi \ln x)\big)$, $|\tau| \leqslant 1$, обладают одинаковыми с f моментами $\alpha_k = \mathbf{E}\,\xi^k = e^{k^2/2}$, $k = 0, 1 \ldots$ Кстати, формальный ряд Маклорена, составленный из этих моментов, имеет нулевой радиус сходимости. ◎

► Способ описания распределения через хар.ф. иногда требует проверки того, что выбранная функция есть хар.ф. некоторой сл.в. (ф.р.).

180| Упр. Используя свойство (X_7), покажите, что функция $e^{-|t|^{2+A}}$ не может быть хар.ф. ни при каком $A > 0$.

Теорема. I) *Пусть $\varphi(t)$, $t \in \mathbb{R}^1$, — непрерывная в нуле комплекснозначная функция, $\varphi(0) = 1$. Функция φ есть хар.ф.*

(Bochner S.): т. т. т. когда φ неотрицательно определена, т.е. для любых наборов действительных чисел $t_1, \ldots, t_n$ и комплексных чисел $z_1, \ldots, z_n$:

$$\sum\nolimits_{j,\,k=1}^{n} \varphi(t_j - t_k) z_j \overline{z}_k \geqslant 0; \tag{7}$$

(Cramér H.): т. т. т. когда φ ограничена и для всех $x \in \mathbb{R}^1$ и $M > 0$

$$\int_0^M \int_0^M \varphi(t-u) e^{-\mathrm{i}x(t-u)}\, dtdu \geqslant 0.$$

II) *(Polya G.) Если φ — действительная исчезающая на бесконечности $(\varphi(\pm\infty) = 0)$ чётная функция, выпуклая книзу при $t > 0$, $\varphi(0) = 1$, то φ — хар.ф. абсолютно непрерывного распределения.*

III) *(Марцинкевич Ю.) Функция вида $\exp\{P_m(t)\}$, где $P_m(t)$ — действительный полином порядка m, есть хар.ф. т. т. т. когда $m \leqslant 2$.*

⊨⇒ (Теорема Бохнера и Крамера.) Во-первых, заметим, что из условия неотрицательной определённости φ следует, что $|\varphi(t)| \leqslant \varphi(0) = 1$, а непрерывность в нуле гарантирует непрерывность для $\forall\, t \in \mathbb{R}^1$. Далее, если в условии (7) для произвольных $t_1, \ldots, t_n$ выбрать $z_j = \exp\{-\mathrm{i}t_j x\}$, $x \in \mathbb{R}^1$, получим неравенство

$$n\varphi(0) + \sum_{j \neq k} \varphi(t_j - t_k) e^{\mathrm{i}(t_j - t_k)x} \ \geqslant\ 0.$$

Проинтегрировав это неравенство по всем $t_j \in [0; M]$, придём к

$$n\varphi(0) M^n + n(n-1) M^{n-2} \int_0^M \int_0^M \varphi(t-u) e^{-\mathrm{i}(t-u)x}\, dtdu \ \geqslant\ 0$$

для $\forall\, M > 0, n \in \mathbb{N}$. В силу произвольности n отсюда получаем условие Крамера. Таким образом, функция

$$f(x; M) \ =\ \frac{1}{M} \int_0^M \int_0^M \varphi(t-u) e^{-\mathrm{i}(t-u)x}\, dtdu \ \geqslant\ 0$$

для $\forall\, x \in \mathbb{R}^1$. Следовательно, после соответствующей нормировки неотрицательная интегрируемая функция $f_T(x;M) = \frac{1}{2\pi T}(T-|x|)f(x;M)$, $|x| \leqslant T$, может выступать в качестве функции плотности.

Произведя замену переменных $t-u \dashrightarrow y, u \dashrightarrow z$ в двойном интеграле, легко показать, что $f(x;M) = \int_{\mathbb{R}} e^{-\mathrm{i}tx}\varphi(t;M)\,dt$, где

$$\varphi(t;M) \;=\; \frac{1}{M}(M-|t|)\varphi(t)\;\dot{\mathrm{I}}\,(|t| \leqslant M).$$

Следовательно, для $u \in \mathbb{R}^1$ функция

$$\begin{aligned}
\varphi_T(u;M) \;:=\; \int\limits_{\mathbb{R}} f_T(x;M)e^{\mathrm{i}ux}\,dx \;=\; \frac{1}{2\pi T}\int\limits_{-T}^{T}(T-|x|)f(x;M)e^{\mathrm{i}ux}\,dx \;= \\
= \frac{1}{2\pi T}\int\limits_{\mathbb{R}}\Big(\int\limits_{-T}^{T}(T-|x|)e^{\mathrm{i}(u-t)x}\,dx\Big)\varphi(t;M)\,dt \;= \\
= \frac{T}{2\pi}\int\limits_{\mathbb{R}}\frac{\sin^2(T(t-u)\frac{1}{2})}{(T(t-u)\frac{1}{2})^2}\varphi(t;M)\,dt \;=\; \frac{1}{\pi}\int\limits_{\mathbb{R}}\frac{\sin^2 z}{z^2}\varphi\big(u+2\tfrac{z}{T};M\big)\,dz,
\end{aligned}$$

где подынтегральная функция по модулю не превосходит интегрируемой на $\mathbb{R}^1$ функции $\dot{\mathrm{I}}(|z| \leqslant 1)\sin^2 z/z^2 + \dot{\mathrm{I}}(|z| > 1)/z^2$. Поэтому здесь можно перейти к пределу под знаком интеграла по $T \to \infty$. Таким образом, предел хар.ф. $\lim_T \varphi_T(u;M)/\varphi_T(0;M) = \varphi(u;M)$ при каждом $u \in \mathbb{R}^1$. Так как по условию функция $\varphi(u;M)$ непрерывна в нуле, то в силу теоремы П. Леви о непрерывном соответствии эта функция также будет хар.ф. некоторого вероятностного закона. Из определения понятно, что в каждой точке $u \in \mathbb{R}^1$ предел $\lim_M \varphi(u;M) = \varphi(u)$, стало быть, и функция φ есть хар.ф. ⫤

Утверждение II) даёт простой способ построения хар.ф. с заданными свойствами. Например, согласно теореме Пойа, функции

$$\begin{aligned}
\varphi_1(t) \;&=\; \big(1-|t|\big)\,\dot{\mathrm{I}}\,(|t| \leqslant 1), \\
\varphi_2(t) \;&=\; \big(1-|t|\big)\,\dot{\mathrm{I}}\,\big(|t| \leqslant \tfrac{1}{2}\,\big) + \tfrac{1}{3}\big(2-|t|\big)\,\dot{\mathrm{I}}\,\big(\,\tfrac{1}{2} \leqslant |t| \leqslant 2\big)
\end{aligned}$$

— это хар.ф. некоторых сл.в. Замечательным здесь является то, что эти функции совпадают на отрезке $[-{}^1/_2;\, {}^1/_2]$, но не совпадают всюду.

► Пример. Пусть дискретная сл.в. $\xi \in \{\pm 2, \pm 3, \ldots\}$ и

$$\mathbf{P}\{\xi = \pm k\} = C\frac{1}{k^2 \ln|k|}, \qquad C = \Big(\sum\nolimits_2^\infty \frac{2}{k^2 \ln k}\Big)^{-1} \approx 0.8257.$$

Применяя интегральный признак сходимости рядов, легко устанавливается, что $\mathbf{E}\,\xi^- = \mathbf{E}\,\xi^+ = +\infty$. С другой стороны, обычными методами анализа можно показать, что хар.ф. этой сл.в.

$$\varphi(t) = 2C \sum\nolimits_2^\infty \frac{\cos kt}{k^2 \ln k}$$

дифференцируема при $t = 0$ и $\varphi'(0) = 0$.

► Приведём ещё ряд важных свойств хар.ф $\varphi(t)$, $t \in \mathbb{R}^1$, сл.в. ξ.

a) $\lim\limits_{|t|\to\infty} |\varphi(t)| = 0$, если распределение абсолютно непрерывно.

b) $\overline{\lim\limits_{|t|\to\infty}} |\varphi(t)| < 1$, если распределение имеет абсолютно непрерывную компоненту (не сингулярно).

c) $|\varphi_\xi(t_0)| = 1$ при некотором $t_0 \neq 0$ т. т. т. когда распределение ξ решётчатое: $\sum_{k=-\infty}^{\infty} \mathbf{P}\{\xi = a + kh\} = 1$, при этом $t_0 = 2\pi/h$.

d) $|\varphi(t)| \leqslant 1 - \frac{1-c^2}{8\delta^2} t^2$ при $|t| < \delta$, если $|\varphi(t)| \leqslant c$ при $|t| \geqslant \delta$.

e) $0 \leqslant (1 - \mathrm{Re}\varphi(2t)) \leqslant 4(1 - \mathrm{Re}\varphi(t)), \quad \forall\, t \in \mathbb{R}^1$.

f) $\frac{1}{2T}\left| \int_{-T}^{T} \varphi(t)\, dt \right| \leqslant \frac{1}{xT} + \big(1 - \frac{1}{xT}\big)\, \mathbf{P}\{|\xi| \leqslant x\}, \qquad T > 0,\ x > \frac{1}{T}$.

181| Упр. Пусть φ — хар.ф. ф.р. F (сл.в. ξ). Показать, что

(а) φ удовлетворяет (7);

(б) справедливо приведённое выше свойство e);

(в) $\lim\limits_{A\to\infty} \dfrac{1}{2A} \displaystyle\int_{-A}^{A} e^{-\mathrm{i}tx} \varphi(t)\, dt = F(x) - F(x-0)$;

(г) $\lim\limits_{A\to\infty} \dfrac{1}{2A} \displaystyle\int_{-A}^{A} \big|\varphi(t)\big|^2 dt = \sum_{x\in\mathbb{R}^1} \big[F(x) - F(x-0)\big]$;

(д) $\dfrac{1}{\pi} \displaystyle\int_{-\infty}^{\infty} \frac{1}{t^2} \big(1 - \mathrm{Re}\,\varphi(t)\big)\, dt = \mathbf{E}\,|\xi|$;

(е) $\displaystyle\int_0^\infty e^{-t} \big(1 - \mathrm{Re}\,\varphi(t)\big)\, dt = \mathbf{E}\left[\frac{\xi^2}{1+\xi^2}\right]$;

(ж) (*равенство Парсеваля*) если $\int_{\mathbb{R}} |\varphi(t)|^2\, dt < \infty$, f — плотность F, то

$$2\pi \int_{\mathbb{R}} f^2(x)\, dx = \int_{\mathbb{R}} \big|\varphi(t)\big|^2\, dt;$$

(з) пусть φ_j — хар.ф., соответствующая ф.р. F_j, $j = 1, 2$, показать, что

$$\int_{-\infty}^{\infty} \varphi_1(t)\, dF_2(t) = \int_{-\infty}^{\infty} \varphi_2(t)\, dF_1(t);$$

(и) привести пример абсолютно непрерывного распределения такого, что $\int_{\mathbb{R}} |\varphi(t)|\,dt = \infty$, а равенство (4) всё же верно, если в нём интеграл понимать по Риману в несобственном смысле;

(к) справедливо приведённое выше свойство c).

✧ **Преобразование Лапласа. Производящие функции.** Формально полагая аргумент хар.ф. $t = \mathrm{i}u$, $u \in \mathbb{R}^1$, получим функцию

$$\mathcal{L}_\xi(u) = \mathbf{E}\, e^{-u\xi},$$

которая называется преобразованием Лапласа. В отличие от хар.ф. преобразование Лапласа существует не для всех сл.в. ξ и не для всех $u \in \mathbb{R}^1$.

Для неотрицательной сл.в. ξ преобразование Лапласа существует при $\forall\, u \geqslant 0$. Для таких сл.в. хар.ф. $\varphi_\xi(z)$ представляет собой аналитическую в области $\operatorname{Im} z > 0$ функцию комплексного аргумента $z = t + \mathrm{i}u$, $u > 0$.

Если $\mathbf{E}\, e^{u\xi} < \infty$ при некотором $u > 0$, то для функции $\overline{F}(x) = \mathbf{P}\{\xi > x\}$ справедлива очевидная оценка

$$e^{ux}\overline{F}(x) \ \leqslant\ \mathbf{E}\left[e^{u\xi}\, \dot{\mathrm{I}}(\xi > x)\right] \ \to\ 0 \quad \text{при } x \to \infty,$$

т.е. правый «хвост» распределения ξ имеет экспоненциальную скорость убывания. В частности, если $m > 0$, то $\mathbf{P}\{(\xi^+)^m > x\} \leqslant e^{-u\sqrt[m]{x}}\,\mathbf{E}\, e^{u\xi}$, и потому сл.в. ξ^+ имеет все моменты: $\mathbf{E}(\xi^+)^m < \infty,\ m \geqslant 0$.

Кроме того, по формуле интегрирования по частям для $\forall\, A > 0$

$$\int_0^A e^{ux}\, dF(x) \ =\ \overline{F}(0) - e^{uA}\overline{F}(0) + u\int_0^A e^{ux}\overline{F}(x)\, dx.$$

Так как $\mathbf{E}\, e^{-u\xi^-} \leqslant 1$ при $u > 0$, то $\mathbf{E}\, e^{u\xi} < \infty$ т. т. т. когда $\int_0^\infty e^{ux}\overline{F}(x)\, dx < \infty$.

182| П р и м е р ы. 1) Преобразование Лапласа для основных распределений:

a) $\xi \backsim \mathcal{E}\mathrm{x}(1) \quad \Rightarrow \quad \mathcal{L}_\xi(u) = \dfrac{1}{1+u}, \qquad u > -1;$

b) $\xi \backsim \mathcal{L}\mathrm{ap}(0,1) \quad \Rightarrow \quad \mathcal{L}_\xi(u) = \dfrac{1}{(1-u^2)}, \qquad |u| < 1;$

c) $\xi \backsim \mathcal{N}(0,1) \quad \Rightarrow \quad \mathcal{L}_\xi(u) = e^{u^2/2}, \qquad u \in \mathbb{R}^1;$

d) $\xi \backsim \mathcal{U}\mathrm{n}(0,1) \quad \Rightarrow \quad \mathcal{L}_\xi(u) = \dfrac{1-e^{-u}}{u}, \qquad u \in \mathbb{R}^1;$

e) $\xi \backsim \mathcal{K}\mathrm{hi}(k) \quad \Rightarrow \quad \mathcal{L}_\xi(u) = (1+2u)^{-k/2}, \quad u > -\dfrac{1}{2}.$

2) Для сл.в. с плотностью $f(x) = \frac{1}{24} e^{-\sqrt[4]{x}}$, $x > 0$, преобразование Лапласа не существует ни для какого отрицательного u, однако все моменты

сл.в. с такой плотностью существуют и конечны. ⊚

Плотность неотрицательной абсолютно непрерывной сл.в. ξ можно восстановить из преобразования Лапласа по формуле Римана–Меллина (формуле обратного преобразования Лапласа):

$$f_\xi(x) \;=\; \mathcal{L}^{-1}(\mathcal{L}_\xi)(x) \;=\; \frac{1}{2\pi\mathtt{i}} \lim_{b\to\infty} \int_{C_{a,b}} e^{xz} \mathcal{L}_\xi(z)\, dz, \tag{8}$$

где $C_{a,b} = \{a + \mathtt{i}u,\ -b < u < b\}$ — отрезок прямой на комплексной плоскости, параллельный мнимой оси, a — произвольное неотрицательное число, для которого плотность $f_\xi(x) \leqslant A\, e^{a\,x}$ для достаточно больших x.

Из предыдущего понятно, что если у знакопеременной сл.в. ξ преобразование $\mathcal{L}(u) < \infty$ в некотором интервале $|u| < t_0$, то у этой сл.в. существуют все моменты. Эти моменты можно найти через функцию $\mathcal{L}$ способом, подобным хар.ф.: $\mathbf{E}\,\xi^k = \partial^k \mathcal{L}(-u)/\partial u^k\big|_{u=0}$. Последнее оправдывает другое название для функции $\mathcal{L}(-u) = \mathbf{E}\, e^{u\xi}$ как *производящей функции моментов* сл.в.

Для дискретных целочисленных сл.в. ξ $(\in \mathbb{N}_0)$ в некоторых случаях вместо хар.ф. удобно рассмотреть так называемую *производящую функцию*:

$$G(t) \;=\; G_\xi(t) \;=\; \mathbf{E}\, t^\xi \;=\; \sum\nolimits_0^\infty p_k t^k,$$

где $|t| \leqslant 1$, $p_k = \mathbf{P}\{\xi = k\}$, $k \in \mathbb{N}_0$.

183| Упр. Покажите, что если $G_\xi(t) = G_\eta(t)$ для $\forall\, |t| < \varepsilon$ с некоторым $\varepsilon > 0$, то $\xi \backsim \eta$.

Хорошо известно, что степенные ряды можно почленно дифференцировать внутри области сходимости. Поэтому

$$G'(t) = \sum\nolimits_1^\infty k p_k\, t^{k-1}, \qquad |t| < 1.$$

Теми же методами, что и для хар.ф., показывается, что $\mu = \mathbf{E}\,\xi = \lim_{t\nearrow 1} G'(t)$, причём это равенство верно как для $\mu < \infty$, так и для $\mu = \infty$. Дальнейшее дифференцирование приводит к так называемым факториальным моментам $\mathbf{E}\left[\xi(\xi-1)\cdots(\xi-r+1)\right] = G^{(r)}(1-)$, которые элементарными формулами могут быть связаны с обычными моментами.

184| Упр. Чему равна дисперсия геометрической сл.в. $\xi \backsim \mathcal{G}\mathsf{eo}(p)$?

B. Некоторые факты из анализа. Семиинварианты

В анализе логарифм комплексного числа $z = u + \mathtt{i}v = |z|(\cos\phi + \mathtt{i}\sin\phi)$, где $|z| = \sqrt{u^2+v^2}$ — модуль z, $\phi = \arg z \in (-\pi;\pi]$ — так называемый аргумент z, определяется как многозначная функция $\mathrm{Ln}\, z = \ln|z| + \mathrm{i}(\phi + 2k\pi)$, $k \in \mathbb{Z}$. Главная ветвь логарифмической функции задаётся в виде $\ln z = \ln|z| + \mathrm{i} \arg z$. Для этой функции могут нарушаться традиционные свойства логарифма. Например, $0 = \ln 1 = \ln((-1)^2) \neq 2\ln(-1) = 2\mathrm{i}\pi$. Равенство $\ln(z_1 z_2) = \ln z_1 + \ln z_2$ выполняется, если $\arg(z_1 z_2) = \arg z_1 + \arg z_2 \in (-\pi; \pi]$, например, если оба числа имеют положительную действительную часть. Кроме того, это свойство справедливо (с любыми сомножителями) для действительной части логарифмической функции: $\mathrm{Re}\ln(z_1 \cdots z_n) = \sum_1^n \mathrm{Re}\ln z_j = \sum_1^n \ln|z_j|$. В частности, $\mathrm{Re}\ln z^n = n\ln|z|$.

Логарифм числа z можно определить также с помощью ряда:

$$\ln z = \sum\nolimits_1^\infty (-1)^{k+1} \frac{(z-1)^k}{k},$$

который сходится в области $|z-1| < 1$, где, как можно показать, оба определения совпадают. Несложно устанавливается, что при $|z-1| < {}^1\!/_2$

$$\ln z \;=\; (z-1) - \frac{1}{2}(z-1)^2 \delta_1(z) \;=\; (z-1) - \frac{1}{2}(z-1)^2 + \frac{1}{3}(z-1)^3 \delta_2(z), \qquad (9)$$

где $|\delta_1|, |\delta_2| < 2$.

Степенную функцию (главную ветвь) можно определить как

$$z^\alpha \;=\; e^{\alpha \ln z} \;=\; |z|^\alpha e^{\mathrm{i}\,\alpha \arg z} \;=\; |z|^\alpha \big(\cos(\alpha\phi) + \mathrm{i}\sin(\alpha\phi)\big), \quad \alpha \in \mathbb{R}^1.$$

Здесь также равенство $(z_1 z_2)^\alpha = z_1^\alpha z_2^\alpha$ имеет место, только если $\arg(z_1 z_2) = \arg z_1 + \arg z_2$. Равенство $\ln z^\alpha = \alpha \ln z$ справедливо при $|\alpha \arg z| < \pi$.

✦ При $t \to 0$ хар.ф. $\varphi(t) = \mathbf{E}\exp(\mathrm{i}t\xi) \to 1$, т.е. $|\varphi(t) - 1| < 1$ в некоторой окрестности $t = 0$, и здесь можно определить функцию $\psi(t) = \ln\varphi(t)$. Если у сл.в. существуют все моменты до m включительно, то, как и хар.ф., функция ψ допускает в окрестности $t = 0$ разложение в ряд:

$$\psi(t) \;=\; \sum\nolimits_1^m \frac{\mathrm{i}^k}{k!}\gamma_k t^k + o(t^m) \;=\; \sum\nolimits_1^{m-1} \frac{\mathrm{i}^k}{k!}\gamma_k t^k + \frac{\mathrm{i}^m}{m!}\psi^{(m)}(\theta t)\, t^m,$$

где $0 \leqslant \theta \leqslant 1$. Коэффициенты γ_k называются *семиинвариантами* (*кумулянтами*) в силу того факта, что при сложении независимых сл.в. соответствующие семиинварианты также складываются. В частности, т.к. для (детерминированной) сл.в. $\xi \equiv \mu$ имеем $\psi(t) = \mathrm{i}\mu t$, т.е. все её старшие семиинварианты $\gamma_k = 0$, $k \geqslant 2$, то для любой сл.в. ξ все старшие семиинварианты совпадают с семиинвариантами сл.в. $\xi - \mu$, где $\mu = \mathbf{E}\xi$. Стало быть, при вычислении старших кумулянт γ_k, $k \geqslant 2$, можно считать $\varphi'\big|_{t=0} = 0$.

Итак, $\varphi\big|_{t=0} = 1,\ \varphi''\big|_{t=0} = \mathtt{i}^2\mathbf{D}\xi = -\sigma^2,\ \varphi'''\big|_{t=0} = -\mathtt{i}\,\mathbf{E}(\xi-\mu)^3$. Поэтому

$$\gamma_1 = \frac{1}{\mathtt{i}}\psi'(t)\big|_{t=0} = \frac{1}{\mathtt{i}}\frac{\varphi'}{\varphi}\big|_{t=0} = \mu = \mathbf{E}\,\xi,$$

$$\gamma_2 = \frac{1}{\mathtt{i}^2}\psi''(t)\big|_{t=0} = -\frac{\varphi''\varphi-(\varphi')^2}{\varphi^2}\big|_{t=0} = \sigma^2 = \mathbf{D}\xi,$$

$$\gamma_3 = \frac{1}{\mathtt{i}^3}\psi'''(t)\big|_{t=0} = -\frac{1}{\mathtt{i}}\,\frac{\varphi'''\varphi^3+\varphi''\varphi'\varphi^2-2\varphi'\varphi''\varphi^2-2\varphi''\varphi^2\varphi'+2(\varphi')^3\varphi}{\varphi^4}\big|_{t=0} = \mathbf{E}(\xi-\mu)^3.$$

Аналогичными вычислениями показывается, что $\gamma_4 = \mathbf{E}(\xi-\mu)^4 - 3\sigma^4$. Следовательно, с точностью до множителя $1/\sigma^3$ третий кумулянт γ_3 совпадает с коэффициентом асимметрии, а четвёртый кумулянт γ_4 с точностью до множителя $1/\sigma^4$ совпадает с коэффициентом эксцесса.

185| Упр. У какого распределения кумулянты $\gamma_k = 1,\ \forall\, k \geqslant 1$?

► Найдём разложение в ряд функции $e^{\mathtt{i}t}$.

Теорема. *Для* $\forall\, t \in \mathbb{R}^1$ *справедливо*

$$e^{\mathtt{i}t} \ = \ \sum\nolimits_0^n \frac{\mathtt{i}^k}{k!}t^k + \frac{t^n}{n!}\delta_n(t), \tag{10}$$

где $|\delta_n(t)| \leqslant \min\big(2,\ \frac{|t|}{(n+1)}\big)$, *следовательно,* $\delta_n(t) \to 0$ *при* $t \to 0$.

⤃ Установим сначала тождество (с $t_0 = t$)

$$(\sharp)\qquad e^{\mathtt{i}t} - \sum\nolimits_{k=0}^n \frac{\mathtt{i}^k}{k!}t^k \ = \ \mathtt{i}^{n+1}\int_0^{t_0}dt_1\int_0^{t_1}dt_2\cdots\int_0^{t_n}e^{\mathtt{i}t_{n+1}}dt_{n+1}.$$

Обозначим правую часть через $J_{n+1},\ n \geqslant 1$. Тогда, очевидно,

$$(*)\qquad J_{n+1} \ = \ \mathtt{i}^n\int_0^{t_0}dt_1\int_0^{t_1}dt_2\cdots\int_0^{t_{n-1}}(e^{\mathtt{i}t_n}-1)dt_n \ = \ J_n - \mathtt{i}^nV_n,$$

$$\text{где}\quad V_n \ = \ \int_0^{t_0}dt_1\cdots\int_0^{t_{n-1}}dt_n \ = \ \int_0^{t_0}dt_1\cdots\int_0^{t_{n-2}}t_{n-1}dt_{n-1} \ = \ldots = \ \frac{t_0^n}{n!} \ = \ \frac{t^n}{n!}.$$

Поэтому $J_{n+1} = J_n - \mathtt{i}^n\frac{t^n}{n!},\ J_1 = \mathtt{i}\int_0^t e^{\mathtt{i}t_1}dt_1 = e^{\mathtt{i}t} - 1$, что и доказывает $(\sharp)$.

Воспользовавшись тем, что $|e^{\mathtt{i}t}| = 1$, из определения получаем неравенство $|J_{n+1}| \leqslant |V_{n+1}| = |t|^{n+1}/((n+1)!)$. Если же в первой части $(*)$ применить неравенство треугольника $|e^{\mathtt{i}t_n}-1| \leqslant |e^{\mathtt{i}t_n}|+1 = 2$, получим неравенство $|J_{n+1}| \leqslant 2|V_n| = 2|t|^n/(n!)$. Наконец,

$$\lim_{t\to 0}|\delta_n(t)| \ = \ \lim_{t\to 0}\frac{|J_{n+1}(t)|\,n!}{|t|^n} \ \leqslant \ \lim_{t\to 0}\frac{|t|}{n+1} \ = \ 0. \qquad ⤂$$

186 **Следствие.** *Для всех* $t \in \mathbb{R}^1$

$$\begin{aligned} e^{it} &= 1 + it + 2t\,\Delta_1^{(0)}(t) &&= 1 + it + \frac{1}{2}t^2\,\Delta_1(t) = \\ &= 1 + it - \frac{1}{2}t^2 + t^2\,\Delta_2^{(0)}(t) &&= 1 + it - \frac{1}{2}t^2 + \frac{1}{6}t^3\,\Delta_2(t), \end{aligned}$$

где $|\Delta_k^{(0)}|$, $|\Delta_k| \leqslant 1$, $\lim\limits_{t\to 0}\Delta_k^{(0)}(t) = 0$, $k = 1, 2$.

► Преобразование Фурье интегрируемой функции исчезает на бесконечности. Точнее, справедливо следующее утверждение.

[*Лемма Римана-Лебега.*] *Если функция* $f(x)$, $x \in \mathbb{R}^1$, *такова, что интеграл Римана* $\int_{\mathbb{R}} |f(x)|\,dx < \infty$, *то* $\lim_{t\to\pm\infty}\int_{-\infty}^{\infty} e^{itx} f(x)\,dx = 0$.

⊨ Если функция f непрерывно дифференцируема, то на любом отрезке $[a;c]$ как сама функция, так и её производная ограничены. Поэтому при $|t| \to \to \infty$ косинус-преобразование (для синус-преобразования аналогично)

$$\int_a^c f(x)\cos(tx)\,dx \;=\; \frac{1}{t}\Big(f(x)\sin(tx)\Big|_a^c - \int_a^c f'(x)\sin(tx)\,dx\Big) \;\to\; 0.$$

Разбивая область интегрирования на соответствующие части, можно показать, что этот факт справедлив и для ступенчатых функций. Дальнейшее доказательство сводится к применению двух известных утверждений: а) для абсолютно интегрируемой функции f при $\forall\varepsilon > 0$ найдётся отрезок $[a;c]$ такой, что $|\int_c^{\infty} + \int_{-\infty}^{a} f(x)\cos(tx)dx| < \varepsilon$, б) интегрируемая на отрезке функция может быть приближена (в смысле $\mathcal{L}_1$) ступенчатыми функциями. ⫤

С. Указания к решению задач

✦ 157, стр. 158. (b) $\mathbf{E}\left[\zeta_1\zeta_2\right] = \mathbf{E}\left[\alpha_1\alpha_2\right] - \mathbf{E}\left[\vartheta_1\vartheta_2\right] + \mathrm{i}\big(\mathbf{E}\left[\alpha_1\vartheta_2\right] - \mathbf{E}\left[\vartheta_1\alpha_2\right]\big)$.

✦ 159, стр. 161. (X_4) $\mathbf{E}\,e^{\mathrm{i}t\eta} = e^{\mathrm{i}t\,b}\,\mathbf{E}\,e^{\mathrm{i}t\xi}$; ($X_5$) $\mathbf{E}\,e^{\mathrm{i}(-t)\xi} = \mathbf{E}\,\overline{e^{\mathrm{i}t\xi}} = \overline{\mathbf{E}\,e^{\mathrm{i}t\xi}}$; ($X_9$) воспользоваться леммой 156, стр. 157.

✦ 161, стр. 163. b) Биномиальная сл.в. есть сумма бернуллиевских; d) воспользоваться $e^z = \sum_0^\infty z^k/k!$; f) $\varphi(t) = \mathbf{E}\cos(t\xi) = \frac{1}{2}\int_{-1}^{1}\cos(tx)\,dx$.

✦ 164, стр. 165. a) Хар.ф. разности независимых одинаково распределённых сл.в. равна произведению хар.ф. $\varphi(t)\varphi(-t)$; б) заметить, что

$$\tfrac{1}{2}\int_{-\infty}^{\infty}\cos(tx)e^{-|x|}\,dx = \mathrm{Re}\int_0^\infty e^{\mathrm{i}tx}e^{-x}\,dx = \mathrm{Re}(1-\mathrm{i}t)^{-1} = \mathrm{Re}(1+\mathrm{i}t)(1+t^2)^{-1}.$$

✦ 165, стр. 165. $\int_{-\pi}^{\pi} e^{-\mathrm{i}mt}\varphi(t)\,dt = \sum_{k=-\infty}^{\infty}\mathbf{P}\{\xi = k\}\int_{-\pi}^{\pi} e^{-\mathrm{i}(m-k)t}\,dt$. Интеграл отличен от нуля только при $m - k = 0$.

✦ 167, стр. 167. Обратное преобразование: a) $\frac{2}{2\pi}\int_0^1\cos(tx)(1-t)\,dt = (1 - -\cos x)/(\pi x^2)$; b) $\frac{2}{2\pi}\int_0^\infty\cos(tx)e^{-t}\,dt$, далее два раза по частям.

✦ 172, стр. 168. Если $\xi_j \backsim \mathtt{Cauch}(\mu_j, \sigma_j)$, то хар.ф. $\varphi_{\xi_j} = e^{\mathtt{i}t\mu_j - \sigma_j|t|}$, поэтому для $S = \xi_1 + \ldots + \xi_n$ хар.ф. $\varphi_S(t) = e^{\mathtt{i}tM - \mathfrak{S}|t|}$ с $M = \mu_1 + \ldots + \mu_n, \mathfrak{S} = \sigma_1 + \ldots + \sigma_n$.

✦ 174, стр. 169. ($\mathrm{X}_{\mathrm{k}2}$) Хар.ф. $\vec{\eta} = (\xi_1, \ldots, \xi_m)$ равна $\mathbf{E}\exp\big(\mathtt{i}(t_1\xi_1 + \ldots + t_m\xi_m + 0 \cdot \xi_{m+1} + \ldots + 0 \cdot \xi_k)\big)$.

($\mathrm{X}_{\mathrm{k}3}$) $\varphi_{\vec{\eta}}(\vec{u}) = e^{\mathtt{i}\,\vec{u}^\flat\,\vec{b}}\,\mathbf{E}\left[e^{\mathtt{i}\,\vec{u}^\flat Q\vec{\xi}}\right]$.

($\mathrm{X}_{\mathrm{k}4}$) Аналогично свойству (X_7).

($\mathrm{X}_{\mathrm{k}5}$) Воспользоваться (10) в виде $e^{\mathtt{i}t\xi_j} = 1 + \mathtt{i}t\xi_j - \frac{1}{2}t^2\xi_j^2 + t^2\delta_{j2}(t)$ с $\mathbf{E}|\delta_{j2}(t)| \to 0$.

✦ 180, стр. 172. При $A > 0$ вторая производная функции в нуле равна нулю, что возможно, только если сл.в. $\xi = 0$ (п.н.).

✦ 181, стр. 174. (б) Доказать неравенство $1 - \cos 2x \leqslant 4(1 - \cos x)$.

(в) Так как $J_A(z) := \dfrac{1}{2A}\int_{-A}^{A} e^{-\mathtt{i}tz}\,dt = \sin(Az)/(Az) \in (-1; 1],\ z \neq 0,$ и $J_A(0) = 1$, то по теореме Фубини

$$\frac{1}{2A}\int_{-A}^{A} e^{-\mathtt{i}tx}\varphi(t)\,dt = \mathbf{E}\,J_A(\xi - x) = \mathbf{E}\left[J_A(\xi - x)\,\dot{\mathrm{I}}(\xi \neq x)\right] + \mathbf{P}\{\xi = x\} \underset{A}{\to} \mathbf{P}\{\xi = x\}.$$

(г) Применить (в) к хар.ф. разности $\xi_1 - \xi_2$, где независимые $\xi_1, \xi_2 \backsim \xi$.

(д) $\int_{-\infty}^{\infty}\big(1 - \cos(tx)\big)t^{-2}\,dt = \pi|x|$.

(е) $\int_0^{\infty} e^{-t}(1 - \cos(tx))\,dt = x^2/(1 + x^2)$.

(ж) Применить формулу обращения к хар.ф. разности двух независимых сл.в.

(з) Применить теорему Фубини.

(и) $\xi \sim \mathtt{Un}(-1, 1)$.

(к) В выражении для хар.ф. решётчатого распределения учесть, что $\exp\{\mathtt{i}2\pi j\} = 1$ для целых j. Обратно, $|\varphi(t_0)| = 1 \Leftrightarrow \varphi(t_0) = \exp\{\mathtt{i}a\}$ с некоторым $a \in \mathbb{R}^1$. Следовательно, $\mathbf{E}\left[e^{\mathtt{i}(t_0\xi - a)}\right] = 1$ или $0 = \mathbf{E}\big[\cos(t_0\xi - a) - 1\big] \leqslant 0$, что возможно, только когда $\mathbf{P}\left\{t_0\xi - a \in \left\langle 2\pi k, k = 0, \pm 1, \ldots\right\rangle\right\} = 1$.

✦ 183, стр. 176. $p_0 = G(0),\ p_1 = (G(t) - p_0)/t\big|_{t=0}, \ldots$ Или так: $G^{(k)}(0) = k!p_k,\ k > 0$.

✦ 184, стр. 176. Найти первый момент и второй факториальный момент через производящую функцию $p/(1 - t(1 - p))$ (см. 160, стр. 161) для $\xi \backsim \mathtt{Geo}(p)$.

✦ 185, стр. 178. $\mathtt{Pois}(1)$.

V Многомерный нормальный закон

Нормальный закон распределения одномерных случайных величин появился как предельный закон в биномиальной модели (теорема Муавра–Лапласа). Аналогичные выкладки, применённые к полиномиальной модели, приведут к многомерному варианту этой теоремы. Мы не будем подробно останавливаться здесь на этом варианте, поскольку с помощью метода характеристических функций устанавливается, что нормальное распределение возникает и в более общих ситуациях. Здесь же мы заблаговременно ознакомимся с многомерной нормальной моделью и её свойствами.

Определение. Сл.вектор $\vec{\xi} = (\xi_1, \ldots, \xi_k)$ имеет *k-мерное нормальное (гауссовское) распределение* с параметрами $(\vec{\mu}, \mathbf{\Sigma})$, где $\vec{\mu} \in \mathbb{R}^k$, $\mathbf{\Sigma} > 0$ — положительно определённая матрица размера $k \times k$, коротко $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma})$, если его функция плотности при $\vec{x} \in \mathbb{R}^k$ имеет вид

$$f_{\vec{\xi}}(\vec{x}) = \frac{1}{(\sqrt{2\pi}\,)^k \sqrt{|\mathbf{\Sigma}|}} \exp\Big\{-\frac{1}{2}(\vec{x} - \vec{\mu})^\flat\, \mathbf{\Sigma}^{-1}(\vec{x} - \vec{\mu})\Big\}. \tag{1}$$

Прежде всего необходимо проверить, что $\int_{\mathbb{R}^k} f_{\vec{\xi}}(\vec{x})\, d\vec{x} = 1$, т.е. (1) действительно есть ф.плотности. Вместо прямого вычисления этого интеграла заметим, что при $\mathbf{\Sigma} = \dot{\mathbb{I}}$ (т.е. с единичной матрицей) функция (1) есть плотность k-мерного вектора с независимыми нормальными компонентами. Покажем, что (1) есть плотность сл.вектора, полученного как линейное преобразование вектора с независимыми нормальными компонентами (см. замечание 189).

§ 1. Линейные преобразования нормального вектора

187| **Лемма.** *Если* $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma})$, *то для любой невырожденной матрицы* $B = B^{k\times k}$ *и* $\forall\, \vec{c} \in \mathbb{R}^k$

$$B\vec{\xi} + \vec{c} \;\backsim\; \mathcal{N}_k(\vec{a}, \mathfrak{S})$$

с параметрами $\vec{a} = B\vec{\mu} + \vec{c}$, $\mathfrak{S} = B\mathbf{\Sigma}B^\flat$.

$\Rrightarrow$ Напомним, что ф.пл. линейного преобразования $\vec{\eta} = B\vec{\xi} + \vec{c}$ связана с ф.пл. $f_{\vec{\xi}}(\vec{x})$ вектора $\vec{\xi}$ равенством $f_{\vec{\eta}}(\vec{y}) = f_{\vec{\xi}}\,(B^{-1}(\vec{y} - \vec{c}))/\big|\,|B|\,\big|$. При подстановке $\vec{x} \dashrightarrow B^{-1}(\vec{y} - \vec{c})$ квадратичная форма в выражении (1)

$$\begin{aligned}(\vec{x} - \vec{\mu})^\flat \mathbf{\Sigma}^{-1}(\vec{x} - \vec{\mu}) \;&=\; \big(B^{-1}(\vec{y} - \vec{c}) - \vec{\mu}\big)^\flat \mathbf{\Sigma}^{-1}\big(B^{-1}(\vec{y} - \vec{c}) - \vec{\mu}\big) \;= \\ &=\; \big(\vec{y} - B\vec{\mu} - \vec{c}\big)^\flat (B^\flat)^{-1}\, \mathbf{\Sigma}^{-1}\, B^{-1}\, \big(\vec{y} - B\vec{\mu} - \vec{c}\big),\end{aligned}$$

где мы вынесли B^{-1} (влево) и воспользовались хорошо известными свойствами операций транспонирования и взятия обратной матрицы: $(AQ)^\flat = Q^\flat A^\flat$, $(AQ)^{-1} = Q^{-1}A^{-1}$, $(A^\flat)^{-1} = (A^{-1})^\flat$.

Поскольку обратная матрица $\mathfrak{S}^{-1} = (B\mathbf{\Sigma}B^\flat)^{-1} = (B^\flat)^{-1}\mathbf{\Sigma}^{-1}B^{-1}$, квадратичная форма относительно переменной $\vec{y}$, стоящая под знаком экспоненты, будет определяться матрицей $\mathfrak{S}^{-1}$ и сдвиговым параметром $\vec{a} = B\vec{\mu} + \vec{c}$. Осталось выяснить, как повлияет якобиан преобразования на константу, стоящую перед знаком экспоненты. Так как модуль определителя $\|B\| = \sqrt{|B|\,|B^\flat|}$, то $\|B\|\,\sqrt{|\mathbf{\Sigma}|} = \sqrt{|B|\,|B^\flat|\,|\mathbf{\Sigma}|} = \sqrt{|\mathfrak{S}|}$. $\Lleftarrow$

✧ **Представление через стандартное распределение.** Как и в одномерном случае, здесь существует линейное преобразование, «стандартизирующее» $\vec{\xi}$. Роль стандартного отклонения играет квадратный корень из матрицы $\mathbf{\Sigma}$, т.е. матрица Q, обозначаемая $\sqrt{\mathbf{\Sigma}}$, такая, что $QQ^\flat = \mathbf{\Sigma}$ (см. стр. 534); причём, очевидно, Q невырожденная т. т. т. когда $\mathbf{\Sigma} > 0$.

188| **Лемма.** I) *Если вектор* $\vec{\eta} = (\eta_1, \ldots, \eta_k)$ *состоит из независимых нормальных* $\mathcal{N}(0,1)$ *сл.в.,* $\vec{\mu} \in \mathbb{R}^k$, *матрица* B *невырожденная, то*

$$\vec{\xi} \;=\; B\vec{\eta} + \vec{\mu} \;\backsim\; \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma}) \quad \text{с} \quad \mathbf{\Sigma} = BB^\flat.$$

II) *Обратно, если* $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma})$, $\mathbf{\Sigma} > 0$, *и матрица* $Q = \sqrt{\mathbf{\Sigma}}$, *то компо-*

ненты случайного вектора $\vec{\eta} = Q^{-1}(\vec{\xi} - \vec{\mu})$ *независимы и имеют нормальное* $(0,1)$ *распределение:* $\eta_j \backsim \mathcal{N}(0,1), \quad j = \overline{1,k}$.

⊨ Для доказательства части II) применим утверждение леммы 187 с $B = Q^{-1}$, $\vec{c} = -Q^{-1}\vec{\mu}$. В этом случае параметр сдвига $\vec{a} = \vec{0}$, а матрица

$$\mathfrak{S} \;=\; Q^{-1}\mathbf{\Sigma}(Q^{\flat})^{-1} \;=\; Q^{-1}QQ^{\flat}(Q^{\flat})^{-1} \;=\; \dot{\mathbb{I}}.$$

Квадратичная форма, определяемая единичной матрицей, равна сумме квадратов компонент вектора. Поэтому функция плотности $\vec{\eta}$

$$f_{\vec{\eta}}(\vec{y}) \;=\; \frac{1}{(\sqrt{2\pi})^k}\exp\Big\{-\sum_{j=1}^{k}\frac{1}{2}y_j^2\Big\} \;=\; \prod_{j=1}^{k}\frac{1}{\sqrt{2\pi}}\exp\Big\{-\frac{1}{2}y_j^2\Big\}, \tag{2}$$

т.е. равна произведению k стандартных нормальных функций плотности, что гарантирует также и независимость соответствующих сл.величин. ⫤

189| ⚠ Первая часть леммы есть не что иное, как частный случай леммы 187, применённой к сл.вектору с единичной матрицей ковариаций. Мы выделили этот факт в отдельное утверждение, т.к. он даёт очень простой способ генерирования многомерных нормальных сл.в. с заданной матрицей ковариаций. Кроме того, т.к. функция (2) (равная произведению плотностей) действительно есть функция плотности, то из первого утверждения леммы следует, что функция плотности сл.вектора $\vec{\xi} = \sqrt{\mathbf{\Sigma}}\,\vec{\eta} + \vec{\mu}$ совпадает с (1).

§2. Моменты нормального закона. Двумерное распределение

Выясним смысл параметров нормального распределения.

190| **Теорема.** *Если вектор* $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma})$, *то*

$$\mathbf{E}\,\vec{\xi} = \vec{\mu}, \quad \mathbf{Cov}(\vec{\xi}) = \mathbf{\Sigma}.$$

⊨ По лемме 188 компоненты вектора $\vec{\eta} = (\sqrt{\mathbf{\Sigma}}\,)^{-1}(\vec{\xi} - \vec{\mu})$ независимы и нормальны $(0,1)$. Поэтому все ковариации $\mathbf{Cov}(\eta_i, \eta_j) = 0, \;\; i \neq j$, а дисперсии $\mathbf{D}\eta_j = 1$. Следовательно, матрица ковариаций $\mathbf{Cov}(\vec{\eta}\,) = \dot{\mathbb{I}}$. Кроме того, вектор средних $\mathbf{E}\,\vec{\eta} = \vec{0}$. Поэтому $\mathbf{E}\,\vec{\xi} = \mathbf{E}(\sqrt{\mathbf{\Sigma}}\,\vec{\eta} + \vec{\mu}) = \sqrt{\mathbf{\Sigma}}\,\vec{0} + \vec{\mu} = \vec{\mu}$, а матрица ковариаций (по лемме 75, стр. 76) $\mathbf{Cov}(\vec{\xi}\,) = \sqrt{\mathbf{\Sigma}}\,\dot{\mathbb{I}}\,\sqrt{\mathbf{\Sigma}}^{\,\flat} = \mathbf{\Sigma}$. ⫤

⚠ Запись $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma})$ читают так: вектор $\vec{\xi}$ имеет k-мерное нормальное распределение с вектором средних $\vec{\mu}$ и матрицей ковариаций $\mathbf{\Sigma}$.

► Исследуем чуть подробнее распределение двумерного нормального вектора. Если вспомнить связь между коэффициентами ковариации и корреляции, то можно записать двумерную матрицу ковариаций в виде

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

где σ_j^2 — дисперсия ξ_j, $j = 1, 2$, ρ — коэффициент корреляции между ξ_1, ξ_2. Определитель матрицы $|\mathbf{\Sigma}| = \sigma_1^2\sigma_2^2(1-\rho^2)$, обратная матрица

$$\mathbf{\Sigma}^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{pmatrix}.$$

Определение. Вектор (ξ_1, ξ_2) имеет двумерное нормальное распределение со средними (μ_1, μ_2), дисперсиями (σ_1^2, σ_2^2) и коэффициентом корреляции ρ, если его функция плотности (при $(x_1, x_2) \in \mathbb{R}^2$)

$$f(x_1, x_2) = \frac{1}{2\pi\,\sigma_1\sigma_2\sqrt{1-\rho^2}} \times$$
$$\times \exp\Big\{-\frac{1}{2(1-\rho^2)}\Big(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)}{\sigma_1}\frac{(x_2-\mu_2)}{\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\Big)\Big\}.$$

На рис. 3 приведены графики двумерной нормальной плотности с нулевыми средними и единичными дисперсиями при коэффициентах корреляции $\rho = 0$ и $\rho = 0.5$ соответственно.

Линии уровней графиков, т.е. линии, задаваемые уравнениями вида $f(x_1, x_2) = c$, $c > 0$, определяют эллипсы с центром в точке (μ_1, μ_2). Эти

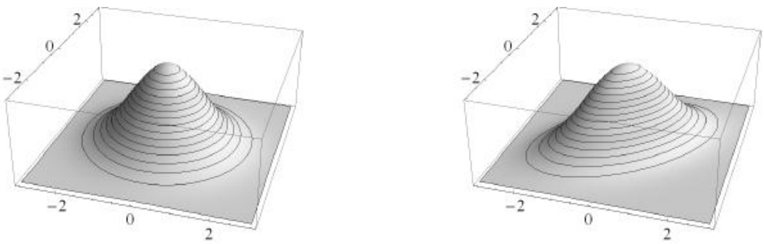


Рис. 3. Плотность двумерного нормального закона

эллипсы обладают интересным свойством: условное распределение $\vec{\xi}$ при условии, что $\vec{\xi}$ принадлежит какому-то эллипсу, равномерно на нём (см. [1]). Другими словами, если выделить реализации вектора, принадлежащие какой-то одной линии уровня, то частота попадания на любую дугу этой линии пропорциональна её длине. Поэтому эллипсы, задающие линии уровней нормальной плотности, называют эллипсами равных вероятностей.

Легко видеть, что при коэффициенте корреляции $\rho = 0$ двумерная функция плотности может быть представлена в виде произведения двух нормальных плотностей, каждая из которых зависит только от одного из аргументов — x_1 или x_2. По известному свойству это означает, что случайные величины ξ_1 и ξ_2 независимы. Другими словами, компоненты двумерного нормального вектора независимы т. т. т. когда они не коррелируют. Позднее мы докажем подобное свойство для общего многомерного нормального распределения.

191 Упр. Матрицу $Q = \sqrt{\mathbf{\Sigma}}$, фигурирующую в лемме 188, можно выбрать многими способами. Приведите варианты такой матрицы для двумерной матрицы ковариаций $\mathbf{\Sigma}$: а) в частном случае, когда дисперсии $\sigma_1^2 = \sigma_2^2 = 1$; б) для произвольных дисперсий.

§ 3. Характеристическая функция. Независимость

Найдём теперь хар.ф. нормального вектора, с помощью которой легко выводятся многие свойства нормального распределения.

192 **Теорема.** *Хар.функция вектора* $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma})$ *при* $\vec{t} \in \mathbb{R}^k$

$$\varphi_{\vec{\xi}}(\vec{t}\,) \;=\; \mathbf{E}\, e^{\mathrm{i}\vec{t}^{\,\flat}\vec{\xi}} \;=\; \exp\Big\{\mathrm{i}\,\vec{t}^{\,\flat}\,\vec{\mu} - \frac{1}{2}\vec{t}^{\,\flat}\,\mathbf{\Sigma}\,\vec{t}\,\Big\}. \tag{3}$$

⊨ Рассмотрим сначала вектор $\vec{\eta}$ с независимыми $\mathcal{N}(0,1)$ компонентами. Очевидно, хар.ф. $\vec{\eta}$ равна произведению хар.ф. компонент:

$$\varphi_{\vec{\eta}}(\vec{u}\,) \;=\; \mathbf{E}\, e^{\mathrm{i}\vec{u}^{\,\flat}\vec{\eta}} \;=\; \prod_{j=1}^{k} \mathbf{E}\, e^{\mathrm{i}\,u_j\eta_j} \;=\; \prod_{j=1}^{k} e^{-\frac{1}{2}u_j^2} \;=\; e^{-\frac{1}{2}\vec{u}^{\,\flat}\vec{u}}.$$

По лемме 188 вектор $\vec{\xi} \backsim Q\vec{\eta} + \vec{\mu}$ с $Q = \sqrt{\mathbf{\Sigma}}$. Поэтому хар.ф. $\vec{\xi}$

$$\varphi_{\vec{\xi}}(\vec{t}\,) \;=\; \mathbf{E}\, e^{\mathrm{i}\,\vec{t}^{\,\flat}(Q\vec{\eta}+\vec{\mu})} \;=\; e^{\mathrm{i}\,\vec{t}^{\,\flat}\vec{\mu}}\;\mathbf{E}\, e^{\mathrm{i}\,(Q^\flat\vec{t}\,)^\flat\,\vec{\eta}} \;=\; e^{\mathrm{i}\vec{t}^{\,\flat}\vec{\mu}}\;\varphi_{\vec{\eta}}(Q^\flat\vec{t}\,).$$

Так как $QQ^\flat = \mathbf{\Sigma}$, то $(Q^\flat\vec{t}\,)^\flat(Q^\flat\vec{t}\,) = \vec{t}^{\,\flat}\,\mathbf{\Sigma}\,\vec{t}$, что и требовалось. ⫤

⚠ Утверждение теоремы может быть положено в основу определения нормального закона. При таком определении можно отказаться от условия невырожденности матрицы ковариаций $\mathbf{\Sigma}$ и требовать только её неотрицательную определённость. В этом случае формально можно считать распределение на числовой прямой, сосредоточенное в одной точке μ, также нормальным $\mathcal{N}(\mu,\sigma^2)$ с дисперсией $\sigma^2 = 0$. Распределение с вырожденной матрицей ковариаций сосредоточено в пространстве меньшей размерности и потому не может быть описано с помощью плотности относительно меры Лебега в k-мерном пространстве.

193| Лемма. (?!) *Для любой неотрицательно определённой матрицы* $\mathbf{\Sigma} \geqslant 0$ *функция (3) есть хар.ф. некоторого сл.вектора.*

Утверждение леммы 187 распространяется теперь на любую матрицу преобразований.

194| Лемма. (?!) *Пусть* $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu},\mathbf{\Sigma})$, B *— действительная* $(m\times k)$*-матрица, вектор* $\vec{r}\in\mathbb{R}^m$. *Тогда сл.вектор* $\vec{\eta} = B\vec{\xi} + \vec{r} \backsim \mathcal{N}_m(B\vec{\mu}+\vec{r}, B\,\mathbf{\Sigma}B^\flat)$.

Если $|\mathbf{\Sigma}| > 0$ *и ранг* $\operatorname{rang}(B) = m$, *то распределение* $\vec{\eta}$ *имеет плотность относительно меры Лебега в* $\mathbb{R}^m$.

Ещё один способ определения многомерного нормального распределения предоставляет следующая

195| Теорема. *Вектор* $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu},\mathbf{\Sigma})$, $\mathbf{\Sigma} > 0$, *т. т. т. когда любая ненулевая линейная комбинация* $\vec{c}^{\,\flat}\vec{\xi} = c_1\xi_1 + \ldots + c_k\xi_k$ *имеет нормальное распределение со средним* $M = \vec{c}^{\,\flat}\vec{\mu}$ *и дисперсией* $\sigma^2 = \vec{c}^{\,\flat}\mathbf{\Sigma}\vec{c} > 0$.

⤇ Пусть $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu},\mathbf{\Sigma})$, тогда для $\forall\,\vec{c}\in\mathbb{R}^k$, $t\in\mathbb{R}^1$, хар.ф. $\eta = \vec{c}^{\,\flat}\vec{\xi}$

$$\varphi_\eta(t) \;=\; \varphi_{\vec{\xi}}(t\vec{c}) \;=\; \exp\Big\{\mathrm{i}(t\vec{c})^\flat\,\vec{\mu} - \frac{1}{2}(t\vec{c})^\flat\,\mathbf{\Sigma}(t\vec{c})\Big\} \;=\; \exp\Big\{\mathrm{i}tM - \frac{1}{2}t^2\sigma^2\Big\}.$$

Обратно, пусть для $\vec{c}\in\mathbb{R}^k$ сл.в. $\eta = \vec{c}^{\,\flat}\vec{\xi} \backsim \mathcal{N}(M,\sigma^2)$ с заявленными в теореме параметрами. Тогда хар.ф. η в точке $t = 1$

$$\varphi_\eta(1) \;=\; e^{\mathrm{i}M - \frac{1}{2}\sigma^2} \;=\; e^{\mathrm{i}\,\vec{c}^{\,\flat}\,\vec{\mu} - \frac{1}{2}\,\vec{c}^{\,\flat}\,\mathbf{\Sigma}\,\vec{c}}.$$

Заметим, что хар.ф. $\vec{\xi}$ в точке $\vec{t} = \vec{c}$ равна хар.ф. η в точке $t = 1$:

$$\varphi_{\vec{\xi}}(\vec{c}) \;\stackrel{\text{df}}{=}\; \mathbf{E}\,e^{\mathrm{i}\vec{c}^{\,\flat}\,\vec{\xi}} \;=\; \mathbf{E}\,e^{\mathrm{i}\cdot 1\cdot\eta} \;=\; \varphi_\eta(1) \;=\; e^{\mathrm{i}\vec{c}^{\,\flat}\,\vec{\mu} - \frac{1}{2}\,\vec{c}^{\,\flat}\,\mathbf{\Sigma}\,\vec{c}}.$$

То есть при $\forall\,\vec{c}$ функция $\varphi_{\vec{\xi}}(\vec{c})$ совпадает с хар.ф. $\mathcal{N}_k(\vec{\mu},\mathbf{\Sigma})$. ⇚

✧ Частные распределения. Из леммы 194 с помощью подбора матрицы B легко вывести, что любой вектор, составленный из компонент нормального вектора, также имеет нормальное распределение.

196| **Лемма.** (?!) *Если* $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma})$, *то*

(a) *любая компонента* $\vec{\xi}$ *имеет одномерное нормальное распределение:* $\xi_j \backsim \mathcal{N}_1(\mu_j, \sigma_j^2)$, *где* $\sigma_j^2 = \lambda_{jj}$ — j*-й диагональный элемент матрицы* $\mathbf{\Sigma}$;

(b) *вектор* $(\xi_{j_1}, \ldots, \xi_{j_m})$, *где* $\langle j_1, \ldots, j_m \rangle$, $m \leqslant k$, — *произвольная выборка из набора индексов* $\overline{1,k}$, *имеет* m*-мерное нормальное распределение с соответствующими параметрами.*

⚠ Вектор, составленный из независимых нормальных компонент, имеет нормальное распределение. В общем случае для нормальности распределения вектора недостаточно, чтобы каждая компонента была нормальна. Простой пример многомерного распределения, отличного от нормального, у которого любая (а не только одномерная) частная ф.р. нормальна, даёт функция

$$F(x, y, z) \;=\; \Phi(x)\Phi(y)\Phi(z)\big(1 - \alpha\Phi(-x)\Phi(-y)\Phi(-z)\big),$$

где $\alpha \in (0; 1)$, Φ — ф.р. нормального $\mathcal{N}(0, 1)$ закона.

Обсудим вопрос независимости компонент нормального вектора. Пусть вектор $\vec{\xi} = (\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_r)$ состоит из подвекторов $\vec{\alpha} = (\alpha_1, \ldots, \alpha_m)$ и $\vec{\beta} = (\beta_1, \ldots, \beta_r)$. Если векторы $\vec{\alpha}, \vec{\beta}$ независимы между собой, то компоненты $\vec{\alpha}$ не коррелируют с компонентами $\vec{\beta}$ и матрица ковариаций $\vec{\xi}$ имеет блочный вид:

$$\mathbf{\Sigma} \;=\; \mathbf{Cov}(\vec{\xi}) \;=\; \begin{pmatrix} \mathbf{\Sigma}_\alpha & \mathbf{0} \\ \mathbf{0}^\flat & \mathbf{\Sigma}_\beta \end{pmatrix} \tag{4}$$

с нулевыми матрицами соответствующих размеров ($m \times r$ и $r \times m$) и матрицами $\mathbf{\Sigma}_\alpha = \mathbf{Cov}(\vec{\alpha})$, $\mathbf{\Sigma}_\beta = \mathbf{Cov}(\vec{\beta})$.

197| **Теорема.** *Пусть* $\vec{\xi} = (\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_r) \backsim \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma})$, $k = r + m$, *тогда подвекторы* $\vec{\alpha}$, $\vec{\beta}$ *независимы между собой т. т. т. когда компоненты вектора* $\vec{\alpha}$ *попарно не коррелируют с компонентами вектора* $\vec{\beta}$.

⇛ Представим числовой вектор $\vec{t}$ в виде $\vec{t} = (u_1, \ldots, u_m, b_1, \ldots, b_r)$. Как для независимых векторов $\vec{\alpha}$, $\vec{\beta}$, так и в случае попарной некоррелированности их компонент, матрица ковариаций вектора $\vec{\xi}$ имеет вид (4). Для такой матрицы квадратичная форма от $\vec{t}$ равна сумме соответствующих квадратичных форм (от $\vec{u}$ и $\vec{b}$) : $\vec{t}^\flat \mathbf{\Sigma}\, \vec{t} = \vec{u}^\flat \mathbf{\Sigma}_\alpha\, \vec{u} + \vec{b}^\flat \mathbf{\Sigma}_\beta\, \vec{b}$. Поэтому хар.ф. совместного вектора $\vec{\xi}$ может быть представлена в виде произведения хар.ф. подвекторов:

$$\varphi_{\vec{\xi}}(\vec{t}) \;=\; \exp\Big\{\mathrm{i}\,\vec{u}^\flat\,\vec{\mu}_\alpha - \frac{1}{2}\,\vec{u}^\flat\,\mathbf{\Sigma}_\alpha \vec{u}\Big\} \times \exp\Big\{\mathrm{i}\,\vec{b}^\flat\,\vec{\mu}_\beta - \frac{1}{2}\,\vec{b}^\flat\,\mathbf{\Sigma}_\beta \vec{b}\Big\}.$$

Заметим, что если рассмотреть произвольный k-мерный вектор $\vec{\eta} = (\vec{\gamma}, \vec{\zeta})$, у которого подвектор $\vec{\gamma} \backsim \vec{\alpha}$, подвектор $\vec{\zeta} \backsim \vec{\beta}$ и эти подвекторы независимы между собой, то правая часть последнего тождества будет совпадать с хар.ф. $\vec{\eta}$. Ввиду взаимной однозначности соответствия между хар.ф. и распределением сл.векторов отсюда следует независимость $\vec{\alpha}$ и $\vec{\beta}$. ⇚

Очевидно, это свойство можно распространить на любое количество подвекторов. В частности, можно утверждать, что независимость всех компонент нормального вектора эквивалентна их попарной некоррелированности.

198| **Следствие.** *Если* $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma})$, *то компоненты вектора* $(\xi_1, \ldots, \xi_k)$ *независимы в совокупности т. т. т. когда они попарно не коррелируют:* $\mathbf{Corr}(\xi_j, \xi_l) = 0, \quad j \neq l, \quad j, l = \overline{1, k}$.

§4. О преобразованиях нормального вектора

✧ Распределение суммы. Одно из важнейших свойств нормальной модели — сохранение нормальности распределения суммы сл.величин.

199| **Теорема.** I) *Если* $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma})$ *и* $\mathfrak{m} = \sum_1^k \mu_i,\; \tau^2 = \sum_{lj} \sigma_{lj}$, *то*

$$S \;:=\; \sum\nolimits_1^k \xi_j \backsim \mathcal{N}_1(\mathfrak{m}, \tau^2).$$

II) *Если* $\xi_1, \ldots, \xi_k$ *— независимые сл.в., распределённые одинаково нормально* $\mathcal{N}_1(\mu,\, \sigma^2)$, *то среднее арифметическое*

$$\frac{1}{k} \sum\nolimits_1^k \xi_j \;\backsim\; \mathcal{N}_1\Big(\mu, \frac{\sigma^2}{k}\Big).$$

⇛ Утверждение I) есть частный случай теоремы 195 с вектором $\vec{c} = (1, \ldots, 1)$. Дисперсию суммы сл.в. проще найти непосредственно:

$$\mathbf{E}(S - \mathfrak{m})^2 \;=\; \mathbf{E}\Big(\sum_{j=1}^k (\xi_j - \mu_j)\Big)^2 \;=\; \sum_{l,j=1}^k \mathbf{E}(\xi_l - \mu_l)(\xi_j - \mu_j) \;=\; \sum_{l,j=1}^k \sigma_{lj}.$$

Утверждение II) легко следует из I). $\Lleftarrow$

Воспользовавшись этой теоремой и теоремой 195, можно доказать аналогичное утверждение о сумме сл.векторов.

200 **Теорема.** *Если* $\vec{\xi}_j \backsim \mathcal{N}_k(\vec{\mu}_j, \mathbf{\Sigma}_j)$, $j = \overline{1, n}$, *— независимые k-мерные сл.векторы,* $\vec{\mathfrak{m}} = \sum_1^n \vec{\mu}_j$, $\mathfrak{S} = \sum_1^n \mathbf{\Sigma}_j$, *то*

$$\vec{S}_n = \sum\nolimits_1^n \vec{\xi}_j \backsim \mathcal{N}_k(\vec{\mathfrak{m}}, \mathfrak{S}).$$

$\Rrightarrow$ Пусть вектор $\vec{c} \in \mathbb{R}^k$. Тогда, по теореме 195, независимые сл.в. $\vec{c}^{\flat}\, \vec{\xi}_j \backsim \mathcal{N}_1(m_j, s_j^2)$ с $m_j = \vec{c}^{\flat}\, \vec{\mu}_j$, $s_j^2 = \vec{c}^{\flat}\, \mathbf{\Sigma}_j\, \vec{c}$, $j = \overline{1, n}$. Из предыдущей теоремы следует, что $\vec{c}^{\flat}\, \vec{S}_n \backsim \mathcal{N}_1(\vec{c}^{\flat}\, \vec{\mathfrak{m}}, \vec{c}^{\flat}\, \mathfrak{S}\, \vec{c})$ с заявленными в теореме параметрами, что завершает доказательство в силу теоремы 195. $\Lleftarrow$

► Ранее мы установили несколько фактов относительно преобразований нормальных сл.в. Пусть $\xi_1, \ldots, \xi_k$ — независимые нормальные $\mathcal{N}(0, 1)$ сл.в. (список вероятностных распределений см. на стр. 107), тогда:

- ✓ $\xi_1^2 + \ldots + \xi_k^2 \backsim \mathcal{G}\mathtt{am}\big(\frac{k}{2}, 2\big) \backsim \mathcal{K}\mathtt{hi}(k)$ (106, стр. 115, 171, стр. 167);
- ✓ $\xi_1^2 + \xi_2^2 \backsim \mathcal{E}\mathtt{x}(2)$ (171, стр. 167);
- ✓ $\dfrac{\xi_k}{\sqrt{\xi_1^2 + \ldots + \xi_{k-1}^2}} \sqrt{k-1} \backsim \mathcal{S}\mathtt{tud}(k-1)$ (стр. 123);
- ✓ $\dfrac{\xi_1}{\xi_2} \backsim \mathcal{C}\mathtt{auch}(0, 1)$ (130, стр. 125);
- ✓ e^{ξ_1} — логнормальная сл.в. (стр. 132).

Кроме того, здесь представляют большой интерес различные способы получения нормальных случайных чисел. Простой способ генерирования случайного вектора из набора независимых стандартных нормальных $\mathcal{N}(0, 1)$ сл.в. предоставляет первая часть леммы 188, стр. 182. Поэтому, в принципе, достаточно только уметь генерировать последовательность независимых сл.в. Универсальный способ здесь состоит в использовании утверждения 41, стр. 54. Для этого, правда, необходимо иметь хороший алгоритм вычисления обратной функции Φ^{-1} к стандартной нормальной ф.р. При отсутствии такового этот способ может потребовать слишком много машинного времени.

В примере 108, стр. 115, приведено следующее утверждение:

$$\left.\begin{array}{l}\theta_1, \theta_2 \backsim \mathcal{U}\mathrm{n}(0,1) \quad \text{— независимые сл.в.,}\\ \xi_1 = \sqrt{-2\ln\theta_1}\,\sin(2\pi\theta_2),\\ \xi_2 = \sqrt{-2\ln\theta_1}\,\cos(2\pi\theta_2),\end{array}\right\} \Rightarrow \quad (\xi_1, \xi_2) \backsim \mathcal{N}_2(\vec{0}, \dot{\mathbb{I}}).$$

Заметим, что этот способ тоже не самый быстрый.

Наконец, последний из способов, который мы здесь рассмотрим, использует центральную предельную теорему 325, стр. 275. Пусть $\left\langle \eta_k \backsim \mathcal{U}\mathrm{n}(-\sqrt{3}, \sqrt{3}) \right\rangle_1^n$ — последовательность независимых равномерных сл.в. Поскольку математическое ожидание $\mathbf{E}\xi_1 = 0$, а дисперсия $\mathbf{D}\xi_1 = 1$, то в соответствии с ЦПТ

$$\frac{1}{\sqrt{n}} \sum\nolimits_1^n \xi_k \rightsquigarrow \mathcal{N}(0,1),$$

т.е. имеет приближённо нормальное распределение. При $n = 9$, а лучше $n = 16$ этот способ даёт вполне пригодный для практики результат при скорости генерации в 1,5 раза выше, чем предыдущий способ.

Заметим, что профессиональные датчики нормальных чисел устроены совсем на других принципах и их скорость генерации в десятки раз выше.

201| Упр. Воспользовавшись приёмом доказательства теоремы 190, покажите, что если $\vec{\xi} \backsim \mathcal{N}_k(\vec{\mu}, \mathbf{\Sigma})$ с невырожденной матрицей ковариаций, то квадратичная форма $(\vec{\xi} - \vec{\mu})^\flat \mathbf{\Sigma}^{-1} (\vec{\xi} - \vec{\mu}) \backsim \mathcal{K}\mathrm{hi}(k)$.

§ 5. Условное распределение

Найдём условное распределение одной компоненты нормального вектора при фиксированных значениях остальных компонент. Рассмотрим сл.вектор $\vec{\theta} = (\zeta, \xi_1, \ldots, \xi_k)$. Матрицу ковариаций $\vec{\theta}$ представим в блочном виде:

$$\mathbf{Cov}(\vec{\theta}) = \begin{pmatrix} \tau^2 & \vec{\psi}^\flat \\ \vec{\psi} & \mathbf{\Sigma} \end{pmatrix}, \tag{5}$$

где $\tau^2 = \mathbf{D}\zeta$ — дисперсия ζ, $\vec{\psi}^\flat = \big(\mathbf{Cov}(\zeta, \xi_1), \ldots, \mathbf{Cov}(\zeta, \xi_k)\big)$ — вектор коэффициентов ковариации ζ со всеми компонентами $\vec{\xi}$, $\mathbf{\Sigma} = \mathbf{Cov}(\vec{\xi})$ — матрица ковариаций сл.вектора $\vec{\xi}$. По правилам обращения с блочными матрицами определитель $\big|\mathbf{Cov}(\vec{\theta})\big| = \big(\tau^2 - \vec{\psi}^\flat \mathbf{\Sigma}^{-1} \vec{\psi}\big)\,|\mathbf{\Sigma}|$.

202 **Теорема.** *Пусть сл.вектор* $\vec{\theta} = (\zeta, \xi_1, \ldots, \xi_k)$ *имеет* $(k+1)$*-мерное нормальное распределение со средними* $m = \mathbf{E}\,\zeta$, $\vec{\mu} = \mathbf{E}\,\vec{\xi}$ *и невырожденной матрицей ковариаций* (5). *Тогда при фиксированных значениях* $\vec{\xi} = \vec{x}$ *условное распределение* ζ *нормально* $\mathcal{N}_1\big(\widehat{\mu}\,(\vec{x}), \mathfrak{s}_k^2\big)$ *с параметрами*

$$\widehat{\mu}\,(\vec{x}) \;=\; m + \vec{\psi}^{\flat}\,\boldsymbol{\Sigma}^{-1}(\vec{x} - \vec{\mu}),$$

$$\mathfrak{s}_k^2 \;=\; \tau^2 - \vec{\psi}^{\flat}\,\boldsymbol{\Sigma}^{-1}\vec{\psi} \quad (> 0).$$

В частности, для одномерного вектора $\vec{\xi} = \xi_1$ *и* $\rho = \mathbf{Corr}(\zeta, \xi_1)$

$$\widehat{\mu}\,(x_1) \;=\; m + \rho\frac{\tau}{\sqrt{\mathbf{D}\xi_1}}(x_1 - \mu_1), \quad \mathfrak{s}_1^2 \;=\; \tau^2(1 - \rho^2).$$

$\Rrightarrow$ Сначала заметим, что по условию теоремы $\mathfrak{s}_k^2\,|\boldsymbol{\Sigma}| \neq 0$, поэтому заявленное условное распределение невырождено и определена функция $\widehat{\mu}$, которая совпадает с функцией линейной регрессии ζ на $\vec{\xi}$ (см. 78, стр. 78). Далее, сл.в. $\eta = \zeta - \widehat{\mu}\,(\vec{\xi})$ не коррелирует со всеми компонентами $\vec{\xi}$:

$$\mathbf{Cov}(\eta, \vec{\xi}^{\flat}) \;=\; \mathbf{E}\Big[\Big(\zeta - m - \vec{\psi}^{\flat}\boldsymbol{\Sigma}^{-1}(\vec{\xi} - \vec{\mu})\Big)(\vec{\xi} - \vec{\mu})^{\flat}\Big] \;=$$

$$=\; \big(\mathbf{Cov}(\vec{\xi}, \zeta)\big)^{\flat} - \vec{\psi}^{\flat}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma} \;=\; \vec{\psi}^{\flat} - \vec{\psi}^{\flat} \;=\; \vec{0}^{\flat}.$$

Дисперсия этой сл.в.

$$\mathbf{D}\eta \;=\; \mathbf{E}\Big[\big(\zeta - m - \vec{\psi}^{\flat}\boldsymbol{\Sigma}^{-1}(\vec{\xi} - \vec{\mu})\big)\big(\zeta - m - \vec{\psi}^{\flat}\boldsymbol{\Sigma}^{-1}(\vec{\xi} - \vec{\mu})\big)^{\flat}\Big] \;=$$

$$=\; \mathbf{E}\Big[(\zeta - m)^2 - \vec{\psi}^{\flat}\boldsymbol{\Sigma}^{-1}(\vec{\xi} - \vec{\mu})(\zeta - m) - (\zeta - m)(\vec{\xi} - \vec{\mu})^{\flat}\boldsymbol{\Sigma}^{-1}\vec{\psi} +$$

$$+\, \vec{\psi}^{\flat}\boldsymbol{\Sigma}^{-1}(\vec{\xi} - \vec{\mu})(\vec{\xi} - \vec{\mu})^{\flat}\boldsymbol{\Sigma}^{-1}\vec{\psi}\Big] \;=$$

$$=\; \mathbf{D}\zeta - 2\vec{\psi}^{\flat}\boldsymbol{\Sigma}^{-1}\vec{\psi} + \vec{\psi}^{\flat}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}\vec{\psi} \;=\; \mathfrak{s}_k^2.$$

Вектор $(\eta, \xi_1, \ldots, \xi_k)$ представляет собой невырожденное линейное преобразование сл.вектора $\vec{\theta}$. По теореме о линейных преобразованиях нормального вектора отсюда следует, во-первых, что сл.в. $\eta \backsim \mathcal{N}(0, \mathfrak{s}_k^2)$ и, во-вторых, η не зависит от $\vec{\xi}$. Таким образом,

$$\mathbf{P}\big\{\zeta \leqslant z \,|\, \vec{\xi} = \vec{x}\big\} = \mathbf{P}\big\{\eta \leqslant z - \widehat{\mu}\,(\vec{\xi}) \,|\, \vec{\xi} = \vec{x}\big\} \overset{(*)}{=} \mathbf{P}\{\eta \leqslant z - \widehat{\mu}\,(\vec{x})\} =$$

$$= \Phi\Big(\frac{z - \widehat{\mu}\,(\vec{x})}{\mathfrak{s}_k}\Big),$$

где Φ — ф.р. стандартного нормального закона и равенство $\stackrel{(*)}{=}$ относительно условной вероятности справедливо для независимых сл.векторов η и $\vec{\xi}$ (теорема 149, стр. 150). Что и требовалось. $\Lleftarrow$

203| Упр. Пусть вектор (ζ, ξ_1, ξ_2) имеет невырожденное нормальное распределение, причём сл.в. ξ_1 и ξ_2 независимы. Показать, что усл.м.о.

$$\mathbf{E}\big[\,\zeta \mid \xi_1, \xi_2\,\big] = \mathbf{E}\big[\,\zeta \mid \xi_1\,\big] + \mathbf{E}\big[\,\zeta \mid \xi_2\,\big] - \mathbf{E}\,\zeta.$$

A. Указания к решению задач

✦ 191, стр. 185. а) $Q = \frac{1}{2}\begin{pmatrix} \sqrt{1+\rho}+\sqrt{1-\rho} & \sqrt{1+\rho}-\sqrt{1-\rho} \\ \sqrt{1+\rho}-\sqrt{1-\rho} & \sqrt{1+\rho}+\sqrt{1-\rho} \end{pmatrix}$ или $Q = \begin{pmatrix} \sqrt{1-\rho^2} & \rho \\ 0 & 1 \end{pmatrix}$;
б) $\widehat{q}_{ij} = \sigma_i q_{ij}$.

✦ 193, стр. 186. Функция (3) есть хар.ф. вектора $Q\vec{\eta} + \vec{\mu}$, где $\vec{\eta} \backsim \mathcal{N}_k(\vec{0}, \dot{\mathbb{I}})$, $Q = \sqrt{\mathbf{\Sigma}}$.

✦ 194, стр. 186. Хар.ф. $\vec{\eta}$ равна $\mathbf{E}\exp\{\mathrm{i}\,\vec{t}^{\flat}(B\vec{\xi} + \vec{r})\}$; ранг произведения матриц $\mathrm{rang}(B\,\mathbf{\Sigma}B^{\flat}) = \mathrm{rang}(B)$, если $|\mathbf{\Sigma}| \neq 0$.

✦ 196, стр. 187. a) $\xi_j = \vec{e}_j^{\,\flat}\,\vec{\xi}$, где $\vec{e}_j$ — единичный вектор с единицей на j-м месте; b) $(\xi_{j_1}, \ldots, \xi_{j_m}) = \Big(\vec{e}_{j_1}, \ldots, \vec{e}_{j_m}\Big)^{\flat}\vec{\xi}$.

✦ 201, стр. 190. $\vec{\eta} = \sqrt{\mathbf{\Sigma}}^{\,-1}(\vec{\xi} - \vec{\mu}) \backsim \mathcal{N}_k(\vec{0}, \dot{\mathbb{I}})$, т.е. компоненты $\eta_j \backsim \mathcal{N}(0,1)$ и независимы, следовательно, $(\vec{\xi} - \vec{\mu})^{\flat}\,\mathbf{\Sigma}^{-1}(\vec{\xi} - \vec{\mu}) = \vec{\eta}^{\,\flat}\,\vec{\eta} = \eta_1^2 + \ldots + \eta_k^2$.

✦ 203, стр. 192. Матрица $\mathbf{\Sigma}$ диагональна, поэтому (в обозначениях теоремы 202) $\widehat{\mu}(x_1, x_2) = \widehat{\mu}(x_1) + \widehat{\mu}(x_2) - m$.

VI Сходимость случайных величин

Как и в математической анализе, в теории вероятностей сходимость случайных величин играет ключевую роль. Имеется несколько видов сходимости, необходимость в которых возникает в тех или иных ситуациях. Если акцентировать внимание на том, что случайные величины суть функции элементарных исходов, то естественным представляется определение сходимости, называемое в анализе поточечной сходимостью. С другой стороны, на практике чаще представляют интерес вопросы о степени «стохастической» близости сл.в., например: насколько велика вероятность их малых различий и можно ли утверждать, что вероятности осуществления одних и тех же событий для этих величин близки между собой. Вопросы первого типа приводят к понятию сходимости по вероятности, второго — к слабой сходимости.

§ 1. Сходимость с вероятностью единица

Пусть $\xi_n : \Omega \mapsto \mathbb{R}^1$, $n \geqslant 0$, — сл.в., заданные на одном вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$.

Определение. Последовательность $\langle \xi_n \rangle_1^\infty$ *сходится почти наверное (п.н.)* или *с вероятностью единица* к сл.в. ξ_0, если

$$\mathbf{P}\{\, \omega : \xi_n(\omega) \not\to \xi_0(\omega) \,\} = 0.$$

Обозначается как $\xi_n \xrightarrow{\text{п.н.}} \xi_0$ или $\xi_n \to \xi_0$ $[\mathbf{P}\text{-п.н.}]$, если нужно выделить меру, относительно которой рассматривается сходимость.

Другими словами, множество исходов ω, для которых $\xi_n(\omega) \to \xi_0(\omega)$, обладает единичной вероятностью, или в вольной интерпретации: при любых обстоятельствах «фортуны» (наверняка) значение сл.в. ξ_n близкó к значению ξ_0. В анализе эквивалентно понятию сходимости почти всюду (п.в.) по мере.

Соотношение $\xi_n(\omega) \not\to \xi_0(\omega)$ для данного ω означает, что события $B_n = \{|\xi_n(\omega) - \xi_0(\omega)| > \varepsilon\}$ с некоторым $\varepsilon > 0$ происходят при $\forall\, n = n_k,\ k \geqslant 1$, для некоторой возрастающей до $+\infty$ последовательности натуральных чисел $\langle n_k \rangle_1^\infty$, или иначе: произойдёт *бесконечное число* (*б.ч.*) событий B_n при изменении n от 1 до $+\infty$, кратко $\{B_n \text{ б.ч.}\}$ (бесконечно часто, см. стр. 520). Для анализа таких последовательностей событий полезна следующая

204| **Лемма.** [*Э. Борель, Г. Кантелли.*] (а) *Если для последовательности событий* $\langle A_n \rangle_1^\infty$ *ряд* $\sum_1^\infty \mathbf{P}\{A_n\} < \infty$, *то* $\mathbf{P}\{A_n \text{ б.ч.}\} = 0$.

(б) *Обратно, если ряд* $\sum_1^\infty \mathbf{P}\{A_n\} = \infty$ *и события* $\langle A_n \rangle_1^\infty$ *независимы в совокупности, то* $\mathbf{P}\{A_n \text{ б.ч.}\} = 1$.

⊨ (а) Событие $\{A_n \text{ б.ч.}\} = \varlimsup\limits_n A_n = \bigcap\limits_{n=1}^\infty \bigcup\limits_{k=n}^\infty A_k = \lim\limits_n{}_{\downarrow} \bigcup\limits_{k=n}^\infty A_k$, поэтому

$$\mathbf{P}\{A_n \text{ б.ч.}\} \;=\; \lim_n \mathbf{P}\Big\{\bigcup_{k=n}^\infty A_k\Big\} \;\leqslant\; \lim_n \sum_{k=n}^\infty \mathbf{P}\{A_k\} \;=\; 0.$$

(б) Применив известное неравенство $1 - x \leqslant e^{-x}$, получим для $\forall\, m < n$ в силу независимости событий (и их дополнений)

$$\mathbf{P}\Big\{\bigcap_{k=m}^n A_k^c\Big\} \;=\; \prod_{k=m}^n (1 - \mathbf{P}\{A_k\}) \;\leqslant\; \exp\Big(-\sum_{k=m}^n \mathbf{P}\{A_k\}\Big) \;\underset{n}{\to}\; 0,$$

т.е. $1 - \mathbf{P}\{A_n \text{ б.ч.}\} = \lim_m \lim_n \mathbf{P}\big\{\bigcap_{k=m}^n A_k^c\big\} = 0$. ⫤

205| **Теорема.** *Пусть* $\langle \xi_n \rangle_1^\infty$ *— действительные сл.в. Для того чтобы*

(I) *существовала сл.в.* ξ_0 *такая, что* $\xi_n \xrightarrow{\text{п.н.}} \xi_0$, *достаточно*

$$\sum\nolimits_1^\infty \mathbf{P}\{|\xi_{n+1} - \xi_n| > \varepsilon_n\} \;<\; \infty$$

для некоторой числовой последовательности $\varepsilon_n > 0$ *такой, что* $\sum_n \varepsilon_n < \infty$;

(II) $\xi_n \xrightarrow{\text{п.н.}} \xi_0$ *с некоторой сл.в.* ξ_0, *достаточно*

$$\sum\nolimits_1^\infty \mathbf{P}\{|\xi_n - \xi_0| > \varepsilon_n\} \;<\; \infty$$

a) *для какой-либо числовой последовательности* $\varepsilon_n \searrow 0$ *или*

b) *для* $\forall\, \varepsilon_n \equiv \varepsilon > 0$.

$\Rrightarrow$ (I) Положим $A_n = \{|\xi_{n+1} - \xi_n| > \varepsilon_n\}$, $n \geqslant 1$. По предположению теоремы и по лемме Бореля–Кантелли $\mathbf{P}\{\ \overline{\lim}_n A_n\ \} = 0$. Для $\forall\, \omega \notin \overline{\lim}_n A_n$, т.е. $\omega \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c$, найдётся номер $n = n(\omega)$ такой, что при всех $k > n$ модуль разности $|\xi_{k+1} - \xi_k| \leqslant \varepsilon_k$. Поскольку последовательность ε_k суммируема, для всех таких ω существует и конечен ряд $\sum_{k=1}^{\infty}(\xi_{k+1} - \xi_k)$. Утверждению теоремы удовлетворяет сл.в. $\xi_0(\omega) = 0$ при $\omega \in \overline{\lim}_n A_n$ и

$$\xi_0(\omega) \;=\; \lim_n \xi_n(\omega) \;=\; \xi_1(\omega) + \lim_n \sum_{k=1}^{n-1} (\xi_{k+1} - \xi_k), \qquad \omega \notin \overline{\lim_n}\, A_n.$$

(II) Запишем множество исходов, где ξ_n не сходится к ξ_0 :

$$\{\omega : \xi_n(\omega) \not\to \xi_0(\omega)\} \;=\; \bigcup_{k=1}^{\infty} \left\{\omega : |\xi_n(\omega) - \xi_0(\omega)| > \frac{1}{k} \ \text{ б.ч. по } n\right\}.$$

По лемме Бореля–Кантелли и условию (b) вероятность каждого события справа (для $\forall\, k \geqslant 1$) равна нулю. (II.a) следует (?!) из (II.b). $\Lleftarrow$

206| ⚠ Говорят, что последовательность сл.в. $\langle\, \xi_n \,\rangle_1^{\infty}$ сходится к ξ_0 *вполне*, если $\sum_1^{\infty} \mathbf{P}\{\, |\xi_n - \xi_0| > \varepsilon \,\} < \infty$ для $\forall\, \varepsilon > 0$. Как следует из приведённой теоремы, сходимость вполне влечёт п.н. сходимость.

207| Пример. *Сильный закон больших чисел Бореля.*

Рассмотрим бесконечную последовательность экспериментов в схеме испытаний Бернулли с вероятностью успеха $p = {}^1\!/_2$. Возможные реализации такой последовательности принадлежат пространству $\Omega = \{\omega = (x_1, x_2, \ldots) : \quad x_i = 0, 1\}$. Не вдаваясь в подробности, скажем, что на этом пространстве можно выделить класс измеримых подмножеств (σ-алгебру) и на этом классе определить вероятностную меру $\mathbf{P}$ так, чтобы вероятности любых событий, описываемых в терминах исходов на конечном числе экспериментов, совпадали с вероятностями, вычисленными по обычным формулам конечной схемы Бернулли. Например, если сл.в. $\xi_n = \xi_n(\omega)$ равна количеству успехов в первых n испытаниях, то $\mathbf{P}\{\, \xi_n = k \,\} = \mathrm{C}_n^k / 2^n$. Легко понять, что для $k < n/2$ эта вероятность возрастает по k. Поэтому для $\forall\, k < n/2$

$$\mathbf{P}\{\xi_n < k\} \;=\; \sum_{j=0}^{k-1} \mathbf{P}\{\, \xi_n = j \,\} \;\leqslant\; k\,\mathbf{P}\{\xi_n = k\} \;=\; \frac{k\mathrm{C}_n^k}{2^n}\,.$$

Применяя формулу Стирлинга, можно показать, что при $0 < \varepsilon < {}^1\!/_2$ и $k = k_n = \llbracket n(1 - 2\varepsilon)/2 \rrbracket$ правая часть предыдущего неравенства имеет порядок малости $\mathcal{O}(\sqrt{n}A(\varepsilon)^{-n/2})$, где $A(\varepsilon) = (1-2\varepsilon)^{1-2\varepsilon}(1+2\varepsilon)^{1+2\varepsilon}$ возрастает по ε и потому $A(\varepsilon) > 1$. Следовательно, для $\forall\,\varepsilon > 0$ ввиду симметрии

$$\sum\nolimits_n \mathbf{P}\Big\{ \Big|\frac{1}{n}\xi_n - \frac{1}{2}\Big| > \varepsilon \Big\} \leqslant 2\sum\nolimits_n \frac{k_n \mathrm{C}_n^{k_n}}{2^n} < \infty.$$

Аналогичные выкладки справедливы для $\forall\, p \in (0;1)$. Таким образом, согласно теореме 205, можно утверждать, что для почти любой последовательности испытаний Бернулли среднее арифметическое $\frac{1}{n}\xi_n$ числа успехов за n экспериментов сходится к вероятности успеха p. ◎

⚠ Утверждения подобного типа возможны только после формального определения единого вероятностного пространства, на котором заданы все сл.в. ξ_n — в примере это пространство всех бесконечных последовательностей бернуллиевских испытаний с соответствующей вероятностной мерой.

✧ **Переход к пределу под знаком интеграла**. По теореме Лебега об ограниченной сходимости 540, стр. 482, соотношение $\lim_n \mathbf{E}\,\xi_n = \mathbf{E}\,\xi_0$ справедливо, если $\xi_n \xrightarrow{\text{п.н.}} \xi_0$ и $|\xi_n| \leqslant \eta$ с некоторой сл.в. η, для которой $\mathbf{E}\,\eta < \infty$. Расширение этого результата на случай последовательностей сл.в., сходящихся по вероятности и слабо, будет установлено ниже. Пока же введём следующее

Определение. Последовательность сл.в. $\langle\, \xi_n \,\rangle_1^\infty$ *сходится в среднем порядка* $q > 0$ (или просто *в среднем,* если $q = 1$, или *в среднем кдадратическом,* если $q = 2$) к сл.в. ξ_0: $\xi_n \xrightarrow{\mathcal{L}_q} \xi_0$, если $\lim\limits_n \mathbf{E}\,|\xi_n - \xi_0|^q = 0$.

⚠ Пример 220, стр. 201, показывает, что сходимость в среднем, вообще говоря, не влечёт п.н. сходимость.

208| Пример. Если ξ_n равно количеству успехов в n испытаниях в схеме Бернулли с вероятностью успеха p, то при $n \to \infty$

$$\mathbf{E}\Big|\frac{\xi_n}{n} - p\Big|^2 = \frac{1}{n^2}\mathbf{D}\xi_n = \frac{1}{n}p(1-p) \to 0,$$

т.е. $\frac{1}{n}\xi_n \xrightarrow{\mathcal{L}_2} p$. Докажите, что $\frac{1}{n}\xi_n \xrightarrow{\mathcal{L}_q} p$ для $\forall\, q > 0$. ◎

209| Упр. Докажите справедливость следующих утверждений:

(а) если $\xi_n \xrightarrow{\mathcal{L}_1} \xi_0$ и $\mathbf{E}\,|\xi_0| < \infty$, то $\lim_n \mathbf{E}\,\xi_n = \mathbf{E}\,\xi_0$ и $\lim_n \mathbf{E}\,|\xi_n| = \mathbf{E}\,|\xi_0|$;

(б) сходимость $\xi_n \xrightarrow{\mathcal{L}_q} \xi_0$ влечёт $\xi_n \xrightarrow{\mathcal{L}_t} \xi_0$ при $0 < t \leqslant q$.

§ 2. Сходимость по вероятности

Определение. Последовательность сл.в. $\langle \xi_n \rangle_1^\infty$, определённых на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$, *сходится по вероятности* к сл.в. ξ_0, определённой там же, если для $\forall \varepsilon > 0$

$$\lim_n \mathbf{P}\left\{ \left|\xi_n - \xi_0\right| > \varepsilon \right\} = 0.$$

Обозначается $\xi_n \xrightarrow{\mathbf{P}} \xi_0$, или $\mathbf{P}\text{-}\lim\limits_n \xi_n = \xi_0$, или $\xi_n \xrightarrow[n]{\mathbf{P}} \xi_0$, если нужно выделить индекс, по которому осуществляется переход к ∞.

⚠ В анализе называется сходимостью по мере.

210| Упр. Пусть $\xi_n \xrightarrow{\mathbf{P}} \zeta$ и $\xi_n \xrightarrow{\mathbf{P}} \eta$. Покажите, что $\mathbf{P}\{\eta \neq \zeta\} = 0$.

Сходимость по вероятности сохраняется при алгебраических операциях над элементами последовательностей.

211| Теорема. *Пусть $\xi_n \xrightarrow{\mathbf{P}} \xi_0$, $\eta_n \xrightarrow{\mathbf{P}} \eta_0$, $\varepsilon_n \xrightarrow{\mathbf{P}} 0$, причём все сл.в. заданы на одном пространстве. Тогда при $n \to \infty$ справедливо:*

(a) $\xi_n + \eta_n \xrightarrow{\mathbf{P}} \xi_0 + \eta_0$; *в частности* $\xi_n + \varepsilon_n \xrightarrow{\mathbf{P}} \xi_0$;

(b) $\xi_n \eta_n \xrightarrow{\mathbf{P}} \xi_0 \eta_0$; *в частности* $\xi_n \varepsilon_n \xrightarrow{\mathbf{P}} 0$ *и* $\xi_n(1 + \varepsilon_n) \xrightarrow{\mathbf{P}} \xi_0$;

(c) $\dfrac{\xi_n}{\eta_n} \xrightarrow{\mathbf{P}} \dfrac{\xi_0}{\eta_0}$, *если* $\mathbf{P}\{\eta_n = 0\} = 0$ *для всех* $n \geqslant 0$.

⊫ Все эти утверждения являются следствиями приводимой ниже теоремы 218 о непрерывном преобразовании сходящихся последовательностей. Мы всё же докажем утверждение пункта (b), дабы продемонстрировать возникающие здесь сложности. Нам понадобятся пять простых фактов:

c_1 : неравенство треугольника: $|x + y| \leqslant |x| + |y|$;

c_2 : если сумма $x + y > u + v$, то либо $x > u$, либо $y > v$;

c_3 : если произведение $xy > uv$ и $x, y > 0$, то либо $x > u$, либо $y > v$;

c_4 : вероятность объединения событий (связанных союзом *или*) не больше суммы вероятностей этих событий;

c_5 : если $a_n \leqslant b_n$ для $\forall n$, то $\overline{\lim\limits_n}\, a_n \leqslant \overline{\lim\limits_n}\, b_n$, $\quad \overline{\lim\limits_n}\, a_n \leqslant \sup\limits_n b_n$.

Обращение (по номеру) к тому или иному факту в следующей цепочке неравенств отмечается посредством соответствующего значка.

Пусть $\varepsilon > 0$, тогда для $\forall M > 0$

$$\mathbf{P}\left\{\left|\xi_n \eta_n - \xi_0 \eta_0\right| > \varepsilon\right\} = \mathbf{P}\left\{ \left|(\xi_n - \xi_0)\eta_n + (\eta_n - \eta_0)\xi_0\right| > \varepsilon \right\} \underset{(1)}{\leqslant}$$

$$\leqslant \mathbf{P}\left\{\left|(\xi_n - \xi_0)\,\eta_n\right| + \left|(\eta_n - \eta_0)\xi_0\right| > \varepsilon\right\} \underset{(2,4)}{\leqslant}$$

$$\leqslant \mathbf{P}\left\{\left|(\xi_n - \xi_0)\right|\left|\eta_n\right| > \tfrac{\varepsilon}{2}\right\} + \mathbf{P}\left\{ \left|(\eta_n - \eta_0)\right|\left|\xi_0\right| > \tfrac{\varepsilon}{2} \right\} \underset{(3,4)}{\leqslant}$$

$$\leqslant \mathbf{P}\left\{\left|\xi_n - \xi_0\right| > \tfrac{\varepsilon}{2M}\right\} + \mathbf{P}\left\{\left|\eta_n\right| > M\right\} +$$

$$+ \mathbf{P}\left\{\left|\eta_n - \eta_0\right| > \tfrac{\varepsilon}{2M}\right\} + \mathbf{P}\left\{\left|\xi_0\right| > M\right\}.$$

Первое и третье слагаемые в последнем выражении стремятся к нулю (при $n \to \infty$) по условию теоремы. Поэтому для $\forall \varepsilon,\ M > 0$

$$\varlimsup_{n} \mathbf{P}\left\{\left|\xi_n \eta_n - \xi_0 \eta_0\right| > \varepsilon\right\} \underset{(5)}{\leqslant} \sup_n \mathbf{P}\left\{\left|\eta_n\right| > M\right\} + \mathbf{P}\left\{\left|\xi_0\right| > M\right\}.$$

Доказательство пункта (b) завершается теперь простой ссылкой на приводимую далее теорему 212 . ⇚

Из курса анализа известно, что любая сходящаяся числовая последовательность ограничена. Близкий по духу факт справедлив и для сходящихся последовательностей сл.в., с той лишь разницей, что ограниченность понимается как стремление к нулю максимальной вероятности принятия больших значений — вероятностная «масса» сл.величин не «уходит на бесконечность».

212| **Теорема.** [*О равномерной плотности.*] *Если последовательность $\zeta_n \xrightarrow{\mathbf{P}} \zeta_0$, то она равномерно плотна (ограничена по вероятности):*

$$\lim_{M\to\infty} \sup_n \mathbf{P}\left\{\left|\zeta_n\right| > M\right\} = 0.$$

⇛ Так как $\zeta_n \xrightarrow{\mathbf{P}} \zeta_0$, то для данного $\varepsilon > 0$ найдётся такое $N > 0$, что для $\forall n > N$ вероятность $\mathbf{P}\left\{\left|\zeta_n - \zeta_0\right| > 1\right\} \leqslant \varepsilon$. С другой стороны, для каждой сл.в. ζ_k можно подобрать число $M_k > 0$ такое, что $\mathbf{P}\left\{\left|\zeta_k\right| > M_k\right\} < \varepsilon$.

Выберем $M = \max\left\langle M_k,\ k = 0, 1, \ldots, N\right\rangle + 1$. Тогда для $0 \leqslant n \leqslant N$ вероятность $\mathbf{P}\left\{\left|\zeta_n\right| > M - 1\right\} \leqslant \varepsilon$. Поскольку $\left|\zeta_n\right| \leqslant \left|\zeta_n - \zeta_0\right| + \left|\zeta_0\right|$, то

$$\mathbf{P}\left\{\left|\zeta_n\right| > M\right\} \leqslant \mathbf{P}\left\{\left|\zeta_n - \zeta_0\right| + \left|\zeta_0\right| > 1 + M - 1\right\} \leqslant 2\varepsilon$$

для $\forall n > N$ (снова в силу фактов c_2, c_4). Таким образом, для $\forall \varepsilon > 0$ можно подобрать число $M > 0$ так, чтобы $\sup\limits_n \mathbf{P}\left\{\left|\zeta_n\right| > M\right\} \leqslant 2\varepsilon$. Так как с ростом

M искомая вероятность не увеличивается, то это доказывает теорему. $\Lleftarrow$

213| Упр. Докажите обобщение (а): $a_n\xi_n + b_n\eta_n + c_n \xrightarrow{\mathbf{P}} a\xi_0 + b\eta_0 + c$ для любых последовательностей действительных чисел $a_n \to a, b_n \to b, c_n \to c$, если $|a|, |b|, |c| < \infty$.

214| Упр. Докажите, что если $\eta_n \xrightarrow{\mathbf{P}} 1$ и $\xi_n \xrightarrow{\mathbf{P}} +\infty$, т.е.

$$\lim_n \mathbf{P}\{\xi_n > M\} = 1 \quad \text{для } \forall\, M > 0,$$

то а) $\xi_n \pm \eta_n \xrightarrow{\mathbf{P}} +\infty$, б) $\xi_n\eta_n \xrightarrow{\mathbf{P}} +\infty$.

✧ Стохастическая O-символика. В анализе символы «o» («о»-малое) и «$\mathcal{O}$» («о»-большое) используются для обозначения сходящихся к нулю и асимптотически ограниченных последовательностей (функций). В теории вероятностей аналогичная символика применяется для обозначения сходящихся по вероятности к нулю и ограниченных по вероятности последовательностей сл.в. (читаются «о»-малое по вероятности и «о»-большое по вероятности).

Определения.

$$\begin{aligned}
\xi_n &= o_{\mathbf{P}}(1) &&\Leftrightarrow\ \xi_n \xrightarrow{\mathbf{P}} 0;\\
\xi_n &= \mathcal{O}_{\mathbf{P}}(1) &&\Leftrightarrow\ \text{семейство } \big\langle\, \xi_n \big\rangle_1^\infty \text{ ограничено по вероятности;}\\
\xi_n &= o_{\mathbf{P}}(\eta_n) &&\Leftrightarrow\ \xi_n = \eta_n\zeta_n, \quad \text{где} \quad \zeta_n \xrightarrow{\mathbf{P}} 0;\\
\xi_n &= \mathcal{O}_{\mathbf{P}}(\eta_n) &&\Leftrightarrow\ \xi_n = \eta_n\zeta_n, \quad \text{где} \quad \zeta_n = \mathcal{O}_{\mathbf{P}}(1).
\end{aligned}$$

215| **Лемма.** *Пусть* $\xi_n \xrightarrow[n]{\mathbf{P}} 0$ *и* $R(x) = o(1),\ x \to 0$. *Тогда* $R(\xi_n) = o_{\mathbf{P}}(1)$.

$\Rrightarrow$ Пусть $\varepsilon > 0$. По определению $o(1)$ найдётся $\delta > 0$ такое, что $\big|R(x)\big| \leqslant \varepsilon$, как только $|x| \leqslant \delta$. Следовательно, событие $\big|R(\xi_n)\big| > \varepsilon$ влечёт $\big|\xi_n\big| > \delta$, поэтому

$$\overline{\lim_n}\, \mathbf{P}\big\{\big|R(\xi_n)\big| > \varepsilon\big\} \ \leqslant\ \lim\nolimits_n \mathbf{P}\big\{\big|\xi_n\big| > \delta\big\} \ =\ 0.$$ $\Lleftarrow$

Многие правила, присущие классическим «о»-символам, имеют место и для их стохастических аналогов.

216| **Лемма.** *Справедливы следующие соотношения:*

(а) $o_{\mathbf{P}}(1) + o_{\mathbf{P}}(1) = o_{\mathbf{P}}(1)$, (б) $\mathcal{O}_{\mathbf{P}}(1) + \mathcal{O}_{\mathbf{P}}(1) = \mathcal{O}_{\mathbf{P}}(1)$,
(в) $\mathcal{O}_{\mathbf{P}}(1) + o_{\mathbf{P}}(1) = \mathcal{O}_{\mathbf{P}}(1)$, (г) $\mathcal{O}_{\mathbf{P}}(1)o_{\mathbf{P}}(1) = o_{\mathbf{P}}(1)$,
(д) $\eta_n o_{\mathbf{P}}(1) = o_{\mathbf{P}}(\eta_n)$, (е) $\eta_n\mathcal{O}_{\mathbf{P}}(1) = \mathcal{O}_{\mathbf{P}}(\eta_n)$.

$\Rrightarrow$ Равенства (д), (е) суть определения $o_{\mathbf{P}}(\eta_n)$ и $\mathcal{O}_{\mathbf{P}}(\eta_n)$. Остальные соотношения легко выводятся с помощью фактов $c_1 - c_5$, стр. 197. Например, если $\xi_n = \mathcal{O}_{\mathbf{P}}(1), \eta_n = o_{\mathbf{P}}(1)$, то для $\forall\,\varepsilon, M > 0$ вероятность $\mathbf{P}\left\{\left|\xi_n\eta_n\right| > \varepsilon\right\} \leqslant \leqslant \mathbf{P}\left\{\left|\xi_n\right| > M\right\} + \mathbf{P}\left\{\left|\eta_n\right| > \varepsilon/M\right\}$. Поэтому для $\forall\,\varepsilon > 0$

(г) $\overline{\lim\limits_n}\,\mathbf{P}\left\{\left|\xi_n\eta_n\right| > \varepsilon\right\} \leqslant \sup\limits_n \mathbf{P}\left\{\left|\xi_n\right| > M\right\} \to 0 \quad (M \to \infty)$. $\Lleftarrow$

✧ **Теорема непрерывности**. Сходимость по вероятности сохраняется при непрерывных преобразованиях. Мы сформулируем это утверждение применительно к последовательности сл.векторов.

Определение. Пусть k-мерные случайные векторы $\left\langle \vec{\xi}_n \right\rangle_{n=0}^{\infty}$ определены на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$. Говорят, что $\vec{\xi}_n$ *сходится по вероятности* к $\vec{\xi}_0$, если для $\forall\,\varepsilon > 0$

$$\lim_n \ \mathbf{P}\left\{ \|\vec{\xi}_n - \vec{\xi}_0\| > \varepsilon \right\} = 0,$$

где $\|\cdot\|$ — норма вектора в евклидовом пространстве $\mathbb{R}^k$.

Весьма важно, что сходимость по вероятности всего вектора $\vec{\xi}_n$ эквивалентна сходимости его компонент.

217| **Лемма.** (?!) *Для сходимости* $(\xi_{1n}, \ldots, \xi_{kn}) \xrightarrow[n]{\mathbf{P}} (\xi_{10}, \ldots, \xi_{k0})$ *необходимо и достаточно, чтобы* $\xi_{jn} \xrightarrow[n]{\mathbf{P}} \xi_{j0}$ *для* $\forall\, j = \overline{1, k}$.

Для пояснения условий следующей теоремы рассмотрим непрерывную функцию двух переменных $H(x,y) = x/y$, определённую при $y \neq 0$. Её можно доопределить в точках $(x, 0)$, положив равной произвольному значению. Подобрать это значение так, чтобы H стала непрерывной всюду, не удастся. С другой стороны, если функция H применяется к какому-либо вектору (ξ, η), то на этот вектор, естественно, накладывается условие $\eta \neq 0$.

218| **Теорема.** [*О непрерывных преобразованиях.*] *Пусть функция* $H: \mathbb{R}^k \mapsto \mathbb{R}^l$ *непрерывна всюду на борелевском множестве* $B \subset \mathbb{R}^k$, *сл.вектор* $\vec{\xi}_0$ *п.н. принадлежит множеству* B: $\mathbf{P}\left\{ \vec{\xi}_0 \in B \right\} = 1$. *Тогда*

$$\vec{\xi}_n \xrightarrow{\mathbf{P}} \vec{\xi}_0 \quad \Rightarrow \quad H(\vec{\xi}_n) \xrightarrow{\mathbf{P}} H(\vec{\xi}_0).$$

$\Rrightarrow$ Зафиксируем $\varepsilon > 0$ и для $m \geqslant 1$ рассмотрим множество

$$W_m = \left\{ \vec{x} \in B : \ \exists \vec{y}, \ \|\vec{y} - \vec{x}\| < \frac{1}{m} \ \& \ \|H(\vec{y}) - H(\vec{x})\| > \varepsilon \right\}.$$

Заметим, что $W_m \searrow \varnothing$. Действительно, если существует $\vec{x} \in \bigcap_1^\infty W_m$, то для $\forall m \geqslant 1$ найдётся $\vec{y}_m: \quad \|\vec{y}_m - \vec{x}\| < \dfrac{1}{m}$. Другими словами, $\lim_m \vec{y}_m = \vec{x}$ и, следовательно (по непрерывности функции H в точке $\vec{x} \in B$), предел $\lim_m H(\vec{y}_m) = H(\vec{x})$, что противоречит условию $\|H(\vec{y}_m) - H(\vec{x})\| > \varepsilon$.

Если $\vec{x} \notin W_m$ и $\|H(\vec{x}) - H(\vec{y})\| > \varepsilon$, то $\|\vec{y} - \vec{x}\| \geqslant \dfrac{1}{m}$. Поэтому

$$\mathbf{P}\bigl\{\, \vec{\xi}_0 \notin W_m,\ \|H(\vec{\xi}_0) - H(\vec{\xi}_n)\| > \varepsilon \,\bigr\} \ \leqslant\ \mathbf{P}\Bigl\{\, \|\vec{\xi}_0 - \vec{\xi}_n\| \geqslant \frac{1}{m} \,\Bigr\} \ \to\ 0$$

при $n \to \infty$ и фиксированном $m > 0$. С другой стороны,

$$\mathbf{P}\bigl\{\, \vec{\xi}_0 \in W_m,\ \|H(\vec{\xi}_0) - H(\vec{\xi}_n)\| > \varepsilon \,\bigr\} \ \leqslant\ \mathbf{P}\bigl\{\, \vec{\xi}_0 \in W_m \,\bigr\}.$$

Таким образом, при любом $m \geqslant 1$ (снова факт c_5)

$$\varlimsup_n \mathbf{P}\bigl\{\, \|H(\vec{\xi}_0) - H(\vec{\xi}_n)\| > \varepsilon \,\bigr\} \ \leqslant\ \mathbf{P}\bigl\{\, \vec{\xi}_0 \in W_m \,\bigr\}.$$

Ввиду непрерывности вероятности правая часть этого неравенства может быть сделана сколь угодно малой, так как $W_m \searrow \varnothing$. ⫤

219| Упр. Докажите утверждения теоремы 211, стр. 197.

✧ **Связь со сходимостью почти наверное.** Начнём с примера последовательности, сходящейся по вероятности, но не сходящейся почти наверное.

220| Пример. Пусть $\Omega = [0;1]$ с борелевской σ-алгеброй и мерой Лебега в качестве вероятности. Определим $\xi_n(\omega) = \dot{\mathrm{I}}_{A_n}(\omega)$ с интервалом $A_n = \bigl[\, k2^{-m}; (k+1)2^{-m} \,\bigr]$, $\quad m = \llbracket \log_2 n \rrbracket$, $k = n - 2^m$. Когда n изменяется от 2^m до $2^{m+1}-1$, интервалы A_n, на которых $\xi_n = 1$, пробегают весь отрезок $[0;1]$. Например, $A_4 = [0; {}^1\!/_4]$, $A_5 = [{}^1\!/_4; {}^2\!/_4]$, $A_6 = [{}^2\!/_4; {}^3\!/_4]$, $A_7 = [{}^3\!/_4; 1]$.

Таким образом, для $\forall\, 0 < \varepsilon < 1$ имеем $\mathbf{P}\bigl\{ |\xi_n| > \varepsilon \bigr\} = 1/2^m \underset{n}{\to} 0$, т.е. $\mathbf{P}\text{-}\lim_n \xi_n = 0$. Кроме того, $\mathbf{E}|\xi_n| = 1/2^m \underset{n}{\to} 0$, т.е. $\xi_n \xrightarrow{\mathcal{L}_1} 0$. С другой стороны, для $\forall\, \omega \in [0;1]$ найдётся подпоследовательность $\xi_{n_j}(\omega) = 1$ при $\forall\, j$, т.е. $\xi_n(\omega) \not\to 0$. ◎

221| Упр. а) При каких $q > 0$ в примере 220 $\xi_n \xrightarrow{\mathcal{L}_q} 0$?
б) Измените пример так, чтобы $\xi_n \xrightarrow{\mathbf{P}} 0$, но $\mathbf{E}|\xi_n| = 1$.

222| Теорема. I) *Если $\xi_n \xrightarrow{\text{п.н.}} \xi_0$, то $\xi_n \xrightarrow{\mathbf{P}} \xi_0$.*
II) *Если $\xi_n \xrightarrow{\mathcal{L}_q} \xi_0$ при некотором $q > 0$, то $\xi_n \xrightarrow{\mathbf{P}} \xi_0$.*
III) *Если $\xi_n \xrightarrow{\mathbf{P}} \xi_0$, то существует подпоследовательность $\xi_{n_k} \xrightarrow{\text{п.н.}} \xi_0$.*
IV) *Если $\xi_n \xrightarrow{\mathbf{P}} \xi_0$ и последовательность $\langle\, \xi_n \,\rangle_n$ монотонна, то $\xi_n \xrightarrow{\text{п.н.}} \xi_0$.*

⇛ I) Последовательность $\xi_n(\omega) \to \xi_0(\omega)$, если «для $\forall\, k \geqslant 1$ найдётся такое $N \geqslant 1$, что при всех $n \geqslant N$ модуль разности $\left|\xi_n(\omega) - \xi_0(\omega)\right| < \frac{1}{k}$ ». Поэтому множество элементарных исходов ω, где $\xi_n(\omega) \to \xi_0(\omega)$, можно записать так:

$$\bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \left\{ \left|\xi_n - \xi_0\right| < \frac{1}{k} \right\} = \lim_{k}{}_{\downarrow} \lim_{N}{}_{\uparrow} \bigcap_{n=N}^{\infty} \left\{ \left|\xi_n - \xi_0\right| < \frac{1}{k} \right\}.$$

По условию вероятность этого события равна единице, поэтому для $\forall\, \varepsilon > 0$

$$1 = \lim_{k}{}_{\downarrow} \lim_{N}{}_{\uparrow} \mathbf{P}\Big\{ \bigcap_{n=N}^{\infty} \left\{ \left|\xi_n - \xi_0\right| < \frac{1}{k} \right\} \Big\} \leqslant \varliminf_{N} \mathbf{P}\left\{ \left|\xi_N - \xi_0\right| < \varepsilon \right\}.$$

III) По определению для $\forall\, k \geqslant 1$ вероятность $\mathbf{P}\left\{ \left|\xi_n - \xi_0\right| > \frac{1}{k} \right\} \leqslant 1/2^k$ при всех достаточно больших n, скажем $n \geqslant n_k$. Таким образом,

$$\sum_{k=1}^{\infty} \mathbf{P}\left\{ \left|\xi_{n_k} - \xi_0\right| > \frac{1}{k} \right\} \leqslant \sum_{k=1}^{\infty} \frac{1}{2^k} < \infty.$$

По теореме 205, стр. 194, отсюда следует $\xi_{n_k} \xrightarrow{\text{п.н.}} \xi_0$.

IV) следует из III). ⇚

223| Упр. Докажите пункт II) теоремы, воспользовавшись неравенством Маркова 289, стр. 254.

224| Упр. Справедливо ли утверждение I) теоремы, если рассмотреть сходимость относительно произвольной меры?

⚠ Кроме монотонности последовательности (пункт IV теоремы) существует ещё одна ситуация, когда сходимость по вероятности влечёт п.н. сходимость. По теореме П. Леви последовательность частичных сумм независимых сл.в. сходится п.н., если она сходится по вероятности (см. [3], гл. 10).

Утверждение III) теоремы 222 допускает обращение.

225| **Лемма.** (?!) *Последовательность сл.в.* $\langle \xi_n \rangle_1^{\infty}$*, заданных на одном вероятностном пространстве, сходится по вероятности к сл.в.* ξ_0 *т. т. т. когда из любой подпоследовательности* $\langle \xi_{n_k} \rangle_{k=1}^{\infty}$ *можно извлечь подпоследовательность* $\xi_{n_{k_m}} \to \xi_0$ (п.н.) .

Кроме того, в силу I) и IV) теоремы 222 справедлива

226| **Лемма.** (?!) *Сл.в.* $\xi_n \xrightarrow{\text{п.н.}} \xi_0$ *т. т. т. когда* $\sup\limits_{k \geqslant n} \left|\xi_k - \xi_0\right| \xrightarrow[n]{\mathbf{P}} 0$.

Утверждение II) теоремы 222, применённое к ситуации примера 208, стр. 196, позволяет сделать вывод, что среднее число успехов в n независимых испытаниях Бернулли с вероятностью успеха p сходится по вероятности к p. Это утверждение носит название закона больших чисел Бернулли. Метод доказательства II) (неравенство Маркова) может быть использован и в более общей ситуации.

227| Теорема. [*Закон больших чисел Чебышёва.*] *Пусть независимые сл.в.* $\left\langle \xi_k \right\rangle_{k=1}^{\infty}$ *имеют одинаковые математические ожидания* $\mu = \mathbf{E}\xi_k$ *и дисперсии* $\sigma^2 = \mathbf{D}\xi_k < \infty$. *Тогда* $\frac{1}{n}\sum_1^n \xi_k \xrightarrow{\mathbf{P}} \mu$ *при* $n \to \infty$.

⇛ По неравенству Маркова (Чебышёва) для $\forall \varepsilon > 0$ при $n \to \infty$

$$\mathbf{P}\left\{ \left| \frac{1}{n}\sum_{k=1}^{n} \xi_k - \mu \right| > \varepsilon \right\} \leqslant \frac{1}{\varepsilon^2}\mathbf{D}\left[\frac{1}{n}\sum_{k=1}^{n} \xi_k \right] = \frac{n\sigma^2}{n^2\varepsilon^2} \to 0.$$ ⇚

✧ **Равномерная интегрируемость.** Утверждение III) теоремы 222 позволяет расширить теорему Лебега о мажорируемой сходимости на последовательности, сходящиеся по вероятности.

228| Лемма. (?!) (*О мажорируемой сходимости по вероятности.*) *Если* $\xi_n \xrightarrow{\mathbf{P}} \xi_0$ *и при всех* n *сл.величины* ξ_n *мажорируются некоторой сл.в.* η: $|\xi_n| \leqslant \eta$ (п.н.), *причём* $\mathbf{E}\eta < \infty$, *то* $\lim_n \mathbf{E}\xi_n = \mathbf{E}\xi_0$.

Для того чтобы усилить теорему Лебега, введём вспомогательное

Определение. Семейство $\left\langle \xi_\alpha \right\rangle$ сл.в.[‡] *равномерно интегрируемо порядка* $q \geqslant 0$ (просто *равномерно интегрируемо* при $q = 1$), если

$$\lim_{M\to\infty} \sup_\alpha \mathbf{E}\left[\left|\xi_\alpha\right|^q \dot{\mathrm{I}} \left(\left|\xi_\alpha\right| > M \right) \right] = 0. \tag{1}$$

Семейство из одной сл.в. ξ равномерно интегрируемо, если математическое ожидание ξ конечно. Действительно, $\left|\xi\right| \geqslant \left|\xi\right| \dot{\mathrm{I}}(\left|\xi\right| > M) \to 0$ при $M \to \infty$, поэтому $\lim_M \mathbf{E}[\left|\xi\right| \dot{\mathrm{I}}(\left|\xi\right| > M)] = 0$ в силу теоремы Лебега.

229| Лемма. a) *Если семейство сл.в.* $\left\langle \xi_\alpha \right\rangle$ *равномерно интегрируемо, то* $\sup_\alpha \mathbf{E}\left|\xi_\alpha\right| < \infty$ *и* $\sup_\alpha \mathbf{E}\left[\left|\xi_\alpha\right| \dot{\mathrm{I}}_A \right] \to 0$ *при* $\mathbf{P}\{A\} \to 0$.

b) *Если* $\left|\xi_\alpha\right| < \eta$ *для* $\forall\alpha$ *и* $\mathbf{E}\eta < \infty$, *то семейство* $\left\langle \xi_\alpha \right\rangle$ *равномерно интегрируемо.*

[‡] Подразумевается любое необходимое семейство сл.в., индексированных α.

c) *Если* $\mathbf{E}\left|\xi_\alpha\right|^s < C < \infty$ *для* $\forall\,\alpha$, *то семейство сл.в.* $\left\langle\,\xi_\alpha\,\right\rangle$ *равномерно интегрируемо порядка* $q < s$.

⤃ a) Так как $\sup_\alpha \mathbf{E}\left[\left|\xi_\alpha\right| \dot{\mathrm{I}}\left(\left|\xi_\alpha\right| > M\right)\right] < \infty$ для всех достаточно больших $M > 0$ и

$$\mathbf{E}\left|\xi_\alpha\right| \;=\; \mathbf{E}\left[\left|\xi_\alpha\right| \dot{\mathrm{I}}\left(\left|\xi_\alpha\right| \leqslant M\right)\right] + \mathbf{E}\left[\left|\xi_\alpha\right| \dot{\mathrm{I}}\left(\left|\xi_\alpha\right| > M\right)\right] \;\leqslant$$
$$\leqslant M + \mathbf{E}\left[\left|\xi_\alpha\right| \dot{\mathrm{I}}\left(\left|\xi_\alpha\right| > M\right)\right],$$

то $\sup_\alpha \mathbf{E}\left|\xi_\alpha\right| < \infty$. Для доказательства второго утверждения a) применим тот же приём: $\sup\limits_\alpha \mathbf{E}\left[\left|\xi_\alpha\right| \dot{\mathrm{I}}_A\right] \leqslant M\,\mathbf{P}\{A\} + \sup\limits_\alpha \mathbf{E}\left[\left|\xi_\alpha\right| \dot{\mathrm{I}}\left(\left|\xi_\alpha\right| > M\right)\right]$. Выбирая здесь сначала M достаточно большим, а затем $\mathbf{P}\{A\} < \varepsilon/(2M)$, можно добиться того, чтобы правая часть стала меньше наперёд заданного $\varepsilon > 0$.

b) Так как, очевидно, $\left|\xi_\alpha\right| \dot{\mathrm{I}}\left(\left|\xi_\alpha\right| > M\right) \leqslant \eta\,\dot{\mathrm{I}}\left(\eta > M\right)$, то

$$\lim_{M\to\infty} \sup_\alpha \mathbf{E}\left[\left|\xi_\alpha\right| \dot{\mathrm{I}}\left(\left|\xi_\alpha\right| > M\right)\right] \;\leqslant\; \lim_{M\to\infty} \mathbf{E}\left[\eta\,\dot{\mathrm{I}}\left(\eta > M\right)\right] \;=\; 0.$$

c) Для $\forall\, q < s$ при $M \to \infty$

$$\mathbf{E}\left[\left|\xi_\alpha\right|^q \dot{\mathrm{I}}\left(\left|\xi_\alpha\right| > M\right)\right] \;\leqslant\; \frac{1}{M^{s-q}}\,\mathbf{E}\left[\left|\xi_\alpha\right|^s \dot{\mathrm{I}}\left(\left|\xi_\alpha\right| > M\right)\right] \;\leqslant\; \frac{C}{M^{s-q}} \;\to\; 0,$$

что доказывает c), поскольку правая часть не зависит от α. ⤂

230| **Теорема.** I) *Если* $\xi_n \xrightarrow{\mathbf{P}} \xi_0$ *и семейство* $\left\langle\,\xi_n\,\right\rangle_1^\infty$ *равномерно интегрируемо, то* $\mathbf{E}\left|\xi_0\right| < \infty$ *и* $\xi_n \xrightarrow{\mathcal{L}_1} \xi_0$, *в частности* $\lim_n \mathbf{E}\,\xi_n = \mathbf{E}\,\xi_0$.

II) *Обратно, если* $\xi_n \xrightarrow{\mathbf{P}} \xi_0$ *и* $\lim_n \mathbf{E}\left|\xi_n\right| = \mathbf{E}\left|\xi_0\right| < \infty$, *то семейство* $\left\langle\,\xi_n\,\right\rangle_1^\infty$ *равномерно интегрируемо.*

⤃ I) В силу утверждения III) теоремы 222 можно выбрать подпоследовательность $\xi_{n_j} \xrightarrow{\text{п.н.}} \xi_0$. По лемме Фату 537, стр. 481, и лемме 229 $\mathbf{E}\left|\xi_0\right| \leqslant \varliminf\limits_j \mathbf{E}\left|\xi_{n_j}\right| \leqslant \sup\limits_n \mathbf{E}\left|\xi_n\right| < \infty$.

Далее, для $\forall\,\varepsilon > 0$ и множества $A_n = \{\left|\xi_n - \xi_0\right| \leqslant \varepsilon\}$ справедливо

$$\mathbf{E}\left|\xi_n - \xi_0\right| \;=\; \mathbf{E}\left[\left|\xi_n - \xi_0\right| \dot{\mathrm{I}}\left(A_n\right)\right] + \mathbf{E}\left[\left|\xi_n - \xi_0\right| \dot{\mathrm{I}}\left(A_n^c\right)\right] \leqslant$$
$$\leqslant\; \varepsilon + \mathbf{E}\left[\left|\xi_n\right| \dot{\mathrm{I}}\left(A_n^c\right)\right] + \mathbf{E}\left[\left|\xi_0\right| \dot{\mathrm{I}}\left(A_n^c\right)\right].$$

Так как $\lim_n \mathbf{P}\{A_n^c\} = 0$, то снова по лемме 229 получаем, что оба математические ожидания справа стремятся к нулю при $n \to \infty$. Следовательно, $\xi_n \xrightarrow{\mathcal{L}_1} \xi_0$ и, как было отмечено в теореме 209, стр. 196, $\mathbf{E}\,\xi_n \to \mathbf{E}\,\xi_0$.

II) Рассмотрим последовательность множеств $A_n = \{|\xi_n| \leqslant M_n\}$ с некоторой последовательностью чисел $M_n \to \infty$. В силу свойств верхнего предела

$$\varlimsup_n \mathbf{E}\left[|\xi_n|\, \dot{\mathrm{I}}(A_n^c)\right] \leqslant \varlimsup_n \mathbf{E}\,|\xi_n| + \varlimsup_n \big(-\mathbf{E}\left[|\xi_n|\, \dot{\mathrm{I}}(A_n)\right]\big) = \\ = \varlimsup_n \mathbf{E}\,|\xi_n| - \varliminf_n \mathbf{E}[|\xi_n|\, \dot{\mathrm{I}}(A_n)].$$

По условию $\varlimsup_n \mathbf{E}\,|\xi_n| = \mathbf{E}\,|\xi_0|$. Покажем, что $\varliminf_n \mathbf{E}[|\xi_n|\, \dot{\mathrm{I}}(A_n)] \geqslant \mathbf{E}\,|\xi_0|$.

Определим непрерывную функцию $\phi_{_M}(x)$:

$$\phi_{_M}(x) = 1\cdot \dot{\mathrm{I}}\left(x; (-\infty, M - \tfrac{1}{M}]\right) + M(M-x)\cdot \dot{\mathrm{I}}\left(x; [M - \tfrac{1}{M}, M]\right) + \\ + 0\cdot \dot{\mathrm{I}}\left(x; [M, \infty)\right).$$

То есть $\phi_{_M}(x) = 1$ при $x \leqslant M - 1/M$ и $\phi_{_M}(x) = 0$ при $x \geqslant M$. Так как сходимость по вероятности сохраняется при непрерывном преобразовании, то $|\xi_n|\phi_{_M}(|\xi_n|) \xrightarrow[n]{\mathbf{P}} |\xi_0|\phi_{_M}(|\xi_0|) \xrightarrow[M\to\infty]{} |\xi_0|$. Следовательно,

$$\mathbf{E}\,|\xi_0| = \lim_{M\to\infty} \lim_{n\to\infty} \mathbf{E}\left[|\xi_n|\phi_{_M}(|\xi_n|)\right],$$

где переход к пределу разрешён по теореме Лебега: в первом случае (по n), поскольку $|\xi_n|\phi_{_M}(|\xi_n|) \leqslant M$, а во втором случае (по M) ввиду неравенства $|\xi_0|\phi_{_M}(|\xi_0|) \leqslant |\xi_0|$.

Пусть $M > 0$, тогда $M_n > M$, начиная с некоторого n, и

$$\mathbf{E}\left[|\xi_n|\, \dot{\mathrm{I}}(A_n)\right] \geqslant \mathbf{E}\left[|\xi_n|\, \dot{\mathrm{I}}(|\xi_n| \leqslant M)\right] \geqslant \mathbf{E}\left[|\xi_n|\phi_{_M}(|\xi_n|)\right],$$

т.к. $\phi_{_M}(x) \leqslant \dot{\mathrm{I}}\left(x; (-\infty, M]\right)$, $x \in \mathbb{R}^1$. Следовательно, $\varliminf_n \mathbf{E}[|\xi_n|\, \dot{\mathrm{I}}(A_n)] \geqslant$ $\geqslant \mathbf{E}\,|\xi_0|$.

Таким образом, предел $\lim_n \mathbf{E}[|\xi_n|\, \dot{\mathrm{I}}(|\xi_n| > M_n)] = 0$ для $\forall\, M_n \to \infty$, что достаточно (?!) для равномерной интегрируемости. ⫤

§ 3. Слабая сходимость

Очень часто, особенно в математической статистике, возникает потребность в вычислении вероятностей попадания сл.в. в определённые области её значений. В этом случае полезны следующие

Определения. Последовательность сл.в. $\left\langle \xi_n \backsim F_n \right\rangle_{n=1}^{\infty}$ *сходится по распределению* (*слабо*) к сл.в. $\xi_0 \backsim F_0$, если функции распределения F_n

сходятся в основном к ф.р. F_0, т.е. $\lim_n F_n(x) = F_0(x)$ в каждой точке x, где ф.р. $F_0(x)$ непрерывна.

⚠ Понятие слабой сходимости в большей степени относится к сходимости законов распределений, а не самих сл.в. — пишут $\xi_n \xrightarrow{low} \xi_0$ (т.е. по закону распределения) или $\xi_n \xrightarrow{\mathrm{d}} \xi_0$ (от *англ.* distribution). Для этого типа сходимости не обязательно, чтобы сл.в. были заданы на одном пространстве. В принципе, можно вовсе отказаться от рассмотрения сл.в. и говорить только о сходимости вероятностных законов (мер) P_n или ф.р. F_n на прямой $(\mathbb{R}^1, \mathscr{B})$ — тогда пишут $\mathsf{P}_n \xrightarrow{\mathrm{w}} \mathsf{P}_0$ (week), $F_n \Rightarrow F_0$ (в основном). Когда сл.в. ξ_n задаются конструктивно, используются записи типа $\xi_n \rightsquigarrow \mathsf{P}_0$ ($\xi_n \rightsquigarrow \mathsf{P}_0$) или $\xi_n \rightsquigarrow \xi_0$, которые читаются как: «ξ_n имеет асимптотическое распределение P_0 (такое же как у ξ_0)». При этом предельный закон P_0 часто заменяется его мнемоническим представлением. Так, в наиболее востребованном случае, когда P_0 — нормальный закон, пишут $\xi_n \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$ и читают как «(последовательность) сл.в. ξ_n при $n \to \infty$ имеет асимптотически нормальное распределение со средним μ и дисперсией σ^2». В этом последнем случае, дабы подчеркнуть приближённую нормальность распределения каждого элемента последовательности, часто употребляют неформальную запись вида $\xi_n \rightsquigarrow \mathcal{N}(\mu_n, \sigma_n^2)$, которая формально трактуется как сходимость $(\xi_n - \mu_n)/\sigma_n \rightsquigarrow \mathcal{N}(0,1)$.

231| Упр. Справедливы ли следующие импликации:

а) $\xi_n \xrightarrow{\mathbf{P}} \xi_0 \;\Rightarrow\; \xi_n - \xi_0 \xrightarrow{\mathbf{P}} 0$; б) $\xi_n \rightsquigarrow \xi_0 \;\Rightarrow\; \xi_n - \xi_0 \rightsquigarrow 0$?

Требование сходимости в основном, а не в каждой точке, расширяет область применения на ситуации, которые вполне подходят под понятие близости распределений. Например, если рассмотреть сл.в. ξ_n с равномерным распределением в отрезке $[-\frac{1}{n}; \frac{1}{n}]$, то естественно считать её близкой к точке $\xi_0 \equiv 0$. В то же время значение ф.р. $F_n(0) = \frac{1}{2} \not\to F_0(0) = 1$.

Следующая лемма даёт ещё одно оправдание необходимости рассмотрения сходимости ф.р. в основном.

232| Лемма. *Пусть ξ, η — две сл.в., заданные на одном вероятностном пространстве, $s > 0$ — действительный параметр. Тогда $\xi + s\,\eta \rightsquigarrow \xi$, $s \to 0$.*

⊫ Заметим, что если сумма двух чисел $a + b \leqslant c + d$, то либо $a \leqslant c$, либо $b \leqslant d$. Поэтому для $\forall\, x \in \mathbb{R}^1, \varepsilon > 0$

$$F_{\xi+s\eta}(x) \;=\; \mathbf{P}\{\xi + s\eta \leqslant x\} \;\leqslant\; \mathbf{P}\{\xi \leqslant x + \varepsilon\} + \mathbf{P}\{s\eta \leqslant -\varepsilon\}\,.$$

Аналогично, $1 - F_{\xi+s\eta}(x) = \mathbf{P}\{\xi + s\eta > x\} \leqslant \mathbf{P}\{\xi > x - \varepsilon\} + \mathbf{P}\{s\eta > \varepsilon\}$.

Так как $s > 0$, то событие $\{s\eta \leqslant -\varepsilon\} = \{\eta \leqslant -\varepsilon/s\} \to \varnothing$ и, аналогично, $\{s\eta > \varepsilon\} = \{\eta > \varepsilon/s\} \to \varnothing$ при $s \to 0$. Таким образом, имеем

$$F_\xi(x-\varepsilon) \leqslant \varliminf_s F_{\xi+s\eta}(x) \leqslant \varlimsup_s F_{\xi+s\eta}(x) \leqslant F_\xi(x+\varepsilon).$$

Если x — точка непрерывности F_ξ, то, полагая $\varepsilon \to 0$, получаем, что ф.р. $F_{\xi+s\eta}(x) \underset{s}{\to} F_\xi(x)$, т.е. $F_{\xi+s\eta} \Rightarrow F_\xi$ при $s \to 0$. ⫤

Введённое определение слабой сходимости не зависит от способа определения функции распределения. Имеет место

233| **Лемма.** *Последовательность* $\xi_n \rightsquigarrow \xi_0$ *т. т. т. когда*

$$\lim_n \mathbf{P}\{\xi_n < x\} = \mathbf{P}\{\xi_0 < x\}$$

в любой точке x, *в которой ф.р.* ξ_0 *непрерывна, т.е.* $\mathbf{P}\{\xi_0 < x\} = \mathbf{P}\{\xi_0 \leqslant x\}$.

⊨ Если x — точка непрерывности ф.р. F_0, то для $\forall \varepsilon > 0$ найдётся точка непрерывности этой ф.р. $y < x$ такая, что $F_0(x) - \varepsilon \leqslant F_0(y)$. Заметим, что $F_n(y) \leqslant \mathbf{P}\{\xi_n < x\}$. Поэтому если $\xi_n \rightsquigarrow \xi_0$, то

$$\begin{aligned} F_0(x) - \varepsilon \leqslant \lim_n F_n(y) \leqslant \varliminf_n \mathbf{P}\{\xi_n < x\} \leqslant \\ \leqslant \varlimsup_n \mathbf{P}\{\xi_n < x\} \leqslant \lim_n \mathbf{P}\{\xi_n \leqslant x\} = F_0(x). \end{aligned}$$

Аналогично устанавливается $\xi_n \rightsquigarrow \xi_0$ при выполнении условий леммы. ⫤

Заметим, что сходимость только в точках непрерывности предельной ф.р. не приводит к некорректной ситуации с несколькими предельными ф.р., поскольку количество точек разрыва любой ф.р. не более чем счётно, а посему, как вытекает из 27, стр. 45, все предельные ф.р. должны совпадать всюду.

234| Упр. Докажите, что если $\xi_n \rightsquigarrow \xi$ и $\xi_n \rightsquigarrow \eta$, то $\xi \backsim \eta$.

Непосредственно из определения могут быть установлены только простейшие свойства последовательностей, сходящихся слабо.

235| Упр. Докажите, что если $\xi_n \rightsquigarrow \xi_0$, то

а) $a\xi_n + b \rightsquigarrow a\xi_0 + b$ для любых констант $a, b \in \mathbb{R}^1$;

б) $H(\xi_n) \rightsquigarrow H(\xi_0)$ для любой непрерывной строго монотонной функции $H : \mathbb{R}^1 \mapsto \mathbb{R}^1$.

Для получения более содержательных результатов, касающихся одной последовательности сл.в., часто удобнее использовать критерий слабой сходимости из следующей теоремы. Установление фактов, относящихся к нескольким последовательностям, требует знания совместного распределения вектора, составленного из соответствующих элементов последовательностей. Для примера: пусть $\xi_n \backsim \mathcal{N}(0,1), \eta_n \backsim \mathcal{N}(0,1)$ для $\forall\, n = 0, 1, \ldots$ Очевидно, что $\xi_n \rightsquigarrow \xi_0$, $\eta_n \rightsquigarrow \eta_0$, но нельзя утверждать, что, скажем, разность $(\xi_n - \eta_n) \rightsquigarrow (\xi_0 - \eta_0)$, поскольку асимптотическое распределение $(\xi_n - \eta_n)$ зависит от совместного распределения ξ_n и η_n. Если, например, при каждом нечётном $n \geqslant 1$ сл.в. ξ_n, η_n независимы и $\xi_n \equiv \eta_n$ при каждом чётном n, то последовательность $(\xi_n - \eta_n)$, очевидно, не имеет предела.

✧ Теорема Хелли–Брэя. Слабую сходимость сл.векторов можно определить также через сходимость ф.р. Однако такое определение не слишком удобно даже в теоретическом плане. Поэтому в следующей теореме мы приведём критерий слабой сходимости, который и кладётся обычно в основу определения слабой сходимости для случайных элементов в произвольных метрических пространствах. Утверждение о предельном переходе под знаком математического ожидания в этой теореме называют теоремой Хелли. Обозначим через $C_b(\mathcal{D})$ класс всех непрерывных ограниченных (continuous bounded) функций, заданных на метрическом пространстве $\mathcal{D}$ и принимающих значения в $\mathbb{R}^1$.

236| **Теорема.** [*E. Helly, H.E. Bray.*] *Последовательность случайных величин* $\xi_n \rightsquigarrow \xi_0$ *т. т. т. когда* $\lim_n \mathbf{E}\, h(\xi_n) = \mathbf{E}\, h(\xi_0)$ *для* $\forall\, h \in C_b(\mathbb{R}^1)$.

⊫ Пусть $\xi_n \rightsquigarrow \xi_0$ и $h(x)$ — ограниченная $(|h(x)| < H)$ непрерывная функция, $\varepsilon > 0$. Выберем $M > 0$ так, чтобы

$$F_0(-M) + 1 - F_0(M) = \mathbf{P}\{\xi_0 \leqslant -M\} + \mathbf{P}\{\xi_0 > M\} < \frac{\varepsilon}{H}\,. \tag{2}$$

Если $(-M)$ — точка непрерывности ф.р. F_0, то по определению

$$\mathbf{P}\{\xi_n \leqslant -M\} = F_n(-M) \quad \to \quad F_0(-M) = \mathbf{P}\{\xi_0 \leqslant -M\}\,.$$

Аналогичный факт справедлив и для вероятности $\mathbf{P}\{\xi_n > M\}$. Следовательно, существует такое $N^* > 0$, что для $\forall\, n > N^*$

$$\mathbf{P}\{\xi_n \leqslant -M\} + \mathbf{P}\{\xi_n > M\} < \frac{\varepsilon}{H}\,.$$

На отрезке $[-M; M]$ функция h равномерно непрерывна, т.е. существует конечное разбиение $(-M; M] = \sum_1^K (y_{j-1}; y_j]$ такое, что разность значений функции внутри каждого интервала не превосходит ε : $\big|h(u) - h(b)\big| < \varepsilon$, $y_{j-1} < u, b \leqslant y_j$, $j = 1, \ldots, K$. Выберем все точки $-M = y_0 < y_1 < \ldots < y_K = M$ так, чтобы они были точками непрерывности ф.р. F_0 (напомним, что количество точек разрыва ф.р. не более чем счётно).

Пусть ζ_j — середина интервала $(y_{j-1}; y_j]$. Определим простую ступенчатую функцию $h_\varepsilon(x) = \sum_{j=1}^K h(\zeta_j)\,\dot{\mathrm{I}}\,(x; (y_{j-1}; y_j])$.

По построению $\bigl|h(x) - h_\varepsilon(x)\bigr| < \varepsilon$ для $\forall\, x \in (-M; M]$. Вне этого интервала $\bigl|h(x) - h_\varepsilon(x)\bigr| = \bigl|h(x)\bigr| < H$. Следовательно, для $\forall\, n \geqslant 0$

$$\mathbf{E}\bigl|h(\xi_n) - h_\varepsilon(\xi_n)\bigr| \;<\; \varepsilon + H\Bigl(\mathbf{P}\{\xi_n \leqslant -M\} + \mathbf{P}\{\xi_n > M\}\Bigr) \;<\; 2\varepsilon.$$

Так как границы интервалов выбраны как точки непрерывности ф.р. F_0, то по условию теоремы $\lim_n \mathbf{P}\{\xi_n \in (y_{j-1}; y_j]\} = \lim_n (F_n(y_j) - F_0(y_{j-1})) = \\ = \mathbf{P}\{\xi_0 \in (y_{j-1}; y_j]\}$. Поэтому

$$\Bigl|\,\mathbf{E}\bigl[h_\varepsilon(\xi_n) - h_\varepsilon(\xi_0)\bigr]\Bigr| \;\leqslant\; \sum\nolimits_1^K \Bigl|\,\mathbf{P}\{\xi_n \in (y_{j-1}; y_j]\} - \mathbf{P}\{\xi_0 \in (y_{j-1}; y_j]\}\Bigr|\bigl|h(\zeta_j)\bigr| \underset{n}{\to} 0,$$

т.е. существует такое $N^{**} > N^*$, что $\mathbf{E}\bigl|h_\varepsilon(\xi_n) - h_\varepsilon(\xi_0)\bigr| < \varepsilon$ для $\forall\, n > N^{**}$. Таким образом, для $\forall\, n > N^{**}$

$$\begin{aligned} \Bigl|\,\mathbf{E}\,h(\xi_n) - \mathbf{E}\,h(\xi_0)\Bigr| \;&\leqslant\; \Bigl|\,\mathbf{E}\bigl[h_\varepsilon(\xi_n) - h_\varepsilon(\xi_0)\bigr]\Bigr| + \\ &+ \mathbf{E}\bigl|h(\xi_n) - h_\varepsilon(\xi_n)\bigr| + \mathbf{E}\bigl|h(\xi_0) - h_\varepsilon(\xi_0)\bigr| \;<\; 5\varepsilon. \end{aligned}$$

Обратно, пусть x — некоторая точка непрерывности F_0 и $\varepsilon > 0$. Рассмотрим непрерывную ограниченную функцию

$$h(t) = 1 \cdot \dot{\mathrm{I}}\,(t; (-\infty; x - \varepsilon]) + \left(1 - \frac{t - (x - \varepsilon)}{\varepsilon}\right) \cdot \dot{\mathrm{I}}\,(t; [x - \varepsilon; x]).$$

Индикатор интервала $(-\infty; x]$ связан с этой функцией неравенствами

$$h(t) \;\leqslant\; \dot{\mathrm{I}}\,(t; (-\infty; x]) \;\leqslant\; h(t - \varepsilon), \quad t \in \mathbb{R}^1.$$

Таким образом, $\mathbf{E}\,h(\xi_n) \leqslant F_n(x) \leqslant \mathbf{E}\,h(\xi_n - \varepsilon)$. Переходя здесь к пределу по $n \to \infty$, получаем в силу условия теоремы

$$\mathbf{E}\,h(\xi_0) \;\leqslant\; \varliminf_n F_n(x) \;\leqslant\; \varlimsup_n F_n(x) \;\leqslant\; \mathbf{E}\,h(\xi_0 - \varepsilon).$$

Поскольку $h(t) \geqslant \dot{\mathrm{I}}\,(t; (-\infty; x - \varepsilon])$ и $h(t - \varepsilon) \leqslant \dot{\mathrm{I}}\,(t; (-\infty; x + \varepsilon])$, то из предыдущих неравенств получаем

$$F_0(x - \varepsilon) \;\leqslant\; \varliminf_n F_n(x) \;\leqslant\; \varlimsup_n F_n(x) \;\leqslant\; F_0(x + \varepsilon).$$

Так как x — точка непрерывности F_0, то, полагая $\varepsilon \to 0$, получаем, что $F_0(x) \;\leqslant\; \varliminf_n F_n(x) \;\leqslant\; \varlimsup_n F_n(x) \;\leqslant\; F_0(x)$. ⟸

237 ⚠ В доказательстве теоремы Хелли (первой части теоремы Хелли–Брэя) существенно используется тот факт, что распределение сл.в. ξ_0 собственное, т.е. для $\forall\varepsilon > 0$ найдётся $M > 0$ такое, что для ф.р. $F_0(-M) + 1 - F_0(M) < \varepsilon$. Теорема Хелли–Брэя справедлива и для несобственных распределений:

последовательность ф.р. F_n сходится в основном к некоторой ограниченной неубывающей непрерывной справа функции F_0 т. т. т. когда $\lim_n \int h(x)\,dF_n(x) = \int h(x)\,dF_0(x)$ для любой ограниченной непрерывной функции h такой, что $h(-\infty) = h(+\infty) = 0$.

Доказательство этого факта отличается от приведённого выше только моментом выбора числа M в (2). Для несобственного распределения F_0 его нужно выбрать так, чтобы

$$\int_{-\infty}^{-M} h(x)\,dF_0(x) + \int_{M}^{\infty} h(x)\,dF_0(x) < \varepsilon.$$

Такое M существует в силу условий на функцию h.

✧ Сходимость мер в метрических пространствах. Теорема 236 позволяет расширить понятие слабой сходимости на случайные элементы, принимающие значения в произвольном метрическом пространстве. Такое расширение удобнее осуществлять, не обращаясь к терминам, связанным со случайными элементами, а оперируя только их вероятностными законами распределения (вероятностными мерами).

Напомним, что в топологическом, в частности метрическом, пространстве $\mathcal{D}$ борелевская σ-алгебра $\mathscr{B}(\mathcal{D})$ определяется как наименьшая σ-алгебра, содержащая все открытые подмножества $\mathcal{D}$. Отсюда следует, что всюду непрерывные функции на $(\mathcal{D}, \mathscr{B}(\mathcal{D}))$ измеримы по Борелю.

Определение. Последовательность вероятностных мер $\left\langle Q_n \right\rangle_1^\infty$ на борелевской σ-алгебре $\mathscr{B}(\mathcal{D})$ слабо сходится к вероятностной мере Q_0 на $\mathscr{B}(\mathcal{D})$ (обозначается $Q_n \xrightarrow{\mathrm{w}} Q_0$), если для $\forall\, h \in C_b(\mathcal{D})$

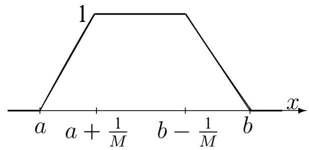
$$\lim_n \int\limits_{\mathcal{D}} h(x)\, Q_n(dx) = \int\limits_{\mathcal{D}} h(x)\, Q_0(dx).$$

Прежде всего необходимо удостовериться, что слабый предел, если он существует, определяется единственным образом. Пусть Z — подмножество метрического пространства $\mathcal{D}$ (с метрикой ρ). Обозначим расстояние от точки

$x \in \mathcal{D}$ до множества Z через
$$\rho(x, Z) := \inf_{y \in Z} \rho(x, y).$$
Легко понять, что для любого открытого множества G расстояние $\rho(x, G^c) >$ > 0 при $x \in G$. Поэтому $\lim\limits_{M \to \infty} \rho(x, G^c) M = \infty$ для $\forall\, x \in G$; $\quad \rho(x, G^c) = 0$ для $\forall\, x \notin G$. Кроме того, поскольку метрика есть непрерывная функция своих аргументов, функция $\rho(x, G^c)$ непрерывна по x.

Определим непрерывную функцию
$$\psi(t) = 0 \cdot \dot{\mathrm{I}}_{(-\infty;0)}(t) + t \cdot \dot{\mathrm{I}}_{[0;1]}(t) + 1 \cdot \dot{\mathrm{I}}_{(1;+\infty)}(t). \tag{3}$$
Таким образом, $\lim_{M \to \infty} \psi(\rho(x, G^c)M) = \dot{\mathrm{I}}(x; G)$ для открытых множеств G, причём $\psi(\rho(x, G^c)M) \leqslant \psi(\rho(x, G^c)N)$ при $M < N$.

Для $G = (a; b)$ непрерывная ограниченная функция $\psi(\rho(x, G^c)M)$, $x \in \mathbb{R}^1$, с достаточно большим $M > 0$ имеет графическое представление в виде равнобокой трапеции.



238| **Лемма.** *Вероятностные меры Q_1, Q_2 на $\mathscr{B}(\mathcal{D})$ совпадают, если $\int_{\mathcal{D}} h\, dQ_1 = \int_{\mathcal{D}} h\, dQ_2$ для $\forall\, h \in C_b(\mathcal{D})$.*

$\Rrightarrow$ Покажем сначала, что меры совпадают на всех открытых множествах. По условию для $\forall\, M > 0$ и любого открытого множества $G \subset \mathcal{D}$
$$\int_{\mathcal{D}} \psi(\rho(x, G^c)M)\, Q_1(dx) = \int_{\mathcal{D}} \psi(\rho(x, G^c)M)\, Q_2(dx).$$
В силу теоремы о монотонной сходимости здесь можно перейти к пределу под знаком интеграла при $M \to \infty$. Таким образом, получаем, что $Q_1(G) = Q_2(G)$ для любого открытого множества $G \subset \mathcal{D}$.

Класс $\mathfrak{O}$ всех открытых множеств, порождающий борелевскую σ-алгебру, замкнут относительно конечных пересечений, поэтому в силу 13, стр. 35, меры совпадают на $\mathscr{B}(\mathcal{D})$. $\Lleftarrow$

Утверждения следующей леммы часто упрощают проверку условий слабой сходимости.

239| **Лемма.** (?!) *Последовательность $Q_n \xrightarrow{\mathrm{w}} Q_0$ при выполнении любого из следующих условий:*

(a) *для любой непрерывной ограниченной функции $h : \mathcal{D} \mapsto \mathbb{R}^1$*
$$\varlimsup_n \int\limits_{\mathcal{D}} h(x)\, dQ_n \leqslant \int\limits_{\mathcal{D}} h(x)\, dQ_0;$$

(b) *из любой подпоследовательности* $\left\langle Q_{n'},\ \{n'\} \subset \mathbb{N} \right\rangle$ *можно извлечь подпоследовательность* $Q_{n''} \xrightarrow{\mathrm{w}} Q_0$ $(\{n''\} \subset \{n'\})$.

✧ **Критерии слабой сходимости.** При выводе различных свойств весьма востребованы характеризации (критерии) слабой сходимости, приводимые в следующей теореме, за которой закрепилось наименование теоремы о «портфеле» (Portmanteau). В случае пространства $\mathbb{R}^k$ в этот портфель включается ещё характеризация через сходимость в основном соответствующих ф.р.

240| **Теорема.** [*О слабой сходимости. А.Д. Александров.*] *Пусть вероятностные меры* Q_n, $n \geqslant 0$, *заданы на борелевской* σ*-алгебре* $\mathscr{B}(\mathcal{D})$ *в метрическом пространстве* $\mathcal{D}$. *Следующие условия эквивалентны*[‡]:

$(\mathrm{W_c})$ $Q_n \xrightarrow{\mathrm{w}} Q_0$, *т.е. для* $\forall\, h \in C_b(\mathcal{D})$

$$\lim_n \int_{\mathcal{D}} h(x)\, Q_n(dx) = \int_{\mathcal{D}} h(x)\, Q_0(dx);$$

$(\mathrm{W_p})$ *для* $\forall$ *неотрицательной непрерывной функции* $h(x)$

$$\varliminf_n \int_{\mathcal{D}} h(x)\, Q_n(dx) \geqslant \int_{\mathcal{D}} h(x)\, Q_0(dx);$$

$(\mathrm{W_t})$ *для* $\forall$ *открытого множества* $G \subset \mathcal{D}$ *(замкнутого множества* Z*)*

$$\varliminf_n Q_n(G) \geqslant Q_0(G) \qquad \Big(\varlimsup_n Q_n(Z) \leqslant Q_0(Z)\Big);$$

$(\mathrm{W_d})$ *для* $\forall\, G \in \mathscr{B}(\mathcal{D})$ *такого, что мера границы* $Q_0(\partial G) = 0$,

$$\lim_n Q_n(G) = Q_0(G).$$

⊫ Схема доказательства: $(\mathrm{W_c}) \Leftrightarrow (\mathrm{W_p}) \Rightarrow (\mathrm{W_t}) \Rightarrow (\mathrm{W_d}) \Rightarrow (\mathrm{W_t}) \Rightarrow (\mathrm{W_p})$. Используем сокращённую запись $\int h\, dQ = \int_{\mathcal{D}} h(x)\, Q(dx)$.

$(\mathrm{W_c}) \Rightarrow (\mathrm{W_p})$. Пусть $h(x)$ — неотрицательная непрерывная функция. Определим для $R > 0$ ограниченную непрерывную функцию

$$h_R(x) = h(x) \cdot \dot{\mathrm{I}}\,(x; h(x) \leqslant R) + R \cdot \dot{\mathrm{I}}\,(x; h(x) \geqslant R), \qquad x \in \mathcal{D}.$$

Поскольку $h(x) \geqslant h_R(x)$ для $\forall\, x$, то $\int h\, dQ_n \geqslant \int h_R\, dQ_n$. По условию $(\mathrm{W_c})$ предел правой части $\lim_n \int h_R\, dQ_n = \int h_R\, dQ_0$. В левой части неравенства предел может не существовать, однако можно утверждать, что

$$\varliminf_n \int_{\mathcal{D}} h(x)\, Q_n(dx) \geqslant \int_{\mathcal{D}} h_R(x)\, Q_0(dx).$$

[‡] Нумерация списка критериев здесь привязана к характеризационному классу: $(\mathrm{W_c})$ — класс непрерывных (continuity) ограниченных функций; $(\mathrm{W_p})$ — класс положительных (positive) функций; $(\mathrm{W_t})$ — класс множеств, определяющих топологию; $(\mathrm{W_d})$ — класс множеств, определяющих распределение (distribution).

Если положить $R \to \infty$, то $h_R(x) \nearrow h(x)$ и по теореме о монотонной сходимости $\lim\limits_R \int h_R\, dQ_0 = \int h\, dQ_0$.

$(\mathrm{W_p}) \Rightarrow (\mathrm{W_c})$. Пусть $h \in C_b(\mathcal{D})$ и $\left|h(x)\right| < H$. Функция $H - h(x)$, очевидно, непрерывна и положительна. Поэтому по условию $(\mathrm{W_p})$

$$\varliminf_n \Big(H - \int_{\mathcal{D}} h(x)\, Q_n(dx)\Big) \geqslant H - \int_{\mathcal{D}} h(x)\, Q_0(dx).$$

Следовательно, $\varlimsup\limits_n \int h\, dQ_n \leqslant \int h\, dQ_0$, что по лемме 239 гарантирует сходимость $Q_n \xrightarrow{\mathrm{w}} Q_0$. Другой способ: применить $(\mathrm{W_p})$ к $h(x) + H \geqslant 0$.

$(\mathrm{W_p}) \Rightarrow (\mathrm{W_t})$. Пусть G — открытое множество. Рассмотрим неотрицательную непрерывную функцию $\psi(\rho(x, G^c)M)$, $x \in \mathcal{D}, M > 0$ (см. (3) и сопутствующие описания перед леммой 238). Справедливо неравенство

$$Q_n(G) = \int_{\mathcal{D}} \dot{\mathrm{I}}\,(x; G)\, Q_n(dx) \geqslant \int_{\mathcal{D}} \psi(\rho(x, G^c)M)\, Q_n(dx).$$

По условию $(\mathrm{W_p})$ предел $\varliminf\limits_n Q_n(G) \geqslant \int \psi(\rho(\cdot, G^c)M)\, dQ_0$. Поскольку $\lim\limits_{M}{}_{\uparrow}\, \psi(\rho(x, G^c)M) = \dot{\mathrm{I}}\,(x; G)$ при $M \nearrow \infty$, утверждение $(\mathrm{W_t})$ следует теперь из теоремы о монотонной сходимости.

Условие относительно замкнутых множеств эквивалентно условию для открытых множеств, поскольку дополнение замкнутого множества открыто и $\varlimsup\limits_n Q_n(Z) = 1 - \varliminf\limits_n Q_n(Z^c)$.

$(\mathrm{W_t}) \Rightarrow (\mathrm{W_d})$. Пусть G — борелевское множество с нулевой вероятностью границы: $Q_0(\partial G) = 0$. Граница $\partial G = [G] \setminus G^o$, где $[G]$ — замыкание множества G, G^o — его внутренняя часть. Другими словами, $[G] = G^o + \partial G$, поэтому $Q_0([G]) = Q_0(G^o) = Q_0(G)$. Применяя условие $(\mathrm{W_c})$ к открытому множеству G^o и замкнутому множеству $[G]$, получаем

$$\begin{aligned} Q_0(G) \;&=\; Q_0(G^o) \;\leqslant\; \varliminf_n Q_n(G^o) \;\leqslant\; \varliminf_n Q_n(G) \;\leqslant \\ &\leqslant\; \varlimsup_n Q_n(G) \;\leqslant\; \varlimsup_n Q_n([G]) \;\leqslant\; Q_0([G]) \;=\; Q_0(G). \end{aligned}$$

Таким образом, $Q_0(G) = \varliminf\limits_n Q_n(G) = \varlimsup\limits_n Q_n(G) = \lim\limits_n Q_n(G)$.

$(\mathrm{W_d}) \Rightarrow (\mathrm{W_t})$. Пусть $Z_r = \big\{x \in \mathcal{D} : \rho(x, Z) < r\big\}$ — окрестность замкнутого множества Z, $r > 0$. Легко понять, что $Z \subset Z_r$ и $Z_r \searrow Z$ при $r \searrow 0$ ввиду замкнутости Z. Граница $\partial Z_r \subset \big\{x \in \mathcal{D} : \rho(x, Z) = r\big\}$, поэтому границы окрестностей с различными r не пересекаются и, значит, найдётся не более чем счётное число r, при которых $Q_0(\partial Z_r) > 0$. Поэтому можно выбрать так

$r_k \searrow 0$, что $Q_0(\partial Z_{r_k}) = 0$ при всех $k \geqslant 1$. По условию $(\mathrm{W_d})$ (применённому к последовательности Z_{r_k})

$$\varlimsup_n Q_n(Z) \ \leqslant \ \lim_n Q_n(Z_{r_k}) \ \stackrel{\mathrm{W_d}}{=} \ Q_0(Z_{r_k}).$$

Полагая $k \to \infty$ $(r_k \to 0)$, получаем утверждение $(\mathrm{W_t})$.

$(\mathrm{W_t}) \Rightarrow (\mathrm{W_p})$. Пусть h — неотрицательная непрерывная функция. Рассмотрим приближение функции h снизу посредством простых функций:

$$h(x) \ = \ \lim_k{}_{\uparrow} h_k(x) = \lim_k{}_{\uparrow} \sum_{j=1}^{K_k} h_{j-1}\, \dot{\mathrm{I}}\,(x; B_j), \quad x \in \mathcal{D},$$

где $B_j = \{x : \ h(x) \in [h_{j-1}; h_j)\}$ и точки $0 = h_0 < h_1 < \ldots < h_{K_k}$ выбраны так, что $Q_0(x : h(x) = h_j) = 0$ — такое всегда возможно, т.к. существует не более чем счётное число точек $y : \ Q_0(x : h(x) = y) > 0$. Таким образом,

$$\int_{\mathcal{D}} h(x)\, Q_n(dx) \ \geqslant \ \sum_{j=1}^{K_k} h_{j-1} Q_n(B_j) \ \geqslant \ \sum_{j=1}^{K_k} h_{j-1} Q_n(\overset{o}{B_j}),$$

где множества $\overset{o}{B_j} = \{x : \ h(x) \in (h_{j-1}; h_j)\}$, $j = 1, \ldots, K_k$, открыты, т.к. представляют собой прообразы открытых множеств при непрерывном отображении. Заметим, что по построению $Q_0(B_j) = Q_0(\overset{o}{B_j})$ для всех j. Применяя к правой части последнего неравенства условие $(\mathrm{W_t})$, получаем

$$\varliminf_n \int_{\mathcal{D}} h(x)\, Q_n(dx) \ \geqslant \ \sum_{j=1}^{K_k} h_{j-1} Q_0(\overset{o}{B_j}) \ = \ \sum_{j=1}^{K_k} h_{j-1} Q_0(B_j) \underset{k\to\infty}{\to} \int_{\mathcal{D}} h\, dQ,$$

где последнее соотношение следует из определения интеграла Лебега. ⋳

► Класс функций, определяющий слабую сходимость, можно уменьшить. Функция h называется *финитной*, если она непрерывна и найдётся компактное множество Z такое, что $h(x) = 0$ для $\forall\, x \notin Z$. Функция h называется *липшицевой*, если $\forall\, x, y \in \mathcal{D}$ разность $\bigl|h(x) - h(y)\bigr| \leqslant L\, \rho(x, y)$ с некоторой константой L.

241| **Лемма.** *Каждое из следующих условий достаточно для того, чтобы* $Q_n \stackrel{\mathrm{W}}{\to} Q_0$:

(W_ℓ) $\int h\, dQ_n \to \int h\, dQ_0$ *для любой липшицевой функции h;*

(W_φ) $\int h\, dQ_n \to \int h\, dQ_0$ *для любой финитной функции h.*

$\models$ $(\mathrm{W}_l) \Rightarrow (\mathrm{W}_t)$. Пусть G — открытое подмножество $\mathcal{D}$, тогда при $m \to \infty$ последовательность функций $g_m(x) = \min\{1, m\,\rho(x, G^c)\} \nearrow \dot{\mathrm{I}}_G$. При $\forall\, m$ функция g_m липшицева с коэффициентом 1, следовательно,

$$\varliminf_n Q_n(G) \ = \ \varliminf_n \int_{\mathcal{D}} \dot{\mathrm{I}}_G \; dQ_n \ \geqslant \ \lim_n \int_{\mathcal{D}} g_m \, dQ_n \ = \ \int_{\mathcal{D}} g_m \, dQ_0.$$

По теореме о монотонной сходимости правая часть этого неравенства при $m \to \infty$, возрастая, сходится к $\int \dot{\mathrm{I}}_G \; dQ_0 = Q_0(G)$.

$(\mathrm{W}_\varphi) \Rightarrow (\mathrm{W}_c)$. Пусть h — непрерывная ограниченная функция: $\bigl|h\bigr| \leqslant H$. По заданному $\varepsilon > 0$ выберем $R > 0$ так, чтобы мера $Q_0(S^c) < \varepsilon/H$, где S — замкнутый шар (с центром в нуле) радиуса R. Снова используем функцию ψ (см. (3)) и рассмотрим финитную функцию $g(x) = 1 - \psi(\rho(x, S))$, которая принимает значение ноль при $\rho(x, 0) > R + 1$. Для финитной функции hg по условию леммы $\int h\, g\, dQ_n \to \int h\, g\, dQ_0$. Кроме того, $\int (1 - g)\, dQ_n \to$ $\to \int (1 - g)\, dQ_0$, т.к. все меры $Q_n, n \geqslant 0$, вероятностные.

Поскольку $0 \leqslant 1 - g(x) \leqslant 1$, причём $1 - g(x) = 0$ только для $x \in S$, то $1 - g(x) \leqslant \dot{\mathrm{I}}\,(x; S^c)$ и, следовательно, $\int (1 - g)\, dQ_0 \leqslant Q_0(S^c) < \varepsilon/H$.

Таким образом, $\bigl|\int h(1 - g)\, dQ_0\bigr| \leqslant \varepsilon$ и

$$\varlimsup_n \bigl|\int_{\mathcal{D}} h(1 - g)\, dQ_n\bigr| \ \leqslant \ H \lim_n \int_{\mathcal{D}} (1 - g)\, dQ_n \ \leqslant \ \varepsilon.$$

Следовательно, для $\forall\, \varepsilon > 0$

$$\begin{aligned} \varlimsup_n \bigl|\int_{\mathcal{D}} h\, dQ_n - \int_{\mathcal{D}} h\, dQ_0\bigr| \ \leqslant \ \varlimsup_n \bigl|\int_{\mathcal{D}} h(1 - g)\, dQ_n\bigr| + \bigl|\int_{\mathcal{D}} h(1 - g)\, dQ_0\bigr| + \\ + \lim_n \bigl|\int_{\mathcal{D}} h\, g\, dQ_n \to \int_{\mathcal{D}} h\, g\, dQ_0\bigr| \ \leqslant 2\, \varepsilon, \end{aligned}$$

что доказывает второе утверждение леммы. $\Leftarrow\!\!\!|$

⚠ Как будет видно из дальнейшего, слабая сходимость — в действительности слабейшая из всех традиционно рассматриваемых типов сходимости сл.элементов. С другой стороны, по знаменитой теореме Ф. Рисса любой линейный непрерывный функционал на $C_b([0;1])$ может быть представлен в виде интеграла $\int_{[0;1]} h(t)\, Q(dt)$, $h \in C_b([0;1])$, относительно некоторого заряда (меры со знаком) Q. То есть слабая сходимость мер на $\mathcal{D} = [0;1]$ эквивалентна так называемой $*$-слабой сходимости в пространстве $C_b^*(\mathcal{D})$ всех линейных непрерывных функционалов на $C_b(\mathcal{D})$.

✧ Слабая сходимость случайных векторов. Слабая сходимость в евклидовом пространстве (сходимость распределений сл.векторов) и все критерии теоремы о портфеле могут быть сформулированы в терминах математических ожиданий и вероятностей попадания в соответствующие множества. Кроме того, здесь, как и в одномерном случае, представляет интерес критерий слабой сходимости, выраженный через ф.р. сл.векторов. Нам понадобится следующая

242| **Лемма.** *Открытое множество $B \subset \mathbb{R}^k$ может быть представлено в виде счётной суммы непересекающихся параллелепипедов: $B = \biguplus_1^\infty (\vec{a}_j; \vec{c}_j]$; причём это представление может быть подобрано так, что все вершины каждого из параллелепипедов будут точками непрерывности заданной ф.р. F.*

⤇ Пусть F_i — маргинальная ф.р. i-й компоненты F. Так как функция F_i имеет не более чем счётное число точек разрыва, то можно построить такую монотонную последовательность разбиений числовой прямой:

$$\delta_M^{(i)} = \left\{x_{nM}^{(i)} : \quad x_{nM}^{(i)} < x_{(n+1)M}^{(i)}, \quad -\infty < n < \infty\right\}, \ \lim_{n\to\pm\infty} x_{nM}^{(i)} = \pm\infty,$$
$$\delta_M^{(i)} \subset \delta_{M+1}^{(i)}, \qquad M = 1, 2, \ldots, \qquad \lim_M \sup_n \left|x_{nM}^{(i)} - x_{(n+1)M}^{(i)}\right| = 0,$$

что для $\forall\, n, M$ точки $x_{nM}^{(i)}$ будут точками непрерывности F_i.

При каждом $M \geqslant 1$ образуем сеточное разбиение $\mathbb{R}^k$ в виде параллелепипедов заданного вида, соответствующее разбиениям $\delta_M^{(i)}$, $i = \overline{1,k}$, координатных пространств. Непрерывность маргинальных ф.р. гарантирует, что каждый узел этой сетки будет точкой непрерывности совместной ф.р. F. Искомое представление множества B строится индуктивно — на шаге с номером M в это представление добавляются все те параллелепипеды, образованные сеткой для выбранного M, которые лежат в B и не принадлежат объединению всех параллелепипедов, выбранных на предыдущих шагах построения.

Поскольку B открыто, для $\forall\, \vec{x}_0 \in B$ найдётся шар с центром в этой точке, лежащий в B, который, в свою очередь, содержит параллелепипед из системы разбиений, содержащий $\vec{x}_0$. ⤆

Пусть $C_b(\mathbb{R}^k)$ обозначает класс непрерывных ограниченных функций, $C(\mathbb{R}^k)$— класс непрерывных функций, $L_b(\mathbb{R}^k)$ — класс ограниченных липшицевых функций, $C_\varphi(\mathbb{R}^k)$ — класс финитных функций на $\mathbb{R}^k$.

243| **Теорема.** [*Критерии слабой сходимости сл.векторов.*] *Пусть $\vec{\xi}_n$, $n \geqslant 0$, — сл.векторы в $\mathbb{R}^k$, F_n — соответствующие им ф.р. Следую-*

щие условия эквивалентны:

(W$_\mathrm{c}$) $\vec{\xi}_n \rightsquigarrow \vec{\xi}_0$, *т.е.* $\lim_n \mathbf{E}\, h(\vec{\xi}_n) = \mathbf{E}\, h(\vec{\xi}_0)$ *для* $\forall\, h \in C_b(\mathbb{R}^k)$

$\{$ (W$_l$) *для* $\forall\, h \in L_b(\mathbb{R}^k)$, (W$_\varphi$) *для* $\forall\, h \in C_\varphi(\mathbb{R}^k)$ $\}$;

(W$_\mathrm{t}$) $\underline{\lim}_n\, \mathbf{P}\{\, \vec{\xi}_n \in G^o \,\} \geqslant \mathbf{P}\{\, \vec{\xi}_0 \in G^o \,\}$ *для* $\forall\, G \subset \mathbb{R}^k$

$\{$ $\overline{\lim}_n\, \mathbf{P}\{\, \vec{\xi}_n \in [G] \,\} \leqslant \mathbf{P}\{\, \vec{\xi}_0 \in [G] \,\}$ *для* $\forall\, G \subset \mathbb{R}^k$ $\}$;

(W$_\mathrm{p}$) $\underline{\lim}_n\, \mathbf{E}\, h(\vec{\xi}_n) \geqslant \mathbf{E}\, h(\vec{\xi}_n)$ *для* $\forall\, h \in C(\mathbb{R}^k)$, $h \geqslant 0$;

(W$_\mathrm{d}$) $\lim_n \mathbf{P}\{\, \vec{\xi}_n \in G \,\} = \mathbf{P}\{\, \vec{\xi}_0 \in G \,\}$ *для* $\forall\, G \in \mathscr{B}^k$ *с вероятностью попадания на границу* $\mathbf{P}\{\, \vec{\xi}_0 \in \partial G \,\} = 0$;

(W$_\mathrm{df}$) $\lim_n F_n(\vec{x}) = F_0(\vec{x})$ *для* $\forall\, \vec{x}$ *— точки непрерывности* F_0.

⫤ Доказательства требует только пункт (W$_\mathrm{df}$).

(W$_\mathrm{d}$) $\Rightarrow$ (W$_\mathrm{df}$). Пусть $\vec{x}_0$ — точка непрерывности ф.р. F_0. Тогда граница множества $G = \big\{\vec{x} : \vec{x} \leqslant \vec{x}_0\big\}$ имеет нулевую F_0-вероятность. Следовательно, по критерию (W$_\mathrm{d}$) справедливо

$$\lim_n F_n(\vec{x}_0) \;=\; \lim_n \mathbf{P}\{\, \vec{\xi}_n \in G \,\} \;=\; \mathbf{P}\{\, \vec{\xi}_0 \in G \,\} \;=\; F_0(\vec{x}_0).$$

(W$_\mathrm{df}$) $\Rightarrow$ (W$_\mathrm{t}$). Пусть B — открытое подмножество $\mathbb{R}^k$. По предыдущей лемме $B = \biguplus_j (\vec{a}_j; \vec{c}_j]$, где все вершины каждого из параллелепипедов являются точками непрерывности F_0. Следовательно, для $\forall\, j \geqslant 1$ $\lim_n \Delta_{\vec{a}_j, \vec{c}_j} F_n = \Delta_{\vec{a}_j, \vec{c}_j} F_0$. Таким образом, для $\forall\, N > 0$ имеют место соотношения:

$$\underline{\lim}_n\, \mathbf{P}\{\, \vec{\xi}_n \in B \,\} \;\geqslant\; \lim_n \sum\nolimits_{j=1}^{N} \mathbf{P}\{\, \vec{\xi}_n \in (\vec{a}_j; \vec{c}_j] \,\} \;=\; \sum\nolimits_{j=1}^{N} \mathbf{P}\{\, \vec{\xi}_0 \in (\vec{a}_j; \vec{c}_j] \,\}.$$

Полагая в правой части этого неравенства $N \to \infty$, получаем, что предел $\underline{\lim}_n\, \mathbf{P}\{\, \vec{\xi}_n \in B \,\} \geqslant \mathbf{P}\{\, \vec{\xi}_0 \in B \,\}$. ⫣

⚠ Сл.векторы могут быть заданы на разных вероятностных пространствах — $\vec{\xi}_n : (\Omega_n, \mathscr{F}_n, \mathbf{P}_n) \mapsto (\mathbb{R}^k, \mathscr{B}^k)$. Поэтому, строго говоря, в теореме каждое утверждение необходимо снабдить соответствующим индексом. Мы отнесли информацию о мере, относительно которой вычисляются вероятности и математические ожидания, в символ соответствующего сл.вектора.

244| Упр. Пусть $\Xi = \big\langle\, \xi_j \backsim \mathcal{U}\mathrm{n}(0,1) \,\big\rangle_1^\infty$ — последовательность независимых равномерных сл.в., $\xi_{(1:n)}, \xi_{(n:n)}$ — соответственно наименьшее и наибольшее значения первых n элементов Ξ. Покажите, что $\big(n\xi_{(1:n)}, n(1 -$

$- \xi_{(n:n)})\big) \rightsquigarrow (\eta_1, \eta_2)$, где сл.в. η_1, η_2 независимы и распределены одинаково по показательному $\mathcal{E}\mathrm{x}(1)$ закону.

245| Пример. Эквивалентность критериев слабой сходимости теоремы о портфеле имеет место только при условии, что все меры (как «допредельные» Q_n, так и предельная Q_0) вероятностные, т.е. для них $Q_n(\mathcal{D}) = 1$ или, в более общем случае, $Q_n(\mathcal{D}) = Const < \infty$. Рассмотрим последовательность ф.р. $F_n(x) = \frac{1}{2}\big(1 - \frac{1}{x}\big)$, $1 \leqslant x < n$, $F_n(x) = 1, x \geqslant n$. Очевидно, $\lim_n F_n(x) = F_0(x) = \frac{1}{2}\big(1 - \frac{1}{x}\big)\, \dot{\mathrm{I}}\,(x; [1; \infty))$ в каждой точке $x \in \mathbb{R}^1$, т.е. $F_n \Rrightarrow F_0$. Функция F_0 есть обобщённая ф.р. — для неё $F(+\infty) = {}^1\!/_2 < 1$.

Для любой измеримой функции h

$$\int_{\mathbb{R}} h\, dF_n = \int_1^n h(x) \frac{1}{2x^2}\, dx + h(n) \frac{1}{2}\big(1 + \frac{1}{n}\big), \qquad n \geqslant 1.$$

Поэтому если функция h — финитная, то $h(n) = 0$ для всех достаточно больших n и $\int h\, dF_n \to \int h\, dF_0$. Если функция $h \geqslant 0$, в частности $h = \dot{\mathrm{I}}_G$, то $\int_{\mathbb{R}} h\, dF_n \geqslant \int_1^n h\, dF_0$. Таким образом, для этой последовательности выполняются условия критериев $(\mathrm{W_p})$, (W_φ) и $(\mathrm{W_t})$ с открытыми множествами.

С другой стороны, условия $(\mathrm{W_c})$, (W_l) для $h(x) = \sin x$, $(\mathrm{W_t})$ для замкнутого $Z = \mathbb{R}^1$ и $(\mathrm{W_d})$ для $G = \mathbb{N}$ не выполняются. ◎

✧ **Свойства слабой сходимости векторов**. Следующее свойство является простым следствием $(\mathrm{W_c})$.

246| **Лемма.** (?!) *Если $\xi_n \rightsquigarrow \xi_0$, $\xi_n \in \mathbb{R}^k, n \geqslant 0$, и функция $H : \mathbb{R}^k \mapsto \mathbb{R}^m$ непрерывна всюду на $\mathbb{R}^k$, то $H(\xi_n) \rightsquigarrow H(\xi_0)$.*

Докажем теперь аналогичное утверждение для функции H, непрерывной только на носителе сл.вектора ξ_0.

247| **Теорема.** [*Непрерывность слабой сходимости.*] *Пусть функция $H : \mathbb{R}^k \mapsto \mathbb{R}^m$ непрерывна во всех точках множества $B \subset \mathbb{R}^k$, причём $\mathbf{P}\big\{ \xi_0 \in B \big\} = 1$. Тогда*

$$\xi_n \rightsquigarrow \xi_0 \quad \Rightarrow \quad H(\xi_n) \rightsquigarrow H(\xi_0).$$

⊫ Рассмотрим замкнутое множество $Z \subset \mathbb{R}^m$. Тогда его прообраз

$$H^{-1}(Z) \ \subset\ [H^{-1}(Z)] \ \subset\ H^{-1}(Z) \cup B^c.$$

Здесь первое включение очевидно. Для доказательства второго выберем $x \in [H^{-1}(Z)]$. Если эта точка не является точкой непрерывности H, то она входит в B^c. В противном случае, т.к. она принадлежит замыканию $[H^{-1}(Z)]$, найдётся последовательность $x_n \in H^{-1}(Z)$ (т.е. $H(x_n) \in Z$) такая, что $x_n \to$ $\to x$. Так как функция H непрерывна в x, то $H(x_n) \to H(x) \in Z$ в силу замкнутости Z. Следовательно, $x \in H^{-1}(Z)$.

Поскольку событие $\{H(\xi_n) \in Z\} = \{\xi_n \in H^{-1}(Z)\}$, то по критерию $(\mathrm{W_t})$

$$\varlimsup_n \mathbf{P}\big\{\, H(\xi_n) \in Z \,\big\} \ \leqslant\ \varlimsup_n \mathbf{P}\big\{\, \xi_n \in [H^{-1}(Z)] \,\big\} \ \leqslant\ \mathbf{P}\big\{\, \xi_0 \in [H^{-1}(Z)] \,\big\}.$$

Так как $\mathbf{P}\big\{\, \xi_0 \in B^c \,\big\} = 0$, то

$$\mathbf{P}\big\{\, \xi_0 \in [H^{-1}(Z)] \,\big\} \ \leqslant\ \mathbf{P}\big\{\, \xi_0 \in H^{-1}(Z) \cup B^c \,\big\} \ =\ \mathbf{P}\big\{\, \xi_0 \in H^{-1}(Z) \,\big\}.$$

Следовательно, $\varlimsup\limits_n \mathbf{P}\big\{\, H(\xi_n) \in Z \,\big\} \leqslant \mathbf{P}\big\{\, H(\xi_0) \in Z \,\big\}$ для любого замкнутого множества Z. ⫤

► Иногда сходимость имеет равномерный характер. Случай равномерной сходимости для всех борелевских множеств рассмотрим позднее при изучении сходимости по вариации. Сейчас же приведём свойство равномерной сходимости по классу интервалов вида $(-\infty; \vec{x}\,]$, т.е. для ф.р.

248 **Теорема.** I) *Если последовательность ф.р.* $F_n \Rightarrow F^*$ *и ф.р.* $F^*(\vec{x})$ *непрерывна во всех точках* $\vec{x} \in \mathbb{R}^k$, *то при* $n \to \infty$

$$\sup_{\vec{x}} \big| F_n(\vec{x}) - F^*(\vec{x}) \big| \ \to\ 0. \tag{4}$$

II) *Утверждение* (4) *в одномерном случае справедливо также для любой ф.р.* F^*, *если* $F_n(x) \to F^*(x)$ *и вдобавок* $F_n(x\pm) \to F^*(x\pm)$ *для* $\forall\, x$.

⊨ I) Упрощая выкладки, рассмотрим только двумерный случай. Как следует из 25, стр. 44, в условиях теоремы одномерные маргинальные ф.р. F_1^*, F_2^* непрерывны. Поэтому для каждой из них найдутся такие точки $-\infty =$ $= x_{j0} < x_{j1} < \ldots < x_{jM} = \infty$, что $F_j^*(x_{jm}) = m/M$. Набор двумерных точек (x_{1m}, x_{2l}), $m, l = \overline{1, M}$, образует конечную сетку, для которой разность значений совместной ф.р. в соседних узлах не превосходит $1/M$. Например,

$$F^*(x_{1m}, x_{2l}) - F^*(x_{1(m-1)}, x_{2l}) \ =\ \mathbf{P}\left\{ x_{1(m-1)} < \xi_1 \leqslant x_{1m}, \xi_2 \ \leqslant\ x_{2,l} \right\} \ \leqslant$$
$$\leqslant \mathbf{P}\left\{ x_{1(m-1)} < \xi_1 \leqslant x_{1m} \right\} \ =\ F_1^*(x_{1m}) - F_1^*(x_{1(m-1)}) \ =\ \frac{1}{M}.$$

Поэтому разность значений F^* в правом верхнем и левом нижнем углах любой ячейки этой сетки $F^*(x_{1m}, x_{2l}) - F^*(x_{1(m-1)}, x_{2(l-1)}) \leqslant 2/M$. Следовательно, если $(x_{1(m-1)}, x_{2(l-1)}) < \vec{x} \leqslant (x_{1m}, x_{2l})$, то в силу монотонности ф.р.

$$F_n(\vec{x}) - F^*(\vec{x}) \leqslant F_n(x_{1m}, x_{2l}) - F^*(x_{1(m-1)}, x_{2(l-1)}) \leqslant$$
$$\leqslant F_n(x_{1m}, x_{2l}) - F^*(x_{1m}, x_{2l}) + \frac{2}{M} \leqslant \max_{i,j} \left| F_n(x_{1i}, x_{2j}) - F^*(x_{1i}, x_{2j}) \right| + \frac{2}{M}.$$

С другой стороны,

$$F_n(\vec{x}) - F^*(\vec{x}) \geqslant F_n(x_{1(m-1)}, x_{2(l-1)}) - F^*(x_{1m}, x_{2l}) \geqslant$$
$$\geqslant F_n(x_{1(m-1)}, x_{2(l-1)}) - F^*(x_{1(m-1)}, x_{2(l-1)}) - \frac{2}{M} \geqslant$$
$$\geqslant -\max_{i,j} \left| F_n(x_{1i}, x_{2j}) - F^*(x_{1i}, x_{2j}) \right| - \frac{2}{M}.$$

Таким образом, $\sup\limits_{\vec{x}} \left| F_n(\vec{x}) - F^*(\vec{x}) \right| \leqslant \max\limits_{i,j} \left| F_n(x_{1i}, x_{2j}) - F^*(x_{1i}, x_{2j}) \right| + \frac{2}{M}$. Поскольку максимум конечного числа сходящихся к нулю последовательностей также сходится к нулю, то по условию имеем

$$\overline{\lim_n} \sup_{\vec{x}} \left| F_n(\vec{x}) - F^*(\vec{x}) \right| \leqslant \frac{2}{M},$$

что доказывает теорему ввиду произвольности $M > 0$.

Часть II) теоремы доказывается теми же рассуждениями. ⫤

249| **Теорема.** [*Гливенко–Кантелли.*] *Пусть* $\Xi = \left\langle \xi_n \backsim F^* \right\rangle_1^\infty$ — *последовательность независимых одинаково распределённых сл.в. на вероятностном пространстве* $(\Omega, \mathscr{F}, \mathbf{P})$,

$$F_n(x) = F_n(x; \omega) = \frac{1}{n} \sum_1^n \dot{\mathrm{I}} \left(\xi_k; (-\infty; x] \right), \quad x \in \mathbb{R}^1,$$

— *эмпирическая ф.р., построенная на первых* n *реализациях* Ξ. *Тогда*

$$\sup_x \left| F_n(x) - F^*(x) \right| \xrightarrow{\text{п.н.}} 0.$$

⊨ Возьмём произвольное x и рассмотрим пространство Q последовательностей вида $\boldsymbol{U} = \big((v_1, z_1), (v_2, z_2), \dots \big)$, $v_j, z_j = 0, 1$, $j \geqslant 1$. Отображение $(v_j, z_j) = \big(\dot{\mathrm{I}}_{(-\infty; x]}(\xi_j), \dot{\mathrm{I}}_{(-\infty; x)}(\xi_j) \big)$, $j \geqslant 1$, из Ω в Q порождает на Q распределение P, для которого последовательности $(v_1, v_2, \dots)$ и $(z_1, z_2, \dots)$ могут рассматриваться как реализации испытаний в схеме Бернулли с вероятностями успеха $p = F^*(x)$ и $\tilde{p} = F^*(x-0)$. В соответствии с сильным законом больших чисел (см. пример 207, стр. 195) для почти любой последовательности $\boldsymbol{U} \in Q$, т.е. для почти всех $\omega \in \Omega$ имеем

$$F_n(x; \omega) \to F^*(x) \quad \text{и} \quad F_n(x-0; \omega) \to F^*(x-0). \tag{5}$$

Так как множество рациональных чисел счётно, то отсюда следует, что найдётся подмножество $A \subset \Omega$ такое, что $\mathbf{P}\{A\} = 1$, и при любом $\omega \in A$ предельные соотношения (5) выполняются для любого рационального x. Поскольку множество рациональных чисел всюду плотно на $\mathbb{R}^1$, легко показать, что для тех же $\omega \in A$ соотношения (5) будут выполняться при $\forall\, x \in \mathbb{R}^1$.

Таким образом, для $\forall\, \omega \in A$ имеем последовательность ф.р. $F_n(\cdot\,;\omega)$, значения которых $F_n(x;\omega)$ и $F_n(x \pm 0;\omega)$ при всех x сходятся к соответствующим значениям ф.р. F^*. В силу части II) теоремы 248 можно утверждать, что эта сходимость равномерная по всем $x \in \mathbb{R}^1$. $\Lleftarrow$

Следующая лемма легко следует из определения слабой сходимости, точнее, из критерия $(\mathrm{W_c})$.

250| **Лемма.** (?!) *Если* $\vec{\xi}_n = (\xi_{1n}, \ldots, \xi_{kn}) \rightsquigarrow \vec{\xi}_0 = (\xi_{10}, \ldots, \xi_{k0})$, *то и каждая компонента* $\xi_{jn} \rightsquigarrow \xi_{j0}, \quad j = \overline{1,k}$.

Как отмечалось ранее, обратное не всегда верно, т.е. покомпонентная слабая сходимость не гарантирует слабую сходимость всего вектора.

251| Упр. Докажите, что если при $\forall\, n \geqslant 1$ сл.в. ξ_n, ζ_n независимы и $\xi_n \rightsquigarrow \xi_0, \quad \zeta_n \rightsquigarrow \zeta_0$, то найдётся такой сл.вектор (α, β) с независимыми компонентами, что $(\xi_n, \zeta_n) \rightsquigarrow (\alpha, \beta)$.

Если один из подвекторов сходится по вероятности к константе, составной вектор уже́ всегда будет иметь соответствующий слабый предел. Докажем сначала одно вспомогательное утверждение, представляющее самостоятельный интерес.

252| **Лемма.** *Пусть* $\langle\, \vec{\xi}_n,\ \vec{\eta}_n \,\rangle_{n=0}^{\infty}$ — *сл.векторы, заданные на одном вероятностном пространстве, причём при* $n \to \infty$

$$\vec{\eta}_n \rightsquigarrow \vec{\xi}_0 \quad \text{и} \quad \|\vec{\xi}_n - \vec{\eta}_n\| \xrightarrow{\mathbf{P}} 0.$$

Тогда $\vec{\xi}_n \rightsquigarrow \vec{\xi}_0$. *Другими словами, справедлива импликация:*

$$\big[\, \vec{\xi}_n \rightsquigarrow \vec{\xi}_0,\ \vec{\varepsilon}_n \xrightarrow{\mathbf{P}} \vec{0}\,\big] \quad \Rightarrow \quad \big[\, \vec{\xi}_n + \vec{\varepsilon}_n \rightsquigarrow \vec{\xi}_0 \big].$$

$\Rrightarrow$ Воспользуемся критерием (W_φ). Пусть $h(\vec{x})$ — финитная функция $(\max|h(\vec{x})| < H)$. Так как $h(\vec{x})$ равномерно непрерывна, то для $\forall\, \varepsilon > 0$ найдётся такое $\delta > 0$, что $\big|h(\vec{x}) - h(\vec{y})\big| < \varepsilon$ при $\|\vec{x} - \vec{y}\| < \delta$. Разбивая область интегрирования на части $\|\vec{\xi}_n - \vec{\eta}_n\| < \delta$ и $\|\vec{\xi}_n - \vec{\eta}_n\| \geqslant \delta$, получаем для $\forall\, \varepsilon > 0$

$$\big|\mathbf{E}\, h(\vec{\xi}_n) - \mathbf{E}\, h(\vec{\eta}_n)\big| \leqslant \varepsilon + 2H\, \mathbf{P}\big\{\, \|\vec{\xi}_n - \vec{\eta}_n\| \geqslant \delta \,\big\}.$$

Так как при $n \to \infty$ второе слагаемое по условию сходится к нулю, то $\mathbf{E}\, h(\vec{\xi}_n)$ и $\mathbf{E}\, h(\vec{\eta}_n)$ имеют одинаковый предел, что и требовалось. ⫤

253| **Теорема.** *Пусть сл.вектор $(\vec{\alpha}_n, \vec{\beta}_n)$ состоит из векторов $\vec{\alpha}_n = (\alpha_{1n}, \ldots, \alpha_{kn})$ размерности k и $\vec{\beta}_n = (\beta_{1n}, \ldots, \beta_{rn})$ размерности r. Если $\vec{\alpha}_n \rightsquigarrow \vec{\alpha}_0$, $\vec{\beta}_n \xrightarrow{\mathbf{P}} \vec{b}_0$, где $\vec{b}_0$ — детерминированный вектор, то*

$$(\vec{\alpha}_n, \vec{\beta}_n) \rightsquigarrow \vec{\xi}_0 = (\vec{\alpha}_0, \vec{b}_0).$$

⫢ Рассмотрим $(k+r)$-мерные векторы $\vec{\xi}_n = (\vec{\alpha}_n, \vec{\beta}_n)$ и $\vec{\eta}_n = (\vec{\alpha}_n, \vec{b}_0)$. Очевидно, для расстояний в пространствах $\mathbb{R}^{k+r}$ и $\mathbb{R}^r$ соответственно справедливо следующее соотношение: $\|\vec{\eta}_n - \vec{\xi}_n\|_{\mathbb{R}^{k+r}} = \|\vec{\beta}_n - \vec{b}_0\|_{\mathbb{R}^r} \xrightarrow[n]{\mathbf{P}} 0$.

Кроме того, $\vec{\eta}_n \rightsquigarrow \vec{\xi}_0$, т.к. по критерию слабой сходимости $(\mathrm{W_c})$, применённому к последовательности $\vec{\alpha}_n$, $\lim_n \mathbf{E}\, h(\vec{\eta}_n) = \lim_n \mathbf{E}\, h(\vec{\alpha}_n, \vec{b}_0) = \mathbf{E}\, h(\vec{\xi}_0)$, поскольку любая непрерывная ограниченная функция $h : \mathbb{R}^{k+r} \mapsto \mathbb{R}^1$ будет также непрерывной ограниченной функцией из $\mathbb{R}^k$ в $\mathbb{R}^1$, если в ней зафиксировать часть (здесь r) переменных. Снова применяя критерий слабой сходимости $(\mathrm{W_c})$ и лемму 252, получаем $\vec{\xi}_n \rightsquigarrow \vec{\xi}_0$. ⫤

✧ Связь со сходимостью по вероятности и в среднем.

254| **Теорема.** *Имеют место следующие соотношения:*

I) *если* $\vec{\xi}_n \xrightarrow{\mathbf{P}} \vec{\xi}_0$, *то* $\vec{\xi}_n \rightsquigarrow \vec{\xi}_0$;

II) *если* $\vec{\xi}_n \rightsquigarrow \vec{C}$ *с константой* $\vec{C}$, *то* $\vec{\xi}_n \xrightarrow{\mathbf{P}} \vec{C}$;

III) *если* $\xi_n \rightsquigarrow \xi_0$, $\mathbf{E}\left|\xi_0\right| < \infty$ *и семейство* $\left\langle \xi_n \right\rangle_1^\infty$ *равномерно интегрируемо, то* $\mathbf{E}\,\xi_n \to \mathbf{E}\,\xi_0$.

⫢ I) Представим $\vec{\xi}_n$ в виде $\vec{\xi}_n = \vec{\xi}_0 + (\vec{\xi}_n - \vec{\xi}_0)$. Утверждение I) следует теперь из леммы 252, поскольку, очевидно, $\vec{\xi}_0 \rightsquigarrow \vec{\xi}_0$.

II) Пусть $\vec{\xi}_0 \equiv \vec{C}$, тогда для $\forall \varepsilon > 0$ вероятность $\mathbf{P}\{\, \|\vec{\xi}_0 - \vec{C}\| < \varepsilon \,\} = 1$ и по критерию слабой сходимости $(\mathrm{W_t})$

$$\varliminf_n \mathbf{P}\{\, \|\vec{\xi}_n - \vec{C}\| < \varepsilon \,\} \geqslant \mathbf{P}\{\, \|\vec{\xi}_0 - \vec{C}\| < \varepsilon \,\} = 1.$$

III) При $\forall M > 0$ функция $h(x) = -M \cdot \dot{\mathrm{I}}_{(-\infty;-M)}(x) + x \cdot \dot{\mathrm{I}}_{[-M;M]}(x) + M \cdot \dot{\mathrm{I}}_{(M;-\infty)}(x)$ непрерывна и ограничена. Поэтому $\mathbf{E}\, h(\xi_n) \to \mathbf{E}\, h(\xi_0)$. Так как $|x - h(x)| \leqslant |x|$, то для любой сл.в. ζ в силу неравенства треугольника

$$\left|\mathbf{E}\,\zeta - \mathbf{E}\, h(\zeta)\right| = \left|\mathbf{E}\left[(\zeta - h(\zeta))\, \dot{\mathrm{I}}(|\zeta| \geqslant M)\right]\right| \leqslant \mathbf{E}\left[|\zeta|\, \dot{\mathrm{I}}(|\zeta| \geqslant M)\right].$$

Поэтому (снова после применения неравенства треугольника)

$$|\mathbf{E}\,\xi_n - \mathbf{E}\,\xi_0| \leqslant \big|\mathbf{E}\,\xi_n - \mathbf{E}\,h(\xi_n)\big| + \big|\mathbf{E}\,h(\xi_n) - \mathbf{E}\,h(\xi_0)\big| + \big|\mathbf{E}\,\xi_0 - \mathbf{E}\,h(\xi_0)\big| \leqslant$$
$$\leqslant \sup_n \mathbf{E}\left[|\xi_n|\,\dot{\mathrm{I}}(|\xi_n| \geqslant M)\right] + \big|\mathbf{E}\,h(\xi_n) - \mathbf{E}\,h(\xi_0)\big| + \mathbf{E}\left[|\xi_0|\,\dot{\mathrm{I}}(|\xi_0| \geqslant M)\right].$$

Здесь первое и третье слагаемые стремятся к нулю при $M \to \infty$ в силу равномерной интегрируемости $\langle \xi_n \rangle$ и $\langle \xi_0 \rangle$. Второе слагаемое стремится к нулю при $n \to \infty$ по критерию слабой сходимости ($\mathrm{W_c}$). ⫤

⚠ Как видно из доказательства III), условие равномерной интегрируемости может быть ослаблено. Вместо этого условия достаточно потребовать, чтобы последовательность ξ_n была асимптотически равномерно интегрируема:

$$\lim_M \overline{\lim_n}\, \mathbf{E}\left[|\xi_n|\,\dot{\mathrm{I}}(|\xi_n| \geqslant M)\right] = 0.$$

Можно показать, что это условие является не только достаточным, но и необходимым для сходимости математических ожиданий.

✧ **Лемма Слуцкого. Дельта-метод.** Теоремы 253 и 254, вку́пе с теоремой о непрерывных преобразованиях, позволяют сформулировать одно из наиболее востребованных, особенно в математической статистике, утверждений.

255| **Лемма.** [*Е. Слуцкий.*] *Пусть* $\xi_n \rightsquigarrow \xi_0$, $\eta_n \rightsquigarrow C$, *где* C — *некоторая константа, тогда имеют место следующие соотношения:*

(a) $\xi_n + \eta_n \rightsquigarrow \xi_0 + C$, (b) $\xi_n\eta_n \rightsquigarrow C\xi_0$, (c) $\dfrac{\xi_n}{\eta_n} \rightsquigarrow \dfrac{\xi_0}{C}$,

если в (c) *константа* $C \neq 0$ *и значения сл.в.* $\eta_n(\omega) \neq 0$ *для* $\forall\, n$ *и* $\forall\, \omega \in \Omega$.

256| Пример. Пусть $\langle \xi_j \rangle_1^\infty$ — независимые одинаково распределённые сл.в. с конечным математическим ожиданием μ и конечной дисперсией σ^2. По закону больших чисел Хинчина 297, стр. 258, при $n \to \infty$

$$S_n^2 = \frac{1}{n}\sum_1^n \xi_j^2 - \Big(\frac{1}{n}\sum_1^n \xi_j\Big)^2 \xrightarrow{\mathbf{P}} \mathbf{E}\,\xi_1^2 - \mu^2 = \sigma^2,$$

т.е. отношение $S_n^2/\sigma^2 \xrightarrow{\mathbf{P}} 1$. С другой стороны, по центральной предельной теореме для среднего арифметического $\overline{X}_n = \frac{1}{n}\sum_1^n \xi_j$ имеем

$$\sqrt{n}\frac{(\overline{X}_n - \mu)}{\sigma} \rightsquigarrow \mathcal{N}(0,1).$$

Таким образом, по лемме Слуцкого статистика Стьюдента имеет асимптотически нормальное распределение:

$$\sqrt{n}\frac{(\overline{X}_n - \mu)}{\sqrt{S_n^2}} \rightsquigarrow \mathcal{N}(0,1).$$

Пусть, кроме того, сл.в. ξ_j имеют конечный четвёртый момент и $\varkappa^2 = \mathbf{D}\left[(\xi_j - \mu)^2\right]$. Легко проверить, что для выборочной дисперсии справедливо представление

$$S_n^2 - \sigma^2 \;=\; \frac{1}{n}\sum_1^n \left[(\xi_j - \mu)^2 - \sigma^2\right] - \left(\overline{X}_n - \mu\right)^2.$$

Математическое ожидание $\mathbf{E}\left[\sqrt{n}\left(\overline{X}_n - \mu\right)^2\right] = \sigma^2/\sqrt{n} \to 0$, поэтому $\sqrt{n}\left(\overline{X}_n - \mu\right)^2 \xrightarrow{\mathbf{P}} 0$. Применяя центральную предельную теорему (к независимым одинаково распределённым сл.в. $(\xi_j - \mu)^2 - \sigma^2, j = 1, 2, \ldots$), получаем

$$\sqrt{n}\,\frac{S_n^2 - \sigma^2}{\varkappa} \rightsquigarrow \mathcal{N}(0,1),$$

снова в силу леммы Слуцкого. ◎

Дельта-метод даёт удобный способ нахождения асимптотического распределения произвольных преобразований элементов последовательности.

257| **Теорема.** [*Дельта-метод.*] *Пусть функция $h(x)$, $x \in \mathbb{R}^1$, дифференцируема в точке $x = a$ и $h'(a) \neq 0$. Если последовательность сл.в. $(\xi_n - a)q_n \rightsquigarrow \xi_0$ с числовой последовательностью $q_n \underset{n}{\to} \infty$, то*

$$\frac{q_n}{h'(a)}\big(h(\xi_n) - h(a)\big) \rightsquigarrow \xi_0.$$

⊫ Покажем сначала, что $\xi_n \xrightarrow{\mathbf{P}} a$. Пусть $\varepsilon > 0$ и точка $M > 0$ является точкой непрерывности ф.р. ξ_0, причём $\mathbf{P}\{\xi_0 > M\} \leqslant \varepsilon$. Тогда для всех n таких, что $\varepsilon q_n > M$

$$\mathbf{P}\{\xi_n - a > \varepsilon\} \;=\; \mathbf{P}\{q_n(\xi_n - a) > \varepsilon q_n\} \;\leqslant\; \mathbf{P}\{q_n(\xi_n - a) > M\}.$$

Следовательно, $\overline{\lim}_n \mathbf{P}\{\xi_n - a > \varepsilon\} \leqslant \mathbf{P}\{\xi_0 > M\} \leqslant \varepsilon$. Аналогичными рассуждениями получаем $\overline{\lim}_n \mathbf{P}\{\xi_n - a < -\varepsilon\} \leqslant \varepsilon$, что и требовалось.

Определим непрерывную в точке $x = a$ функцию

$$H(x) \;=\; \frac{1}{h'(a)}\frac{h(x) - h(a)}{x - a}\,\dot{\mathrm{I}}_{x \neq a}(x) + \dot{\mathrm{I}}_{x = a}(x).$$

По теореме 218, стр. 200, $H(\xi_n) \xrightarrow{\mathbf{P}} 1$. Применяя лемму Слуцкого, получаем

$$\frac{q_n}{h'(a)}\big(h(\xi_n) - h(a)\big) \;=\; H(\xi_n)(\xi_n - a)q_n \;\rightsquigarrow\; \xi_0. \qquad \Lleftarrow$$

258| Примеры. 1) Пусть $\xi_1, \ldots, \xi_n$, $n = 2k+1$, — выборка независимых наблюдений из непрерывного распределения F, $\mathrm{med}_n = \xi_{(k+1)}$ — выборочная медиана. Как следует из леммы 40, стр. 53, $F\big(\mathrm{med}_n\big)$ есть выборочная медиана n независимых реализаций равномерных $\mathcal{U}\mathtt{n}(0,1)$ сл.в. Поэтому в силу утверждений примера 134, стр. 128, $\big(F\big(\mathrm{med}_n\big) - \frac{1}{2}\big)2\sqrt{n} \rightsquigarrow \mathcal{N}(0,1)$. Предположим, что ф.р. F имеет обратную F^{-1}, дифференцируемую в точке $t = {}^1\!/_2$: $(F^{-1})'({}^1\!/_2) = 1/F'\big(F^{-1}({}^1\!/_2)\big) = 1/f(\mathfrak{m})$, где $f = F'$, и по определению $\mathfrak{m} = F^{-1}({}^1\!/_2)$ совпадает с медианой распределения F. Итак, если $f(\mathfrak{m}) \neq 0$, то для выборочной медианы имеет место (в неформальной записи)

$$\mathrm{med}_n \rightsquigarrow \mathcal{N}(\mathfrak{m}, \sigma_n^2)$$

с асимптотической дисперсией $\sigma_n^2 = 1/4f^2(\mathfrak{m})\,n$.

2) Пусть $\big\langle \xi_j \big\rangle_1^n$ — независимые сл.в. с общей (показательной) ф.пл. $\theta e^{-\theta x}$, $x \geqslant 0$. Оценка максимального правдоподобия для неизвестного параметра θ равна $\widehat{\theta}_n = 1/\overline{X}_n$, где $\overline{X}_n = \frac{1}{n}\sum_1^n \xi_n$. Поскольку $\mathbf{E}\,\xi_1 = 1/\theta$, $\mathbf{D}\xi_1 = 1/\theta^2$ (см. стр. 107), то по центральной предельной теореме $q_n(\overline{X}_n - a) \rightsquigarrow \mathcal{N}(0,1)$ с $q_n = \theta\sqrt{n}, a = 1/\theta$. Применяя дельта-метод к функции $h(x) = 1/x$, получаем $\sqrt{n}\big(\widehat{\theta}_n - \theta\big)/\theta \;\rightsquigarrow\; \mathcal{N}(0,1)$. ◎

259| Упр. I) Покажите, что если в условиях теоремы первая производная $h'(a) = 0$, а вторая производная $h''(a) \neq 0$, то $\frac{2q_n^2}{h''(a)}\big(h(\xi_n) - h(a)\big) \rightsquigarrow \xi_0^2$.

II) Пусть $\xi_n \rightsquigarrow \mathcal{N}\big(\mu, \sigma_n^2\big)$, $\sigma_n \to 0$, т.е. $(\xi_n - \mu)/\sigma_n \rightsquigarrow \mathcal{N}\big(0,1\big)$, найдите асимптотическое распределение ξ_n^2 при а) $\mu \neq 0$ и б) $\mu = 0$.

Классические предельные теоремы для одномерных сл.в. можно переформулировать в духе слабой сходимости. К таким относится, например, теорема Пуассона (1), стр. 18, утверждающая, что для сл.в. $\xi_n \backsim \mathcal{B}\mathtt{in}(n, \lambda/n)$, $n = 1, 2\ldots$, предел $\lim_n \mathbf{P}\left\{\xi_n = m\right\} = \mathbf{P}\left\{\xi_0 = m\right\}$, $m \geqslant 0$, где $\xi_0 \backsim \mathcal{P}\mathtt{ois}(\lambda)$ — пуассоновская сл.в.

260| Упр. Докажите, что из теоремы Пуассона следует утверждение о слабой сходимости $\mathcal{B}\mathtt{in}(n, \lambda/n) \rightsquigarrow \mathcal{P}\mathtt{ois}(\lambda)$.

Вообще, для дискретных распределений такой метод доказательства слабой сходимости абсолютно естествен.

261| **Теорема.** *Пусть* $\mathscr{X}$ *— конечное или счётное множество и* $\mathbf{P}\{\,\xi_n \in \mathscr{X}\,\} = 1$ *для* $\forall n \geqslant 0$. *Если* $\lim_n \mathbf{P}\{\,\xi_n = x\,\} = \mathbf{P}\{\,\xi_0 = x\,\}$ *для* $\forall x \in \mathscr{X}$, *то последовательность* $\xi_n \rightsquigarrow \xi_0$.

⊨ Пусть $\mathscr{X} = \{x_k,\ \ k = \overline{1,K}\}$, $K \leqslant \infty$. Тогда для любой неотрицательной функции $h(x)$ и любого $M < K$ (если $K < \infty$, то $M = K$)

$$\mathbf{E}\,h(\xi_n) \;=\; \sum_{k=1}^{K} h(x_k)\,\mathbf{P}\{\,\xi_n = x_k\,\} \;\geqslant\; \sum_{k=1}^{M} h(x_k)\,\mathbf{P}\{\,\xi_n = x_k\,\}.$$

В последней сумме $M < \infty$, поэтому здесь можно положить $n \to \infty$:

$$\varliminf_{n} \mathbf{E}\,h(\xi_n) \;\geqslant\; \sum_{k=1}^{M} h(x_k) \lim_{n} \mathbf{P}\{\,\xi_n = x_k\,\} \;=\; \sum_{k=1}^{M} h(x_k)\,\mathbf{P}\{\,\xi_0 = x_k\,\}.$$

Полагая $M \to K$, окончательно получаем $\varliminf_{n} \mathbf{E}\,h(\xi_n) \geqslant \mathbf{E}\,h(\xi_0)$, что гарантирует слабую сходимость по критерию $(\mathrm{W_p})$. ⫤

✧ Сходимость по вариации. Лемма Шеффе. Часто слабая сходимость $\xi_n \rightsquigarrow \xi_0$ эквивалентна сходимости $\mathbf{P}\{\xi_n \in B\} \to \mathbf{P}\{\xi_0 \in B\}$ для $\forall B \in \mathscr{B}$. Если эта сходимость равномерна по $B \in \mathscr{B}$, то говорят о сходимости по вариации.

Определения. Пусть ξ, η — случайные элементы, принимающие значения в измеримом пространстве $(\mathcal{D}, \mathscr{B})$. *Полной вариацией* разности их распределений называется величина $\mathcal{V}(\xi,\eta) = \sup\limits_{B\in\mathscr{B}} \bigl|\,\mathbf{P}\{\xi \in B\} - \mathbf{P}\{\eta \in B\}\,\bigr|$.

Последовательность $\xi_n \in \mathcal{D}$ *сходится по вариации* к случайному элементу ξ_0: $\xi_n \xrightarrow{\text{var}} \xi_0$, если $\lim_n \mathcal{V}(\xi_n, \xi_0) = 0$.

262| **Следствие.** *Если* $\xi_n \xrightarrow{\text{var}} \xi_0$, *то* $\xi_n \rightsquigarrow \xi_0$. *Более того, если* $\xi_n \in \mathbb{R}^k$ *с ф.р.* F_n, $n \geqslant 0$, *то* $\sup\limits_{\vec{x}} \bigl|F_n(\vec{x}) - F_0(\vec{x})\bigr| \to 0, \qquad n \to \infty$.

⚠ Сравните с утверждением теоремы 248, стр. 219.

Для сходимости по вариации достаточно поточечной сходимости соответствующих плотностей (функций вероятностей).

263| **Лемма.** [*Шеффе.*] *Пусть* f_n *— плотность (производная Радона–Никодима) распределения сл.элемента* ξ_n, $n \geqslant 0$, *относительно меры* μ. *Последовательность* $\xi_n \xrightarrow{\text{var}} \xi_0$, *если* $f_n(x) \to f_0(x)$ *для* μ*-почти всех* x.

$\Rrightarrow$ По условию теоремы μ-почти всюду

$$0 \leqslant \left(f_n + f_0 - \left|f_n - f_0\right|\right) \to 2f_0,$$

где неравенство очевидно. Следовательно, в силу леммы Фату 537, стр. 481, и по свойству плотности вероятностной меры ($\int_{\mathcal{D}} f_n\, d\mu = 1,\ \forall\, n \geqslant 0$) имеем

$$2 \ = \ \int\limits_{\mathcal{D}} 2f_0\, d\mu \ \leqslant \ \varliminf_n \int\limits_{\mathcal{D}} \left(f_n \ + \ f_0 \ - \ \left|f_n \ - \ f_0\right|\right) d\mu \ = \ 2 \ - \ \varlimsup_n \int\limits_{\mathcal{D}} \left|f_n \ - \ f_0\right| d\mu.$$

Поэтому $0 \leqslant \varliminf\limits_n \int_{\mathcal{D}} \left|f_n - f_0\right| d\mu \leqslant \varlimsup\limits_n \int_{\mathcal{D}} \left|f_n - f_0\right| d\mu \leqslant 0$. Для завершения доказательства нам понадобится утверждение следующей леммы. $\Lleftarrow$

264| **Лемма.** *Если f_1, f_2 — плотности распределений сл.элементов ξ_1, ξ_2 относительно меры μ, то полная вариация*

$$\mathcal{V}(\xi_1, \xi_2) = \frac{1}{2} \int_{\mathcal{D}} \left|f_1 - f_2\right| d\mu.$$

$\Rrightarrow$ Рассмотрим множества $A_1 = \left\{x : f_1(x) \geqslant f_2(x)\right\}$ и $A_2 = \left\{x : f_2(x) \geqslant f_1(x)\right\}$. Для любого множества B справедливы соотношения:

$$\int_B (f_1 - f_2)\, d\mu \ = \ \int_{BA_1} (f_1 - f_2)\, d\mu + \int_{BA_2} (f_1 - f_2)\, d\mu \ \leqslant \ \int_{A_1} (f_1 - f_2)\, d\mu.$$

И, аналогично, $\int_B (f_1 - f_2)\, d\mu = -\int_B (f_2 - f_1)\, d\mu \geqslant -\int_{A_2} (f_2 - f_1)\, d\mu$.

Заметим теперь, что $\int_{A_1} (f_1 - f_2)\, d\mu = \int_{A_2} (f_2 - f_1)\, d\mu$. Действительно,

$$\int_{A_1} (f_1 - f_2)\, d\mu - \int_{A_2} (f_2 - f_1)\, d\mu \ = \ \int_{A_1 \cup A_2} (f_1 - f_2)\, d\mu \ = \ \int_{\mathcal{D}} (f_2 - f_1)\, d\mu \ = \ 0.$$

С другой стороны, сумма этих интегралов $\int_{A_1} (f_1 - f_2)\, d\mu + \int_{A_2} (f_2 - f_1)\, d\mu = \int_{\mathcal{D}} \left|f_1 - f_2\right| d\mu$, что и доказывает лемму. $\Lleftarrow$

265| Примеры. 1) Плотность распределения Стьюдента

$$f_k(x) \ = \ \frac{\Gamma((k+1)/2)}{\sqrt{\pi k}\, \Gamma(k/2)} \left(1 + \frac{x^2}{k}\right)^{-\frac{1}{2}(k+1)}, \qquad x \in \mathbb{R}^1.$$

При увеличении числа степеней свободы k ф.пл. $f_k(x)$ поточечно сходится к плотности нормального $(0,1)$ закона. Следовательно, при $k \to \infty$ сл.в. ζ_k с плотностью f_k асимптотически нормальна: $\zeta_k \rightsquigarrow \mathcal{N}(0,1)$.

2) Распределение с плотностью $f_n(x) = 1 - \cos(2\pi n x)$, $x \in [0;1]$, сходится

слабо к равномерному на $[0;1]$ распределению:

$$F_n(x) = x - \frac{1}{2\pi n}\sin(2\pi n x) \underset{n}{\to} x, \quad 0 \leqslant x \leqslant 1.$$

Однако очевидно, что для $\forall\, x \in (0;1)$ предел $\lim_n f_n(x)$ не существует. ◎

266| Упр. Докажите утверждение об асимптотической нормальности распределения Стьюдента, исходя из определения стьюдентовской сл.в. $\zeta_k = \sqrt{k}\,\xi/\sqrt{\eta_k}$, где независимые сл.в. $\xi \backsim \mathcal{N}(0,1)$, $\eta_k \backsim \mathcal{G}\mathrm{am}(k/2, 2)$.

В качестве доминирующей меры μ в теореме Шеффе может выступать как мера Лебега — и мы имеем утверждение, аналогичное утверждению об асимптотической нормальности распределения Стьюдента, так и, например, считающая мера — и мы получим доказательство утверждения 261, стр. 226, в более сильном, равномерном варианте.

267| **Следствие.** *Пусть $\mathscr{X}$ — конечное или счётное подмножество $\mathcal{D}$ и $\mathbf{P}\{\xi_n \in \mathscr{X}\} = 1$ для $\forall\, n \geqslant 0$. Если $\lim_n \mathbf{P}\{\xi_n = x\} = \mathbf{P}\{\xi_0 = x\}$ для $\forall\, x \in \mathscr{X}$, то $\xi_n \xrightarrow{\text{var}} \xi_0$; в частности, для сл.векторов $\xi_n = \vec{\xi}_n \in \mathbb{R}^k$*

$$\lim_n \sup_{\vec{x}\in\mathbb{R}^k} \left|\mathbf{P}\{\vec{\xi}_n \leqslant \vec{x}\} - \mathbf{P}\{\vec{\xi}_0 \leqslant \vec{x}\}\right| = 0.$$

⊨ Пусть μ — мера, считающая точки $\mathcal{X}$, т.е. мера любого подмножества $B \subset \mathcal{D}$ равна количеству точек $\mathcal{X}$, попавших в B. Если $\mu(B) = 0$, то $B \cap \mathscr{X} = \varnothing$ и $\mathbf{P}\{\vec{\xi}_n \in B\} = 0$, т.е. распределение ξ_n абсолютно непрерывно относительно μ с плотностью (производной Радона–Никодима) $f_n(x) = \mathbf{P}\{\xi_n = x\}$ для $\forall\, n \geqslant 0$. По условию $f_n(x) \to f_0(x)$ при $\forall\, x \in \mathcal{D}$ ($f_n(x) = 0$ при $x \notin \mathscr{X}$), что доказывает следствие ввиду леммы Шеффе. ⫤

268| Примеры. 1) Если $\mathbf{P}\{\xi_n = 0\} = \mathbf{P}\{\xi_n = 1/n\} = 1/2$, то $\xi_n \rightsquigarrow \xi_0 \equiv 0$. Для этой последовательности нарушено условие предыдущего следствия — $\mathbf{P}\{\xi_n = 0\} = 1/2 \not\to \mathbf{P}\{\xi_0 = 0\} = 1$. В точке разрыва $x = 0$ ф.р. сл.в. ξ_0 ф.р. $F_n(0) = \mathbf{P}\{\xi_n \leqslant 0\} = 1/2 \not\to F_0(0) = \mathbf{P}\{\xi_0 \leqslant 0\} = 1$.

2) Если $\mathbf{P}\{\xi_n = 1/n\} = 1 - \mathbf{P}\{\xi_n = 0\} = 1/n$, то, выбирая $\mathscr{X} = \langle 0, 1, 1/2, 1/3, \ldots \rangle$, имеем $\xi_n \xrightarrow{\text{var}} 0$. Легко видеть непосредственно, что максимум расхождения $\sup\limits_x \left|F_n(x) - F_0(x)\right| = 1/n \underset{n}{\to} 0$. ◎

Хотя расстояние по вариации в большей степени относится не к самим сл.в., а к их распределениям, в ситуации, когда сл.в. заданы на одном вероятностном пространстве, это расстояние может быть оценено сверху вероятностью несовпадения самих сл.в.

269| **Лемма.** [*Неравенство каплинга.*] *Для любых случайных элементов* $\xi_1, \xi_2 : (\Omega, \mathscr{F}, \mathbf{P}) \mapsto (\Upsilon, \mathscr{A})$

$$\sup_{A \in \mathscr{A}} \big| \mathbf{P}\{\xi_1 \in A\} - \mathbf{P}\{\xi_2 \in A\} \big| \leqslant \mathbf{P}\{\xi_1 \neq \xi_2\}.$$

⇛ Для любого измеримого подмножества $A \in \mathscr{A}$ и $j = 1, 2$

$$\mathbf{P}\{\xi_j \in A\} = \mathbf{P}\{\xi_j \in A, \xi_1 = \xi_2\} + \mathbf{P}\{\xi_j \in A, \xi_1 \neq \xi_2\} =$$
$$= \mathbf{P}\{\xi_{3-j} \in A, \xi_1 = \xi_2\} + \mathbf{P}\{\xi_j \in A, \xi_1 \neq \xi_2\} \leqslant \mathbf{P}\{\xi_{3-j} \in A\} + \mathbf{P}\{\xi_1 \neq \xi_2\},$$

что и доказывает требуемое неравенство. ⇚

270| Пример. Заметим, что левая и правая части неравенства каплинга 269 (от *англ.* coupling — сцепка) могут относиться к совершенно разным сл.в, лишь бы соответствующие сл.в. имели одинаковые распределения. Можно попытаться так подобрать сл.в. в правой части неравенства, чтобы вероятность их несовпадения была по возможности наименьшей. Рассмотрим для примера сл.в. $\xi \backsim \mathcal{B}\mathtt{in}(n, q)$, $\eta \backsim \mathcal{P}\mathtt{ois}(\lambda)$ с константой $\lambda = nq$. Пусть $\big\langle (\gamma_j, \zeta_j) \big\rangle_1^n$ — независимые одинаково распределённые сл.векторы, заданные на одном вероятностном пространстве, причём $\gamma_1 \backsim \mathcal{B}\mathtt{ern}(q)$, $\zeta_1 \backsim \mathcal{P}\mathtt{ois}(q)$. В соответствии с теоремой 126, стр. 122,

$$\xi \backsim \widetilde{\xi} = \sum\nolimits_1^n \gamma_j \quad \text{и} \quad \eta \backsim \widetilde{\eta} = \sum\nolimits_1^n \zeta_j.$$

По неравенству каплинга и свойству полуаддитивности вероятности

$$\sup_{B \in \mathscr{B}} \big| \mathbf{P}\{\xi \in B\} - \mathbf{P}\{\eta \in B\} \big| \leqslant \mathbf{P}\Big\{\widetilde{\xi} \neq \widetilde{\eta}\Big\} \leqslant$$
$$\leqslant \sum\nolimits_1^n \mathbf{P}\{\gamma_j \neq \zeta_j\} = n\,\mathbf{P}\{\gamma \neq \zeta\},$$

где сл.вектор $(\gamma, \zeta) \backsim (\gamma_1, \zeta_1)$. Зададим сл.в. γ, ζ на пространстве $\Omega = [0; 1]$ с мерой Лебега на борелевской σ-алгебре. Определим бернуллиевскую сл.в. как $\gamma(\omega) = 0$, если $0 \leqslant \omega \leqslant 1 - q$, и $\gamma(\omega) = 1$ в противном случае. Для определения пуассоновской сл.в. рассмотрим последовательность $P_k = \sum_0^k q^j e^{-q}/j!$, $P_{-1} = 0$, $k = 0, 1 \ldots$, и положим $\zeta(\omega) = k$, если $P_{k-1} \leqslant \omega \leqslant P_k$. Так как $P_0 = e^{-q} \geqslant 1 - q$, то эти сл.в. совпадают при $0 \leqslant \omega \leqslant 1 - q$ $(\gamma = \zeta = 0)$ или $P_0 \leqslant \omega \leqslant P_1$ $(\gamma = \zeta = 1)$. Таким образом,

$$\mathbf{P}\{\gamma \neq \zeta\} = 1 - \big(1 - \frac{\lambda}{n} + \frac{\lambda}{n} e^{-\lambda/n}\big) = \frac{\lambda}{n}\big(1 - e^{-\lambda/n}\big) \leqslant \frac{\lambda^2}{n^2}.$$

Следовательно, в условиях теоремы Пуассона расстояние по вариации между биномиальным и пуассоновским распределением

$$\sup_{B\in\mathscr{B}} \left| \mathbf{P}\{\xi\in B\} - \mathbf{P}\{\eta\in B\} \right| \leqslant \frac{\lambda^2}{n}.$$

Численные эксперименты показывают, что для $n \geqslant 10$, $\lambda \geqslant 1$ константу λ^2 в этом неравенстве можно заменить на $\lambda/5$. ◎

§4. Метод характеристических функций. Теорема П. Леви

Слабая сходимость доказывается чаще всего при помощи метода характеристических функций. Чтобы воспользоваться этим методом, необходимо установить связь между слабой сходимостью сл.в. (ф.р.) и их хар.ф.

271| **Теорема.** [*О непрерывном соответствии. Поль Леви.*] *Пусть $\langle \varphi_n(\vec t),\ n \geqslant 1 \rangle$ — последовательность хар.ф., $\langle F_n(\vec x),\ n \geqslant 1 \rangle$ — соответствующая ей последовательность ф.р.*

I) *Если $F_n \Rightarrow F_0$ с некоторой ф.р. $F_0(\vec x)$, то для $\forall \vec t \in \mathbb{R}^k$ существует предел $\lim_n \varphi_n(\vec t) = \varphi_0(\vec t)$, где φ_0 — хар.ф. F_0.*

II) *Если для $\forall \vec t \in \mathbb{R}^k$ существует предел $\lim_n \varphi_n(\vec t) = \varphi_0(\vec t)$, причём функция $\varphi_0(\vec t)$ непрерывна в точке $\vec t = \vec 0$, то здесь существует ф.р. $F_0(\vec x)$ такая, что $F_n \Rightarrow F_0$ и φ_0 — хар.ф. F_0.*

272| Упр. Докажите утверждение I).

⚠ Зачастую можно сразу видеть, что предельная функция φ_0 есть хар.ф. некоторого вероятностного закона — проверка её непрерывности в этом случае не требуется. Этот вариант теоремы мы и докажем сейчас. Этапы доказательства общего случая будут рассмотрены ниже.

$\Rrightarrow$ Пусть $\vec\xi_n \backsim F_n$ — сл.векторы (вообще говоря, фиктивные) с ф.р. F_n и хар.ф. φ_n. Рассмотрим произвольную финитную функцию h ($|h| < H$), равную нулю вне некоторого множества $S \subset \mathbb{R}^k$. Применим приём сглаживания, использованный при доказательстве свойства единственности для хар.ф.

(см. стр. 164). Справедливо равенство

$$\begin{aligned}\mathbf{E}\,h(\vec{\xi}_n) - \mathbf{E}\,h(\vec{\xi}_0) \;=\; \Big(\mathbf{E}\,h(\vec{\xi}_n) - \mathbf{E}\,h(\vec{\xi}_n + \sigma\vec{\eta})\Big) + \\ + \Big(\mathbf{E}\,h(\vec{\xi}_0 + \sigma\vec{\eta}) - \mathbf{E}\,h(\vec{\xi}_0)\Big) + \Big(\mathbf{E}\,h(\vec{\xi}_n + \sigma\vec{\eta}) - \mathbf{E}\,h(\vec{\xi}_0 + \sigma\vec{\eta})\Big),\end{aligned} \tag{6}$$

где $\vec{\eta} \backsim \mathcal{N}_k(\vec{0}, \dot{\mathbb{I}})$ — стандартный нормальный вектор, который не зависит от $\vec{\xi}_n$ для $\forall\, n \geqslant 0$, параметр $\sigma > 0$ (в дальнейшем $\sigma \to 0$). Терминология, связанная со сл.векторами, применена здесь исключительно с целью упрощения записи — последнее соотношение есть соотношение между интегралами относительно соответствующих ф.р.

Зафиксируем $\varepsilon > 0$. Поскольку финитная функция равномерно непрерывна, то найдётся такое $\delta > 0$, что $\big|h(\vec{x}) - h(\vec{x} + \vec{y})\big| < \varepsilon$ при любых $\|\vec{y}\| < \delta$ и любых $\vec{x}$. Выберем $\sigma > 0$ так, чтобы $\mathbf{P}\{\|\vec{\eta}\| \geqslant \delta/\sigma\} < \varepsilon/(2H)$. Тогда, разбивая математическое ожидание на области $\{\|\vec{\eta}\| < \delta/\sigma\}$ и $\{\|\vec{\eta}\| \geqslant \delta/\sigma\}$, получаем по неравенству треугольника для всех $n \geqslant 0$, что

$$\Big|\mathbf{E}\left[h(\vec{\xi}_n) - h(\vec{\xi}_n + \sigma\vec{\eta}\right]\Big| \;\leqslant\; \varepsilon + 2H\,\mathbf{P}\{\|\vec{\eta}\| \geqslant \delta/\sigma\} \;<\; 2\varepsilon,$$

т.е. модуль суммы первых двух слагаемых (6) не превосходит 4ε.

Для анализа третьего слагаемого (6) обозначим плотность $\vec{\eta}$ через $\psi(\vec{t}\,) = \exp(-\vec{t}^{\,\flat}\,\vec{t}/2)/(\sqrt{2\pi})^k$. Функция $\psi(\vec{t}\,)(\sqrt{2\pi})^k$ совпадает с хар.ф. $\vec{\eta}$ (теорема 192, стр. 185), поэтому $\psi(\vec{t}\,) = (\sqrt{2\pi})^{-k}\int_{\mathbb{R}^k} e^{\mathtt{i}\vec{t}^{\flat}\,\vec{u}}\,\psi(\vec{u})\,d\vec{u}$ и, следовательно,

$$\begin{aligned}\mathbf{E}\,h(\vec{\xi}_n + \sigma\vec{\eta}) \;&=\; \int\limits_{\mathbb{R}^k}\int\limits_{\mathbb{R}^k} h(\vec{x} + \sigma\vec{y})\psi(\vec{y})\,d\vec{y}\,dF_n(\vec{x}) \;= \\ &=\; \frac{1}{\sigma^k}\int\limits_{\mathbb{R}^k}\int\limits_{S} h(\vec{y})\psi\Big(\frac{\vec{y} - \vec{x}}{\sigma}\Big)\,d\vec{y}\,dF_n(\vec{x}) \;= \\ &=\; \frac{1}{(\sqrt{2\pi}\sigma)^k}\int\limits_{\mathbb{R}^k}\int\limits_{S}\int\limits_{\mathbb{R}^k} h(\vec{y})\exp\Big\{\mathtt{i}\,\frac{1}{\sigma}(\vec{y} - \vec{x})^\flat\,\vec{u}\Big\}\psi(\vec{u})\,d\vec{u}\,d\vec{y}\,dF_n(\vec{x}).\end{aligned}$$

Так как S — ограниченное множество, то подынтегральная функция абсолютно интегрируема по $(\vec{u}, \vec{y}, \vec{x}) \in \mathbb{R}^k \times S \times \mathbb{R}^k$ и по теореме Фубини можно заменить порядок интегрирования. Заметим, что по определению $\int_{\mathbb{R}^k}\exp\{-\mathtt{i}\,\frac{1}{\sigma}\,\vec{u}^\flat\,\vec{x}\}\,dF_n(\vec{x}) = \varphi_n(-\vec{u}/\sigma)$. Таким образом, для $\forall\, n \geqslant 0$

$$\mathbf{E}\,h(\vec{\xi}_n + \sigma\vec{\eta}) \;=\; \frac{1}{(\sqrt{2\pi}\sigma)^k}\int_S\int_{\mathbb{R}^k} h(\vec{y})\exp\Big\{\mathtt{i}\,\frac{1}{\sigma}\,\vec{y}^\flat\,\vec{u}\Big\}\varphi_n\big(-\frac{\vec{u}}{\sigma}\big)\psi(\vec{u})\,d\vec{u}\,d\vec{y}.$$

В силу ограниченности h подынтегральная функция в последнем выражении при любых $n \geqslant 1$ не превосходит интегрируемой по области $(\vec{y} \in S) \times (\vec{u} \in \mathbb{R}^k)$ функции $H\psi(\vec{u})/(\sqrt{2\pi}\sigma)^k$. Следовательно, по теореме Лебега об ограниченной сходимости здесь можно перейти к пределу $\lim_n \varphi_n = \varphi_0$. Таким образом, для $\forall\, \sigma > 0$ (в частности, для выбранного ранее) предел $\lim\limits_n \mathbf{E}\, h(\vec{\xi}_n + \sigma\vec{\eta}) = \mathbf{E}\, h(\vec{\xi}_0 + {} + \sigma\vec{\eta})$, т.е. третье слагаемое (6) по модулю может быть сделано меньше ε.

Итак, $\mathbf{E}\, h(\vec{\xi}_n) - \mathbf{E}\, h(\vec{\xi}_0) \to 0$, что по критерию слабой сходимости (W_φ) гарантирует сходимость $\vec{\xi}_n \rightsquigarrow \vec{\xi}_0$. ⫤

✧ Лемма Хелли о выборе. Доказательство теоремы П. Леви в общем случае опирается на ряд вспомогательных утверждений, которые сами по себе столь же важны для теории вероятностей, как и теорема П. Леви.

273| **Лемма.** [*О выборе. Э. Хэлли.*] *Из любого бесконечного семейства ф.р.* $\langle F_\alpha \rangle$ *в* $\mathbb{R}^k$ *можно выделить последовательность*[‡] $F_n \Rightarrow F_0$ *с некоторой обобщённой ф.р.* F_0.

⊨ Рассмотрим одномерный случай, однако, помня о многомерной ситуации, будем везде, где это необходимо, заменять знак строгого неравенства «$>$» на знак строгого векторного неравенства «$\gg$». Занумеруем каким-либо способом все действительные рациональные числа $\mathbb{Q} = \langle q_1, q_2, \dots \rangle$.

Из ограниченного бесконечного семейства чисел $\langle 0 \leqslant F_\alpha(q_1) \leqslant 1 \rangle_\alpha$ по теореме Больцано–Вейерштрасса можно выделить сходящуюся подпоследовательность $F_{1n}(q_1) \underset{n}{\to} y_1 =: G(q_1) \in [0;1]$. Аналогично, из ограниченного набора чисел $\langle F_{1n}(q_2) \rangle$ выделим подпоследовательность $F_{2n}(q_2) \underset{n}{\to} y_2 =: G(q_2) \in {} \in [0;1]$. Заметим, что $F_{2n}(q_1) \underset{n}{\to} G(q_1)$. Продолжая подобным образом, получим систему вложенных последовательностей $\langle F_{(j-1)n} \rangle_{n=1}^\infty \supset \langle F_{jn} \rangle_{n=1}^\infty$, $j = {} = 2, 3 \dots$, и функцию $G(q)$, $q \in \mathbb{Q}$, со следующими свойствами:

(A) $F_{jn}(q_i) \underset{n}{\to} G(q_i)$ для $\forall\, i \leqslant j$;

(B) если $q_i < q_j$ $(\in \mathbb{Q})$, то $0 \leqslant G(q_i) \leqslant G(q_j) \leqslant 1$.

Из свойства (A) следует, что для диагональной последовательности $\langle F_{nn} \rangle_{n=1}^\infty$ предел $\lim_n F_{nn}(q) = G(q)$ для $\forall\, q \in \mathbb{Q}$.

Отталкиваясь от G, определим при всех $x \in \mathbb{R}^1$ функцию

$$F_0(x) = \inf \langle G(q) : \ q \gg x \rangle \tag{7}$$

и покажем, что F_0 — искомая обобщённая ф.р. и $F_{nn} \Rightarrow F_0$.

[‡] Отношение $F_n \Rightarrow F_0$ означает сходимость F_n к F_0 во всех точках непрерывности предельной функции F_0 (в основном). Эквивалентность такой сходимости слабой имеет место только при выполнении условия нормировки всеми ф.р.: $F_n(-\infty) = 1 - F_n(+\infty) = 0, n \geqslant 0$.

Отметим, что по построению F_0 для $\forall x \in \mathbb{R}^1$ найдётся такая последовательность рациональных чисел $q_j \gtrdot x$, $q_j \to x$, что $F_0(x) = \lim_j G(q_j)$. Учитывая монотонность G (свойство (B)), можно сделать вывод, что

$$F_0(x) = \lim_{q \to x+} G(q) \tag{8}$$

для любых рациональных $q \to x$, $q \gtrdot x$. Отсюда сразу следует, что $0 \leqslant F_0 \leqslant 1$. Кроме того, F_0 не убывает, т.к. если $x' < x''$, то для $x'' \lessdot q \to x''$, $q \in \mathbb{Q}$, по определению $F_0(x') \leqslant G(q) \to F_0(x'')$.

Покажем, что F_0 непрерывна справа. Пусть $x \in \mathbb{R}^1$ и $x' \to x$, оставаясь справа: $x' \geqslant x$. По определению инфимума для $\forall \varepsilon > 0$ найдётся рациональное $q \gtrdot x$ такое, что $F_0(x) + \varepsilon > G(q)$. Поэтому из определения F_0 при всех $x \leqslant x' \lessdot q$ имеем $F_0(x) \leqslant F_0(x') \leqslant G(q) < F_0(x) + \varepsilon$, что и требовалось.

Осталось показать, что $F_{nn} \Rightarrow F_0$. Если x — точка непрерывности F_0, то для $\forall \varepsilon > 0$ найдутся действительные точки x', x'' (а по ним произвольные рациональные точки y', y'') такие, что $x' \lessdot y' \lessdot x \lessdot y'' \lessdot x''$ и

$$G(y'') - \varepsilon \underset{(a)}{\leqslant} F_0(x'') - \varepsilon \underset{(b)}{\leqslant} F_0(x) \underset{(c)}{\leqslant} F_0(x') + \varepsilon \underset{(d)}{\leqslant} G(y') + \varepsilon,$$

где неравенства (b), (c) следуют из непрерывности F_0, неравенство (d) — из определения, неравенство (a) — также из определения и монотонности G. Таким образом, ввиду монотонности ф.р. F_{nn}

$$\begin{gathered}\overline{\lim_n} F_{nn}(x) - \varepsilon \leqslant \lim_n F_{nn}(y'') - \varepsilon = G(y'') - \varepsilon \leqslant \\ \leqslant F_0(x) \leqslant G(y') + \varepsilon = \lim_n F_{nn}(y') + \varepsilon \leqslant \underline{\lim_n} F_{nn}(x) + \varepsilon.\end{gathered}$$

Полагая $\varepsilon \to 0$, получаем доказательство теоремы.

Здесь мы использовали только факт, по которому между любыми двумя действительными числами найдётся рациональное число. Аналогичное свойство справедливо и в $\mathbb{R}^k$. Поэтому предыдущее доказательство дословно переносится на многомерный случай, если в множество $\mathbb{Q} = \langle \vec{q} \rangle$ включить все векторы с рациональными координатами. При $\vec{a}, \vec{c} \in \mathbb{Q}$, $\vec{a} \leqslant \vec{c}$, разностный оператор $\Delta_{\vec{a},\vec{c}} G \geqslant 0$ (см. (4), стр. 35) в силу справедливости этого свойства для всех ф.р. F_{nn}. Для функции F_0 это свойство выполняется, т.к. по построению $F_0(\vec{x}) = \lim_{\vec{q}} G(\vec{q})$ для некоторой последовательности $\vec{q} \to \vec{x}$, $\vec{q} \gtrdot \vec{x}$, $\vec{q} \in \mathbb{Q}$. ⫤

274| Упр. Пусть $G(x)$, $x \in \mathbb{R}^1$, — неубывающая функция. Докажите, что функция $F(x) = \inf_{y \geqslant x} G(y)$ не убывает и непрерывна справа.

⚠ Обобщённые распределения можно трактовать как распределения сл.в., принимающих с положительной вероятностью значения $\pm\infty$. Такие сл.в. (распределения) называют ещё *несобственными*. Следующая фундаментальная теорема Ю.В. Прохорова утверждает, что из любого ограниченного по вероятности семейства сл.в. можно выделить подпоследовательность, слабо сходящуюся к сл.в. с собственным распределением.

✧ Теорема Прохорова об относительной компактности.

275| **Теорема.** [*Об относительной компактности. Ю.В.Прохоров.*]

I) *Если* $\vec{\xi}_n \rightsquigarrow \vec{\xi}_0$, *то последовательность* $\left\langle \vec{\xi}_n \right\rangle_1^\infty$ *равномерно плотна (ограничена по вероятности):*

$$\lim_{M\to\infty} \sup_n \mathbf{P}\left\{ \|\vec{\xi}_n\| > M \right\} = 0. \tag{9}$$

II) *Если семейство* $\left\langle \vec{\xi}_n \right\rangle_1^\infty$ *равномерно плотно (ограничено по вероятности), то найдётся сл.вектор* $\vec{\xi}_0$ *и подпоследовательность* $\vec{\xi}_{n_k} \rightsquigarrow \vec{\xi}_0$, $k \to \infty$.

⊫ I) Зададим произвольное $\varepsilon > 0$ и выберем $M_0 > 0$ так, чтобы $\mathbf{P}\left\{ \|\vec{\xi}_0\| \geqslant M_0 - 1 \right\} < \varepsilon$. По критерию $(\mathrm{W_t})$ теоремы о портфеле $\overline{\lim\limits_n}\, \mathbf{P}\left\{ \|\vec{\xi}_n\| \geqslant M_0 - 1 \right\} \leqslant \mathbf{P}\left\{ \|\vec{\xi}_0\| \geqslant M_0 - 1 \right\}$, т.е. начиная с некоторого N вероятность $\mathbf{P}\left\{ \|\vec{\xi}_n\| \geqslant M_0 - 1 \right\} < 2\varepsilon$ для $\forall n \geqslant N$. Для первых N векторов $\vec{\xi}_n, n = \overline{1,N}$, определим $M_n > 0$ так, что $\mathbf{P}\left\{ \|\vec{\xi}_n\| \geqslant M_n \right\} < 2\varepsilon$. Тогда, очевидно, для $\forall M > \max\left\langle M_0, M_1, \ldots, M_N \right\rangle$ при всех $n \geqslant 1$

$$\mathbf{P}\left\{ \|\vec{\xi}_n\| > M \right\} < 2\varepsilon.$$

II) Пусть F_n — ф.р. сл.вектора $\vec{\xi}_n$. При доказательстве леммы Хелли была построена функция $G(\vec{q})$ на множестве $\mathbb{Q}$ точек с рациональными координатами и подпоследовательность (обозначим её тем же индексом) $F_n(\vec{q}) \to G(\vec{q}), \forall \vec{q} \in \mathbb{Q}$. При этом на основе функции G была определена в (7) обобщённая ф.р. F_0, для которой $F_n \Rightarrow F_0$. Покажем, что F_0 — полноценная ф.р., т.е. $F_0(\vec{x}) \to 1$, если все координаты $x_j \to +\infty$, и $F_0(\vec{x}) \to 0$, если хотя бы одна из координат $x_j \to -\infty$.

По условию для $\varepsilon > 0$ найдётся $M > 0$ такое, что $\sup_n \mathbf{P}\left\{ \|\vec{\xi}_n\| > M' \right\} < \varepsilon$ для $\forall M' > M$. Выберем точку с рациональными координатами $\vec{q} \in \mathbb{Q}$ так, чтобы область $\{\vec{x} \leqslant \vec{q}\}$ полностью покрывала шар $\{\|\vec{x}\| \leqslant M\}$. Ясно, что

$\|\vec{q}\| \geqslant M$. Таким образом, $F_n(\vec{q}) \geqslant \mathbf{P}\{\, \|\vec{\xi}_n\| \leqslant M \,\} \geqslant 1 - \varepsilon$ для $\forall\, n$. Следовательно, $G(\vec{q}) \to 1$, если каждая координата $q_j \to +\infty$. Так же показывается, что $G(\vec{q}) \to 0$, если хотя бы одна координата $q_j \to -\infty$. Ввиду (8) отсюда следует, что аналогичное свойство справедливо и для функции F_0.

Таким образом, F_0 есть ф.р., а соответствующий ей сл.вектор $\vec{\xi}_0$ будет удовлетворять утверждению теоремы. ⇚

276| ⚠ Утверждение II) теоремы как раз означает относительную компактность (в слабой топологии) ограниченного по вероятности семейства сл.в. (сравните с теоремой Больцано–Вейерштрасса). Теорема Прохорова верна и для случайных элементов, принимающих значения в некотором метрическом пространстве $\mathcal{D}$. Равномерная плотность произвольного семейства вероятностных мер $\langle\, \mathbf{P}_\alpha \,\rangle$ на борелевской σ-алгебре $\mathcal{D}$ определяется следующим образом:

$$\forall\, \varepsilon > 0 \;\exists \text{ компакт } K \subset \mathcal{D} : \sup_\alpha \mathbf{P}_\alpha\{K^c\} < \varepsilon.$$

Доказательство теоремы Прохорова в общем случае можно найти, например, в [2, стр. 54].

Условие равномерной плотности (9) идентично условию равномерной интегрируемости порядка $q = 0$ (см. (1), стр. 203).

277| **Следствие.** *Если $F_n \Rightarrow F_0$ и при некотором $q > 0$ математическое ожидание $\int_{\mathbb{R}^k} \|\vec{x}\|^q \, dF_n < C < \infty$ для $\forall\, n \geqslant 1$, то F_0 есть ф.р. и $\int_B \|\vec{x}\|^s \, dF_n \to \int_B \|\vec{x}\|^s \, dF_0$ для $\forall\, 0 < s < q$ и $\forall\, B \in \mathscr{B}^k$.*

⇛ По лемме 229, c), стр. 203, последовательность F_n равномерно плотна ввиду отмеченной выше связи с равномерной интегрируемостью порядка 0 (‡). В силу теоремы Прохорова предельная функция F_0 должна быть функцией распределения. Вторая часть утверждения следует из той же леммы 229, стр. 203, и теоремы 254, стр. 222. ⇚

Для завершения доказательства теоремы П. Леви осталось показать, что в условиях теоремы семейство сл.в. равномерно плотно.

278| **Лемма.** *Если последовательность хар.ф. $\varphi_n(\vec{t}\,) \to \varphi_0(\vec{t}\,)$ при любых $\vec{t} \in \mathbb{R}^k$, причём функция φ_0 непрерывна в точке $\vec{t} = \vec{0}$, то семейство сл.в., соответствующих $\langle\, \varphi_n \,\rangle$, равномерно плотно.*

(‡) Доказать равномерную плотность здесь можно также простой ссылкой на неравенство Маркова.

$\Rrightarrow$ Рассмотрим одномерный случай. Воспользуемся следующим неравенством, связывающим поведение распределения сл.в. при больших значениях с поведением хар.ф. в окрестности нуля (доказательство ниже):

$$\mathbf{P}\left\{\left|\xi\right| > M\right\} \leqslant M\int_0^{2/M}\big(1 - \operatorname{Re}\varphi_\xi(t)\big)\,dt. \tag{10}$$

В силу непрерывности φ_0 в нуле (и равенства $\varphi_0(0) = 1$) для $\forall\,\varepsilon > 0$ можно подобрать $0 < \varepsilon' < \varepsilon$ так, чтобы $\sup\limits_{0<t<\varepsilon'}\big(1 - \operatorname{Re}\varphi_0(t)\big) < \varepsilon/4$. Поэтому

$$M\int_0^{2/M}\big(1 - \operatorname{Re}\varphi_0(t)\big)\,dt < \frac{\varepsilon}{2} \qquad \text{для } \forall\,M > M' = \frac{2}{\varepsilon'}.$$

Так как $\varphi_n(t) \underset{n}{\to} \varphi_0(t)$, то по теореме Лебега $\int_0^{2/M'}\big(1 - \operatorname{Re}\varphi_n(t)\big)\,dt \underset{n}{\to}$ $\underset{n}{\to} \int_0^{2/M'}\big(1 - \operatorname{Re}\varphi_0(t)\big)\,dt$ и, следовательно,

$$\mathbf{P}\left\{\left|\xi_n\right| > M'\right\} \leqslant M'\int_0^{2/M'}\big(1 - \operatorname{Re}\varphi_n(t)\big)\,dt < \varepsilon$$

при всех $n > n'$ для некоторого $n' \geqslant 1$.

Ясно, что $\exists M'' > M'$ такое, что $\mathbf{P}\big\{\,|\xi_n| > M''\,\big\} < \varepsilon$ при $1 \leqslant n \leqslant n'$. Таким образом, для $\forall\,\varepsilon > 0$ найдётся такое M'', что $\sup_n \mathbf{P}\big\{\,|\xi_n| > M\,\big\} <$ $< \varepsilon$ при $M > M''$, т.е. семейство $\big\langle\,\xi_n\,\big\rangle_1^\infty$ ограничено по вероятности.

В k-мерном случае вместо (10) можно использовать неравенство

$$\mathbf{P}\big\{\,\|\vec{\xi}\| > M\sqrt{k}\,\big\} \leqslant M\sum\nolimits_1^k\int_0^{2/M}\big(1 - \operatorname{Re}\varphi_{\vec{\xi}}^{(j)}(\vec{t})\big)\,dt_j,$$

где $\varphi_{\vec{\xi}}^{(j)}(\vec{t}) = \varphi(t_1,\ldots,t_k)\big|_{t_i=0,i\neq j}$ — хар.ф. j-й компоненты ξ_j вектора $\vec{\xi}$ в точке t_j, которое следует из (10), т.к. неравенство $\sum_1^k \xi_j^2 > kM^2$ влечёт выполнение неравенства $|\xi_j| > M$ хотя бы при одном $j = 1,\ldots,k$. $\Lleftarrow$

$\Rrightarrow$ Докажем (10). Если F — ф.р., определяющая хар.ф. φ, то $\operatorname{Re}\varphi(t) =$ $= \int_{\mathbb{R}}\cos(tx)\,dF(x)$. По теореме Фубини

$$\frac{2}{\varepsilon}\int\limits_0^\varepsilon\big(1 - \operatorname{Re}\varphi(t)\big)\,dt = \int\limits_{-\infty}^\infty\Big(\frac{2}{\varepsilon}\int\limits_0^\varepsilon\big(1 - \cos(tx)\big)\,dt\Big)\,dF(x) =$$
$$= \int\limits_{-\infty}^\infty 2\Big(1 - \frac{\sin(\varepsilon x)}{\varepsilon x}\Big)\,dF(x) \geqslant \int\limits_{|\varepsilon x|>2} 2\Big(1 - \frac{\sin(\varepsilon x)}{\varepsilon x}\Big)\,dF(x).$$

Так как $2(1 - (\sin y)/y) \geqslant (2 - 2/|y|) \geqslant 1$ в области $|y| \geqslant 2$, то при $\varepsilon = 2/M$

$$M\int_0^{2/M}(1-\operatorname{Re}\varphi(t))\,dt \;\geqslant\; \int_{|x|>M} dF(x) \;=\; \mathbf{P}\{\,|\xi|>M\,\}\,. \qquad \Lleftarrow$$

Теперь мы можем завершить доказательство теоремы П. Леви.

$\Rrightarrow$ В силу предыдущей леммы в условиях теоремы П. Леви последовательность сл.векторов, соответствующих заданным хар.ф., ограничена по вероятности. Поэтому по теореме Прохорова из любой подпоследовательности ф.р. $F_{n'}$ можно выбрать подпоследовательность $F_{n''} \xrightarrow{\text{w}} F''_0$, где F''_0 — ф.р. некоторого сл.вектора $\vec{\xi}''_0$ и $\vec{\xi}_{n''} \rightsquigarrow \vec{\xi}''_0$.

Из части I) теоремы П. Леви следует, что φ_0 — хар.ф. $\vec{\xi}''_0$, следовательно, ф.р. F''_0 не зависит от выбранной подпоследовательности $F_{n''}$. Таким образом, из любой подпоследовательности ф.р. $F_{n'}$ можно выбрать подпоследовательность $F_{n''} \xrightarrow{\text{w}} F_0$, где F_0 — ф.р. некоторого сл.вектора $\vec{\xi}_0$ (с хар.ф. φ_0), что в силу утверждения b) леммы 239, стр. 211, доказывает теорему П. Леви. $\Lleftarrow$

Можно предложить более короткий путь доказательства теоремы Леви, использующий вариант теоремы Хелли–Брэя, описанный в 237, стр. 210.

$\Rrightarrow$ Воспользуемся тем, что функция $e^{-\frac{1}{2\sigma^2}x^2}$ есть хар.ф. (в точке x) нормального $\mathcal{N}(0,\sigma^{-2})$ распределения. Поскольку для $\forall t\in\mathbb{R}^1$, $\sigma>0$ предел $\lim_n \varphi_n(t)\exp\{-\frac{1}{2}t^2\sigma^2\} = \varphi_0(t)\exp\{-\frac{1}{2}t^2\sigma^2\}$, то по теореме Лебега об ограниченной сходимости

$$\begin{aligned}\int_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}x^2}\,dF_n(x) &= \int_{\mathbb{R}}\Big(\int_{\mathbb{R}} e^{\mathrm{i}xt}\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{1}{2}\sigma^2t^2}\,dt\Big)dF_n(x) = \\ &= \frac{\sigma}{\sqrt{2\pi}}\int_{\mathbb{R}}\varphi_n(t)e^{-\frac{1}{2}\sigma^2t^2}\,dt \;\underset{n}{\to}\; \frac{\sigma}{\sqrt{2\pi}}\int_{\mathbb{R}}\varphi_0(t)e^{-\frac{1}{2}\sigma^2t^2}\,dt = \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\varphi_0\Big(\frac{t}{\sigma}\Big)e^{-\frac{1}{2}t^2}\,dt.\end{aligned}$$

Согласно лемме Хелли о выборе найдётся подпоследовательность $F_{n_k} \Rightarrow F_0$, где F_0, возможно, несобственная ф.р. Тогда, по указанному выше варианту предельной теоремы Хелли–Брэя, с учётом предыдущих построений

$$\int_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}x^2}\,dF_0(x) \;=\; \lim_k \int_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}x^2}\,dF_{n_k}(x) \;=\; \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\varphi_0\Big(\frac{t}{\sigma}\Big)e^{-\frac{1}{2}t^2}\,dt.$$

Полагая $\sigma\to\infty$ и учитывая непрерывность φ_0 при $t=0$, по теореме Лебега получаем, что $\int_{\mathbb{R}} dF_0 = 1$, т.е. ф.р. F_0 собственная. По теореме Хелли–Брэя φ_0 есть хар.ф. F_0. Другими словами, из любой подпоследовательности ф.р. $F_{n'}$ можно выделить подпоследовательность, слабо сходящуюся к одной и той же ф.р. F_0, что по лемме 239, стр. 211, гарантирует $F_n \Rightarrow F_0$. ⊲

279| Пример. Установим связь между дискретным геометрическим распределением и непрерывным показательным распределением. Пусть эксперименты в схеме Бернулли происходят часто в моменты времени $k\Delta$, $k = 1, 2, \ldots$, где Δ — малый промежуток времени ($\Delta \approx 0$). При этом вероятность наступления события также мала — $p \approx 0$. Другими словами, в каждый момент времени может произойти некое событие с малой вероятностью. Если это событие приводит к необратимым последствиям (например, к отказу электроприбора), то геометрическая сл.в. ξ описывает время жизни прибора.

Классическая геометрическая сл.в. $\eta \backsim \mathcal{G}\mathrm{eo}(p)$ сосредоточена на множестве натуральных чисел $k = 1, 2, \ldots$ Наша сл.в. связана с ней простым соотношением: $\xi = \eta\Delta$. Понятно, что интересные результаты получаются, только если между вероятностью p и длиной промежутка времени Δ имеется некоторая связь. Пусть $p = \lambda\Delta$ с некоторым $\lambda > 0$. Параметр $\lambda = p/\Delta$, т.е. равен относительной частоте наступления события в единицу времени. Естественно назвать его интенсивностью событий (отказов).

Найдём сначала асимптотическое (при $\Delta \to 0$) распределение ξ. Для этого определим X_Δ как округлённое вверх число x/Δ и вычислим функцию распределения ξ для $x > 0$:

$$F_\xi(x) \;=\; 1 - \mathbf{P}\{\xi > x\} \;=\; 1 - \mathbf{P}\big\{\,\eta \geqslant X_\Delta\,\big\} \;=\; 1 - \sum\nolimits_{k=X_\Delta}^{\infty} (1-p)^{k-1} p.$$

Суммируя бесконечную геометрическую прогрессию, получаем по «второму замечательному пределу», что для $\forall x > 0$ ф.р.

$$F_\xi(x) \;=\; 1 - (1-p)^{X_\Delta - 1} \;\to\; 1 - e^{-\lambda x}$$

при $p = \lambda\Delta$ и $\Delta \to 0$. Таким образом, $\xi \rightsquigarrow \mathcal{E}\mathrm{x}\left({}^1\!/_\lambda\right)$.

Получим теперь этот же результат с помощью теоремы П. Леви. Хар.ф. ξ (см. лемму 160, стр. 161) равна $\varphi_\xi(t) \;=\; \mathbf{E}\, e^{\mathrm{i}t\Delta\eta} = \varphi_\eta(t\Delta) = p/(p - 1 + e^{-\mathrm{i}t\Delta})$. По правилу Лопиталя

$$\lim_{\Delta\to 0} \varphi_\xi(t) \;=\; \lim_{\Delta\to 0} \frac{\lambda\Delta}{\lambda\Delta - 1 + e^{-\mathrm{i}t\Delta}} \;=\; \frac{\lambda}{\lambda - \mathrm{i}t},$$

что совпадает с хар.ф. показательного закона $\mathcal{E}\mathrm{x}\left({}^1\!/_\lambda\right)$.

Заметим, что если при $\Delta \to 0$ отношение $p/\Delta \to \infty$, то предел ф.р. $\lim_\Delta F_\xi(x) = \lim_\Delta \mathbf{P}\{\xi \leqslant x\} = 1$ для $\forall x > 0$ (при $x \leqslant 0$ ф.р. $F_\xi(x) \equiv 0$). Другими словами, если интенсивность отказов слишком велика, то прибор почти сразу выходит из строя. При этом хар.ф. $\varphi_\xi(t) \to 1$ для любого $t \in \mathbb{R}^1$. В этом случае обе функции — и ф.р. F_0, и хар.ф. $\varphi_0 \equiv 1$ — описывают хотя и «собственную», но вырожденную сл.в., сосредоточенную в точке $x = 0$.

Обратно, при $p/\Delta \to 0$ предел ф.р. $\lim_\Delta F_\xi(x) = 0$ для $\forall\, x \geqslant 0$, т.е. прибор никогда не выйдет из строя. Функция $F_0(x) \equiv 0$ обладает всеми свойствами ф.р., кроме свойства $\lim\limits_{x\to\infty} F_0(x) = 1$. Таким образом, предельная функция есть обобщённая, «несобственная» ф.р., сосредоточенная полностью в $+\infty$. В этом случае предел хар.ф. $\lim_\Delta \varphi_\xi(t) = \varphi_0(t)$, где $\varphi_0(t) = 0$ при $t \neq 0$ и $\varphi_0(0) = 1$. Таким образом, предел хар.ф. терпит разрыв в точке $t = 0$. ◎

► В некоторых ситуациях (как показывает пример 265, стр. 227, не во всех) из сходимости хар.ф. можно сделать вывод не только о слабой сходимости ф.р., но и о сходимости соответствующих плотностей. Для этого нужно потребовать, чтобы хар.ф. сходились в смысле $\mathcal{L}_1$. Напомним, что у распределений с абсолютно интегрируемой хар.ф. существует ограниченная непрерывная плотность и справедлива формула обращения (4), стр. 166.

280| **Теорема.** (?!) *Пусть* φ_n, $n \geqslant 0$, — *хар.ф., соответствующие плотностям* f_n, *удовлетворяющие всюду (4), стр. 166. Если* $\int_{\mathbb{R}} |\varphi_n(t) - \varphi_0(t)|\, dt \underset{n}{\to} 0$, *то*

$$\sup_x |f_n(x) - f_0(x)| \underset{n}{\to} 0.$$

Для последовательности произвольных вариантов ф.пл. утверждение теоремы требует уточнения в следующем виде:

$$\operatorname*{ess\,sup}_{\lambda}\ \sup_x |f_n(x) - f_0(x)| \underset{n}{\to} 0,$$

где так называемый существенный (относительно меры Лебега λ) супремум функции $h(x)$, $x \in \mathbb{R}^1$, определяется как $\inf\{C : \lambda\big(h(x) \geqslant C\big) = 0\}$.

► Последовательность нормальных случайных величин может сходиться только к нормальной (быть может, вырожденной) сл.в. Справедлива

281| **Теорема.** *Если для последовательности* $\left\langle\, \vec{\xi}_n \backsim \mathcal{N}_k(\vec{\mu}_n, \mathbf{\Sigma}_n),\ n \geqslant 1 \right\rangle$ *с некоторым набором математических ожиданий* $\vec{\mu}_n$ *и матриц ковариаций* $\mathbf{\Sigma}_n$ *имеет место* $\vec{\xi}_n \rightsquigarrow \vec{\xi}_0$, *то* $\exists\ \vec{\mu}_0 = \lim_n \vec{\mu}_n$, $\mathbf{\Sigma}_0 = \lim_n \mathbf{\Sigma}_n$ *(покомпонентно) и* $\vec{\xi}_0 \backsim \mathcal{N}_k(\vec{\mu}_0, \mathbf{\Sigma}_0)$.

$\Rrightarrow$ По условию теоремы для $\forall \vec{t} \in \mathbb{R}^k$ последовательность хар.ф $\vec{\xi}_n$

$$\exp\{\mathrm{i}\vec{t}^\flat \vec{\mu}_n\} \exp\{-\frac{1}{2}\vec{t}^\flat \boldsymbol{\Sigma}_n \vec{t}\} \;\to\; \mathbf{E}\, e^{\mathrm{i}\vec{t}^\flat \vec{\xi}_0}.$$

Последовательность комплексных чисел имеет предел т. т. т. когда последовательность из аргументов $\vec{t}^\flat \vec{\mu}_n$ и последовательность из модулей $\exp\{-\frac{1}{2}\vec{t}^\flat \boldsymbol{\Sigma}_n \vec{t}\}$ этих чисел имеют пределы. Следовательно, для любого $\vec{t}$ существуют пределы $\lim_n \vec{t}^\flat \vec{\mu}_n$ и $\lim_n \vec{t}^\flat \boldsymbol{\Sigma}_n \vec{t}$. Перебирая различные $\vec{t}$, например $\vec{t} = (1, 0, \ldots, 0)^\flat$, отсюда легко получаем, что для каждой компоненты последовательностей $\vec{\mu}_n$ и $\boldsymbol{\Sigma}_n$ существуют пределы. Другими словами, найдутся $\vec{\mu}_0 = \lim_n \vec{\mu}_n$ и $\boldsymbol{\Sigma}_0 = \lim_n \boldsymbol{\Sigma}_n$, причём, очевидно, $\boldsymbol{\Sigma}_0 \geqslant 0$. Следовательно,

$$\mathbf{E}\, e^{\mathrm{i}\vec{t}^\flat \vec{\xi}_0} \;=\; \exp\{\mathrm{i}\vec{t}^\flat \vec{\mu}_0\} \exp\{-\frac{1}{2}\vec{t}^\flat \boldsymbol{\Sigma}_0 \vec{t}\},$$

что и требовалось. $\Lleftarrow$

✧ Принцип Крамера–Волда. Следующий способ помогает упростить асимптотический анализ многомерных законов.

282| **Теорема.** [*Принцип Крамера-Волда.*]

I) *Последовательность k-мерных случайных векторов $\vec{\xi}_n \rightsquigarrow \vec{\xi}_0$ в том и только в том случае, если для $\forall \vec{c} \in \mathbb{R}^k$ линейная комбинация $\vec{c}^\flat \vec{\xi}_n \rightsquigarrow \vec{c}^\flat \vec{\xi}_0$.*

II) *Последовательность сл.векторов $\vec{\xi}_n \rightsquigarrow \mathcal{N}_k(\vec{\mu}, \boldsymbol{\Sigma})$ т. т. т. когда любая ненулевая линейная комбинация $\vec{c}^\flat \vec{\xi}_n \rightsquigarrow \mathcal{N}(\mathrm{m}_c, \sigma_c^2)$ со средним $\mathrm{m}_c = \vec{c}^\flat \vec{\mu}$ и дисперсией $\sigma_c^2 = \vec{c}^\flat \boldsymbol{\Sigma}\, \vec{c}$.*

$\Rrightarrow$ Пусть $\vec{c} \in \mathbb{R}^k, t \in \mathbb{R}^1$ и сл.в. $\zeta_n = \vec{c}^\flat \vec{\xi}_n$, $n \geqslant 0$. Хар.ф. сл.в. ζ_n связана с хар.ф. вектора $\vec{\xi}_n$ равенством

$$\varphi_{\zeta_n}(t) \;=\; \mathbf{E}\, e^{\mathrm{i}t\zeta_n} \;=\; \mathbf{E}\, e^{\mathrm{i}t\,\vec{c}^\flat\, \vec{\xi}_n} \;=\; \varphi_{\vec{\xi}_n}(t\vec{c}\,).$$

Следовательно, если $\zeta_n \rightsquigarrow \zeta_0$, то $\varphi_{\zeta_n}(1) \to \varphi_{\zeta_0}(1)$ и потому $\varphi_{\vec{\xi}_n}(\vec{c}\,) \to \varphi_{\vec{\xi}_0}(\vec{c}\,)$. Так как этот факт имеет место для $\forall \vec{c}$, то $\vec{\xi}_n \rightsquigarrow \vec{\xi}_0$. Аналогично доказывается обратное утверждение.

Вторая часть теоремы сразу следует из первой, если воспользоваться утверждением теоремы 195, стр. 186. $\Lleftarrow$

Докажем многомерный вариант центральной предельной теоремы.

283| **Теорема.** [*Многомерная ЦПТ.*] *Пусть $\langle \vec{\xi}_k \in \mathbb{R}^m \rangle_1^\infty$ — независимые одинаково распределённые сл.векторы с вектором конечных средних*

$\vec{\mu} = \mathbf{E}\,\vec{\xi}_k$ *и невырожденной матрицей ковариаций* $\mathbf{\Sigma} = \mathbf{Cov}(\vec{\xi}_k) > 0$. *Тогда*

$$\sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n}\vec{\xi}_k - \vec{\mu}\right) \rightsquigarrow \mathcal{N}_m(\vec{0}, \mathbf{\Sigma}).$$

⊫ Пусть $\vec{c}$ — произвольный ненулевой вектор. Тогда сл.в. $\eta_k = \vec{c}^{\flat}\,\vec{\xi}_k$, $k = 1, 2, \ldots$, образуют последовательность независимых одинаково распределённых сл.в. с общим средним $\mathfrak{m} = \vec{c}^{\flat}\,\vec{\mu}$ и дисперсией $\sigma^2 = \vec{c}^{\flat}\mathbf{\Sigma}\,\vec{c} > 0$. Следовательно, в соответствии с одномерной ЦПТ (стр. 275)

$$\vec{c}^{\flat}\left(\sqrt{n}\left(\frac{1}{n}\sum\nolimits_1^n\vec{\xi}_k - \vec{\mu}\right)\right) = \sqrt{n}\left(\frac{1}{n}\sum\nolimits_1^n\eta_k - \mathfrak{m}\right) \rightsquigarrow \mathcal{N}(0, \vec{c}^{\flat}\,\mathbf{\Sigma}\,\vec{c}).$$

Утверждение теоремы следует теперь из принципа Крамера–Волда. ⫤

284| **Теорема.** [*Многомерный дельта-метод.*] *Пусть функция* $h(\vec{x})$, $\vec{x} \in \mathbb{R}^m$, *дифференцируема в точке* $\vec{x} = \vec{a}$; $\vec{h}_a = (h'_{a_1}, \ldots, h'_{a_m})^{\flat}$ — *(ненулевой) вектор-столбец её частных производных в* $\vec{a}$. *Если последовательность сл.векторов* $(\vec{\xi}_n - \vec{a})q_n \rightsquigarrow \mathcal{N}_m(\vec{0}, \mathbf{\Sigma})$ *с некоторой последовательностью* $q_n \to \infty$ *и* $\mathbf{\Sigma} > 0$, *то*

$$\big(h(\vec{\xi}_n) - h(\vec{a})\big)q_n \rightsquigarrow \mathcal{N}_1\big(0, \vec{h}_a^{\flat}\,\mathbf{\Sigma}\vec{h}_a\big).$$

⊫ Дифференцируемость h в точке $\vec{a}$ означает, что

$$h(\vec{x}) - h(\vec{a}) = \vec{h}_a^{\flat}\,(\vec{x} - \vec{a}) + r(\|\vec{x} - \vec{a}\|),$$

где $r(z) = z\,\delta(z)$, $\delta(z) = o(1)$, $z \to 0$. Как и при доказательстве одномерного варианта дельта-метода 257, стр. 224, из условий теоремы следует $\vec{\xi}_n \xrightarrow{\mathbf{P}} \vec{a}$, что в силу 215, стр. 199, обеспечивает $\delta(\|\vec{\xi}_n - \vec{a}\|) = o_{\mathbf{P}}(1)$, $n \to \infty$.

По теореме Прохорова 275, стр. 234, последовательность $\|\vec{\xi}_n - \vec{a}\|q_n$ ограничена по вероятности, поэтому в силу 216, стр. 199, $r(\|\vec{\xi}_n - \vec{a}\|)q_n = \|\vec{\xi}_n - \vec{a}\|q_n\,\delta(\|\vec{\xi}_n - \vec{a}\|) = o_{\mathbf{P}}(1) \xrightarrow{\mathbf{P}} 0$. Кроме того, в соответствии с принципом Крамера–Волда асимптотическая нормальность $\vec{\xi}_n$ влечёт

$$\vec{h}_a^{\flat}\,(\vec{\xi}_n - \vec{a})q_n \rightsquigarrow \mathcal{N}_1\big(0, \vec{h}_a^{\flat}\,\mathbf{\Sigma}\vec{h}_a\big).$$

Применение леммы Слуцкого завершает доказательство теоремы. ⫤

285| Примеры. 1) Найдём асимптотическое распределение вектора $(\overline{X}_n, S_n^2)$, состоящего из выборочного среднего и выборочной дисперсии:

$$\overline{X}_n = \frac{1}{n}\sum\nolimits_1^n X_i, \qquad S_n^2 = \frac{1}{n}\sum\nolimits_1^n (X_i - \overline{X}_n)^2 = \frac{1}{n}\sum\nolimits_1^n X_i^2 - \big(\overline{X}_n\big)^2,$$

где $\langle\, X_i \backsim X,\ i=1,2,\dots\,\rangle$ — последовательность независимых сл.в., распределённых как сл.в. X.

Перейдём к новым сл.в. $\eta_i = X_i - \mu \backsim \eta = X - \mu$, где $\mu = \mathbf{E}X$, и пусть существуют (конечные) моменты $\alpha_j = \mathbf{E}\eta^j$, $j \leqslant 4$. Тогда $\alpha_1 = 0$, $\alpha_2 = \sigma^2 = \mathbf{D}X$ и моменты $\alpha_3 = \gamma\sigma^3$, $\alpha_4 = (\varkappa+3)\sigma^4$, где γ — коэффициент асимметрии X, $\varkappa$ — коэффициент эксцесса X. В этих обозначениях

$$\overline{X}_n = \overline{\eta}_n + \mu, \qquad S_n^2 = \frac{1}{n}\sum\nolimits_1^n (\eta_i - \overline{\eta}_n)^2 = \frac{1}{n}\sum\nolimits_1^n \eta_i^2 - \left(\overline{\eta}_n\right)^2.$$

У вектора $\vec{\theta} = (\eta, \eta^2)^\flat$ математическое ожидание $\mathbf{E}\,\vec{\theta} = (0,\sigma^2)^\flat$, а матрица ковариаций $\mathbf{\Sigma}$, очевидно, состоит из строк $(\sigma^2, \alpha_3) = \sigma^2(1,\gamma\sigma)$ и $(\alpha_3, \alpha_4 - \sigma^4) = \sigma^2(\gamma\sigma, (\varkappa+2)\sigma^2)$. По теореме 283 для последовательности $\vec{\xi}_n = \frac{1}{n}\sum_1^n \vec{\theta}_i = \left(\frac{1}{n}\sum_1^n \eta_i, \frac{1}{n}\sum_1^n \eta_i^2\right)^\flat$ имеем $\sqrt{n}\left(\vec{\xi}_n - (0,\sigma^2)^\flat\right) \rightsquigarrow \mathcal{N}_2(\vec{0}, \mathbf{\Sigma})$.

Пусть c_1, c_2 — произвольные детерминированные константы. Из предыдущего ясно, что линейная комбинация $c_1\overline{X}_n + c_2 S_n^2$ есть преобразование $h(\vec{\xi}_n)$ сл.вектора $\vec{\xi}_n$ с функцией $h(x_1,x_2) = c_1(x_1+\mu) + c_2(x_2 - x_1^2)$. В точке $\vec{a} = (0,\sigma^2)$ значение $h(\vec{a}) = c_1\mu + c_2\sigma^2$ и градиент $\vec{h}_a = (c_1, c_2)^\flat$. Применяя дельта-метод, получаем $\sqrt{n}\left(c_1(\overline{X}_n - \mu) + c_2(S_n^2 - \sigma^2)\right) \rightsquigarrow \mathcal{N}_1(0, \vec{c}^{\,\flat}\,\mathbf{\Sigma}\,\vec{c})$, что в соответствии с принципом Крамера–Волда влечёт

$$\sqrt{n}\left(\overline{X}_n - \mu, S_n^2 - \sigma^2\right) \ \rightsquigarrow\ \mathcal{N}_2(\vec{0}, \mathbf{\Sigma}), \qquad \mathbf{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \gamma\sigma \\ \gamma\sigma & (\varkappa+2)\sigma^2 \end{pmatrix}.$$

Как следствие, отсюда получаем, что при $n \to \infty$

$$\frac{1}{\sqrt{n}}\sum\nolimits_1^n \left(\left(\frac{X_i - \overline{X}_n}{\sigma}\right)^2 - 1\right) \ \rightsquigarrow\ \mathcal{N}_1(0, \varkappa+2).$$

При коэффициенте асимметрии $\gamma = 0$ выборочные среднее и дисперсия асимптотически независимы (при конечном n эти величины некоррелированы — см. упр. А3, стр. 106). При выборе из нормального распределения независимость выборочных моментов имеет место для $\forall\, n < \infty$.

2) Пусть $\xi_{(1)}, \xi_{(n)}$ — соответственно наименьшее и наибольшее значение последовательности $\xi_1,\dots,\xi_n$ независимых равномерных $\mathcal{U}\mathtt{n}(0,1)$ сл.в. Найдём асимптотическое распределение размаха $\rho_n = \xi_{(n)} - \xi_{(1)}$. Как было установлено в 244, стр. 217, $\left(n\xi_{(1)}, n(1-\xi_{(n)})\right) \rightsquigarrow (\eta_1, \eta_2)$, где сл.в. η_1, η_2 независимы и распределены показательно $\mathcal{E}\mathtt{x}(1)$. Поскольку $n(1-\rho_n) = \vec{a}^{\,\flat}\,(n\xi_{(1)}, n(1-\xi_{(n)}))$ с вектором $\vec{a}^{\,\flat} = (1,1)$, то в соответствии с принципом Крамера–Волда $n(1-\rho_n) \rightsquigarrow \eta_1 + \eta_2 \backsim \mathcal{G}\mathtt{am}(2,1)$.

К аналогичному выводу придём, воспользовавшись леммой Шеффе

263, стр. 226. Плотность размаха ρ_n при равномерном распределении наблюдений равна $f(x) = n(n-1)x^{n-2}(1-x)$, $0 \leqslant x \leqslant 1$ (см. формулу (10), стр. 127). Отсюда плотность сл.в. $\eta_n = n(1-\rho_n)$ (см. пример 100, стр. 112)

$$f_n(y) = \frac{n(n-1)}{n}\,\frac{y}{n}\left(1 - \frac{y}{n}\right)^{n-2} \underset{n}{\to} y\,e^{-y}, \qquad \forall\, y > 0. \qquad \circledcirc$$

§ 5. Слабая сходимость в функциональных пространствах

Исследование сходимости по распределению последовательности случайных элементов, принимающих значения в пространстве непрерывных функций $\mathcal{C}([0;1])$, возможно только путём анализа всевозможных конечномерных распределений этих случайных элементов. К сожалению, класс конечномерных цилиндров не определяет слабую сходимость в $\mathcal{C}([0;1])$. Пусть мера P_n сосредоточена в единственной точке (функции) $x_n(t) = nt\,\dot{\mathrm{I}}\left(t \leqslant \frac{1}{n}\right) + (2-nt)\,\dot{\mathrm{I}}\left(\frac{1}{n} \leqslant t \leqslant \frac{2}{n}\right)$, $t \in [0;1]$. Отметим, что $x_n(t) \underset{n}{\to} x_0(t) \equiv 0$ для $\forall t \in [0;1]$, однако $\|x_n - x_0\| = \sup_t \bigl|x_n(t) - x_0(t)\bigr| = 1 \not\to 0$ и, более того, из этой последовательности нельзя выделить никакую подпоследовательность, сходящуюся по норме к x_0. Слабая сходимость P_n к мере P_0, сосредоточенной в x_0, также не имеет места, т.к., например, для открытого множества $B = \{x : \|x - x_0\| < {}^1\!/_2\}$ мера $P_n(B) = 0$ для $\forall n \geqslant 1$, в то время как $P_0(B) = 1$ (см. критерий сходимости ($\mathrm{W_t}$) из теоремы о портфеле 240, стр. 212). С другой стороны, для любых фиксированных $0 < t_1 < \ldots < t_k$ при всех $n > 2/t_1$ функции $x_n(t_j) = 0 = x_0(t_j)$, $j = \overline{1,k}$, т.е. любое цилиндрическое множество содержит или не содержит одновременно функцию x_0 и все функции x_n, $n \geqslant 2/t_1$. Другими словами, мера $P_0(B) = P_n(B)$, $n \geqslant 2/t_1$, для любого цилиндрического множества B. Следовательно, конечномерные распределения P_n сходятся к конечномерным распределениям P_0, но $P_n \not\Rightarrow P_0$.

По теореме Прохорова (см. замечание 276, стр. 235) сходимость конечномерных распределений обеспечивает слабую сходимость т. т. т. когда последовательность мер P_n равномерно плотна. Поэтому для анализа слабой сходимости в $\mathcal{C}([0;1])$ необходимо иметь удобное средство для установления равномерной плотности семейств вероятностных распределений. В метрических пространствах это понятие тесно связано с понятием компактности. Удобный критерий компактности в $\mathcal{C}([0;1])$ описан в теореме Арцела́–Асколи, гласящей, что подмножество пространства $\mathcal{C}([0;1])$ относительно компактно т. т. т. когда оно ограничено и равностепенно непрерывно (см. [13], гл. II, §7).

Полезным инструментом анализа непрерывных функций является так называемый модуль непрерывности:

$$w(x\,;\,\delta) = \sup_{|s-t|<\delta} \bigl|x(s) - x(t)\bigr|, \qquad x \in \mathcal{C}([0;1]),\ \delta \in (0;1).$$

Модуль непрерывности не убывает по δ. Так как любая непрерывная функция на компакте равномерно непрерывна, то $\lim_{\delta\to 0} w(x;\delta) = 0$. Равностепенная непрерывность семейства функций $B \subset \mathcal{C}([0;1])$ означает, что $\lim_{\delta\to 0}\sup_{x\in B} w(x;\delta) = 0$. Кроме того, легко видеть, что $\bigl|w(x;\delta) - w(y;\delta)\bigr| \leqslant$ $\leqslant 2\|x-y\|$, поэтому модуль непрерывности при $\forall\,\delta > 0$ непрерывен как отображение из $\mathcal{C}([0;1])$ в $\mathbb{R}^1$, а следовательно, и измерим по Борелю.

Предположим, что для $\forall\, t \in [0;1]$ определена сл.в. $\xi(t\,;\,\cdot) : (\Omega, \mathscr{F}, \mathbf{P}) \mapsto$ $\mapsto (\mathbb{R}^1, \mathscr{B}^1)$. Если при каждом $\omega \in \Omega$ функция $\xi(t) := \xi(t\,;\,\omega)$, $t \in [0;1]$, непрерывна, то мы получаем отображение из вероятностного пространства в пространство непрерывных функций $\mathcal{C}([0;1])$. Поскольку борелевская σ-алгебра на $\mathcal{C}([0;1])$ совпадает с цилиндрической σ-алгеброй, то для измеримости отображения $\xi : \Omega \mapsto \mathcal{C}([0;1])$ необходимо и достаточно, чтобы были измеримы по Борелю отображения $\xi(t) : \Omega \mapsto \mathbb{R}^1$ при $\forall\, t \in [0;1]$. Кроме того, т.к. непрерывная функция полностью определяется своими значениями на счётном всюду плотном множестве рациональных точек $t \in [0;1]$, то любые преобразования типа $\sup_{t\in[0;1]} \xi(t)$ будут также измеримы по Борелю, в частности, модуль непрерывности $w(\xi\,;\,\delta)$ представляет собой измеримое по Борелю отображение Ω в $\mathbb{R}^1$, т.е. $w(\xi\,;\,\delta)$ — случайная величина.

286| **Теорема.** *Семейство* $\boldsymbol{\xi} = \bigl\langle\, \xi_\lambda(t) \,\bigr\rangle_{\lambda\in\Lambda}$ *случайных элементов в* $\mathcal{C}([0;1])$ *плотно т. т. т. когда*

i) *плотно семейство сл.в.* $\bigl\langle\, \xi_\lambda(0) \,\bigr\rangle_\lambda$, *т.е. для* $\forall\,\varepsilon > 0\,\exists\alpha$ *такое, что*

$$\sup_\lambda \mathbf{P}\bigl\{\bigl|\xi_\lambda(0)\bigr| > \alpha\bigr\} \ \leqslant\ \varepsilon; \tag{11}$$

ii) *для* $\forall\,\varepsilon, \gamma > 0$ *существует* $\delta \in (0;1)$ *такое, что*

$$\sup_\lambda \mathbf{P}\bigl\{\, w(\xi_\lambda\,;\,\delta) > \gamma \,\bigr\} \ \leqslant\ \varepsilon. \tag{12}$$

$\Rrightarrow$ $(\Leftarrow)$ Пусть $\varepsilon > 0$, тогда в соответствии с (11) можно подобрать α так, чтобы $\sup_\lambda \mathbf{P}\{\xi_\lambda \notin B\} \leqslant \varepsilon/2$, где множество

$$B \;=\; \{x \in \mathcal{C}([0;1]):\; |x(0)| \leqslant \alpha\}.$$

Аналогично, в силу (12) для $\forall k \;\geqslant\; 1$ найдётся δ_k такое, что $\sup_\lambda \mathbf{P}\{\xi_\lambda \notin B_k\} \leqslant \varepsilon/2^{k+1}$, для множества

$$B_k \;=\; \Big\{x \in \mathcal{C}([0;1]):\; w(x\,;\,\delta_k) < \frac{1}{k}\Big\}.$$

Множество функций $Q = B\bigcap_1^\infty B_k$ равностепенно непрерывно, т.к. при $\forall\, a > 0$ для $\forall\,\delta \leqslant \delta_k$ с некоторым $k < 1/a$ имеем $w(x\,;\,\delta) < \dfrac{1}{k} < a$ для всех $x \in Q$. Очевидно, что множество Q $(\subset B)$ равномерно ограничено, поэтому в силу теоремы Арцелá–Асколи замыкание $K = [Q]$ компактно. С другой стороны, для $\forall\,\lambda \in \Lambda$

$$\mathbf{P}\{\xi_\lambda \notin K\} \;\leqslant\; \mathbf{P}\{\xi_\lambda \notin B\} + \sum\nolimits_{k=1}^{\infty} \mathbf{P}\{\xi_\lambda \notin B_k\} \;\leqslant\; \varepsilon,$$

т.е. семейство $\boldsymbol{\xi}$ равномерно плотно.

$(\Rightarrow)$ Зафиксируем $\varepsilon, \gamma > 0$, и пусть компактное подмножество K из $\mathcal{C}([0;1])$ таково, что $\sup_\lambda \mathbf{P}\{\xi_\lambda \notin K\} < \varepsilon$. По теореме Арцелá–Асколи $K \subset$ $\subset \{x:\; |x(0)| \leqslant \alpha\}$ при некотором $\alpha > 0$ и, кроме того, $K \subset \{x : w(x\,;\,\delta) \leqslant \gamma\}$ при некотором $\delta > 0$. Следовательно, условия i) и ii) выполняются. ⫤

A. Дополнения

► Проверку сходимости почти наверное, по вероятности и в среднем можно осуществлять с помощью аналога критерия Коши.

Определения. Последовательность сл.в. $\langle\,\xi_n\,\rangle_1^\infty$ фундаментальна:

a) почти наверное, если $\mathbf{P}\big\{\lim_{k,n} \big|\xi_k - \xi_n\big| = 0\big\} = 1$;

b) по вероятности, если $\lim_{k,n} \mathbf{P}\big\{\big|\xi_k - \xi_n\big| > \varepsilon\big\} = 0,\ \forall\,\varepsilon > 0$;

c) в среднем порядка $q > 0$, если $\lim_{k,n} \mathbf{E}\big|\xi_k - \xi_n\big|^q = 0$.

[*Критерий Коши.*] *Последовательность* $\Xi = \langle\,\xi_n\,\rangle_1^\infty$ *сходится к некоторой сл.в.* ξ_0 *почти наверное (по вероятности, в среднем порядка* q*) т. т. т. когда* Ξ *фундаментальна почти наверное (по вероятности, в среднем порядка* q*).*

⫤ a) Зафиксируем $\omega \in \Omega$, тогда по критерию Коши последовательность действительных чисел $\xi_n(\omega) \to \xi_0(\omega)$ т. т. т. когда она фундаментальна. Поэтому если $\xi_n \to \xi_0$ (п.н.), то выполняется условие п.н. фундаментальности. Обратно, положим $\xi_0(\omega) = \widetilde{\xi}\,\dot{\mathrm{I}}\,(\omega; |\widetilde{\xi}\,| < \infty)$, где $\widetilde{\xi} = \overline{\lim\limits_{n}}\,\xi_n$. Из условия п.н. фундаментальности следует, что ξ_0 есть сл.в. и $\xi_n \xrightarrow{\text{п.н.}} \xi_0$.

b) Пусть $\xi_n \xrightarrow{\mathbf{P}} \xi_0$, тогда для $\forall\,\varepsilon > 0$

$$\mathbf{P}\{\,|\xi_k - \xi_n| > \varepsilon\,\} \;\leqslant\; \mathbf{P}\{\,|\xi_k - \xi_0| > \frac{\varepsilon}{2}\,\} + \mathbf{P}\{\,|\xi_n - \xi_0| > \frac{\varepsilon}{2}\,\} \underset{k,n\to\infty}{\longrightarrow} 0.$$

Обратно, из условия фундаментальности по вероятности следует, что найдётся подпоследовательность ξ_{n_k} такая, что

$$\mathbf{P}\{\,|\xi_{n_k} - \xi_{n_{k+1}}| > \frac{1}{2^k}\,\} \;<\; \frac{1}{2^k}.$$

В силу I) теоремы 205, стр. 194, существует сл.в. $\xi_0 = \lim_k \xi_{n_k}$ (п.н.) и $\xi_0 = \mathbf{P}\text{-}\lim_k \xi_{n_k}$, ввиду известной связи сходимости п.н. и по вероятности. Поэтому снова в силу условия фундаментальности

$$\mathbf{P}\{\,|\xi_n - \xi_0| > \varepsilon\,\} \;\leqslant\; \mathbf{P}\{\,|\xi_n - \xi_{n_k}| > \frac{\varepsilon}{2}\,\} + \mathbf{P}\{\,|\xi_{n_k} - \xi_0| > \frac{\varepsilon}{2}\,\} \underset{k,n\to\infty}{\longrightarrow} 0.$$

c) Заметим, что если $q \leqslant 1$, то при неотрицательных a, b справедливо неравенство $(a+b)^q \leqslant a^q + b^q$ (поскольку функция $(1+x)^q - x^q$ не возрастает при $x \geqslant 0$). Если же $q \geqslant 1$, то ввиду выпуклости функции x^q справедливо неравенство $(a+b)^q \leqslant 2^{q-1}(a^q + b^q)$. Другими словами, $(a+b)^q \leqslant C_q(a^q + b^q)$ с некоторым $C_q > 0$. Поэтому если последовательность $\langle\,\xi_n\,\rangle_1^\infty$ сходится в среднем порядка q, то она и фундаментальна в среднем порядка q:

$$\mathbf{E}\,|\xi_k - \xi_n|^q \;\leqslant\; C_q\big(\mathbf{E}\,|\xi_k - \xi_0|^q + \mathbf{E}\,|\xi_n - \xi_0|^q\big) \underset{k,n\to\infty}{\longrightarrow} 0.$$

Обратно, если последовательность $\langle\,\xi_n\,\rangle_1^\infty$ фундаментальна в среднем порядка q, то по неравенству Маркова она будет фундаментальна и по вероятности. Стало быть, ввиду b) найдётся сл.в. $\xi_0 = \mathbf{P}\text{-}\lim_n \xi_n$. По теореме 222, стр. 201, можно выделить подпоследовательность $\xi_{n_k} \xrightarrow{\text{п.н.}} \xi_0$ при $n_k \to \infty$. В силу леммы Фату

$$\begin{aligned}\lim_n \mathbf{E}\,|\xi_n - \xi_0|^q \;&=\; \lim_n \mathbf{E}\,\big[\,\lim_k |\xi_n - \xi_{n_k}|^q\big] \;\leqslant \\ &\leqslant\; \lim_n \varliminf_k \mathbf{E}\,|\xi_n - \xi_{n_k}|^q \;\leqslant\; \lim_{k,n} \mathbf{E}\,|\xi_k - \xi_n|^q \;=\; 0\end{aligned}$$

по условию фундаментальности в среднем порядка q. ⟸

► Слабая сходимость и сходимость по вероятности метризуемы.

Теорема. *Пусть $\rho(\vec{x}, \vec{y})$ — некоторая метрика в евклидовом пространстве $\mathbb{R}^k$, например $\rho(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$. Тогда*

a) *функция $d(\vec{\xi}, \vec{\zeta}) = \mathbf{E}\left[\dfrac{\rho(\vec{\xi}, \vec{\zeta})}{1 + \rho(\vec{\xi}, \vec{\zeta})}\right]$ задаёт метрику в пространстве всех* *сл.векторов на* $(\Omega, \mathscr{F}, \mathbf{P})$ *со значениями в* $\mathbb{R}^k$, *т.е.*

($\checkmark$) $d(\vec{\xi}, \vec{\zeta}) = 0 \quad \Leftrightarrow \quad \vec{\xi} = \vec{\zeta}$ (п.н.), ($\checkmark$) $d(\vec{\xi}, \vec{\zeta}) \leqslant d(\vec{\xi}, \vec{\eta}) + d(\vec{\zeta}, \vec{\eta})$;

b) *последовательность* $\rho(\vec{\xi}_n, \vec{\xi}_0) \xrightarrow{\mathbf{P}} 0 \quad \Leftrightarrow \quad d(\vec{\xi}_n, \vec{\xi}_0) \to 0$.

⊫ Заметим, что функция $f(x) = x/(1+x)$ при $x \geqslant 0$ строго возрастает от $f(0) = 0$ до $f(+\infty) = 1$. Поэтому для любой неотрицательной сл.в. $\chi \geqslant 0$ справедливы неравенства (здесь множество $A = \{\omega : \chi(\omega) > \varepsilon\}$):

$$\mathbf{P}\{\chi > \varepsilon\} = \mathbf{E}[\dot{\mathrm{I}}_A] \leqslant \frac{1}{f(\varepsilon)} \mathbf{E}[\dot{\mathrm{I}}_A f(\chi)] \leqslant \frac{1}{f(\varepsilon)} \mathbf{E}[f(\chi)],$$

$$\mathbf{E}[f(\chi)] = \mathbf{E}[\dot{\mathrm{I}}_A f(\chi)] + \mathbf{E}[\dot{\mathrm{I}}_{A^c} f(\chi)] \leqslant \mathbf{P}\{\chi > \varepsilon\} + \varepsilon.$$

Для доказательства пункта b) достаточно выбрать $\chi = \rho(\vec{\xi}_n, \vec{\xi}_0)$. ⫤

Определение. Пусть $F(x), G(x)$, $x \in \mathbb{R}^1$, — функции распределения на числовой прямой. *Метрикой П. Леви* между ф.р. называется функционал

$$\Lambda(F, G) = \inf\{\varepsilon > 0 : \ F(x - \varepsilon) - \varepsilon \leqslant G(x) \leqslant F(x + \varepsilon) + \varepsilon, \ \forall\, x \in \mathbb{R}^1\}.$$

⚠ Расстояние Леви с точностью до множителя $1/\sqrt{2}$ равно максимальному расстоянию между точками графиков ф.р., вычисляемому вдоль линий, проходящих под углом 135° к оси абсцисс; в точках разрыва линии графиков вертикальны (как на приведённом далее рисунке 4 у ступенчатой ф.р. G).

На рисунке справа центральная штрих пунктирная линия изображает ф.р. F нормальной $\mathcal{N}(0,1)$ сл.в., две пунктирные линии — функции $F(x \pm \varepsilon) \pm \varepsilon$ с $\varepsilon = 0.1$, а сплошная ступенчатая линия — ф.р. G сл.в. $(\xi - np)/\sqrt{np(1-p)}$, где $\xi \backsim \mathcal{B}\mathrm{in}(n,p)$, $n = 25, p = {}^3/_4$.

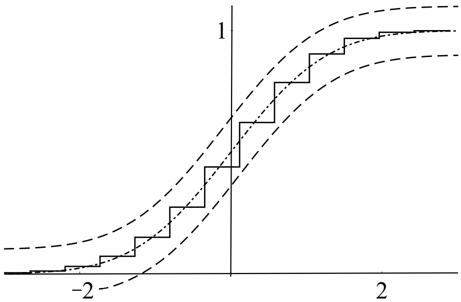


Рис. 4. К определению расстояния Леви

При указанных значениях параметров $\Lambda(F, G) \approx 0.077$.

Теорема. I) *Функционал* $\Lambda(F, G)$ *задаёт метрику в пространстве всех ф.р. на* $\mathbb{R}^1$.

II) *Последовательность ф.р.* $F_n \Rightarrow F_0$, *где* F_0 *также ф.р., т. т. т. когда* $\Lambda(F_n, F_0) \to 0$.

⊫ II) Пусть $C(F_0)$ — множество точек непрерывности ф.р. F_0.
($\Rightarrow$) Пусть $\varepsilon > 0$, тогда $\exists\, l, r \in C(F_0) :\ \max\{F_0(l), 1 - F_0(r)\} < \frac{1}{2}\varepsilon$.

- Найдутся $\langle a_j \rangle_0^m \subset C(F_0)$: $a_0 = l < a_1 < \ldots < a_m = r$, $a_j - a_{j-1} < \varepsilon$.
- Определим N_j : $\left|F_n(a_j) - F_0(a_j)\right| < \frac{1}{2}\varepsilon$, $\forall n > N_j$, и $N = \max\limits_{0 \leqslant j \leqslant m} N_j$.
- Если $x \leqslant a_0$ (аналогично для $x \geqslant a_m$), то для $\forall n > N$

$$F_n(x) \leqslant F_n(a_0) < F_0(a_0) + \frac{1}{2}\varepsilon \leqslant \varepsilon + F_0(x+\varepsilon),$$
$$F_n(x) \geqslant 0 > F_0(x) - \frac{1}{2}\varepsilon \geqslant F_0(x-\varepsilon) - \varepsilon.$$

- Если $a_{j-1} \leqslant x \leqslant a_j$, то $x - \varepsilon \leqslant a_{j-1} < a_j \leqslant x + \varepsilon$ и при $n > N$

$$F_n(x) \leqslant F_n(a_j) < F_0(a_j) + \frac{1}{2}\varepsilon \leqslant F_0(x+\varepsilon) + \varepsilon,$$
$$F_n(x) \geqslant F_n(a_{j-1}) > F_0(a_{j-1}) - \frac{1}{2}\varepsilon \geqslant F_0(x-\varepsilon) - \varepsilon.$$

($\Leftarrow$) Если $x \in C(F_0)$, то для $\forall \varepsilon > 0$

$$\exists \delta > 0 : \left|F_0(u) - F_0(x)\right| < \varepsilon, \ \forall \left|u - x\right| \leqslant \varepsilon.$$

- Выберем N : $\Lambda(F_n, F_0) < z := \min\{\varepsilon, \delta\}$, $\forall n \geqslant N$.
- Для $\forall n \geqslant N$

$$F_n(x) < F_0(x+z) + z \leqslant F_0(x+\delta) + \varepsilon < F_0(x) + 2\varepsilon,$$
$$F_n(x) > F_0(x-z) - z \geqslant F_0(x-\delta) - \varepsilon > F_0(x) - 2\varepsilon.$$

I) Соотношение $\left[\Lambda(F,G) = 0 \ \Leftrightarrow \ F \equiv G\right]$ следует из II (выбрать $F_n = F$) и того факта, что для совпадения двух ф.р. достаточно, чтобы они совпадали на множестве, дополнение которого не более чем счётно.

Предположим от противного, что $\Lambda(F,G) > \Lambda(F,Q) + \Lambda(G,Q)$.

- Найдутся $\varepsilon_1 > \Lambda(F,Q), \varepsilon_2 > \Lambda(G,Q)$ такие, что $\Lambda(F,G) > \varepsilon_1 + \varepsilon_2$.
- По определению $\Lambda(F,G)$ существует такая точка x, что

$$G(x) > F(x + \varepsilon_1 + \varepsilon_2) + \varepsilon_1 + \varepsilon_2 \tag{$\sharp$}$$

(или $G(x) < F(x - \varepsilon_1 - \varepsilon_2) - \varepsilon_1 - \varepsilon_2$).

- По определению $\Lambda(F,Q)$, $\Lambda(G,Q)$

$$F(x + \varepsilon_1 + \varepsilon_2) + \varepsilon_1 + \varepsilon_2 \geqslant Q(x + \varepsilon_2) + \varepsilon_2,$$
$$G(x) \leqslant Q(x + \varepsilon_2) + \varepsilon_2,$$

что противоречит ($\sharp$). Аналогично со вторым неравенством.

Равенство $\Lambda(F,G) = \Lambda(G,F)$ следует из предыдущих свойств. ⇚

► Так как, с одной стороны, существуют последовательности, сходящиеся по вероятности, но не сходящиеся п.н., а с другой стороны — из любой такой последовательности можно выделить сходящуюся п.н. подпоследовательность, то сходимость почти наверное метризуема только на вероятностных пространствах, для которых понятия сходимости почти наверное и по вероятности эквивалентны. Можно показать, что это справедливо только для атомических пространств. Событие $A \in \mathscr{F}$ называется *атомом* вероятностной меры $\mathbf{P}$ на измеримом пространстве $(\Omega, \mathscr{F})$, если вероятность $\mathbf{P}\{BA\} = 0$ или $\mathbf{P}\{BA\} = \mathbf{P}\{A\}$ для $\forall B \in \mathscr{F}$. Вероятностное пространство $(\Omega, \mathscr{F}, \mathbf{P})$ называется *атомическим*, если $\Omega = \biguplus_1^N A_k,\ N \leqslant \infty,\ \langle A_k \rangle_1^N$ — атомы $\mathbf{P}$.

✧ Метод одного пространства. Иногда полезен следующий приём, позволяющий рассматривать слабую сходимость как сходимость почти наверное. Пусть ф.р. $F_n \xrightarrow{\mathrm{w}} F_0$, где F_0 также является ф.р. Для каждой из этих ф.р. определим (псевдо)обратную функцию $F_n^{-1},\ n \geqslant 0$ (см. стр. 54). Зададим вероятностное пространство $\big((0;1), \mathscr{B}, \lambda\big)$ с борелевской σ-алгеброй и мерой Лебега. Рассмотрим сл.в. на этом вероятностном пространстве $\xi_n'(t) = F_n^{-1}(t),\ t \in (0;1),\ n \geqslant 0$. Как следует из 41, стр. 54, сл.в. $\xi_n' \backsim F_n,\ n \geqslant 0$.

Так как функции F_n, F_n^{-1} не убывают, то найдётся не более чем счётный набор точек $U = \{u \in (0;1)\}$ такой, что функция $F_0^{-1}(t)$ непрерывна при $t \notin U$ и, кроме того, все функции F_n непрерывны в точках $F_n^{-1}(t),\ n \geqslant 0$. Покажем, что для $\forall t \notin U$ имеет место сходимость $\xi_n'(t) \to \xi_0'(t)$.

Для $t \notin U$ по построению $F_n\big(F_n^{-1}(t)\big) = t$, поэтому если $F_{n_k}^{-1}(t) > F_0^{-1}(t) + \varepsilon$ при некотором $\varepsilon > 0$ для бесконечной последовательности n_k, то, выбирая ε так, чтобы $F_0^{-1}(t) + \varepsilon$ также была точкой непрерывности F_0, получаем

$$t \ \geqslant \ F_{n_k}\big(F_0^{-1}(t) + \varepsilon\big) \ \to \ F_0\big(F_0^{-1}(t) + \varepsilon\big) \ \geqslant \ F_0\big(F_0^{-1}(t + \delta)\big) \ \geqslant \ t + \delta$$

при некотором $\delta > 0$ (ввиду непрерывности F_0^{-1} в точке t). Аналогично, если $F_{n_k}^{-1}(t) < F_0^{-1}(t) - \varepsilon$ при некотором $\varepsilon > 0$ и $n_k \to \infty$, то

$$t \ \leqslant \ F_{n_k}\big(F_0^{-1}(t) - \varepsilon\big) \ \to \ F_0\big(F_0^{-1}(t) - \varepsilon\big),$$

т.е. $F_0^{-1}(t) \leqslant F_0^{-1}(t) - \varepsilon$. Стало быть, $\xi_n'(t) \to \xi_0'(t)$ для всех t из множества U^c, мера которого равна единице, т.е. $\xi_n' \xrightarrow{\text{п.н.}} \xi_0'$, в то же время $\xi_n' \rightsquigarrow \xi_0' \backsim F_0$.

287 Упр. Докажите вариант леммы Фату 537, стр. 481, для последовательности сл.в., сходящейся слабо:

$$\langle\, \xi_n \geqslant 0 \,\rangle_1^\infty, \quad \xi_n \rightsquigarrow \xi_0 \quad \Rightarrow \quad \mathbf{E}\,\xi_0 \leqslant \varliminf_n \mathbf{E}\,\xi_n.$$

► Другой способ изучения асимптотического поведения распределений сл.в., активно использованный Марковым при доказательстве предельных теорем, основан на анализе моментов этих сл.в. Справедлива

[*Теорема Frechet–Shohat.*] *Пусть распределение сл.в.* ξ_0 *единственным образом определяется её моментами* $\alpha_k = \mathbf{E}\xi_0^k$, $k \in \mathbb{N}$. *Если для последовательности сл.в.* $\langle \xi_n \rangle_1^\infty$ *моменты* $\mathbf{E}\xi_n^k \underset{n}{\to} \alpha_k$ *при всех* $k \in \mathbb{N}$, *то* $\xi_n \rightsquigarrow \xi_0$.

288| Упр. Докажите теорему Frechet–Shohat, воспользовавшись леммой Хелли о выборе и следствием 277, стр. 235.

Как следует из формулировки теоремы, здесь возникает ещё так называемая проблема моментов — проблема идентификации распределения по набору всех его моментов. Кратко эта проблема рассмотрена в разделе хар.функций, стр. 171, где приведены примеры различающихся ф.р. $F_1 \neq F_2$, у которых совпадают моменты $\int_{\mathbb{R}} x^k\,dF_1(x) = \int_{\mathbb{R}} x^k\,dF_2(x)$ для $\forall\, k = 1, 2, \ldots$

► Слабая сходимость сл.в. обеспечивает в некотором смысле равномерную сходимость соответствующих хар.ф. Точнее, если последовательность хар.ф. $\varphi_n(t) \to \varphi_0(t)$ в любой точке t и φ_0 — хар.ф., то эта сходимость равномерная на любом интервале $|t| \leqslant T$: $\sup_{|t|\leqslant T} |\varphi_n(t) - \varphi_0(t)| \underset{n}{\to} 0$.

Действительно, в силу теоремы П. Леви для соответствующих сл.в. имеет место слабая сходимость $\xi_n \rightsquigarrow \xi_0$. По теореме Прохорова для $\forall\, \varepsilon > 0$ найдётся $A > 0$ такое, что $\mathbf{P}\{|\xi_n| > A\} < \varepsilon$ для $\forall\, n \geqslant 0$. Так как $|e^{\mathrm{i}a} - e^{\mathrm{i}b}| \leqslant |a - b|$, то отрезок $[-A; A]$ можно разбить точками u_k, $k = 0, 1, \ldots, K$, так, что для $\forall\, |t| \leqslant T$ и $u_{j-1} < x \leqslant u_j$ справедливо неравенство $|e^{\mathrm{i}tx} - e^{\mathrm{i}Tx_j}| \leqslant \varepsilon$ и, кроме того, все эти точки суть точки непрерывности ф.р. ξ_0. Разбивая теперь область интегрирования в определении хар.ф. $\varphi_n(t) = \int_{\mathbb{R}} x\,dF_n(x)$ на участки $|x| > A, u_{j-1} < x \leqslant u_j$, $j = 1, \ldots, K$, можно легко показать, что при всех достаточно больших $n > n_0$ разность $|\varphi_n(t) - \varphi_0(t)| < 3\varepsilon$ для $\forall\, |t| \leqslant T$.

B. Упражнения

Упр. B.1. Пусть детерминированная последовательность $q_n \to 0$. Показать, что (a) $q_n\xi \xrightarrow{\text{п.н.}} 0$ для любой сл.в. ξ, (b) $q_n\xi_n \xrightarrow{\mathbf{P}} 0$ для любой последовательности $\langle \xi_n \rangle_1^\infty$ одинаково распределённых сл.в.

Упр. B.2. Привести пример последовательности, сходящейся п.н., но не сходящейся вполне (см. замечание 206, стр. 195).

Упр. B.3. Пусть $\xi_n \xrightarrow{\mathbf{P}} \eta$, $\zeta_n \xrightarrow{\mathbf{P}} \theta$ и $\mathbf{P}\{\eta \neq \theta\} = 0$. Показать, что $\xi_n - \zeta_n \xrightarrow{\mathbf{P}} 0$.

Упр. B.4. Пусть $\sum_1^\infty \mathbf{E}\,|\xi_n|^r < \infty$ для некоторого $r > 0$. Показать, что $\xi_n \xrightarrow{\text{п.н.}} 0$.

Упр. B.5. Доказать, что последовательность независимых сл.в. может сходиться (почти наверное или по вероятности) только к вырожденной (собственной) сл.в.

С. Указания к решению задач

✦ 205, стр. 194. II.a) Если $\varepsilon_n \to 0$, то для $\forall \varepsilon > 0$ найдётся $N : \varepsilon_n < \varepsilon$ для $\forall n > N$. Поэтому

$$\sum\nolimits_1^\infty \mathbf{P}\{|\xi_n - \xi_0| > \varepsilon_n\} \leqslant \sum\nolimits_1^N \mathbf{P}\{|\xi_n - \xi_0| > \varepsilon_n\} + \sum\nolimits_{N+1}^\infty \mathbf{P}\{|\xi_n - \xi_0| > \varepsilon\} < \infty.$$

✦ 209, стр. 196. (а) Использовать неравенство треугольника $\big||a| - |b|\big| \leqslant \big|a - b\big|$. (б) Применить неравенство Ляпунова 549, стр. 487.

✦ 210, стр. 197. $\mathbf{P}\{|\eta - \zeta| > \varepsilon\} \leqslant \mathbf{P}\Big\{|\eta - \xi_n| > \frac{1}{2}\,\varepsilon\Big\} + \mathbf{P}\Big\{|\zeta - \xi_n| > \frac{1}{2}\,\varepsilon\Big\} \underset{n}{\to} 0$.

✦ 213, стр. 199. Последовательно применить свойства (b) и (a).

✦ 214, стр. 199. б) $\mathbf{P}\{\xi_n\eta_n \leqslant M\} \leqslant \mathbf{P}\{\eta_n \leqslant {}^1/_2\} + \mathbf{P}\{\xi_n \leqslant 2M\} \underset{n}{\to} 0$.

✦ 217, стр. 200. Для $\forall \varepsilon > 0$ имеем

$$\mathbf{P}\Big\{(\xi_{jn} - \xi_{j0})^2 > \varepsilon^2\Big\} \leqslant \mathbf{P}\Big\{\sum_{i=1}^k (\xi_{in} - \xi_{i0})^2 > \varepsilon^2\Big\} \leqslant \sum_{i=1}^k \mathbf{P}\Big\{(\xi_{in} - \xi_{i0})^2 > \frac{1}{k}\varepsilon^2\Big\}.$$

✦ 219, стр. 201. В силу 217, стр. 200, вектор $(\xi_n, \eta_n) \xrightarrow{\mathbf{P}} (\xi_0, \eta_0)$; теперь достаточно применить теорему о непрерывном преобразовании к непрерывным функциям $H(x, y) = x + y$, $H(x, y) = xy$, $H(x, y) = x/y$.

✦ 221, стр. 201. а) при $\forall q > 0$; б) $\xi_n = 2^m\,\dot{\mathrm{I}}_{A_n}$.

✦ 224, стр. 202. Рассмотреть $\xi_n = \dot{\mathrm{I}}_{(n;n+1)}(\omega)$ при $\omega \in \mathbb{R}^1$ с мерой Лебега.

✦ 226, стр. 202. $\xi_n \xrightarrow{\text{п.н.}} 0 \Leftrightarrow \sup_{k \geqslant n} |\xi_k| \xrightarrow{\text{п.н.}} 0 \Leftrightarrow \sup_{k \geqslant n} |\xi_k| \xrightarrow{\mathbf{P}} 0$.

✦ 228, стр. 203. Пусть $\mathbf{E}\xi_{n_j} > \mathbf{E}\xi_0 + \varepsilon$ для $\forall j$, тогда $\exists \langle n_m \rangle_{m=1}^\infty \subset \langle n_j \rangle_{j=1}^\infty$, что $\xi_{n_m} \to \xi_0$ (п.н.). По теореме Лебега приходим к противоречию: $\mathbf{E}\xi_0 = \lim_m \mathbf{E}\xi_m \geqslant \mathbf{E}\xi_0 + \varepsilon$.

✦ 231, стр. 206. а) Да. б) Нет (привести пример).

✦ 234, стр. 207. Так как $F_\xi(x) = F_\eta(x)$ во всех точках непрерывности обеих ф.р., то эти функции совпадают всюду.

✦ 235, стр. 207. а) Пусть x — точка непрерывности ф.р. $a\xi_0 + b$, тогда $(x - b)/a$ — точка непрерывности ф.р. сл.в. ξ_0, поэтому $F_n((x - b)/a) \underset{n}{\to} F_0((x - b)/a)$, следовательно, $F_{a\xi_n + b}(x) \underset{n}{\to} F_{a\xi_0 + b}(x)$. б) Функция H имеет обратную H^{-1}. Если H убывает, то $H(x) \leqslant y \Leftrightarrow x \geqslant H^{-1}(y)$ и $H(x) < y \Leftrightarrow x > H^{-1}(y)$. Пусть y — точка непрерывности ф.р. сл.в. $H(\xi_0)$, тогда $\mathbf{P}\{\xi_0 \geqslant H^{-1}(y)\} = \mathbf{P}\{H(\xi_0) \leqslant y\} = \mathbf{P}\{H(\xi_0) < y\} = \mathbf{P}\{\xi_0 > H^{-1}(y)\}$, т.е. $H^{-1}(y)$ — точка непрерывности ξ_0. Далее аналогично предыдущему пункту.

✦ 239, стр. 211. (a) Условие леммы справедливо как для функции h, так и для функции $-h$, поэтому $\overline{\lim_n} \int h\,dQ_n \leqslant \int h\,dQ_0 \leqslant \underline{\lim_n} \int h\,dQ_n$. (b) Предположение, что найдутся $\varepsilon > 0$, непрерывная ограниченная функция h и подпоследовательность $Q_{n'}$ такие, что $\int h\,dQ_{n'} - \int h\,dQ_0 > \varepsilon$, противоречит условию леммы.

✦ 244, стр. 217. Вместо традиционной ф.р. рассмотреть связанную с ней функцию $\mathbf{P}\{n\xi_{(1:n)} \geqslant x,\, n(1-\xi_{(n:n)}) \geqslant y\} = \left(1 - \dfrac{x}{n} - \dfrac{y}{n}\right)^n \to \exp(-x-y) = \mathbf{P}\{\eta_1 \geqslant x\}\,\mathbf{P}\{\eta_2 \geqslant y\}$ (ср. с замечанием 131, стр. 126).

✦ 246, стр. 218. Если h — непрерывная ограниченная функция, то таковой будет также суперпозиция $h \circ H$.

✦ 251, стр. 221. Пусть $F_0(x,y) = F_{\xi_0}(x)F_{\zeta_0}(y)$. Если (x,y) — точка непрерывности F_0 и $F_0(x,y) \neq 0$, то x, y — точки непрерывности F_{ξ_0}, F_{ζ_0} соответственно, поэтому $F_{\xi_n}(x)F_{\zeta_n}(y) \to F_{\xi_0}(x)F_{\zeta_0}(y)$. Если же, например, $F_{\xi_0}(x) = 0$, то, очевидно, x — точка непрерывности F_{ξ_0}, поэтому $\lim_n F_{\xi_n}(x) = 0$ и, следовательно, $F_{\xi_n}(x)F_{\zeta_n}(y) \to 0 = F_{\xi_0}(x)F_{\zeta_0}(y)$.

✦ 259, стр. 225. I) Рассмотреть $H(x) = \dfrac{2}{h''(a)}\,\dfrac{h(x)-h(a)}{(x-a)^2}\,\dot{\mathrm{I}}_{x\neq a}(x) + \dot{\mathrm{I}}_{x=a}(x)$. II) а) $\xi_n^2 \rightsquigarrow \mathcal{N}(\mu^2, 4\mu^2\sigma_n^2)$, б) $\xi_n^2/\sigma_n^2 \rightsquigarrow \mathcal{K}\mathtt{hi}(1)$.

✦ 260, стр. 225. Любое нецелое число x есть точка непрерывности ф.р. Пуассона. Для любой такой точки

$$F_n(x) = \sum_{j=0}^{\llbracket x\rrbracket} \mathbf{P}\{\xi_n = j\} \underset{n}{\to} \sum_{j=0}^{\llbracket x\rrbracket} \mathbf{P}\{\xi_0 = j\},$$

поскольку количество слагаемых здесь конечно.

✦ 274, стр. 233. Монотонность F очевидна. Для доказательства непрерывности заметить, что по определению инфимума $F(x) = \lim_n G(x_n)$ с $x_n \searrow x$, и для любой последовательности $y_n \searrow x$ найдётся последовательность $n_k \nearrow \infty$ такая, что $x_{n_{k+1}} \leqslant y_{n_k} \leqslant x_{n_k}$.

✦ 266, стр. 228. Математическое ожидание $\mathbf{E}\eta_k = k$, а дисперсия $\mathbf{D}\eta_k = 2k$, поэтому $\mathbf{D}[\eta_k/k] = 2/k \to 0$, т.е. $\eta_k/k \xrightarrow{\mathcal{L}_2} 1$; далее применить лемму Слуцкого.

✦ 287, стр. 249. Воспользоваться методом одного пространства.

✦ 288, стр. 250. Из любой подпоследовательности ф.р. F_{n_k} можно выбрать подпоследовательность $F_{n_m} \Rightarrow G$ с некоторой функцией G. В силу 277, стр. 235, функция G — собственная ф.р. некоторой сл.в. η. По условию моменты $\mathbf{E}\eta^k = \mathbf{E}\xi_0^k$, следовательно, $\eta \backsim \xi_0$ независимо от подпоследовательности F_{n_k}.

✦ B.1. (a) Для $\forall\omega \in \Omega\,\forall\varepsilon > 0\exists N > 1 : |\xi(\omega)| < \varepsilon/|q_n|$ при $\forall n > N$; (b) если $\xi_n \backsim F(x)$, то $\mathbf{P}\{|q_n\xi_n| > \varepsilon\} = F(-\varepsilon/|q_n|) + 1 - F(\varepsilon/|q_n|) \underset{n}{\to} 0$.

✦ B.2. Рассмотреть $\Omega = [0;1]$ и $\xi_n = \dot{\mathrm{I}}(\omega, (a_n; b_n))$ с $b_n - a_n \to 0$.

✦ B.3. $\mathbf{P}\{\xi_n - \zeta_n < -\varepsilon\} \leqslant \mathbf{P}\{\xi_n - \eta < -\varepsilon/2\} + \mathbf{P}\{\theta - \zeta_n < -\varepsilon/2\} + \mathbf{P}\{\eta \neq \theta\} \underset{n}{\to} 0$.

✦ B.4. Воспользоваться теоремой 205, стр. 194, и неравенством Маркова.

✦ B.5. Пусть $\xi_n \xrightarrow{\text{п.н.}} \xi_0$. Событие $A = \{\omega : \xi_n(\omega) \text{ не сходится}\} = \cup_{N=1}^{\infty} \cap_{n=k}^{\infty} \cup_{m=1}^{\infty} \{|\xi_{n+m} - \xi_n| > 1/N\}$, $k \geqslant 1$, остаточное относительно σ-алгебры, порождённой $\langle \xi_n \rangle_1^{\infty}$; применить закон нуля или единицы к остаточному событию $\{\omega : \omega \notin A,\ \xi_0(\omega) \leqslant x\} = \cap_{N=1}^{\infty} \cup_{n=k}^{\infty} \cap_{m=n}^{\infty} (\{\xi_m \leqslant x + 1/N\}A^c)$, $k \geqslant 1$. Если $\xi_n \xrightarrow{\mathbf{P}} \xi_0$, то существует подпоследовательность $\xi_{n_k} \xrightarrow{\text{п.н.}} \xi_0$.

VII Предельные теоремы

Анализ асимптотического поведения сумм случайных величин, в особенности сумм независимых сл.в., входит в круг первостепенных задач теории вероятностей. Здесь мы обсудим некоторые методы доказательства так называемого закона больших чисел (далее — ЗБЧ), когда специальным образом нормированная сумма сл.в. сходится к константе, а также методы доказательства центральной предельной теоремы (ЦПТ), когда в пределе распределение такой суммы совпадает с нормальным законом. На протяжении всей этой главы мы будем интересоваться последовательностями сл.в. $\left\langle \xi_k,\ k \in \mathbb{N} \right\rangle$, заданных на одном вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$; конечную сумму первых n членов этих сл.в. будем обозначать через $S_n = \sum_1^n \xi_k$. Вопрос существования таких последовательностей здесь рассматриваться не будет (см. по этому поводу главу VIII пособия, посвящённую сл.процессам).

Основной метод установления предельных законов для сумм сл.в. опирается на теорему П. Леви о непрерывном соответствии 271, стр. 230, и свойства хар.ф. 158, стр. 158. Кроме того, нам понадобится здесь представление показательной функции:

$$\exp(-x) = 1 - x + \frac{1}{2}x^2\theta(x), \qquad |\theta(x)| \leqslant 1,\ x > 0, \tag{1}$$

а также неравенство (27), стр. 302, для комплексных чисел $|a_k|, |b_k| \leqslant 1$, которое в частном случае $|a|, |b| \leqslant 1$ принимает вид

$$|a^n - b^n| = |(a-b)(a^{n-1} + a^{n-2}b + \ldots + ab^{n-2} + b^{n-1}| \leqslant n|a-b|. \tag{2}$$

§ 1. Слабые законы больших чисел

Просто доказываются варианты ЗБЧ, утверждающие сходимость по вероятности средних значений. Поскольку сходимость по вероятности к константе эквивалентна слабой сходимости, такие ЗБЧ называют слабыми законами — в противовес усиленным законам, утверждающим сходимость почти наверное. Один из самых простых с точки зрения доказательства — закон больших чисел в форме Чебышёва, использующий неравенство Чебышёва.

289| **Лемма.** (?!) а) [*Неравенство Маркова.*] *Пусть сл.в.* $\eta \geqslant 0$, *тогда при* $\forall Y > 0$

$$\mathbf{P}\{\eta > Y\} \leqslant \frac{1}{Y}\,\mathbf{E}\left[\eta\,\dot{\mathrm{I}}(\eta > Y)\right] \leqslant \frac{1}{Y}\mathbf{E}\,\eta.$$

б) [*Неравенство Чебышёва.*] *Для любой сл.в.* η *с конечным математическим ожиданием* $m = \mathbf{E}\,\eta$

$$\mathbf{P}\left\{\left|\eta - m\right| > \varepsilon\right\} \leqslant \frac{\mathbf{D}\eta}{\varepsilon^2}, \qquad \varepsilon > 0.$$

290| **Теорема.** [*Закон больших чисел Чебышёва.*] *Пусть* $\left\langle \xi_k \right\rangle_1^\infty$ — *одинаково распределённые некоррелированные (например, независимые) сл.в. с конечными математическим ожиданием* $\mu = \mathbf{E}\,\xi_k$ *и дисперсией* $\sigma^2 = \mathbf{D}\xi_k$, $k \in \mathbb{N}$. *Тогда*

$$\mathbf{P}\left\{\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \mu\right| > \varepsilon\right\} \leqslant \frac{\sigma^2}{n\varepsilon^2}, \tag{3}$$

следовательно, при $n \to \infty$

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow{\mathbf{P}} \mu. \tag{4}$$

$\Rrightarrow$ Легко понять, что математическое ожидание $\mathbf{E}\left[S_n/n\right] = \mu$, и ввиду некоррелированности слагаемых

$$\mathbf{D}\left[\frac{1}{n}S_n\right] = \frac{1}{n^2}\sum\nolimits_1^n \mathbf{D}\xi_k = \frac{\sigma^2}{n}.$$

Оценка (3), а вместе с ней и утверждение теоремы очевидным образом следуют теперь из неравенства Чебышёва. $\Lleftarrow$

Из доказательства теоремы видно, что $\mathbf{D}[S_n/n] \underset{n}{\to} 0$. Другими словами, в условиях ЗБЧ Чебышёва имеет место $\mathcal{L}_2$-сходимость последовательности средних арифметических к математическому ожиданию μ :

$$\frac{1}{n}\sum\nolimits_1^n \xi_k \xrightarrow{\mathcal{L}_2} \mu.$$

Неравенство Чебышёва в доказательстве теоремы выполняет связующую роль между двумя типами сходимости. Использовать соотношение (3), называемое иногда неравенством Чебышёва–Бьенэме, для оценки вероятности отклонения среднего арифметического от истинного среднего при конечном числе слагаемых, немного опрометчиво, поскольку это неравенство очень грубое. Так, для бернуллиевских сл.в. с вероятностью успеха $p = {}^1\!/_2$ правая часть (3) при $n = 1000$ и $\varepsilon = 0.01$ равна 2.5, что даёт, конечно, верную, но бесполезную оценку для вероятности. В этих условиях, кстати, левая часть равна 0.527. При $\varepsilon = 0.1$ неравенство выглядит так: $2.3 \cdot 10^{-10} < 2.5 \cdot 10^{-2}$.

Если не предполагать существования дисперсии, но потребовать достаточно быстрое убывание хвоста распределения ξ (см. 308, стр. 264), можно доказать сходимость в среднем (в $\mathcal{L}_1$).

291| **Следствие.** (Из теоремы 308.) *Пусть* $\langle\, \xi_k \,\rangle_1^\infty$ *— попарно независимые одинаково распределённые сл.в. с конечным математическим ожиданием* $\mu = \mathbf{E}\,\xi_1$. *Если*

$$\lim_n \frac{1}{n}\,\mathbf{E}\left[(\xi_1-\mu)^2\,\dot{\mathrm{I}}\,(\xi_1 \leqslant n)\right] = 0,$$

то $\frac{1}{n}S_n \xrightarrow{\mathcal{L}_1} \mu$ *и, следовательно,* $\frac{1}{n}S_n \xrightarrow{\mathbf{P}} \mu$.

На этом же пути можно получить различные обобщения ЗБЧ Чебышёва, например, для случая некоррелированных, но различно распределённых сл.в.

292| **Теорема.** (?!) *Пусть* $\langle\, \xi_k \,\rangle_1^\infty$ *— некоррелированные сл.в. с математическими ожиданиями* $\mu_k = \mathbf{E}\,\xi_k$ *и дисперсиями* $\mathbf{D}\xi_k < d^2 < \infty$, $k \in \mathbb{N}$. *Тогда*

$$\frac{1}{n}\sum_{k=1}^n \xi_k - \frac{1}{n}\sum_{k=1}^n \mu_k \xrightarrow[n]{\mathbf{P}} 0. \tag{5}$$

Кроме того, здесь можно отказаться и от условия независимости (некоррелированности). Сформулируем сначала три вспомогательных утверждения (доказательство см. на стр. 302).

293| **Лемма.** *Пусть* $\langle\, a_k, b_k, \,\rangle_{k=0}^\infty$ *— действительные числа. Тогда* (а) *если* $\lim_k a_k = A,\ \lim_k b_k = B,\ |A|, |B| < \infty$, *то*

$$(\checkmark)\quad \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} a_k = A, \qquad (\checkmark)\quad \lim_{n\to\infty} \frac{1}{n^2}\sum_{j,k=1}^{n} a_j b_{|k-j|} = AB;$$

(б) [*Кронекер*] *если ряд* $\sum_1^\infty \frac{a_k}{b_k}$ *сходится и* $b_k \nearrow \infty$, *то*

$$\lim_n \frac{1}{b_n}\sum_{k=1}^{n} a_k = 0.$$

294| **Теорема.** *Пусть* $\langle \xi_k \rangle_1^\infty$ *— сл.в. с конечными математическими ожиданиями* $\mu_k = \mathbf{E}\xi_k$ *и дисперсиями* $\mathbf{D}\xi_k,\ k \in \mathbb{N}$, *такими, что* $\sum_1^n \mathbf{D}\xi_k/n^2 \underset{n}{\to} 0$. *Если коэффициент ковариации* $\mathbf{Cov}(\xi_i, \xi_j) = \gamma_{|j-i|}$ *зависит только от разности номеров* $j-i$ *и при этом* $\lim_k \gamma_k = 0$, *то справедлив ЗБЧ* (5).

⤃ Как и выше, достаточно показать, что $\mathbf{D}\left[\frac{1}{n}S_n\right] \to 0$. Имеем:

$$\mathbf{D}\left[\frac{1}{n}S_n\right] = \frac{1}{n^2}\sum_{k,j=1} \mathbf{E}\left[(\xi_k - \mu_k)(\xi_j - \mu_j)\right] \leqslant$$

$$\leqslant \frac{1}{n^2}\sum_{k=1}^{n} \mathbf{D}\xi_k + \frac{1}{n^2}\sum_{k,j=1}^{n} a_j b_{|k-j|},$$

где $a_j \equiv 1,\ b_0 = 0, b_m = |\gamma_m|,\ m > 0$. По условию теоремы и в силу утверждения (а) леммы 293 дисперсия $\mathbf{D}\left[\frac{1}{n}S_n\right] \to 0$, что и требовалось. ⤆

295| Упр. Докажите справедливость ЗБЧ для зависимых сл.в., если $\sum_1^n \mathbf{D}\xi_k/n^2 \underset{n}{\to} 0$ и все коэффициенты ковариации неположительны: $\mathbf{Cov}(\xi_k, \xi_j) \leqslant 0,\ k \neq j$.

Освободимся теперь от условия существования вторых моментов и исследуем сначала случай одинаково распределённых сл.в. Если сл.в. независимы в совокупности, можно воспользоваться методом характеристических функций. В случае попарно независимых сл.в. ключевое свойство хар.ф. уже не работает, поэтому приходится прибегать к более подробным выкладкам. Аналогичные выкладки применяются и для установления обратного результата, а также для доказательства ЗБЧ (5) для разнораспределённых сл.в. Основной приём, используемый здесь, состоит в отсечении больших значений сл.в. Определим для любой сл.в. ζ усечение на уровне h (> 0)

$$\zeta^{(h)} = \zeta\, \dot{\mathrm{I}}\,(|\zeta| \leqslant h).$$

Ясно, что $|\zeta^{(h)}| \leqslant |\zeta|$ и при $h \to \infty$ поточечно (т.е. для $\forall \omega$) $\zeta^{(h)} \to \zeta$. В силу теоремы Лебега об ограниченной сходимости отсюда следует, что $\mathbf{E}\,\zeta^{(h)} \to \mathbf{E}\,\zeta$, если $\mathbf{E}\,|\zeta| < \infty$, $h \to \infty$. Последнее, правда, при доказательстве сформулированного далее ЗБЧ Хинчина нам не понадобится. Здесь важнее характер поведения распределения интегрируемых сл.в. на «хвостах» числовой оси.

296| **Лемма.** *Для любой неотрицательной сл.в.* ζ :

(a) $\lim\limits_{x\to\infty} x\,\mathbf{P}\{\zeta \geqslant x\} = 0$, *если* $\mathbf{E}\,\zeta < \infty$;

(b) *если* (a), *то* $\lim\limits_{x\to\infty} \dfrac{1}{x}\,\mathbf{E}\left[\zeta^2\,\dot{\mathrm{I}}(\zeta < x)\right] = 0$;

(c) $\mathbf{E}\,\zeta = \int_0^\infty \mathbf{P}\{\zeta \geqslant x\}\,dx$;

(d) $\sum\limits_{n=1}^{\infty} \mathbf{P}\{\zeta \geqslant n\} \;\leqslant\; \mathbf{E}\,\zeta \;\leqslant\; \sum\limits_{n=0}^{\infty} \mathbf{P}\{\zeta \geqslant n\}$.

$\Rrightarrow$ Утверждение (a) непосредственно следует из первой части неравенства Маркова 289, стр. 254, и теоремы Лебега. Для доказательства (b) положим $G_j = \mathbf{P}\{\zeta \geqslant j\}$. Очевидно, $G_0 = 1$ и $\mathbf{P}\{j-1 \leqslant \zeta < j\} = G_{j-1} - G_j$. Пусть $n_x = \llbracket x \rrbracket + 1$ $\quad (1 < n_x/x \to 1$ при $x \to \infty)$, тогда

$$\mathbf{E}\left[\zeta^2\,\dot{\mathrm{I}}(\zeta < x)\right] \;\leqslant\; \sum_{j=1}^{n_x} \mathbf{E}\left[\zeta^2\,\dot{\mathrm{I}}(j-1 \leqslant \zeta < j)\right] \;\leqslant$$

$$\leqslant\; \sum_{j=1}^{n_x} j^2\bigl(G_{j-1} - G_j\bigr) \;=\; \sum_{j=0}^{n_x-1} (j+1)^2 G_j - \sum_{j=1}^{n_x} j^2 G_j \;=$$

$$=\; G_0 - n_x^2 G_{n_x} + \sum_{j=1}^{n_x-1} (2j+1) G_j \;\leqslant\; 1 + \sum_{j=1}^{n_x-1} (2j+1)\,\mathbf{P}\{\zeta \geqslant j\}\,.$$

Так как в силу (a) $\lim_j \left[(2j+1)\,\mathbf{P}\{\zeta \geqslant j\}\,\right] = 0$, то $\frac{1}{x}\,\mathbf{E}\left[\zeta^2\,\dot{\mathrm{I}}(\zeta < x)\right] \to 0$, $x \to \infty$, по первой части леммы 293. (c) Совпадает с упр. 44, стр. 55.

(d) Положим $\mathbf{J}_n = [n; n+1)$, $n \geqslant 0$, и воспользуемся (c):

$$\mathbf{E}\,\zeta = \int_0^\infty \mathbf{P}\{\zeta \geqslant x\}\,dx = \sum_{n=0}^{\infty} \int_{\mathbf{J}_n} \mathbf{P}\{\zeta \geqslant x\}\,dx \quad \begin{matrix} \leqslant\; \sum_0^\infty \mathbf{P}\{\zeta \geqslant n\} \\[2ex] \geqslant\; \sum_0^\infty \mathbf{P}\{\zeta \geqslant n+1\} \end{matrix}$$

в силу очевидного неравенства $\mathbf{P}\{\zeta \geqslant a\} \leqslant \mathbf{P}\{\zeta \geqslant b\}$ для $a \geqslant b$. $\Lleftarrow$

Заметим, что ряд в (d) конечен т. т. т. когда $\mathbf{E}\,\zeta < \infty$.

297| **Теорема.** [*Закон больших чисел Хинчина–Феллера.*] *Пусть независимые (попарно или в совокупности) сл.в.* $\langle \xi_k \rangle_1^\infty$ *одинаково распределены с конечным ожиданием* $\mu = \mathbf{E}\,\xi_k$, $k \in \mathbb{N}$. *Тогда справедлив ЗБЧ* (4).

$\Rrightarrow$ [*Хинчин.*] Пусть сл.в. $\langle \xi_k \rangle_1^\infty$ независимы в совокупности; $\varphi(t)$ — хар.ф. сл.в. ξ_k. По свойству (X_8) для хар.ф. (стр. 158)

$$\varphi(t) = 1 + \mathrm{i}\mu t + t\delta(t)$$

с остаточным членом $\delta(t) \to 0$ при $t \to 0$. В силу свойств (X_4) и (X_9) хар.ф. среднего $\zeta_n = \frac{1}{n} S_n$ независимых в совокупности сл.в.

$$\varphi_{\zeta_n}(u) = \Big(\varphi\Big(\frac{u}{n}\Big)\Big)^n = \Big(1 + \mathrm{i}\mu\frac{u}{n} + \frac{u}{n}\delta\Big(\frac{u}{n}\Big)\Big)^n, \qquad \forall\, u \in \mathbb{R}^1.$$

Так как $\lim_n \big(\mathrm{i}\mu\frac{u}{n} + \frac{u}{n}\delta\big(\frac{u}{n}\big)\big)n = \mathrm{i}\mu u$, то по «второму замечательному пределу» имеем $\lim_n \varphi_{\zeta_n}(u) = \exp\{\mathrm{i}\mu u\}$ для $\forall\, u \in \mathbb{R}^1$. Очевидно, функция $\exp\{\mathrm{i}\mu u\}$ есть хар.ф. детерминированной сл.в. $\xi_0 \equiv \mu$, что доказывает ЗБЧ Хинчина в силу теоремы П. Леви о непрерывном соответствии для хар.ф.

[*Феллер.*] Пусть сл.в. попарно независимы. Ясно, что можно считать $\mu = 0$. Так как для $\forall \varepsilon > 0$ в силу леммы 296, (a) при $n \to \infty$

$$\mathbf{P}\Big\{\Big|\frac{1}{n}\sum\nolimits_1^n \big(\xi_k - \xi_k\,\dot{\mathrm{I}}_{\{|\xi_k|<n\}}\big)\Big| > \varepsilon\Big\} \leqslant \mathbf{P}\Big\{\bigcup\nolimits_1^n \big(|\xi_k| \geqslant n\big)\Big\} \leqslant$$
$$\leqslant n\,\mathbf{P}\{|\xi_1| > n\} \underset{n}{\to} 0,$$

то

$$\frac{1}{n}\sum\nolimits_1^n \xi_k \;-\; \frac{1}{n}\sum\nolimits_1^n \xi_k\,\dot{\mathrm{I}}_{\{|\xi_k|<n\}} \xrightarrow[n]{\mathbf{P}} 0. \tag{6}$$

Поэтому для доказательства теоремы достаточно показать, что второе слагаемое здесь стремится к нулю по вероятности (или в $\mathcal{L}_2$). Имеем, во-первых,

$$\mathbf{E}\Big[\frac{1}{n}\sum\nolimits_1^n \xi_k\,\dot{\mathrm{I}}_{\{|\xi_k|<n\}}\Big] = \frac{1}{n}\sum\nolimits_1^n \mathbf{E}\Big[\xi_k\,\dot{\mathrm{I}}_{\{|\xi_k|<n\}}\Big] =$$
$$= \mathbf{E}\Big[\xi_1\,\dot{\mathrm{I}}_{\{|\xi_1|<n\}}\Big] \underset{n}{\to} \mu = 0.$$

Во-вторых, в силу попарной независимости и ввиду 296, (b) дисперсия

$$\mathbf{D}\Big[\frac{1}{n}\sum\nolimits_1^n \xi_k\,\dot{\mathrm{I}}_{\{|\xi_k|<n\}}\Big] = \frac{1}{n}\mathbf{D}\big[\xi_1\,\dot{\mathrm{I}}_{\{|\xi_1|<n\}}\big] \underset{n}{\to} 0. \qquad \Lleftarrow$$

298| Упр. Анализ доказательства показывает, что ЗБЧ в форме

$$\frac{1}{n} S_n - \mathbf{E}\left[\xi_1 \, \dot{\mathrm{I}}\,(|\xi_1| < n)\right] \xrightarrow[n]{\mathbf{P}} 0$$

будет справедлив, если потребовать лишь $\lim_n n\,\mathbf{P}\{|\xi_1| \geqslant n\} = 0$. Приведите пример сл.в., удовлетворяющей этому условию, но не имеющей математического ожидания.

⚠ Конечность математического ожидания потребовалась при доказательстве теоремы Хинчина только для представления хар.ф. в виде отрезка ряда Тейлора. К такому же результату можно прийти, если предположить, что хар.ф. непрерывно дифференцируема в некоторой окрестности нуля и первая производная $\varphi'(0) = 0$. Напомним, что последнее не гарантирует существование конечного первого момента.

✧ Необходимое условие для слабого ЗБЧ. Для доказательства обратного утверждения примéним метод симметризации. Суть этого метода состоит в рассмотрении наряду с последовательностью сл.в. ξ_n разностей $\xi_n - \xi_n'$, где ξ_n' — искусственно построенная последовательность сл.в. такая, что все сл.в. $\left\langle \xi_n,\ \xi_n' \right\rangle_1^\infty$ независимы в совокупности и $\xi_n' \backsim \xi_n$ при каждом $n \geqslant 1$. В этом случае, как легко понять, вектор $(\xi_n, \xi_n') \backsim (\xi_n', \xi_n)$, и поэтому разность $\xi_n - \xi_n'$ имеет симметричное распределение, т.е. $\mathbf{P}\{\xi_n - \xi_n' \leqslant -x\} = \mathbf{P}\{\xi_n - \xi_n' \geqslant x\}$ при $\forall x \in \mathbb{R}^1$. В качестве центра распределения симметричной сл.в. можно взять точку $x = 0$, т.к. эта точка есть одна из медиан распределения: $\mathbf{P}\{\xi_n - \xi_n' \leqslant 0\} = \mathbf{P}\{\xi_n - \xi_n' \geqslant 0\}$ $(= 1/2$, если ф.р. ξ_n непрерывна в нуле). Связь между распределениями исходной сл.в. и её симметризованного аналога раскрывают так называемые неравенства симметризации (доказательство см. на стр. 303).

299 **Лемма.** [*Неравенства симметризации.*] *Пусть $\zeta_1, \ldots, \zeta_n$ — независимые сл.в., тогда $\forall t > 0$:*

(a) *если $\zeta_1 \backsim \zeta_2$, то* $\mathbf{P}\{|\zeta_1| \geqslant t\} \ \geqslant\ \frac{1}{2}\mathbf{P}\{|\zeta_1 - \zeta_2| \geqslant 2t\}$;

(b) *если m_2 — медиана ζ_2, то*

$$\mathbf{P}\{|\zeta_1 - m_2| \geqslant t\} \ \leqslant\ 2\,\mathbf{P}\{|\zeta_1 - \zeta_2| \geqslant t\};$$

(c) *если сл.в. $\zeta_1, \ldots, \zeta_n$ имеют симметричные распределения, то распределение $S = \zeta_1 + \ldots + \zeta_n$ также симметрично и*

$$\mathbf{P}\{|S| \geqslant t\} \ \geqslant\ \tfrac{1}{2}\mathbf{P}\Big\{\max_k |\zeta_k| \geqslant t\Big\} \ =\ \tfrac{1}{2}\Big(1 - \prod_{k=1}^{n} \mathbf{P}\{\,|\zeta_k| < t\,\}\Big). \tag{7}$$

300| **Теорема.** [ЗБЧ. Необходимое условие.] *Пусть* $\langle \xi_k \rangle_1^\infty$ *— независимые в совокупности одинаково распределённые сл.в. Если выполняется ЗБЧ* (4) *с некоторой конечной константой* μ, *то*

$$(\text{I}) \quad \lim_{n\to\infty} n\,\mathbf{P}\{|\xi_1| \geqslant n\} = 0, \qquad (\text{II}) \quad \mu = \lim_{n\to\infty} \mathbf{E}\left[\xi_1\,\dot{\mathrm{I}}(|\xi_1| < n)\right].$$

В частности, если математическое ожидание $\mathbf{E}|\xi_1| < \infty$, *то* $\mu = \mathbf{E}\xi_1$.

⤇ Предположим сначала, что распределение сл.в. ξ_k симметрично. В этом случае правая часть (II) равна нулю. По условию теоремы

$$0 \;=\; \lim_n \mathbf{P}\Big\{\,\frac{1}{n}|S_n| > 1\,\Big\} \;\geqslant\; \overline{\lim_n}\,\frac{1}{2}\Big(1 - \big(\mathbf{P}\{\,|\xi_1| < n\,\}\big)^n\Big)$$

в силу неравенства (7). Итак, $\big(1 - \mathbf{P}\{\,|\xi_1| \geqslant n\,\}\big)^n = \big(\mathbf{P}\{\,|\xi_1| < n\,\}\big)^n \underset{n}{\to} 1$. Хорошо известно, что это возможно т. т. т. когда справедливо (I), т.е. $n\,\mathbf{P}\{\,|\xi_1| \geqslant n\,\} \underset{n}{\to} 0$. Отсюда следует (6) (см. доказательство (6)). Математическое ожидание второго слагаемого в (6) равно нулю (в силу симметрии), а дисперсия этого слагаемого стремится к нулю (см. доказательство теоремы 297). Таким образом, $\sum_1^n \xi_k/n \xrightarrow[n]{\mathbf{P}} 0$, т.е. $\mu = 0$, что доказывает (II).

В общем случае рассмотрим такую последовательность сл.в. $\langle\,\xi'_k,\ k \in \mathbb{N}\,\rangle$, что все сл.в. $\langle\,\xi_k, \xi'_k,\ k \in \mathbb{N}\,\rangle$ независимы в совокупности и одинаково распределены. Тогда по неравенству симметризации 299 , (a)

$$0 \;=\; \lim_n \mathbf{P}\Big\{\,\frac{1}{n}\Big|\sum_1^n(\xi_k - \mu)\Big| \geqslant \frac{1}{2}\varepsilon\,\Big\} \;\geqslant\; \lim_n \frac{1}{2}\,\mathbf{P}\Big\{\,\frac{1}{n}\Big|\sum_1^n(\xi_k - \xi'_k)\Big| \geqslant \varepsilon\,\Big\}.$$

Следовательно, ввиду уже рассмотренного симметричного случая и неравенства симметризации 299 , (b)

$$0 \;=\; \lim_n n\,\mathbf{P}\{\,|\xi_1 - \xi'_1| \geqslant n\,\} \;\geqslant\; \lim_n \frac{1}{2}\,n\,\mathbf{P}\{\,|\xi_1 - m_1| \geqslant n\,\},$$

что, очевидно, эквивалентно (I). Таким образом (см. 298), имеем сходимость

$$\frac{1}{n}S_n - \mathbf{E}\left[\xi_1\,\dot{\mathrm{I}}(|\xi_1| < n)\right] \;\xrightarrow[n]{\mathbf{P}}\; 0,$$

что доказывает (II), т.к. разность последовательностей, сходящихся по вероятности к одной величине, сходится к нулю. ⤆

✧ Бесконечное математическое ожидание. Рассмотрим теперь ситуацию с бесконечным математическим ожиданием слагаемых. Напомним, что равенство $\mathbf{E}\xi = +\infty$ означает конечность математического ожидания $\mathbf{E}\xi^- < \infty$ для отрицательной части $\xi^- = -\min\{\xi, 0\}$ и равенство $\mathbf{E}\xi^+ = \infty$ для положительной части $\xi^+ = \max\{\xi, 0\}$.

301| **Теорема.** *Пусть* $\langle\, \xi_k \,\rangle_1^\infty$ *— последовательность попарно независимых одинаково распределённых сл.в. с математическим ожиданием* $\mathbf{E}\,\xi_k = +\infty$. *Тогда* $\frac{1}{n}\sum_1^n \xi_k \xrightarrow[n]{\mathbf{P}} +\infty$, *т.е. для* $\forall\, M > 0$

$$\lim_{n\to\infty} \mathbf{P}\Big\{\, \frac{1}{n}\sum_{k=1}^n \xi_k > M \,\Big\} \;=\; 1.$$

⊨> Пусть $\xi \backsim \xi_1$. Так как $\mathbf{E}\,\xi^+ = \infty$, то для $\forall\, M > 0$ найдётся такое $h = h_{_M}$, что $\mathbf{E}\left[(\xi^+)^{(h)}\right] > 2M$. По ЗБЧ Чебышёва для ограниченных сл.в.

$$\mathbf{P}\Big\{\, \frac{1}{n}\sum_{k=1}^n \xi_k^+ > M \,\Big\} \;\geqslant\; \mathbf{P}\Big\{\, \frac{1}{n}\sum_{k=1}^n (\xi_k^+)^{(h)} - \mathbf{E}\left[(\xi^+)^{(h)}\right] > -M \,\Big\} \;\to\; 1$$

при $n \to \infty$.

Для среднего отрицательных частей $\sum_1^n \xi_k^- / n$ по теореме 297 справедлив ЗБЧ с $\mu = \mathbf{E}\,\xi_k^-$. Утверждение теоремы следует из того, что сумма сходящихся по вероятности последовательностей сходится к сумме соответствующих пределов, возможно равных $+\infty$ (см. 213 и 214, стр. 199). <=|

302| ⚠ Пусть функция плотности независимых сл.величин ξ_k равна $f(x) = x^{-2},\ x > 1$. Из курса математического анализа известно, что существует интеграл Римана $\mu = \int_1^\infty x \sin x\, f(x)\, dx \approx 0.625$. Это значение не может считаться математическим ожиданием $\mathbf{E}\,\zeta_k$ для сл.в. $\zeta_k = \xi_k \sin \xi_k$, поскольку также хорошо известно, что интеграл Римана $\int_1^\infty |x \sin x| f(x)\, dx$ расходится. Моделирование показывает, что хотя ф.р. ζ_1 имеет весьма экзотический вид, однако ф.р. среднего арифметического $\frac{1}{n}\sum_1^n \zeta_k$ при $n \geqslant 50$ неотличима от ф.р. модели Коши $\mathfrak{Cauch}(\mu, 1)$, т.е. ЗБЧ для этих сл.в. не выполняется. Это демонстрирует важность для теории вероятностей определения математического ожидания именно как интеграла Лебега, а не как интеграла Римана.

В ситуациях с бесконечным математическим ожиданием закону больших чисел можно придать несколько иную формулировку.

303| **Теорема.** *Пусть* $\langle\, \xi, \xi_k,\ k \in \mathbb{N} \,\rangle$ *— неотрицательные попарно независимые одинаково распределённые сл.в. с математическим ожиданием* $\mathbf{E}\,\xi_k = +\infty$. *Если найдётся последовательность* $h_n \nearrow \infty$ *такая, что*

$n\,\mathbf{P}\{\xi \geqslant h_n\} \underset{n}{\to} 0$ *и* $h_n/(n\mu_n) \underset{n}{\to} 0$ *с* $\mu_n = \mathbf{E}\left[\xi\,\dot{\mathrm{I}}(\xi < h_n\right]$, *то*

$$\frac{1}{n\,\mu_n}\sum_1^n \xi_k \xrightarrow[n]{\mathbf{P}} 1. \tag{8}$$

⚠ Как замечено в [22], существование указанной последовательности h_n обеспечивает условие

$$\lim_{t\to\infty}\frac{t\,\mathbf{P}\{\xi \geqslant t\}}{\mathbf{E}\left[\xi\,\dot{\mathrm{I}}(\xi < t)\right]} = 0. \tag{9}$$

⇛ Для $\forall\, h_n > 0$ рассмотрим усечённые сл.в. $\xi_{kn} = \xi_k\,\dot{\mathrm{I}}(\xi_k < h_n)$ и соответствующую сумму $S_{nh_n} = \sum_{k=1}^n \xi_{kn}$. Очевидно, $\mu_n = \mathbf{E}\,\xi_{kn}$. Заметим, что любое событие B можно представить в виде объединения $B = \left[B\bigcap_k(\xi_k = \xi_{kn})\right] \uplus \left[B\bigcup_k(\xi_k \neq \xi_{kn})\right]$. Поэтому для $\forall\,\varepsilon > 0$ по неравенству Чебышёва

$$\mathbf{P}\big\{\,|S_n - n\mu_n| \geqslant n\mu_n\varepsilon\,\big\} \;\leqslant\; \mathbf{P}\big\{\,|S_{nh_n} - n\mu_n| \geqslant n\mu_n\varepsilon\,\big\} + \sum\nolimits_{k=1}^n \mathbf{P}\{\xi_k \neq \xi_{kn}\} \;\leqslant$$
$$\leqslant\; \frac{1}{n\varepsilon^2\mu_n^2}\,\mathbf{E}\left[\xi^2\,\dot{\mathrm{I}}(\xi < h_n)\right] + n\,\mathbf{P}\{\xi \geqslant h_n\} \;\leqslant$$
$$\leqslant\; \frac{h_n}{n\varepsilon^2\mu_n} + n\,\mathbf{P}\{\xi \geqslant h_n\}\,.$$

По условию теоремы правая часть здесь стремится к нулю. ⇚

304| Пример. Пусть ф.р. ξ равна $F(x) = 1 - x^{-1}, x \geqslant 1$, т.е. плотность $f(x) = x^{-2},\ x \geqslant 1$. Тогда $\mathbf{E}\left[\xi\,\dot{\mathrm{I}}(\xi < t)\right] = \int_1^t x\frac{1}{x^2}\,dx = \ln t \underset{t\to\infty}{\longrightarrow} \infty$. Условие (9) выполняется:

$$\lim_{t\to\infty}\frac{t\,\mathbf{P}\{\xi \geqslant t\}}{\mathbf{E}\left[\xi\,\dot{\mathrm{I}}(\xi < t)\right]} = \lim_{t\to\infty}\frac{t\cdot 1/t}{\ln t} = 0.$$

Если взять $h_n = n\ln^q n$ с $0 < q < 1$, то

$$n\,\mathbf{P}\{\xi \geqslant h_n\} = \frac{n}{n\ln^q n} \underset{n}{\to} 0, \qquad \frac{h_n}{n\mu_n} = \frac{n\ln^q n}{n\ln(n\ln^q n)} = \frac{\ln^q n}{\ln n + q\ln(\ln n)} \underset{n}{\to} 0.$$

Таким образом, здесь (8) справедливо с $\mu_n = \ln n + q\ln(\ln n)$ с произвольным $q \in (0;1)$. Поскольку, очевидно, $\mu_n/\ln n \underset{n}{\to} 1$, то (8) справедливо и для $\mu_n = \ln n$, т.е. при $q = 0$, хотя выбор q порядка $0.9 - 0.95$ повышает точность асимптотического утверждения. ◎

⚠ Можно показать (см. Феллер, [22]), что условие (9) является и необходимым для того, чтобы (8) имело место для какой-либо последовательности констант μ_n.

✧ **Схема серий**. Прежде чем переходить к обсуждению ЗБЧ для сл.в., имеющих различное распределение, заметим следующее. Среднее арифметическое при фиксированном n можно записать в виде суммы:

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k = \sum_{k=1}^{n}\eta_{nk}$$

с $\eta_{nk} = \xi_k/n,\ k = 1,\ldots,n$. Как видно из представленных доказательств, анализ поведения средних связан с асимптотическим поведением набора сл.в. $\eta_{nk},\ k = 1,\ldots,K_n,\ K_n \to \infty$ при $n \to \infty$. Такие наборы принято называть *сериями* (по n) или треугольными массивами. При этом с изменением n сл.в., входящие в различные серии, будут иметь различные распределения. Дабы не усложнять формулировки теорем, будем считать, что $K_n = n,\ n \geqslant 1$.

Почти дословно повторив вторую часть доказательства ЗБЧ 297, можно установить справедливость следующей теоремы (в части достаточности; о необходимости условий см. [32]).

305 **Теорема.** *Пусть* $\left\langle \xi_{nk},\ k = \overline{1,n} \right\rangle_{n=1}^{\infty}$ *— последовательность серий попарно независимых в каждой серии сл.в.,* $S_n = \sum_{k=1}^{n}\xi_{nk}$.

I) *Если*

$$\text{(a)}\quad \lim_{n\to\infty}\sum_{k=1}^{n}\mathbf{P}\left\{|\xi_{nk}| \geqslant 1\right\} = 0, \qquad \text{(b)}\quad \lim_{n\to\infty}\sum_{k=1}^{n}\mathbf{D}\left[\xi_{nk}^{(1)}\right] = 0, \tag{10}$$

то

$$S_n - \mu_n \xrightarrow[n]{\mathbf{P}} 0,$$

где $\mu_n = \sum_{k=1}^{n}\mathbf{E}\left[\xi_{nk}^{(1)}\right] = \sum_{k=1}^{n}\mathbf{E}\left[\xi_{nk}\,\dot{\mathrm{I}}\left(|\xi_{nk}| < 1\right)\right]$.

II) *Обратно, если* $S_n - \mu_n \xrightarrow[n]{\mathbf{P}} 0$ *с некоторыми* μ_n *и* $\max_{1\leqslant k\leqslant n}|m(\xi_{nk})| \xrightarrow[n]{} 0$, *где* $m(\xi_{nk})$ *— медиана сл.в.* ξ_{nk}, *то справедливы соотношения* (10); *при этом разность* $\mu_n - \sum_{k=1}^{n}\mathbf{E}\left[\xi_{nk}^{(1)}\right] \xrightarrow[n]{} 0$.

306 Пример. Пусть $\left\langle \eta_k \backsim \mathrm{Cauch}(0,\sigma_k), \sigma_k > 0 \right\rangle_{k=1}^{\infty}$ — независимые сл.в. Коши. Если масштабный параметр распределения Коши постоянен ($\sigma_k = \sigma$), то, воспользовавшись хар.ф. $\exp(-\sigma|t|)$ для распределения Коши, легко понять, что среднее арифметическое $\frac{1}{n}\sum_{1}^{n}\eta_k \backsim \mathrm{Cauch}(0,\sigma)$, т.е. ЗБЧ заведомо не выполняется (заметим, что математическое ожидание $\mathbf{E}\eta_k$ не существует). Определим $\xi_{nk} = \frac{1}{n}\eta_k,\ k = 1,\ldots,n$. Тогда $\xi_{nk}^{(1)} = \frac{1}{n}\eta_k^{(n)}$ и, следовательно,

$$\text{(a)}\quad \sum_{k=1}^{n}\mathbf{P}\{|\xi_{nk}| \geqslant 1\} = \sum_{k=1}^{n}\mathbf{P}\{|\eta_k| \geqslant n\} = \sum_{k=1}^{n}\Big(1 - \frac{2}{\pi}\operatorname{arctg}\big(\frac{n}{\sigma_k}\big)\Big),$$

$$\text{(b)}\quad \sum_{k=1}^{n}\mathbf{D}\big[\xi_{nk}^{(1)}\big] = \frac{1}{n^2}\sum_{k=1}^{n}\mathbf{D}\big[\eta_k^{(n)}\big] = \frac{2}{\pi n^2}\sum_{k=1}^{n}\sigma_k^2\Big(\frac{n}{\sigma_k} - \operatorname{arctg}\big[\frac{n}{\sigma_k}\big]\Big).$$

Таким образом, при $\sigma_k \equiv \sigma$ предел в (a) равен $\lim_n n(1 - \frac{2}{\pi}\operatorname{arctg}(\frac{n}{\sigma})) = 2\sigma/\pi$ и совпадает с аналогичным пределом в (b).

Несложными рассуждениями показывается, что условия (10) здесь выполняются, если $\lim_k \sigma_k = 0$; при этом, очевидно, $\mu_n = 0$. Таким образом, ЗБЧ может иметь место и для сл.в. Коши, если характеристика разброса σ с каждым проведённым наблюдением уменьшается (в пределе до нуля). Совсем просто этот результат получается, если воспользоваться методом хар.функций:

$$\varphi_{S_n}(t) = \exp\Big\{-\frac{1}{n}\sum\nolimits_1^n \sigma_k|t|\Big\} \to e^0 = 1$$

по первой части леммы 293. ◎

307| Упр. Покажите, что для независимых сл.в. $\langle \xi_k \rangle_1^\infty$ с $\mathbf{P}\{\xi_k = \pm k\} = 1/2k^2$, $\mathbf{P}\{\xi_k = 0\} = 1 - 1/k^2$ справедливо $\frac{1}{\sqrt{n}}\sum_1^n \xi_k \xrightarrow[n]{\mathbf{P}} 0$.

Немного усилив условия теоремы, можно доказать простыми методами не только сходимость по вероятности, но даже сходимость в среднем.

308| **Теорема.** *Пусть $\langle \xi_{nk},\ k = \overline{1,n}\rangle_{n=1}^\infty$ — последовательность серий попарно независимых в каждой серии сл.в., $S_n = \sum_{k=1}^n \xi_{nk}$. Если выполняются следующие условия:*

a) $\mathbf{E}\xi_{nk} = 0$ *для* $\forall\, k, n$;

b) $\lim\limits_{n\to\infty}\sum_{k=1}^n \mathbf{E}\big[|\xi_{nk}|\ \dot{\mathrm{I}}(|\xi_{nk}| > 1)\big] = 0$;

c) $\lim\limits_{n\to\infty}\sum_{k=1}^n \mathbf{E}\big[\xi_{nk}^2\ \dot{\mathrm{I}}(|\xi_{nk}| \leqslant 1)\big] = 0$,

то $\mathbf{E}\big|S_n\big| \underset{n}{\to} 0$, *т.е.* $S_n \xrightarrow{\mathcal{L}_1} 0$.

⇛ Как обычно, положим $\eta^{(1)} = \eta\,\dot{\mathrm{I}}(|\eta| \leqslant 1)$; тогда $\eta - \eta^{(1)} = \eta\,\dot{\mathrm{I}}(|\eta| > 1)$. Для суммы $S_n' = \sum \xi_{nk}^{(1)}$ (здесь и далее суммирование ведётся по $k = \overline{1,n}$)

имеем по неравенству Коши–Буняковского

$$\mathbf{E}\left|S'_n - \mathbf{E}\,S'_n\right| \leqslant \sqrt{\mathbf{E}\left[\left(S'_n - \mathbf{E}\,S'_n\right)^2\right]} = \sqrt{\sum \mathbf{D}\xi^{(1)}_{kn}} \leqslant$$

$$\leqslant \sqrt{\sum \mathbf{E}\left[\left(\xi^{(1)}_{kn}\right)^2\right]} \xrightarrow{\text{c)}} 0.$$

Для дополнительной суммы $S''_n = \sum(\xi_{kn} - \xi^{(1)}_{nk})$ по неравенству треугольника

$$\mathbf{E}\left|S''_n - \mathbf{E}\,S''_n\right| \leqslant \mathbf{E}\left|S''_n\right| + \left|\mathbf{E}\,S''_n\right| \leqslant 2\sum \mathbf{E}\left[|\xi_{nk}|\ \dot{\mathrm{I}}\left(|\xi_{nk}| > 1\right)\right] \xrightarrow{\text{b)}} 0.$$

Доказательство теоремы следует теперь из того, что по условию a) математическое ожидание $\mathbf{E}\,S_n = 0 = \mathbf{E}\,S'_n + \mathbf{E}\,S''_n$. ⫤

§2. Усиленные законы больших чисел

Название *усиленный закон больших чисел* (УЗБЧ) закрепилось за законами, в которых утверждается факт стабилизации сумм сл.в. в пределе около детерминированных величин в смысле сходимости почти наверное. Отметим, что для УЗБЧ важно, чтобы все входящие в сумму сл.в. были заданы на одном вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$; фраза о том, что некоторое утверждение выполняется почти наверное ((п.н.), [**P**-п.н.]), относится к событиям именно этого пространства. В дальнейшем в этом разделе мы будем, не оговариваясь, предполагать наличие этого условия.

ЗБЧ Бореля, который мы привели в 207, стр. 195, в качестве демонстрации возможностей следствия 205, стр. 194, из леммы Бореля–Кантелли 204, стр. 194, есть один из примеров УЗБЧ. При доказательстве этого закона использовались свойства биномиального распределения. Доказательство можно значительно сократить, если воспользоваться неравенством Гёфдинга 358, стр. 309. Так как бернуллиевская сл.в. $|\xi_k| \leqslant X = 1$, то, выбирая в этом неравенстве $t = \varepsilon_n = \sqrt{4n^{-1}\ln n} \underset{n}{\to} 0$, имеем для независимых бернуллиевских сл.в. с вероятностью успеха p:

$$\mathbf{P}\left\{\left|\frac{1}{n}\sum\nolimits_{k=1}^{n}(\xi_k - p)\right| > \varepsilon_n\right\} \leqslant 2\exp\left\{-\frac{n^2\varepsilon_n^2}{2n}\right\} = \frac{2}{n^2}.$$

Поскольку $\sum_1^\infty n^{-2} < \infty$, то в силу 205 получаем, что для последовательности $\left\langle \xi_k \right\rangle_1^\infty$ независимых бернуллиевских сл.в. с вероятностью успеха p имеет место УЗБЧ $\sum_1^n \xi_k/n \underset{n}{\to} p$ (п.н.) .

309| Примеры. Приведём две ситуации, когда справедлив слабый, но не усиленный ЗБЧ. Заметим сначала, что если для сл.в. $\langle \xi_k \rangle_1^\infty$ при некотором ω имеем $\frac{1}{n} S_n(\omega) \to 0$, то $\frac{1}{n}\xi_n(\omega) = \frac{1}{n}(S_n(\omega) - S_{n-1}(\omega)) \to 0$. Следовательно, при $\forall\, \varepsilon > 0$ события $|\xi_n(\omega)| > n\varepsilon$ не могут происходить бесконечно часто. Поэтому если для независимых сл.в. $\mathbf{P}\{S_n/n \to 0\} = 1$, то по второй части леммы Бореля–Кантелли $\sum_1^\infty \mathbf{P}\{|\xi_n| > n\varepsilon\} < \infty$. Следовательно, если $\sum_1^\infty \mathbf{P}\{|\xi_n| > n\varepsilon\} = \infty$, то $\mathbf{P}\{S_n/n \not\to 0\} = 1$.

1) Пусть $\langle \eta_k \backsim \mathfrak{Cauch}(0, \lambda_k), \lambda_k > 0 \rangle_{k=1}^\infty$ — независимые сл.в. Коши. Как показано в 306, стр. 263, при $\lambda_k \to 0$ справедлив ЗБЧ: $\frac{1}{n}\sum_1^n \xi_k \xrightarrow{\mathbf{P}} 0$. С другой стороны, с помощью правила Лопиталя легко показать, что $\lambda_n = n/\operatorname{tg}\left(\frac{\pi}{2}(1 - \delta_n)\right) \to 0$ для $\delta_n = 1/(n \ln n)$. Поэтому $\mathbf{P}\{|\xi_n| > n\varepsilon\} = \left(1\frac{2}{\pi}\operatorname{arctg}(n/\lambda_n)\right) = \delta_n$, т.е. $\sum_1^\infty \mathbf{P}\{|\xi_n| > n\varepsilon\} = \infty$.

2)[20]. Пусть независимые дискретные сл.в. $\langle \xi_k \rangle_3^\infty$ таковы, что $\mathbf{P}\{\xi_k = \pm k\} = 1 - \frac{1}{2}\mathbf{P}\{\xi_k = 0\} = 1/(2k \ln k)$. Аналогично предыдущему примеру показывается, что $\mathbf{P}\{S_n/n \not\to 0\} = 1$. Так как $\mathbf{D}\xi_k = k/\ln k$ и функция $x/\ln x$ возрастает при $x \geqslant 3$, то

$$\frac{1}{n^2}\sum\nolimits_3^n \mathbf{D}\xi_k = \frac{1}{n^2}\sum\nolimits_3^n \frac{k}{\ln k} \leqslant \frac{1}{n^2}\frac{n}{\ln n}(n-3) \underset{n}{\to} 0.$$

В соответствии с теоремой 294, стр. 256, $\frac{1}{n} S_n \xrightarrow{\mathbf{P}} 0$. ◎

С помощью 205 можно доказать более общее утверждение, если учесть наличие четвёртого момента у слагаемых.

310| **Теорема.** [*Кантелли.*] *Пусть $\langle \xi_k \rangle_1^\infty$ — независимые сл.в. с конечными математическими ожиданиями $\mu_k = \mathbf{E}\,\xi_k$ и равномерно ограниченными четвёртыми моментами $\mathbf{E}(\xi_k - \mu_k)^4 < \varkappa < \infty$, $k \geqslant 1$. Тогда*

$$\lim_n \left(\frac{1}{n}\sum_{k=1}^n \xi_k - \frac{1}{n}\sum_{k=1}^n \mu_k\right) = 0 \quad (\text{п.н.}). \tag{11}$$

⤇ Как всегда в таких случаях, можно считать $\mu_k = 0$. Рассмотрим величину четвёртого момента суммы $S_n = \sum_1^n \xi_k$:

$$\mathbf{E}\left[S_n^4\right] = \sum_{k,i,j,m=1}^n \mathbf{E}\left[\xi_k \xi_i \xi_j \xi_m\right].$$

Ясно, что все слагаемые здесь конечны и по модулю не превосходят $\varkappa$:

$$\left|\,\mathbf{E}\,\xi_k\xi_i\xi_j\xi_m\right| \leqslant \sqrt{\mathbf{E}\left[\xi_k^2\xi_i^2\right]\,\mathbf{E}\left[\xi_j^2\xi_m^2\right]} \leqslant \sqrt{\sqrt{\mathbf{E}\,\xi_k^4\;\mathbf{E}\,\xi_i^4}\,\sqrt{\mathbf{E}\,\xi_j^4\;\mathbf{E}\,\xi_m^4}} < \varkappa$$

в силу неравенства Коши–Буняковского. Более того, если один из индексов в слагаемом не совпадает ни с одним из остальных, то такое слагаемое равно нулю. Например, если $m \neq i, m \neq j, m \neq k$, то ввиду независимости $\mathbf{E}\left[\xi_k\xi_i\xi_j\xi_m\right] = \mathbf{E}\left[\xi_k\xi_i\xi_j\right]\mathbf{E}\left[\xi_m\right] = 0$. Очевидно, имеется всего n ненулевых слагаемых вида $\mathbf{E}\left[\xi_k^4\right]$ и $\mathrm{C}_n^2\,\mathrm{C}_4^2 = 3n(n-1)$ слагаемых вида $\mathbf{E}\left[\xi_k^2\xi_j^2\right]$. Итак,

$$\mathbf{E}\left[S_n^4\right] < \varkappa(n + 3n(n-1)) < 3\varkappa n^2.$$

Следовательно, для $\forall\varepsilon > 0$ по неравенству Маркова

$$\sum_{n=1}^{\infty}\mathbf{P}\Big\{\Big|\frac{1}{n}\sum_{k=1}^{n}(\xi_k - \mu_k)\Big| > \varepsilon\Big\} \leqslant \frac{1}{\varepsilon^4}\sum_{n=1}^{\infty}\frac{1}{n^4}\mathbf{E}\left[S_n^4\right] \leqslant \frac{3\varkappa}{\varepsilon^4}\sum_{n=1}^{\infty}\frac{1}{n^2} < \infty.$$

Утверждение настоящей теоремы следует теперь из 205, стр. 194. ⇚

✧ **ЗБЧ Колмогорова для разнораспределённых сл.величин.** Можно потребовать лишь существования второго момента. Заметим, что метод доказательства предыдущей теоремы, применённый к $\mathbf{E}\big(S_n/n\big)^2$, не даст желаемого результата, поскольку получаемый в этом случае ряд $\sum_1^\infty \frac{1}{n} = \infty$. Основная идея доказательства следующей теоремы, как и многих других, близких к УЗБЧ, эксплуатирует лемму Кронекера 293, стр. 255. В упрощённом варианте леммы утверждается, что равенство $\lim_n \frac{1}{n}\sum_{k=1}^n a_k = 0$ имеет место, если сходится ряд $\sum_{k=1}^\infty (a_k/k)$ к какой-либо конечной величине. Другой инструмент, часто применяемый при анализе распределений сумм независимых сл.в., — неравенство Колмогорова, представляющее собой усиление неравенства Чебышёва.

311| **Лемма.** [*Неравенство Колмогорова.*] *Если* $\langle\,\xi_k\,\rangle_1^n$ *— независимые сл.в. с нулевым математическим ожиданием и* $S_k = \xi_1 + \ldots + \xi_k$, $k = \overline{1,n}$, *то для* $\forall\varepsilon > 0$

$$\mathbf{P}\Big\{\max_{1\leqslant k\leqslant n}\big|S_k\big| > \varepsilon\Big\} \leqslant \frac{1}{\varepsilon^2}\mathbf{D}S_n = \frac{1}{\varepsilon^2}\sum\nolimits_1^n\mathbf{E}\left[\xi_k^2\right]. \tag{12}$$

⇛ Очевидно, искомое событие $A = \{\max_{1\leqslant k\leqslant n}\big|S_k\big| > \varepsilon\} = \biguplus_1^n B_k$, где

$$B_k = \{\big|S_k\big| > \varepsilon,\ \max_{1\leqslant j<k}\big|S_j\big| \leqslant \varepsilon\}.$$

В силу независимости любых функций вида $h(\xi_1, \ldots, \xi_k)$ от функций вида $g(\xi_{k+1}, \ldots, \xi_n)$ и равенства нулю математических ожиданий $\mathbf{E}\,\xi_j = 0$, $j \geqslant 1$, имеем

$$\mathbf{E}\left[S_k\,\dot{\mathrm{I}}_{B_k}(S_n - S_k)\right] \;=\; \mathbf{E}\left[S_k\,\dot{\mathrm{I}}_{B_k}\right]\,\mathbf{E}\left[(S_n - S_k)\right] \;=\; 0.$$

Поэтому

$$\mathbf{E}\left[S_n^2\,\dot{\mathrm{I}}_{B_k}\right] \;=\; \mathbf{E}\left[(S_k + (S_n - S_k))^2\,\dot{\mathrm{I}}_{B_k}\right] \;=\; \mathbf{E}\left[S_k^2\,\dot{\mathrm{I}}_{B_k}\right] + \mathbf{E}\left[(S_n - S_k)^2\,\dot{\mathrm{I}}_{B_k}\right] \;\geqslant$$
$$\geqslant\; \mathbf{E}\left[S_k^2\,\dot{\mathrm{I}}_{B_k}\right] \;\geqslant\; \varepsilon^2\,\mathbf{P}\{B_k\},$$

т.е.

$$\mathbf{P}\{A\} \;=\; \sum\nolimits_1^n \mathbf{P}\{B_k\} \;\leqslant\; \frac{1}{\varepsilon^2}\,\mathbf{E}\left[S_n^2 \sum\nolimits_1^n \dot{\mathrm{I}}_{B_k}\right] \;=$$
$$=\; \frac{1}{\varepsilon^2}\,\mathbf{E}\left[S_n^2\,\dot{\mathrm{I}}_A\right] \;\leqslant\; \frac{1}{\varepsilon^2}\,\mathbf{E}\left[S_n^2\right]. \qquad \Lleftarrow$$

312| **Теорема.** [*А. Н. Колмогоров.*] *Пусть* $\langle\,\xi_k\,\rangle_1^\infty$ *— независимые в совокупности сл.в. с конечными математическими ожиданиями* $\mu_k = \mathbf{E}\,\xi_k$ *и дисперсиями* $\sigma_k^2 = \mathbf{D}\xi_k,\; k \geqslant 1$, *такими, что* $\sum_1^\infty \frac{\sigma_k^2}{k^2} < \infty$. *Тогда справедлив УЗБЧ* (11).

$\Rrightarrow$ Снова можно считать $\mu_k = 0$. Определим вспомогательную сумму $V_n = \sum_1^n \xi_k/k,\;\; n \geqslant 1,\;\; V_0 = 0$. Как замечено выше, по лемме Кронекера достаточно показать, что почти наверное существует конечный предел $V = \lim_n V_n$.

Докажем фундаментальность (п.н.) последовательности $V_n,\; n \geqslant 1$. Рассмотрим множество $A_n(\varepsilon) = \{\omega :\; \sup_{m,\,k\geqslant n}\left|V_k(\omega) - V_m(\omega)\right| > \varepsilon\}$. Последовательность $V_n,\; n \geqslant 1$, не удовлетворяет условию Коши (условию фундаментальности) в тех $\omega \in \Omega$, которые принадлежат подмножеству

$$A \;=\; \bigcup\nolimits_{l=1}^\infty \bigcup\nolimits_{n=1}^\infty A_n\left({}^1\!/_l\right) \;=\; \lim_{l}{}_{\uparrow}\; \lim_{n}{}_{\uparrow}\, A_n\left({}^1\!/_l\right).$$

Заметим, что $\max\limits_{m,\,k}\left|V_k - V_m\right| \leqslant \max\limits_{m}\left|V_m - V_n\right| + \max\limits_{k}\left|V_k - V_n\right|$. Поэтому если $\max\limits_{m,\,k}\left|V_k - V_m\right| > \varepsilon$, то $\max\limits_{k}\left|V_k - V_n\right| > \frac{1}{2}\varepsilon$ либо $\max\limits_{m}\left|V_m - V_n\right| > \frac{1}{2}\varepsilon$. Вероятности последних двух событий (в ситуации, рассмотренной далее) совпадают. Следовательно, для $\forall\,\varepsilon > 0,\; n \geqslant 1$

$$\mathbf{P}\Big\{\sup_{m,\,k\geqslant n}\left|V_k - V_m\right| > \varepsilon\Big\} \;=\; \lim_{K\nearrow\infty}\mathbf{P}\Big\{\max_{n\leqslant m,\,k\leqslant K}\left|V_k - V_m\right| > \varepsilon\Big\} \;\leqslant$$
$$\leqslant\; 2\lim_{K\nearrow\infty}\mathbf{P}\Big\{\max_{n\leqslant k\leqslant K}\left|V_k - V_n\right| > \tfrac{1}{2}\varepsilon\Big\}.$$

Применяя к приращениям $V_k - V_n = \sum\limits_{j=n+1}^{k} \frac{\xi_j}{j}$, $k = \overline{n+1, K}$, неравенство Колмогорова (12), получаем

$$\mathbf{P}\Big\{ \sup_{m,\,k\geqslant n} \big|V_k - V_m\big| > \varepsilon \Big\} \ \leqslant\ \frac{8}{\varepsilon^2} \lim_{K\to\infty} \mathbf{D}\Big[\sum_{j=n+1}^{K} \frac{\xi_j}{j} \Big] \ <\ \frac{8}{\varepsilon^2} \sum_{j=n+1}^{\infty} \frac{\sigma_j^2}{j^2} \ \underset{n}{\to}\ 0$$

ввиду сходимости ряда $\sum_1^\infty \sigma_j^2/j^2$. Таким образом,

$$\mathbf{P}\{A\} \ =\ \lim_l \Big(\lim_n \mathbf{P}\Big\{ \sup_{m,\,k\geqslant n} \big|V_k - V_m\big| > 1/l \Big\}\Big) \ =\ 0.$$

То есть для всех $\omega \notin A$ последовательность $V_n(\omega)$, $n \geqslant 1$, фундаментальна, и, следовательно, существует конечное $V(\omega) = \lim_n V_n(\omega)$. Функция $V(\omega)\,\dot{\mathrm{I}}_{A^c}(\omega)$, $\omega \in \Omega$ измерима как предел измеримых функций $V_n(\omega)\,\dot{\mathrm{I}}_{A^c}(\omega)$, и последовательность сл.в. V_n почти наверное сходится к $V\,\dot{\mathrm{I}}_{A^c}$. ⇚

При доказательстве сходимости последовательности V_n использовались только центрированность слагаемых и сходимость ряда из дисперсий слагаемых, входящих в определение V_n. Другими словами, по ходу доказательства предыдущей теоремы была установлена справедливость так называемой теоремы о двух рядах.

313| **Теорема.** [О двух рядах. Колмогоров–Хинчин.] *Пусть* $\langle \xi_k \rangle_1^\infty$ — *независимые сл.в. Ряд* $\sum_1^\infty \xi_k$ *сходится п.н., если сходятся два ряда из математических ожиданий и из дисперсий:* $\big|\sum_1^\infty \mathbf{E}\,\xi_k\big| < \infty$, $\sum_1^\infty \mathbf{D}\xi_k < \infty$.

Практически повторив предыдущие рассуждения, можно доказать следующее утверждение.

314| **Лемма.** (?!) *Пусть* $\langle \xi_k \rangle_1^\infty$ — *независимые сл.в. с конечными математическими ожиданиями* $\mu_k = \mathbf{E}\,\xi_k$ *и дисперсиями* $\sigma_k^2 = \mathbf{D}\xi_k$, $k \geqslant 1$, *такими, что* $\sum_1^\infty \frac{\sigma_k^2}{a_k^2} < \infty$, $a_k \to \infty$. *Тогда справедливо*

$$\lim_n \frac{1}{a_n} \sum\nolimits_1^n (\xi_k - \mu_k) \ =\ 0 \quad (\text{п.н.}).$$

315| Пример. Если дисперсии $\sigma_k^2 < D < \infty$, то для центрированных сл.в., наряду с $\frac{1}{n}\sum_1^n \xi_k \underset{n}{\to} 0$ (п.н.), можно утверждать, что и $\frac{1}{\sqrt{n}\ln n}\sum_1^n \xi_k \underset{n}{\to} 0$ (п.н.), т.к. ряд $\sum_2^\infty 1/(n\ln^2 n) < \infty$. ◎

В классическом УЗБЧ Колмогорова не предполагается даже наличия второго момента у слагаемых. Вместо УЗБЧ Колмогорова мы приведём здесь сначала чуть более сильное утверждение относительно среднего попарно независимых сл.в., доказательство которого основано на представлении сл.в. в виде разности двух неотрицательных сл.в. Предварим это утверждение двумя вариантами УЗБЧ для различно распределённых некоррелированных слагаемых, усиливающих теорему 312.

316| **Лемма.** *Пусть* $\left\langle \xi_k \right\rangle_1^\infty$ *— неотрицательные попарно некоррелированные сл.в. с конечными математическими ожиданиями* $\mu_k = \mathbf{E}\xi_k$ *и дисперсиями* $\sigma_k^2 = \mathbf{D}\xi_k$ *такими, что*

$$\text{a)}\quad \sup_{k\geqslant 1} \mathbf{E}\xi_k \;<\; B \;<\; \infty, \qquad\qquad \text{b)}\quad \sum_{k=1}^{\infty} \frac{1}{k^2}\mathbf{D}\xi_k \;<\; \infty. \tag{13}$$

Тогда справедлив УЗБЧ (11).

$\Rrightarrow$ Пусть $\lambda > 1$ и $n_k = \llbracket \lambda^k \rrbracket$ (здесь округление вниз). Заметим, что

$$\sum_{k=1}^{\infty} \frac{1}{n_k^2} \sum_{j=1}^{n_k} \sigma_j^2 \;\overset{*}{=}\; \sum_{j=1}^{\infty} \sigma_j^2 \sum_{k\geqslant \log_\lambda j} \frac{1}{n_k^2} \;\leqslant\; \sum_{j=1}^{\infty} \frac{\sigma_j^2}{j^2} \sum_{m=0}^{\infty} \frac{1}{\lambda^{2(m-1)}} \;<\; \infty,$$

где $\overset{*}{=}$ получено путём изменения порядка суммирования, а последующее неравенство — с помощью замены переменной суммирования $k \dashrightarrow \llbracket \log_\lambda j \rrbracket + m$, $m \geqslant 0$, и очевидного неравенства $n_k = \llbracket \lambda^{\lfloor \log_\lambda j \rfloor + m} \rrbracket \geqslant j\lambda^{m-1}$. Следовательно, после применения неравенства Чебышёва, имеем

$$\sum_{k=1}^{\infty} \mathbf{P}\Big\{ \frac{1}{n_k} \big| S_{n_k} - \mathbf{E}\, S_{n_k} \big| > \varepsilon \Big\} \;\leqslant\; \frac{1}{\varepsilon^2} \sum_{k=1}^{\infty} \frac{1}{n_k^2} \sum_{j=1}^{n_k} \sigma_j^2 \;<\; \infty.$$

По следствию 205, стр. 194, из леммы Бореля–Кантелли получаем из этого неравенства, что $\frac{1}{n_k}\big(S_{n_k} - \mathbf{E}\, S_{n_k}\big) \underset{k}{\to} 0$ для всех $\omega \notin \Omega_\lambda \subset \Omega$, где $\mathbf{P}\{\Omega_\lambda\} = 0$. В дальнейшем мы можем выбирать λ из (счётного) множества рациональных чисел Q. Ясно, что для объединения $\Omega_0 = \bigcup_{\lambda\in Q} \Omega_\lambda$ мера $\mathbf{P}\{\Omega_0\} = 0$ и вне Ω_0 указанная сходимость имеет место при любом рациональном $\lambda > 1$.

Зафиксируем $\omega \notin \Omega_0$ и для заданного $\varepsilon > 0$ выберем рациональное $\lambda > 1$ так, чтобы $(\lambda - 1)B < \frac{1}{2}\varepsilon$. Заметим, что

$$\begin{aligned} \varlimsup_k \frac{1}{n_k}\, \mathbf{E}(S_{n_{k+1}} - S_{n_k}) \;&=\; \varlimsup_k \frac{1}{n_k}\, \mathbf{E} \textstyle\sum_{j=n_k+1}^{n_{k+1}} \xi_j \;\leqslant\; \displaystyle\varlimsup_k \frac{1}{n_k} (n_{k+1} - n_k) B \;= \\ &=\; \varlimsup_k \frac{1}{\llbracket \lambda^k \rrbracket} (\llbracket \lambda^{k+1} \rrbracket - \llbracket \lambda^k \rrbracket) B \;=\; (\lambda - 1) B \;<\; \frac{1}{2}\varepsilon\,. \end{aligned}$$

Ввиду неотрицательности рассматриваемых сл.в. при любых $n_k \leqslant n \leqslant n_{k+1}$

$$\frac{1}{n}\big(S_n - \mathbf{E}\,S_n\big) \leqslant \frac{1}{n}\big(S_{n_{k+1}} - \mathbf{E}\,S_{n_k}\big) = \\ = \frac{1}{n_{k+1}}\big(S_{n_{k+1}} - \mathbf{E}\,S_{n_{k+1}}\big)\frac{n_{k+1}}{n} + \frac{1}{n_k}\mathbf{E}\big(S_{n_{k+1}} - S_{n_k}\big)\frac{n_k}{n}.$$

Так как $n_{k+1}/n \leqslant \lambda$, то первое слагаемое здесь стремится к нулю в силу выбора $\omega \notin \Omega_0$. Верхний предел второго слагаемого, как отмечалось выше, не превосходит $\frac{1}{2}\varepsilon$, поскольку $n_k \leqslant n$. Таким образом, $\overline{\lim\limits_n}\,\big(S_n - \mathbf{E}\,S_n\big)/n < \varepsilon$.

Аналогично показывается, что $\underline{\lim\limits_n}\,\frac{1}{n}\big(S_n - \mathbf{E}\,S_n\big) > -\varepsilon$. ⫤

Условие (13), b) обеспечивает в УЗБЧ также сходимость в смысле $\mathcal{L}_2$.

317| **Теорема.** *Пусть* $\langle \xi_k \rangle_1^\infty$ *— попарно некоррелированные сл.в. с конечными математическими ожиданиями* $\mu_k = \mathbf{E}\,\xi_k$ *и дисперсиями* $\sigma_k^2 = \mathbf{D}\xi_k$ *такими, что* $\sum_{k=1}^\infty \sigma_k^2/k^2 < \infty$, *тогда*

$$\lim_n \mathbf{E}\left[\frac{1}{n}\sum\nolimits_{k=1}^n(\xi_k - \mu_k)\right]^2 = 0.$$

⫢ Так как ряд $Z = \sum_1^\infty \sigma_k^2/k^2$ сходится, то для $\forall\,\varepsilon > 0$ найдётся такое N_0, что $\sum_n^\infty \sigma_k^2/k^2 < \varepsilon/2$ при всех $n > N_0$. Для дисперсии $S_n = \sum_1^n(\xi_k - \mu_k)$ при $n > \sqrt{2N_0/\varepsilon Z}$ в силу некоррелированности слагаемых верна оценка

$$\mathbf{E}\left[\frac{1}{n}S_n\right]^2 = \frac{1}{n^2}\sum_{k=1}^n \sigma_k^2 = \frac{N_0^2}{n^2}\sum_{k=1}^{N_0}\frac{1}{N_0^2}\sigma_k^2 + \sum_{j=N_0+1}^n \frac{1}{n^2}\sigma_j^2 \leqslant \\ \leqslant \frac{N_0^2}{n^2}\sum_{k=1}^\infty \frac{1}{k^2}\sigma_k^2 + \sum_{j=N_0+1}^\infty \frac{1}{j^2}\sigma_j^2 < \varepsilon,$$

что и требовалось. ⫤

318| **Теорема.** [*Н. Этимади.*] *Пусть* $\langle \xi_k \rangle_1^\infty$ *— попарно независимые сл.в. с конечными математическими ожиданиями* $\mu_k = \mathbf{E}\,\xi_k$ *такие, что*

$$\text{a)}\ \sup_{k\geqslant 1}\mathbf{E}\,|\xi_k| < B < \infty, \qquad \text{b)}\ \sum_{k=1}^\infty \frac{1}{k^2}\mathbf{E}\left[\xi_k^2\,\dot{\mathrm{I}}\,(|\xi_k| < k)\right] < \infty, \\ \text{c)}\ \sum_{k=1}^\infty \frac{1}{k}\mathbf{E}\left[|\xi_k|\,\dot{\mathrm{I}}\,(|\xi_k| \geqslant k)\right] < \infty. \tag{14}$$

Тогда справедлив УЗБЧ (11).

$\Rrightarrow$ Рассмотрим сначала случай $\xi_k \geqslant 0$. Тогда условия a), b) гарантируют выполнение условий (13) предыдущей леммы последовательностью $\eta_k = \xi_k \,\dot{\mathrm{I}}\,(\xi_k < k)$. Таким образом,

$$\frac{1}{n}\sum_{k=1}^{n}\Big(\eta_k - \mathbf{E}\big[\xi_k \,\dot{\mathrm{I}}\,(\xi_k < k)\big]\Big) \underset{n}{\rightarrow} 0 \quad [\mathbf{P}\text{-п.н.}]\,. \tag{15}$$

Заметим, что если $\omega \in \Omega$ таково, что $\overline{\lim}_n \frac{1}{n}\sum_{k=1}^{n}\big(\xi_k(\omega) - \eta_k(\omega)\big) \neq 0$, то события $\{\xi_k \neq \eta_k\} = \{\xi_k \geqslant k\}$, $k \geqslant 1$, происходят бесконечно часто. Поэтому

$$\begin{aligned}\mathbf{P}\Big\{\overline{\lim_n}\frac{1}{n}\sum_{k=1}^{n}(\xi_k - \eta_k) \neq 0\Big\} &\leqslant \lim_n \mathbf{P}\Big\{\bigcup_{k=n}^{\infty}(\xi_k \geqslant k)\Big\} \leqslant \\ &\leqslant \lim_n \sum_{k=n}^{\infty}\mathbf{P}\{\xi_k \geqslant k\} = \lim_n \sum_{k=n}^{\infty}\mathbf{E}\big[\dot{\mathrm{I}}\,(\xi_k \geqslant k)\big] \leqslant \\ &\leqslant \lim_n \sum_{k=n}^{\infty}\frac{1}{k}\mathbf{E}\big[\xi_k \,\dot{\mathrm{I}}\,(\xi_k \geqslant k)\big] = 0\end{aligned}$$

в силу условия c). Другими словами, (15) будет справедливо, если в нём заменить сл.в. η_k на сл.в. ξ_k, $k \geqslant 1$.

Далее, из условия c) и леммы Кронекера 293, стр. 255, следует, что

$$0 = \lim_n \frac{1}{n}\sum\nolimits_1^n \mathbf{E}\big[\xi_k \,\dot{\mathrm{I}}\,(\xi_k \geqslant k)\big] = \lim_n \frac{1}{n}\sum\nolimits_1^n\Big(\mathbf{E}\,\xi_k - \mathbf{E}\big[\xi_k \,\dot{\mathrm{I}}\,(\xi_k < k)\big]\Big).$$

Поэтому в (15) можно заменить математические ожидания усечённых сл.в. $\mathbf{E}\,\eta_k$ на математические ожидания $\mathbf{E}\,\xi_k$.

В общем случае сл.в. ξ_k представим в виде разности неотрицательных сл.в. $\xi_k = \xi_k^+ - \xi_k^-$. Поскольку $\xi_k^2 = (\xi_k^+)^2 + (\xi_k^-)^2$ и $|\xi_k| = \xi_k^+ + \xi_k^-$, условия (14) будут выполняться и для последовательностей, составленных из положительных и отрицательных частей сл.в., т.е. если $\mu_k^{\pm} = \mathbf{E}\,\xi_k^{\pm}$, то

$$\frac{1}{n}\sum\nolimits_1^n(\xi_k - \mu_k) = \frac{1}{n}\sum\nolimits_1^n(\xi_k^+ - \mu_k^+) - \frac{1}{n}\sum\nolimits_1^n(\xi_k^- - \mu_k^-) \underset{n}{\rightarrow} 0 \quad (\text{п.н.}),$$

что и требовалось. $\Lleftarrow$

319| **Теорема.** [*А. Н. Колмогоров, Н. Этимади.*] *Пусть $\langle \xi_k \rangle_1^{\infty}$ — попарно независимые одинаково распределённые сл.в. с конечным математическим ожиданием $\mu = \mathbf{E}\,\xi_k$. Тогда справедлив УЗБЧ*

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k \underset{n}{\rightarrow} \mu \quad (\text{п.н.})\,.$$

⚠ Вариант УЗБЧ со сл.в., независимыми в совокупности, установлен А. Н. Колмогоровым.

⊨⇒ В условиях настоящей теоремы выполняются только первые два соотношения (14) УЗБЧ Этимади. Условие a) вполне очевидно. Ввиду одинаковой распределённости слагаемых в условии b) везде под знаком математического ожидания можно перейти к одной сл.в., например $\xi \backsim \xi_k$:

$$\sum_{k=1}^{\infty} \frac{1}{k^2} \mathbf{E} \left[\xi^2 \, \dot{\mathrm{I}} (|\xi| < k)\right] \; = \; \sum_{k=1}^{\infty} \frac{1}{k^2} \sum_{j=1}^{k} \mathbf{E} \left[\xi^2 \, \dot{\mathrm{I}} (j-1 \leqslant |\xi| < j)\right].$$

Произведём изменение порядка суммирования (что разрешено для рядов с положительными членами) и применим известные из анализа факты:

$$\textstyle\sum_1^{\infty} \dfrac{1}{k^2} \; = \; c \; = \; \pi^2/6 \quad \text{и} \quad j \textstyle\sum_{k=j}^{\infty} \dfrac{1}{k^2} \; \searrow \; 1, \; j \to \infty.$$

В результате исследуемая сумма преобразуется к виду

$$\sum_{j=1}^{\infty} \mathbf{E} \left[\xi^2 \, \dot{\mathrm{I}} (j-1 \leqslant |\xi| < j)\right] \sum_{k=j}^{\infty} \frac{1}{k^2} \; \leqslant \; c \, \mathbf{E} \left[\sum_{j=1}^{\infty} \frac{1}{j} \xi^2 \, \dot{\mathrm{I}} (j-1 \leqslant |\xi| < j) \right] \; \leqslant$$

$$\leqslant c \, \mathbf{E} \left[\sum_{j=1}^{\infty} |\xi| \, \dot{\mathrm{I}} (j-1 \leqslant |\xi| < j) \right] \; = \; c \, \mathbf{E} \, |\xi| < \infty,$$

где смена порядка суммирования и взятия математического ожидания также разрешена ввиду неотрицательности компонент.

Условие c), как показывает нижеследующий пример, может не выполняться, однако заметим, что при доказательстве теоремы Этимади это условие понадобилось только для установления следующих двух соотношений:

$$\infty \; > \; \sum_{k=1}^{\infty} \mathbf{P} \{ |\xi_k| \geqslant k \} \; = \; \sum_{k=1}^{\infty} \mathbf{P} \{ |\xi| \geqslant k \} \, ,$$

$$0 \; = \; \lim_n \frac{1}{n} \sum_{k=1}^{n} \mathbf{E} \left[|\xi_k| \, \dot{\mathrm{I}} (|\xi_k| \geqslant k) \right] \; = \; \lim_n \frac{1}{n} \sum_{k=1}^{n} \mathbf{E} \left[|\xi| \, \dot{\mathrm{I}} (|\xi| \geqslant k) \right],$$

первое из которых верно здесь в силу леммы 296 , (d), стр. 257. Второе следует из леммы 293 , (a), стр. 255, и свойства математического ожидания $\mathbf{E} \left[|\xi| \, \dot{\mathrm{I}} (|\xi| \geqslant k) \right] \xrightarrow{k} 0$, справедливого для интегрируемых сл.в. ⇚

320 Пример. Пусть распределение ξ_k задаётся плотностью $f(x) = (x \ln x)^{-2}/C, \; x \geqslant 2$, с константой $C = \int_2^{\infty} (x \ln x)^{-2} \, dx \approx 0.3787$. Математическое ожидание сл.в. ξ_k конечно: $\mathbf{E}\xi_k = \int_2^\infty xf(x)dx = \frac{1}{C\ln 2}$. Можно показать, что условие (14), b) здесь также справедливо.

С другой стороны, условие (14), c) не выполняется:

$$\sum\nolimits_{k=2}^{\infty} \frac{1}{k}\, \mathbf{E}\left[\xi_k\, \dot{\mathrm{I}}(\xi_k \geqslant k)\right] \;=\; \frac{1}{C} \sum\nolimits_{k=2}^{\infty} \frac{1}{k\ln k} \;=\; \infty. \qquad \circledcirc$$

Существование математического ожидания для УЗБЧ Колмогорова является необходимым условием.

321| **Следствие.** *Пусть $\langle\, \xi_k \,\rangle_1^\infty$ — независимые в совокупности одинаково распределённые сл.в. Если имеет место сходимость*

$$\frac{1}{n}\sum_{k=1}^{n} \xi_k \underset{n}{\to} \mu \quad (\text{п.н.})$$

к некоторой конечной константе μ, то $\mu = \mathbf{E}\xi_1$.

$\Rrightarrow$ По условию

$$\frac{\xi_n}{n} \;=\; \frac{1}{n}(S_n - S_{n-1}) \;=\; \frac{1}{n}S_n - \frac{n-1}{n}\frac{1}{n-1}S_{n-1} \underset{n}{\to} 0 \quad (\text{п.н.}),$$

т.е. почти наверное события $|\xi_n| \geqslant n$, $n \geqslant 1$, могут произойти лишь конечное число раз: $\mathbf{P}\{|\xi_n| \geqslant n \text{ б.ч.}\} = 0$. По лемме Бореля–Кантелли отсюда следует, что ряд $\sum_1^\infty \mathbf{P}\{|\xi_1| \geqslant k\} = \sum_1^\infty \mathbf{P}\{|\xi_k| \geqslant k\} < \infty$, что, в свою очередь, ввиду неравенств (d) леммы 296, стр. 257, обеспечивает конечность математического ожидания $\mathbf{E}|\xi_1|$. Следовательно, здесь применим УЗБЧ Колмогорова, т.е. $\frac{1}{n}S_n \to \mathbf{E}\xi_1$ (п.н.), что, очевидно, возможно т. т. т. когда $\mu = \mathbf{E}\xi_1$. $\Lleftarrow$

Для случая, когда математическое ожидание существует, но бесконечно, также справедлив УЗБЧ.

322| **Теорема.** *Пусть $\langle\, \xi_k \,\rangle_1^\infty$ — неотрицательные попарно независимые одинаково распределённые сл.в. с математическим ожиданием $\mathbf{E}\xi_k = +\infty$. Тогда $\frac{1}{n}\sum_1^n \xi_k \underset{n}{\to} +\infty$ (п.н.).*

$\Rrightarrow$ Рассмотрим усечение $\xi_k^{(N)} = \min\{\xi_k, N\}$ на уровне $N \geqslant 1$. Математическое ожидание $\mu^{(N)} = \mathbf{E}\xi_k^{(N)} < \infty$, и по теореме о монотонной сходимости $\mu^{(N)} \nearrow \infty$ при $N \nearrow \infty$. Для попарно независимых сл.в. $\xi_k^{(N)}$ справедлив УЗБЧ, т.е. с вероятностью единица справедливо

$$\underline{\lim_{n}}\, \frac{1}{n}\sum_{k=1}^{n} \xi_k \;\geqslant\; \lim_{n} \frac{1}{n}\sum_{k=1}^{n} \xi_k^{(N)} \;=\; \mu^{(N)}.$$

Откуда в силу монотонности последовательности $\mu^{(N)}, N = 1, 2, \ldots,$

$$\mathbf{P}\Big\{ \lim_n \frac{1}{n} \sum_{k=1}^n \xi_k = \infty \Big\} \;=\; \lim_N \mathbf{P}\Big\{ \underline{\lim_n}\, \frac{1}{n} \sum_{k=1}^n \xi_k \geqslant \mu^{(N)} \Big\} \;=\; 1. \qquad \Lleftarrow$$

Подобным же методом можно доказать следующее утверждение для случая, когда математическое ожидание слагаемых не существует.

323| **Теорема.** [*Kesten.*] *Пусть* $\langle \xi_k \rangle_1^\infty$ *— независимые одинаково распределённые сл.в. с* $\mathbf{E}\xi_k^+ = \mathbf{E}\xi_k^- = \infty$*. Тогда справедливо одно из трёх:*

a) $\lim_n \frac{1}{n} \sum_{k=1}^n \xi_k \;=\; +\infty$ (п.н.); b) $\lim_n \frac{1}{n} \sum_{k=1}^n \xi_k \;=\; -\infty$ (п.н.);

c) $\overline{\lim_n}\, \frac{1}{n} \sum_{k=1}^n \xi_k \;=\; +\infty, \quad \underline{\lim_n}\, \frac{1}{n} \sum_{k=1}^n \xi_k \;=\; -\infty$ (п.н.) .

324| Пример. Для распределения Коши, очевидно, выполняется вариант c). Таким образом, в этом случае в любой (почти) последовательности чисел $\frac{1}{n}\sum_1^n \xi_k$, $n = 1, 2, \ldots$, встретятся как сколь угодно большие положительные, так и сколь угодно большие отрицательные числа. ⊚

§ 3. Центральная предельная теорема

Центральная предельная теорема (ЦПТ) играет первостепенную роль в теории вероятностей. Термин «центральная» указывает на то место, которое занимает нормальное распределение как предельный закон для нормированных сумм случайных величин.

✧ **Одинаково распределённые сл.величины**. Начнём с простого варианта асимптотического утверждения о распределении суммы независимых одинаково распределённых сл.в. с конечной дисперсией.

325| **Теорема.** [*Центральная предельная теорема.*] *Пусть независимые в совокупности одинаково распределённые сл.в.* $\langle \xi_k \rangle_1^\infty$ *имеют конечные математическое ожидание* $\mu = \mathbf{E}\xi_k$ *и дисперсию* $\sigma^2 = \mathbf{D}\xi_k$, $k \in \mathbb{N}$*. Тогда*

$$\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n (\xi_k - \mu) \;\rightsquigarrow\; \mathcal{N}(0, 1), \tag{16}$$

т.е.

$$\lim_{n\to\infty} \mathbf{P}\Big\{ \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n (\xi_k - \mu) \leqslant x \Big\} \;=\; \Phi(x) \;:=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}y^2} dy. \tag{17}$$

⤃ Пусть $\varphi(t)$ — хар.ф. $\eta_k = (\xi_k - \mu)/\sigma$ (одинаковая для всех $k \geqslant 1$). Заметим, что $\alpha_1 = \mathbf{E}\eta_k = 0, \alpha_2 = \mathbf{D}\eta_k = 1$. Тогда по свойству ($\mathrm{X}_8$) для хар.ф.

$$\varphi(u) \;=\; 1 + \mathrm{i}\alpha_1 u - \frac{1}{2}\alpha_2 u^2 + \frac{1}{2}u^2\delta(u) \;=\; 1 - \frac{1}{2}u^2 + \frac{1}{2}u^2\delta(u),$$

с $\delta(u) \underset{u\to 0}{\longrightarrow} 0$. По свойствам ($\mathrm{X}_4$), ($\mathrm{X}_9$) хар.ф. сл.в. $\zeta_n = \frac{1}{\sqrt{n}}\sum_1^n \frac{1}{\sigma}(\xi_k - \mu)$

$$\varphi_{\zeta_n}(t) \;=\; \Big(\varphi\Big(\frac{t}{\sqrt{n}}\Big)\Big)^n \;=\; \Big(1 - \frac{1}{2}\frac{t^2}{n} + \frac{1}{2}\frac{t^2}{n}\delta\Big(\frac{t}{\sqrt{n}}\Big)\Big)^n.$$

В силу неравенства (2), стр. 253, и представления (1), стр. 253, для $\forall\, t \in \mathbb{R}^1$

$$\Big|\varphi_{\zeta_n}(t) - \Big(\exp\Big(-\frac{t^2}{2n}\Big)\Big)^n\Big| \;\leqslant\; n\Big|1 - \frac{t^2}{2n} + \frac{t^2}{2n}\delta\Big(\frac{t}{\sqrt{n}}\Big) - 1 + \frac{t^2}{2n} - \frac{t^4}{8n^2}\theta\Big| \;\leqslant$$

$$\leqslant\; \frac{t^2}{2}\Big|\delta\Big(\frac{t}{\sqrt{n}}\Big)\Big| + \frac{t^4}{8n} \;\underset{n}{\to}\; 0,$$

что по критерию слабой сходимости П. Леви доказывает теорему. ⟸

⚠ Поскольку ф.р. нормального закона непрерывна, сходимость в (17) по теореме 248, стр. 219, носит равномерный характер:

$$\lim_{n\to\infty}\sup_x \Big|\mathbf{P}\Big\{\frac{1}{\sigma\sqrt{n}}\sum\nolimits_1^n(\xi_k - \mu) \leqslant x\Big\} - \Phi(x)\Big| \;=\; 0.$$

⚠ Доказательство аналогичного факта для последовательности сл.векторов можно осуществить с помощью применения принципа Крамера–Волда (см. теорему 283, стр. 240).

В связи с доказанной теоремой возникает естественный вопрос о сходимости моментов нормированных сумм. Справедлива следующая теорема, доказательство которой можно найти, например, в [32].

326| **Теорема.** *Пусть независимые в совокупности одинаково распределённые сл.в.* $\langle \xi_k \rangle_1^\infty$ *с математическим ожиданием* $\mathbf{E}\xi_1 = 0$ *и дисперсией* $\mathbf{D}\xi_1 = 1$ *имеют* $\mathbf{E}\xi_1^{2m} < \infty$ *при некотором* $m \geqslant 1$. *Тогда для* $\forall\, 1 \leqslant j \leqslant m$

$$\lim_n \frac{1}{n^j}\,\mathbf{E}(S_n)^{2j} = \frac{(2j)!}{2^j j!}, \qquad \lim_n \frac{1}{n^{j-1/2}}\,\mathbf{E}(S_n)^{2j-1} = 0.$$

В монографии В. В. Петрова [17] (см. также [8]) приведены разнообразные необходимые и достаточные условия справедливости ЦПТ, доказанные при помощи более тонких методов анализа поведения хар.ф. Одно из утверждений (теорема 19, стр. 122), адаптированное на случай независимых одинаково распределённых сл.в., формулируется следующим образом.

327| **Теорема.** *Пусть* $\langle \xi_k \rangle_1^\infty$ *— независимые в совокупности одинаково распределённые сл.в. Для того чтобы имело место* (16) *с* $\mu = 0$ *и* $\sigma = 1$, *необходимо и достаточно выполнение при* $n \to \infty$ *следующих условий:*

a) $n\,\mathbf{P}\{|\xi_1| \geqslant \varepsilon\sqrt{n}\} \;\to\; 0, \quad \forall \varepsilon > 0;$

b) $\sqrt{n}\,\mathbf{E}\left[\xi_1\,\dot{\mathrm{I}}(|\xi_1| < \sqrt{n})\right] \;\to\; 0;$

c) $\mathbf{E}\left[\xi_1^2\,\dot{\mathrm{I}}(|\xi_1| < \sqrt{n})\right] \;\to\; 1.$

Из условия c) следует, что $\mathbf{E}\xi_1^2 = \lim_n \mathbf{E}\left[\xi_1^2\,\dot{\mathrm{I}}(|\xi_1| < \sqrt{n})\right] = 1$. По свойству математического ожидания это влечёт существование $\mathbf{E}|\xi_1| < \infty$. Условие b) в этом случае может выполняться только при $\mathbf{E}\xi_1 = 0$. В силу леммы 296, стр. 257, тот факт, что $\mathbf{E}\xi_1^2 < \infty$, делает условие a) излишним. Оно приведено здесь, чтобы сохранить структуру теоремы из указанной монографии, в которой утверждение сформулировано для разнораспределённых сл.в.

328| **Следствие.** *Для независимых одинаково распределённых сл.в. ЦПТ* (16) *имеет место т. т. т. когда* $\mu = \mathbf{E}\xi_1$, $\sigma^2 = \mathbf{D}\xi_1$.

329| Упр. Как отмечалось, условие $\mathbf{E}\xi_1^2 < \infty$ влечёт a). Докажите, что a) гарантирует $\mathbf{E}|\xi_1| < \infty$. Приведите пример, когда выполняется условие a), но $\mathbf{E}\xi_1^2 < \infty$.

✧ Разнораспределённые сл.величины. Теорема Линдеберга. Асимптотическая нормальность имеет место и для сумм различно распределённых сл.в. Однако в этом случае требуются чуть большее условие, чем просто наличие вторых моментов у слагаемых. Прежде всего заметим, что в (16) нормирующий множитель $\sqrt{n}\sigma$ равен стандартному отклонению суммы сл.в., т.е. $n\sigma^2 = \mathbf{D}(S_n)$. Этот множитель появляется в тех вариантах предельных теорем, в которых слагаемые независимы и имеют конечные вторые моменты. Приведём здесь ставшую уже классической теорему об асимптотической нормальности распределения сумм сл.в. с условием в форме Линдеберга. Говорят, что последовательность сл.в. $\langle \xi_k \rangle_1^\infty$ с нулевыми математическими ожиданиями и дисперсиями $\sigma_k^2 = \mathbf{D}\xi_k < \infty$, $k \geqslant 1$, удовлетворяет условию Линдеберга, если

$$\frac{1}{D_n}\sum_{k=1}^{n}\mathbf{E}\big[\xi_k^2\,\dot{\mathrm{I}}\,\big(|\xi_k|\geqslant\varepsilon\sqrt{D_n}\,\big)\big]\ \underset{n}{\to}\ 0,\qquad\forall\,\varepsilon>0,\tag{18}$$

$$D_n=\textstyle\sum_1^n\sigma_k^2.$$

Условие Линдеберга влечёт условие *равномерной малости* слагаемых:

$$\frac{1}{D_n}\max_{1\leqslant k\leqslant n}\sigma_k^2\ \underset{n}{\to}\ 0.\tag{19}$$

Действительно, если обозначить левую часть (18) через $L_n(\varepsilon)$, то получим

$$\sigma_k^2\ =\ \mathbf{E}\xi_k^2\ =\ \mathbf{E}\left[\xi_k^2\,\dot{\mathrm{I}}\,(|\xi_k|<\varepsilon\sqrt{D_n})\right]+\mathbf{E}\left[\xi_k^2\,\dot{\mathrm{I}}\,(|\xi_k|\geqslant\varepsilon\sqrt{D_n})\right]\ \leqslant$$
$$\leqslant\varepsilon^2D_n+L_n(\varepsilon)D_n.$$

Поэтому для $\forall\,\varepsilon>0$ предел $\overline{\lim}_n\max_{1\leqslant k\leqslant n}(\sigma_k^2/D_n)\leqslant\varepsilon^2$. Понятно, что для (19) и, следовательно, для условия Линдеберга необходимо, чтобы $D_n\to\infty$.

330| **Теорема.** [*Линдеберг.*] *Пусть* $\langle\,\xi_k\,\rangle_1^\infty$ *— последовательность независимых в совокупности сл.в. с нулевыми математическими ожиданиями и конечными дисперсиями* $\sigma_k^2=\mathbf{D}\xi_k<\infty$, $k\geqslant 1$. *Тогда*

I) *если выполняется условие Линдеберга* (18), *то*

$$\frac{1}{\sqrt{D_n}}\sum\nolimits_1^n\xi_k\rightsquigarrow\mathcal{N}(0,1);\tag{20}$$

II) [*Феллер*] *если имеет место* (20) *и, кроме того,* $D_n\to\infty$, $\sigma_n^2/D_n\to 0$, *что влечёт справедливость* (19), *то выполняется условие Линдеберга.*

⊨⇒ I) Пусть φ_k — хар.ф. сл.в. ξ_k, тогда хар.ф. $\zeta_n=S_n/\sqrt{D_n}$

$$\varphi_{\zeta_n}(t)\ =\ \prod_{k=1}^{n}\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big).$$

Так как $\exp(-\frac12t^2)=\prod_1^n\exp\big(-\frac{t^2}{2D_n}\sigma_k^2\big)$, то в силу неравенства (27), стр. 302,

$$\big|\varphi_{\zeta_n}(t)-e^{-t^2/2}\big|\ \leqslant\ \sum_{k=1}^{n}\Big|\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big)-\exp\Big(-\frac12\frac{t^2\sigma_k^2}{D_n}\Big)\Big|\ \leqslant$$
$$\leqslant\ \sum_{k=1}^{n}\Big|\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big)-1+\frac12\frac{t^2\sigma_k^2}{D_n}\Big|+\sum_{k=1}^{n}\Big|\exp\Big(-\frac12\frac{t^2\sigma_k^2}{D_n}\Big)-1+\frac12\frac{t^2\sigma_k^2}{D_n}\Big|.$$

По свойству показательной функции вторая сумма здесь не превосходит

$$t^4 \sum_{k=1}^{n} \frac{\sigma_k^4}{D_n^2} \leqslant t^4 \Big[\frac{1}{D_n} \max_{1\leqslant k\leqslant n} \sigma_k^2\Big] \sum_{k=1}^{n} \frac{\sigma_k^2}{D_n} = \Big[\frac{1}{D_n} \max_{1\leqslant k\leqslant n} \sigma_k^2\Big]\, t^4 \underset{n}{\to} 0$$

в силу (19). В первой сумме слагаемые суть математические ожидания (напомним, что $\mathbf{E}\,\xi_k = 0$) сл.в. $\Delta_k = \exp(\mathtt{i}t\xi_k/\sqrt{D_n}) - 1 - \mathtt{i}t\xi_k/\sqrt{D_n} + \frac{1}{2}\left(t\xi_k/\sqrt{D_n}\right)^2$. Для их оценки представим функцию $e^{\mathtt{i}x}$ с действительным x в виде двух отрезков ряда Маклорена (см. (10), стр. 178):

$$e^{\mathtt{i}x} = 1 + \mathtt{i}x - \frac{x^2}{2}\delta_2(x) = 1 + \mathtt{i}x - \frac{x^2}{2} + \frac{x^3}{6}\delta_3(x), \tag{21}$$

где $|\delta_2|, |\delta_3| \leqslant 1$. Поэтому $|\Delta_k| \leqslant \min\left(t^2\xi_k^2/D_n,\ t^3\xi_k^3/D_n^{3/2}\right)$ и, следовательно,

$$\begin{aligned}
\big|\,\mathbf{E}\,\Delta_k\big| &\leqslant \mathbf{E}\left[|\Delta_k|\,\dot{\mathtt{I}}\,(|\xi_k| \geqslant \varepsilon\sqrt{D_n}\,)\right] + \mathbf{E}\left[|\Delta_k|\,\dot{\mathtt{I}}\,(|\xi_k| < \varepsilon\sqrt{D_n}\,)\right] \leqslant \\
&\leqslant \frac{t^2}{D_n}\,\mathbf{E}\left[\xi_k^2\,\dot{\mathtt{I}}\,(|\xi_k| \geqslant \varepsilon\sqrt{D_n}\,)\right] + \frac{t^3}{D_n^{3/2}}\,\mathbf{E}\left[|\xi_k^3|\,\dot{\mathtt{I}}\,(|\xi_k| < \varepsilon\sqrt{D_n}\,)\right] \leqslant \\
&\leqslant \frac{t^2}{D_n}\,\mathbf{E}\left[\xi_k^2\,\dot{\mathtt{I}}\,(|\xi_k| \geqslant \varepsilon\sqrt{D_n}\,)\right] + \varepsilon\frac{t^3}{D_n}\,\mathbf{E}\left[\xi_k^2\right].
\end{aligned}$$

Таким образом,

$$\sum_{k=1}^{n} |\,\mathbf{E}\,\Delta_k| \leqslant \sum_{k=1}^{n} \frac{t^2}{D_n}\,\mathbf{E}\left[\xi_k^2\,\dot{\mathtt{I}}\,(|\xi_k| \geqslant \varepsilon\sqrt{D_n}\,)\right] + \sum_{k=1}^{n} \varepsilon\frac{|t|^3}{D_n}\,\mathbf{E}\left[\xi_k^2\right].$$

Вторая часть последнего выражения $\varepsilon|t|^3 \sum_{k=1}^{n} \mathbf{E}\left[\xi_k^2\right]/D_n = \varepsilon\,|t|^3$ может быть сделана сколь угодно малой за счёт выбора $\varepsilon > 0$. Первая часть при выбранном ε стремится к нулю по условию Линдеберга. Итак, хар.ф. $S_n/\sqrt{D_n}$ в пределе совпадает с хар.ф. стандартного нормального закона.

II) Условие (19) действительно выполняется, т.к. для $\forall\, m \geqslant 1$

$$\frac{1}{D_n} \max_{1\leqslant k\leqslant n} \sigma_k^2 \leqslant \frac{1}{D_n} \max_{1\leqslant k\leqslant m} \sigma_k^2 + \max_{m\leqslant k\leqslant n} \frac{\sigma_k^2}{D_k}.$$

Выбирая достаточно большое m, можно добиться того, что второе слагаемое здесь станет меньше $\forall\,\varepsilon > 0$, т.к. $\sigma_k^2/D_k \to 0$. Первое слагаемое может быть сделано меньше ε по условию $D_n \nearrow \infty$.

Сохраним обозначения для хар.ф. из доказательства первой части. Зафиксируем $\varepsilon > 0$ и выберем произвольное $t:\ t^2 > 8/\varepsilon^2$. Из первого представления в (21) с учётом $\mathbf{E}\xi_k = 0$ следует, что

$$\Big|\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big) - 1\Big| \;=\; \Big|\mathbf{E}\Big[\exp\Big(\mathrm{i}\frac{t}{\sqrt{D_n}}\,\xi_k\Big) - 1 - \mathrm{i}\frac{t}{\sqrt{D_n}}\,\xi_k\Big]\Big| \;\leqslant\; \frac{t^2}{D_n}\sigma_k^2.$$

Таким образом, ввиду (19) имеем

$$\max_{1\leqslant k\leqslant n}\Big|\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big) - 1\Big| \underset{n}{\to} 0, \qquad \sum\nolimits_{k=1}^{n}\Big|\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big) - 1\Big| \;\leqslant\; \frac{t^2}{2}.$$

Поэтому при всех достаточно больших n хар.ф. $|\varphi_k\big(t/\sqrt{D_n}\big) - 1| < 1/2$, $1 \leqslant k \leqslant n$. Следовательно, в силу (9), стр. 177,

$$\Big|\sum\nolimits_{k=1}^{n}\ln\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big) - \varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big) + 1\Big| \;\leqslant\; \sum\nolimits_{k=1}^{n}\Big|\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big) - 1\Big|^2 \;\leqslant$$

$$\leqslant\; \max_{1\leqslant k\leqslant n}\Big|\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big) - 1\Big|\;\sum\nolimits_{k=1}^{n}\Big|\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big) - 1\Big| \underset{n}{\to} 0, \quad \text{т.е.}$$

$$(*) \qquad \sum\nolimits_{k=1}^{n}\mathrm{Re}\Big(\ln\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big) - \varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big) + 1\Big) \underset{n}{\to} 0.$$

По предположению теоремы логарифм модуля хар.ф. $\ln|\varphi_{\zeta_n}(t)| \underset{n}{\to} -\frac{1}{2}t^2$. Так как по определению $\mathrm{Re}(\ln\varphi) = \ln|\varphi|$, то

$$\sum\nolimits_{1}^{n}\mathrm{Re}\Big(\ln\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big)\Big) \;=\; \sum\nolimits_{1}^{n}\ln\Big|\varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big)\Big| \;=\; \ln|\varphi_{\zeta_n}(t)| \underset{n}{\to} -\frac{t^2}{2} \quad \text{и из } (*)$$

$$\sum\nolimits_{k=1}^{n}\mathrm{Re}\Big(1 - \varphi_k\Big(\frac{t}{\sqrt{D_n}}\Big)\Big) \underset{n}{\to} \frac{t^2}{2}.$$

По определению хар.ф. последнее можно записать в виде (с некоторым $\gamma_n \underset{n}{\to} 0$)

$$\sum\nolimits_{k=1}^{n}\mathbf{E}\Big[1 - \cos\Big(\frac{t\xi_k}{\sqrt{D_n}}\Big)\Big] \;=\; \frac{t^2}{2} + \gamma_n.$$

Откуда

$$\sum\nolimits_{k=1}^{n}\mathbf{E}\Big[\,\dot{\mathrm{I}}\,(|\xi_k| \geqslant \varepsilon\sqrt{D_n})\,\frac{\varepsilon^2}{2}\Big(1 - \cos\Big(\frac{t\xi_k}{\sqrt{D_n}}\Big)\Big)\Big] \;=$$

$$=\; \frac{t^2\varepsilon^2}{4}\Big(1 - \sum\nolimits_{k=1}^{n}\mathbf{E}\Big[\,\dot{\mathrm{I}}\,(|\xi_k| < \varepsilon\sqrt{D_n})\,\frac{2}{t^2}\Big(1 - \cos\Big(\frac{t\xi_k}{\sqrt{D_n}}\Big)\Big)\Big]\Big) \;+\; \frac{\varepsilon^2}{2}\gamma_n(t).$$

Под знаком математического ожидания в левой части $\frac{\varepsilon^2}{2}(1 - \cos(\cdot)) \leqslant \varepsilon^2 \leqslant$

$\leqslant \xi_k^2/D_n$; в правой части $2(1-\cos(u)) \leqslant u^2$. Поэтому в правой части

$$1-\sum\nolimits_{k=1}^{n} \mathbf{E}\left[\,\dot{\mathrm{I}}\,(|\xi_k|<\varepsilon\sqrt{D_n})\frac{2}{t^2}\Big(1-\cos\Big(\frac{t\xi_k}{\sqrt{D_n}}\Big)\Big)\right] \;\geqslant$$

$$\geqslant\; 1-\frac{1}{D_n}\sum\nolimits_{k=1}^{n} \mathbf{E}\left[\xi_k^2\,\dot{\mathrm{I}}\,(|\xi_k|<\varepsilon\sqrt{D_n})\right] \;=\; \frac{1}{D_n}\sum\nolimits_{k=1}^{n} \mathbf{E}\left[\xi_k^2\,\dot{\mathrm{I}}\,(|\xi_k|\geqslant\varepsilon\sqrt{D_n})\right],$$

ибо $\sum_1^n \mathbf{E}\,\xi_k^2/D_n = 1$. Напомним, что $t^2 > 8/\varepsilon^2$, поэтому

$$\begin{aligned}\frac{1}{D_n}\sum_{k=1}^{n} \mathbf{E}\left[\xi_k^2\,\dot{\mathrm{I}}\,(|\xi_k|\geqslant\varepsilon\sqrt{D_n})\right] \;&\geqslant\; \frac{t^2\varepsilon^2}{4D_n}\sum_{k=1}^{n} \mathbf{E}\left[\xi_k^2\,\dot{\mathrm{I}}\,(|\xi_k|\geqslant\varepsilon\sqrt{D_n})\right]+\frac{\varepsilon^2}{2}\gamma_n\\ &\geqslant\; 2\frac{1}{D_n}\sum_{k=1}^{n} \mathbf{E}\left[\xi_k^2\,\dot{\mathrm{I}}\,(|\xi_k|\geqslant\varepsilon\sqrt{D_n})\right]+\frac{\varepsilon^2}{2}\gamma_n,\end{aligned}$$

т.е. $$0 \;\leqslant\; \frac{1}{D_n}\sum\nolimits_{k=1}^{n} \mathbf{E}\left[\xi_k^2\,\dot{\mathrm{I}}\,(|\xi_k|\geqslant\varepsilon\sqrt{D_n})\right] \;\leqslant\; -\frac{\varepsilon^2}{2}\gamma_n \underset{n}{\to} 0. \qquad \Lleftarrow$$

331| Примеры. 1) Для одинаково распределённых сл.в. условие Линдеберга выполняется, если $\sigma^2 = \mathbf{D}\xi_1 < \infty$, поскольку в этом случае $D_n = n\sigma^2$ и левая часть (18) равна $\mathbf{E}\left[\xi_1^2\,\dot{\mathrm{I}}\,\left(|\xi_1|\geqslant\varepsilon n\sigma^2\right)\right]\underset{n}{\to}0$.

2) Пусть $\xi_k \backsim \mathcal{N}(0,1/k^2)$, тогда $S_n/\sqrt{D_n} \backsim \mathcal{N}(0,1)$ при каждом $n \geqslant 1$. С другой стороны, в этом случае $D_n = \sum_1^n 1/k^2 \underset{n}{\to} \pi^2/6 \neq \infty$, т.е. ЦПТ справедлива вместе с условием $\sigma_n^2/D_n \to 0$, однако условие Линдеберга и соотношение (19) не выполняются.

3) Если накопленная дисперсия $D_n \to \sigma^2 < \infty$ и имеет место сходимость $S_n/\sqrt{D_n} \rightsquigarrow \mathcal{N}(0,1)$, то по теореме Слуцкого $S_n \rightsquigarrow \mathcal{N}(0,\sigma^2)$. Поэтому для хар.ф. φ_k сл.в. ξ_k справедливо равенство

$$\varphi_1(t)\prod\nolimits_2^{\infty}\varphi_k(t) \;=\; \exp\left\{-\tfrac{1}{2}t^2s^2\right\}.$$

По теореме Г. Крамера это возможно, только если φ_1 и $\prod_2^{\infty}\varphi_k(t)$ суть хар.ф. нормальных законов. Продолжая по индукции, получаем, что вариант ЦПТ с $D_n \to \sigma^2 < \infty$ справедлив только при нормальных слагаемых.

4) Пусть $\left\langle \zeta_k, \eta_k \right\rangle_1^{\infty}$ независимые в совокупности сл.в., где все $\zeta_k \backsim \zeta_1$, $k \geqslant 1$, т.е. одинаково распределены, и $\mathbf{E}\,\zeta_1 = 0$, $\mathbf{D}\zeta_1 = 1$, а сл.в. η_k, $k \geqslant 1$, таковы, что $\mathbf{P}\{\eta_k = \pm k\} = 1/2k^2$, $\mathbf{P}\{\eta_k = 0\} = 1-1/k^2$ ($\mathbf{E}\,\eta_k = 0$, $\mathbf{D}\eta_k = 1$, см. 307, стр. 264). Положим $\xi_k = \zeta_k + \eta_k$. Из 307 следует, что

$\sum_1^n \eta_k/\sqrt{n} \xrightarrow{\mathbf{P}} 0$, поэтому

$$\frac{1}{\sqrt{n}}\xi_k \;=\; \frac{1}{\sqrt{n}}\zeta_k + \frac{1}{\sqrt{n}}\eta_k \;\rightsquigarrow\; \mathcal{N}(0,1).$$

Условие Линдеберга здесь не выполняется, т.к. в противном случае асимптотическая нормальность имела бы место с множителем $1/\sqrt{2n}$. ◎

По-видимому, наиболее простое требование, гарантирующее выполнение условия Линдеберга, принадлежит Ляпунову. Отметим, что Ляпунов доказывал ЦПТ, не обращаясь к теореме Линдеберга.

332| **Теорема.** [*Ляпунов.*] *Пусть* $\langle \xi_k \rangle_1^\infty$ *— последовательность независимых в совокупности сл.в. с нулевыми математическими ожиданиями и конечными дисперсиями* $\sigma_k^2 = \mathbf{D}\xi_k < \infty,\ k \geqslant 1$. *Если для некоторого* $\delta > 0$

$$\frac{1}{D_n^{1+\delta}} \sum_{k=1}^n \mathbf{E}\,|\xi_k|^{2+2\delta} \underset{n}{\to} 0,$$

то выполнено условие Линдеберга (18), *и потому* $\dfrac{1}{\sqrt{D_n}} \sum_1^n \xi_k \rightsquigarrow \mathcal{N}(0,1)$.

⊫ Доказательство очевидно:

$$\frac{1}{D_n}\sum_1^n \mathbf{E}\left[\xi_k^2\, \dot{\mathrm{I}}\,(|\xi_k| \geqslant \varepsilon\sqrt{D_n})\right] \;\leqslant\; \frac{1}{D_n}\sum_1^n \mathbf{E}\left[\xi_k^2 \frac{|\xi_k|^{2\delta}}{\varepsilon^{2\delta}D_n^\delta}\, \dot{\mathrm{I}}\,(|\xi_k| \geqslant \varepsilon\sqrt{D_n})\right] \;\leqslant$$

$$\leqslant\; \frac{1}{\varepsilon^{2\delta}D_n^{1+\delta}} \sum_{k=1}^n \mathbf{E}\,|\xi_k|^{2+2\delta} \underset{n}{\to} 0.$$ ⫤

333| Упр. Покажите справедливость ЦПТ для последовательности независимых бинарных сл.в. с $\mathbf{P}\{\xi_k = \pm k\} = {}^1/_2$.

► Наиболее полно вопрос об асимптотическом распределении сумм сл.в. раскрывается в классических монографиях [8], [17]. Приведём здесь (без доказательства) одну теорему о предельном законе для сумм сл.в., образованных сериями. Снова в целях сокращения записи число элементов серии $K_n = n$. Заметим, что условия Линдеберга влекут за собой выполнение условий этой теоремы.

334| **Теорема.** *Пусть* $\langle \xi_{nk},\ 1 \leqslant k \leqslant n \rangle_{n=1}^\infty$ *— последовательность серий независимых в каждой серии сл.в.*

I) *Если для* $\forall\,\varepsilon>0$ *при* $n\to\infty$

$$(\natural)\qquad \begin{aligned} &(\checkmark)\quad \sum\nolimits_{k=1}^{n}\mathbf{P}\left\{|\xi_{nk}|\geqslant\varepsilon\right\}\ \to\ 0;\\ &(\checkmark)\quad \sum\nolimits_{k=1}^{n}\mathbf{E}\left[\xi_{nk}\,\dot{\mathrm{I}}\left(|\xi_{nk}|<\varepsilon\right)\right]\ \to\ \mu;\\ &(\checkmark)\quad \sum\nolimits_{k=1}^{n}\left\{\mathbf{E}\left[\xi_{nk}^2\,\dot{\mathrm{I}}\left(|\xi_{nk}|<\varepsilon\right)\right]-\Big(\mathbf{E}\left[\xi_{nk}\,\dot{\mathrm{I}}\left(|\xi_{nk}|<\varepsilon\right)\right]\Big)^2\right\}\ \to\ \sigma^2,\end{aligned}$$

то
$$S_n=\sum\nolimits_{k=1}^{n}\xi_{nk}\ \rightsquigarrow\ \mathcal{N}(\mu,\sigma^2)\,.$$

II) *Если* $S_n\rightsquigarrow\mathcal{N}(\mu,\sigma^2)$ *и выполняется условие бесконечной малости*

$$\max_{1\leqslant k\leqslant n}\mathbf{P}\left\{|\xi_{nk}|\geqslant\varepsilon\right\}\ \underset{n}{\to}\ 0,$$

то выполняются условия $(\natural)$.

335| Упр. Пусть $\left\langle\,\xi_k\,\right\rangle_1^\infty$ — независимые сл.в. с плотностью $f(x)=|x^{-3}|,\ |x|\geqslant1$. Покажите, что $(n\ln n)^{-1/2}\sum_1^n\xi_k\rightsquigarrow\mathcal{N}(0,1)$.

Наконец, дадим вариант ЦПТ без предположения существования моментов у слагаемых (см., например, [32], [17], теорема 19; ср. с 303, стр. 261).

336| **Теорема.** *Пусть* $\left\langle\,\xi_k\,\right\rangle_1^\infty$ — *независимые одинаково распределённые сл.в.* ($\xi_k\backsim\xi$). *Утверждение* $\frac{1}{D_n}S_n-A_n\rightsquigarrow\mathcal{N}(0,1)$ *имеет место т. т. т. когда*

$$\lim_{C\to\infty}\frac{C^2\,\mathbf{P}\left\{|\xi|\geqslant C\right\}}{\mathbf{E}\left[\xi^2\,\dot{\mathrm{I}}\left(|\xi|<C\right)\right]}=0;$$

при этом можно выбрать

$$D_n\ =\ \sup\{C:\ \mathbf{E}\left[\xi^2\,\dot{\mathrm{I}}\left(|\xi|<C\right)\right]\ \geqslant\ \frac{C^2}{n}\},\qquad A_n\ =\ \frac{n}{D_n}\mathbf{E}\left[\xi\,\dot{\mathrm{I}}\left(|\xi|<D_n\right)\right].$$

337| Примеры. 1) Пусть ф.р. ξ_k определяется соотношениями $F(-x)=1-F(x)=1/(2|x|^m),\ x\geqslant1$. Если $m>2$, то здесь существует конечная дисперсия. Поэтому (см. 315, стр. 269) $S_n/(\sqrt{n}\ln n)\ \to\ 0$ (п.н.). При $m=2$ в соответствии с приведённой теоремой следует выбрать $D_n=C$, где C — наибольшее решение уравнения $\ln C^2=C^2/n$. Несложно показать, что $D_n=\delta_n\,n\ln n$, где $\delta_n\to1$. Поэтому $S_n/\sqrt{n\ln n}\rightsquigarrow\mathcal{N}(0,1)$ (ср. с 335). Следовательно, снова $S_n/(\sqrt{n}\ln n)\ \xrightarrow{\mathbf{P}}\ 0$, хотя дисперсия и не существует.

2) В ситуации с различно распределёнными слагаемыми сходимость к нормальному закону возможна и при отсутствии математических ожиданий у

слагаемых. Пусть $\langle \vartheta_k, \eta_k,\ k \geqslant 1 \rangle$ — независимые в совокупности нормальные сл.в. $\vartheta_k \backsim \mathcal{N}(0,1)$ и сл.в. Коши $\eta_k \backsim \mathtt{Cauch}(0,\ {}^1/_k)$. Тогда, очевидно, сл.в. $\xi_k = \vartheta_k + \eta_k$ не имеет математического ожидания. С другой стороны, хар.ф. такой сл.в. (см. 160, стр. 161) равна $\varphi_k(t) = \exp\big(-\frac{1}{2}t^2 - \frac{1}{k}|t|\big)$. Следовательно, $\frac{1}{\sqrt{n}}\sum_1^n \xi_k \rightsquigarrow \mathcal{N}(0,1)$, т.к.

$$\prod_1^n \varphi_k\Big(\frac{t}{\sqrt{n}}\Big) \;=\; \exp\big(-\tfrac{1}{2}t^2 - \frac{1}{\sqrt{n}}\sum_1^n \frac{1}{k}|t|\big) \;\underset{n}{\to}\; \exp\big(-\tfrac{1}{2}t^2\big).$$

Отметим, что в рассмотренных здесь случаях приемлемая точность нормального приближения достигается при «запредельно» больших $n > 10^7$. ◎

✧ **Локальные предельные теоремы.** Так называются утверждения об асимптотическом поведении вероятностей попадания сумм сл.в. в ограниченные области, например вероятности принятия какого-либо значения $\mathbf{P}\{S_n = x\}$ или вероятности попадания в интервал $\mathbf{P}\{S_n \in (a;b)\}$. Самая первая в этом направлении теорема Муавра–Лапласа относится к сумме независимых бернуллиевских сл.в. $\langle \xi_j \backsim \mathcal{B}\mathtt{ern}(p) \rangle_1^\infty$. Напомним, что для бернуллиевской сл.в. математическое ожидание $\mathbf{E}\xi_1 = p$, а дисперсия $\mathbf{D}\xi_1 = p(1-p)$.

В доказательстве теоремы Муавра–Лапласа используется формула Стирлинга для факториалов больших чисел (доказательство см. стр. 304).

338| **Лемма.** [*Формула Стирлинга.*] *Для* $\forall n \geqslant 1$

$$n! \;=\; \sqrt{2\pi n}\, n^n\, e^{-n+\theta(n)}, \qquad \frac{1}{12n+1} < \theta(n) < \frac{1}{12n}.$$

Определим функцию $H(x) = x\ln\frac{x}{p} + (1-p)\ln\frac{1-x}{1-p}$, $x \in (0;1)$. Следующее утверждение получается простым применением формулы Стирлинга.

339| **Теорема.** *Пусть* $n, k, (n-k) \to \infty$ *и* $p^* = k/n$, *тогда*

$$\mathbf{P}\{S_n = k\} \;=\; \mathrm{C}_n^k p^k (1-p)^{n-k} \;=\; \frac{1}{\sqrt{2\pi n p^*(1-p^*)}}\exp\{-nH(p^*)\}(1+\varepsilon(k,n)),$$

где $1+\varepsilon(k,n) \;=\; e^{\gamma(k,n)},\ |\gamma(k,n)| < \dfrac{1}{12np^*(1-p^*)}$.

Дальнейшее преобразование этой формулы связано с представлением функции H при различных значениях p^*. Во-первых, разложим H в ряд

Тейлора до второй степени в окрестности точки $x = p$:

$$H(x) = \frac{1}{2p(1-p)}(x-p)^2 + \mathcal{O}(|x-p|^3), \quad x \to p.$$

340| **Теорема.** [*Муавра–Лапласа локальная.*] *Если* $p^* = k_n/n \to p$ *и* $n(p^*-p)^3 \to 0$, *то для* $\sigma^2 = p(1-p)$

$$\mathbf{P}\{S_n = k_n\} \ = \ \frac{1}{\sqrt{2\pi}\sigma}\exp\Big\{-\frac{n}{2\sigma^2}(p^*-p)^2\Big\}\frac{1}{\sqrt{n}}\big(1+\varepsilon(k_n,n)\big), \qquad (22)$$

где $\varepsilon(k_n, n) \to 0$.

Из локальной теоремы Муавра–Лапласа несложно получается так называемая интегральная теорема Муавра–Лапласа, представляющая собой вариант центральной предельной теоремы для сумм бернуллиевских сл.в. Доказательства этих теорем имеются в большинстве учебников по теории вероятностей (см., например, [3], [27]). Более подробным анализом функции H (см. [3]) можно показать, что при $|p^*-p| \leqslant \frac{1}{2}\min(p, 1-p)$ в правой части (22)

$$\big(1+\varepsilon(k_n,n)\big) \ = \ \exp\Big\{\gamma\Big(\frac{|k_n-np|^3}{3n^2\sigma^4}\Big) + \big(|k_n-np| + \tfrac{1}{6}\big)\frac{1}{n\sigma^2}\Big)\Big\}, \quad |\gamma| < 1.$$

Из локальной предельной теоремы Муавра–Лапласа можно получить центральную предельную теорему для бернуллиевских сл.в. с оценкой погрешности. Пусть Φ — ф.р. стандартного нормального закона. Определим решётку Q_n чисел, представимых в виде $(k-np)/\sqrt{n}\sigma$, $k = 0, 1, \ldots$, $\sigma^2 = p(1-p)$.

341| **Теорема.** *Пусть* $b > a$, $a, b \in Q_n$, *и* $c = \max(|a|, |b|)$. *Тогда*

$$\mathbf{P}\Big\{a < \frac{S_n - np}{\sqrt{n}\sigma} \leqslant b\Big\} \ = \ \big(\Phi(b) - \Phi(a)\big)\Big(1 + \theta_1\frac{1}{\sigma}\frac{c}{\sqrt{n}}\Big)(1 + 2\theta_2\rho)$$

в предположении, что $\rho = \dfrac{1}{3\sigma}\dfrac{c^3+3c}{\sqrt{n}} + \dfrac{1}{6n\sigma^2} \leqslant \dfrac{1}{2}$, где $|\theta_j| < 1, j = 1, 2$.

Как видно их теоремы, вероятность попадания нормированной суммы S_n в интервал $(a; b]$ можно приближённо вычислять через нормальный закон не только при фиксированных a, b, но и в случае, когда $|a|, |b| \to \infty$, т.е. в ситуации больших уклонений (см. далее), когда и вероятность, и её оценка близки к нулю и важна не их абсолютная, а их относительная близость.

342| **Следствие.** *Если* $a = a_n \to +\infty, b = b_n \to +\infty$ *так, что* $b = o(n^{1/6})$, *то*

$$\mathbf{P}\Big\{\, a < \frac{S_n - np}{\sqrt{n}\sigma} \leqslant b \,\Big\} \;=\; \big(\Phi(b) - \Phi(a)\big)\big(1 + \varepsilon_n\big),$$

где $\varepsilon_n \to 0$.

Один вариант возможных утверждений об асимптотическом поведении плотности частичной суммы (в ситуации, когда такая плотность существует) приведён в следующей теореме; другой вариант см. 346, стр. 293.

343| **Теорема.** *Пусть* $\langle\, \xi_k \,\rangle_1^\infty$ *— абсолютно непрерывные, независимые и одинаково распределённые сл.в.,* $\mathbf{E}\,\xi_1 = 0$, $\mathbf{D}\xi_1 = 1$. *Если хар.ф.* φ *сл.в.* ξ_1 *абсолютно интегрируема, то плотность* f_n *распределения величины* $\eta_n = S_n/\sqrt{n}$ *сходится равномерно к плотности* ϕ *нормального* $\mathcal{N}(0,1)$ *закона:*

$$\sup_x |f_n(x) - \phi(x)| \underset{n}{\to} 0. \qquad (\natural)$$

$\Rrightarrow$ Ясно, что и хар.ф. $\varphi_n(x)\big(= \varphi_{\eta_n}(x)\big) = \varphi^n(x/\sqrt{n})$ абсолютно интегрируема. Поэтому можно для представления плотностей f_n и ϕ применить формулу обращения (4), стр. 166:

$$|f_n(x) - \phi(x)| \;\leqslant\; \frac{1}{2\pi}\int_{-\infty}^{\infty} \big|e^{-\mathrm{i}tx}\big|\Big|\varphi^n\Big(\frac{t}{\sqrt{n}}\Big) - e^{-\frac{1}{2}t^2}\Big|\,dt. \qquad (23)$$

Поскольку по центральной предельной теореме $\big|\varphi^n\big(t/\sqrt{n}\big) - e^{-\frac{1}{2}t^2}\big| \underset{n}{\to} 0$ для $\forall\, t \in \mathbb{R}^1$, то в силу теоремы Лебега для любого фиксированного $\delta > 0$

$$\int_{|t|\leqslant\delta} \Big|\varphi^n\Big(\frac{t}{\sqrt{n}}\Big) - e^{-\frac{1}{2}t^2}\Big|\,dt \;\underset{n}{\to}\; 0.$$

Воспользуемся теперь разложением в ряд Тейлора для хар.ф.:

$$|\varphi(t)| \;=\; \Big|1 - \frac{1}{2}t^2 + t^2\varepsilon(t)\Big| \;\leqslant\; 1 - \frac{1}{2}t^2 + t^2\big|\varepsilon(t)\big|, \quad t^2 \leqslant 2,$$

где $\lim_{t\to 0}\varepsilon(t) \to 0$ (свойство (X_8) для хар.ф.). Стало быть, можно подобрать $A < \sqrt{2}$ такое, что $\big|\varepsilon(t)\big| < 1/4$ при $|t| \leqslant A$ и, следовательно,

$$|\varphi(t)| \;\leqslant\; \exp\{-\frac{1}{4}t^2\}$$

ввиду известного неравенства $1 + x \leqslant e^x$. Таким образом, при достаточно большом (фиксированном) δ и всех $n > \delta^2/A^2$

$$\int\limits_{\delta<|t|\leqslant A\sqrt{n}} \Big|\varphi^n\Big(\frac{t}{\sqrt{n}}\Big)\Big|\,dt \;\leqslant\; 2\sqrt{4\pi}\big(\Phi(A\sqrt{n}) - \Phi(\delta)\big) \;\leqslant\; 2\sqrt{4\pi}\big(1 - \Phi(\delta)\big) \;\leqslant\; \varepsilon.$$

Аналогично для ϕ.

Поскольку ξ_1 — абсолютно непрерывная сл.в., то для хар.ф. $|\varphi(t)| < 1$ для $\forall t \neq 0$ (см. стр. 174). Кроме того, по лемме Римана–Лебега (стр. 179) $\lim_{t\to\infty} |\varphi(t)| = 0$. Поэтому $\sup_{|t|>A} |\varphi(t)| = \gamma < 1$. Следовательно,

$$\int\limits_{|t|>A\sqrt{n}} \left|\varphi^n\Big(\frac{t}{\sqrt{n}}\Big) - e^{-\frac{1}{2}t^2}\right| dt \leqslant \gamma^{n-1} \int_{-\infty}^{\infty} \left|\varphi\Big(\frac{t}{\sqrt{n}}\Big)\right| dt + 2\sqrt{2\pi}\big(1 - \Phi(A\sqrt{n})\big)$$

$$= \gamma^{n-1}\sqrt{n} \int_{-\infty}^{\infty} |\varphi(t)|\, dt + 2\sqrt{2\pi}\big(1 - \Phi(A\sqrt{n})\big) \underset{n}{\to} 0.$$

Итак, все части интеграла (23) при $|t| \leqslant \delta$, $\delta < |t| \leqslant A\sqrt{n}$ и $|t| > A\sqrt{n}$ можно сделать сколь угодно малыми за счёт выбора δ, A и $n \to \infty$. ⫤

⚠ Утверждение (♮) справедливо и для решётчатых распределений, т.е. для одинаково распределённых сл.в., для которых $\sum_{k=-\infty}^{\infty} \mathbf{P}\{\xi = b + kh\} = 1$ при некоторых $b \in \mathbb{R}^1, h > 0$ (с максимально возможным шагом h). Для таких сл.в. имеет место формула обращения, аналогичная 165, стр. 165. Поэтому здесь можно применить метод доказательства только что доказанной теоремы; роль функции плотности выполняет нормированная вероятность:

$$f_n(x) = \frac{\sqrt{n}\sigma}{h} \mathbf{P}\Big\{ \frac{1}{\sigma\sqrt{n}} S_n = x \Big\},$$

и супремум берётся по всем x вида $\dfrac{1}{\sqrt{n}}(nb + kh)$, $k = 0, \pm 1, \ldots$

§ 4. Функциональная предельная теорема

Рассмотрим ситуацию, приводящую к классической центральной предельной теореме, с последовательностью $\langle \xi_k \rangle_1^\infty$ независимых одинаково распределённых сл.в. с математическим ожиданием $\mu = 0$ и дисперсией $\sigma^2 = 1$; положим $S_n = \sum_1^n \xi_k$, $S_0 := 0$. При каждом фиксированном $n \geqslant 1$ график функции $\widetilde{S}_n(t) := S_{\lfloor nt \rfloor}$, $t \in [0;1]$, представляет собой ступенчатую линию со случайными скачками в точках $t = \frac{k}{n}$, $k = \overline{1,n}$. Другими словами, $\widetilde{S}_n$ есть случайный процесс, траектории которого принадлежат пространству $\mathcal{D}([0;1])$ функций на $[0;1]$, непрерывных справа и имеющих предел слева.

Для фиксированного $t \in [0;1]$ обозначим $t_n = \llbracket nt \rrbracket$, тогда при $n \to \infty$

монотонно $t_n \nearrow \infty$, $t_n/n \to t$, и по центральной предельной теореме

$$\frac{1}{\sqrt{n}}\widetilde{S}_n(t) \ = \ \frac{\sqrt{t_n}}{\sqrt{n}}\frac{1}{\sqrt{t_n}}S_{t_n} \ \rightsquigarrow \ \mathcal{N}(0,t).$$

Пусть теперь $u < t$ и c_1, c_2 — произвольные константы, тогда

$$\frac{c_1}{\sqrt{n}}\widetilde{S}_n(u) + \frac{c_2}{\sqrt{n}}\widetilde{S}_n(t) \ = \ \frac{\sqrt{u_n}}{\sqrt{n}}\frac{c_1+c_2}{\sqrt{u_n}}S_{u_n} + \frac{\sqrt{t_n-u_n}}{\sqrt{n}}\frac{c_2}{\sqrt{t_n-u_n}}(S_{t_n}-S_{u_n}).$$

Применяя утверждение 251, стр. 221, и принцип Крамера–Волда 282, стр. 240, к вектору независимых сл.в. $\bigl(S_{u_n}, S_{t_n}-S_{u_n}\bigr)$, аналогично предыдущему получаем:

$$c_1\frac{1}{\sqrt{n}}\widetilde{S}_n(u) + c_2\frac{1}{\sqrt{n}}\widetilde{S}_n(t) \ \rightsquigarrow \ \mathcal{N}\bigl(0, (c_1+c_2)^2u + c_2^2(t-u)\bigr).$$

Легко видеть, что дисперсия $(c_1+c_2)^2u + c_2^2(t-u) = (c_1,c_2)\boldsymbol{\Sigma}(c_1,c_2)^\flat$ с матрицей $\boldsymbol{\Sigma} = \left(\begin{smallmatrix} u & u \\ u & t\end{smallmatrix}\right)$. Таким образом, снова в силу принципа Крамера–Волда это доказывает, что двумерный вектор $\bigl(\widetilde{S}_n(u), \widetilde{S}_n(t)\bigr)/\sqrt{n} \ \rightsquigarrow \ \mathcal{N}_2(\vec{0}, \boldsymbol{\Sigma})$.

Заметим, что $\boldsymbol{\Sigma}$ совпадает с матрицей ковариаций двумерного распределения стандартного винеровского процесса $\boldsymbol{w}(t)$, $t \in [0;1]$ (стр. 350). Аналогичными рассуждениями показывается, что любые конечномерные распределения процесса $\widetilde{\boldsymbol{Z}}_n = \widetilde{S}_n/\sqrt{n}$ сходятся к соответствующим конечномерным распределениям процесса Винера. В монографии [2] показано, что последовательность сл.процессов $\bigl\langle \widetilde{\boldsymbol{Z}}_n \bigr\rangle_1^\infty$ равномерно плотна в $\mathcal{D}([0;1])$. В соответствии с теоремой Прохорова 275, стр. 234, из любой подпоследовательности $\widetilde{\boldsymbol{Z}}_{n_k}$ (ввиду равномерной плотности $\widetilde{\boldsymbol{Z}}_n$) можно выделить подпоследовательность $\widetilde{\boldsymbol{Z}}_{k_m} \rightsquigarrow \boldsymbol{Z}$, где $\boldsymbol{Z}$ — некоторый процесс в $\mathcal{D}([0;1])$. Поскольку конечномерные распределения $\boldsymbol{Z}$ должны совпадать с конечномерными распределениями винеровского процесса $\boldsymbol{w}$, то по теореме Каратеодори $\boldsymbol{Z} \backsim \boldsymbol{w}$. В силу леммы 239, стр. 211, окончательно получаем, что $\widetilde{\boldsymbol{Z}}_n = \widetilde{S}_n/\sqrt{n} \rightsquigarrow \boldsymbol{w}$.

Удобнее вместо ступенчатой случайной функции $\widetilde{S}_n$ рассматривать непрерывную функцию, полученную из $\widetilde{S}_n$ соединением ступенек прямой линией:

$$S_n^*(t) = S_{\lfloor nt \rfloor} + \bigl(nt - [\![nt]\!]\bigr)\xi_{\lfloor nt \rfloor + 1}.$$

Аналогично предыдущему показывается, что все конечномерные распределения $\boldsymbol{Z}_n^* := S_n^*/\sqrt{n}$ сходятся к соответствующим конечномерным распределениям процесса Винера, и чуть попроще, что последовательность $\bigl\langle \boldsymbol{Z}_n^* \bigr\rangle_1^\infty$ равномерно плотна в пространстве непрерывных функций $\mathcal{C}([0;1])$.

344| **Теорема.** [*Функциональная предельная теорема. М. Донскер.*] *Пусть* $\langle \xi_k \rangle_1^\infty$ — *независимые одинаково распределённые сл.в. с математическим ожиданием* $\mu = 0$ *и дисперсией* $0 < \sigma^2 < \infty$, $\boldsymbol{w} = \langle w(t),\ t \in [0;1] \rangle$ — *винеровский процесс. Тогда для последовательности случайных функций* $\widetilde{S}_n(t) = S_{\lfloor nt \rfloor}$, $t \in [0;1]$, $n \geqslant 1$ *справедливо*

$$\frac{1}{\sigma\sqrt{n}}\widetilde{S}_n \rightsquigarrow \boldsymbol{w} \ \text{ в } \ \mathcal{D}([0;1]),$$

т.е. $\mathbf{E}\, h\bigl(\widetilde{S}_n/(\sigma\sqrt{n})\bigr) \underset{n}{\to} \mathbf{E}\, h(\boldsymbol{w})$ *для любого непрерывного ограниченного отображения* $h : \mathcal{D}([0;1]) \mapsto \mathbb{R}^1$, *в частности,* $h\bigl(\widetilde{S}_n/(\sigma\sqrt{n})\bigr) \rightsquigarrow h(\boldsymbol{w})$ *для любого непрерывного отображения* h. *Аналогично, последовательность случайных функций* $\dfrac{1}{\sigma\sqrt{n}} S_n^*(t) = \dfrac{1}{\sigma\sqrt{n}}\bigl(S_{\lfloor nt \rfloor} + (nt - \llbracket nt \rrbracket)\xi_{\lfloor nt \rfloor + 1}\bigr) \rightsquigarrow \boldsymbol{w}$ *в* $\mathcal{C}([0;1])$.

⚠ Функциональная предельная теорема была обобщена Прохоровым на последовательность разнораспределённых сл.в. с условием Линдеберга. Поэтому эту теорему часто называют теоремой Донскера–Прохорова.

По теореме Донскера асимптотическое распределение любой непрерывной функции, построенной на траекториях частичных сумм последовательности независимых одинаково распределённых сл.в., совпадает с распределением этой функции, вычисленной от траекторий винеровского процесса. Последнее распределение может быть найдено либо непосредственно на основе свойств процесса Винера, либо снова асимптотически с помощью специально построенной последовательности частичных сумм, например на основе частичных сумм бернуллиевских сл.в. Такой приём называется принципом инвариантности.

345| Примеры. 1) Легко понять, что функция $h(x) = \sup_t x(t) : \mathcal{C}([0;1]) \mapsto \mathbb{R}^1$ непрерывна. Чтобы найти распределение $\sup_t w(t)$, рассмотрим последовательность независимых сл.в. $\xi_1, \xi_2, \ldots$, для которых $\mathbf{P}\{\xi_j = -1\} = \mathbf{P}\{\xi_j = +1\} = {}^1\!/_2$, $j = 1, 2, \ldots$ Очевидно, что $\{\sup_t S_n^* > b\} = \{\max_{1 \leqslant k \leqslant n} S_k > b\}$. Обозначим $M_j = \max_{1 \leqslant k \leqslant j} S_k$, $j = \overline{1,n}$, и покажем, что для любого целого $m > 0$ справедливо равенство

$$\mathbf{P}\{M_n \geqslant m, S_n < m\} \ = \ \mathbf{P}\{M_n \geqslant m, S_n > m\}. \tag{24}$$

В рассматриваемой ситуации вероятность любого события равна количеству n-векторов вида $(\pm 1, \ldots, \pm 1)$, для которых удовлетворяется соответствующее условие, делённому на 2^n. Если $k = \min\{j : \ S_j = m\}$, т.е. событие

$M_n \geqslant m$ впервые произошло на k-м шаге ($k < m$), то любому вектору $(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n)$, входящему в событие левой части (24), будет соответствовать вектор $(x_1, \ldots, x_k, -x_{k+1}, \ldots, -x_n)$, входящий в событие правой части, что и доказывает их равновероятность.

Поскольку событие $\{S_n = m\} \subset \{M_n \geqslant m\}$ и $\{S_n > m\} \subset \{M_n \geqslant m\}$, то

$$\mathbf{P}\Big\{\max_k S_k \geqslant m\Big\} - \mathbf{P}\{S_n = m\} = \mathbf{P}\{M_n \geqslant m\} - \mathbf{P}\{S_n = m\} =$$
$$= \mathbf{P}\{M_n \geqslant m, S_n > m\} + \mathbf{P}\{M_n \geqslant m, S_n < m\} =$$
$$= 2\,\mathbf{P}\{M_n \geqslant m, S_n > m\} = 2\,\mathbf{P}\{S_n > m\}.$$

Следовательно, для $\forall\, x > 0$

$$\mathbf{P}\Big\{\max_{1\leqslant k\leqslant n} \frac{1}{\sqrt{n}} S_k \geqslant x\Big\} = 2\,\mathbf{P}\big\{S_n > [\![x\sqrt{n}\,]\!]\big\} + \mathbf{P}\big\{S_n = [\![x\sqrt{n}\,]\!]\big\},$$

где $[\![a\,]\!]$ — наименьшее целое, не меньше a (округление вверх). В рассматриваемом симметричном случае $\mathbf{P}\{S_n = [\![x\sqrt{n}\,]\!]\} \leqslant \mathbf{P}\{S_n = 0\} = \frac{1}{2^n} \mathrm{C}_n^{n/2} \underset{n}{\to} 0$. В свою очередь, по центральной предельной теореме

$$\lim_n \mathbf{P}\Big\{\frac{1}{\sqrt{n}} S_n > \frac{1}{\sqrt{n}} [\![x\sqrt{n}\,]\!]\Big\} = 1 - \Phi(x),$$

где Φ — ф.р. стандартного нормального закона. Таким образом, по теореме Донскера ф.р. максимума винеровского процесса задаётся формулой

$$\mathbf{P}\Big\{\sup_{t\in[0;1]} \boldsymbol{w}(t) \leqslant x\Big\} = 2\Phi(x) - 1 = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{1}{2}u^2}\,du, \quad x \geqslant 0.$$

Отсюда, в силу принципа инвариантности, предельное распределение максимума (нормированных) частных сумм для любой последовательности независимых одинаково распределённых сл.в. с нулевым математическим ожиданием и единичной дисперсией определяется ф.р. $2\Phi(x) - 1$, $x \geqslant 0$.

2) Аналогичным образом можно найти распределение интеграла Римана от траекторий винеровского процесса $\int_0^1 \boldsymbol{w}(t)\,dt$. Во-первых, заметим, что интеграл представляет собой непрерывное отображение $\mathcal{C}([0;1])$ в $(\mathbb{R}^1, \mathscr{B}^1)$. Поэтому это отображение измеримо, следовательно, $\int_0^1 \boldsymbol{w}(t)\,dt$ есть случайная величина. Рассмотрим последовательность независимых стандартных нормальных сл.в. $\xi_k \backsim \mathcal{N}(0,1)$, $k \geqslant 1$. Напомним, что график функции $S_n^*(t)$ на участке $\frac{1}{n}(k-1) \leqslant t \leqslant \frac{1}{n}k$ представляет собой отрезок прямой, соединяющей S_{k-1} и $S_k = S_{k-1} + \xi_k$. Поэтому интеграл от $S_n^*(t)$ на этом участке равен

$\frac{1}{n}\big(S_k - \frac{1}{2}\,\xi_k\big)$. Следовательно,

$$\frac{1}{\sqrt{n}}\int_0^1 S_n^*(t)\,dt = \frac{1}{n\sqrt{n}}\sum\nolimits_1^n \big(S_k - \tfrac{1}{2}\,\xi_k\big) = \frac{1}{n\sqrt{n}}\sum\nolimits_1^n \big(n-k+\tfrac{1}{2}\big)\xi_k.$$

Последняя сумма имеет нормальное распределение со средним 0 и дисперсией

$$\frac{1}{4n^3}\sum\nolimits_1^n (2(n-k)+1)^2 = \frac{1}{4n^3}\,\frac{n(4n^2-1)}{3} \underset{n}{\to} \frac{1}{3}$$

в силу известной формулы для суммы квадратов целых нечётных чисел. Стало быть, $\int_0^1 \boldsymbol{w}(t)\,dt \sim \mathcal{N}(0, {}^1\!/_3)$ (ср. с 25, стр. 364). ◎

§ 5. О точности аппроксимации в ЦПТ

✧ **Неравенство Берри–Эссеена. Большие уклонения.** Обсудим вопрос о точности аппроксимации в ЦПТ. Результаты будем формулировать только для случая независимых одинаково распределённых сл.в., хотя для большинства из приводимых фактов соответствующие утверждения справедливы и для случая разнораспределённых сл.в., и для схемы серий.

Обозначим через $F_n(x) = \mathbf{P}\big\{\sum_1^n \xi_k \leqslant x\sigma\sqrt{n}\big\}$ ф.р. нормированной суммы независимых сл.в. с нулевым математическим ожиданием и дисперсией σ^2. Здесь можно выделить два направления оценки отклонения $|F_n(x) - \Phi(x)|$ ф.р. F_n от ф.р. Φ стандартного нормального закона. Первое восходит к Ляпунову, который при доказательстве ЦПТ (в предположении существования третьих моментов) нашёл, что $\sup_x |F_n(x) - \Phi(x)|$ при каждом n не превосходит $C\ln n\,/\sqrt{n}$. Затем усилиями Берри, Эссеена и Крамера были получены результаты, подобные следующей теореме ([21], [17]).

[*Неравенство Берри–Эссеена.*] *Если* $\langle\,\xi_k\,\rangle_1^\infty$ *— независимые одинаково распределённые сл.в. с* $\mathbf{E}\,\xi_1 = 0,\ 0 < \sigma^2 = \mathbf{D}\xi_1 < \infty$, *то*

$$\sup_x \big|F_n(x) - \Phi(x)\big|\,\frac{\sigma^3}{\gamma_3}\sqrt{n} \leqslant C^*, \tag{25}$$

где $\gamma_3 = \mathbf{E}\,|\xi_1^3|$.

Доказательство неравенства приведено в дополнении A к этой главе. Универсальная, т.е. пригодная для любых моделей распределения ξ_k, константа $C^* \geqslant \frac{1}{\sqrt{2\pi}} \approx 0.399$. Эссеен показал, что $C^* < 7.6$. Дальнейшими исследованиями эта граница всё более уточнялась. К настоящему моменту известно, что $C^* < 0.71$. Кроме того, установлена асимптотическая верхняя граница:

$$\overline{\lim_n}\ \sup_F\ \sup_x \left|F_n(x) - \Phi(x)\right| \frac{\sigma_F^3}{\gamma_{3F}}\sqrt{n} \;=\; \frac{\sqrt{10}+3}{6\sqrt{2\pi}} \approx 0.410,$$

где σ_F^2, γ_{3F} — моменты, соответствующие ф.р. F сл.в. ξ_1.

Некоторое уточнение неравенства Берри–Эссеена даёт неравенство

$$\left|F_n(x) - \Phi(x)\right| \frac{\sigma^3}{\gamma_3}\sqrt{n} \;\leqslant\; C^* \frac{1}{1+|x^3|}, \qquad x \in \mathbb{R}^1.$$

► Отметим, что нормированная невязка в левой части неравенства Берри–Эссеена для конкретных распределений ξ_k и конкретных n может быть значительно меньше универсальной границы $1/\sqrt{2\pi}$. Так, при равномерном распределении ξ_k и $n \geqslant 10$ отклонение не превышает 0.007. Этот недостаток в какой-то степени исправляет подход к оценке точности ЦПТ, предложенный Чебышёвым и развитый в трудах Шарлье, Эджворта, Крамера. Связан этот подход с представлением расхождения между ф.р. в виде ряда по степеням $\frac{1}{\sqrt{n}}$ и полиномам относительно x, зависящим от моментов сл.в. Следует сказать, что здесь результаты разнятся в зависимости от того, принадлежит распределение ξ_k к решётчатому типу (т.е. можно значения ξ_k представить в виде $a, a \pm h, a \pm 2h, \ldots$) или нет. Приведём только один результат ([8], [21, т. 2]).

[*Разложение Эджворта.*] *Пусть* $\langle\, \xi_k \,\rangle_1^\infty$ *— независимые одинаково распределённые (не решётчатые) сл.в. с нулевым математическим ожиданием и единичной дисперсией, которые имеют конечные моменты* $\alpha_m = \mathbf{E}\,\xi_1^m$, $m \leqslant 3\ (4)$. *Тогда для ф.р.* $F_n(x) = \mathbf{P}\left\{\sum_1^n \xi_k \leqslant x\sqrt{n}\right\}$ *справедливо*

$$\sup_x \left| F_n(x) - \Phi(x) - \frac{\alpha_3(1-x^2)}{6}\phi(x)\frac{1}{\sqrt{n}} \right| \;=\; o\Big(\frac{1}{\sqrt{n}}\Big),$$

$$\sup_x \left| F_n(x) - \Phi(x) - \frac{\alpha_3(1-x^2)}{6}\phi(x)\frac{1}{\sqrt{n}} + Q_4(x)\phi(x)\frac{1}{n} \right| \;=\; o\Big(\frac{1}{n}\Big),$$

где $\phi(x) = \exp(-\frac{1}{2}x^2)/\sqrt{2\pi}$ *— плотность нормального распределения,*

$Q_4(x) \;=\; \frac{1}{24}\gamma_4(x^3-3x) + \frac{1}{72}\alpha_3^2(x^5 - 10x^3 + 15x)$,

$\gamma_4 \;=\; \alpha_4 - 3\alpha_2^2 \;=\; \alpha_4 - 3$ *— параметр эксцесса.*

⚠ Обратим внимание, что если распределение ξ_k симметрично ($\alpha_3 = 0$) и его параметр эксцесса также равен нулю, то ошибка нормальной апроксимации в ЦПТ имеет порядок $o\big({}^1\!/_n\big)$, а не $\mathcal{O}\big({}^1\!/_{\sqrt{n}}\big)$, как обещает неравенство Берри–Эссеена. С другой стороны, следует отметить, что результаты этой теоремы носят асимптотический характер. Поэтому при малых n указанные здесь добавки не обязательно улучшают аппроксимацию F_n с помощью Φ.

► **Большие уклонения.** В ситуации, когда необходимо вычислить значение $F_n(x)$ при больших положительных (или отрицательных) значениях x, соотношение $F_n(x) \approx \Phi(x)$ малоинформативно: $1 \approx 1$ (или $0 \approx 0$). Здесь важна относительная ошибка аппроксимации (см. 342, стр. 285), т.е. рассматриваемые при $x = x_n \to +\infty$ отношения $F_n(-x)/\Phi(-x)$ и $(1 - F_n(x))/(1 - \Phi(x))$. В этой задаче важную роль играет так называемое условие Крамера, предполагающее, что математическое ожидание $\mathbf{E}\, e^{t\xi} < \infty$ при некотором $t > 0$.

Рассмотрим функцию $\psi(t) = \mathbf{E}\, e^{t\xi}$ в области $t \in [0; t_+)$, где $t_+ = \sup\{t \geqslant 0 :\ \psi(t) < \infty\}$, и определим при $a > 0$

$$\Lambda(a) := \sup\big\{a\lambda - \ln\big[\psi(\lambda)\big] :\ \lambda \in [0; t_+)\big\}.$$

Пусть $\lambda(a)$ — точка достижения указанного максимума. Можно показать (см. [3], гл. 8, § 8*), что функции $\psi, \ln\psi$ выпуклые, поэтому для $\forall\, a$ существует единственное $\lambda(a)$, причём $\Lambda(a) = \int_0^a \lambda(u)du$. Положим $\sigma^2(a) = \big(\ln\psi(t)\big)''\big|_{t=\lambda(a)}$ (> 0 в силу указанной выпуклости).

Для ф.р. F_n сл.в. $S_n/\sqrt{n} = \sum_1^n \xi_j/\sqrt{n}$ справедлива

346| **Теорема.** ([3], гл. 8, § 8*) *Пусть $\big\langle\, \xi_j \,\big\rangle_1^\infty$ — независимые одинаково распределённые сл.в., удовлетворяющие условию Крамера с $t_+ > 0$, причём функция $\varphi_{S_m}(t)/(1 + |t|^\alpha)$ абсолютно интегрируема при некотором $m \geqslant 1$, где φ_{S_m} — хар.ф. S_m. Тогда для любой последовательности $x_n,\ n \geqslant 1$, положительных чисел таких, что $\overline{\lim\limits_n}\, x_n/\sqrt{n} < A_+ := \sup_{t<t_+}\big(\ln\psi(t)\big)'$, имеют место следующие утверждения:*

a) *при $\alpha = 0$ частичные суммы $S_n,\ n \geqslant m$, обладают плотностью, для которой справедливо асимптотическое (при $n \to \infty$) равенство*

$$f_{S_n}(\sqrt{n}x_n) \ \asymp\ \frac{1}{\sigma_n\sqrt{2\pi n}}\, e^{-n\Lambda(x_n/\sqrt{n})},$$

где $\sigma_n^2 = \sigma^2(x_n/\sqrt{n})$;

b) *при $\alpha = 1$*

$$F_n(x_n + a_n) - F_n(x_n) \ \asymp\ \frac{1}{\sigma_n\lambda_n\sqrt{2\pi n}}\, e^{-n\Lambda(x_n/\sqrt{n})}\big(1 - e^{-\sqrt{n}a_n\Lambda(x_n/\sqrt{n})}\big),$$

где $\lambda_n = \lambda(x_n/\sqrt{n})$ и положительные $a_n = o(1)$;

c) *при* $\alpha = 1$ *и* $x_n \to \infty$

$$1 - F_n(x_n) \asymp \frac{1}{\sigma_n \lambda_n \sqrt{2\pi n}} e^{-n\Lambda(x_n/\sqrt{n})}.$$

347| Примеры. 1) Пусть $\xi_j \backsim \mathcal{N}(0,1)$, тогда функция $\psi(t) = \varphi_\xi(\mathrm{i}t) = e^{\frac{1}{2}t^2} < \infty$ для $\forall t > 0$, т.е. $t_+ = \infty$. Функция $a\lambda - \ln\bigl[\psi(\lambda)\bigr] = \frac{1}{2}\bigl(a^2 - (\lambda - a)^2\bigr)$. Следовательно,

$$\Lambda(a) = \frac{1}{2}a^2, \quad \lambda(a) = a, \quad \sigma_a^2 = 1, \quad \lambda_n = \frac{x_n}{\sqrt{n}}.$$

Таким образом, в утверждении a) теоремы знак асимптотического равенства «$\asymp$» следует заменить на знак обычного равенства «$=$». Утверждения b), c) повторяют асимптотику для нормальной ф.р. (см. стр. 311).

2) Пусть теперь $\xi_j + \gamma \backsim \mathcal{E}\mathrm{x}(\gamma)$ с $\gamma = 1$ (для упрощения). Таким образом, $\mathbf{E}\,\xi_j = 0$, $\mathbf{D}\xi_j = 1$. При $\forall t \geqslant 1$ в условии Крамера $\int_0^\infty e^{tx}e^{-x}dx = +\infty$, т.е. $t_+ = 1$ и $\psi(t) = \frac{e^{-t}}{1-t} < \infty$ для $t < 1$. Легко видеть, что здесь $A_+ = +\infty$ и

$$\Lambda(a) = a - \ln(a+1), \quad \lambda(a) = \frac{a}{a+1}, \quad \sigma_a^2 = (1+a)^2.$$

Если $a \to 0$, то $\lambda(a) \asymp a$, $\sigma_a \to 1$, $\Lambda(a) = \frac{1}{2}a^2 + \mathcal{O}(a^3)$. Поэтому для любой последовательности $x_n \to +\infty$ такой, что $x_n = o(n^{1/6})$, имеем

$$1 - F_n(x_n) = \mathbf{P}\Bigl\{\frac{1}{\sqrt{n}}S_n \geqslant x_n\Bigr\} \asymp \frac{1}{x_n\sqrt{2\pi}}e^{-\frac{1}{2}x_n^2} \asymp 1 - \Phi\bigl(x_n\bigr). \qquad \circledcirc$$

Аналогичный результат справедлив в общем случае.

$\bigl[$*О больших уклонениях.*$\bigr]$ ([21], §6, гл. XVI) *Пусть* $\langle \xi_k \rangle_1^\infty$ *— независимые одинаково распределённые сл.в. с* $\mathbf{E}\,\xi_1 = 0$, $\mathbf{D}\xi_1 = 1$, *удовлетворяющие условию Крамера* $\mathbf{E}\,e^{t\xi_1} < \infty$ *при некотором* $t > 0$. *Тогда справедливо следующее:*

a) *если* $\frac{1}{\sqrt[6]{n}}x_n \underset{n}{\to} 0$, *то* $\dfrac{1 - F_n(x_n)}{1 - \Phi(x_n)} = 1 + O\Bigl(\frac{x_n^3}{\sqrt{n}}\Bigr);$

b) *если* $\frac{1}{\sqrt[6]{n}}x_n \underset{n}{\to} \infty$, $\frac{1}{\sqrt[4]{n}}x_n \underset{n}{\to} 0$, *то для* $\mu_3 = \mathbf{E}(\xi_1 - \mathbf{E}\,\xi_1)^3$

$$\frac{1 - F_n(x_n)}{1 - \Phi(x_n)}\exp\Bigl(-\frac{\mu_3}{6}\cdot\frac{x_n^3}{\sqrt{n}}\Bigr) \to 1.$$

§ 6. Закон повторного логарифма

Рассмотрим последовательность независимых одинаково распределённых сл.в. $\langle \xi_k \rangle_1^\infty$ с нулевым математическим ожиданием и конечной дисперсией. Пусть, как всегда, $S_n = \sum_1^n \xi_k$. Из УЗБЧ следует, что нормированная на $b_n = n$ частичная сумма S_n/b_n сходится к нулю почти наверное. Как замечено в 315, стр. 269, даже при $b_n = \sqrt{n}\ln n$ имеем $S_n/b_n \to 0$ (п.н.). В то же время при $b_n = \sqrt{n}$ по ЦПТ $S_n/b_n \rightsquigarrow \mathcal{N}(0,1)$, т.е. имеет в пределе собственное распределение на всей числовой прямой. В 1913 г. Хаусдорф доказал, что для бернуллиевских сл.в. (с $p = {}^1\!/_2$) для $\forall \varepsilon > 0$ частичные суммы $S_n = \mathcal{O}(\sqrt{n}n^\varepsilon)$ (п.н.). Через год в 1914 г. Харди и Литтлвуд уточнили асимптотику для этого симметричного случая: $S_n = \mathcal{O}(\sqrt{n\ln n})$ (п.н.). Хинчин сначала в 1923 г. нашёл, что для общего класса бернуллиевских сл.в. $S_n = \mathcal{O}\big(\sqrt{n\ln(\ln n)}\big)$ (п.н.), а затем через год получил и константу $\sqrt{2}\sigma$ в этом асимптотическом утверждении. Вскоре в 1929 г. Колмогоров обобщил результат Хинчина на случай независимых ограниченных сл.в. Из всех вариантов закона повторного логарифма приведём здесь наиболее просто формулируемый закон, доказанный Хартманом и Винтнером в 1941 г.

348 **Теорема.** [*Закон повторного логарифма. Hartman–Wintner.*] *Пусть $\langle \xi_k \rangle_1^\infty$ — независимые одинаково распределённые сл.в., $\mathbf{E}\xi_k = 0$. Если $\sigma_n^2 = n\mathbf{D}\xi_1 < \infty$, то имеют место соотношения*

$$\varlimsup_n \frac{S_n}{\sqrt{\sigma_n^2 \ln(\ln n)}} = \sqrt{2} \quad (\text{п.н.}), \qquad \varliminf_n \frac{S_n}{\sqrt{\sigma_n^2 \ln(\ln n)}} = -\sqrt{2} \quad (\text{п.н.}). \tag{26}$$

Более того, почти наверное множество предельных точек последовательности $S_n/\sqrt{2\sigma_n^2 \ln(\ln n)}$ совпадает с отрезком $[-1;1]$.

Обратно, если соотношения (26) выполняются для некоторой последовательности $\sigma_n^2 \to \infty$, то $\mathbf{D}\xi_1 < \infty$.

Приведём доказательство прямого утверждения этой теоремы, позаимствованное из работы [30]. Нам понадобятся вспомогательные утверждения, которые докажем позже. Везде в дальнейшем

$$S_n = \sum\nolimits_1^n \xi_j, \quad LL(n) = \ln(\ln n), \quad \lambda(n) = \sqrt{2nLL(n)}, \ n > 2,$$

$LL(1) = LL(2) = 1$. Кроме того, можно положить $\mathbf{D}\xi_1 = 1$.

349| **Лемма.** *Справедливо неравенство* $\sum_{k=1}^{n} \dfrac{1}{\lambda(k)} \leqslant \dfrac{4n}{\sqrt{nLL(n)}}$.

350| **Лемма.** [*Неравенство Оттавиани.*] *Для любой последовательности независимых сл.в.* $\langle \xi_k \rangle_1^n$ *при* $\forall\, b \in \mathbb{R}^1,\ \varepsilon > 0$

$$\mathbf{P}\Big\{ \max_{1\leqslant j\leqslant n} S_n > b+\varepsilon \Big\}\Big(\min_{1\leqslant m<n} \mathbf{P}\left\{S_{m,n} > -\varepsilon\right\} \Big) \ \leqslant\ \mathbf{P}\left\{S_n > b\right\},$$

где $S_{m,n} = \sum_{k=m+1}^{n} \xi_k$.

351| **Лемма.** *Пусть* $\langle \xi_k \rangle_1^\infty$ *— независимые сл.в. такие, что*

(1) $\dfrac{1}{\lambda(n)} S_n \xrightarrow{\mathbf{P}} 0$,

(2) *существуют* $a > 1,\ b, c > 0, n_0 \geqslant 1$, *при которых*

$$\mathbf{P}\Big\{ \frac{1}{\lambda(n)} S_n > b \Big\} \ \leqslant\ c\, e^{-a\, LL(n)}, \quad \forall\, n \geqslant n_0.$$

Тогда $\mathbf{P}\Big\{ \varlimsup_n \dfrac{1}{\lambda(n)} S_n \leqslant b \Big\} \ =\ 1$.

352| **Лемма.** *Пусть* $\langle \xi_k \rangle_1^\infty$ *— независимые сл.в. такие, что*

(1) $\mathbf{E}\,\xi_j = 0,\ \rho = \sqrt{\sup_j \mathbf{D}\xi_j} < \infty$,

(2) $|\xi_j| \leqslant \tau\sqrt{j/LL(j)}$ (п.н.) *при некотором* $\tau > 0$.

Тогда для $\forall\, t > 0,\ a \geqslant \rho,\ n \geqslant 1$

$$\mathbf{P}\Big\{ \frac{1}{\lambda(n)} S_n > t \Big\} \ \leqslant\ \exp\Big\{ -\frac{t^2}{a^2}\big(2 - e^{\sqrt{2}\,\tau t/a^2}\big) LL(n) \Big\}.$$

353| **Лемма.** *Пусть* $\langle \xi_k \rangle_1^\infty$ *— независимые одинаково распределённые сл.в. с* $\mathbf{D}\xi_1 < \infty$. *Определим для* $\forall\, \tau > 0$ *сл.в.* $\eta_j = \xi_j\, \dot{\mathrm{I}}\,\big(|\xi_j| > \tau\sqrt{j/LL(j)}\,\big)$ *и* $T_n = \sum_1^n \eta_j$. *Тогда*

$$\text{i)}\ \sum\nolimits_1^\infty \frac{1}{\lambda(j)}\, \mathbf{E}\,|\eta_j| \ <\ \infty, \qquad \text{ii)}\ \mathbf{P}\Big\{ \lim_n \frac{T_n}{\lambda(n)} = 0 \Big\} \ =\ 1.$$

354| **Лемма.** *Пусть* $\langle \xi_k \rangle_1^\infty$ *— независимые одинаково распределённые сл.в. с* $\mathbf{E}\,\xi_1 = 0,\ \mathbf{D}\xi_1 = 1$. *Пусть* $n_k \in \mathbb{N},\ \alpha_k > 0,\ \alpha_k/m_k \to 0,\ \alpha_k/m_k^2 \to \infty$. *Тогда для* $\forall\, \varepsilon > 0,\ b \in \mathbb{R}^1$

$$\varliminf_n \frac{m_k}{\alpha_k^2} \ln\Big(\mathbf{P}\Big\{ \Big|\frac{1}{\alpha_k} S_{m_k} - b\Big| < \varepsilon \Big\}\Big) \ \geqslant\ -\frac{1}{2} b^2.$$

⊫ (Закон повторного логарифма.) Покажем сначала, что

$$(ll_1) \qquad \mathbf{P}\Big\{\ \varlimsup_n \frac{1}{\lambda(n)} S_n \leqslant 1\ \Big\} \ = \ 1.$$

Пусть $\delta > 0$, выберем достаточно малое $\tau > 0$, чтобы

$$\alpha \ = \ (1+\delta)^2\big(2 - e^{\sqrt{2}\,(1+\delta)\tau}\big) \ > \ 1.$$

Определим сл.в. $\xi'_j = \xi_j\,\dot{\mathrm{I}}\,\big(|\xi_j| \leqslant \frac{\tau}{2}\sqrt{j/LL(j)}\,\big)$ и $\eta_j = \xi'_j - \mathbf{E}\,\xi'_j$.

355| Упр. Пусть $\mathbf{E}\,\zeta = 0$, покажите, что $\mathbf{D}\big[\zeta\,\dot{\mathrm{I}}\,(|\zeta| < c)\big] \leqslant \mathbf{D}\zeta$.

Применим лемму 352 к сл.в. η_j с $a = 1$ $(= \mathbf{D}\xi_j)$, $t = 1+\delta$ и $S'_n = \sum_1^n \eta_j$:

$$\mathbf{P}\Big\{\ \frac{1}{\lambda(n)} S'_n > 1+\delta\ \Big\} \ \leqslant \ e^{-\alpha LL(n)}.$$

В силу неравенства Чебышёва и ввиду 355

$$\mathbf{P}\Big\{\ \frac{1}{\lambda(n)}\big|S'_n\big| > \varepsilon\ \Big\} \ \leqslant \ \frac{n}{\varepsilon^2\lambda(n)^2} \ = \ \frac{n}{\varepsilon^2 2nLL(n)} \ \underset{n}{\to} \ 0.$$

В соответствии с леммой 351 предел $\varlimsup_n \dfrac{1}{\lambda(n)} S'_n \ \leqslant \ 1+\delta$ (п.н.).

Положим $\zeta_j = \xi_j - \xi'_j$. Поскольку $\mathbf{E}\,\xi'_j = -\,\mathbf{E}\,\zeta_j$, то в силу леммы 353

$$\sum\nolimits_1^\infty \frac{1}{\lambda(j)}\big|\,\mathbf{E}\,\xi'_j\big| \ \leqslant \ \sum\nolimits_1^\infty \frac{1}{\lambda(j)}\,\mathbf{E}\,\big|\zeta_j\big| \ < \ \infty.$$

По лемме Кронекера отсюда получаем

$$\Big|\frac{1}{\lambda(n)}\sum\nolimits_1^n \mathbf{E}\,\xi'_j\Big| \ \leqslant \ \frac{1}{\lambda(n)}\sum\nolimits_1^n \big|\,\mathbf{E}\,\xi'_j\big| \ \to \ 0.$$

Воспользовавшись вторым утверждением леммы 353, окончательно получаем

$$\varlimsup_n \frac{1}{\lambda(n)} S_n \ = \ \varlimsup_n \frac{1}{\lambda(n)}\Big(S'_n + \sum\nolimits_1^n \mathbf{E}\,\xi'_j + \sum\nolimits_1^n \zeta_j\Big) \ \leqslant \ (1+\delta) \quad (\text{п.н.}),$$

что в силу произвольности δ доказывает (ll_1). Применяя полученный результат к сл.в. $-\xi_j$, заключаем, что п.н.

$$(ll_2) \qquad -1 \ \leqslant \ \varliminf_n \frac{1}{\lambda(n)} S_n \ \leqslant \ \varlimsup_n \frac{1}{\lambda(n)} S_n \ \leqslant \ 1.$$

Докажем теперь, что для $\forall\,|b| < 1$ п.н. найдётся подпоследовательность $\lim_{n'} S_{n'}/\lambda(n') = b$ или, по-другому, $\varliminf_n |S_n/\lambda(n) - b| = 0$. Положим $n_k = k^k$, тогда, очевидно, $\lambda(n_{k-1})/\lambda(n_k) \to 0,\ k \to \infty$. Поэтому в неравенстве

$$(ll_3) \qquad \Big|\frac{1}{\lambda(n_k)} S_{n_k} - b\Big| \leqslant \Big|\frac{1}{\lambda(n_k)} S_{n_{k-1}}\Big| + \Big|\frac{1}{\lambda(n_k)}(S_{n_k} - S_{n_{k-1}}) - b\Big|$$

первое слагаемое п.н. стремится к нулю в силу доказанного (ll_2). Для анализа второго слагаемого воспользуемся леммой 354 с $m_k = n_k - n_{k-1},\ \alpha_k = \lambda(n_k)$. При таком выборе, очевидно, $m_k \asymp n_k$ и

$$\frac{\alpha_k}{m_k} \asymp \frac{\sqrt{2k^k \ln(\ln k^k)}}{k^k} \to 0, \qquad \frac{\alpha_k^2}{m_k} \asymp \frac{2k^k \ln(\ln k^k)}{k^k} = 2LL(n_k) \to \infty.$$

В силу леммы 354 для $\forall\,\delta > 0$, начиная с некоторого k_0, т.е. при $\forall\,k > k_0$

$$\mathbf{P}\Big\{ \frac{1}{\alpha_k}|S_{m_k} - b| < \varepsilon \Big\} \geqslant \exp\Big\{ -\frac{\alpha_k^2}{2n_k}(b^2+\delta)\Big\} \geqslant \exp\{-LL(k^k)(b^2+\delta)\},$$

где учтено также, что $m_k/n_k \to 1$. Выберем δ так, чтобы $\theta = b^2+\delta < 1$. Тогда для событий $A_k = \{\big|\frac{1}{\lambda(n_k)}(S_{n_k} - S_{n_{k-1}}) - b\big| < \varepsilon\}$ сумма вероятностей

$$\sum\nolimits_{k=k_0}^{\infty} \mathbf{P}\{A_k\} \geqslant \sum\nolimits_{k=k_0}^{\infty} \frac{1}{k^\theta\big[\ln(k)\big]^\theta} = \infty.$$

Так как события A_k независимы, то по лемме Бореля–Кантелли для $\forall\,\varepsilon > 0$

$$\mathbf{P}\Big\{ \varlimsup_k \Big|\frac{1}{\lambda(n_k)}(S_{n_k} - S_{n_{k-1}}) - b\Big| < \varepsilon \Big\} = 1,$$

что доказывает существование п.н. подпоследовательности, для которой второе слагаемое в правой части (ll_3) будет стремиться к нулю. ⫤

Перейдём теперь к доказательству вспомогательных лемм.

⊨ (349) По формуле Тейлора для любой дифференцируемой функции приращение $h(x) - h(x+1) = -h'(x+\theta)$, где $0 \leqslant \theta \leqslant 1$. Легко проверить, что для функции $h(x) = 1/\sqrt{x\,LL(x)}$ производная в точке $\widetilde{x} = x + \theta$

$$-h'(\widetilde{x}) \leqslant \frac{1}{2\widetilde{x}\sqrt{\widetilde{x}\,LL(\widetilde{x})}} \leqslant \frac{1}{2x\sqrt{x\,LL(x)}}.$$

Полагая $a_k = 1/\sqrt{k\,LL(k)}$ и замечая, что $(k+1)/k \leqslant 3/2,\ \forall\,k \geqslant 2$, отсюда приходим к неравенству $(k+1)(a_k - a_{k+1}) \leqslant \frac{3}{4}a_k,\ \forall\,k \geqslant 1$. Следовательно,

$$\sum_{k=1}^{n} a_k = \sum_{k=1}^{n-1}(k+1)(a_k - a_{k+1}) + (n+1)a_n - a_1 < \frac{3}{4}\sum_{k=1}^{n} a_k + \big(n + \tfrac{1}{4}\big)a_n.$$

Таким образом, искомый ряд не превосходит $(4n+1)a_n/\sqrt{2} < 4n/\sqrt{nLL(n)}$.

(350) Пусть $\nu = \min\{m \geqslant 1 : \; S_m > b+\varepsilon\}$, тогда, очевидно, событие $\{\max_{1\leqslant j\leqslant n} S_j > b+\varepsilon\} = \{\nu \leqslant n\}$. Кроме того (с $S_{n,n}=0$),

$$\biguplus_{j=1}^{n}\{\nu = j, S_{j,n} > -\varepsilon\} \subset \{S_n > b\}.$$

В силу независимости событий $\{\nu = j\}$ и $S_{j,n} > -\varepsilon$ (для $\forall\, j \leqslant n$) отсюда получаем, что

$$\begin{aligned}\mathbf{P}\{S_n > b\} &\geqslant \sum\nolimits_{j=1}^{n}\mathbf{P}\{\nu = j\}\,\mathbf{P}\{S_{j,n} > -\varepsilon\} \geqslant \\ &\geqslant \min_{1\leqslant j\leqslant n}\mathbf{P}\{S_{j,n} > -\varepsilon\}\,\mathbf{P}\Big\{\max_{1\leqslant j\leqslant n} S_j > b+\varepsilon\Big\}.\end{aligned}$$

(351) Пусть $\gamma > 1$, рассмотрим последовательность натуральных чисел $n_k = [\![\gamma^{2k}]\!]$, $k \geqslant 1$. Очевидно, что $n_k < n_{k+1}-1 < \gamma^2 n_k$. Поэтому $n_{k+1}/n_k \to \gamma^2$ и $\lambda(n_{k+1}) \asymp \gamma\lambda(n_k)$ при $k\to\infty$. По неравенству Оттавиани и условию (2)

$$\mathbf{P}\Big\{\frac{1}{\lambda(n_k)}\max_{1\leqslant j\leqslant n_k} S_{n_k} > \gamma b\Big\} \leqslant \frac{1}{C_k}\,\mathbf{P}\Big\{\frac{1}{\lambda(n_k)}S_{n_k} > b\Big\} \leqslant \frac{1}{C_k}e^{-aLL(n_k)},$$

где $C_k = \min_{1\leqslant j\leqslant n_k}\mathbf{P}\{S_j/\lambda(n_k) > -(\gamma-1)b\}$. По условию (1) леммы $C_k \to 1$ при $n_k\to\infty$. Поэтому можно считать, что $1/C_k \leqslant 2$. Так как $n_k \geqslant \gamma^{2k}$, то

$$\sum\nolimits_1^\infty e^{-aLL(n_k)} \leqslant \sum\nolimits_1^\infty \frac{1}{k^a(\ln(\gamma^2))^a} < \infty.$$

Вероятность того, что события $\big\{\frac{S_n}{\lambda(n)} > \gamma^2 b\big\}_{n=1}^{\infty}$ произойдут бесконечно часто, равна нулю, т.к.

$$\begin{aligned}\mathbf{P}\big\{S_n > b\gamma^2\lambda(n) \;\text{ б.ч. по } n\big\} &\leqslant \mathbf{P}\Big\{\max_{n_{k-1}<j\leqslant n_k} S_j > b\gamma^2\lambda(n_{k-1}) \;\text{ б.ч. по } k\Big\} \leqslant \\ &\leqslant \mathbf{P}\Big\{\max_{1\leqslant j\leqslant n_k} S_j > b\gamma\lambda(n_k) \;\text{ б.ч. по } k\Big\} = 0\end{aligned}$$

в силу леммы Бореля–Кантелли и доказанных выше соотношений. Таким образом, п.н. $\overline{\lim}_n S_n/\lambda(n) \leqslant \gamma^2 b$ при $\forall\,\gamma > 1$, что и требовалось.

(352) По условию леммы при $j \leqslant n$

$$\frac{1}{\lambda(n)}|\xi_j| \leqslant \frac{1}{\lambda(n)}\tau\sqrt{\frac{j}{LL(j)}} \leqslant \frac{1}{\sqrt{2nLL(n)}}\tau\sqrt{\frac{n}{LL(n)}} = \frac{\tau}{\sqrt{2}LL(n)}.$$

Применяя неравенство $e^x \leqslant 1+x+\frac{x^2}{2}e^{|x|}$, $x\in\mathbb{R}^1$, и вспоминая, что по условию леммы $\mathbf{E}\xi_j = 0, \mathbf{D}\xi_j \leqslant a^2$, получаем для $\forall\,\gamma > 0$

$$\mathbf{E}\exp\Big(\frac{\gamma\xi_j}{\lambda(n)}\Big) \leqslant 1 + \frac{\gamma^2a^2}{2\lambda^2(n)}\exp\Big(\frac{\gamma\tau}{\sqrt{2}LL(n)}\Big) \leqslant \exp\Big\{\frac{\gamma^2a^2}{2\lambda^2(n)}\exp\Big(\frac{\gamma\tau}{\sqrt{2}LL(n)}\Big)\Big\}.$$

По неравенству Маркова и в силу независимости сл.в.

$$\mathbf{P}\Big\{\,\frac{1}{\lambda(n)}S_n > t\,\Big\} \;\leqslant e^{-\gamma t}\,\mathbf{E}\exp\Big\{\frac{1}{\lambda(n)}\gamma S_n\Big\} \;\leqslant$$

$$\leqslant\; \exp\Big\{-\gamma t + \frac{\gamma^2 a^2}{4LL(n)}\exp\Big(\frac{\gamma\tau}{\sqrt{2}LL(n)}\Big)\Big\}.$$

Полагая здесь $\gamma = 2t\,LL(n)/a^2$, придём к требуемому соотношению.

(353) По лемме Кронекера для доказательства `ii)` достаточно доказать, что п.н. сходится ряд $\sum_1^\infty \eta_j/\lambda(j)$. Здесь пригодится следующее утверждение.

356| Упр. Если $\sum_1^\infty \mathbf{E}\,|\zeta_j| < \infty$, то п.н. $\sum_1^\infty |\zeta_j| < \infty$.

Итак, `ii)` следует из `i)`. Пусть $\gamma_j = \sqrt{j/LL(j)}$, тогда в силу 349

$$\Lambda_k \;:=\; \sum\nolimits_{j=1}^{k}\frac{1}{\lambda(j)} \;\leqslant\; 4\gamma_k.$$

Следовательно, после замены порядка суммирования получаем

$$\sum_{j=1}^{\infty}\frac{1}{\lambda(j)}\,\mathbf{E}\,|\eta_j| \;=\; \sum_{j=1}^{\infty}\frac{1}{\lambda(j)}\Big(\sum_{k=j}^{\infty}\mathbf{E}\left[|\xi_j|\;\dot{\mathrm{I}}\,(\tau\gamma_k < |\xi_j| \leqslant \tau\gamma_{k+1})\right)\;\leqslant$$

$$\leqslant\; \sum_{j=1}^{\infty}\frac{1}{\lambda(j)}\Big(\sum_{k=j}^{\infty}\tau\gamma_{k+1}\,\mathbf{P}\left\{\tau\gamma_k < |\xi_j| \leqslant \tau\gamma_{k+1}\right\}\Big) \;=$$

$$=\; \sum_{k=1}^{\infty}\tau\gamma_{k+1}\,\mathbf{P}\left\{\tau\gamma_k < |\xi_j| \leqslant \tau\gamma_{k+1}\right\}\Lambda_k \;\leqslant$$

$$\leqslant\; \frac{4}{\tau}\sum_{k=1}^{\infty}\tau^2\gamma_{k+1}\gamma_k\,\mathbf{P}\left\{\tau\gamma_k < |\xi_j| \leqslant \tau\gamma_{k+1}\right\} \;\leqslant\; \frac{8}{\tau}\,\mathbf{D}\xi_1 \;<\; \infty,$$

поскольку, как можно заметить, $\gamma_{k+1} < 2\gamma_k$.

(354) Здесь применяется одно свойство стандартного нормального закона (с плотностью $\phi(x)$). Обозначим через $\nu(B) = \int_B \phi(x)\,dx$ вероятность множества B, определяемую нормальным распределением. Тогда при $t > 0, b \in \mathbb{R}^1$

$$\nu(tb - t\varepsilon; tb + t\varepsilon) \;=\; \int_{tb-t\varepsilon}^{tb+t\varepsilon}\phi(x)\,dx \;=\; \int_{-t\varepsilon}^{t\varepsilon}e^{tbx}\phi(x)\,dx\,\exp\Big\{-\frac{1}{2}t^2b^2\Big\}.$$

Применяя неравенство Йенсена (относительно соответствующего симметричного условного распределения), получаем

$$\nu(tb - t\varepsilon; tb + t\varepsilon) \;\geqslant\; \nu(-t\varepsilon; t\varepsilon)\exp\Big\{-\frac{1}{2}t^2b^2\Big\}.$$

Следовательно, т.к. $\ln\big(\nu(-t\varepsilon;+t\varepsilon)\big)\to 0$ при $t\to\infty$, то

$$(N_t)\qquad \frac{1}{t^2}\ln(\nu(tb-t\varepsilon;tb+t\varepsilon)) \;\geqslant\; -\frac{1}{2}b^2+\frac{1}{t^2}\ln(\nu(-t\varepsilon;+t\varepsilon)) \;\;\to\;\; -\frac{1}{2}b^2.$$

Очевидно, что для $\forall\,A,B\in\mathbb{R}^1$ и любой возрастающей последовательности натуральных чисел $n_j,\ 1\leqslant j\leqslant z,\ n_0=0$,

$$\mathbf{P}\Big\{\bigcap\nolimits_{j=0}^{z-1}\big(A<\sum\nolimits_{i=n_j+1}^{n_{j+1}}\xi_i<B\big)\Big\}\;\leqslant\;\mathbf{P}\Big\{kA<\sum\nolimits_{i=1}^{n_z}\xi_i<kB\big)\Big\}.$$

Поэтому для любых целых n_k,z_k и действительных $r_k\in\mathbb{R}^1$

$$\left(\mathbf{P}\left\{A<\frac{1}{r_k}S_{n_k}<B\right\}\right)^{z_k}\;\leqslant\;\mathbf{P}\left\{A<\frac{1}{z_kr_k}S_{n_kz_k}<B\right\}$$

в силу независимости и одинаковой распределённости слагаемых. Выберем здесь $n_k=\llbracket m_k^2t^2/\alpha_k^2\rrbracket,\ z_k=\llbracket\alpha_k^2/t^2m_k\rrbracket$, тогда, очевидно, $n_kz_k\leqslant m_k$ и по условию $(m_k-n_kz_k)/\alpha_k^2\to 0$. По неравенству Чебышёва

$$\gamma_k:=\;\mathbf{P}\left\{\frac{1}{\alpha_k}\big|S_{m_k}-S_{n_kz_k}\big|>\varepsilon\right\}\;\leqslant\;\frac{1}{\varepsilon^2}\frac{(m_k-n_kz_k)}{\alpha_k^2}\;\to\;0.$$

В соответствии с предыдущим неравенством

$$\mathbf{P}\left\{b-\varepsilon<\frac{1}{\alpha_k}S_{n_kz_k}<b+\varepsilon\right\}\;\geqslant\;\left(\mathbf{P}\left\{tb-t\varepsilon<\frac{tz_k}{\alpha_k}S_{n_k}<tb+t\varepsilon\right\}\right)^{z_k}.$$

Заметим, что $\alpha_k/tz_k\asymp\sqrt{n_k}$, поэтому здесь применима центральная предельная теорема. Поскольку

$$\mathbf{P}\left\{b-2\varepsilon<\frac{1}{\alpha_k}S_{m_k}<b+2\varepsilon\right\}\;\geqslant\;\mathbf{P}\left\{b-\varepsilon<\frac{1}{\alpha_k}S_{n_kz_k}<b+\varepsilon\right\}(1-\gamma_k),$$

то из предыдущего, с учётом того, что $m_kz_k/\alpha_k^2\to 1$, получаем

$$\begin{aligned}
&\varliminf_k \ln\Big(\mathbf{P}\Big\{\,b-2\varepsilon<\frac{1}{\alpha_k}S_{m_k}<b+2\varepsilon\,\Big\}\Big)\frac{m_k}{\alpha_k^2}\;\geqslant\\
&\quad\geqslant\;\varliminf_k \ln\Big(\mathbf{P}\left\{tb-t\varepsilon<\frac{tz_k}{\alpha_k}S_{n_kz_k}<tb+t\varepsilon\right\}\Big)\frac{m_kz_k}{\alpha_k^2}+\varliminf_k\ln\big(1-\gamma_k\big)\frac{m_k}{\alpha_k^2}\;=\\
&\quad=\;\frac{1}{t^2}\ln\big(\nu(tb-t\varepsilon;tb+t\varepsilon)\big).
\end{aligned}$$

Полагая здесь $t\to\infty$, в силу (N_t) приходим к утверждению леммы. $\Lleftarrow$

A. Доказательства вспомогательных утверждений

✦ Из тождества $a_1a_2-b_1b_2=(a_1-b_1)a_2+(a_2-b_2)b_1$ следует по неравенству треугольника, что $|a_1a_2-b_1b_2| \leqslant |(a_1-b_1)||a_2|+|(a_2-b_2)||b_1| \leqslant |(a_1-b_1)|+|(a_2-b_2)|$ для любых комплексных чисел $|a_2|, |b_1| \leqslant 1$. Легко проверяется по индукции, что справедливо более общее неравенство

$$\Big|\prod_{k=1}^{n} a_k-\prod_{k=1}^{n} b_k\Big| \leqslant \sum_{k=1}^{n}|a_k-b_k|, \qquad |a_k|,|b_k| \leqslant 1,\ k=\overline{1,n}. \tag{27}$$

⤇ [Лемма 293, стр. 255] (a) (Вторая часть.) Зафиксируем $\varepsilon>0$ и выберем n_0 так, что $(1-\varepsilon)a<a_n<(1+\varepsilon)a$ и $(1-\varepsilon)b<b_n<(1+\varepsilon)b$ при всех $n>n_0$ (в случае $a=0$ или $b=0$ неравенства заменяются на $-\varepsilon \leqslant a_n\ (b_n)<\varepsilon$). Ввиду симметрии индексов последовательности $b_{|n-k|}$

$$\frac{1}{N^2}\sum_{n,k=1}^{N} a_n b_{|n-k|} \;=\; \frac{2}{N^2}\sum_{n=1}^{N-1} a_n \sum_{k=n+1}^{N} b_{k-n}+\frac{1}{N^2} b_0 \sum_{n=1}^{N} a_n.$$

Последовательность a_n ограничена, скажем, числом A. Поэтому второе слагаемое здесь по модулю не превосходит $A|b_0|/N \to 0$ при $N \to \infty$.

В первом слагаемом произведём замену переменных $n \dashrightarrow x, k \dashrightarrow n+y$. Легко понять, что область суммирования примет вид $1 \leqslant x \leqslant N-1, 1 \leqslant y \leqslant N-x$ и что в этой области всего $N(N-1)/2$ точек. Выделим в этой области подобласть Ω_0, содержащую точки $n_0<x \leqslant N-n_0, n_0<y \leqslant N-n_0$. Количество точек в этой области $N_0=(N-2n_0)(N-2n_0+1)/2$. Заметим, что $2N_0/N^2 \to 1$ при $N \to \infty$. Количество точек в остальной части (Ω_1) области суммирования равно $N_1=N(N-1)/2-N_0$, и, следовательно, $2N_1/N^2 \to 0$.

В области Ω_0, где $x, y>n_0$, имеем

$$2(1-\varepsilon)^2 ab\frac{N_0}{N^2} \;\leqslant\; \frac{2}{N^2}\sum_{(x,y)\in\Omega_0} a_x b_y \;\leqslant\; 2(1+\varepsilon)^2 ab\frac{N_0}{N^2}$$

с очевидными изменениями в случае $ab=0$. Полагая здесь $N \to \infty$, получаем

$$(1-\varepsilon)^2 ab \;\leqslant\; \varliminf_{N} \frac{2}{N^2}\sum_{(x,y)\in\Omega_0} a_x b_y \;\leqslant\; \varlimsup_{N} \frac{2}{N^2}\sum_{(x,y)\in\Omega_0} a_x b_y \;\leqslant\; (1+\varepsilon)^2 ab.$$

Для суммы по индексам $(x,y) \in \Omega_1$ ввиду ограниченности сходящихся последовательностей $(a_n < A, b_n < B)$ при $N \to \infty$

$$\frac{2}{N^2} \sum_{(x,y)\in\Omega_1} a_x b_y \leqslant AB\frac{2N_1}{N^2} \to 0.$$

Поскольку ε произвольно, вторая часть леммы доказана. Первая часть доказывается (?!) аналогично.

(б) Введём вспомогательную сумму $V_n = \sum\limits_{j=n}^{\infty} \frac{1}{b_j} a_j,\ n \geqslant 1$. Тогда

$$\frac{1}{b_n}\sum_{k=1}^{n} a_k = \frac{1}{b_n}\sum_{k=1}^{n} b_k \frac{a_k}{b_k} = \frac{1}{b_n}\sum_{k=1}^{n} b_k(V_k - V_{k+1}) =$$
$$= \frac{1}{b_n}\big(b_1V_1 - b_1V_2 + b_2V_2 - b_2V_3 + \ldots + b_nV_n - b_nV_{n+1}\big) =$$
$$= \frac{1}{b_n}\sum_{k=1}^{n-1} V_k(b_{k+1} - b_k) + \frac{b_1}{b_n}V_1 - V_{n+1}.$$

Так как $V_1 < \infty$, то $V_n \underset{n}{\to} 0$, т.е. для любого $\varepsilon > 0$ можно подобрать N_0 так, чтобы $|V_n| < \varepsilon,\ \forall k \geqslant N_0$. Поэтому в силу монотонности b_n для $\forall n > N_0$

$$\left|\frac{1}{b_n}\sum_{k=N_0}^{n-1} V_k(b_{k+1} - b_k)\right| \leqslant \varepsilon\frac{1}{b_n}\sum_{k=N_0}^{n-1}(b_{k+1} - b_k) = \varepsilon\Big(1 - \frac{b_{N_0}}{b_n}\Big) \leqslant \varepsilon.$$

Слагаемое b_1V_1/b_n и сумма $\sum_{k=1}^{N_0-1} V_k(b_{k+1} - b_k)/b_n$ могут быть сделаны меньше ε ввиду $b_n \to \infty$ при $n \to \infty$. ⫤

⫢ [Лемма 299, стр. 259] (a) Здесь достаточно заметить, что неравенство $|\zeta_1 - \zeta_2| \geqslant 2t$ влечёт выполнение хотя бы одного из неравенств $|\zeta_1| \geqslant t$ или $|\zeta_2| \geqslant t$. В условиях леммы вероятности этих двух событий совпадают.

(b) По определению медианы $\mathbf{P}\{\zeta_2 \geqslant m_2\} \geqslant \frac{1}{2} \leqslant \mathbf{P}\{\zeta_2 \leqslant m_2\}$. Поэтому

$$\frac{1}{2}\,\mathbf{P}\big\{\,|\zeta_1 - m_2| \geqslant t\,\big\} = \frac{1}{2}\,\mathbf{P}\big\{\,\zeta_1 - m_2 \geqslant t\,\big\} + \frac{1}{2}\,\mathbf{P}\big\{\,\zeta_1 - m_2 \leqslant -t\,\big\} \leqslant$$
$$\leqslant \mathbf{P}\big\{\,\zeta_1 \geqslant m_2 + t,\ m_2 \geqslant \zeta_2\,\big\} + \mathbf{P}\big\{\,\zeta_1 \leqslant m_2 - t,\ m_2 \leqslant \zeta_2\,\big\} \leqslant$$
$$\leqslant \mathbf{P}\big\{\,\zeta_1 - \zeta_2 \geqslant t\,\big\} + \mathbf{P}\big\{\,\zeta_1 - \zeta_2 \leqslant -t\,\big\} = \mathbf{P}\big\{\,|\zeta_1 - \zeta_2| \geqslant t\,\big\}$$

в силу независимости ζ_1, ζ_2.

(c) Распределение сл.в. симметрично т. т. т. когда её хар.ф. есть действительная функция. Хар.ф. суммы независимых сл.в. равна произведению соответствующих хар.ф., т.е. остаётся действительной функцией. Введём сл.в.

$M = \max(|\zeta_1|, \ldots, |\zeta_n|)$ и $\nu = \min\{k : |\zeta_k| = M\}$. Хотя они зависят от всего вектора $(\zeta_1, \ldots, \zeta_n)$, однако понятно, что они не изменяются при замене $(\zeta_1, \ldots, \zeta_n) \dashrightarrow (\pm\zeta_1, \ldots, \pm\zeta_n)$ с любым сочетанием знаков. Кроме того, воспользовавшись методом хар.ф., легко показать, что в рассматриваемом случае с независимыми и симметричными сл.в. $(\zeta_1, \ldots, \zeta_n) \backsim (\pm\zeta_1, \ldots, \pm\zeta_n)$, т.е. вероятности любых событий, связанных со сл.вектором $(\zeta_1, \ldots, \zeta_n)$, совпадают с аналогичными вероятностями относительно $(\pm\zeta_1, \ldots, \pm\zeta_n)$. Поэтому

$$\mathbf{P}\{\zeta_\nu \geqslant t, S - \zeta_\nu \leqslant 0\} = \sum_{k=1}^{n} \mathbf{P}\Big\{\nu = k, \zeta_k \geqslant t, \sum_{j \neq k} \zeta_j \leqslant 0\Big\} =$$

$$= \sum_{k=1}^{n} \mathbf{P}\Big\{\nu = k, \zeta_k \geqslant t, \sum_{j \neq k} (-\zeta_j) \leqslant 0\Big\} = \mathbf{P}\{\zeta_\nu \geqslant t, S - \zeta_\nu \geqslant 0\}.$$

Таким образом,

$$\mathbf{P}\{\zeta_\nu \geqslant t\} \leqslant \mathbf{P}\{\zeta_\nu \geqslant t, S - \zeta_\nu \leqslant 0\} + \mathbf{P}\{\zeta_\nu \geqslant t, S - \zeta_\nu \geqslant 0\} =$$
$$= 2\,\mathbf{P}\{\zeta_\nu \geqslant t, S - \zeta_\nu \geqslant 0\} \leqslant 2\,\mathbf{P}\{S \geqslant t\}.$$

Аналогично, $\mathbf{P}\{\zeta_\nu \leqslant -t\} \leqslant 2\,\mathbf{P}\{S \leqslant -t\}$, что и требовалось. ⇚

⇛ [*Формула Стирлинга* 338, стр. 284] В известном разложении в ряд Тейлора логарифмической функции ($|x| < 1$)

$$\ln\frac{1+x}{1-x} = \ln(1+x) - \ln(1-x) = 2x \sum\nolimits_0^\infty \frac{x^{2k}}{2k+1}$$

выберем $x = X_n = 1/(2n+1)$. Получим

$$\frac{1}{2X_n} \ln\frac{1+X_n}{1-X_n} = 1 + \left(\frac{X_n^2}{3} + \frac{X_n^4}{5} + \ldots\right) =: 1 + \delta_n.$$

Очевидно,

$$\delta_n < \frac{X_n^2}{3}(1 + X_n^2 + X_n^4 + \ldots) = \frac{X_n^2}{3(1 - X_n^2)} = \frac{1}{12n(n+1)} = \frac{1}{12}\Big(\frac{1}{n} - \frac{1}{n+1}\Big).$$

Поэтому $C := \sum_1^\infty \delta_k < \infty$ и, кроме того, $\varepsilon_n = \sum_n^\infty \delta_k < \frac{1}{12n}$. С другой стороны, $\delta_n > \frac{1}{3} X_n^2 > \frac{1}{12n+1} - \frac{1}{12(n+1)+1}$, следовательно,

$$\frac{1}{12n+1} < \varepsilon_n < \frac{1}{12n}.$$

Заметим, что

$$\prod_{k=1}^{n} \frac{(k+1)^{k+1/2}}{k^{k+1/2}} = \prod_{k=1}^{n} \frac{(k+1)^{k+1}}{k^k k^{1/2}(k+1)^{1/2}} = \frac{(n+1)^{n+1/2}}{n!} .$$

Стало быть,

$$\ln \frac{(n+1)^{n+1/2}}{n!} = \sum\nolimits_1^n \ln\big(1+\tfrac{1}{k}\big)^{k+1/2} = \sum\nolimits_1^n \frac{1}{2X_k} \ln \frac{1+X_k}{1-X_k} = n + C - \varepsilon_{n+1}.$$

Следовательно,

$$(n+1)! = (n+1)n! = (n+1)^{n+3/2}\, e^{-n-C+\,\varepsilon_{n+1}}$$

или $n! = \gamma\, n^{n+1/2}\, \exp\{-n+\varepsilon_n\}$, где $\gamma = e^{1-C}$.

Покажем, что $\gamma = \sqrt{2\pi}$. Воспользуемся тригонометрическим представлением для бета-функции:

$$2\int_0^{\pi/2} \sin^k x\, dx = \mathrm{B}\big(\frac{k+1}{2}, \frac{1}{2}\big) = \frac{\Gamma((k+1)/2)\,\Gamma(1/2)}{\Gamma((k+2)/2)} ,$$

из которого следует, что эта функция убывает по k, т.е.

$$\mathrm{B}\big(\frac{k+2}{2}, \frac{1}{2}\big) < \mathrm{B}\big(\frac{k+1}{2}, \frac{1}{2}\big) < \mathrm{B}\big(\frac{k}{2}, \frac{1}{2}\big) .$$

Поскольку $\Gamma(1/2) = \sqrt{\pi}$ и при чётном $k = 2n$ гамма-функция

$$\Gamma(2n/2) = (n-1)!, \qquad \Gamma((2n+1)/2) = \frac{(2n-1)!!\sqrt{\pi}}{2^n} = \frac{(2n)!\sqrt{\pi}}{2^{2n} n!} ,$$

из монотонности бета-функции получаем

$$\frac{(2^n n!)^2}{(2n+1)!} < \frac{(2n+1)!\,\pi}{(2n+1)(2^n n!)^2} < \frac{(2n+1)(2^n n!)^2}{2n(2n+1)!} .$$

Отсюда имеем

$$\frac{(2n+1)(2^n n!)^4}{\big((2n+1)!\big)^2} < \pi < \frac{(2n+1)^2(2^n n!)^4}{2n\big((2n+1)!\big)^2} .$$

Крайние члены этих неравенств, очевидно, эквивалентны между собой, поэтому их предел (при $n \to \infty$) равен π. Воспользовавшись полученным выше представлением для $n!$, получим, что $\pi = \gamma^2/2$, т.е. $\gamma = \sqrt{2\pi}$. ⟽

✧ **Доказательство неравенства Берри–Эссеена** (25), стр. 291. Сначала с помощью метода сглаживания докажем неравенство Эссеена.

357| **Лемма.** [*Неравенство Эссеена.*] *Пусть φ_j — хар.ф., отвечающие ф.р. F_j, $j = 1, 2$, причём для $\forall\, x$ существует производная $F_2'(x)$ и*

$C = \sup_x F_2'(x) < \infty$. *Тогда для* $\forall\, T > 0$

$$\sup_x \bigl|F_1(x) - F_2(x)\bigr| \;\leqslant\; \frac{1}{\pi}\int_{-T}^{T}\left|\frac{\varphi_1(t)-\varphi_2(t)}{t}\right| dt + \frac{24C}{\pi T}. \tag{28}$$

⚠ Из приведённого далее доказательства следует, что константа $24/\pi$ в правой части неравенства может быть заменена на 4.45.

⤇ Пусть η_j сл.в. с ф.р. F_j и хар.ф. φ_j, $j = 1,2$. Рассмотрим вспомогательную сл.в. ζ, ф.р. которой G имеет плотность g. Тогда по формуле свёртки ф.р. суммы $\eta_j + \zeta$ (для $x \in \mathbb{R}^1$)

$$\widetilde{F}_j(x) \;=\; \int_{-\infty}^{\infty} F_j(x-y)g(y)\,dy \;=\; \int_{-\infty}^{\infty} F_j(u)g(x-u)\,du$$

после очевидной замены. В свою очередь, ф.р. разности $\eta_j - \zeta$

$$\widehat{F}_j(x) \;=\; \int_{-\infty}^{\infty} F_j(x+y)g(y)\,dy \;=\; \int_{-\infty}^{\infty} F_j(u)g(u-x)\,du.$$

Пусть теперь $b, T > 0$ и $\zeta = \zeta_0 + b/T$, где сл.в. ζ_0 имеет плотность $g_0(x) = \dfrac{1}{\pi T x^2}(1-\cos(Tx))$ с хар.ф. $\psi_0(t) = \dfrac{1}{T}(T - |t|)\,\dot{\mathtt{I}}\,(|t| \leqslant T)$ (см. задачу 167, стр. 167). Тогда хар.ф. ζ равна $\psi(t) = \psi_0(t)\exp(\mathtt{i}tb/T)$, а ф.пл.

$$g(x) = \frac{T}{2\pi}\left(\frac{\sin\bigl((Tx-b)/2\bigr)}{(Tx-b)/2}\right)^2, \qquad x \in \mathbb{R}^1.$$

Хар.ф. разности $\eta_j - \zeta$ равна $\varphi_j(t)\psi(-t)$, причём эта функция, очевидно, интегрируема, следовательно, по формуле обращения 166, стр. 166, и ввиду равенства нулю $\psi(t)$ при $|t| > T$ имеем для $\forall\, x > y$ тождество

$$\widehat{F}_j(x) - \widehat{F}_j(y) \;=\; \frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-\mathtt{i}tx} - e^{-\mathtt{i}ty}}{-\mathtt{i}t}\varphi_j(t)\psi(-t)\,dt.$$

Для доказываемого неравенства интересен только случай, когда $\int_{-T}^{T}|\varphi_1(t) - \varphi_2(t)|/|t|\,dt < \infty$. В этом случае можно записать

$$\begin{aligned}(\widehat{F}_1(x) - \widehat{F}_2(x)) &- (\widehat{F}_1(y) - \widehat{F}_2(y)) \;= \\ &= \frac{1}{2\pi}\int_{-T}^{T}\frac{\varphi_1(t)-\varphi_2(t)}{-\mathtt{i}t}e^{-\mathtt{i}tx}\psi(-t)\,dt - \\ &\qquad - \frac{1}{2\pi}\int_{-T}^{T}\frac{\varphi_1(t)-\varphi_2(t)}{-\mathtt{i}t}e^{-\mathtt{i}ty}\psi(-t)\,dt.\end{aligned}$$

В этом представлении интеграл во втором слагаемом есть преобразование Фурье абсолютно интегрируемой на $\mathbb{R}^1$ функции $\frac{\mathtt{i}}{t}(\varphi_1(t) - \varphi_2(t))\psi(-t)$. По известной теореме Римана (см. стр. 179) это преобразование стремится к нулю при $y \to \pm\infty$. Так как $\widehat{F}_1(-\infty) = \widehat{F}_2(-\infty) = 0$, то получаем, что

$$\Big|\int_{-\infty}^{\infty} \big(F_1(u) - F_2(u)\big) g(u-x)\,du\Big| \;=\; \Big|\frac{1}{2\pi}\int_{-T}^{T} \frac{\varphi_1(t)-\varphi_2(t)}{-\mathtt{i}t}\, e^{-\mathtt{i}tx}\psi(-t)\,dt\Big| \;\leqslant$$

$$(*) \qquad \leqslant\; \frac{1}{2\pi}\int_{-T}^{T} \Big|\frac{\varphi_1(t)-\varphi_2(t)}{t}\Big|\,dt.$$

Определим

$$B \;=\; \mathbf{P}\Big\{\, 0 \leqslant \zeta \leqslant \frac{2b}{T} \,\Big\} \;=\; \int_0^{2b/T} g(y)dy \;=\; \frac{2}{\pi}\int_0^{b/2} \frac{\sin^2 y}{y^2}\,dy \;=\; \mathbf{P}\Big\{\, |\zeta_0| \leqslant \frac{b}{T} \,\Big\}.$$

Поскольку ф.р. F_1 не убывает, то для $\forall\, x$

$$(\sharp) \qquad F_1(x) \;\leqslant \frac{1}{B}\int\limits_x^{x+2b/T} F_1(u)g(u-x)\,du \;\leqslant\; F_2(x) +$$

$$+\frac{1}{B}\int\limits_0^{2b/T} |F_2(x+y) - F_2(x)| g(y)\,dy + \frac{1}{B}\int\limits_x^{x+2b/T} (F_1(u) - F_2(u))g(u-x)\,du.$$

По формуле конечных приращений $|F_2(x+y) - F_2(x)| = F_2'(y_x^*)\,y \leqslant Cy,\ y > 0, x \leqslant y_x^* \leqslant x+y$. Поэтому с учётом определения B

$$\frac{1}{B}\int_0^{2b/T} |F_2(x+y) - F_2(x)| g(y)\,dy \;\leqslant\; \frac{C}{B}\int_0^{2b/T} y\, g(y)\,dy \;=$$

$$=\; \frac{C}{B}\frac{1}{\pi T}\int_{-b/2}^{b/2} (2u+b)\frac{\sin^2 u}{u^2}du \;=\; \frac{Cb}{T}.$$

Далее, полагая $\Delta = \sup_x |F_1(x) - F_2(x)|$, получаем из $(*)$

$$\Big|\int\limits_x^{x+2b/T} (F_1(u) - F_2(u))g(u-x)\,du\Big| \;\leqslant$$

$$\leqslant\; \frac{1}{2\pi}\int\limits_{-T}^{T} \Big|\frac{\varphi_1(t)-\varphi_2(t)}{t}\Big|\,dt + \Delta\int\limits_{-\infty}^{x} g(u-x)\,du + \Delta\int\limits_{x+2b/T}^{\infty} g(u-x)\,du \;=$$

$$=\; \frac{1}{2\pi}\int\limits_{-T}^{T} \Big|\frac{\varphi_1(t)-\varphi_2(t)}{t}\Big|\,dt + \Delta(1-B).$$

Следовательно, ввиду $(\sharp)$ при $\forall\, x \in \mathbb{R}^1$

$$F_1(x) - F_2(x) \;\leqslant\; \frac{1}{2\pi B}\int_{-T}^{T} \Big|\frac{\varphi_1(t)-\varphi_2(t)}{t}\Big|\,dt + \frac{Cb}{T} + \frac{1-B}{B}\Delta.$$

Проведя аналогичные построения с использованием ф.р. $\widetilde{F}_j$, получим, что разность $F_1(x)-F_2(x)$ не меньше правой части последнего выражения, умноженной на -1. Отсюда можно сделать вывод, что при $2B>1$

$$\Delta \leqslant \frac{1}{\pi 2(2B-1)}\int_{-T}^{T}\left|\frac{\varphi_1(t)-\varphi_2(t)}{t}\right|dt+\frac{CbB}{T(2B-1)}.$$

Выбирая $b=3$ и воспользовавшись какой-либо программой вычисления интегрального синуса, получаем $B\approx 0.7546$, т.е. $2(2B-1)>1$ и $bB/(2B-1)\approx$ ≈ 4.446. Следовательно,

$$\Delta \leqslant \frac{1}{\pi}\int_{-T}^{T}\left|\frac{\varphi_1(t)-\varphi_2(t)}{t}\right|dt+\frac{4.45\,C}{T}. \qquad \Lleftarrow$$

Приступим к доказательству неравенства Берри–Эссеена. Пусть φ — хар.ф. ξ_k, $\gamma=\mathbf{E}\,|\xi_k|^3$, $k\geqslant 1$. Покажем, что в области $|u|\leqslant a/\gamma$ с $a={}^3/_4$ хар.ф. $|\varphi(u)|\leqslant \exp(-qu^2)$, где $q=543/2048<{}^1/_2$. По свойству (X_8), стр. 159, $\varphi(u)=1-\frac{u^2}{2}+\frac{u^3}{6}\delta(u)$, где $|\delta(u)|\leqslant\gamma$. Так как $|z|^2=z\overline{z}$, то

$$|\varphi(u)|^2 = \Big(1-\frac{u^2}{2}+\frac{u^3}{6}\delta(u)\Big)\Big(1-\frac{u^2}{2}+\frac{u^3}{6}\overline{\delta(u)}\Big) =$$
$$= 1-u^2+\frac{u^3}{6}(\delta(u)+\overline{\delta(u)})+\frac{u^4}{4}-\frac{u^5}{12}(\delta(u)+\overline{\delta(u)})+\frac{u^6}{36}\delta(u)\overline{\delta(u)}.$$

Заметим, что все слагаемые здесь — действительные числа, поэтому для них допустимы все операции сравнения, в частности можно использовать известное неравенство $1+x\leqslant e^x$. Стало быть, справедливо неравенство

$$|\varphi(u)|^2 \leqslant \exp\Big(-u^2+\frac{|u|^3}{3}\gamma+\frac{u^4}{4}+\frac{|u|^5}{6}\gamma+\frac{u^6}{36}\gamma^2\Big).$$

По неравенству Ляпунова 549, стр. 487, $\gamma=\mathbf{E}\,|\xi_1|^3\geqslant(\mathbf{E}\,\xi_1^2)^{3/2}=1$. Поэтому

$$|\varphi(u)|^2 \leqslant \exp\Big(-u^2+\frac{au^2}{3}+\frac{a^2u^2}{4}+\frac{a^3u^2}{6}+\frac{a^4u^2}{36}\Big) = e^{-2qu^2}$$

при $|u|\leqslant\frac{a}{\gamma}$ $(\leqslant a)$. Полагая $u=\frac{t}{\sqrt{n}}$, получаем $|\varphi(\frac{t}{\sqrt{n}})|\leqslant\exp\big(-\frac{q}{n}t^2\big)$ для $|t|\leqslant\frac{a}{\gamma}\sqrt{n}$. Применим теперь (27), предварительно вынеся e^{-qt^2} за знак модуля (здесь $q<{}^1/_2$ и потому $\exp\big(-t^2/2n+qt^2/n\big)\leqslant 1$):

$$\Big|\varphi^n\Big(\frac{t}{\sqrt{n}}\Big)-\exp\Big(-\frac{t^2}{2n}\Big)^n\Big| \leqslant$$
$$\leqslant n\Big|\varphi\Big(\frac{t}{\sqrt{n}}\Big)\exp(\tfrac{q}{n}t^2)-\exp\Big(-\frac{t^2}{2n}+\frac{q}{n}t^2\Big)\Big|\exp\big(-qt^2\big) \leqslant$$

$$\leqslant\ n\Big|\varphi\Big(\frac{t}{\sqrt{n}}\Big)-\exp\Big(-\frac{t^2}{2n}\Big)\Big|\,e^{-qt^2+qa^2},$$

где в последнем неравенстве снова учтено, что $t^2/n\leqslant a^2$. Так как

$$\varphi\Big(\frac{t}{\sqrt{n}}\Big)-\exp\Big(-\frac{t^2}{2n}\Big)\ =\ 1-\frac{t^2}{2n}+\frac{t^3}{6n\sqrt{n}}\,\delta\Big(\frac{t}{\sqrt{n}}\Big)-1+\frac{t^2}{2n}-\frac{t^4}{8n^2}\,\theta(t),$$

где $|\delta|\leqslant\gamma,|\theta|\leqslant 1$, то при $|t|\leqslant\frac{a}{\gamma}\sqrt{n}\leqslant a\sqrt{n}\,\gamma,\ a={}^3/_4$, для хар.ф. $\varphi_1(t)= =\varphi^n\big(\frac{t}{\sqrt{n}}\big),\ \varphi_2(t)=\exp(-t^2/2)$ имеем

$$\Big|\frac{\varphi_1(t)-\varphi_2(t)}{t}\Big|\ \leqslant\ \frac{1}{\sqrt{n}}\Big(\frac{\gamma}{6}+\frac{|t|}{8\sqrt{n}}\Big)\,t^2e^{-qt^2+qa^2}\ \leqslant$$
$$\leqslant\ \frac{\gamma}{\sqrt{n}}\Big(\frac{1}{6}+\frac{a}{8}\Big)\,t^2e^{-qt^2+qa^2}\ =\ \frac{\gamma}{\sqrt{n}}\frac{25}{96}\,t^2e^{-qt^2+qa^2}.$$

Поскольку интеграл (при $a=3/4,\ q=543/2048$)

$$\int_{-T}^{T}t^2e^{-qt^2+qa^2}\,dt\ \leqslant\ \int_{-\infty}^{\infty}t^2e^{-qt^2+qa^2}\,dt\ =\ e^{qa^2}\sqrt{\pi}/2q\sqrt{q}\ \leqslant\ 7.6,$$

то неравенство (28) (с $T=\frac{3}{4\gamma}\sqrt{n},\ C=\max_x\Phi'(x)=\frac{1}{\sqrt{2\pi}}$ и уточнённой константой 4.45 вместо $24/\pi$) даёт

$$\sup_x\big|F_n(x)-\Phi(x)\big|\ \leqslant\ \frac{25\cdot 7.6}{96\,\pi}\,\frac{\gamma}{\sqrt{n}}+\frac{4\cdot 4.45}{3\sqrt{2\pi}}\,\frac{\gamma}{\sqrt{n}}\ =\ Q\frac{\gamma}{\sqrt{n}}$$

с константой $Q\approx 2.9971<3$.

✧ Неравенство Гёфдинга.

358| **Лемма.** [*Hoeffding.*] *Пусть независимые сл.в.* $\langle\,\xi_k\,\rangle_1^n$ *имеют нулевые математические ожидания и ограничены по модулю константой* X, *т.е.* $\mathbf{P}\big\{\big|\xi_k\big|\leqslant X\big\}=1$. *Тогда для* $S_n=\sum_1^n\xi_k$ *при* $\forall\,t>0$ *имеем:*

$$\mathbf{P}\big\{\,S_n\geqslant t\,\big\}\ \leqslant\ \exp\Big\{-\frac{t^2}{2nX^2}\Big\},\qquad \mathbf{P}\big\{\,\big|S_n\big|\geqslant t\,\big\}\ \leqslant\ 2\exp\Big\{-\frac{t^2}{2nX^2}\Big\}.$$

⚠ Если $\mathbf{P}\big\{\big|\xi_k\big|\leqslant X_k\big\}=1$, то можно положить $nX^2=\sum_1^n X_k^2$.

⤇ Очевидно, достаточно считать $X=1$. Заметим сначала, что в силу выпуклости книзу показательной функции для $\forall\,x\in[-1;1],a\in\mathbb{R}^1$

$$e^{ax}=\exp\Big\{a\frac{(1+x)}{2}-a\frac{(1-x)}{2}\Big\}\leqslant\frac{(1+x)}{2}e^a+\frac{(1-x)}{2}e^{-a}=\cosh(a)+x\sinh(a),$$

где $\cosh, \sinh$ — гиперболические косинус и синус соответственно. Имеем

$$\cosh(a) \;=\; \sum\nolimits_1^\infty \frac{a^{2k}}{(2k)!} \;\leqslant\; \sum\nolimits_1^\infty \frac{a^{2k}}{2^k\, k!} \;=\; \exp\bigl\{\tfrac{1}{2}a^2\bigr\},$$

поскольку, очевидно, $(2k)! \geqslant 2^k\, k!$ при $\forall\, k \geqslant 1$. Следовательно, для $\forall\, a \in \mathbb{R}^1$

$$\mathbf{E}\, e^{a\xi_k} \;\leqslant\; \cosh(a) + \sinh(a)\, \mathbf{E}\, \xi_k \;=\; \cosh(a) \;\leqslant\; \exp\bigl\{\tfrac{1}{2}a^2\bigr\}.$$

Таким образом, для $\forall\, a > 0$ в силу неравенства Маркова (и с учётом независимости сомножителей)

$$\mathbf{P}\bigl\{\, S_n \geqslant t \,\bigr\} \;=\; \mathbf{P}\bigl\{\, e^{aS_n} \geqslant e^{at} \,\bigr\} \;\leqslant\; e^{-at}\, \mathbf{E}\, e^{aS_n} \;=\; e^{-at}\Bigl(\prod_1^n \mathbf{E}\, e^{a\xi_k}\Bigr) \;\leqslant$$
$$\leqslant\; e^{-at} \exp\bigl\{n\tfrac{1}{2}a^2\bigr\} \quad \Bigl(=\; \exp\Bigl\{-\,\frac{t^2}{2n}\Bigr\} \;\text{ при } a = {}^{t}\!/_{n}\Bigr).$$

Применяя полученное неравенство к сл.величинам $-\xi_k$, получаем оценку для вероятности двустороннего неравенства. ⫤

✧ Экспоненциальное неравенство Колмогорова.

359| **Лемма.** *Пусть независимые сл.в.* $\langle\, \xi_k \,\rangle_1^n$ *имеют нулевые математические ожидания. Тогда для* $S_j = \sum_1^j \xi_k,\ j = \overline{1,n}$, *при* $\forall\, t, \varepsilon > 0$*:*

I) $\mathbf{P}\bigl\{\, \max\limits_{1\leqslant j\leqslant n} S_j \geqslant \varepsilon \,\bigr\} \;\leqslant\; e^{-t\varepsilon}\, \mathbf{E}\, e^{tS_n}$;

II) *если* $\mathbf{P}\,\{\xi_k \leqslant d\} \;=\; 1,\ \forall\, k = \overline{1,n}$, *и* $\sigma_n^2 \;=\; \sum_1^n \mathbf{D}\xi_j$, *то*

$$\mathbf{P}\bigl\{\, \max_{1\leqslant j\leqslant n} S_j \geqslant \varepsilon \,\bigr\} \;\leqslant\; \exp\{-t\varepsilon + \sigma_n^2 t^2 g(td)\}, \tag{29}$$

где функция $g(x) \;=\; \dfrac{1}{x^2}(e^x - 1 - x),\ x \in \mathbb{R}^1$, *строго возрастает и выпукла книзу.*

⫢ I) Рассмотрим представление

$$A := \{\max\nolimits_{1\leqslant j<n} S_j \geqslant \varepsilon\} = \biguplus_1^n B_k$$

с $B_k \;=\; \{S_k \;\geqslant\; \varepsilon,\ \max\nolimits_{1\leqslant j<k} S_j \;<\; \varepsilon\}$. Заметим, что по неравенству Йенсена в силу выпуклости книзу функции e^x математическое ожидание $\mathbf{E}\, e^{t\xi_j} \geqslant$ $\geqslant \exp\{\mathbf{E}\, t\xi_j\} = 1$. Поэтому, учитывая независимость сл.в., получаем

$$\mathbf{E}\,\bigl[e^{tS_n}\, \dot{\mathrm{I}}\,(B_k)\bigr] \;=\; \mathbf{E}\,\bigl[e^{t(S_n-S_k)} e^{tS_k}\, \dot{\mathrm{I}}\,(B_k)\bigr] \;=\; \mathbf{E}\,\bigl[e^{t(S_n-S_k)}\bigr]\, \mathbf{E}\,\bigl[e^{tS_k}\, \dot{\mathrm{I}}\,(B_k)\bigr] \;=$$

$$= \Big(\prod_{i=k+1}^{n} \mathbf{E}\, e^{t\xi_i} \Big) \, \mathbf{E} \left[e^{tS_k} \, \dot{\mathrm{I}}\,(B_k) \right] \;\geqslant\; e^{t\varepsilon} \, \mathbf{P}\{ B_k \} \, .$$

Дальнейшие выкладки идентичны доказательству 12, стр. 267:

$$\mathbf{P}\{A\} \;=\; \sum\nolimits_1^n \mathbf{P}\{B_k\} \;\leqslant\; e^{-t\varepsilon}\, \mathbf{E}\Big[e^{tS_n} \, \dot{\mathrm{I}}\, \big(\sum\nolimits_1^n B_k \big) \Big] \;\leqslant\; e^{-t\varepsilon}\, \mathbf{E}\, e^{tS_n}.$$

II) Заявленные свойства функции g следуют из легко проверяемого непосредственным интегрированием тождества $g(x) = \int_0^1 \big(\int_0^y e^{xz} dz \big) dy$. Далее, поскольку $x^2 g(x) + 1 + x = e^x$ и по условию $\mathbf{E}\,\xi_j = 0$, то

$$\mathbf{E}\, e^{t\xi_j} \;=\; 1 + t^2\, \mathbf{E}\left[\xi_j^2 g(t\xi_j) \right] \;\leqslant\; 1 + t^2 g(td)\, \mathbf{E}\,\xi_j^2 \;\leqslant\; \exp\{ t^2 g(td)\, \mathbf{E}\,\xi_j^2 \}$$

ввиду монотонности g и условия ограниченности ξ_j. Таким образом, в силу независимости слагаемых S_n

$$\mathbf{E}\, e^{tS_n} \;\leqslant\; \exp\Big\{ t^2 g(td) \sum\nolimits_1^n \mathbf{D}\xi_j \Big\}.$$

В сочетании с I) это доказывает II). ⫤

► Аппроксимации нормальной ф.р.

360| **Лемма.** *Пусть* $\Phi(x) = \int_{-\infty}^{x} \phi(u)\, du$ *— стандартная нормальная ф.р.,* $\phi(u) = e^{-\frac{1}{2} u^2} / \sqrt{2\pi}$. *Тогда для* $\forall\, x > 0$

$$\frac{1}{x} \Big(1 - \frac{1}{x^2} \Big) \phi(x) \;<\; \frac{x^2}{1 + x^2} \, \frac{1}{x} \phi(x) \;<\; 1 - \Phi(x) \;<\; \frac{1}{x} \phi(x),$$

стало быть,

$$1 - \Phi(x) \asymp \frac{1}{x\sqrt{2\pi}} \, e^{-\frac{1}{2} x^2}, \quad x \to \infty.$$

⫢ Интегрированием по частям легко показать, что

$$\frac{1}{x^2} \big(1 - \Phi(x) \big) \;>\; \int_x^{\infty} \frac{1}{u^2} \phi(u)\, du \;=\; \frac{1}{x} \phi(x) - \big(1 - \Phi(x) \big) \;>\; 0,$$

откуда сразу следуют оба неравенства относительно $1 - \Phi(x)$. ⫤

Изумительно точное приближение значения $1 - \Phi(x)$ дают границы

$$\frac{2}{x + \sqrt{x^2 + 4}} \, \phi(x) \;\leqslant\; 1 - \Phi(x) \;\leqslant\; \frac{2}{x + \sqrt{x^2 + (4 - \Delta)}} \, \phi(x),$$

где первое неравенство справедливо для $\forall x \geqslant 0$, а второе — для $x \geqslant (2 - \Delta)/\sqrt{\Delta}$, $0 < \Delta \leqslant 2$. Так, при $\Delta = 1$ разность между $1 - \Phi(x)$ и правой границей не превосходит 0.003 для $\forall x \geqslant 1$. При $\Delta = 4 - 8/\pi$ ошибка приближения ф.р. $\Phi(x)$ с помощью функции $F(x) = 1 - \dfrac{2\phi(x)}{x + \sqrt{x^2 + 8/\pi}}$, $x \geqslant 0$ ($1 - F(-x)$ при $x \leqslant 0$), не превышает 0.016 для $\forall x \in \mathbb{R}^1$, с правильной асимптотикой на «хвостах» числовой прямой.

$\Rrightarrow$ Легко проверить, что производная функции $G(x) = \big(1 - \Phi(x)\big)/\phi(x)$ удовлетворяет уравнению $G'(x) = xG(x) - 1$ для $\forall x$. Для функции $g_B(x) = 2/\big(x + \sqrt{x^2 + B}\big)$ имеем:

$$g_B'(x) = \Big(-2 - \frac{2x}{\sqrt{x^2 + B}}\Big)\frac{1}{\big(x + \sqrt{x^2 + B}\big)^2} \quad \text{и} \quad x\,g_B(x) - 1 = \frac{-B}{\big(x + \sqrt{x^2 + B}\big)^2}.$$

Очевидно, $g_4'(x) \geqslant xg_4(x) - 1$ при $\forall x \geqslant 0$. Для $B \in [2;4)$ имеет место противоположное неравенство $g_B'(x) \leqslant x\,g_B(x) - 1$ при $\forall x \geqslant (B-2)/\sqrt{4 - B}$. В частности, для $B = 2$ это неравенство выполняется при $\forall x \geqslant 0$.

Предположим, что для некоторого $x > 0$ первое из неравенств не выполняется. Тогда в этой точке имеем $g_4'(x) - G'(x) \geqslant x(g_4(x) - G(x)) > 0$. То есть, если (непрерывная) функция $g_4(x) - G(x) > 0$ в какой-то точке x, то в этой точке она возрастает и, значит, во всех точках справа также будет положительной и будет возрастать. Последнее противоречит тому, что $\lim_{x\to\infty}\big(g_4(x) - G(x)\big) = 0$. Аналогично доказывается второе заявленное неравенство. $\Lleftarrow$

B. Указания к решению задач

✦ 295, стр. 256. В условиях задачи $\mathbf{D}(S_n/n) \leqslant \sum_1^n \mathbf{D}\xi_k/n^2$.

✦ 296, стр. 257, (a). Применить неравенства Маркова $x\,\mathbf{P}\{\zeta \geqslant x\} \leqslant \mathbf{E}\big[\zeta\,\dot{\mathrm{I}}(\zeta \geqslant x)\big]$; далее воспользоваться теоремой Лебега.

✦ 298, стр. 258. Рассмотреть ф.р. $F(x) = 1 - \dfrac{1}{x\ln x}$ при $x > 2$.

✦ 307, стр. 264. Проверить условия теоремы 305, стр. 263, для сл.в. $\xi_{kn} = \xi_k/\sqrt{n}$.

✦ 329, стр. 277. Математическое ожидание $\mathbf{E}\,|\xi| \leqslant 1 + \sum_1^\infty \dfrac{1}{n^2} n^2\,\mathbf{P}\{|\xi| \geqslant n\} < \infty$; рассмотреть ф.р. $F(x) = 1 - x^{-5/2}$, $x \geqslant 1$.

✦ 333, стр. 282. Моменты $\mathbf{E}\,|\xi_k|^m = k^m$. По известным формулам $\sum_1^n k^2 = \dfrac{1}{6}\,n(n+1)(2n+1)$, $\sum_1^n k^3 = \dfrac{1}{4}\,n^2(n+1)^2$; проверить условие Ляпунова для $\mathbf{E}\,|\xi_k|^3$ ($\delta = {}^1/_2$).

✦ 335, стр. 283. Проверить условия теоремы 334, стр. 282.

✦ 355, стр. 297. Пусть $\zeta_c = \zeta\,\dot{\mathrm{I}}(|\zeta| < c)$, тогда $\mathbf{D}\zeta_c = \mathbf{E}\,\zeta_c^2 - \big(\mathbf{E}\,\zeta_c\big)^2 = \mathbf{E}\,\zeta^2 - \mathbf{E}\,(\zeta - \zeta_c)^2 - \big(\mathbf{E}\,\zeta_c\big)^2 \leqslant \mathbf{E}\,\zeta^2 = \mathbf{D}\,\zeta$.

✦ 356, стр. 300. $\sum_1^\infty \mathbf{E}\,|\xi_j| = \mathbf{E}\big[\lim_n \sum_1^n |\xi_j|\big] = \mathbf{E}\big[\sum_1^\infty |\xi_j|\big]$.

VIII Случайные процессы

Случайный процесс (сл.проц.) неформально можно определить как семейство случайных величин ξ_t, индексированных некоторым параметром $t \in \mathbb{T}$. Такого определения вполне достаточно, если нас интересуют вопросы, связанные со значениями этого процесса в конечном числе точек t, — тогда мы можем применить методы анализа распределений случайных векторов. Однако чаще всего при изучении того или иного процесса интерес представляют ответы на вопросы о значениях ξ_t на множестве индексов $t \in T_0$, которое может быть даже несчётным. В этом случае неформального определения уже может оказаться мало. Здесь мы вынуждены считаться с тем, что семейство сл.в. представляет собой функцию по параметру $t \in \mathbb{T}$.

361| Пример. Для того чтобы очертить круг проблем, возникающих в связи со сл.процессами, рассмотрим задачу о разорении одного из игроков в последовательности игр, каждая из которых приводит либо к пополнению капитала игрока на одну условную единицу (ставку), либо к уменьшению его капитала на ту же единицу. Путь a — капитал (количество у.е., ставок) интересующего нас игрока A перед началом серии игр, b — капитал его противника. Предположим, что каждая игра с вероятностью p заканчивается победой A; $q = (1 - p)$ — вероятность поражения A в одной игре. Вся серия игр заканчивается естественным образом, если после какой-то из игр капитал одного из игроков превысит суммарный начальный капитал $a + b$. Пусть

ω_k — исход k-й игры ($\omega_k = \pm 1$), тогда доход ξ_t игрока A после t игр

$$\xi_t = \sum\nolimits_1^t \omega_k.$$

Распределение ξ_t, если t зафиксировано, можно описать простой схемой Бернулли, т.е. рассмотреть пространство элементарных исходов $\Omega_t = \{-1, +1\}^t$, состоящее из t-мерных векторов вида $(\pm 1, \ldots, \pm 1)$, и положить вероятность каждого элементарного исхода $\omega = (\omega_1, \ldots, \omega_t)$ равной $p^z q^{t-z}$, где $z = z(\omega) = \frac{1}{2}(t + \omega_1 + \ldots + \omega_t)$ — количество игр, закончившихся победой A. Таким образом, $\xi_t = \xi_t(\omega)$ есть сл.в. на Ω_t, и для всех $m \in [-t; t]$ таких, что $z = \frac{1}{2}(m + t)$ целое число, вероятность $\mathbf{P}\{\xi_t = m\} = \mathrm{C}_t^z p^z q^{t-z}$. Понятно, что серия игр может продолжаться бесконечно. Можно описать весь эксперимент как набор всевозможных векторов конечной или бесконечной размерности. Однако проще считать, что в любом случае эксперимент продолжается бесконечно, без изменения капитала игроков после момента первого достижения накопленным капиталом одного из них нулевого уровня.

Итак, пространство элементарных исходов состоит из всевозможных двоичных последовательностей: $\Omega_\infty = \{\omega : \ \omega = (\omega_1, \omega_2, \ldots)\}$, где $\omega_k = \pm 1$, $k \geqslant 1$. Набор измеримых подмножеств этого пространства и их вероятности необходимо задать так, чтобы вероятности любого конечномерного события, т.е. события, описываемого исходами конечного числа экспериментов, совпадали с вероятностями, вычисляемыми по формулам для конечной схемы Бернулли. Например, вероятность того, что в первой игре победит игрок A, а во второй B, в пространстве Ω_2 вычисляется как $p(1-p)$. Заметим, что любое t-мерное событие может рассматриваться как $(t+1)$-мерное, $(t+2)$-мерное, ... и даже как бесконечномерное событие — достаточно в описании этого события разрешить исходам, начиная с $(t+1)$-го эксперимента, принимать любые возможные значения. При этом вероятность такого события, вычисляемая в биномиальной схеме пространства Ω_t, должна совпадать с вероятностями этого же события в более широких пространствах (такое свойство называется согласованностью). Например, в пространстве Ω_3 вероятность рассмотренного ранее события, т.е. события, что в первой игре победит игрок A, во второй — B, а в третьей игре — либо A, либо B, равна $p(1-p)p + p(1-p)(1-p) = p(1-p)$. Очевидно, что во всех более широких пространствах эта вероятность останется неизменной. Кроме того, вероятность не изменится, если переставить последовательность игр в формулировке события: во второй игре победит B, а в первой А или в третьей игре победит либо А, либо В, во второй В, а в первой А (свойство перестановочности).

Конечномерные подмножества, вложенные в бесконечномерное пространство описанным выше способом, называются цилиндрическими множествами (сравните с трёхмерным цилиндром, в котором две координаты должны принадлежать некоторому кругу, а третья координата может принимать произвольные значения). Минимальная σ-алгебра в Ω_∞, содержащая все конечномерные цилиндры, называется цилиндрической. По теореме Колмогорова 496, стр. 451, на такой σ-алгебре существует вероятностная мера, совпадающая на конечномерных событиях с биномиальными вероятностями. Только после всех этих построений мы можем поставить вопрос о вероятности разорения игрока А, поскольку события подобного рода принадлежат бесконечномерному пространству. Способ отыскания этой вероятности приведён в дополнении к главе; см. также § 5, посвящённый мартингалам. ◎

§ 1. Случайные элементы в пространстве функций

✧ **Пространство функций и цилиндрическая σ-алгебра.** Не стремясь к широкому обобщению, рассмотрим только случай, когда $t \in \mathbb{T} \subset \mathbb{R}^1$. По сложившейся традиции параметр t интерпретируют как «время» (момент времени), прошедшее от некоторого начала отсчёта $t = t_0$ — чаще всего $t_0 = 0$. Пространство $\mathscr{X}$ возможных значений сл.величин ξ_t называют *фазовым пространством* (здесь также ограничимся случаем $\mathscr{X} \subset \mathbb{R}^1$ или, в специальном случае, $\mathscr{X}$, совпадающим с комплексной плоскостью $\mathbb{C}$ с борелевской σ-алгеброй). Пусть $\boldsymbol{\xi} = \left\langle \xi_t, t \in \mathbb{T} \right\rangle$ — семейство сл.в. на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$. При каждом фиксированном ω это семейство представляет собой функцию $\xi_t(\omega)$ параметра $t \in \mathbb{T}$, которую называют выборочной функцией (траекторией, реализацией) $\boldsymbol{\xi}$. Другими словами, $\boldsymbol{\xi}$ есть отображение Ω в пространство $\mathscr{X}^{\mathbb{T}}$ всех функций $X = X(t)$, заданных на $\mathbb{T}$ и принимающих значения в $\mathscr{X}$. Фиксированная реализация вероятностного эксперимента в виде элементарного исхода ω полностью определяет всю траекторию $\xi_t(\omega), t \in \mathbb{T}$. Как мы видели в предыдущем примере, элементарный исход ω может быть вовсе даже не «элементарным» и совпадать, по существу, с совокупностью всех реализаций процесса. Измеримость такого отображения предполагает введение на $\mathscr{X}^{\mathbb{T}}$ некоторой σ-алгебры подмножеств.

В главе IX описана схема построения так называемой цилиндрической σ-алгебры $\mathscr{B}^{\mathbb{T}}$, в основе которой лежит система всех конечномерных цилиндров (цилиндрических множеств) с борелевскими основаниями:

$$B \times \mathscr{X}^{\mathbb{T}\setminus\tau} := \{X \in \mathscr{X}^{\mathbb{T}} : (X(t_1), \ldots, X(t_k)) \in B\}, \quad B \in \mathbb{R}^k,$$

где набор параметров $\tau = (t_1, \ldots, t_k) \subset \mathbb{T}$, т.е. τ состоит из k различных элементов $\mathbb{T}$ с фиксированным порядком записи — будем обозначать совокупность всех таких наборов $\mathbb{T}^{*k}$. Таким образом, цилиндр $B \times \mathscr{X}^{\mathbb{T}\setminus\tau}$ с основанием в виде борелевского прямоугольника $B = B_1 \times \cdots \times B_k \in \mathscr{B}^k$ состоит из всех тех функций (траекторий), которые в заданный момент времени $t_j \in \tau$ проходят через «ворота» B_j, $j = 1, \ldots, k$. *Цилиндрической σ-алгеброй* называется минимальная σ-алгебра $\mathscr{B}^{\mathbb{T}} = \sigma\big(\mathscr{B}^k \times \mathscr{X}^{\mathbb{T}\setminus\tau}, \tau \in \mathbb{T}^{*k}, k \geqslant 1\big)$, содержащая все конечномерные цилиндры.

Поскольку конечномерный вектор функций измерим по Борелю т. т. т. когда измеримы компоненты этого вектора, то с учётом 507, стр. 464, имеет место

362| **Лемма.** (?!) *Семейство $\boldsymbol{\xi} = \big\langle \xi_t, t \in \mathbb{T} \big\rangle$ измеримых отображений пространства $(\Omega, \mathscr{F})$ на борелевскую прямую $(\mathbb{R}^1, \mathscr{B})$ измеримо как отображение $\boldsymbol{\xi} : (\Omega, \mathscr{F}) \mapsto (\mathscr{X}^{\mathbb{T}}, \mathscr{B}^{\mathbb{T}})$.*

Это обстоятельство позволяет определять процесс как набор случайных величин, зависящих от параметра $t \in \mathbb{T}$.

Необходимость рассмотрения именно цилиндрической σ-алгебры обусловлена тем, что единственно доступный нам способ построения вероятностной меры на подмножествах $\mathbb{R}^{\mathbb{T}}$ с бесконечным параметрическим множеством $\mathbb{T}$ предполагает задание, например с помощью функции распределения, этой вероятности на конечномерных подмножествах и продолжение её на бесконечномерные подмножества с помощью обращения к теореме Каратеодори. При счётном $\mathbb{T}$ цилиндрическая σ-алгебра совпадает с σ-алгеброй, порождённой открытыми множествами в топологии Тихонова. Это свойство позволяет уделять пониженное внимание вопросам измеримости отображений, заданных на $\mathscr{X}^{\mathbb{T}}$, поскольку непрерывные отображения становятся автоматически измеримыми. При несчётном множестве $\mathbb{T}$ цилиндрическая σ-алгебра совпадает с σ-алгеброй, порождённой только базой тихоновской топологии. В этом случае многие утверждения о поведении сл.функции $\boldsymbol{\xi}$ не будут измеримы. Например, невозможно ответить на вопрос о вероятности разорения —

$\inf_t \xi_t < 0$, для ξ_t, равного капиталу финансовой компании к моменту времени t, если не предполагать выполнения каких-либо дополнительных свойств у процесса. Например, если известно, что значения ξ_t меняются в фиксированные моменты времени $t_1, t_2, \ldots$, или если по каким-либо соображениям можно утверждать, что функция ξ_t непрерывна, то указанное событие становится измеримым. Поэтому обычно придерживаются определения случайного процесса как набора сл.величин с последующим описанием его свойств, позволяющих вычислять вероятности большинства интересных событий. К таким свойствам относятся как свойства параметрического пространства $\mathbb{T}$ (например, процесс с дискретным временем), так и свойства его траекторий (непрерывность или без разрывов второго рода) или свойства его вероятностного распределения.

Пусть $\boldsymbol{\xi} = \left\langle \xi_t,\, t \in \mathbb{T} \right\rangle$ — семейство сл.в. на одном вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$, т.е. семейство измеримых отображений пространства $(\Omega, \mathscr{F})$ в борелевскую прямую $(\mathbb{R}^1, \mathscr{B})$. Таким образом, $\boldsymbol{\xi}$ есть функция двух аргументов $\xi(t, \omega)$, $\omega \in \Omega$, $t \in \mathbb{T}$. Поскольку нас будут интересовать вопросы поведения $\boldsymbol{\xi}$ по параметру t, мы чаще всего не будем писать аргумент ω и, более того, зависимость от t будем указывать в виде $\xi(t)$. Кстати, иногда термин «процесс» применяют только к семейству $\boldsymbol{\xi}$, в котором параметр t можно интерпретировать как время, в противном случае говорят о случайной функции. Напомним, что сечением функции двух аргументов называется та же функция, рассматриваемая только от одного из аргументов при фиксированном другом. В теории сл.проц. это наименование сохранилось для сечения по времени (т.е. при фиксированном t); при фиксированном конечном наборе $(t_1, \ldots, t_n)$ значений параметра t говорят о конечномерном сечении $(\xi(t_1), \ldots, \xi(t_n))$, представляющем собой конечномерный сл.вектор.

► Обсудим вопрос измеримости $\boldsymbol{\xi}$ как функции двух аргументов (ω, t) со значениями в фазовом пространстве $\mathscr{X}$. Необходимость в этом возникает, например, в связи с применением формулы Фубини в равенстве вида $\mathbf{E}\left[\int \xi(t)\, dt\right] = \int \mathbf{E}\left[\xi(t)\right] dt$, которое, кроме всего прочего, предполагает совместную измеримость функции $\xi(t, \omega)$.

Определение. Пусть $\mathbb{T}$ — интервал числовой прямой $\mathbb{R}^1$ с борелевской σ-алгеброй $\mathscr{B}$, $(\Omega, \mathscr{F})$ — измеримое пространство. Сл.проц. $\left\langle \xi_t,\ t \in \mathbb{T} \right\rangle$ на $(\Omega, \mathscr{F})$ называется *измеримым*, если отображение $\xi_t(\omega) : \Omega \times \mathbb{T} \mapsto \mathbb{R}^1$ измеримо по Борелю относительно прямого произведения σ-алгебр $\mathscr{F} \otimes \mathscr{B}$.

Для неизмеримых сл.процессов следующий факт может не выполняться.

363| **Лемма.** (?!) *Пусть сл.проц.* $\left\langle \xi_t(\omega),\ t \in (\mathbb{T}, \mathscr{B}),\ \omega \in (\Omega, \mathscr{F}) \right\rangle$ *измерим и* $\vartheta : (\Omega, \mathscr{F}) \mapsto (\mathbb{T}, \mathscr{B})$ *— некоторая сл.в. Тогда* ξ_ϑ *— случайная величина.*

Самый надёжный, но редко осуществимый способ конструирования процесса с заданными свойствами, в частности измеримого процесса, состоит в прямом описании его как функции параметра t и некоторых случайных компонент.

Определение. Пусть $h(t, x_1, \ldots, x_k)$ — измеримая действительная функция на $\mathbb{T} \times \mathbb{R}^k$, $(\beta_1, \ldots, \beta_k)$ — сл.вектор на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$. Сл.проц. $\xi(t) = h(t; \beta_1, \ldots, \beta_k)$ называется *квазидетерминированным* (или *элементарным*).

Поведение и свойства траекторий элементарного процесса полностью определяются единственной реализацией сл.вектора $\vec{\beta}$, задающей параметры зависимости процесса от времени.

364| Примеры. 1) Некий кредитор раздал все свои средства в равных долях трём заёмщикам. По условиям договоров (которые обязательно будут выполнены) возврат средств должен произойти в любое время от 0 до 13 часов первого апреля. Наличные средства кредитора в этот промежуток можно описать квазидетерминированным процессом с функцией $h(t; x_1, x_2, x_3) = \dot{\mathrm{I}}(x_1 \leqslant t) + \dot{\mathrm{I}}(x_2 \leqslant t) + \dot{\mathrm{I}}(x_3 \leqslant t)$ и вектором сл.в. $(\beta_1, \beta_2, \beta_3)$, соответствующих моментам возврата средств заёмщиками. Траектории этого процесса представляют собой непрерывную справа ступенчатую функцию с единичными скачками в случайных точках $\beta_{(1)} \leqslant \beta_{(2)} \leqslant \beta_{(3)}$.

2) Пусть $h(t; x_1, x_2) = x_1^2 - 2x_2 t + t^2$, $t \in [0; \infty)$, и независимые сл.в. $\beta_1, \beta_2 \backsim \mathfrak{Un}(0,1)$, т.е. распределены равномерно на $[0; 1]$. Траектория процесса $\xi(t) = h(t; \beta_1, \beta_2)$ представляет собой параболу на правой полуплоскости. Требуется найти вероятность того, что хотя бы в одной точке траектория процесса пересечёт ось абсцисс: $\mathbf{P}\{\exists t \geqslant 0 : \xi(t) = 0\}$. Так как ветви параболы направлены вверх и $\xi(0) = \beta_1^2 \geqslant 0$, то искомым требованиям удовлетворяют все траектории, у которых минимальное значение меньше нуля и достигается при $t > 0$: $\{\min_{t \geqslant 0} \xi(t) = \beta_1^2 - \beta_2^2 < 0\} \cap \{\beta_2 > 0\}$. В условиях примера вероятность этого равна $1/2$. Заметим, что для сл.проц. общего вида множество $\{\omega : \exists t \geqslant 0 : \xi_t(\omega) = 0\}$ может быть неизмеримым. ⊚

✧ Распределение случайного процесса. Если способ прямого задания процесса $\boldsymbol{\xi} = \left\langle \xi_t, t \in \mathbb{T} \right\rangle$ как функции (ω, t) недоступен, то так же, как и при рассмотрении обычных сл.в. (векторов), мы можем попытаться описать процесс с помощью его распределения в пространстве значений $\mathscr{X}^{\mathbb{T}}$. Во-первых, заметим, что процесс $\boldsymbol{\xi} : (\Omega, \mathscr{F}, \mathbf{P}) \mapsto \mathscr{X}^{\mathbb{T}}$, заданный на вероятностном пространстве, порождает распределение $\mathrm{P}_{\boldsymbol{\xi}}$ на $(\mathscr{X}^{\mathbb{T}}, \mathscr{B}^{\mathbb{T}})$ по обычной схеме: $\mathrm{P}_{\boldsymbol{\xi}}\{B\} = \mathbf{P}\big\{ \omega : \xi_t(\omega) \in B \big\}$, $B \in \mathscr{B}^{\mathbb{T}}$. Единственно возможный способ понять структуру этого распределения — описать его через конечномерные цилиндры, причём основания этих цилиндров достаточно выбирать в виде борелевских прямоугольников: $\forall\, B^{(k)} = B_1 \times \cdots \times B_k \in \mathscr{B}^k$, $\tau = (t_1, \ldots, t_k)$,

$$\mathrm{P}_{\boldsymbol{\xi}}\{B^{(k)} \times \mathscr{X}^{\mathbb{T} \setminus \tau}\} = \mathbf{P}\Big\{ \bigcap\nolimits_1^n \big(\xi(t_k) \in B_k\big) \Big\}. \tag{1}$$

Поскольку класс таких цилиндрических множеств замкнут относительно операции пересечения, то согласно лемме 13, стр. 35, любая другая вероятностная мера на $(\mathscr{X}^{\mathbb{T}}, \mathscr{B}^{\mathbb{T}})$, совпадающая с мерой $\mathrm{P}_{\boldsymbol{\xi}}$ на всех цилиндрических множествах, будет совпадать с ней всюду на $\mathscr{B}^{\mathbb{T}}$.

Ясно, что для каждого фиксированного $\tau = (t_1, \ldots, t_k)$ правило $\mathrm{P}\{B^{(k)}; \tau\} := \mathrm{P}_{\boldsymbol{\xi}}\left\{B^{(k)} \times \mathscr{X}^{\mathbb{T} \setminus \tau}\right\}$, $B^{(k)} = B_1 \times \cdots \times B_k \in \mathscr{B}^k$, задаёт вероятностную меру на борелевской σ-алгебре $\mathscr{X}^k$ (сначала для прямоугольников $B^{(k)}$, затем по теореме Каратеодори для $\forall\, B \in \mathscr{B}^k$). Осталось решить обратный вопрос о минимальном наборе требований на семейство конечномерных распределений $\mathcal{P} = \left\langle \mathrm{P}\{B; \tau\},\ B \in \mathscr{B}^{\tau}, \tau \subset \mathbb{T} \right\rangle$, при которых существует продолжение этого семейства на $(\mathscr{X}^{\mathbb{T}}, \mathscr{B}^{\mathbb{T}})$. Теорема Колмогорова 496, стр. 451, 503, стр. 459, утверждает, что достаточно только двух естественных требований: симметрии (перестановочности) (9), стр. 458, и согласованности (10), стр. 459, причём для счётных пространств $\mathbb{T}$ с фиксированным порядком параметров достаточно только свойства согласованности.

365 **Теорема.** [*А. Н. Колмогоров.*] *Для перестановочного и согласованного семейства конечномерных распределений*

$$\mathcal{P} = \left\langle \mathrm{P}\{B; \tau\},\ B \in \mathscr{B}^{\tau},\ \tau \subset \mathbb{T} \right\rangle$$

найдётся единственное вероятностное распределение $\widetilde{\mathrm{P}}$ *на цилиндрической* σ*-алгебре* $\mathscr{B}^{\mathbb{T}}$ *пространства функций* $\mathbb{R}^{\mathbb{T}}$, *для которого семейство конечномерных распределений совпадает с* $\mathcal{P}$, *и сл.проц.* $\boldsymbol{\xi}$ *на некотором вероятностном пространстве такой, что его распределение* $\mathrm{P}_{\boldsymbol{\xi}} = \widetilde{\mathrm{P}}$.

$\Rrightarrow$ Для доказательства последней части теоремы рассмотрим $\Omega = \mathscr{X}^{\mathbb{T}}$ с σ-алгеброй $\mathscr{F} = \mathscr{B}^{\mathbb{T}}$ и вероятностной мерой $\mathbf{P} = \widetilde{\mathrm{P}}$. В этом вероятностном пространстве элементарный исход ω есть некоторая функция $X(t) \in \mathscr{X}^{\mathbb{T}}$. Очевидно, процесс $\boldsymbol{\xi}(\omega) = \omega = X(t)$, с соответствующей функцией $X(t)$, имеет распределение $\mathrm{P}_{\boldsymbol{\xi}} = \widetilde{\mathrm{P}}$. $\Lleftarrow$

⚠ Если семейство $\mathcal{P}$ конечномерных распределений задаётся только для упорядоченных наборов $\tau: \ t_1 < t_2 < \ldots < t_k$, то на все остальные наборы это семейство распространяется с помощью соотношения, описывающего условие перестановочности. Однако в этом случае условие согласованности должно быть проверено в расширенном варианте (11), стр. 459.

В конечномерном пространстве распределение может быть задано с помощью ф.р. или ф.плотности, если распределение абсолютно непрерывно. Пусть

$$\mathcal{F} \ = \ \left\langle \, F_\tau(\vec{x}) = F(\vec{x}\,;\tau), \ \vec{x} \in \mathbb{R}^k, \ \tau = (t_1 < \ldots < t_k) \subset \mathbb{T} \, \right\rangle \qquad (2)$$

— некоторое семейство функций, каждая из которых при заданном упорядоченном наборе $\tau = (t_1 < \ldots < t_k)$ удовлетворяет всем свойствам ф.р. в $\mathbb{R}^k$. Если τ — произвольный набор параметров, не обязательно расположенных по возрастанию, то для такого набора положим $F(\vec{x}\,;\tau) = F(\pi\,\vec{x}\,;\pi\,\tau)$, где $\pi = (j_1, \ldots, j_k)$ — перестановка, при которой элементы $\pi\,\tau$ расположены по возрастанию: $t_{j_1} < \ldots < t_{j_k}$.

Для любого набора из k элементов $\alpha = (a_1, \ldots, a_k)$ определим $\alpha_{(j,\infty)}$ как набор α, в котором $a_j = +\infty$; набор $\alpha^{(j)}$ из $(k-1)$ элементов как набор α, в котором отсутствует j-й элемент $a_j, j = \overline{1,k}$. Условие согласованности означает, что если на каком-то месте $1 \leqslant j \leqslant k$ вектора $\vec{x}$ подставить $x_j = +\infty$ (перейти к соответствующему пределу), то получившаяся маргинальная ф.р. будет совпадать с ф.р. из семейства $\mathcal{F}$ с набором параметров $\tau^{(j)}$:

$$F(\vec{x}_{(j,\infty)}\,;\tau) \ = \ F(\vec{x}^{(j)}\,;\tau^{(j)}), \quad \forall\,\tau = (t_1 < \ldots < t_k), \ 1 \leqslant j \leqslant k, \ \vec{x} \in \mathbb{R}^k, \ k \geqslant 1.$$

Например, для трёхмерной ф.р.:

$$\begin{aligned} F(x_1, x_2, +\infty\,; t_1, t_2, t_3) \ &= \ F(x_1, x_2\,; t_1, t_2), \\ F(x_1, +\infty, x_3\,; t_1, t_2, t_3) \ &= \ F(x_1, x_3\,; t_1, t_3) \ \text{ и т.п.} \end{aligned}$$

Теорему Колмогорова можно сформулировать в терминах ф.р.:

семейство конечномерных ф.р. (2) *задаёт распределение некоторого сл.проц. в* $(\mathscr{X}^{\mathbb{T}}, \mathscr{B}^{\mathbb{T}})$ *т. т. т. когда оно перестановочно и согласовано.*

366 Упр. Любое конечномерное распределение однозначно описывается соответствующей характеристической функцией:

$$\varphi(\vec{u}\,;\tau) = \mathbf{E}\left[\,\exp\{\mathtt{i}\,(u_1\xi(t_1) + \ldots + u_n\xi(t_k))\}\right], \qquad \tau \subset \mathbb{T}\,, \vec{u} \in \mathbb{R}^k.$$

Сформулируйте условия симметрии и согласованности в терминах хар.ф.

✧ **Процессы со свойством независимости сечений.** Предположим, что сечения процесса независимы в совокупности и сл.в. $\xi(t) \backsim F_t$ с некоторой (одномерной) ф.р., $t \in \mathbb{T}$. Определим n-мерную ф.р.

$$F(\vec{x}\,;\tau) = \prod\nolimits_1^k F_{t_j}(x_j), \qquad \tau = (t_1, \ldots, t_k) \subset \mathbb{T}\,,\ \vec{x} \in \mathbb{R}^k.$$

Очевидно, такое семейство удовлетворяет условиям симметрии и согласованности. Следовательно, по теореме Колмогорова процесс с независимыми сечениями существует. Особенно важен случай, когда все одномерные распределения совпадают.

Определение. Случайный процесс, у которого все сечения независимы в совокупности и одинаково распределены, называется *чисто случайной функцией*. Если параметрическое множество дискретно, например $\mathbb{T} = \mathbb{N}_0$, и все сечения имеют нулевое среднее и ненулевую конечную дисперсию, то такая чисто сл.функция называется *белым шумом*.

Название «шум» пришло из радиофизики, когда полезный сигнал улавливается прибором не в чистом виде, а при наличии некоторой «зашумлённости», причём считается, что эту зашумлённость можно объяснить присутствием постоянно действующих случайных, но неизменных факторов. «Белый» он потому, что его спектр (см. далее) подобен спектру белого цвета.

Для процессов с непрерывным временем ($\mathbb{T} \subset \mathbb{R}^1$) название «белый шум» не используется. Считается, что он физически неосуществим. Однако как математическую абстракцию его можно вводить в модели, описывающие вполне реальные явления.

Рассмотрим процесс $\xi_t = A\cos(Lt + \phi) + \varkappa_t,\ t \in [0;1]$, в котором квазидетерминированная составляющая со случайной фазой колебания $\phi \backsim \mathfrak{Un}(0, 2\pi)$ наблюдается с чисто случайной помехой $\varkappa_t$; здесь L — нормирующая константа, A — амплитуда, возможно также случайные. Картинка траектории процесса существенно зависит от степени дискретизации моментов времени, в которых мы фиксируем его значения. На рис. 5 приведены графики одной

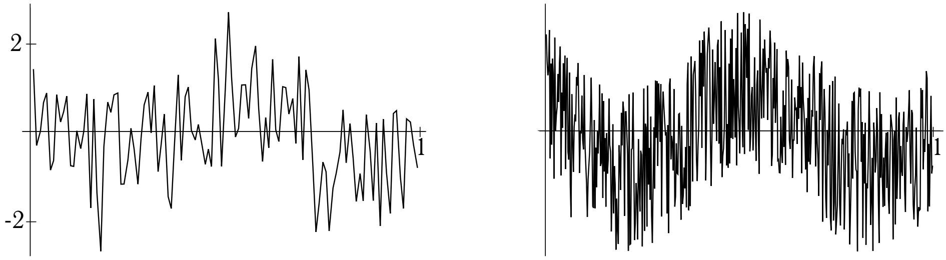


Рис. 5. Траектории процесса $\xi_t = \cos(10t + \frac{1}{3}\pi) + \varepsilon_t$

и той же траектории, зафиксированной при разных разбиениях временно́го отрезка $\mathbb{T} = [0; 1]$. Глядя на первый из графиков, можно подумать, что траектории этого процесса непрерывны. Второй рисунок нас в этом абсолютно разубеждает.

367| **Лемма.** (?!) *Пусть* $\left\langle \varkappa_t,\ t \in \mathbb{T} = [a; b] \right\rangle$ *— чисто случайная функция, тогда для* $\forall t$

$$\mathbf{P}\text{-}\lim_{u \to t} |\varkappa_u - \varkappa_t| \neq 0.$$

368| Упр. Как бы выглядела траектория процесса $\xi_t = \cos(2\pi t) + \varkappa_t,\ t \in$ $\in [0; 1]$, если бы нам удалось зафиксировать его значения при всех t в случае, когда сечения чисто случайной функции $\varkappa_t \backsim \mathcal{U}\mathrm{n}(-1, 1)$?

369| Пример. Формулировка многих предельных теорем теории вероятностей начинается с фразы типа: пусть $\xi_1, \xi_2, \ldots$ — последовательность независимых в совокупности сл.в. Только после установления возможности осуществления такой последовательности эта фраза становится осмысленной. Предыдущие построения показывают, что такая последовательность всегда существует, правда, пространство Ω, на котором задана эта последовательность, может иметь сложную структуру $\mathbb{R}^{\mathbb{N}}$. ◎

► Процессы с независимыми приращениями образуют ещё один важный класс сл.проц.

Определение. Сл.проц. $\left\langle \xi(t),\ t \in \mathbb{T} \subset \mathbb{R}^1 \right\rangle$ имеет *независимые приращения*, если для $\forall n \geqslant 2$ и $\forall t_1 < \ldots < t_n$ $(\in \mathbb{T})$ сл.в. $\xi(t_1), \xi(t_2) -$ $- \xi(t_1), \ldots, \xi(t_n) - \xi(t_{n-1})$ независимы в совокупности.

Рассмотрим случай, когда $\mathbb{T} \subset \mathbb{R}^1_+$ (для удобства можно считать, что $0 \in \mathbb{T}$). Идея способа описания распределения такого процесса опирается на возможность представления любого набора чисел $x_0, x_1, \ldots$ в виде суммы разностей нескольких соседних чисел: $x_0, (x_1 - x_0) + x_0, (x_2 - x_1) + (x_1 - x_0) + x_0, \ldots$ Естественно, там, где имеются суммы независимых сл.в., прямой путь к изучению распределений лежит через их хар.ф.

Проиллюстрируем эту идею на примере четырёх сл.в. $\xi_0, \xi_1, \xi_2, \xi_3$, обладающих свойством независимости приращений — для сокращения записи здесь выбрано $t_i = i$, $i = 0, 1, 2, 3$. Пусть φ_0 — хар.ф. ξ_0, φ_{ij} — хар.ф. разности $(\xi_j - \xi_i)$, $i, j = 0, 1, 2, 3$, тогда с учётом условия независимости приращений хар.ф. вектора (ξ_1, ξ_2, ξ_3) можно записать в виде

$$\begin{aligned}\varphi_{(\xi_1,\xi_2,\xi_3)}(u_1, u_2, u_3) &= \mathbf{E}\exp\bigl\{\mathtt{i}(u_1\xi_1 + u_2\xi_2 + u_3\xi_3)\bigr\} = \\ &= \mathbf{E}\exp\bigl\{\mathtt{i}\bigl(u_3(\xi_3 - \xi_2) + (u_2 + u_3)(\xi_2 - \xi_1) + \\ &\qquad + (u_1 + u_2 + u_3)(\xi_1 - \xi_0) + (u_1 + u_2 + u_3)\xi_0\bigr)\bigr\} = \\ &= \varphi_0(u_1 + u_2 + u_3)\varphi_{01}(u_1 + u_2 + u_3)\varphi_{12}(u_2 + u_3)\varphi_{23}(u_3). \qquad (3)\end{aligned}$$

Заметим, что (3) определяет некоторую хар.ф., т.к. может рассматриваться как хар.ф. линейного преобразования $(\eta_2 + \eta_1, \eta_3 + \eta_2 + \eta_1, \eta_4 + \eta_3 + \eta_2 + \eta_1)$ вектора из независимых сл.в. $\vec{\eta} = (\xi_0, \xi_1 - \xi_0, \xi_2 - \xi_1, \xi_3 - \xi_2)$. Как известно, хар.ф. маргинального распределения, скажем (ξ_1, ξ_3), получается отсюда присвоением нулевого значения аргументу, соответствующему исключаемой сл.в. Эта хар.ф. должна совпадать с хар.ф., вычисленной непосредственно из распределения (ξ_1, ξ_3). Таким образом,

$$\begin{aligned}\varphi_{(\xi_1,\xi_3)}(u_1, u_3) &= \varphi_0(u_1 + u_3)\varphi_{01}(u_1 + u_3)\varphi_{13}(u_3) = \\ = \varphi_{(\xi_1,\xi_2,\xi_3)}(u_1, 0, u_3) &= \varphi_0(u_1 + u_3)\varphi_{01}(u_1 + u_3)\varphi_{12}(u_3)\varphi_{23}(u_3),\end{aligned}$$

для чего достаточно, чтобы $\varphi_{12}(u)\varphi_{23}(u) = \varphi_{13}(u)$ для любых u.

370 **Теорема.** *Пусть* $\langle \varphi_{st},\ s, t \geqslant 0 \rangle$ *— семейство хар.ф., удовлетворяющее условию*

$$\varphi_{rs}(u)\varphi_{st}(u) = \varphi_{rt}(u), \qquad \forall\, 0 \leqslant r < s < t,\ u \in \mathbb{R}^1.$$

Тогда существует сл.проц. $\langle \xi(t),\ t \geqslant 0 \rangle$ *с независимыми приращениями, для которого* φ_{st} *— хар.ф.* $\xi(t) - \xi(s)$, $0 \leqslant s < t$.

⊨ Пусть φ_0 — произвольная хар.ф. Определим для каждого набора чисел $0 = t_0 \leqslant t_1 < \ldots < t_n$, $n \geqslant 1$, n-мерную функцию

$$\varphi(\vec{u} \mid t_1, \ldots, t_n) = \varphi_0\bigl(u_1 + \ldots + u_n\bigr) \prod_{j=1}^{n} \varphi_{t_{j-1}t_j}\bigl(u_j + \ldots + u_n\bigr) \qquad (4)$$

с вектором $\vec{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$. Как было отмечено выше, эта функция есть хар.ф. некоторого n-мерного вероятностного распределения, а всё семейство (по $t_1 < \ldots < t_n$, $n \geqslant 1$) удовлетворяет условию согласованности. В силу теоремы Колмогорова это семейство описывает конечномерные распределения некоторого сл.проц. Очевидно, φ_0 здесь есть хар.ф. начальной сл.в. ξ_0. ⫤

371| Упр. Докажите независимость приращений семейства сл.в., распределение которых определяется хар.ф. (4).

Условие независимости приращений обеспечивает независимость приращения за каждый фиксированный промежуток времени от всей траектории процесса до начала этого промежутка.

372| **Лемма.** *Пусть $\mathscr{F}_s = \sigma\big(\xi(t), t \leqslant s\big)$ — σ-алгебра, порождённая сечениями до момента s процесса $\big\langle\, \xi(t),\ t \in \mathbb{T}\, \big\rangle$ с независимыми приращениями. Тогда для $u > s$ сл.в. $\xi(u) - \xi(s)$ не зависит от σ-алгебры $\mathscr{F}_s$.*

⫢ По теореме 38, стр. 50, достаточно доказать, что $\xi(u) - \xi(s)$ не зависит от любого конечного набора $\xi(t_1), \ldots, \xi(t_k)$ с $t_1 < \ldots < t_k \leqslant s$. Дальнейшие рассуждения слабо зависят от значения k, поэтому рассмотрим только случай $k = 2$. Хар.ф. вектора $(\xi(t_1), \xi(t_2), \xi(u) - \xi(s))$ вычисляется так:

$$\begin{aligned}
&\mathbf{E}\exp\{\mathrm{i}\big(a_1\xi(t_1) + a_2\xi(t_2) + a_3(\xi(u) - \xi(s))\big)\} = \\
&= \mathbf{E}\exp\{\mathrm{i}\big((a_1 + a_2)\xi(t_1) + a_2(\xi(t_2) - \xi(t_1)) + a_3(\xi(u) - \xi(s))\big)\} = \\
&= \mathbf{E}\exp\{\mathrm{i}\big((a_1 + a_2)\xi(t_1) + a_2(\xi(t_2) - \xi(t_1))\big)\}\,\mathbf{E}\{\mathrm{i}a_3\big(\xi(u) - \xi(s)\big)\} = \\
&= \mathbf{E}\exp\{\mathrm{i}\big(a_1\xi(t_1) + a_2\xi(t_2)\big)\}\,\mathbf{E}\{\mathrm{i}a_3\big(\xi(u) - \xi(s)\big)\}
\end{aligned}$$

в силу независимости приращений. То есть искомая хар.ф. равна произведению хар.ф. $(\xi(t_1), \xi(t_2))$ и $(\xi(u) - \xi(s))$. Что и требовалось. ⫤

► Важным свойством процессов с независимыми приращениями является свойство их марковости. Рассмотрим сечения процесса в моменты $p_2 < p_1 < r < f_1 < f_2$. Если r — настоящий момент времени, то p_2, p_1 — два момента из «прошлого», f_1, f_2 — два момента из «будущего». В силу условия независимости приращений, векторы $(\xi_{p_2}, \xi_{p_1} - \xi_{p_2}, \xi_r - \xi_{p_1})$ и $(\xi_{f_2} - \xi_{f_1}, \xi_{f_1} - \xi_r)$ независимы. Так как измеримые преобразования любого из независимых векторов оставляют их независимыми, то отсюда следует независимость вектора с «настоящим» и «прошлым» $\vec{\eta} = (\xi_{p_2}, \xi_{p_1}, \xi_r)$ и вектора будущих при-

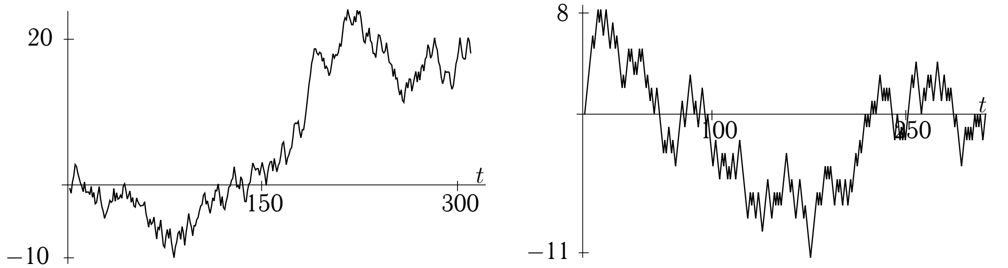


Рис. 6. Траектории процесса случайного блуждания

ращений $\vec{\zeta} = (\xi_{f_1} - \xi_r\,,\, \xi_{f_2} - \xi_r)$ относительно настоящего (сравните с леммой 372). Обозначим ф.р. вектора $\vec{\zeta}$ через F. Тогда условную ф.р. $(\xi_{f_1},\, \xi_{f_2})$ при фиксированном значении $\vec{\eta} = \vec{y} = (y_2, y_1, y_r)$ в силу указанной независимости (и в соответствии с теоремой 149, стр. 150) можно представить так:

$$\mathbf{P}\left\{ \left(\xi_{f_1}, \xi_{f_2}\right) \leqslant (x_1, x_2) \,\middle|\, \vec{\eta} = \vec{y} \right\} \;=\; \mathbf{P}\left\{ \vec{\zeta} \leqslant \left(x_1 - \xi_r, x_2 - \xi_r\right) \,\middle|\, \vec{\eta} = \vec{y} \right\}$$
$$=\; F(x_1 - y_r, x_2 - y_r).$$

Другими словами, при фиксированных значениях процесса в прошлом и настоящем распределение будущего зависит только от настоящего. Выбор по два момента из прошлого и будущего был сделан только в демонстрационных целях и не влияет на общий вывод.

373| П р и м е р. Процесс *случайного блуждания* образован по той же схеме, что и накопленный одним из игроков капитал в примере на стр. 313. Пусть $\varkappa_1, \varkappa_2, \ldots$ — последовательность независимых в совокупности сл.в., тогда последовательность частичных сумм

$$\xi_0 = 0, \quad \xi_1 = \varkappa_1, \quad \xi_2 = \varkappa_1 + \varkappa_2, \;\ldots\;, \; \xi_n = \xi_{n-1} + \varkappa_n, \;\; \ldots,$$

очевидно, образует сл.проц. (с дискретным временем $\mathbb{T} = \mathbb{N}_0$) с независимыми приращениями. Пара (k, ξ_k), $k = 0, 1, \ldots$, может рассматриваться как координата изучаемого объекта на пространственно-временной плоскости.

На двух слайдах рис. 6 приведены графики траекторий *случайного блуждания* для двух различных распределений $\varkappa_k$, $k \geqslant 1$, с единичными дисперсиями: на первом слайде приращения имеют равномерное распределение в отрезке $[-\sqrt{3}; \sqrt{3}]$, на втором — приращения принимают значения ± 1 с равными вероятностями.

Если при случайном блуждании все сл.в. ε_k одинаково распределены с ф.р. F, то, как легко понять, для фиксированного $m \in \mathbb{N}$ распределение приращения $\xi(t+m) - \xi(t) = \varkappa_{t+1} + \ldots + \varkappa_{t+m}$ не зависит от t (т.е. процесс однородный) и совпадает с m-кратной свёрткой $F * \cdots * F$. ◎

✧ **Эквивалентность процессов**. В связи с возможностью описания процесса через его распределение возникает вопрос: можно ли по заданному семейству конечномерных распределений определить характер поведения траекторий процесса ? Ответ будет отрицательным, если потребовать, чтобы то или иное свойство выполнялось для любого сл.проц. с этим распределением, даже если соответствующее утверждение измеримо. И часто положительным, если поставить вопрос о существовании такого процесса.

374| П р и м е р. а) Пусть $\mathbb{T} = [-1; 1]$ и $\xi(t) = t^2\eta$, где η — сл.в. (на некотором вероятностном пространстве) с ф.р. $F_\eta(x)$, $x \in \mathbb{R}^1$. Любая траектория этого сл.проц. (при $\eta \neq 0$) представляет собой параболу с вершиной в т. $(0,0)$, вид которой полностью определяется единственной реализацией сл.в. η. Для $\tau = \big\langle t_1, \ldots, t_n \big\rangle$, $t_i \neq 0$, $i = \overline{1,n}$, конечномерная ф.р. сл.проц.

$$F(\vec{x} \mid \tau) \;=\; \mathbf{P}\left\{\eta t_1^2 \leqslant x_1, \ldots, \eta t_n^2 \leqslant x_n\right\} \;=\; \mathbf{P}\Big\{\, \eta \leqslant \min_i \frac{x_i}{t_i^2} \,\Big\} = F_\eta\Big(\min_i \frac{x_i}{t_i^2}\Big).$$

Очевидно, это распределение сосредоточено на n-мерной прямой $l = \{\vec{x} : x_1 / t_1^2 = \ldots = x_n / t_n^2\}$.

Заметим, что все траектории этого процесса дифференцируемы всюду на $\mathbb{T}$ и интегрируемы на любом отрезке $\mathbb{T}$. Причём как производная функция $\xi'(t) = 2t\eta$, так и интеграл, скажем, от 0 до t : $\tilde{\xi}(t) = \frac{1}{3}t^3\eta$ также определяют некоторый квазидетерминированный сл.проц.

б) Изменим сл.проц. $\boldsymbol{\xi}$ всего в одной случайной точке. Пусть сл.в. $\zeta \backsim \backsim \mathfrak{Un}(-1,1)$ не зависит от η. Определим $\hat{\xi}(t) = \xi(t)\, \dot{\mathrm{I}}(\zeta \neq t)$. Легко понять, что т.к. для $\forall t$ вероятность $\mathbf{P}\{\zeta = t\} = 0$, то все конечномерные распределения $\hat{\boldsymbol{\xi}}$ совпадают с конечномерными распределениями $\boldsymbol{\xi}$. С другой стороны, почти все траектории $\hat{\boldsymbol{\xi}}$ терпят разрыв в какой-либо точке. ◎

Для дальнейшего нам необходимо определиться с тем, какие сл.проц. мы можем считать эквивалентными, т.е. описывающими одно и то же случайное явление. Наиболее естественным, на первый взгляд, здесь представляется подход, ориентированный на то, что реализации сл.проц. суть элементы измеримого пространства $(\mathscr{X}^{\mathbb{T}}, \mathscr{B}^{\mathbb{T}})$, что приводит нас к следующему определению.

Определение. Процессы $\xi(t)$, $\eta(t)$, $t \in \mathbb{T}$, называются *неотличимыми* (эквивалентными в узком смысле), если

$$\mathbf{P}\Big\{ \xi(t) = \eta(t) \quad \text{для} \quad \forall\, t \in \mathbb{T} \Big\} = 1$$

при условии, что указанное событие измеримо.

Ясно, что при счётном параметрическом пространстве $\mathbb{T}$ два сл.проц. будут неотличимы, если они неотличимы в каждый фиксированный момент времени.

375| **Лемма.** (?!) *Пусть $\mathbb{T}$ — не более чем счётное множество, тогда процессы $\xi(t)$, $\eta(t)$, $t \in \mathbb{T}$, неотличимы т. т. т. когда*

$$\mathbf{P}\Big\{ \xi(t) = \eta(t) \Big\} \;=\; 1 \quad \textit{для} \;\; \forall\, t \in \mathbb{T}. \tag{5}$$

Определение. Сл.проц. $\xi(t)$, $\eta(t)$, $t \in \mathbb{T}$, для которых имеет место (5), называются (стохастически) *эквивалентными* или *модификациями* (относительно друг друга).

Пример 374 показывает, что эквивалентные процессы могут и не быть неотличимыми. В примере это связано с наличием у $\hat{\xi}(t)$ слишком плохого разрыва в случайном месте, что, надеемся, никогда не наблюдается в природе. Пусть $\mathbb{T}$ — некоторый интервал $\mathbb{R}^1$. Обозначим через $\mathcal{C}(\mathbb{T})$ ($\mathcal{D}(\mathbb{T})$) подпространство $\mathscr{X}^{\mathbb{T}}$, включающее в себя все непрерывные (соответственно непрерывные справа и имеющие предел слева[‡]) функции на $\mathbb{T}$.

376| **Лемма.** (?!) *Пусть сл.проц. $\xi(t)$, $\eta(t)$, $t \in \mathbb{T}$, таковы, что почти все их траектории принадлежат $\mathcal{C}(\mathbb{T})$: $\mathbf{P}\{\boldsymbol{\xi} \in \mathcal{C}(\mathbb{T})\} = \mathbf{P}\{\boldsymbol{\eta} \in \mathcal{C}(\mathbb{T})\} = 1$. Тогда эти процессы неотличимы т. т. т. когда они эквивалентны. Аналогичное утверждение справедливо для пространства $\mathcal{D}(\mathbb{T})$.*

Определение. Сл.проц. $\xi(t)$, $\eta(t)$, $t \in \mathbb{T}$, распределения которых на $(\mathscr{X}^{\mathbb{T}}, \mathscr{B}^{\mathbb{T}})$ совпадают: $\mathsf{P}_{\boldsymbol{\xi}} = \mathsf{P}_{\boldsymbol{\eta}}$, называются *эквивалентными в широком смысле*. По лемме 13, стр. 35, для этого достаточно совпадения конечномерных распределений.

377| **Лемма.** *Если два процесса $\xi(t)$, $\eta(t)$, $t \in \mathbb{T}$, стохастически эквивалентны, то они эквивалентны в широком смысле.*

[‡] В метрике Скорохода (стр. 462) это пространство сепарабельно; для его обозначения часто используется аббревиатура «cadlag»— от *фр.* Continue A Droite, Limite A Gauche.

$\Rrightarrow$ Так как пересечение конечного числа событий вероятности единица также имеет вероятность единица, то для $\forall \langle t_1, \ldots, t_n \rangle$ векторы $\vec{\xi} = (\xi(t_1), \ldots, \xi(t_n))$, $\vec{\eta} = (\eta(t_1), \ldots, \eta(t_n))$ неотличимы: $\mathbf{P}\{\, \vec{\xi} = \vec{\eta} \,\} = 1$. Следовательно, для $\forall B \in \mathscr{B}^n$ вероятность

$$\mathbf{P}\{\, \vec{\xi} \in B \,\} \;=\; \mathbf{P}\{\, \vec{\xi} = \vec{\eta}, \vec{\xi} \in B \,\} \;=\; \mathbf{P}\{\, \vec{\eta} \in B \,\}. \qquad \Lleftarrow$$

378 Пример. Пусть $\Omega = [0;1]$ и $\mathbf{P}$ — вероятностная мера Лебега на $(\Omega, \mathscr{B})$. Сл.в. $\xi(\omega) = \omega$, $\eta(\omega) = 1 - \omega$ имеют одно и то же равномерное $\mathcal{U}\mathtt{n}(0,1)$ распределение, однако $\mathbf{P}\{\xi = \eta\} = \mathbf{P}\{\omega = {}^1\!/_2\} = 0$. ◎

В общем случае нельзя гарантировать, что процесс, заданный с помощью конечномерных распределений, будет иметь траектории с теми или иными свойствами. Вопрос должен быть поставлен несколько по-другому: существуют ли модификации процесса, для которых выполняются эти свойства? В качестве примера приведём классический результат А. Н. Колмогорова о непрерывной модификации. Но сначала — несколько определений.

Определения. Сл.процесс $\xi(t)$, $t \in \mathbb{T} = [a;b]$, называется:

a) *почти наверное непрерывным* (имеющим почти наверное непрерывные траектории), если для почти всех $\omega \in \Omega$ функция $\xi(t)$ непрерывна при всех $t \in [a;b]$;

b) *стохастически непрерывным* (непрерывным по вероятности) в точке t, если

$$\left|\xi(s) - \xi(t)\right| \xrightarrow[s\to t]{\mathbf{P}} 0;$$

c) непрерывным в *среднем квадратическом* (или в $\mathcal{L}_2$) в точке t, если

$$\lim_{s\to t} \mathbf{E}\left(\xi(s) - \xi(t)\right)^2 = 0.$$

✧ **Теорема Колмогорова о непрерывной модификации.** Оба процесса, рассмотренные в примере 374, непрерывны всюду как по вероятности, так и в среднем квадратическом, если, конечно, $\mathbf{E}\eta^2 < \infty$. В то же время почти все траектории процесса из пункта б) этого примера имеют по одному разрыву. Чисто случайная функция не будет непрерывной ни по вероятности, ни в среднем квадратическом ни в одной точке.

379 **Теорема.** [*А.Н. Колмогоров.*] *Пусть конечномерные распределения сл.процесса* $\boldsymbol{\xi} = \langle\, \xi(t), t \in [a;b] \,\rangle$ *таковы, что для некоторых* $\beta, z, C > 0$

при $\forall\, t, s \in [a; b]$

$$\mathbf{E}\left|\xi(t) - \xi(s)\right|^{\beta} \leqslant C|t-s|^{1+z}. \tag{6}$$

Тогда для процесса $\boldsymbol{\xi}$ *существует модификация* $\boldsymbol{\eta}$, *у которой все траектории непрерывны.*

$\Rrightarrow$ Заметим сначала, что по неравенсту Маркова для $\forall\, \varepsilon > 0$

$$\mathbf{P}\left\{\left|\xi(t) - \xi(s)\right| > \varepsilon\right\} \;\leqslant\; \frac{1}{\varepsilon^{\beta}}\,\mathbf{E}\left|\xi(t) - \xi(s)\right|^{\beta} \;\leqslant\; \frac{C}{\varepsilon^{\beta}}\,|t-s|^{1+z}$$

в силу условия теоремы. Отсюда ясно, что процесс $\boldsymbol{\xi}$ стохастически непрерывен в любой точке $\forall\, t \in [a; b]$.

Положим $[a; b] = [0; 1]$. Рассмотрим покрытие отрезка $[0; 1]$ интервалами вида $J_{n,k} = \left[\frac{k}{2^n}; \frac{k+1}{2^n}\right]$, $k = 0, 1 \ldots, 2^n - 1$, $n \geqslant 1$. Для каждой фиксированной траектории процесса $\boldsymbol{\xi}$ зададим кусочно-линейную непрерывную функцию $\xi_n(t) = \sum_{k=0}^{2^{n-1}} \xi_{n,k}(t)\, \dot{\mathrm{I}}\, \left(t \in J_{n,k}\right)$, где

$$\xi_{n,k}(t) \;=\; \xi\Big(\frac{k}{2^n}\Big) + 2^n\Big(t - \frac{k}{2^n}\Big)\Big[\xi\Big(\frac{k+1}{2^n}\Big) - \xi\Big(\frac{k}{2^n}\Big)\Big].$$

Другими словами, функция $\xi_n(t)$ на участке $t \in \left[\frac{k}{2^n}; \frac{k+1}{2^n}\right]$ представляет собой отрезок прямой и совпадает на краях этого интервала с $\xi(t)$. В то же время функция $\xi_{n+1}(t)$ на этом же интервале $t \in \left[\frac{k}{2^n}; \frac{k+1}{2^n}\right]$ $\left(= \left[\frac{2k}{2^{n+1}}; \frac{2k+2}{2^{n+1}}\right]\right)$ состоит из двух отрезков прямой с изломом в средней точке $(2k+1)/2^{n+1}$. Стало быть, максимум расхождения

$$\begin{aligned}
\max_{t \in J_{n,k}} |\xi_n(t) - \xi_{n+1}(t)| \;&=\; \Big|\xi_n\Big(\frac{2k+1}{2^{n+1}}\Big) - \xi_{n+1}\Big(\frac{2k+1}{2^{n+1}}\Big)\Big| \;= \\
&=\; \Big|\frac{1}{2}\xi\Big(\frac{2k}{2^{n+1}}\Big) - \frac{1}{2}\xi\Big(\frac{2k+2}{2^{n+1}}\Big) - \xi\Big(\frac{2k+1}{2^{n+1}}\Big)\Big| \;\leqslant \\
&\leqslant\; \frac{1}{2}\Big(\Big|\xi\Big(\frac{2k+1}{2^{n+1}}\Big) - \xi\Big(\frac{2k}{2^{n+1}}\Big)\Big| + \Big|\xi\Big(\frac{2k+1}{2^{n+1}}\Big) - \xi\Big(\frac{2k+2}{2^{n+1}}\Big)\Big|\Big).
\end{aligned}$$

Заметим, что первое равенство в этих соотношениях гарантирует измеримость рассматриваемого максимума, как функции ω.

По неравенству Маркова из условия (6) получаем для первого слагаемого

$$\begin{aligned}
\mathbf{P}\left\{\frac{1}{2}\Big|\xi\Big(\frac{2k+1}{2^{n+1}}\Big) - \xi\Big(\frac{2k}{2^{n+1}}\Big)\Big| > \frac{1}{2^{nz/2\beta+1}}\right\} \;&\leqslant\; 2^{nz/2}\,\mathbf{E}\Big|\xi\Big(\frac{2k+1}{2^{n+1}}\Big) - \xi\Big(\frac{2k}{2^{n+1}}\Big)\Big|^{\beta} \;\leqslant \\
&\leqslant\; C\,2^{-(1+z)(n+1)+nz/2} \;=\; C\,\frac{1}{2^{n+1}}\,\frac{1}{2^{nz/2}}.
\end{aligned}$$

Аналогично для второго слагаемого. Таким образом,

$$\mathbf{P}\Big\{ \max_{t\in[0;1]} |\xi_n(t)-\xi_{n+1}(t)| > \frac{1}{2^{nz/2\beta+1}} \Big\} \leqslant$$
$$\leqslant \sum\nolimits_{k=0}^{2^{n-1}} \mathbf{P}\Big\{ \max_{t\in J_{n,k}} |\xi_n(t)-\xi_{n+1}(t)| > \frac{1}{2^{nz/2\beta+1}} \Big\} \leqslant C\,\frac{1}{2^{nz/2}}.$$

Поскольку ряд $\sum 2^{-nz/2} < \infty$, то по лемме Бореля–Кантелли почти наверное события $\big\{\max_{t\in[0;1]} |\xi_n(t)-\xi_{n+1}(t)| > \frac{1}{2^{nz/2\beta+1}}\big\}$ не могут происходить бесконечное число раз, т.е. для почти всех ω найдётся такой номер $n(\omega)$, что $\max_{t\in[0;1]} |\xi_n(t)-\xi_{n+1}(t)| \leqslant \frac{1}{2^{nz/2\beta+1}}$ при всех $n > n(\omega)$. Отсюда следует, что для почти всех ω ряд (с $\xi_0 \equiv 0$)

$$\eta(t) \;=\; \sum_{n=1}^{\infty} \big(\xi_n(t)-\xi_{n-1}(t)\big)$$

абсолютно сходится при любом $t\in[0;1]$. При этом для $m > n(\omega)$

$$\max_{t\in[0;1]} \big|\eta(t)-\xi_m(t)\big| \;\leqslant\; \sum_{n=m}^{\infty} \frac{1}{2^{nz/2\beta+1}} \underset{m\to\infty}{\longrightarrow} 0.$$

Таким образом, функция $\eta(t)$ для почти всех ω непрерывна, т.к. есть равномерный предел непрерывных функций. Осталось показать стохастическую эквивалентность $\boldsymbol{\eta}$ и $\boldsymbol{\xi}$.

Так как процесс $\boldsymbol{\xi}$ стохастически непрерывен, то для $\forall\, t\in[0;1]$ и любой последовательности двоично-рациональных точек $s_n = k_n/2^n \underset{n}{\to} t$ имеет место сходимость $\xi(s_n) \underset{n}{\to} \xi(t)$ по вероятности. Следовательно, найдётся подпоследовательность s_{n_k} такая, что эта сходимость имеет место и почти наверное. По построению $\eta(s_{n_k}) = \xi(s_{n_k})$, стало быть, и $\eta(t)=\xi(t)$ почти наверное. ⫤

Рассматриваемый далее процесс Пуассона $\xi(t)$, подсчитывающий количество «редких» событий (интенсивности λ), произошедших к моменту t, не удовлетворяет условиям теоремы Колмогорова, поскольку $\mathbf{E}\,|\xi(t+h)-\xi(t)|^\beta \asymp \asymp h,\ h\to 0$. Однако этот процесс можно определить так, чтобы все его траектории были непрерывны справа.

380| **Теорема.** [*Колмогоров–Ченцов.*] *Пусть конечномерные распределения сл.процесса* $\boldsymbol{\xi} = \big\langle\, \xi(t), t\in[a;b] \,\big\rangle$ *таковы, что для некоторых* $\alpha,\beta,z,C > > 0$ *при* $\forall\, a \leqslant t-h_1 < t < t+h_2 \leqslant b$

$$\mathbf{E}\Big[\big|\xi(t)-\xi(t-h_1)\big|^{\alpha}\big|\xi(t+h_2)-\xi(t)\big|^{\beta}\Big] \;\leqslant\; C(h_1+h_2)^{1+z}.$$

Тогда для процесса $\boldsymbol{\xi}$ существует модификация $\boldsymbol{\eta}$, у которой все траектории принадлежат пространству $\mathcal{D}(\mathbb{T})$.

Для процесса Пуассона при любых $t, h_1, h_2 > 0$ сл.в. $(\xi(t) - \xi(t-h_1))$ и $(\xi(t+h_2) - \xi(t))$ независимы (в силу независимости приращений). Поэтому

$$\mathbf{E}\left[|\xi(t) - \xi(t-h_1)||\xi(t+h_2) - \xi(t)|\right] \;=\; \lambda^2 h_1 h_2 \;\leqslant\; \lambda^2 (h_1 + h_2)^2,$$

что в соответствии с предыдущей теоремой гарантирует наличие модификации процесса, все траектории которой принадлежат пространству $\mathcal{D}(\mathbb{T})$.

✧ **Основные характеристики распределения процесса**. Любому процессу, представляющему собой набор сл.в., можно поставить в соответствие набор его числовых характеристик. Часто эти характеристики полностью описывают процесс.

Определения. Пусть $\boldsymbol{\xi} = \left\langle\, \xi(t), t \in \mathbb{T} \,\right\rangle$ — сл.процесс.

Функция $\mu_{\boldsymbol{\xi}}(t)$, $t \in \mathbb{T}$, определяемая соотношением

$$\mu_{\boldsymbol{\xi}}(t) \;=\; \mathbf{E}\left[\xi(t)\right],$$

называется *математическим ожиданием* (*средним значением*) процесса.

Функция $\mathbb{V}_{\boldsymbol{\xi}}(t)$, $t \in \mathbb{T}$, определяемая соотношением

$$\mathbb{V}_{\boldsymbol{\xi}}(t) \;=\; \mathbf{D}\left[\xi(t)\right] \;=\; \mathbf{E}\left[\xi(t)^2\right] - \left(\mathbf{E}\left[\xi(t)\right]\right)^2,$$

называется *дисперсионной функцией (дисперсией)* процесса; для её существования необходимо, чтобы процесс был *гильбертовым*, т.е. $\mathbf{E}\left[\xi(t)^2\right] < \infty, t \in \mathbb{T}$.

Функция $R_{\boldsymbol{\xi}}(t,s)$, $t, s \in \mathbb{T}$, определяемая соотношением

$$R_{\boldsymbol{\xi}}(t,s) \;=\; \mathbf{Cov}\left(\xi(t), \xi(s)\right) \;=\; \mathbf{E}\left[\xi(t)\xi(s)\right] - \mathbf{E}\left[\xi(t)\right]\mathbf{E}\left[\xi(s)\right],$$

называется *ковариационной функцией* (гильбертова) процесса. При $s = t$ ковариационная функция совпадает с дисперсионной функцией: $R_{\boldsymbol{\xi}}(t,t) = \mathbb{V}_{\boldsymbol{\xi}}(t)$.

⚠ Иногда (точнее, очень часто) ковариационную функцию называют корреляционной, имея в виду, что она показывает степень связей между сечениями сл.проц. В классических разделах теории вероятностей корреляция, как известно, отличается от ковариации наличием нормирующего коэффициента $\left(\mathbb{V}_{\boldsymbol{\xi}}(t)\mathbb{V}_{\boldsymbol{\xi}}(s)\right)^{-1/2}$. Здесь этот коэффициент только усложнит анализ.

⚠ Наименование «гильбертов» связано с тем, что пространство всех сл.величин $\xi : (\Omega, \mathscr{F}, \mathbf{P}) \mapsto (\mathbb{R}^1, \mathscr{B})$, интегрируемых с квадратом, образует гильбертово пространство (см. стр. 530) со скалярным произведением $(\xi, \eta) = \int_\Omega \xi\eta\, d\,\mathbf{P}$. Чтобы в полной мере воспользоваться богатствами этого пространства, удобнее не ограничиваться действительными сл.в., а считать, что фазовое пространство процесса $\mathscr{X} \subset \mathbb{C}$, т.е. все сечения суть комплекснозначные сл.в. Скалярное произведение для таких сл.в. определяется равенством $(\xi, \eta) = \int_\Omega \xi\,\overline{\eta}\, d\,\mathbf{P}$, где, как всегда в комплексном анализе, $\overline{z} = (a - \mathrm{i}b)$ — сопряжённое к числу $z = (a + \mathrm{i}b)$. В гильбертовом пространстве $\mathbb{R}^2$ равенство нулю скалярного произведения геометрически означает ортогональность элементов. В теории вероятностей центрированные сл.в. также часто называют *ортогональными*, если их коэффициент ковариации равен нулю.

Набору моментов времени $\tau = (t_1, \ldots, t_n)$ можно сопоставить n-мерную матрицу ковариаций гильбертова процесса

$$R_{\boldsymbol{\xi}}(\tau) = \Big(R_{\boldsymbol{\xi}}(t_i, t_j)\Big)_{i,j=1}^{n}, \tag{7}$$

которая совпадает с матрицей ковариаций сл.вектора $(\xi(t_1), \ldots, \xi(t_n))$. Из хорошо известного свойства ковариационных матриц (см. 75, стр. 76) следует, что эта матрица неотрицательно определена. Следовательно, функция $R_{\boldsymbol{\xi}}(s,t)$ неотрицательно определена, т.е. $\forall\,(t_1, \ldots, t_n) \subset \mathbb{T}$, $z_j \in \mathbb{R}^1$ $(\mathbb{C})$, $j = \overline{1,n}$,

$$\sum_{i,j=1}^{n} R_{\boldsymbol{\xi}}(t_i, t_j) z_i z_j \geqslant 0.$$

Из дальнейшего будет видно, что любая неотрицательно определённая функция $R(s,t)$ на $\mathbb{T} \times \mathbb{T}$ может служить ковариацией некоторого сл.процесса.

✧ **Стационарные процессы. Спектральное разложение.** Стационарные процессы образуют один из самых глубоко исследованных классов сл.процессов.

Определение. Сл.проц. $\left\langle\, \xi(t), t \in \mathbb{T}\,(\subset \mathbb{R}^1)\,\right\rangle$ называется *стационарным* (в узком смысле), если конечномерное распределение процесса не зависит от сдвига по времени:

$$\big(\xi(t_1), \ldots, \xi(t_n)\big) \backsim \big(\xi(t_1 + m), \ldots, \xi(t_n + m)\big)$$

для $\forall\, t_i, t_i + m \in \mathbb{T}$, $i = \overline{1,n}$, $n \geqslant 1$, $m \geqslant 0$.

⚠ Распределение стационарного процесса с независимыми приращениями и дискретным временем ($\mathbb{T} = \mathbb{N}_0$) полностью определяется распределением F_0 начальной сл.в. ξ_0 и распределением F_1 приращения за один шаг $\varkappa = \xi_1 - \xi_0$. Действительно, для $\forall\, t \geqslant 1$

$$\xi_t \;=\; \xi_0 + \sum\nolimits_{k=1}^{t} (\xi_k - \xi_{k-1}) \;=\; \xi_0 + \sum\nolimits_{k=1}^{t} \varkappa_k \quad\backsim\quad F_0 * F_1^{*t},$$

где $\varkappa_1, \varkappa_2, \ldots$ — независимые сл.в., распределённые как $\varkappa$; здесь «$*$» — операция свёртки. Аналогично записывается любое конечномерное распределение.

Для стационарного процесса $\boldsymbol{\xi} = \big\langle\, \xi(t),\ t \in \mathbb{T}\, \big\rangle$ сл.в. $\xi(t) \backsim \xi(t + m)$, $\forall t, t + m \in \mathbb{T}, m \geqslant 0$. Поэтому $\mathbf{E}\big[\xi(t)\big]^k = \mathbf{E}\big[\xi(t+m)\big]^k$, если эти моменты существуют. Следовательно, у стационарного процесса функция среднего значения и дисперсионная функция постоянны: $\mu_{\boldsymbol{\xi}}(t) = \mu$, $\mathbb{V}_{\boldsymbol{\xi}}(t) = \lambda^2$, $\forall t \in \mathbb{T}$. Кроме того, для стационарного процесса распределение сл.вектора $(\xi(s), \xi(t)) \backsim (\xi(t_0), \xi(t - s + t_0))$, $t_0 \in \mathbb{T}, \forall 0 \leqslant s < t$, поэтому ковариационная функция такого процесса зависит только от разности аргументов: $R_{\boldsymbol{\xi}}(s,t) =: \mathcal{R}(t - s)$, $\quad \forall\, t, s \in \mathbb{T}$.

Сильное условие стационарности, определяющее всю структуру конечномерных распределений, можно ослабить, введя следующее

Определение. Гильбертов сл.процесс называется *стационарным в широком смысле*, если:

i) его функция среднего и дисперсионная функция постоянны;

ii) ковариационная функция зависит от разности аргументов.

Выше отмечалось, что ковариационная функция процесса неотрицательно определена. С подобными функциями мы уже встречались вскользь при рассмотрении свойств характеристических функций, где была приведена теорема Бохнера, утверждающая, что при соответствующих условиях нормировки и непрерывности неотрицательно определённая функция есть характеристическая функция некоторого распределения вероятностей. Подобное утверждение справедливо и для ковариационных функций стационарных процессов.

381| **Теорема.** $\big[$*Герглотц,* [5].$\big]$ *Пусть* $R(s,t) = \mathcal{R}(t - s)$ *— ковариационная функция стационарного в широком смысле сл.проц. с дискретным временем,* $s, t \in \mathbb{T} = \mathbb{Z} = \{0, \pm 1, \ldots\}$. *Тогда на борелевской* σ*-алгебре* $\mathscr{B}$ *интервала* $\Theta = (-\pi; \pi]$ *существует однозначно определённая конечная мера* $\mathscr{G}$ *такая, что*

$$\mathcal{R}(t-s)=\int_{\Theta} e^{\mathrm{i}(t-s)\,\theta}\,\mathscr{G}(d\theta),\qquad s,t\in\mathbb{Z}.$$

Мера $\mathscr{G}$ в этом представлении называется *спектральной мерой*. Спектральную меру можно описать с помощью *спектральной функции* — аналога ф.распределения: $F(\theta)=\mathscr{G}(-\pi;\theta]$. Если спектральная мера (функция) абсолютно непрерывна относительно меры Лебега, то соответствующая плотность $g(\theta)$ называется *спектральной плотностью* и в представлении для ковариации можно положить $\mathscr{G}(d\theta)=g(\theta)\,d\theta$. Заметим, что для действительных процессов плотность симметрична около нуля (чётна), поскольку ковариационная функция $\mathcal{R}(s-t)=\mathcal{R}(t-s)$. По некоторым соображениям дисперсию $\mathcal{R}(0)$ можно воспринимать как среднюю мощность стационарного случайного сигнала (например, мощность звукового сигнала, изменяющегося во времени, или мощность процесса теплоотдачи электрического проводника). Тогда ввиду равенства $\mathbb{V}_{\boldsymbol{\xi}}(t)=\mathcal{R}(0)=\int_{-\pi}^{\pi}g(\theta)\,d\theta$ спектральную плотность можно охарактеризовать как плотность распределения мощности сигнала по частотному спектру.

Числа $\mathcal{R}_{\boldsymbol{\xi}}(n)$ равны коэффициентам разложения плотности $g(\theta)$ в ряд Фурье по системе ортогональных функций $e^{\mathrm{i}n\theta}$, $n\in\mathbb{Z}$, $\theta\in(-\pi;\pi]$. Обратно, если $\sum_n\big|\mathcal{R}_{\boldsymbol{\xi}}(n)\big|<\infty$, то ряд Фурье с коэффициентами $\mathcal{R}_{\boldsymbol{\xi}}(n)$ абсолютно сходится к спектральной плотности $g_{\boldsymbol{\xi}}(\theta)$ стационарного случ.проц. $\boldsymbol{\xi}$:

$$g_{\boldsymbol{\xi}}(\theta)=\frac{1}{2\pi}\sum_{n\in\mathbb{Z}}e^{-\mathrm{i}n\theta}\mathcal{R}_{\boldsymbol{\xi}}(n). \tag{8}$$

382| Примеры. 1) Простыми вычислениями показывается, что для сигнала, у которого мощность на всех участках спектра одинакова, т.е. спектральная мера $\mathscr{G}\backsim\mathtt{Un}(-\pi,\pi)$ (спектральная плотность $g(\theta)\equiv\dfrac{1}{2}\pi$), ковариационная функция при ненулевом целом аргументе $t\in\mathbb{N}$ равна $\mathcal{R}(t)=0$ и $\mathcal{R}(0)=1$. Таким образом, постоянная спектральная плотность соответствует дискретному стандартному белому шуму (если процесс ещё и центрирован). Отсюда название «белый шум» — белый цвет есть результат смешения всех цветов спектра с равной интенсивностью (мощностью).

2) Рассмотрим процесс белого шума $\boldsymbol{\varkappa}=\big\langle\,\varkappa_t\,\big\rangle_0^{\infty}$. Из отрезка этого процесса, состоящего из реализаций в последние $m+1$ моментов времени (по отношению к моменту t), образуем процесс $\boldsymbol{\xi}$ в виде линейной комбинации

$$\xi(t) \;=\; \sum_{k=0}^{m} c_k\, \varkappa(t-k) \;=\; c_m\, \varkappa(t-m) + \ldots + c_0\, \varkappa(t), \qquad t \geqslant m,$$

с некоторым набором весовых констант $c_0, \ldots, c_m$. С физической точки зрения этот процесс соответствует сигналу на выходе некоторого прибора (линейного фильтра), обрабатывающего входной сигнал, когда на вход подаётся белый шум. Для анализа удобнее в представлении этого процесса сделать замену переменной суммирования $k \dashrightarrow t-k$ и записать его в виде двустороннего бесконечного ряда

$$\xi(t) = \sum_{k=-\infty}^{\infty} c_{t-k}\, \varkappa(k), \tag{9}$$

где в нашем случае $c_k = 0$ при $k < 0$ или $k > m$, а $\varkappa(k)$ при отрицательном индексе может быть выбрано произвольным образом, например, $\varkappa(k) = 0$ или $\varkappa(k) = \varkappa'(-k)$, $k < 0$, где $\boldsymbol{\varkappa}'$ — процесс, не зависящий от $\boldsymbol{\varkappa}$. Поскольку только конечное число слагаемых отлично от нуля, вопрос сходимости бесконечного ряда не возникает. В общем случае с $m = \infty$ сходимость этого ряда (в смысле $\mathcal{L}_2$) можно обеспечить при условии, что $\sum_k c_k^2 < \infty$. Среднее этого процесса, очевидно, равно нулю. Кроме того, легко понять, что если $t-s > m$, т.е. когда отрезки белого шума, образующие значения $\xi(s)$ и $\xi(t)$, не пересекаются, ковариация $\mathbf{Cov}(\xi(s), \xi(t)) = 0$. В противном случае

$$\begin{aligned} R(s,t) \;&=\; \mathbf{Cov}(\xi(s), \xi(t)) \;=\; \sum\nolimits_{k=-\infty}^{\infty} c_{s-k} c_{t-k} \;= \\ &=\; \sum\nolimits_{k=t-m}^{s} c_{s-k} c_{t-k} \;=\; \sum\nolimits_{j=t-s}^{m} c_{j-(t-s)} c_j, \end{aligned}$$

в частности $\mathbf{D}\xi(t) = \sum_0^m c_k^2$. Таким образом, процесс $\boldsymbol{\xi}$ стационарный в широком смысле. Ясно, что он стационарный и в узком смысле. Его спектральная плотность равна $|\varphi(\theta)|^2 = \varphi(\theta)\overline{\varphi(\theta)}$, где

$$\varphi(\theta) \;=\; \frac{1}{\sqrt{2\pi}} \sum\nolimits_{-\infty}^{\infty} c_k e^{-\mathtt{i}k\theta}$$

при тех же условиях на коэффициенты c_k.

Действительно, для любых $s \leqslant t$ произведение

$$e^{\mathtt{i}(t-s)\theta}|\varphi(\theta)|^2 \;=\; \frac{1}{2\pi}\varphi(\theta)e^{\mathtt{i}t\theta}\overline{\varphi(\theta)e^{\mathtt{i}s\theta}} \;=\; \sum\nolimits_{k,j=-\infty}^{\infty} \frac{c_k c_j}{2\pi}\, e^{\mathtt{i}(t-s-k+j)\theta}.$$

Интеграл от $e^{\mathtt{i}(t-s-k+j)\theta}$ по области $\theta \in (-\pi, \pi]$, очевидно, равен $\sin\big((t-s- - k+j)\pi\big)/(t-s-k+j) = 0$, если $(t-s-k+j) \neq 0$. В противном случае

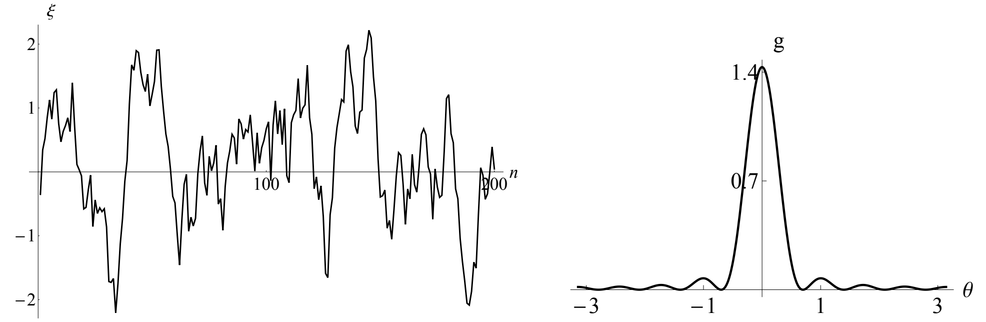


Рис. 7. Траектория процесса скользящего среднего ($m = 9$) и график соответствующей спектральной плотности

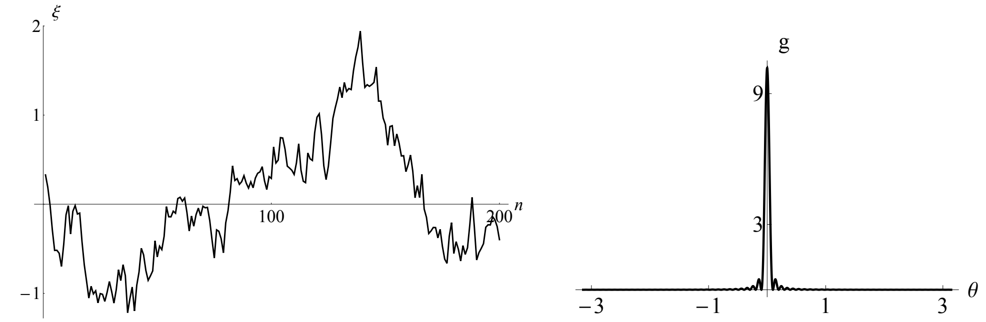


Рис. 7′. Траектория процесса скользящего среднего ($m = 64$) и график соответствующей спектральной плотности

этот интеграл равен 2π. Таким образом,

$$\int_{-\pi}^{\pi} e^{\mathtt{i}(t-s)\theta}|\varphi(\theta)|^2\,d\theta \;=\; \sum_{j=t-s}^{m} c_{j-(t-s)}c_j \;=\; R(s,t).$$

Аналогичный результат справедлив и в случае, когда $m = \infty$, $\sum_k c_k^2 < \infty$ и сходимость случайного ряда (9) понимается в среднем квадратическом.

При выборе $c_k = 1/(m+1)$ значения $\xi(t)$ представляют собой среднее арифметическое значений последних $m+1$ реализаций процесса $\boldsymbol{\varkappa}$. Поэтому процессы подобного типа называют *процессами скользящего среднего*. На рисунках 7 и 7' представлены графики траекторий двух процессов скользящего среднего вида $\sum_1^m \varkappa_i/\sqrt{m}$, где независимые сл.в. $\varkappa_i \backsim \mathtt{Un}(-\sqrt{3},\sqrt{3}), i = \overline{1,m}$, так что среднее процесса равно нулю, а дисперсия равна единице. Для понимания характера зависимости поведения траекторий от вида спектральной плотности здесь также приведены графики соответствующих спектральных плотностей процессов.

з) Иногда можно предположить, что значения процесса в данный момент времени t линейно зависят от части предыдущих значений. Пусть $\langle \varkappa_n,\ n \in \mathbb{Z} \rangle$ — двусторонний дискретный белый шум с единичной дисперсией (последнее только для сокращения записи). Говорят, что сл.проц. ξ_n удовлетворяет уравнению *авторегрессии* первого порядка, если

$$\xi_n = \lambda \xi_{n-1} + \varkappa_n. \tag{10}$$

383 **Теорема.** *Если* $|\lambda| < 1$, *то существует стационарный в широком смысле сл.проц., удовлетворяющий* (10), *который может быть представлен в виде одностороннего скользящего среднего:*

$$\xi_n = \sum\nolimits_{j=0}^{\infty} \lambda^j \varkappa_{n-j}.$$

⇛ По теореме Колмогорова–Хинчина о двух рядах 313, стр. 269, этот ряд сходится с вероятностью единица. Более того, он сходится и в среднем квадратическом:

$$\mathbf{E}\left[\sum\nolimits_{j=k}^{m} \lambda^j \varkappa_{n-j}\right]^2 = \sum\nolimits_{j=k}^{m} \lambda^{2j} \underset{k,m\to\infty}{\longrightarrow} 0.$$

В силу этого математическое ожидание $\mathbf{E}\xi_n = 0$, а ковариация (для $k \leqslant n$)

$$\mathbf{E}\left[\xi_n \xi_k\right] = \sum\nolimits_{j=0}^{\infty} \sum\nolimits_{m=0}^{\infty} \lambda^{j+m} \mathbf{E}\left[\varkappa_{n-j}\varkappa_{k-m}\right] = \sum\nolimits_{m=0}^{\infty} \lambda^{n-k+2m} = \frac{\lambda^{n-k}}{1-\lambda^2},$$

т.к. множители $\mathbf{E}\left[\varkappa_{n-j}\varkappa_{k-m}\right] = 1$ при $j = n-k+m$; в остальном — равны нулю. Таким образом, процесс $\boldsymbol{\xi} = \langle \xi_n,\ n = 0, \pm 1, \dots \rangle$ стационарный.

Легко проверить, что процесс $\boldsymbol{\xi}$ удовлетворяет уравнению (10). Обратно, любой другой стационарный процесс, удовлетворяющий авторегрессионному уравнению, почти наверное совпадает с $\boldsymbol{\xi}$. Действительно, последовательно применяя уравнение авторегрессии, получаем

$$\xi_n = \lambda^k \xi_{n-k} + \sum\nolimits_{j=0}^{k-1} \lambda^j \varkappa_{n-j} \underset{k\to\infty}{\longrightarrow} \sum\nolimits_{j=0}^{\infty} \lambda^j \varkappa_{n-j} \quad (\text{п.н.}),$$

где $\lambda^k \xi_{n-k} \to 0$ (п.н.), т.к. по предположению процесс ξ_n стационарный и, следовательно, $\sum_k \mathbf{P}\left\{|\lambda^k \xi_{n-k}| > \lambda^{k/2}\right\} \leqslant \mathbf{D}\xi_1 \sum_k \lambda^k < \infty$ (см. теорему 205, стр. 194). ⇚

Спектральную плотность этого процесса, как процесса скользящего среднего, можно представить следующим образом (см. предыдущий пример 2):

$$|\phi(\theta)|^2 \;=\; \frac{1}{2\pi}\Big|\sum\nolimits_{-\infty}^{\infty}\lambda^k e^{-\mathrm{i}k\theta}\Big|^2 \;=\; \frac{1}{2\pi}\frac{1}{|1-\lambda e^{-\mathrm{i}\theta}|^2}\,.$$

Обобщением авторегрессионного процесса первого порядка (10) служит *процесс авторегрессии порядка* $q \geqslant 1$:

$$\xi_n \;=\; \lambda_1\xi_{n-1} + \ldots + \lambda_q\xi_{n-q} + \varkappa_n. \tag{11}$$

Такой процесс также может быть записан в виде скользящего среднего. Для этого рассмотрим оператор, действующий на двустороннюю бесконечную последовательность $\boldsymbol{x} = (\ldots, x_{-1}, x_0, x_1, \ldots)$ путём сдвига на одну координату вправо:

$$\Upsilon(\boldsymbol{x}) = (\ldots, x_{-2}, x_{-1}, x_0, \ldots).$$

Уравнение (11) в терминах этого оператора можно переписать в виде

$$\boldsymbol{\xi} = \lambda_1\Upsilon(\boldsymbol{\xi}) + \ldots + \lambda_q\Upsilon^q(\boldsymbol{\xi}) + \boldsymbol{\varkappa} \qquad \text{или}$$

$$\Big(\mathbf{1} - \sum\nolimits_1^q \lambda_j\Upsilon^j\Big)(\boldsymbol{\xi}) = \boldsymbol{\varkappa},$$

где $\mathbf{1}$ — единичный оператор. Таким образом, $\boldsymbol{\xi} = \Big(\mathbf{1} - \sum_1^q \lambda_j\Upsilon^j\Big)^{-1}\boldsymbol{\varkappa}$, если существует обратный оператор.

Поиск обратного оператора здесь сродни поиску разложения в степенной ряд (по степеням γ) дробно-рациональной функции:

$$\frac{1}{1-\lambda_1\gamma-\ldots-\lambda_q\gamma^q} \;=\; C\prod_1^q\frac{1}{\gamma_j-\gamma} \;=\; \prod_1^q\frac{1}{1-\gamma/\gamma_j}\,,$$

у которой такое разложение существует только в области $|\gamma| < \min_j|\gamma_j|$, где γ_j — корни уравнения $1-\lambda_1\gamma-\ldots-\lambda_q\gamma^q = 0$. Если все корни этого уравнения по модулю больше 1, то существует представление обратного оператора

$$\Big(\mathbf{1} - \sum\nolimits_1^q \lambda_j\Upsilon^j\Big)^{-1} \;=\; \sum\nolimits_0^{\infty}\beta_j\Upsilon^j, \qquad \beta_0 \;=\; 1,$$

и стационарная последовательность авторегрессии порядка q

$$\xi_n \;=\; \varkappa_n + \sum\nolimits_1^{\infty}\beta_j\varkappa_{n-j},$$

где ряд сходится в среднем квадратическом.

Рассмотрим для примера процесс авторегрессии $\xi_n = \frac{1}{4}\xi_{n-1} + \frac{1}{8}\xi_{n-2} + \varkappa_n$.

Уравнение $1-\frac{1}{4}\gamma-\frac{1}{8}\gamma^2=0$ имеет два корня $\gamma_1=-4, \gamma_2=2$. Поэтому

$$\frac{1}{1-\frac{1}{4}\gamma-\frac{1}{8}\gamma^2} = \frac{1}{1-\frac{1}{2}\gamma}\frac{1}{1+\frac{1}{4}\gamma} = \sum_{j=0}^{\infty}\sum_{m=0}^{\infty}(-1)^m\frac{\gamma^{j+m}}{2^{j+2m}} \overset{(j+m \dashrightarrow k)}{=}$$

$$= 1+\sum_{k=1}^{\infty}\frac{1}{2^k}\gamma^k\sum_{m=0}^{k}\Big((-1)^m\frac{1}{2^m}\Big).$$

Таким образом, $\xi_n=\varkappa_n+\sum_1^\infty\beta_k\varkappa_{n-k}$, где $\beta_k=\frac{1}{3\cdot 2^{k-1}}\big(1-(-1/2)^{k+1}\big)$.

4) Рассмотрим теперь двусторонний комплекснозначный процесс

$$\xi_n = \sum_{k=1}^{m}\eta_k e^{\mathrm{i}\theta_k n}, \qquad n\in\mathbb{Z}=\{0,\pm1,\pm2,\ldots\}, \tag{12}$$

где $\eta_1,\ldots,\eta_m$ — некоррелированные и центрированные комплексные сл.в., т.е. $\mathbf{E}\,\eta_k=0$ и $\exists\,\mathbf{Cov}(\eta_k,\eta_j)=\mathbf{E}\,\eta_k\overline{\eta}_j=0,\ k\neq j\in\{1,\ldots,m\}$. Точки $\theta_1,\ldots,\theta_m$ — произвольные различные точки из интервала $(-\pi;\pi]$. В физике функция $Ae^{\mathrm{i}\theta t}$ описывает элементарное гармоническое колебание с частотой (гармоникой) θ и амплитудой A. Таким образом, процесс $\boldsymbol{\xi}$ есть результат наложения m гармонических колебаний с различными частотами. Набор гармоник $\theta_k,\ k=\overline{1,m}$, и соответствующих им дисперсий амплитуд $\lambda_k^2=\mathbf{D}\eta_k,\ k=\overline{1,m}$, называется частотным спектром колебания $\boldsymbol{\xi}$. Покажем, что спектральная функция

$$F(\theta) = \sum_{\theta_k\leqslant\theta}\lambda_k^2. \tag{13}$$

Другими словами, мера $\mathscr{G}$ сосредоточена в точках $\theta_1,\ldots,\theta_m$ и $\mathscr{G}(\{\theta_k\})=\lambda_k^2$. Очевидно, у такой меры нет плотности относительно меры Лебега.

Установим сначала стационарность процесса. Ясно, что функция среднего $\mathbf{E}\,\xi_n=0,\ \forall\,n\geqslant 1$. Ковариация (для $\forall\,n,t\in\mathbb{Z}$)

$$\mathbf{Cov}(\xi_n,\xi_t) = \mathbf{E}\Big(\sum_{k,j=1}^{m}\eta_k e^{\mathrm{i}\theta_k n}\overline{\eta_j e^{\mathrm{i}\theta_j t}}\Big) = \sum_{k,j=1}^{m}\mathbf{Cov}(\eta_k,\eta_j)e^{\mathrm{i}(\theta_k n-\theta_j t)} =$$

$$= \sum_{k=1}^{m}\lambda_k^2 e^{\mathrm{i}\theta_k(n-t)} = \mathcal{R}(n-t),$$

т.е. зависит только от разности $n-t$. Заметим, что дисперсия процесса $\mathbf{D}\xi_n=\mathcal{R}(0)=\sum_1^m\sigma_k^2$. Интеграл Лебега–Стилтьеса относительно (дискретной) спектральной функции (13), как известно, равен сумме с соответствующими весами значений подынтегральной функции в точках сосредоточения меры:

$$\int\limits_{(-\pi;\pi]} e^{\mathrm{i}\theta n}\, dF(\theta) \;=\; \sum_{k=1}^{m} \lambda_k^2 e^{\mathrm{i}\theta_k n} \;= \mathcal{R}(n),$$

что и требовалось. ◎

✧ **Ортогональные меры. Стохастический интеграл.** Равенство (12), определяющее процесс, может быть обобщено следующим образом. Для фиксированных значений вектора $(\eta_1, \ldots, \eta_m)$ определим функцию множеств (в данном случае меру) на подмножествах интервала $\Theta = (-\pi; \pi]$ соотношением

$$Z(B) = \sum\nolimits_1^m \eta_k \, \dot{\mathrm{I}}\,(\theta_k \in B), \qquad B \subset \Theta. \tag{14}$$

Тогда, очевидно, имеет место равенство

$$\xi_n = \int_\Theta e^{\mathrm{i}\theta n}\, Z(d\theta). \tag{15}$$

Пусть $\mathcal{A}$ — класс интервалов и их конечных объединений пространства $\Theta = (-\pi; \pi]$. Для каждого $B \in \mathcal{A}$ функция $Z(B)$ есть сл.в. на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$, для которой справедливы нижеследующие свойства.

(M$_1$) Математическое ожидание $\mathbf{E}\, Z(B) = 0$.

(M$_2$) Дисперсия $\mathbf{D}\, Z(B) < \infty$.

(M$_3$) Если $\langle B_n \rangle \subset \mathcal{A}$ и $B_n \searrow \varnothing$, то $\mathbf{E}\big|Z(B_n)\big|^2 \to 0$.

(M$_4$) Для любых непересекающихся множеств $B_1 \cap B_2 = \varnothing$:

a) (аддитивность) $Z(B_1 + B_2) = Z(B_1) + Z(B_2)$;

b) ковариация $\mathbf{Cov}\big(Z(B_1), Z(B_2)\big) \;=\; \mathbf{E}\big[Z(B_1)\overline{Z(B_2)}\,\big] = 0$, т.е. как элементы гильбертова пространства эти сл.в. ортогональны.

Определение. Функция $Z(B) = Z(B, \omega)$, $B \in \mathcal{A}$, $\omega \in \Omega$, заданная на классе подмножеств $\mathcal{A}$ пространства Θ и вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$, называется *ортогональной стохастической мерой*, если для $\forall\, B \in \mathcal{A}$ она измерима по Борелю и удовлетворяет свойствам (M$_1$) – (M$_4$).

384| **Лемма.** *Пусть $Z(B) = Z(B, \omega)$, $B \in \mathcal{A}$, $\omega \in \Omega$, — ортогональная стохастическая мера на кольце $\mathcal{A}$.*

I) *Мера Z σ-аддитивна в среднеквадратическом смысле:*

$$\lim_{n\to\infty} \mathbf{E}\Big|Z\Big(\biguplus\nolimits_1^\infty B_k\Big) - \sum\nolimits_1^n Z(B_k)\Big|^2 = 0, \quad \langle B_k \rangle_1^\infty,\ \biguplus\nolimits_1^\infty B_k \subset \mathcal{A}.$$

II) *Функция $M_Z(B) = \mathbf{D}\, Z(B) = \mathbf{E}\big|Z(B)\big|^2$, $B \in \mathcal{A}$, есть (детерминиро-*

ванная) конечная мера на $\mathcal{A}$.

III) *Существует единственное продолжение* M_Z *на минимальную* σ*-алгебру* $\sigma(\mathcal{A})$, *порождённую* $\mathcal{A}$.

⊨ Так как $\sum\limits_{k>n} B_k \searrow \varnothing$, то I) следует из (M_3), (M_4).

II) Равенство $M_Z(\varnothing) = 0$ следует из условия $(\mathsf{a}, \mathsf{M}_4)$ и (почти наверное) конечности Z. Непрерывность меры буквально прописана в свойстве (M_3). Аддитивность есть следствие аддитивности $(\mathsf{a}, \mathsf{M}_4)$ и ортогональности $(\mathsf{b}, \mathsf{M}_4)$ меры Z:

$$M_Z(B_1 + B_2) \;=\; \mathbf{D}\big[Z(B_1) + Z(B_2)\big] \;=\; \mathbf{D}Z(B_1) + \mathbf{D}Z(B_2)$$

в силу хорошо известного свойства дисперсии суммы некоррелированных сл.в. Это объясняет, кстати, причину введения условия ортогональности. Таким образом, M_Z есть аддитивная непрерывная функция множеств на кольце, т.е. мера (σ-аддитивная функция множеств, см. 468, стр. 430).

III) Следует из теоремы Каратеодори и предыдущего пункта. ⫤

Определение. Мера $\mathfrak{m}(B)$, $B \in \sigma(\mathcal{A})$, называется *структурной функцией* (*мерой*) ортогональной стохастической меры Z, если $\mathfrak{m}(B) = \mathbf{D}\big[Z(B)\big] = \mathbf{E}\,|Z(B)|^2$ для $\forall\, B \in \mathcal{A}$.

Наша цель — описать представление, подобное (15), для любого стационарного процесса. Но прежде необходимо придать точный смысл понятию интеграла от детерминированной функции по случайной мере. Рассмотрим гильбертово пространство $L_2(\Theta, \sigma(\mathcal{A}), \mathfrak{m})$ всех комплекснозначных функций на Θ, интегрируемых с квадратом относительно структурной функции $\mathfrak{m}$:

$$h \in L_2(\Theta, \sigma(\mathcal{A}), \mathfrak{m}) \quad \Leftrightarrow \quad \|h\|^2_\Theta := \int\limits_\Theta h(\theta)\overline{h(\theta)}\,\mathfrak{m}(d\theta) < \infty. \tag{16}$$

Рассмотрим простую функцию $h(\theta) = \sum_1^N b_k\, \dot{\mathrm{I}}_{B_k}(\theta)$ с комплексными b_k и непересекающимися множествами $B_k \in \mathcal{A}$, $k = 1, \ldots, N$, образующими разбиение Θ: $\biguplus_1^N B_k = \Theta$. Обратим внимание на то, что эта простая функция измерима относительно $\mathcal{A}$. Чтобы подчеркнуть это обстоятельство, будем называть такую функцию $\mathcal{A}$-простой. Определим для $\mathcal{A}$-простой функции интеграл по мере Z с помощью формулы

$$J(h) \;=\; \int_\Theta h(\theta)\, Z(d\theta) \;:=\; \sum\nolimits_1^n b_k Z(B_k).$$

По построению $J(h)$ есть сл.величина на $(\Omega, \mathscr{F}, \mathbf{P})$. Её математическое ожидание $\mathbf{E}\, J(h) = 0$, а дисперсия

$$\mathbf{D}\, J(h) \;=\; \sum\nolimits_1^N |b_k|^2 \mathfrak{m}(B_k) \;=\; \int_\Theta |h(\theta)|^2\, \mathfrak{m}(d\theta) \;<\; \infty$$

в силу свойств меры Z. Таким образом, сл.в. $J(h)$ принадлежит гильбертову пространству $L_2(\Omega, \mathscr{F}, \mathbf{P})$ всех комплесных сл.в. с конечной дисперсией. Норма $\|\cdot\|_\Omega$ в пространстве $L_2(\Omega, \mathscr{F}, \mathbf{P})$ определяется скалярным произведением $(\xi, \eta) := \mathbf{Cov}(\xi, \eta) = \mathbf{E}\left[\xi\overline{\eta}\right]$. Следовательно, норма интеграла (см. (16))

$$\|J(h)\|_\Omega \;=\; \sqrt{\mathbf{D}\, J(h)} \;=\; \|h\|_\Theta.$$

Известно, что для $\forall\, h \in L_2(\Theta, \sigma(\mathcal{A}), \mathfrak{m})$ существует последовательность $\mathcal{A}$-простых функций h_n, сходящаяся к h по норме:

$$\|h_n - h\|^2_\Theta \underset{n}{\to} 0.$$

В частности, эта последовательность $\mathcal{L}_2$-фундаментальна:

$$\|h_n - h_k\|^2_\Theta \to 0, \quad n, k \to \infty.$$

В силу предыдущих утверждений отсюда следует, что и последовательность интегралов от простых функций будет фундаментальной уже в $L_2(\Omega, \mathscr{F}, \mathbf{P})$:

$$\|J(h_n) - J(h_k)\|^2_\Omega \;=\; \|h_n - h_k\|^2_\Theta \;\to\; 0, \quad n, k \to \infty.$$

Как отмечалось, пространство $L_2(\Omega, \mathscr{F}, \mathbf{P})$ гильбертово, т.е. полное. Поэтому при $n \to \infty$ существует предел $J(h_n) \xrightarrow{\mathcal{L}_2} J(h)$.

Определение. Сл.в. $J(h)$ называется *стохастическим интегралом* на Θ от функции $h \in L_2(\Theta, \sigma(\mathcal{A}), \mathfrak{m})$ по ортогональной стохастической мере Z со структурной функцией $\mathfrak{m}$ и обозначается

$$J(h) = \int_\Theta h(\theta)\, Z(d\theta).$$

385| ⚠ Стохастический интеграл определяется с точностью до множества меры нуль. Если бы этот интеграл определялся для каждой функции h единственным образом (без добавки п.н.), то, подставляя вместо h индикаторную функцию $\dot{\mathtt{I}}_B(\theta)$ любого множества $B \in \sigma(\mathcal{A})$, мы могли бы утверждать, что функция $J(\dot{\mathtt{I}}_B)$, $B \in \sigma(\mathcal{A})$, задаёт продолжение меры Z с кольца $\mathcal{A}$ на σ-алгебру $\sigma(\mathcal{A})$. Увы, ввиду отмеченного обстоятельства не всё так просто. Аналогичная проблема возникает и при построении условного распределения. Здесь также можно доказать, что существует такое продолжение меры Z до ортогональной меры $\widetilde{Z}$, что $\widetilde{Z}(B) = J(\dot{\mathtt{I}}_B)$ (п.н.), $\forall B \in \sigma(\mathcal{A})$.

Непосредственно из определения следуют основные свойства стохастического интеграла.

386| **Теорема.** *Пусть* $h, g, h_n \in L_2(\Theta, \sigma(\mathcal{A}), \mathfrak{m})$, *где* $\mathfrak{m}$ *— структурная функция ортогональной стохастической меры* Z *на* $(\Theta, \mathcal{A}) \times (\Omega, \mathscr{F}, \mathbf{P})$. *Тогда*

1) $\mathbf{E}\, J(h) = 0$;

2) $J(ah + bg) = aJ(h) + bJ(g)$ [$\mathbf{P}$-п.н.], $\quad \forall\, a, b \in \mathbb{C}$;

3) *скалярное произведение (в* $L_2(\Omega, \mathscr{F}, \mathbf{P})$*)*

$$(J(h), J(g))_\Omega \ = \ \mathbf{Cov}(J(h), J(g)) \ = \ \int_\Theta h\overline{g}\, d\mathfrak{m};$$

в частности $\mathbf{D}\, J(h) = \int_\Theta |h(\theta)|^2\, \mathfrak{m}(d\theta)$;

4) *если* $h_n \xrightarrow{\mathcal{L}_2} h$ *в* $L_2(\Theta, \sigma(\mathcal{A}), \mathfrak{m})$, *то имеет место сходимость по норме* $L_2(\Omega, \mathscr{F}, \mathbf{P})$: $\quad J(h_n) \xrightarrow{\mathcal{L}_2} J(h)$.

387| Пример. Легко видеть, что структурная функция ортогональной стохастической меры (14)

$$\mathfrak{m}(B) \ = \ \sum\nolimits_1^m \lambda_k^2\, \dot{\mathtt{I}}\,(\theta_k \in B).$$

Эта мера сосредоточена в точках θ_k с весами λ_k^2, $k = \overline{1, m}$, (стохастическая мера Z сосредоточена в этих же точках только со случайными весами). Интеграл от любой функции относительно такой (дискретной) меры равен

$$\int_\Theta h(\theta)\, \mathfrak{m}(d\theta) = \sum\nolimits_1^m \lambda_k^2 h(\theta_k).$$

Поэтому $\|h_n - h\|_\Theta^2 \underset{n}{\to} 0$ т. т. т. когда $h_n(\theta_k) \underset{n}{\to} h(\theta_k)$ для $\forall\, k = 1, \ldots, m$. Зафиксируем $s \in \langle 0, \pm 1, \pm 2, \ldots \rangle$ и рассмотрим функцию $h(\theta) = e^{\mathrm{i}\theta s}$. Выберем произвольное разбиение множества $\Theta = (-\pi; \pi] = \biguplus_1^m B_k$, при котором $\theta_k \in B_k$, $k = \overline{1, m}$. Тогда при любом $n \geqslant 1$ для простой функции

$$h_n(\theta) = \sum\nolimits_{k=1}^m e^{\mathrm{i}\theta_k s}\, \dot{\mathtt{I}}\,(\theta \in B_k)$$

норма разности $\|h_n - h\|_\Theta = 0$, т.е. можно сказать, что $h_n \xrightarrow{\mathcal{L}_2} h$. Интеграл от h_n по мере Z, совпадающий с интегралом от h, по определению равен $\int_\Theta h_n\, dZ = \sum_1^m \eta_k e^{\mathrm{i}\theta_k s}$, что совпадает с (12). ◎

Теперь мы можем сформулировать утверждение о спектральном представлении стационарных процессов ([5, стр. 240]).

388| **Теорема.** *Пусть* $\left\langle \xi_n,\ n = 0, \pm 1, \pm 2, \ldots \right\rangle$ *— стационарная в широком смысле последовательность комплекснозначных сл.в., определённых на вероятностном пространстве* $(\Omega, \mathscr{F}, \mathbf{P})$. *Тогда на этом же вероятностном пространстве существует ортогональная стохастическая мера* Z, *заданная на борелевской* σ*-алгебре* $\mathscr{B}(\Theta)$ *интервала* $\Theta = (-\pi; \pi]$, *такая, что*

$$\xi_n = \int_\Theta e^{i\theta n} Z(d\theta), \qquad n \in \{0, \pm 1, \pm 2, \ldots\}. \tag{17}$$

При этом ковариационная функция процесса

$$R(s,t) = \int_\Theta e^{i(t-s)\theta} \mathfrak{m}(d\theta), \qquad s,t \in \{0, \pm 1, \pm 2, \ldots\},$$

где спектральная мера $\mathfrak{m}$ *совпадает со структурной функцией* Z.

Аналогичное представление имеет место и для стационарных сл.проц. с непрерывным временем и непрерывными (в среднем квадратическом) траекториями. Отметим, что стационарный процесс непрерывен в среднем квадратическом, если его ковариационная функция непрерывна в нуле.

389| **Лемма.** (?!) *Стационарный в широком смысле процесс* $\left\langle \xi(t),\ t \in \mathbb{R}^1 \right\rangle$ *непрерывен в среднем квадратическом, т.е.*

$$\lim_{s \to t} \mathbf{E} \left[\xi(s) - \xi(t)\right]^2 = 0 \text{ для } \forall\, t \in \mathbb{R}^1,$$

т. т. т. когда ковариационная функция $\mathcal{R}(u) = \mathbf{Cov}\left(\xi(t), \xi(t+u)\right)$ *непрерывна в точке* $u = 0$.

390| **Теорема.** [*Крамер, Бохнер, Хинчин.*] *Пусть центрированный стационарный в широком смысле процесс* $\left\langle \xi(t),\ t \in \mathbb{R}^1 \right\rangle$ *задан на вероятностном пространстве* $(\Omega, \mathscr{F}, \mathbf{P})$. *Если ковариационная функция процесса непрерывна в нуле, то на* $(\Omega, \mathscr{F}, \mathbf{P})$ *существует ортогональная стохастическая мера* Z, *определённая на борелевских подмножествах* $\mathscr{B}(\mathbb{R}^1)$, *такая, что*

$$\xi(t) = \int_\mathbb{R} e^{it\theta} Z(d\theta), \qquad t \in \mathbb{R}^1. \tag{18}$$

При этом для ковариации этого процесса справедливо

$$\mathcal{R}(t) = \int_\mathbb{R} e^{it\theta} \mathfrak{m}(d\theta), \qquad t \in \mathbb{R}^1,$$

где $\mathfrak{m}(\cdot) = \mathbf{E}\, Z^2(\cdot)$ *— структурная функция* Z.

✧ Эргодическое свойство. Для стационарных процессов с дискретным временем справедлив аналог закона больших чисел.

391 **Теорема.** *Пусть $\xi(t)$ — стационарный в широком смысле процесс со средним $\mathbf{E}\xi(t) = \mu$, $t \in \mathbb{N}_0$. Если спектральная мера $\mathscr{G}$ процесса не имеет атома в нуле — $\mathscr{G}(\{0\}) = 0$, то*

$$\frac{1}{n}\sum_{k=0}^{n-1}\xi_k \xrightarrow{\mathcal{L}_2} \mu \quad \text{при} \quad n \to \infty.$$

⤃ Воспользуемся спектральным представлением Герглотца 381, стр. 333, для ковариации стационарного процесса. Имеем

$$\mathbf{D}\Big(\frac{1}{n}\sum_{k=0}^{n-1}\xi_k\Big) = \frac{1}{n^2}\sum_{k,j=0}^{n-1}\mathbf{Cov}(\xi_k,\xi_j) = \frac{1}{n^2}\int\limits_{(-\pi;\pi]}\Big(\sum_{k,j=0}^{n-1}e^{\mathrm{i}(k-j)\theta}\Big)\mathscr{G}(d\theta) =$$

$$= \frac{1}{n^2}\int\limits_{(-\pi;\pi]}\Big(\sum_{k=0}^{n-1}e^{\mathrm{i}k\theta}\cdot\sum_{j=0}^{n-1}e^{-\mathrm{i}j\theta}\Big)\mathscr{G}(d\theta) =$$

$$= \frac{1}{n^2}\int\limits_{(-\pi;\pi]}\frac{1-e^{\mathrm{i}n\theta}}{1-e^{\mathrm{i}\theta}}\frac{1-e^{-\mathrm{i}n\theta}}{1-e^{-\mathrm{i}\theta}}\mathscr{G}(d\theta) = \int\limits_{(-\pi;\pi]}\frac{\sin^2(n\theta/2)}{n^2\sin^2(\theta/2)}\mathscr{G}(d\theta).$$

Подынтегральное выражение здесь не превосходит 1, и при $n \to \infty$ предел этого выражения равен 0 для всех $\theta \neq 0$. По теореме Лебега 540, стр. 482, отсюда следует, что если (и только если) мера $\mathscr{G}(\{0\}) = 0$, то предел дисперсии искомого среднего арифметического равен нулю. ⤆

Условие теоремы выполняется, например, если процесс обладает спектральной плотностью.

Заметим, что μ есть среднее каждого сечения по случайной составляющей процесса ω. Из приведённого утверждения следует, что это среднее можно приближённо вычислять, усредняя не по большому числу реализаций какого-то одного конкретного сечения, а по реализациям сечений из одной траектории процесса. Такое свойство называется *эргодичностью* (по среднему) процесса. Широко применяется в физике, при этом предполагается, что оно должно выполняться для всех моментов случайного процесса.

Для процессов с непрерывным временем среднее арифметическое может как стабилизироваться к детерминированной константе, так и оставаться случайной величиной. Первая возможность осуществляется, если, например, при фиксированных $a, b \in \mathbb{R}^1$ рассмотреть среднее

$$\frac{1}{n}\sum\nolimits_{k=0}^{n-1}\xi(ak+b).$$

В этом случае закон больших чисел следует из того, что процесс с дискретным временем $\left\langle\, \xi(at+b),\ t \in \mathbb{N}_0 \right\rangle$ будет также стационарным.

Вторая из описанных выше возможностей реализуется, если сечения, в которых подсчитывается среднее, выбираются достаточно «плотно». Для упрощения описания предположим, что процесс непрерывен не только в среднем квадратическом, но и все его траектории непрерывны всюду на $\mathbb{T}$. Тогда для любой траектории можно вычислить интеграл Римана по любому конечному отрезку. Рассмотрим отрезок $[0;t]$ и выберем для фиксированного $n \geqslant 2$ точки $t_k = \dfrac{k}{n}t,\ k = 0, \ldots, n-1$. Интегральная сумма от ξ

$$S_n \;=\; \frac{t}{n}\sum\nolimits_{k=0}^{n-1}\xi\Big(\frac{kt}{n}\Big) \;=\; t\overline{X}_n.$$

По определению интеграла Римана, для непрерывных траекторий процесса

$$\overline{X}_n \;=\; \frac{1}{n}\sum\nolimits_{k=0}^{n-1}\xi\Big(\frac{kt}{n}\Big) \;\underset{n}{\to}\; \frac{1}{t}\int_0^t \xi(u)\,du.$$

Для стационарного процесса со средним μ математическое ожидание $\mathbf{E}\,\overline{X}_n = \mu$. Если ковариационная функция процесса $\mathcal{R}(t)$ непрерывна в нуле (а значит, и всюду), то дисперсия (по определению интеграла Римана)

$$\mathbf{D}S_n \;=\; \frac{t^2}{n^2}\sum_{k,j=0}^{n-1}\mathcal{R}\Big(\frac{kt}{n}-\frac{jt}{n}\Big) \;\to\; \int\limits_0^t\int\limits_0^t \mathcal{R}(u-v)\,dudv.$$

В случае, когда последний интеграл больше нуля, мы получаем, что предельное значение среднего арифметического $\overline{X}_n$ также будет невырожденной сл.в. со средним μ и дисперсией $\int_0^t\int_0^t \mathcal{R}(u-v)\,dudv/t^2$. Проведённые построения позволяют сформулировать интегральный вариант свойства эргодичности (по среднему) стационарных процессов с непрерывным временем.

392| **Теорема.** *Пусть $\xi(t),\ t \geqslant 0,$ — стационарный в широком смысле процесс со средним $\mathbf{E}\xi(t) = \mu$ и ковариацией $\mathbf{Cov}(\xi(s), \xi(t)) = \mathcal{R}(t-s),\ 0 \leqslant \leqslant s \leqslant t$. Если при $t \to \infty$ предел $\lim\limits_{t\to\infty}\frac{1}{t^2}\int_0^t\int_0^t \mathcal{R}(u-v)\,dudv = 0,$ то*

$$\frac{1}{t}\int_0^t \xi(u)\,du \;\xrightarrow{\mathcal{L}_2}\; \mu.$$

Условие теоремы выполняется, если, например, $\lim\limits_{t\to\infty}\mathcal{R}(t) = 0$. Как и для последовательностей, можно показать, что это условие выполняется, когда спектральная мера, соответствующая $\mathcal{R}$ (см. 390), не имеет атома в нуле.

✧ **Ковариация процесса с независимыми приращениями.** Наконец, изучим структуру ковариации процесса $\langle\, \xi(t),\ t \geqslant 0 \,\rangle$ с независимыми приращениями. Поскольку математическое ожидание не имеет принципиального значения при анализе свойств ковариации, будем считать, что функция среднего $\mu(t) \equiv 0$. Рассмотрим три числа $s < t < u$. Тогда ковариации приращений

$$\begin{aligned}
\mathbf{Cov}\,\big(\xi(s), \xi(t) - \xi(s)\big) &= \mathbf{E}\big[\xi(s)\,(\xi(t) - \xi(s))\big] = R(s,t) - R(s,s),\\
\mathbf{Cov}\,\big(\xi(u) - \xi(t), \xi(t) - \xi(s)\big) &= \mathbf{E}\,\big[(\xi(t) - \xi(s))(\xi(u) - \xi(t))\big] =\\
&= R(t,u) - R(s,u) + R(s,t) - R(t,t).
\end{aligned}$$

Если приращения процесса независимы, то из первого равенства

$$R(s,t) = R(s,s) = \mathbb{V}\big(\min(s,t)\big), \qquad \forall\, s,t \in \mathbb{T}, \tag{19}$$

где $\mathbb{V}$ — дисперсионная функция процесса. При выполнении этого условия, очевидно, и $\mathbf{Cov}\,\big(\xi(u) - \xi(t), \xi(t) - \xi(s)\big) = 0$. Заметим, что в этом случае $\mathbf{D}\big(\xi(t) - \xi(s)\big) = \mathbb{V}(t) + \mathbb{V}(s) - 2R(s,t) = \mathbb{V}(t) - \mathbb{V}(s) \geqslant 0$ для $s < t$. Другими словами, у процессов с независимыми приращениями дисперсия не убывает.

Как уже отмечалось, равенство нулю ковариации центрированных сл.в. геометрически интерпретируется как ортогональность этих сл.в. Отсюда

Определение. Сл.проц. $\langle\, \xi(t) \,\rangle_{t \geqslant 0}$ с нулевой функцией среднего $\mu(t) = \mathbf{E}\,\xi(t) \equiv 0$ имеет *ортогональные приращения*, если

$$\mathbf{Cov}\,(\xi(v) - \xi(u), \xi(t) - \xi(s)) = \mathbf{E}\,\big[(\xi(v) - \xi(u))(\xi(t) - \xi(s))\big] = 0$$

для любых $0 \leqslant s \leqslant t \leqslant u \leqslant v$.

Центрированный гильбертов процесс с независимыми приращениями имеет ортогональные приращения. Обратное, конечно, не всегда верно.

393| **Лемма.** (?!) *Если ковариация процесса* $R(s,t) = \mathbb{V}\big(\min(s,t)\big)$, *где* $\mathbb{V}(t)$, $t \geqslant 0$, *— возрастающая положительная функция, то процесс с нулевой функцией среднего имеет ортогональные приращения. Обратно, если процесс выходит из нуля —* $\xi(0) = 0$, *имеет нулевую функцию среднего и ортогональные приращения, то ковариация процесса* $R(s,t) = \mathbb{V}\big(\min(s,t)\big)$ *с возрастающей дисперсионной функцией* $\mathbb{V}(t) = \mathbf{D}\xi(t)$, $t \geqslant 0$.

§2. Гауссовские процессы

Необходимость отдельного изучения процессов с нормальным распределением сечений объясняют следующие соображения. Рассмотрим процесс $\left\langle\, \xi_t,\ t \in [0;b] \subset \mathbb{R}^1 \,\right\rangle$ с непрерывным временем и независимыми приращениями. Разобьём интервал $[0;t]$ на n частей точками $0 = t_0 < t_1 < \ldots < t_n = t$. Тогда имеем равенство

$$\xi(t) - \xi(0) = \sum\nolimits_{k=1}^{n} \big(\xi(t_k) - \xi(t_{k-1})\big),$$

в котором все слагаемые независимы. В соответствии с центральной предельной теоремой можно ожидать, что при определённых условиях на приращения $\xi(t_k) - \xi(t_{k-1})$ распределение процесса будет нормальным.

Определение. Сл.проц. $\boldsymbol{\xi} = \left\langle\, \xi_t,\ t \in \mathbb{T} \,\right\rangle$ называется *гауссовским*, если его конечномерные распределения нормальны, т.е. хар.ф. процесса для любого $\tau = (t_1, \ldots, t_n) \in \mathcal{T}, \vec{u} \in \mathbb{R}^n$, представляется формулой

$$\varphi_{\boldsymbol{\xi}}(\vec{u} \mid \tau) = \exp\Big\{ \mathrm{i} \sum_{k=1}^{n} u_k \mu(t_k) - \frac{1}{2} \sum_{k,j=1}^{n} R(t_k, t_j) u_k u_j \Big\} \qquad (20)$$

с функцией среднего $\mu(t)$ и ковариацией $R(s,t)$, $s, t \in \mathbb{T}$.

Если функция $R(s,t)$ неотрицательно определена, то для $\forall\, \tau = (t_1, \ldots, t_n)$ выражение (20) представляет собой хар.ф. (например, линейного преобразования $\sqrt{R(\tau)}\, \vec{\eta} + \vec{\mu}(\tau)$ сл.вектора $\vec{\eta} = (\eta_1, \ldots, \eta_n)$ независимых стандартных нормальных $\mathcal{N}(0{,}1)$ сл.в.). Легко проверить, что семейство таких хар.ф. будет задавать согласованное и симметричное семейство конечномерных распределений. Отсюда, согласно теореме Колмогорова, справедлива

394| **Лемма.** *Если функция $R(s,t)$, $s, t \in \mathbb{R}^1$, неотрицательно определена, то существует гауссовский процесс с ковариационной функцией $R(s,t)$.*

Соображения, приведённые в начале параграфа, формализует следующая теорема Дуба (см. доказательство [5], стр. 332).

395| **Теорема.** [*Дж. Дуб.*] *Если сл.проц. $\left\langle\, \xi(t),\ t \in [a;b] \subset \mathbb{R}^1 \,\right\rangle$ с независимыми приращениями имеет почти наверное непрерывные траектории, то $\left\langle\, \xi(t) - \xi(a),\ t \in [a;b] \,\right\rangle$ — гауссовский процесс.*

396| Пример. Пусть сл.вектор $\vec{\eta} = (\eta_1, \eta_2) \backsim \mathcal{N}_2(\vec{\mu}, \boldsymbol{\Sigma})$ и $h_1(t)$, $h_2(t)$ — произвольные детерминированные функции. Рассмотрим процесс

$\xi(t) = \eta_1 h_1(t) + \eta_2 h_2(t) = \vec{h}(t)^\flat \, \vec{\eta}$; здесь $\cdot^\flat$ — операция транспонирования. Очевидно, функция среднего этого процесса равна $\mu_1 h_1(t) + \mu_2 h_2(t)$. Чтобы найти конечномерные распределения, выберем произвольные $t_1, \ldots, t_n$. Тогда вектор-столбец $(\xi(t_1), \ldots, \xi(t_n))^\flat$ можно представить в виде $(\xi(t_1), \ldots, \xi(t_n))^\flat = H^\flat \, \vec{\eta}$ с $(2 \cdot n)$-матрицей $H = (\vec{h}(t_1), \ldots, \vec{h}(t_n))$. Таким образом, любое конечномерное распределение нормально с ковариационной матрицей $H^\flat \boldsymbol{\Sigma} H \geqslant 0$. Как известно, ранг произведения матриц не превосходит рангов сомножителей. Поэтому при $n > 2$ все n-мерные распределения вырождены (сосредоточены в некотором двумерном подпространстве $\mathbb{R}^n$) и не имеют плотности.

Если одна из функций h_1 или h_2 терпит разрыв, то почти все траектории процесса будут разрывны, причём даже не существует непрерывных модификаций этого процесса. ◎

Поскольку распределение гауссовского процесса полностью определяется функцией среднего и ковариационной функцией, для такого процесса понятия стационарности в узком и широком смысле эквивалентны.

397| **Лемма.** *Гауссовский процесс стационарен в узком смысле т. т. т. когда его функция среднего постоянна:* $\mu(t) = m, \ \forall t \in \mathbb{T}$, *а ковариационная функция зависит только от разности параметров:* $R(s,t) = \mathcal{R}(|t-s|), \ s, t \in \mathbb{T}$.

⫤ Применяя (20), получаем, что хар.ф. конечномерного распределения стационарного в широком смысле гауссовского процесса равна $(\vec{u} \in \mathbb{R}^n)$

$$\varphi_{\boldsymbol{\xi}}(\vec{u} \mid \tau) = \exp\Big\{ \mathrm{i} m \sum\nolimits_1^n u_j - \frac{1}{2} \sum\nolimits_{k,j=1}^n \mathcal{R}(|t_k - t_j|) u_k u_j \Big\},$$

т.е. не изменяется при замене $t_k \dashrightarrow t_k + h, \ k = 1, \ldots, n$. ⫣

Учтя вид ковариации (19) для процесса с независимыми приращениями, можно сформулировать следующее утверждение.

398| **Лемма.** *Гауссовский процесс с дисперсионной функцией* $\mathbb{V}(t), \ t \in \mathbb{R}^1$, *имеет независимые приращения т. т. т. когда его ковариационная функция* $R(s,t) = \mathbb{V}\big(\min(s,t)\big)$. *При этом для любых* $s < t$ *приращение* $\xi(t) - \xi(s)$ *имеет нормальное распределение с дисперсией* $\mathbb{V}(t) - \mathbb{V}(s)$.

Обратно, рассмотрим последовательность чисел $0 = b_0 < b_1 < \ldots < b_n$. По известному способу вычисления определителя

$$\begin{vmatrix} b_1 & b_1 & b_1 & \dots & b_1 \\ b_1 & b_2 & b_2 & \dots & b_2 \\ b_1 & b_2 & b_3 & \dots & b_3 \\ & & \dots & & \\ b_1 & b_2 & b_3 & \dots & b_n \end{vmatrix} = \begin{vmatrix} b_1 & 0 & 0 & \dots & 0 \\ 0 & b_2-b_1 & b_2-b_1 & \dots & b_2-b_1 \\ 0 & b_2-b_1 & b_3-b_1 & \dots & b_3-b_1 \\ & & \dots & & \\ 0 & b_2-b_1 & b_3-b_1 & \dots & b_n-b_1 \end{vmatrix} =$$

$$= \begin{vmatrix} b_1 & 0 & 0 & \dots & 0 \\ 0 & b_2-b_1 & 0 & \dots & 0 \\ 0 & 0 & b_3-b_2 & \dots & b_3-b_2 \\ & & \dots & & \\ 0 & 0 & b_3-b_2 & \dots & b_n-b_2 \end{vmatrix} = \dots = \prod_{k=1}^{n}(b_k - b_{k-1}).$$

Откуда следует положительность всех главных миноров такой матрицы, что по критерию Сильвестра гарантирует её положительную определённость. Таким образом, справедлива

399| **Лемма.** *Пусть* $\mathbb{V}(t)$, $t \geqslant 0$, *— возрастающая функция. Тогда существует гауссовский процесс с ковариацией* $R(s,t) = \mathbb{V}\big(\min(s,t)\big)$. *Этот процесс имеет независимые приращения, и все его конечномерные распределения абсолютно непрерывны относительно меры Лебега.*

Особенно важен случай, когда $R(s,t) = \lambda^2 s$ для $s < t$.

400| **Теорема.** *Пусть дисперсионная функция гауссовского процесса с независимыми приращениями* $\mathbb{V}(t) = \lambda^2 t$, *начальная сл.в.* $\xi(0) = 0$ (п.н.) *и функция среднего* $\mu(t) \equiv 0$, *тогда*

a) *плотность* n*-мерного распределения процесса*

$$f(\vec{x} \mid \tau) \;=\; \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi\lambda^2 |t_k - t_{k-1}|}} \exp\Big\{ -\frac{1}{2\lambda^2}\frac{(x_k - x_{k-1})^2}{|t_k - t_{k-1}|}\Big\}, \tag{21}$$

$\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, $\tau = (t_1, \dots, t_n) \in \mathcal{T}$, $t_0 = 0, x_0 = 0$;

b) *условное распределение* $\xi(t)$ *при фиксированном значении* $\xi(s) = y$ $(s < t)$ *нормально* $\mathcal{N}(y, \lambda^2(t-s))$.

$\Rrightarrow$ a) Пусть $0 = t_0 < t_1 < \dots < t_n$. Так как процесс имеет независимые приращения, то ввиду 398 вектор $\vec{\eta} = (\eta_1, \dots, \eta_n)$ с

$$\eta_k = \xi(t_k) - \xi(t_{k-1}), \quad k = 1, \dots, n,$$

распределён по n-мерному нормальному закону с плотностью

$$f_\eta(\vec{y}) \;=\; \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi\lambda^2 |t_k - t_{k-1}|}} \exp\Big\{ -\frac{y_k^2}{2\lambda^2 |t_k - t_{k-1}|}\Big\}.$$

В соответствии с правилом вычисления плотности линейного преобразования, плотность $(\xi(t_1), \ldots, \xi(t_n))$ в точке $\vec{x} = (x_1, \ldots, x_n)$ получается отсюда подстановкой $y_k = (x_k - x_{k-1})$ с $x_0 = 0$.

b) Ввиду (21) условная плотность для $x \in \mathbb{R}^1$ представляется так:

$$f_t\big(x \,|\, \xi(s) = y\big) \;=\; \frac{f(y, x \,|\, s, t)}{f(y \,|\, s)} \;=$$

$$= \;\frac{1}{\sqrt{2\pi\lambda^2 s}} \frac{1}{\sqrt{2\pi\lambda^2 |t-s|}} \sqrt{2\pi\lambda^2 s}\, \exp\Big\{ -\frac{y^2}{2\lambda^2 s} - \frac{(x-y)^2}{2\lambda^2 (t-s)} + \frac{y^2}{2\lambda^2 s} \Big\},$$

т.е. совпадает с ф.пл. нормального $\mathcal{N}(y, \lambda^2(t-s))$ закона. ⩤

Определения. Гауссовский процесс $\boldsymbol{w} = \big\langle\, w(t),\ t \geqslant 0,\ w(0) = 0 \big\rangle$ с математическим ожиданием $\mu_w(t) = 0$ и ковариацией $R_{\boldsymbol{w}}(s,t) = \lambda^2 \min(s,t),\ s, t \geqslant 0$, (плотностью конечномерных распределений (21)) называется *винеровским процессом* (или *процессом броуновского движения*). Коэффициент λ^2 называется *коэффициентом диффузии*; винеровский процесс с коэффициентом $\lambda^2 = 1$ называется *стандартным*.

✧ Свойства винеровского процесса.

Для процесса Винера четвёртый момент приращения

$$\mathbf{E}\big[w(t) - w(s)\big]^4 \;=\; 3\big(\mathbf{D}\big[w(t) - w(s)\big]\big)^2 \;=\; 3\lambda^4 (t-s)^2.$$

В силу теоремы Колмогорова 379, стр. 328, для $\forall T > 0$ существует непрерывная модификация $\big\langle\, w(t),\ t \in [0;T] \,\big\rangle$. Ввиду этого требование непрерывности траекторий винеровского процесса часто включают составной частью его определения. Мы поступим так же.

Существование непрерывной модификации для процессов, определённых для всех $t \geqslant 0$, можно показать, например, с помощью «склейки» независимых непрерывных винеровских процессов на отрезках $[n; n+1),\ n \geqslant 0$. Кроме того, существуют конкретные реализации процесса броуновского движения с помощью последовательностей независимых стандартных нормальных сл.в. $\xi_n \backsim \mathcal{N}(0,1),\ n \geqslant 0$. Винер и Пэли предложили следующую реализацию:

$$\xi_0 t + \sqrt{2} \sum_{n=1}^{\infty} \sum_{k=2^{n-1}}^{2^n - 1} \frac{\sin(\pi k t)}{\pi k} \xi_k.$$

По поводу других реализаций см. [5], гл. III, [27], гл. II.

✦ Приращение $w(t) - w(s) \backsim \mathcal{N}(0, \lambda^2 |t-s|)$, т.е. его распределение зависит только от разности $t-s$. Процессы с таким свойством приращений называются *однородными*. Можно показать, что выходящий из нуля однородный процесс с ортогональными приращениями и непрерывной дисперсионной функцией имеет ковариационную функцию вида $R(s,t)=\lambda^2 \min(s,t)$.

✦ Учитывая утверждение теоремы Дуба 395, можно охарактеризовать винеровский процесс как

выходящий из нуля, центрированный, однородный гильбертов процесс с независимыми приращениями и непрерывными траекториями.

✦ В 1827 году ботаник R. Brown наблюдал через микроскоп непрерывные хаотичные движения частиц цветочной пыльцы в водном растворе и вначале интерпретировал это движение как наличие жизни у пыльцы. Вскоре было высказано мнение, что это движение, названное впоследствии броуновским, вызвано столкновениями частиц с молекулами жидкости, обусловленными тепловым движением молекул.

Рассмотрим только одну координату пространственного движения частицы. Несколько упрощая ситуацию, предположим, что за время t от начала наблюдения через равные промежутки времени $\Delta=\frac{t}{n}$ частица испытывает n соударений с молекулами воды, каждое из которых приводит к смещению частицы влево или вправо от её текущего положения на малую величину ε. Итак, можно считать, что положение частицы в момент t на оси представляется в виде суммы смещений:

$$X(t)=\sum\nolimits_1^n \varepsilon\,\xi_j,$$

где $\xi_1,\ldots,\xi_n$ — независимые бинарные сл.в., принимающие значения ± 1 с равными вероятностями, $t=n\Delta$. Процесс $X(t)$ есть процесс случайного блуждания (см. пример 373, стр. 325), т.е. он имеет независимые приращения и однороден. По центральной предельной теореме распределение сл.в. $X(t)$ приближённо может быть описано нормальным законом с нулевым средним и дисперсией $n\varepsilon^2=t\left(\varepsilon^2/\Delta\right)$. Поведение $X(t)$ адекватно поведению броуновской частицы, если отношение $\lambda^2=\varepsilon^2/\Delta$ не слишком мало́ и не слишком велико́ — при $\varepsilon^2\ll\Delta$ частица будет оставаться на месте, при $\Delta\ll\varepsilon^2$ частица будет почти мгновенно смещаться на край пробирки. Если $\varepsilon\approx\lambda\sqrt{\Delta}$ и соударения происходят достаточно часто $(\Delta\to 0,\ n\to\infty)$, то

$$X(t)=\lambda\sqrt{\frac{t}{n}}\sum\nolimits_1^n \xi_j \ \rightsquigarrow\ \mathcal{N}(0,\lambda^2 t).$$

Применяя многомерную центральную предельную теорему, теми же рассуждениями легко устанавливается, что и любое конечномерное распределение $X(t)$ в пределе будет нормальным с ковариацией $\mathbf{Cov}(X(s), X(t)) \to \to \lambda^2 \min(s,t)$ (последнее, конечно, ввиду независимости приращений).

Можно доказать и более сильный результат — так называемый принцип инвариантности Донскера–Прохорова (см. стр. 289). Рассмотрим отрезок $[0;1]$, который разобьём на n равных частей. Определим процесс

$$X_n(t) = \Big(S_{k-1} + \Big(t - \frac{k-1}{n}\Big)\xi_k\Big)\frac{1}{\sqrt{n}}, \quad t \in \Big(\frac{k-1}{n}; \frac{k}{n}\Big],$$

где $S_k = \sum_{i=1}^k \xi_i$, $k = 1, \ldots, n$, $S_0 = 0$. Другими словами, функция $X_n(t)$, $t \in [0;1]$, представляет собой непрерывную ломаную линию с узлами $(k/n, S_k/\sqrt{n})$. Принцип инвариантности (другое название «функциональная предельная теорема») утверждает, что последовательность процессов $\mathcal{X}_n = \big\langle X_n(t),\ t \in [0;1] \big\rangle$, $n = 1, 2, \ldots$, как последовательность случайных элементов в пространстве непрерывных функций $\mathcal{C}[0;1]$, сходится по распределению к винеровскому процессу $\boldsymbol{w}$. То есть для любого непрерывного ограниченного функционала h на $\mathcal{C}[0;1]$

$$\lim_n \mathbf{E}\, h(\mathcal{X}_n) = \mathbf{E}\, h(\boldsymbol{w}).$$

Ясно, что отрезок $[0;1]$ может быть заменён на отрезок $[0;T]$.

Заметим, что подобная модель описывает движение частицы в вязкой среде, без инерции. При наличии инерции, как вариант вероятностного описания движения броуновской частицы, можно использовать процесс Орнштейна–Уленбека (см. ниже).

✦ Процесс броуновского движения — процесс диффузионного типа, т.е. это однородный марковский (ввиду независимости приращений) процесс с непрерывными траекториями и переходным распределением, т.е. условным распределением $w(t)$ при фиксированном $w(s) = y$, $s < t$, таким, что

$$(\mathrm{D}_1) \quad \lim_{t \to s+0} \mathbf{P}\Big\{ \frac{|w(t) - y|}{t - s} > \varepsilon \mid w(s) = y \Big\} \ = \ 0;$$

$$(\mathrm{D}_2) \quad \lim_{t \to s+0} \mathbf{E}\Big[\frac{w(t) - y}{t - s}\, \dot{\mathrm{I}}\, \big(|w(t) - y| \leqslant \varepsilon\big) \mid w(s) = y \Big] \ = \ \mu(y);$$

$$(\mathrm{D}_3) \quad \lim_{t \to s+0} \mathbf{E}\Big[\frac{(w(t) - y)^2}{t - s}\, \dot{\mathrm{I}}\, \big(|w(t) - y| \leqslant \varepsilon\big) \mid w(s) = y \Big] \ = \ \delta^2(y)$$

для $\forall \varepsilon > 0$, $y \in \mathbb{R}^1$, где в данном случае коэффициент сноса $\mu(y) \equiv 0$, а коэффициент диффузии $\delta^2(y) \equiv \lambda^2$.

⚠ Легко проверить, что одномерная плотность винеровского процесса $f(x \mid t)$ $\big(\backsim \mathcal{N}(0,\lambda^2 t)\big)$ удовлетворяет уравнению диффузии $\partial f(x \mid t)/\partial t - \frac{\lambda^2}{2}\partial^2 f(x \mid t)/\partial x^2 = 0$. Как известно, это уравнение описывает изменение во времени концентрации одного вещества в процессе диффузии (проникновения) его в другое вещество.

Для доказательства свойств (D$_1$) – (D$_3$) воспользуемся утверждением b) теоремы 400. В соответствии с этим утверждением (полагая для упрощения $s=0$) при $t \to +0$ отношение

$$\frac{\mathbf{P}\left\{\left|w(t)-y\right| > \delta \mid w(0)=y\right\}}{t} = \frac{2\big(1-\Phi\big(\delta(\lambda\sqrt{t})^{-1}\big)\big)}{t} < \frac{2\lambda e^{-\delta^2/2\lambda^2 t}}{\delta\sqrt{2\pi t}} \to 0$$

в силу известного неравенства $1-\Phi(x) < e^{-x^2/2}/(x\sqrt{2\pi})$, $x>0$, (360, стр. 311) для ф.р. $\Phi(x)$ стандартного нормального закона.

Далее, применяя замену $(x-y)/\lambda\sqrt{t} \dashrightarrow z$, получаем

$$\frac{1}{t}\,\mathbf{E}\big[\,(w(t)-y)\,\dot{\mathrm{I}}\,\big(\left|w(t)-y\right| \leqslant \delta\big) \mid w(0)=y\,\big] =$$

$$= \int\limits_{y-\delta}^{y+\delta} \frac{(x-y)}{t\lambda\sqrt{2\pi t}}\, e^{-(x-y)^2/2\lambda^2 t}\,dx = \frac{\lambda}{\sqrt{2\pi t}} \int\limits_{-\delta/\lambda\sqrt{t}}^{\delta/\lambda\sqrt{t}} z e^{-z^2/2}\,dz = 0.$$

Аналогично, при $t \to +0$

$$\frac{1}{n}t\,\mathbf{E}\big[\,(w(t)-y)^2\,\dot{\mathrm{I}}\,\big(\left|w(t)-y\right| \leqslant \delta\big) \mid w(0)=y\,\big] = \int\limits_{y-\delta}^{y+\delta} \frac{(x-y)^2}{t\lambda\sqrt{2\pi t}} e^{-(x-y)^2/2\lambda^2 t}\,dx =$$

$$= \frac{\lambda^2}{\sqrt{2\pi}} \int\limits_{-\delta/\lambda\sqrt{t}}^{\delta/\lambda\sqrt{t}} z^2 e^{-z^2/2}\,dz \nearrow \frac{\lambda^2}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} z^2 e^{-z^2/2}\,dz = \lambda^2.$$

✦ Исследуем вариацию траектории (стандартного) винеровского процесса на конечном отрезке. Рассмотрим отрезок $[0;1]$ и выберем равномерные разбиения этого отрезка: $t_k = k/n$, $n \geqslant 2$. Положим

$$\delta_n = \sum_{k=1}^{n} \Big|w\big(\tfrac{k}{n}\big) - w\big(\tfrac{k-1}{n}\big)\Big|.$$

Так как $\eta := w\big(\frac{k}{n}\big) - w\big(\frac{k-1}{n}\big) \backsim \mathcal{N}\big(0,\frac{1}{n}\big)$, то $\mathbf{E}\left|\eta\right| = \sqrt{2}/\sqrt{n\pi}$, а дисперсия

$\mathbf{D}|\eta| = (\pi-2)/(n\pi)$. Поэтому в силу независимости приращений

$$\mathbf{E}\,\delta_n = \sqrt{\frac{2}{\pi}\,n} \quad \text{и} \quad \mathbf{D}\delta_n = \frac{\pi-2}{\pi}.$$

Следовательно, для $\forall\, M>0$ при $\forall\, n>2M^4$ имеем $\mathbf{E}\,\delta_n > M^2$ и

$$\mathbf{P}\{\delta_n < M\} \;=\; \mathbf{P}\Big\{\,\delta_n - \mathbf{E}\,\delta_n < M - \mathbf{E}\,\delta_n\Big\} \;\leqslant$$
$$\leqslant\; \mathbf{P}\Big\{\,\big(\delta_n - \mathbf{E}\,\delta_n\big)^2 > \big(M - \mathbf{E}\,\delta_n\big)^2\Big\} \;\leqslant\; \frac{\pi-2}{\pi(M-\mathbf{E}\,\delta_n)^2} \;\underset{n}{\to}\; 0.$$

Стало быть, вариация процесса Винера на любом отрезке равна ∞.

✦ Для стандартного винеровского процесса на отрезке $[0;t]$ при любом разбиении $0=x_0<x_1<\ldots<x_n=t$ математическое ожидание квадратической вариации

$$\mathbf{E}\,\Delta_n^2(t) \;:=\; \mathbf{E}\sum\nolimits_1^n \big(w(x_k)-w(x_{k-1})\big)^2 \;=\; \sum\nolimits_1^n (x_k - x_{k-1}) \;=\; t.$$

При равномерном разбиении с $x_k = \dfrac{k}{n}\,t$ дисперсия квадратической вариации

$$\mathbf{D}\Delta_n^2(t) \;=\; \sum\nolimits_{k=1}^n \mathbf{D}\big(w(x_k)-w(x_{k-1})\big)^2 \;=\; \sum\nolimits_{k=1}^n 2\frac{t^2}{n^2} \;=\; 2\frac{t^2}{n} \;\underset{n}{\to}\; 0, \qquad (22)$$

в силу независимости приращений и равенства $\mathbf{D}\eta^2 = 2(\gamma^2)^2$ для $\eta \backsim \mathcal{N}(0,\gamma^2)$ (?!). Другими словами, при $n\to\infty$ квадратическая вариация $\Delta_n^2(t) \xrightarrow{\mathcal{L}_2} t$.

✦ Поскольку нормированное приращение $\big(w(t+h)-w(t)\big)/\sqrt{h} \backsim \mathcal{N}(0,1)$, то для $\forall\, M>0$ при $h\to 0$

$$\mathbf{P}\Big\{\,\Big|\frac{w(t+h)-w(t)}{h}\Big| < M\,\Big\} \;=\; 2\Phi\big(M\sqrt{h}\big)-1 \;\to\; 0.$$

Поэтому $\Big|(w(t+h)-w(t))/h\Big| \xrightarrow{\mathbf{P}} \infty,\ h\to 0$. Этот факт означает, что в любой точке $t\geqslant 0$ траектории винеровского процесса недифференцируемы. Можно доказать и более сильное утверждение (см. [5, стр. 76]):

с вероятностью единица траектории винеровского процесса имеют бесконечную вариацию на любом конечном отрезке и недифференцируемы ни в одной точке полуоси $t\geqslant 0$.

✦ Винеровский процесс не обладает свойством стационарности. Однако можно утверждать, что при любом фиксированном $v\geqslant 0$ процесс $w(t+ +\, v)-w(t)$, $t\geqslant 0$, будет стационарным. Действительно, т.к. любые линейные преобразования оставляют процесс гауссовским, то ввиду утверждения 397

достаточно вычислить ковариацию сечений для $s<t$:

$$\mathbf{E}\left[\big(w(s+v)-w(s)\big)\big(w(t+v)-w(t)\big)\right] \;=\; (s+v)-s-\min(s+v,t)+s.$$

Последнее выражение равно 0, если $v\leqslant t-s$, и равно $v-(t-s)$ в противном случае, т.е. всегда зависит только от разности $t-s$.

✦ Процесс броуновского движения обладает свойством так называемой *автомодельности*: для $\forall c>0$ процесс $\left\langle \sqrt{c}\, w(t/c),\ t\geqslant 0\right\rangle$ также является броуновским движением (?!).

✦ *Инверсия времени*: случайная функция $z(t)=tw(1/t)\,\dot{\mathrm{I}}\,(t>0)$ образует винеровский процесс. Гауссовость конечномерных распределений z очевидна, а ковариация для $s<t$ легко вычисляется:

$$\mathbf{Cov}\,(z(s),z(t)) \;=\; st\min\left\{\frac{1}{s},\frac{1}{t}\right\} \;=\; st\frac{1}{t} \;=\; s.$$

Осталось показать только непрерывность (п.н.) при $t\to 0$.

401| Упр. Докажите непрерывность $z(t)$ в нуле в среднем квадратическом и по вероятности.

Докажем сначала неравенство Колмогорова для винеровского процесса:

$$\mathbf{P}\Big\{\sup_{0\leqslant s\leqslant t}|w(s)|>\varepsilon\Big\} \;\leqslant\; \frac{t}{\varepsilon^2}.$$

В силу непрерывности траекторий броуновского движения

$$\mathbf{P}\Big\{\sup_{0\leqslant s\leqslant t}|w(s)|>\varepsilon\Big\} \;=\; \lim_{n}{}_{\uparrow}\,\mathbf{P}\Big\{\sup_{1\leqslant k\leqslant 2^n}\Big|w\Big(\frac{kt}{2^n}\Big)\Big|>\varepsilon\Big\} \;=$$

$$=\; \lim_{n}{}_{\uparrow}\,\mathbf{P}\Big\{\sup_{1\leqslant k\leqslant 2^n}\Big|\sum\nolimits_{j=1}^{k}\Big[w\Big(\frac{tj}{2^n}\Big)-w\Big(\frac{(j-1)t}{2^n}\Big)\Big]\Big|>\varepsilon\Big\} \;\leqslant\; \frac{t}{\varepsilon^2},$$

где последнее неравенство есть результат применения классического неравенства Колмогорова для суммы независимых центрированных сл.в. — здесь дисперсия суммы при $k=2^n$ равна $\mathbf{D}w(t)=t$. Далее, для $n\geqslant 1$ введём события $A_n=\big\{\sup_{2^{n-1}\leqslant s\leqslant 2^n}\frac{1}{s}|w(s)|>\varepsilon\big\}$. Очевидно, что если $\overline{\lim\limits_{t\to\infty}}\,\frac{1}{t}|w(t)|>\varepsilon$, то произойдёт бесконечное число (б.ч.) событий $A_n,\ n\geqslant 1$. Таким образом,

$$\mathbf{P}\Big\{\overline{\lim_{t\to\infty}}\,\frac{1}{t}|w(t)|>\varepsilon\Big\} \;\leqslant\; \mathbf{P}\,\{A_n\ \text{б.ч.}\}\,.$$

Воспользовавшись неравенством Колмогорова, получаем

$$\sum_1^\infty \mathbf{P}\{\, A_n \,\} \;\leqslant\; \sum_1^\infty \mathbf{P}\Big\{ \sup_{2^{n-1}\leqslant s\leqslant 2^n} |w(s)| > 2^{n-1}\varepsilon \Big\} \;\leqslant\; \sum_1^\infty \frac{1}{\varepsilon^2 2^{n-2}} < \infty.$$

По лемме Бореля–Кантелли отсюда следует, что $\mathbf{P}\{A_n \text{ б.ч.}\} = 0$. Стало быть,

$$\mathbf{P}\Big\{ \lim_{u\to 0} z(u) = 0 \Big\} \;=\; \mathbf{P}\Big\{ \lim_{t\to\infty} \frac{1}{t} w(t) = 0 \Big\} \;=\; 1.$$

✦ При доказательстве предыдущего свойства попутно нами доказан усиленный закон больших чисел для винеровского процесса:

$$\lim_{t\to\infty} \frac{1}{t} w(t) = 0 \quad (\text{п.н.})\,.$$

✦ Неравенство Колмогорова в данной ситуации может быть уточнено. Без доказательства примем следующий факт симметрии броуновского движения (называемый ещё принципом отражения по аналогии с симметричным бернуллиевским случайным блужданием). Пусть $\nu_b = \inf\{u > 0 : w(u) = b\}$ — момент первого достижения процессом заданного уровня $b > 0$. Тогда справедливо следующее тождество:

$$\mathbf{P}\left\{ w(t) > b \,|\, \nu_b < t \right\} \;=\; \mathbf{P}\left\{ w(t) < b \,|\, \nu_b < t \right\} \;=\; \tfrac{1}{2}.$$

Другими словами, если когда-то до момента t процесс достигал уровня b, то в момент t с равными вероятностями процесс будет либо выше, либо ниже этого уровня. Поскольку в силу непрерывности траекторий событие $\{w(t) > b\} \subset \{\nu_b < t\}$, отсюда следует, что

$$\mathbf{P}\Big\{ \sup_{0\leqslant s\leqslant t} w(s) > b \Big\} \;=\; \mathbf{P}\{\nu_b < t\} \;=\; 2\,\mathbf{P}\{w(t) > b\} \;=\; 2\Phi(-b/\sqrt{t}).$$

Отметим как полезный факт, что ввиду доказанного сл.в. ν_b собственная, т.е. $\nu_b < \infty$ (п.н.) . С другой стороны, её математическое ожидание бесконечно. Действительно, с помощью правила Лопиталя легко вывести, что

$$\mathbf{P}\{\nu_b \geqslant t\} \;=\; 1 - 2\Phi\Big(-\frac{b}{\sqrt{t}}\Big) \;\asymp\; \frac{2b}{\sqrt{2\pi}}\frac{1}{\sqrt{t}}, \quad t\to\infty. \tag{23}$$

Поэтому $\mathbf{E}\,\nu_b = \int_0^\infty \mathbf{P}\{\nu_b \geqslant t\}\, dt = \infty$.

Для отыскания распределения времени нахождения траекторий винеровского процесса в полосе $[a; b]$ можно применить мартингальные методы (подробности см. § 5, стр. 402). Обозначим $\mathscr{F}_t = \sigma\big(w(s), s \leqslant t\big) := \sigma\big(\bigcup_{s\leqslant t} w_s^{-1}(\mathscr{B})\big)$

минимальную σ-алгебру, относительно которой измеримы все сечения винеровского процесса до момента t. Понятно, что для $t < u$ имеем $\mathscr{F}_t \subset \mathscr{F}_u \subset \mathscr{F}$, т.е. семейство $\mathscr{F}^{\boldsymbol{w}} = \left\langle \mathscr{F}_t,\ t \geqslant 0 \right\rangle$ образует *поток σ-алгебр* или *фильтрацию* (естественную фильтрацию для процесса $\boldsymbol{w}$). Напомним, что измеримость любой функции $h(\omega)$, $\omega \in \Omega$, относительно σ-алгебры $\mathscr{F}_t$ означает, что её значения полностью определяются траекторией $\boldsymbol{w}$ до момента t. Заметим, что по лемме 372, стр. 324, приращение $w(u) - w(t)$ не зависит от $\mathscr{F}_t$ для $u > t$. Поэтому для $u > t$ усл.м.о.

$$\mathbf{E}\big[\, w(u) \,|\, \mathscr{F}_t \,\big] = \mathbf{E}\big[\, w(t) + w(u) - w(t) \,|\, \mathscr{F}_t \,\big] = w(t) + \mathbf{E}\,\big[w(u) - w(t)\big] = w(t).$$

Другими словами, винеровский процесс образует мартингал относительно естественной фильтрации $\mathscr{F}^{\boldsymbol{w}}$.

Марковским моментом относительно фильтрации $\mathcal{G} = \left\langle \mathscr{G}_t \right\rangle_0^\infty$ называется сл.в. $\nu : \Omega \mapsto \mathbb{R}_+^1 \cup \{+\infty\}$, для которой события $\{\nu \leqslant t\} \in \mathscr{G}_t$ при $\forall\, t \geqslant 0$. Если $\mathbf{P}\,\{\nu < \infty\} = 1$, то ν называется *моментом остановки*. Другими словами, решение о прекращении наблюдений за процессом после момента времени t принимается только на основе информации, доступной к моменту t. Легко понять, что если ν — марковский момент, то $\nu \wedge t = \min\{\nu, t\}$ — момент остановки для $\forall\, t > 0$.

402| Пример. Момент первого достижения некоторого заданного уровня $b > 0$ винеровским процессом

$$\nu_b = \inf\{u > 0 : w(u) = b\}$$

представляет собой момент остановки относительно естественной фильтрации $\mathscr{F}^{\boldsymbol{w}}$. Действительно, в силу непрерывности траекторий процесса

$$\{\nu_b > t\} = \bigcup_{k=1}^{\infty} \bigcap_{q \leqslant t} \left\{ b - w(q) \geqslant \tfrac{1}{k} \right\} \in \mathscr{F}_t,$$

где пересечение берётся по всем рациональным $q \leqslant t$. Кроме того, в силу (23) вероятность $\mathbf{P}\,\{\nu_b < \infty\} = 1$.

Аналогично показывается, что при $-a < 0 < b$ марковским моментом (то, что это момент остановки, требует дополнительной проверки) будет

$$\nu_{(-a,b)} = \inf\{u > 0 : w(u) = b \ \text{ или } \ w(u) = -a\}. \qquad \circledcirc$$

Первая теорема о свободной остановке (optional stoping (или sampling) theorem) гласит, что для любого марковского момента ν относительно некоторой фильтрации $\mathcal{G} = \left\langle \mathscr{G}_t \right\rangle_0^\infty$ и любого мартингала $U(t)$ относительно этой же фильтрации $\mathcal{G}$, удовлетворяющего условию $\mathbf{E}|U(t)| < \infty$, справедливо равенство

$$\mathbf{E}\,U(\nu \wedge n) = \mathbf{E}\,U(0).$$

Пусть $\boldsymbol{w}$ — стандартный винеровский процесс, тогда процесс $U(t) = w^2(t) - t$ образует мартингал относительно естественной фильтрации $\mathscr{F}^{\boldsymbol{w}}$ со средним $\mathbf{E}\,U(t) = 0$. Действительно, $\mathbf{E}\,|U(t)| \leqslant \mathbf{E}\,w^2(t) + t = 2t < \infty$, и

$$\begin{aligned}\mathbf{E}\big[\,U(t)\mid \mathscr{F}_s\big] \;&=\; \mathbf{E}\big[\,w^2(s)\mid \mathscr{F}_s\big] + \mathbf{E}\Big[\,\big(w(t)-w(s)\big)^2 \mid \mathscr{F}_s\Big] + \\ &\quad + 2\,\mathbf{E}\Big[\,\big(w(t)-w(s)\big)w(s)\mid \mathscr{F}_s\Big] - t \;= \\ &=\; w^2(s) + (t-s) + 0 - t = U(s)\end{aligned}$$

в силу независимости $w(t)-w(s)$ от σ-алгебры $\mathscr{F}_s$ и измеримости $w(s) \Subset \mathscr{F}_s$.

Таким образом, для $\forall\, n > 1$ по теореме о свободной остановке имеем

$$\mathbf{E}\left[n \wedge \nu_{(-a,b)}\right] \;=\; \mathbf{E}\left[w^2\big(n \wedge \nu_{(-a,b)}\big)\right] \;\leqslant\; \min\{a^2, b^2\}. \tag{24}$$

Следовательно, по теореме Б. Леви о монотонной сходимости

$$\mathbf{E}\left[\nu_{(-a,b)}\right] \;=\; \lim_{n}{\uparrow}\; \mathbf{E}\left[n \wedge \nu_{(-a,b)}\right] \;\leqslant\; \min\{a^2, b^2\}.$$

Стало быть, $\mathbf{P}\big\{\,\nu_{(-a,b)} < \infty\,\big\} = 1$, откуда $w^2\big(n \wedge \nu_{(-a,b)}\big) \xrightarrow[n]{} w^2\big(\nu_{(-a,b)}\big)$ (п.н.) . Из первого равенства (24) в силу теоремы Лебега получаем тождество

$$\mathbf{E}\left[\nu_{(-a,b)}\right] \;=\; \mathbf{E}\left[w^2\big(\nu_{(-a,b)}\big)\right] \;=\; b^2 p_b + a^2 p_a,$$

где p_a, p_b — вероятности достижения соответственно нижней (верхней) границы $-a$ (b) раньше, чем противоположной границы. Так как $\nu_{(-a,b)}$ — момент остановки, то по аналогичным соображениям, применённым к мартингалу $\boldsymbol{w}$, получаем $0 = \mathbf{E}\,w\big(\nu_{(-a,b)}\big) = bp_b - ap_a$, откуда, т.к. $p_a + p_b = 1$,

$$p_b = \frac{a}{a+b}\,.$$

С учётом предыдущих вычислений приходим к тому, что среднее время нахождения траекторий броуновского движения в полосе $(-a; b)$, $a, b > 0$, равно

$$\mathbf{E}\left[\nu_{(-a,b)}\right] = ab.$$

✦ Винеровский процесс представляет собой обобщение на непрерывный случай последовательности частичных сумм $S_t = \sum_1^t \xi_k,\ t \geqslant 1$, независимых

$\mathcal{N}(0,1)$ сл.в. Если же слагаемые в этих суммах $\xi_k \backsim \mathcal{N}(\mu,\sigma^2)$, то частичная сумма $S_t \backsim \mathcal{N}(\mu t, \sigma^2 t)$.

Определение. Гауссовский процесс $\boldsymbol{w}_\mu = \left\langle\, w_\mu(t),\ t \in \mathbb{R}^1 \,\right\rangle$, называется винеровским процессом со сносом, если его функция среднего $\mathbf{E}\, w_\mu(t) = \mu t,\ \mu \in \mathbb{R}^1$, а ковариационная функция $R(s,t) = \sigma^2 \min\{s,t\}$.

Понятно, что если $\boldsymbol{w}$ — стандартный винеровский процесс, то $\sigma w(t) + \mu t$ — винеровский процесс со сносом, и наоборот, $(w_\mu - \mu t)/\sigma$ — стандартный винеровский процесс.

Найдём, используя мартингальные методы, среднее значение момента $\nu_{(-a,b)}$ выхода из полосы $(-a,b)$ траектории винеровского процесса w_μ со сносом. Не ограничивая общности, можно положить $\sigma^2 = 1$ (очевидно, общий случай получится, если в полученных результатах заменить вектор параметров $(a,b,\mu) \dashrightarrow \frac{1}{\sigma}(a,b,\mu)$). Рассмотрим мартингал (для $\forall\,\lambda \in \mathbb{R}^1$)

$$V(t) \;=\; \exp\left\{\lambda w_\mu(t) - (\lambda\mu + \tfrac{1}{2}\,\lambda^2)t\right\}.$$

Для проверки того, что это действительно мартингал, достаточно заметить, что для $t > s$

$$\mathbf{E}\exp\left\{\lambda\eta\right\} \;=\; \exp\left\{(t-s)\big(\lambda\mu + \tfrac{1}{2}\,\lambda^2\big)\right\},$$

где $\eta \backsim \mathcal{N}\big(\mu(t-s), (t-s)\big)$. Выбирая $\lambda = -2\mu$, после применения теоремы о свободном выборе 458, стр. 415, получаем

$$1 \;=\; \mathbf{E}\,V(0) \;=\; \mathbf{E}\exp\left\{-2\mu\, w_\mu\big(\nu_{(-a,b)}\big)\right\} \;=\; p_b e^{-2b\mu} + p_a e^{2a\mu}.$$

Откуда

$$p_b \;=\; \frac{\exp\{-2\mu a\} - 1}{\exp\{-2\mu b\}\exp\{-2\mu a\} - 1}.$$

При отрицательном сносе ($\mu < 0$) вероятность того, что процесс w_μ достигнет положительного уровня $b > 0$, меньше единицы, т.к.

$$\mathbf{P}\Big\{\sup_{t\geqslant 0} w_\mu(t) > b\Big\} \;=\; \lim_{a\to\infty} p_b \;=\; \exp\{-2|\mu| b\} \;<\; 1.$$

При $\mu > 0$ эта вероятность, конечно, всегда равна единице.

Опишем распределение ν_b первого достижения фиксированного уровня $b > 0$ винеровским процессом со сносом. Пусть $\lambda \geqslant 0$ таково, что $\theta = \lambda\mu +$

$+\frac{1}{2}\lambda^2 > 0$. Тогда для мартингала $V(t) = \exp\{\lambda w_\mu(t) - \theta t\}$ имеем

$$1 = \mathbf{E}\big[V(n\wedge\nu_b)\big] = \mathbf{E}\exp\{\lambda w(n\wedge\nu_b) - (n\wedge\nu_b)\theta\}.$$

По построению $V(n\wedge\nu_b) \leqslant e^{\lambda b}$. При этом, если $\nu_b < \infty$, то $\lambda w(n\wedge\nu_b) - (n\wedge \wedge\,\nu_b)\theta \underset{n}{\to} \lambda b - \theta\nu_b$, в противном случае $\lambda w(n\wedge\nu_b) - (n\wedge\nu_b)\theta \underset{n}{\to} -\infty$. Поэтому, в силу теоремы Лебега о мажорируемой сходимости,

$$1 = e^{\lambda b}\,\mathbf{E}\,e^{-\theta\nu_b}, \text{ следовательно,}$$

$$\mathbf{E}\,e^{-\theta\nu_b} = e^{-\lambda b} = \exp\{-b(\sqrt{\mu^2+2\theta}-\mu)\}.$$

Последнее выражение представляет собой не что иное, как преобразование Лапласа неотрицательной сл.в. ν_b. Полагая $\theta\to 0$, получим

$$\lim_{\theta\to 0}\mathbf{E}\,e^{-\theta\nu_b} = \mathbf{P}\{\nu_b<\infty\} = \begin{cases} 1, & \mu\geqslant 0,\\ e^{2\mu b}\ (<1), & \mu<0,\end{cases}$$

т.е. ν_b имеем несобственное распределение при $\mu<0$: $\mathbf{P}\{\nu_b=\infty\}>0$.

При $\mu>0$ можно воспользоваться обратным преобразованием Лапласа (8), стр. 176, чтобы получить плотность распределения ν_b.

403 **Теорема.** *Пусть $w_\mu(t) = \sigma w(t) + \mu t$ — винеровский процесс со сносом $\mu\geqslant 0$ и коэффициентом диффузии $\sigma^2>0$, $\nu_b = \inf\{s\geqslant 0 : w_\mu(s) = = b\}$ — момент первого достижения траекторией процесса уровня $b>0$. Тогда плотность распределения ν_b*

$$f(t;\ b,\mu,\sigma) = \frac{b}{\sigma\sqrt{2\pi t^3}}\exp\Big\{-\frac{(b-\mu t)^2}{2\sigma^2 t}\Big\},\qquad t>0.$$

Математическое ожидание времени нахождения ниже уровня b можно найти, как обычно, из выражения для функции плотности. Однако проще это сделать, воспользовавшись равенством

$$\mathbf{E}\,\nu_b = -\Big(\mathbf{E}\,e^{-\theta\nu_b}\Big)'_{\theta=0} = b\,\frac{\exp\{-b(\sqrt{\mu^2+2\theta}-\mu)\}}{\sqrt{\mu^2+2\theta}}\bigg|_{\theta=0} = \frac{b}{\mu}.$$

Заметим, что при $\mu\searrow 0$ математическое ожидание $\mathbf{E}\,\nu_b\nearrow\infty$, что вполне согласуется с выводом, сделанным на основе (23).

✦ Винеровский процесс сохраняет свои вероятностные свойства, если начать обзор его траектории с достигнутого уровня после какого-то момента времени v. При этом поведение траектории после этого момента не зависит от всей истории процесса при $t \leqslant v$.

404| **Лемма.** *Пусть $\langle w(t),\ t \geqslant 0 \rangle$ — стандартный винеровский процесс. Тогда для любого фиксированного $v \geqslant 0$ процесс $\tilde{w}(t) = w(t+v) - w(v)$, $t \geqslant 0$, есть стандартный винеровский процесс. При любых $t_1, \ldots, t_n \geqslant 0, s_1, \ldots, s_m < v$ вектор $(\tilde{w}(t_1), \ldots, \tilde{w}(t_n))$ не зависит от вектора $(w(s_1), \ldots, w(s_m))$.*

⊨ Нормальность конечномерных распределений $\tilde{w}$ следует из того, что распределение любого линейного преобразования компонент нормального вектора снова нормально. Поскольку $\tilde{w}(0) = 0$ и среднее значение $\mathbf{E}\,\tilde{w}(t) \equiv 0$, осталось показать, что ковариация $\mathbf{Cov}(\tilde{w}(s), \tilde{w}(t)) = s$ для $s < t$:

$$\begin{aligned} \mathbf{Cov}(\tilde{w}(s), \tilde{w}(t)) &= \mathbf{E}\left[\big(w(s+v) - w(v)\big)\big(w(t+v) - w(v)\big)\right] = \\ &= (s+v) - v - v + v = s. \end{aligned}$$

Вектор $(\tilde{w}(t_1), \ldots, \tilde{w}(t_n), w(s_1), \ldots, w(s_m))$ имеет нормальное распределение, поэтому для доказательства независимости указанных в лемме векторов достаточно показать попарную некоррелированность их компонент. Последнее следует из независимости приращений винеровского процесса: $\mathbf{Cov}\,\big(w(t+v) - w(v), w(s)\big) = 0$, $\forall\, t \geqslant 0$, $s \leqslant v$ (см. также 372, стр. 324). ⫤

✦ Предыдущее свойство сохраняется, если вместо фиксированного момента времени v взять случайный марковский момент. Строго марковское свойство винеровского процесса гласит, что для любого момента остановки ν относительно потока $\mathscr{F}^{w}$ процесс $w_\nu(t) = w(t+\nu) - w(\nu)$, $t \geqslant 0$ (для простоты считаем здесь, что $\nu(\omega) < \infty$ для $\forall\, \omega \in \Omega$), также будет винеровским процессом, который не зависит от реализаций исходного процесса w, произошедших до момента ν. Последнее означает независимость семейства сл.в. $w_\nu(t)$, $t \geqslant 0$, от σ-алгебры

$$\mathscr{F}_\nu = \big\langle B \in \mathscr{F} :\ B \cap \{\nu \leqslant t\} \in \mathscr{F}_t \ \text{ для } \ \forall\, t \geqslant 0 \big\rangle.$$

✦ С винеровским процессом тесно связано понятие броуновского моста. Этот процесс возникает в ситуациях, подобных следующей. Пусть $F_n(x; \xi^{(n)})$, $x \in [0;1]$, — эмпирическая ф.р., построенная по выборке $\xi^{(n)}$ из равномерного $\mathcal{U}\mathtt{n}(0;1)$ распределения. Тогда предельное (при $n \to \infty$) распределение $\sqrt{n}\big(F_n(x; \xi^{(n)}) - x\big)$ как элемента пространства функций $\mathcal{D}[0;1]$

совпадает с распределением некоторого гауссовского процесса $\beta(x)$, $x \in \in [0;1]$. Понятно, что математическое ожидание $\mathbf{E}\beta(x) = 0$, а дисперсия $\mathbf{D}\beta(x) = x(1-x)$. Из построения ясно, что $\beta(0) = \beta(1) = 0$, т.е. траектория процесса образует своего рода мост между «двух берегов» отрезка $[0;1]$.

Определение. Гауссовский процесс $\beta(t)$, $t \in [0;1]$, с непрерывными траекториями называется *броуновским мостом*, если его математическое ожидание $\mathbf{E}\beta(t) = 0$, а ковариационная функция $\mathbf{Cov}\big(\beta(s), \beta(t)\big) = s(1-t)$ для $s \leqslant t$.

Обратим внимание, что $\mathbf{D}\beta(0) = \mathbf{D}\beta(1) = 0$, т.е. $\beta(0) = \beta(1) = 0$ (п.н.) .

Связь между броуновским мостом и винеровским процессом объясняют следующие два факта.

405| **Лемма.** (?!) *Если* $w(t)$, $t \in [0;1]$, *— винеровский процесс, то случайная функция* $w(t) - tw(1)$, $t \in [0;1]$, *образует броуновский мост.*

Кроме того, распределение броуновского моста (в пространстве непрерывных функций) совпадает с условным распределением винеровского процесса при условии, что $w(1) = 0$. Это утверждение требует уточнения, поскольку вероятность условия равна нулю.

406| **Лемма.** *Пусть* Π *— распределение броуновского моста в пространстве* $\mathcal{C}[0;1]$, W_ε *— условное распределение винеровского процесса, определяемое равенством*

$$W_\varepsilon(B) \;=\; \mathbf{P}\big\{\boldsymbol{w} \in B \,\big|\, |w(1)| < \varepsilon\big\}, \quad B \subset \mathcal{C}[0;1],\ \varepsilon > 0.$$

Тогда при $\varepsilon \to 0$ *имеет место слабая сходимость семейства мер* $W_\varepsilon \Rightarrow \Pi$.

$\Rrightarrow$ Заметим сначала, что процесс $\beta(t) = w(t) - tw(1)$ не зависит от $w(1)$. Действительно, ковариация $\mathbf{Cov}\big(\beta(t), w(1)\big) = \mathbf{E}\big(w(t) - tw(1)\big)w(1) = = t - t = 0$ для $\forall t \in [0;1]$. С другой стороны, любой конечномерный вектор $\big(\beta(t_1), \ldots, \beta(t_k), w(1)\big)$ есть линейное преобразование нормального вектора, следовательно, некоррелированность его компонент эквивалентна независимости. Представим винеровский процесс в виде $\boldsymbol{w} = \boldsymbol{\beta} + \boldsymbol{\eta}$, где процесс $\boldsymbol{\eta} = \big\langle\, t\,w(1),\ t \in [0;1]\,\big\rangle$.

Для любого непрерывного и ограниченного функционала $H : \mathcal{C}[0;1] \mapsto \mathbb{R}^1$

математическое ожидание $H(\boldsymbol{w})$ относительно W_ε равно усл.м.о.:

$$\mathbf{E}\big[\,H(\boldsymbol{w})\,\big|\,|w(1)|<\varepsilon\,\big] \;=\; \mathbf{E}\big[\,H(\boldsymbol{\beta}+\boldsymbol{\eta})\,\big|\,|w(1)|<\varepsilon\,\big] \;= \\ =\; \mathbf{E}^{\beta}\Big(\mathbf{E}^{\eta}\big[H(\boldsymbol{\beta}+\boldsymbol{\eta})\,\big|\,|w(1)|<\varepsilon\big]\Big),$$

где символы $\mathbf{E}^{\beta}, \mathbf{E}^{\eta}$ указывают на случайные элементы, относительно которых производится вычисление математического ожидания, и последнее равенство ввиду независимости $\boldsymbol{\beta}$ и $w(1)$ следует из свойств усл.м.о. Так как при $\varepsilon \to 0$ условное распределение $\boldsymbol{\eta}$ сходится к вырожденному в нуле распределению, то при любых реализациях процесса $\boldsymbol{\beta}$ усл.м.о. $\mathbf{E}^{\eta}\big[H(\boldsymbol{\beta}+\boldsymbol{\eta})\,\big|\,|w(1)|<\varepsilon\big] \to H(\boldsymbol{\beta})$, что и требовалось. ⫤

✧ **Интегральное исчисление относительно процесса Винера.** Рассмотрим стандартный процесс Винера $\boldsymbol{w} = \big\langle\, w(t),\ t \geqslant 0\,\big\rangle$ с непрерывными траекториями.

В силу непрерывности для каждой траектории винеровского процесса может быть вычислен интеграл Римана по любому конечному отрезку. Выберем отрезок $[0;t]$, тогда по определению интеграла Римана

$$\int\limits_0^t w(u)\,du = \lim_{n\to\infty}\frac{t}{n}\sum_{k=0}^{n-1} w\Big(k\frac{t}{n}\Big).$$

По свойству нормального распределения для $\forall n \geqslant 2$ интегральная сумма здесь есть реализация нормальной сл.в. с нулевым средним (напомним, что среднее винеровского процесса равно нулю) и дисперсией

$$\mathbf{D}\Big[\frac{t}{n}\sum_{k=0}^{n-1} w\Big(k\frac{t}{n}\Big)\Big] \;=\; \frac{t^2}{n^2}\sum_{k,j=0}^{n-1} R\Big(k\frac{t}{n}, j\frac{t}{n}\Big) \;\underset{n}{\to}\; \int\limits_0^t\int\limits_0^t R(u,v)\,du dv \;=\; \frac{t^3}{3},$$

где мы использовали вид ковариации $R(u,v) = \min(u,v)$. Опираясь на метод хар.функций, легко понять, что предел нормальных сл.в. имеет нормальное распределение (см. 281, стр. 239). Таким образом, для винеровского процесса

$$\int_0^t w(u)\,du \backsim \mathcal{N}\big(0,\ \tfrac{1}{3}t^3\big). \tag{25}$$

Более того, аналогично показывается, что процесс $\big\langle\, \int_0^t w(u)\,du,\ t \geqslant 0\,\big\rangle$ есть гауссовский процесс с непрерывными траекториями.

407| Упр. Покажите, что ковариация процесса $\xi(t) = \int_0^t w(u)\,du,\ t \geqslant 0,$

для $s \leqslant t$ равна $R(s,t) = (3s^2 t - s^3)/6$.

Подсказка. Рассмотрите интеграл $\int_0^t g(u)w(u)\,du$ с $g(u) = \dot{\mathrm{I}}_{\{u \leqslant s\}}$.

✦ Определим для каждого конечного интервала $(a;c] \subset \mathbb{R}^1_+$ функцию множеств $W(a;c] = w(c) - w(a)$. Легко проверить, что для такой функции выполняются все свойства $(\mathrm{M}_1)-(\mathrm{M}_4)$ ортогональной стохастической меры (стр. 340). Структурная функция для этой меры $\mathfrak{m}(a;c] = \mathbf{D}\big[w(c) - w(a)\big] = c - a$, т.е. совпадает с мерой Лебега. Поэтому для любой интегрируемой с квадратом по Риману функции h на $\mathbb{R}^1_+$ существует стохастический интеграл

$$J_w(h) \;=\; \int_{\mathbb{R}^1_+} h(u)\,W(du) \;=:\; \int_{\mathbb{R}^1_+} h(u)\,dw(u), \tag{26}$$

где последнее обозначение не следует воспринимать как интеграл Стилтьеса — это просто указание на то, что мера, по которой производится интегрирование, порождена процессом Винера (сравните обозначения интеграла Лебега–Стилтьеса по мере и по функции распределения).

По теореме 386, стр. 343, математическое ожидание этого интеграла равно нулю, а дисперсия $\mathbf{D}\big[J_w(h)\big] = \int_{\mathbb{R}^1_+} h^2(u)\,du$. Поскольку стохастический интеграл от детерминированной функции равен пределу (в смысле $\mathcal{L}_2$) конечных линейных комбинаций приращений винеровского процесса, имеющих нормальное распределение, то и сам этот интеграл будет иметь нормальное распределение:

$$\int_{\mathbb{R}^1_+} h(u)\,dw(u) \;\backsim\; \mathcal{N}\big(0, \mathbf{D}\big[J_w(h)\big]\big).$$

В частности, для $h(u) = u\,\dot{\mathrm{I}}\,(u \leqslant t)$ дисперсия $\mathbf{D}\big[J_w(h)\big] = t^3/3$.

Интересно, что введённые выше два интеграла можно связать формулой, подобной формуле интегрирования по частям:

$$\int_0^t w(u)\,du + \int_0^t u\,dw(u) \;\backsim\; tw(t), \tag{27}$$

где стохастический интеграл в левой части вычисляется от функции $h(u) = u\,\dot{\mathrm{I}}_t(u)$, с индикатором $\dot{\mathrm{I}}_t$ отрезка $[0;t]$. Далее мы установим, что для стохастического интеграла Ито $(\int u\,dw)$ совпадение сл.в. в левой и правой частях здесь имеет место не только по распределению, но и почти наверное. Пока же интегралы в этой формуле представляют собой сл.в., построенные различными способами. Если первый из них $(\int w\,du)$ есть поточечный для каждой реализации $\omega \in \Omega$ предел некоторых интегральных сумм, то попытка определения стохастического интеграла для каждой траектории винеровского процесса как обычного интеграла Стилтьеса обречена на провал, поскольку вариация любой траектории на любом отрезке равна бесконечности и, следовательно, интегральные суммы Стилтьеса не будут сходиться.

Покажем справедливость (27). Поскольку ясно, что обе части этого соотношения имеют нормальное распределение с нулевым средним, достаточно показать, что дисперсия левой части равна дисперсии правой части $\mathbf{D}\big[tw(t)\big] = t^3$. Не стремясь к абсолютной строгости, при вычислении обоих интегралов разобьём отрезок $[0;t]$ на равное число интервалов. Следовательно, ковариация между интегралами вычисляется следующим образом:

$$\lim_n \sum_{k=0}^{n-1}\sum_{j=0}^{n-1} \frac{kt}{n}\Big(\frac{(j+1)t}{n} - \frac{jt}{n}\Big)\,\mathbf{E}\Big[w\Big(\frac{jt}{n}\Big)\Big(w\Big(\frac{(k+1)t}{n}\Big) - w\Big(\frac{kt}{n}\Big)\Big)\Big] =$$

$$= \lim_n \sum_{k=0}^{n-1}\sum_{j=k+1}^{n-1} \frac{kt}{n}\Big(\frac{(j+1)t}{n} - \frac{jt}{n}\Big)\Big(\frac{(k+1)t}{n} - \frac{kt}{n}\Big) =$$

$$= \int_0^t x \int_x^t ds\,dx = \frac{t^3}{6},$$

где $\mathbf{E}\Big[w\big(\frac{jt}{n}\big)\Big(w\big(\frac{(k+1)t}{n}\big) - w\big(\frac{kt}{n}\big)\Big)\Big] = 0$ для $j \leqslant k$. Таким образом,

$$\mathbf{D}\Big[\int_0^t w(u)\,du + \int_0^t u\,dw(u)\Big] = \frac{t^3}{3} + \frac{t^3}{3} + 2\frac{t^3}{6} = t^3.$$

408| Упр. Проверьте равенство распределений $\int_0^t w(u)\,dg(u) + \int_0^t g(u)\,dw(u)$ и $g(t)w(t)$:

а) для функции $g(u) = u^2$;

б) для непрерывной функции g с ограниченной вариацией.

Как отмечалось в 385, стр. 342, существует продолжение меры W на борелевские подмножества $\mathbb{R}^1_+$. Структурная функция W совпадает с мерой Лебега. Это объясняет значимость меры W для теории стохастического интегрирования, сравнимую со значимостью меры Лебега для теории интегрирования действительных функций.

Определим теперь интеграл, подобный (26), для случайных функций h, т.е. для процессов. Пусть $\boldsymbol{\xi} = \big\langle\, \xi(t),\ t \in \mathbb{T} = [0;T] \,\big\rangle$ — некоторый сл.процесс. Действуя по схеме определения интеграла Римана, рассмотрим интегральную сумму

$$S_n = \sum_{k=1}^{n} \xi(\tilde{t}_k)\big(w(t_k) - w(t_{k-1})\big), \tag{28}$$

где $0 = t_0 \leqslant \tilde{t}_1 \leqslant t_1 \leqslant \ldots \leqslant t_{n-1} \leqslant \tilde{t}_n \leqslant t_n = T$. Точное распределение этой суммы зависит от $\boldsymbol{\xi}$ и в общем случае произвольно. Единственное, что мы можем попытаться найти, — её числовые характеристики. Очевидно, для вычисления математического ожидания $\mathbf{E}\,S_n$ необходимо обладать информацией о ковариации между сл.в. $\xi(\tilde{t}_k)$ и $w(t_j)$. Из вида интегральной суммы понятно, что при построении общей теории удобнее всего потребовать, чтобы $\xi(\tilde{t}_k)$ не зависело от приращения $\big(w(t_k) - w(t_{k-1})\big)$. Для этого можно выбрать $\tilde{t}_k = t_{k-1}$ и рассматривать такие сл.процессы, для которых сечения в каждой точке t зависят только от реализаций винеровского процесса до момента t. Напомним, что зависимость одной сл.в. ζ от значений другой сл.в. η означает, что ζ измерима относительно σ-алгебры $\sigma(\eta)$, порождённой η. Обозначим $\mathscr{F}_t = \sigma\big(w(s), s \leqslant t\big) := \sigma\big(\bigcup_{s\leqslant t} w_s^{-1}(\mathscr{B})\big)$ минимальную σ-алгебру, относительно которой измеримы все сечения винеровского процесса до момента t. Как уже отмечалось, семейство $\big\langle\, \mathscr{F}_t,\ t \geqslant 0 \,\big\rangle$ образует поток σ-алгебр, или фильтрацию. Дабы избежать ситуаций, при которых из двух совпадающих почти наверное сл.в. одна измерима относительно некоторой σ-алгебры, а другая неизмерима, удобно каждую σ-алгебру $\mathscr{F}_t$ и σ-алгебру $\mathscr{F}$ пополнить, т.е. добавить к ним все подмножества нулевых множеств (см. стр. 436).

Определение. Сл.функция $\big\langle\, \xi(t),\ t \geqslant 0 \,\big\rangle$ называется *неупреждающей* относительно потока σ-алгебр $\big\langle\, \mathscr{F}_t,\ t \geqslant 0 \big\rangle$, или *согласованной с фильтрацией*, если для $\forall\, t \geqslant 0$ сл.в. $\xi(t)$ измерима относительно $\mathscr{F}_t$.

Значения сл.функции, согласованной с потоком σ-алгебр, порождённым винеровским процессом, в каждый момент времени t полностью определяются отрезком траектории $w(s)$, $s \in [0;t]$. Для такой сл.функции можно вычислить характеристики интегральной суммы вида (28) с $\tilde{t}_k = t_{k-1}$. По свойству условного математического ожидания для $\forall\, k = 0, \ldots, n-1$ ковариация

$$\mathbf{E}\left[\xi(t_k)\big(w(t_{k+1}) - w(t_k)\big)\right] = \mathbf{E}\Big[\xi(t_k)\,\mathbf{E}\big[\,\big(w(t_{k+1}) - w(t_k)\big) \mid \mathscr{F}_{t_k}\big]\Big] \overset{*}{=}$$
$$\overset{*}{=} \mathbf{E}\left[\xi(t_k)\,\mathbf{E}\,\big(w(t_k) - w(t_{k-1})\big)\right] = 0,$$

где равенство $\overset{*}{=}$ справедливо в силу независимости приращения $w(t_{k+1}) -$

$-\ w(t_k)$ от $\mathscr{F}_{t_k}$. Аналогично, для $j<k$

$$\begin{aligned}\mathbf{E}\big[\ \xi(t_j)\big(w(t_{j+1})-w(t_j)\big)\xi(t_k)\big(w(t_{k+1})-w(t_k)\big)\big]\ &=\\ =\ -\mathbf{E}\Big[\xi(t_j)\big(w(t_{j+1})-w(t_j)\big)\xi(t_k)\,\mathbf{E}\big[\,\big(w(t_{k+1})-w(t_k)\big)\mid \mathscr{F}_{t_k}\,\big]\Big]\ &=\ 0,\end{aligned}$$

снова, т.к. множитель $\xi(t_j)\big(w(t_{j+1})-w(t_j)\big)\xi(t_k)$ измерим относительно $\mathscr{F}_{t_k}$, а приращение $\big(w(t_{k+1})-w(t_k)\big)$ не зависит от $\mathscr{F}_{t_k}$. По тем же соображениям

$$\begin{aligned}\mathbf{E}\left[\xi^2(t_k)\big(w(t_{k+1})\ -w(t_k)\big)^2\right]\ &=\ \mathbf{E}\Big[\xi^2(t_k)\,\mathbf{E}\big[\,\big(w(t_{k+1})-w(t_k)\big)^2\mid \mathscr{F}_{t_k}\,\big]\Big]\ =\\ =\ \mathbf{E}\left[\xi^2(t_k)\ \mathbf{E}\big(w(t_k)-w(t_{k-1})\big)^2\right]\ &=\ \mathbf{E}\,\xi^2(t_k)(t_{k+1}-t_k).\end{aligned}$$

Следовательно, математическое ожидание $\mathbf{E}\,S_n=0$, а дисперсия

$$\mathbf{D}S_n=\sum_{k=0}^{n-1}\mathbf{E}\,\xi^2(t_k)(t_{k+1}-t_k)\qquad\Big(\approx\int_0^T\mathbf{E}\,\xi^2(t)\,dt\Big).$$

Обозначим предел в метрике пространства $\mathcal{L}_2(\Omega,\mathscr{F},\mathbf{P})$ через $\mathrm{l.i.m.}_\Omega$ (от *англ.* limit in mean square).

409| **Теорема.** *Пусть* $\boldsymbol{\xi}=\big\langle\,\xi(t),\ t\in[0;T]\,\big\rangle$ *— процесс, согласованный с фильтрацией, определяемой винеровским процессом* $\big\langle\,w(t),\ t\in[0;T]\,\big\rangle$. *Если функция* $\mathbf{E}\,\xi^2(t)$ *непрерывна на отрезке* $t\in[0;T]$, *то для последовательности интегральных сумм вида (28) существует*

$$J_w(\boldsymbol{\xi})\ =\ \int_0^T\xi(t)\,dw(t)\ :=\ \mathrm{l.i.m.}_\Omega\,S_n$$

для любой последовательности разбиений $0=t_0=\tilde t_{1n}<t_{1n}=\tilde t_{2n}<\ldots<$ $<t_{(n-1)n}=\tilde t_{nn}<t_{nn}=T$ *такой, что* $\max\{t_{(k+1)n}-t_{kn},\ k=0,\ldots,n-1\}\to 0$ *при* $n\to\infty$.

410| Пример. Покажем, что

$$J_w(\boldsymbol{w})\ =\ \int\limits_0^T w(t)\,dw(t)\ =\ \frac{w^2(T)}{2}-\frac{T}{2}. \tag{29}$$

Выберем $t_{kn}=kT/n, k=0,\ldots,n$; тогда интегральная сумма

$$\begin{aligned}S_n\ &=\ \sum_{k=1}^n w\Big(\frac{(k-1)T}{n}\Big)\Big(w\Big(\frac{kT}{n}\Big)-w\Big(\frac{(k-1)T}{n}\Big)\Big)\ =\\ &=\ \frac12\sum_{k=1}^n\left[\Big(w^2\Big(\frac{kT}{n}\Big)-w^2\Big(\frac{(k-1)T}{n}\Big)\Big)-\Big(w\Big(\frac{kT}{n}\Big)-w\Big(\frac{(k-1)T}{n}\Big)\Big)^2\right]\ =\end{aligned}$$

$$= \tfrac{1}{2}\big(w^2(T) - w^2(0)\big) - \tfrac{1}{2}\sum_{k=1}^{n}\left[\left(w\Big(\frac{kT}{n}\Big) - w\Big(\frac{(k-1)T}{n}\Big)\right)^2\right].$$

Как отмечалось в (22), среднее значение последней суммы равно T, а дисперсия $T^2/n \underset{n}{\to} 0$, т.е. $\mathrm{l.i.m.}_\Omega\, S_n = (w^2(T) - T)/2$. Обратим внимание, что только с «неожиданной» добавкой $-T/2$ математическое ожидание правой части (29), как и положено, равно нулю. ◎

⚠ Определяя стохастический интеграл $\int \xi(t)\,dw(t)$, мы выбирали точки $\tilde{t}_k$ для сечений процесса $\xi(t)$ на левом краю интервалов разбиения области интегрирования. Такой способ приводит к построению так называемого *интеграла Ито*. Для некоторых сл.функций (например, $\xi(t) = h(w(t))$) можно выбирать $\tilde{t}_k$ и на правом краю интервалов разбиения, а также положить $\tilde{t}_k$ равными серединам интервалов. Последний вариант называется *интегралом Стратановича*, и для него $\int_0^T w(t)\,dw(t) = \dfrac{1}{2}\,w^2(T)$.

⚠ Понятие интеграла Ито допускает расширение. Для этого расширения необходимо на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$:

✦ задать некоторую фильтрацию $\mathcal{F} = \big\langle\, \mathscr{F}_t, t \in [0;T]\,\big\rangle$, удовлетворяющую *стандартным условиям*:

i) σ-алгебры $\mathscr{F}, \mathscr{F}_t, t \in [0;T]$ являются $\mathbf{P}$-полными;

ii) $\mathscr{F}_0$ содержит все $\mathbf{P}$-нулевые множества $\mathscr{F}$;

iii) $\mathcal{F}$ непрерывна справа, т.е. $\mathscr{F}_t = \mathscr{F}_{t+} = \bigcap_{s>t} \mathscr{F}_s,\quad t \in [0;T)$;

✦ рассмотреть процесс Винера $w(t), t \in [0;T]$, согласованный с фильтрацией $\mathcal{F}$, приращения которого $w(t) - w(s)$ для $t > s$ не зависят от $\mathscr{F}_s$;

✦ рассмотреть случайные функции (процессы), *прогрессивно измеримые* относительно фильтрации $\mathcal{F}$, т.е. процессы $\xi(t,\omega), t \in [0;T], \omega \in \Omega$, которые при каждом s измеримы по обеим переменным t и ω как функции $\xi(t,\omega)$: $\big([0;s] \times \Omega, \mathscr{B}[0;s] \otimes \mathscr{F}_s\big) \mapsto (\mathbb{R}^1, \mathscr{B})$. Заметим, что прогрессивно измеримыми являются процессы, согласованные с фильтрацией, у которых все траектории непрерывны справа (или слева).

Интеграл Ито $\int_0^T \xi(t)\,dw(t)$ определяется по схеме, близкой к схеме определения интеграла Лебега, для любого прогрессивно измеримого процесса с $\int_0^T \mathbf{E}\,\xi^2(t)\,dt < \infty$ (или $\mathbf{P}\big\{ \int_0^T \xi^2(t)\,dt < \infty \big\} = 1$).

✧ **Свойства интеграла Ито. Стохастический дифференциал.** Рассмотрим стандартный винеровский процесс $\boldsymbol{w} = \big\langle\, w(t),\ t \in [0;T]\,\big\rangle$. Пусть

$\xi(t), \eta(t)$ — неупреждающие сл.функции (относительно естественного потока σ-алгебр $\mathcal{F}^{\boldsymbol{w}} = \big\langle \mathscr{F}_t = \sigma\big(w(u), u \leqslant t\big), t \in [0;T] \big\rangle\big)$, для которых функции вторых моментов $\mathbf{E}\,\xi^2(t)$, $\mathbf{E}\,\eta^2(t)$ непрерывны на $t \in [0;T]$. Имеют место следующие свойства (для $\forall\, a, b \in \mathbb{R}^1$) :

(✓) $\int_0^T \big(a\xi(t) + b\eta(t)\big)\, dw(t) = a \int_0^T \xi(t)\, dw(t) + b \int_0^T \eta(t)\, dw(t);$

(✓) $\mathbf{E}\Big[\int_0^T \xi(t)\, dw(t)\Big] = 0;$

(✓) ковариация

$$\mathbf{Cov}\,\Big(\int_0^T \xi(t)\, dw(t), \int_0^T \eta(t)\, dw(t)\Big) = \int_0^T \mathbf{E}\,\big[\xi(t)\eta(t)\big]\, dt,$$

в частности $\mathbf{D}\Big[\int_0^T \xi(t)\, dw(t)\Big] = \int_0^T \mathbf{E}\,\xi^2(t)\, dt < \infty\,;$

(✓) сл.функция

$$\zeta(s) = \int_0^s \xi(t)\, dw(t) := \int_0^T \big(\,\dot{\mathrm{I}}_{\{t \leqslant s\}}(t)\xi(t)\big)\, dw(t),\ \ s \in [0;T],$$

имеет ортогональные приращения, почти наверное непрерывные траектории и, кроме того, образует мартингал относительно фильтрации $\mathcal{F}^{\boldsymbol{w}}$, т.е. для $\forall\, s \leqslant t$ условное математическое ожидание $\mathbf{E}\big[\,\zeta(t)\,\big|\,\mathscr{F}_s\big] = \zeta(s)$ (п.н.) .

► Полагая в (29) $T = t$ и $Y(t) = w^2(t)/2$, получаем

$$Y(t) = \int_0^t w(s)\, dw(s) + \frac{1}{2}\, t. \tag{30}$$

В классическом анализе это уравнение (с учётом непрерывности $w(t)$ и условия $w(0) = 0$) эквивалентно дифференциальному уравнению

$$dY(t) = w(t)\, dw(t) + \frac{1}{2}\, dt.$$

В стохастическом анализе поступают аналогичным образом: говорят, что сл.функция $\boldsymbol{\xi} = \big\langle\, \xi(t), t \in [0;T]\,\big\rangle$ удовлетворяет *стохастическому дифференциальному уравнению* (или *имеет стохастический дифференциал*)

$$d\xi(t) = f(t)\, dw(t) + g(t)\, dt,$$

где сл.функции f, g согласованы с естественной фильтрацией $\mathcal{F}^{\boldsymbol{w}}$, если с вероятностью единица для $\forall\, t \in [0;T]$

$$\xi(t) - \xi(0) = \int_0^t f(s)\, dw(s) + \int_0^t g(s)\, ds.$$

Минимальные требования, при которых можно обосновать корректность такого определения, заключаются в выполнении с вероятностью единица условий:

$$\int_0^T |g(s)|\,ds < \infty, \qquad \int_0^T f^2(s)\,ds < \infty.$$

Винеровский процесс, очевидно, имеет стохастический дифференциал $dw(t) = 1 \cdot dw(t) + 0 \cdot dt$.

411 **Теорема.** [*Формула Ито.*] *Пусть* $h(t,x), t \in [0;T],\ x \in \mathbb{R}^1$, — *непрерывная функция с непрерывными производными*

$$h'_t(t,x) = \frac{\partial}{\partial t} h(t,x), \quad h'_x(t,x) = \frac{\partial}{\partial x} h(t,x), \quad h''_{xx}(t,x) = \frac{\partial^2}{\partial x^2} h(t,x).$$

Предположим, что сл.функция $\xi(t)$ *имеет стохастический дифференциал*

$$d\xi(t) = f(t)\,dw(t) + g(t)\,dt,$$

тогда сл.функция $h(t,\xi_t)$ *имеет стохастический дифференциал*

$$dh(t,\xi_t) = \Big[h'_t(t,\xi_t) + g(t)h'_x(t,\xi_t) + \frac{1}{2} f^2(t) h''_{xx}(t,\xi_t)\Big]\,dt \ + \ f(t)h'_x(t,\xi_t)\,dw_t\,.$$

В частности, если $h(t,x) = H(x)$, *то*

$$dH(\xi_t) = \big[g(t)H'(\xi_t) + \frac{1}{2} f^2(t) H''(\xi_t)\big]\,dt + f(t)H'(\xi_t)\,dw(t),$$
$$dH(w_t) = \frac{1}{2} H''(w_t)\,dt + H'(w_t)\,dw(t)$$

при соблюдении соответствующих условий к функции H.

⚠ В классическом анализе, когда функция $w(t)$ имеет ограниченную вариацию, слагаемое со второй производной h по пространственной переменной отсутствует. Приведём неформальные соображения (для функции $H(x)$), показывающие причину появления этого слагаемого при стохастическом исчислении. Рассмотрим разбиение отрезка $[0;t]$ на интервалы Δt_i, тогда по формуле Тейлора

$$\begin{aligned} H(\xi(t_i)) - H(\xi(t_{i-1})) &= H'\Delta\xi(t_i) + o\big(|\Delta\xi(t_i)|\big) = \\ &= H' f(t_i)\Delta w(t_i) + H' g(t_i)\Delta t_i + o\big(|\Delta t_i| + |\Delta w(t_i)|\big). \end{aligned}$$

Таким образом,

$$H(\xi(t)) = \sum_i H' g(t_i)\Delta t_i + \sum_i H' f(t_i)\Delta w(t_i) + \sum_i o\big(|\Delta t_i|\big) + \sum_i o\big(|\Delta w(t_i)|\big).$$

Ввиду бесконечной вариации функции $w(t)$ вторая сумма с o здесь уходит на ∞, что не позволяет перейти к пределу.

Проведём аналогичные построения, воспользовавшись разложением в ряд Тейлора функции $H(x)$ до второй производной:

$$H(\xi(t_i)) - H(\xi(t_{i-1})) = H'\big(\xi(t_i)\big)\Delta\xi(t_i) + \frac{1}{2}H''\big(\xi(t_i)\big)\big[\Delta\xi(t_i)\big]^2 + o\big((\Delta\xi(t_i))^2\big) =$$
$$= H'\big(\xi(t_i)\big)g(t_i)\Delta t_i + \frac{1}{2}H''\big(\xi(t_i)\big)f^2(t_i)(\Delta w(t_i))^2 + H'\big(\xi(t_i)\big)f(t_i)\Delta w(t_i) + Z_i$$

с соответствующим остатком Z_i. Как и при вычислении квадратической вариации (см. (22), стр. 355), можно показать, что математическое ожидание

$$\mathbf{E}\Big[\sum\nolimits_i H''\big(\xi(t_i)\big)f^2(t_i)(\Delta w(t_i))^2 - \sum\nolimits_i H''\big(\xi(t_i)\big)f^2(t_i)\Delta t_i\Big] = 0,$$

а дисперсия этой разности стремится к нулю при измельчении разбиений. Итак,

$$\mathrm{l.i.m.}_{\Omega}\sum\nolimits_i H''\big(\xi(t_i)\big)f^2(t_i)\big(\Delta w(t_i)\big)^2 = \int_0^t H''\big(\xi(s)\big)f^2(s)\,ds,$$

что и объясняет появление дополнительного слагаемого. Если бы функция $w(t)$ имела ограниченную вариацию, то последний предел равнялся бы нулю.

412| Примеры. 1) Пусть $H(x) = x^2/2$. Тогда $H' = x$, $H'' = 1$. Согласно формуле Ито, стохастический дифференциал для процесса $H(w(t)) = w^2(t)/2$ равен

$$d\frac{1}{2}w^2(t) = w(t)\,dw(t) + \frac{1}{2}\,dt,$$

что совпадает с (30).

2) Для функции $h = tx$ производные $h'_t = x$, $h'_x = t$, $h''_{xx} = 0$. В соответствии с формулой Ито $d(tw(t)) = w(t)\,dt + t\,dw(t)$, т.е.

$$tw(t) = \int_0^t w(s)\,ds + \int_0^t s\,dw(s),$$

что усиливает (27).

3) В классическом анализе решение дифференциального уравнения $dy(t) = y(t)b(t)dt$ даёт функция $y(t) = \exp\big\{\int_0^t b(s)\,ds\big\}$. При стохастическом исчислении аналогичное уравнение приводит к так называемой стохастической экспоненте. Пусть $b(t)$ — процесс, согласованный с естественной фильтрацией $\mathscr{F}^{\boldsymbol{w}}$, для которого п.н. $\int_0^T b^2(t)dt < \infty$, $\xi(t)$ — процесс, удовлетворяющий стохастическому дифференциальному уравнению

$$d\xi(t) = b(t)\,dw(t) - \frac{1}{2}b^2(t)\,dt.$$

Тогда по формуле Ито, применённой к функции e^x, получаем, что процесс

$$\eta(t) = \exp\left\{ \int_0^t b(s)\,dw(s) - \frac{1}{2}\int_0^t b^2(s)\,ds \right\}, \qquad t \in [0;T],$$

удовлетворяет стохастическому дифференциальному уравнению $d\eta(t) = \eta(t)b(t)\,dw(t)$.

При некоторых дополнительных условиях к процессу $b(t)$ можно показать (см., например, [4], предложение 6.1), что для $\forall\, t_1 < t_2$

$$\mathbf{E}\big[\,\eta(t_2)\,\big|\,\mathscr{F}_{t_1}\big] = \eta(t_1),$$

т.е. процесс $\boldsymbol{\eta}$ образует мартингал (см. ниже).

4) Рассмотрим уравнение Ланжевена, описывающее изменение скорости движения $V(t)$ частицы в некоторой вязкой среде в присутствии внешней силы:

$$dV(t) = -\alpha V(t)dt + \lambda\,dw(t), \tag{31}$$

где коэффициент α пропорционален вязкости среды, dw — «шумовая» составляющая, обусловленная внешним воздействием на частицу, λ — коэффициент интенсивности внешней силы. Пусть V_0 — начальная скорость частицы. Если бы «шум» был не случайным, т.е. $dw(t) = g(t)dt$ с детерминированной функцией g, то метод решения такого уравнения хорошо известен: рассматривается однородное уравнение $dV(t) = -\alpha V(t)dt$, которое даёт $V(t) = ce^{-\alpha t}$. Подставляя это решение с множителем $c = c(t)$ в исходное уравнение, получаем

$$V(t) = V_0 e^{-\alpha t} + \lambda \int_0^t e^{\alpha(s-t)} g(s)\,ds.$$

Покажем, что, произведя здесь замену $g(s)\,ds \dashrightarrow dw(s)$, мы получим решение стохастического дифференциального уравнения (31).

Пусть $\xi(t) = \int_0^t e^{\alpha s} dw(s)$, т.е. стохастический дифференциал $d\xi(t) = e^{\alpha t} dw(t) + 0dt$. Применим формулу Ито к функции $h(t,x) = \lambda e^{-\alpha t} x$, для чего вычислим необходимые для этого производные:

$$h'_t = -\alpha h, \quad h'_x = \lambda e^{-\alpha t}, \quad h''_{xx} = 0.$$

Для сл.функции $Y(t) = h(t,\xi_t) = \lambda \int_0^t e^{-\alpha(t-s)} dw(s)$ по формуле Ито получаем

$$dY(t) \;=\; -\alpha h(t,\xi_t)dt + e^{\alpha t}\lambda e^{-\alpha t} dw(t) \;=\; -\alpha Y(t)dt + \lambda dw(t).$$

Если сл.в. V_0 измерима относительно σ-алгебры $\mathscr{F}_0 = \sigma(w(0))$, то, очевидно, процесс $X(t) = V_0 e^{-\alpha t}$ неупреждающий и имеет дифференциал

$$dX(t) = -\alpha X(t)dt \ + 0\, dw(t).$$

Легко видеть из свойства линейности стохастических интегралов, что стохастический дифференциал суммы двух сл.функций имеет стохастический дифференциал с суммой соответствующих коэффициентов. Поэтому процесс

$$V(t) = V_0 e^{-\alpha t} + \lambda \int_0^t e^{-\alpha(t-s)}\, dw(s)$$

даёт решение уравнения (31) с начальным условием $V(0) = V_0$, если сл.в. V_0 измерима относительно σ-алгебры $\mathscr{F}_0$, т.е. начальная скорость полностью определяется начальным импульсом силы $w(0)$, например, для стандартного винеровского процесса V_0 есть детерминированная величина.

Не вдаваясь в подробности, скажем, что если $V_0 \backsim \mathcal{N}\big(0, {}^{\lambda^2}\!/_{2\alpha}\big)$, то процесс $V(t)$ гауссовский с нулевым средним и ковариацией

$$\mathbf{Cov}(V_t, V_s) = \frac{\lambda^2}{2\alpha} e^{-\alpha|t-s|}.$$

Гауссовский процесс с такой ковариационной функцией называется процессом Орнштейна–Уленбека. Напомним, что интеграл от функции мгновенной скорости объекта даёт координату положения этого объекта в рассматриваемый момент времени. Другими словами, потраекторный интеграл от процесса Орнштейна–Уленбека

$$\tilde{w}(t) = \int_0^t V(s)ds, \qquad t \geqslant 0,$$

можно рассматривать как вариант описания процесса движения броуновских частиц с учётом инерции. При этом процесс $\tilde{w}$ становится не только непрерывным, но и дифференцируемым. ◎

413| Упр. Покажите, что если $w(t)$ — стандартный винеровский процесс, то $\boldsymbol{V} = \left\langle\, e^{-\alpha t}\, w\big(\frac{\lambda^2}{2\alpha} e^{\alpha t}\big),\ t \in \mathbb{R}^1 \,\right\rangle$ есть процесс Орнштейна–Уленбека.

§ 3. Пуассоновский процесс

Пуассоновский процесс описывает поведение так называемого простого потока событий. Под потоком событий понимается, возможно неоднократное, появление события какого-то фиксированного типа в случайные моменты времени за наблюдаемый период. При этом учитываются только факт и время осуществления события без какой-либо дополнительной «окраски», например, взятие ворот той или иной командой за 60 мин хоккейного матча — название клуба, фамилия автора гола и другая информация игнорируются. В такой ситуации с потоком событий можно связать сл.процесс $\left\langle \xi(t), t\in[0;T] \right\rangle$, где $\xi(t)$ — число событий, произошедших к моменту t (включительно). Понятно, что все траектории процесса не убывают и имеют вид непрерывной справа ступенчатой линии со скачками единичной высоты в случайных точках.

Случайный процесс, описывающий *простой поток событий*, должен иметь независимые приращения и удовлетворять условию однородности (т.е. распределение приращения $\xi(t+\Delta)-\xi(t)$ зависит только от Δ и не зависит от t) и условию ординарности — разреженности. Под последним понимается «почти невозможность» осуществления более одного события за малый промежуток времени.

Определение. Неубывающий сл.проц. $\left\langle \xi(t), t\in[0;T] \right\rangle$ с дискретным фазовым пространством $\mathscr{X}=\mathbb{N}_0$ удовлетворяет условию *ординарности* (*разреженности*), если для $\forall t\in[0;T]$ и $\Delta\to 0$

$$\begin{aligned}
&\text{i)} & \mathbf{P}\{\xi(t+\Delta)-\xi(t)\geqslant 2\} &= o(\Delta),\\
&\text{ii)} & \mathbf{P}\{\xi(t+\Delta)-\xi(t)=1\} &= \lambda\Delta+o(\Delta),\\
&\text{iii)} & \mathbf{P}\{\xi(t+\Delta)-\xi(t)=0\} &= 1-\lambda\Delta-o(\Delta).
\end{aligned}$$

✧ **Распределение простого потока событий**. Можно, конечно, считать начальное число событий также случайным; мы рассмотрим только ситуации, когда $\xi(0)\equiv 0$. Для любого возрастающего набора временны́х параметров и любых значений из пространства $\mathscr{X}$ можно записать (с $t_0=0, x_0=0$):

$$\mathbf{P}\Big\{\bigcap\nolimits_1^k \big(\xi(t_j)=x_j\big)\Big\} \;=\; \mathbf{P}\Big\{\bigcap\nolimits_1^k \big(\xi(t_j)-\xi(t_{j-1})=x_j-x_{j-1}\big)\Big\}.$$

Во-первых, заметим, что в силу однородности процесса вероятность появления заданного количества событий за временной промежуток длины s не зависит от начала отсчёта:

$$\mathbf{P}\{\xi(t+s)-\xi(t)=k\} \;=\; \mathbf{P}\{\xi(s)-\xi(0)=k\} \;=:p_k(s),$$

для $\forall t,s\in[0;T], k=0,1,\dots$ Поэтому, учитывая независимость приращений, можно представить искомую конечномерную вероятность в виде

$$\mathbf{P}\{\xi(t_1)=x_1,\ldots,\xi(t_k)=x_k\} \;=\; \prod_{j=1}^{k} p_{x_j-x_{j-1}}(t_j-t_{j-1}), \tag{32}$$

$$0=t_0<t_1<\ldots<t_k,\quad 0=x_0\leqslant x_1\leqslant\ldots\leqslant x_k\,;$$

для остальных значений аргументов конечномерная вероятность или равна нулю, если порядок расположения моментов времени не совпадает с порядком количеств событий (например, $t_1<t_2$, но $x_1>x_2$), или вычисляется по той же формуле после предварительного упорядочения моментов времени.

Дискретное распределение (32) удовлетворяет условию перестановочности по построению. Для проверки условия согласованности требуется маргинальное распределение. Как известно, для его вычисления необходимо просуммировать вероятности (32) по всем значениям исключаемых сл.величин. Например, если исключить сл.в. $\xi(t_1)$, то условие согласованности с учётом равенства нулю соответствующих вероятностей будет выглядеть следующим образом:

$$\sum_{x_1=0}^{x_2}\prod_{j=1}^{k} p_{x_j-x_{j-1}}(t_j-t_{j-1}) \;=\; p_{x_2}(t_j)\prod_{j=3}^{k} p_{x_j-x_{j-1}}(t_j-t_{j-1}).$$

Аналогичное соотношение верно и для других (одномерных и многомерных) распределений. Поэтому в соответствии с теоремой Колмогорова справедлива

414| **Теорема.** *Пусть $p_x(s)$, $s>0$, $x=0,1,\ldots,$ — семейство дискретных распределений на $\mathbb{N}_0$: $p_x(s)\geqslant 0$, $\sum_{x=0}^{\infty}p_x(s)=1$. Если*

$$\sum_{k=0}^{x} p_k(s)p_{x-k}(t-s)=p_x(t),\qquad \forall\, 0<s\leqslant t,\ x\in\mathbb{N}_0,$$

то это семейство определяет неубывающий дискретный однородный процесс с независимыми приращениями, конечномерные распределения которого находятся по формуле (32).

415| Упр. Проверьте справедливость условий теоремы для пуассоновского семейства $p_x(s)=e^{-s}s^x/x!$, $x\in\mathbb{N}_0$, $s>0$.

Покажем теперь, что условие ординарности процесса автоматически приводит к пуассоновскому распределению сечений.

416| **Теорема.** *Если неубывающий дискретный однородный процесс*

$\left\langle \xi(t),\ t \geqslant 0 \right\rangle$ *имеет независимые приращения и удовлетворяет условию ординарности, то*[‡]

$$p_x(t) \ \stackrel{\mathrm{df}}{=}\ \mathbf{P}\{\xi(t) = x\} \ = \ \frac{(\lambda t)^x e^{-\lambda t}}{x!}, \qquad x = 0, 1 \ldots$$

$\Rrightarrow$ Легко понять, что для любого целого положительного x справедливо

$$\{\xi(t+\Delta) = x\} \ = \ \biguplus_{k=0}^{x} \Big(\{\xi(t) = x-k\} \bigcap \{\xi(t+\Delta) - \xi(t) = k\} \Big).$$

Отсюда в силу независимости приращений и однородности имеем

$$p_x(t+\Delta) = \sum\nolimits_{k=0}^{x} p_{x-k}(t) p_k(\Delta).$$

Воспользовавшись ординарностью, перепишем равенство в виде

$$\begin{aligned} p_k(t+\Delta) = {} & p_x(t)(1-\lambda\Delta) + p_{x-1}(t)\lambda\Delta + \sum\nolimits_2^x p_{x-k}(t)p_k(\Delta) + o(\Delta) \ = \\ = {} & p_x(t) + \lambda(p_{x-1}(t) - p_x(t))\Delta + o(\Delta) \end{aligned}$$

при $\Delta \to 0$. Таким образом,

$$\frac{\partial}{\partial t} p_x(t) \ = \ \lim_{\Delta\to 0} \frac{1}{\Delta}\big(p_x(t+\Delta) - p_x(t)\big) \ = \ \lambda\big(p_{x-1}(t) - p_x(t)\big). \qquad (33)$$

Аналогичное равенство $\frac{\partial}{\partial t} p_0(t) = -\lambda p_0(t)$ справедливо и для $x = 0$, если положить $p_{-1}(t) = 0$. Решение этого простейшего дифференциального уравнения — $p_0(t) = Ce^{-\lambda t}$. Так как $\xi(0) \equiv 0$, то $C = p_0(0) = 1$. Применим теперь метод индукции по x. Пусть равенство справедливо при некотором натуральном x. Положим $H(t) = p_{x+1}(t)e^{\lambda t}$, тогда $p_{x+1}(t) = H(t)e^{-\lambda t}$ и

$$\frac{\partial}{\partial t} p_{x+1}(t) \ = \ H'(t)e^{-\lambda t} - \lambda p_{x+1}(t).$$

Из уравнения (33) при $x \dashrightarrow x+1$ по предположению индукции

$$H'(t)e^{-\lambda t} - \lambda p_{x+1}(t) \ = \ -\lambda p_{x+1}(t) + \lambda \frac{(\lambda t)^x e^{-\lambda t}}{x!},$$

т.е. $H'(t) = \lambda^{x+1} t^x / x!$, поэтому $H(t) = C + \frac{1}{(x+1)!}(\lambda t)^{x+1}$. Так как $H(0) = {}$ $= p_{x+1}(0)e^{\lambda 0} = 0$, то $C = 0$, что и требовалось. $\Lleftarrow$

С учётом утверждений 414 и 415 можно ввести несколько эквивалентных определений пуассоновского процесса — пуассоновского потока событий.

(‡) По договорённости $0^0 = 1$.

Определения. Неубывающий процесс $\boldsymbol{\xi} = \left\langle \xi(t),\ t \in [0;T] \right\rangle$ со значениями в множестве неотрицательных целых чисел ($\xi(0) = 0$) есть процесс Пуассона, если

а) $\xi(t)$ равно числу событий, произошедших за время t в простом потоке событий, т.е. $\boldsymbol{\xi}$ есть однородный процесс с независимыми приращениями, удовлетворяющий условию разреженности, или

б) $\boldsymbol{\xi}$ есть однородный процесс с независимыми приращениями, для которого каждое сечение имеет распределение Пуассона:

$$\mathbf{P}\{\xi(t) = x\} \ = \ \frac{(\lambda t)^x}{x!} e^{-\lambda t}, \qquad x \in \mathbb{N}_0,\ t > 0,$$

в) или конечномерное распределение $\boldsymbol{\xi}$ для $0 < t_1 < \ldots < t_n$ задаётся равенством

$$\mathbf{P}\{\xi(t_1) = x_1, \ldots, \xi(t_n) = x_n\} \ = \ \lambda^{x_n} e^{-\lambda t_n} \prod_{j=1}^{n} \frac{(t_j - t_{j-1})^{x_j - x_{j-1}}}{(x_j - x_{j-1})!},$$
$$0 \leqslant x_1 \leqslant \ldots \leqslant x_k, \qquad t_0 = 0,\ x_0 = 0.$$

Параметр λ называется *интенсивностью* потока.

⚠ В дальнейшем процесс Пуассона будет охарактеризован как однородная цепь Маркова (с непрерывным временем), для которой переходное распределение вероятностей за любой промежуток времени длины t описывается посредством распределения Пуассона с параметром λt.

417| Упр. Проверьте, что i) процесс из пункта б) определения удовлетворяет условию ординарности; ii) процесс с распределением из пункта в) имеет независимые приращения, однороден и разрежен.

✧ Свойства процесса Пуассона.

Из свойств распределения Пуассона следует, что функция среднего и дисперсионная функция процесса Пуассона равны соответственно $\mu(t) = \lambda t$, $\mathbb{V}(t) = \lambda t$. Таким образом, интенсивность потока λ равна среднему числу появившихся событий в единицу времени.

418| Упр. Воспользовавшись видом двумерного распределения $(\xi(s), \xi(t))$, покажите, что $\mathbf{E}\left[\xi(s)\xi(t)\right] = \lambda s + \lambda^2 st$ для $s \leqslant t$.

✦ Из утверждения 418 следует, что ковариация

$$\mathbf{Cov}(\xi(s), \xi(t)) = \lambda \min(s, t).$$

Впрочем, необходимо заметить, что в силу 393, стр. 347, этот факт является следствием независимости приращений процесса и предыдущего свойства.

✦ Для любого потока событий интересно время ожидания осуществления события, например время до первой поломки прибора.

419| **Теорема.** *Пусть $\tau = \min\{t : \ \xi(t) \geqslant 1\}$ — время до первого осуществления события пуассоновского потока. Тогда*

$$\tau \backsim \mathcal{E}\mathrm{x}\big({}^1/_\lambda\big), \quad \text{т.е.} \quad \mathbf{P}\{\tau \leqslant t\} \ = \ 1 - e^{-\lambda t}, \ t \geqslant 0.$$

⤇ Событие $t-\Delta < \tau \leqslant t$ означает, что $\xi(t-\Delta) = 0$ и $\xi(t)-\xi(t-\Delta) \geqslant 1$. Следовательно, при $\Delta \to 0$

$$\begin{aligned}\mathbf{P}\{t - \Delta < \tau \leqslant t\} \ &= \ p_0(t-\Delta)\big(p_1(\Delta) + \mathbf{P}\{\xi(\Delta) \geqslant 2\}\big) \ = \\ &= \ e^{-\lambda(t-\Delta)}(\lambda \Delta e^{-\lambda\Delta} + o(\Delta)).\end{aligned}$$

Другими словами, ф.р. τ имеет производную в любой т. t:

$$F'_\tau(t) \ = \ \lim_{\Delta\to 0} \frac{1}{\Delta} \mathbf{P}\{t - \Delta < \tau \leqslant t\} \ = \ \lambda e^{-\lambda t},$$

что характеризует показательный закон. ⤆

420| **Следствие.** *Среднее время ожидания события в простом потоке событий обратно пропорционально интенсивности потока:* $\mathbf{E}\tau = 1/\lambda$.

✦ В более общем виде: пусть $\mathcal{T}_k = \min\{t : \ \xi(t) \geqslant k\}$ — время ожидания осуществления k событий потока. Тогда $\tau_k = \mathcal{T}_k - \mathcal{T}_{(k-1)}$ — время между двумя последовательными ($(k-1)$-м и k-м) появлениями событий.

421| **Теорема.** *Для пуассоновского потока событий справедливы следующие утверждения:*

а) сл.в. $\tau_1, \ldots, \tau_n$ независимы в совокупности для $\forall\, n > 1$;

б) сл.в. $\tau_k \backsim \mathcal{E}\mathrm{x}\big({}^1/_\lambda\big)$.

⤇ В целях сокращения записи будем считать $\lambda = 1$. Необходимо показать, что плотность совместного распределения $(\tau_1, \ldots, \tau_n)$ равна $\prod_{k=1}^n e^{-t_k}$, $\vec{t} \geqslant 0$. Согласно 60, стр. 63, в качестве искомой плотности может выступать «правая» производная

$$F^{(n)}(t_1,\dots,t_n) \;=\; \lim_{\Delta\to 0}\frac{1}{\Delta^n}\,\mathbf{P}\left\{t_1<\tau_1\leqslant t_1+\Delta,\dots,t_n<\tau_n\leqslant t_n+\Delta\right\},$$

если $\int_{\mathbb{R}^n} F^{(n)}(\vec{t}\,)\,d\vec{t} \geqslant 1$. Положим $T_k = t_1+\ldots+t_k,\ k=\overline{1,n}$. Ясно, что при достаточно малом $\Delta>0$ условие под знаком вероятности будет выполнено, если в каждом из интервалов $(T_k;T_k+\Delta],\ k=\overline{1,n}$, произойдёт по одному событию из потока, а вне этих интервалов таких событий не будет; см. таблицу, в которой приведены также вероятности соответствующих событий:

	t_1	$t_1+\Delta$	t_1+t_2	$t_1+t_2+\Delta$	$t_1+t_2+t_3$	$t_1+t_2+t_3+\Delta$
0	1	0	1	0	1	$\cdots$
e^{-t_1}	$\Delta e^{-\Delta}$	$e^{t_2-\Delta}$	$\Delta e^{-\Delta}$	$e^{t_3-\Delta}$	$\Delta e^{-\Delta}$	$\cdots$

Другими словами, вероятность

$$\begin{aligned}\mathbf{P}\left\{t_1<\tau_1\leqslant t_1+\Delta,\dots,t_n<\tau_n\leqslant t_n+\Delta\right\} \;&\geqslant\\ \geqslant e^{-t_1}\prod_{k=1}^{n}\left(\Delta e^{-\Delta}\right)\prod_{k=2}^{n}e^{-(t_k-\Delta)} &= \Delta^n e^{-\Delta}\prod_{k=1}^{n}e^{-t_k}.\end{aligned}$$

Таким образом, производная $F^{(n)}(t_1,\dots,t_n) \geqslant \prod_1^n e^{-t_k}$. В правой части здесь стоит плотность n-мерного закона с независимыми показательными $\mathcal{E}\mathrm{x}(1)$ компонентами. Следовательно, условия теоремы 60, стр. 63, выполнены и $F^{(n)}(t_1,\dots,t_n)$ есть плотность вектора $(\tau_1,\dots,\tau_n)$. Ясно, что строгое неравенство $F^{(n)}(t_1,\dots,t_n) > \prod_1^n e^{-t_k}$ между плотностями в этом случае может иметь место только на множестве лебеговой меры нуль. Поэтому в качестве плотности $(\tau_1,\dots,\tau_n)$ можно взять $\prod_1^n e^{-t_k}$. ⊲

✦ Справедлив и обратный результат. Как всегда, минимум пустого множества полагаем равным $+\infty$.

422| **Теорема.** *Пусть $\tau_k \backsim \mathcal{E}\mathrm{x}({}^1/_\lambda)$, $k=1,2,\dots,$ — последовательность независимых одинаково распределённых показательных сл.в. Сл.процесс*

$$\xi(t) = \min\left\{n:\ \textstyle\sum_1^n \tau_k > t\right\} - 1 \tag{34}$$

есть процесс Пуассона с интенсивностью λ.

⚠ Процессы, определяемые соотношениями типа (34), называются процессами *восстановления*, что связано со следующей интерпретацией. Пусть

τ_k — время безаварийной работы прибора после предыдущей поломки (и окончания ремонта). Тогда $\xi(t)$ — количество остановок на ремонт (восстановление) за время эксплуатации прибора t без учёта времени на ремонт.

$\Rrightarrow$ Положим $\lambda = 1$ и рассмотрим только два момента времени $t_1 = s$ и $t_2 = t$ (в общем случае $t_1 < \ldots < t_n$, $n \geqslant 2$, $\lambda > 0$, применяются аналогичные построения). Найдём совместное распределение сл.в. $\xi(s)$ и приращения процесса $\xi(t) - \xi(s)$ для $s < t$. Для произвольных целых $m > 0, k > 1$ определим сл.в. $\zeta_m = \tau_1 + \ldots + \tau_m$, которая имеет гамма-распределение $\mathcal{G}\mathtt{am}(m, 1)$, и сл.в. $\beta_{k-1} = \tau_{m+2} + \ldots + \tau_{m+k} \backsim \mathcal{G}\mathtt{am}(k-1, 1)$. Кроме того, пусть $\alpha_1 = \tau_{m+1}$, $\gamma_1 = \tau_{m+k+1}$. Тогда, очевидно, событие $\{\xi(s) = m\} \bigcap \{\xi(t) - \xi(s) = k\}$ эквивалентно следующим неравенствам:

$$\zeta_m \leqslant s < \zeta_m + \alpha_1 \leqslant \zeta_m + \alpha_1 + \beta_{k-1} \leqslant t < \zeta_m + \alpha_1 + \beta_{k-1} + \gamma_1,$$

где сумма индексов сл.в. как раз показывает количество переменных τ_j, накопленная сумма которых попадает в соответствующие области: m от 0 до s, $(m+k)$ от s до t, $(m+k+1)$ после t. В свою очередь, эта система неравенств эквивалентна системе

$$\left.\begin{array}{rcccl} t - \zeta_m - \alpha_1 - \beta_{k-1} & < & \gamma_1 & < & \infty, \\ 0 & \leqslant & \beta_{k-1} & \leqslant & t - \zeta_m - \alpha_1, \\ s - \zeta_m & < & \alpha_1 & < & t - \zeta_m, \\ 0 & \leqslant & \zeta_m & \leqslant & s. \end{array}\right\} \tag{35}$$

Таким образом, искомая вероятность равна интегралу по области (35) от произведения плотностей (ввиду независимости), участвующих в её определении сл.величин. В целях большей наглядности переменные интегрирования в следующих соотношениях совпадают со сл.величинами, а индексы опущены:

$$\int\limits_0^s \frac{\zeta^{m-1}}{(m-1)!} e^{-\zeta}\, d\zeta \int\limits_{s-\zeta}^{t-\zeta} e^{-\alpha}\, d\alpha \int\limits_0^{t-\zeta-\alpha} \frac{\beta^{k-2}}{(k-2)!} e^{-\beta}\, d\beta \int\limits_{t-\zeta-\alpha-\beta}^{\infty} e^{-\gamma}\, d\gamma =$$

$$= \frac{1}{(k-2)!(m-1)!} e^{-t} \int\limits_0^s \zeta^{m-1}\, d\zeta \int\limits_{s-\zeta}^{t-\zeta} d\alpha \int\limits_0^{t-\zeta-\alpha} \beta^{k-2}\, d\beta =$$

$$= \frac{1}{(k-1)!(m-1)!} e^{-t} \int\limits_0^s \zeta^{m-1}\, d\zeta \int\limits_{s-\zeta}^{t-\zeta} (t - \zeta - \alpha)^{k-1}\, d\alpha =$$

$$= \frac{1}{k!(m-1)!} e^{-t}(t-s)^k \int\limits_0^s \zeta^{m-1}\, d\zeta = \frac{1}{k!m!} e^{-t}(t-s)^k s^m =$$

$$= \frac{1}{m!} s^m\, e^{-s} \cdot \frac{1}{k!}(t-s)^k\, e^{-(t-s)}.$$

Следовательно, сл.в. $\xi(s), \big(\xi(t)-\xi(s)\big)$ имеют распределение Пуассона с параметрами s и $t-s$ соответственно и независимы. Кроме того, распределение приращения зависит только от длины интервала, т.е. процесс однороден, что и требовалось (см. определение 6), стр. 378). ⇚

✦ Другая характеристика, которой часто интересуются в связи с процессом, считающим события простого потока, — время достижения заданного уровня. Например, при движении по трассе стирание протектора автошины (скажем, на одинаковую величину) происходит при резких торможениях. Дальность поездки тогда напрямую зависит от количества таких торможений.

423| **Теорема.** *Пусть* $\mathscr{T} = \min\{t :\ \xi(t) \geqslant m\}$ *— время первого осуществления* m $(\in \mathbb{N})$ *событий в простом потоке интенсивности* λ. *Тогда* $\mathscr{T} \backsim \mathrm{Gam}(m, {}^1/_\lambda)$, *т.е.* $\mathscr{T}$ *имеет гамма-распределение с параметрами формы* m *и масштаба* ${}^1/_\lambda$.

⇛ Легко понять, что неравенство $\mathscr{T} \leqslant t$ выполняется т. т. т. когда $\xi(t) \geqslant m$. Поэтому ввиду пуассоновского распределения $\xi(t)$ (с параметром $u = \lambda t$) функция распределения

$$F(t) := \mathbf{P}\{\mathscr{T} \leqslant t\} = 1 - \mathbf{P}\{\xi(t) \leqslant m-1\} =$$

$$= 1 - e^{-u} - ue^{-u} - \frac{u^2}{2}e^{-u} - \ldots - \frac{u^{(m-1)}}{(m-1)!}e^{-u}.$$

Производная (по u) этой функции

$$F'_u(t) = e^{-u} - e^{-u} + ue^{-u} - ue^{-u} + \ldots - \frac{u^{(m-2)}}{(m-2)!}e^{-u} + \frac{u^{(m-1)}}{(m-1)!}e^{-u} =$$

$$= \frac{u^{(m-1)}}{(m-1)!}e^{-u} = \frac{(\lambda t)^{(m-1)}}{(m-1)!}e^{-\lambda t},$$

после очевидных сокращений. Следовательно, производная $F'_t(t) = \lambda F'_u(t)$ есть плотность гамма-распределения с параметром формы m и параметром масштаба $1/\lambda$. ⇚

✦ Предположим, что заявки на станцию обслуживания поступают как простой поток событий интенсивности λ. Бригады рабочих на станции меняются через определённые часы. Естественно ожидать, что бригады находятся в равных условиях. Пусть $\left\langle \xi(t),\ t \in \mathbb{R}^1_+ \right\rangle$ — пуассоновский процесс, t_0 — время окончания первой смены и начала второй. Тогда $\eta(s) = \xi(s + t_0) - \xi(t_0)$ — количество заявок, поступивших во вторую смену за время s.

424| Упр. Покажите, что $\left\langle \eta(s),\ s \geqslant 0 \right\rangle$ — пуассоновский процесс интенсивности λ.

Предположим теперь, что бригады меняются по достижении определённого уровня m поступивших заявок. Пусть, как и выше, $\mathscr{T}$ — время ожидания m-й заявки (случайное время смены бригад). Тогда количество заявок $\eta(s) = (\xi(s + \mathscr{T}) - m)$, поступивших во вторую смену за время s, также образует пуассоновский процесс интенсивности λ. Действительно, пусть $\left\langle \tau_k,\ k = 1, 2, \ldots \right\rangle$ — последовательность независимых одинаково распределённых сл.в., описывающих время между соответствующими событиями пуассоновского потока — $\mathscr{T} = \tau_1 + \ldots + \tau_m$. Тогда равенство $\xi(t) = k\ (> 0)$ эквивалентно неравенствам $\tau_1 + \ldots + \tau_k \leqslant t < \tau_1 + \ldots + \tau_{k+1}$ $(t < \tau_1$ для $k = 0)$. Аналогично, равенство $\eta(s) = k\ (> 0)$ эквивалентно неравенствам $\tau_1 + \ldots + \tau_{m+k} \leqslant s + \mathscr{T} < \tau_1 + \ldots + \tau_{m+k+1}$ или $\tau_{m+1} + \ldots + \tau_{m+k} \leqslant s < \tau_{m+1} + \ldots + \tau_{m+k+1}$. Ввиду одинаковой распределённости сл.величин τ распределение $\boldsymbol{\xi}$ эквивалентно распределению $\boldsymbol{\eta}$.

✦ Если количество событий n простого потока в интервале $[0; T]$ зафиксировано, то моменты $\mathscr{T}_1 \leqslant \mathscr{T}_2 \leqslant \ldots \leqslant \mathscr{T}_n \leqslant T$, в которые появились события из потока, можно интерпретировать как реализации независимых сл.величин с равномерным распределением $\mathfrak{Un}(0, T)$. Точнее, рассмотрим вариационный ряд $\upsilon_{(1)} \leqslant \ldots \leqslant \upsilon_{(n)}$, полученный из n независимых равномерных $\mathfrak{Un}(0, T)$ сл.в. $\upsilon_1, \ldots, \upsilon_n$ с помощью упорядочения. Совместная ф.р. $(\upsilon_{(1)}, \ldots, \upsilon_{(n)})$ находится по формуле (12), стр. 129:

$$\begin{gathered}\mathbf{P}\left\{\upsilon_{(i)} \leqslant u_i,\ i = 1, \ldots, n\right\} = \frac{n!}{T^n} \int\limits_0^{u_1} ds_1 \int\limits_{s_1}^{u_2} ds_2 \cdots \int\limits_{s_{n-1}}^{u_n} ds_n, \\ 0 \leqslant u_1 \leqslant \ldots \leqslant u_n \leqslant T.\end{gathered} \tag{36}$$

Поскольку неравенство $\upsilon_{(k)} \leqslant u$ влечёт справедливость неравенств $\upsilon_{(j)} \leqslant u$ для любых $j < k$, то для произвольных $0 \leqslant u_i \leqslant T$, $i = \overline{1, n}$, ф.р. вычисляется по той же формуле с заменой $u_i \dashrightarrow u'_i = \min\{u_n, \ldots, u_i\}$, $i = \overline{1, n}$. Плотность этого распределения равна $n!\,/\,T^n$ в указанной области; в остальном эта плотность равна нулю. Имеет место

425| **Теорема.** *Пусть $\mathscr{T}_k$ — время осуществления k-го события в пуассоновском потоке, тогда условное распределение $(\mathscr{T}_1, \ldots, \mathscr{T}_n)$ при условии, что к моменту T произошло ровно n событий, описывается правой частью* (36), *т.е. совпадает с распределением вариацинного ряда из равномерного $\mathcal{U}\mathtt{n}(0,T)$ распределения.*

$\Rrightarrow$ Пусть τ_k — время между $(k-1)$-м и k-м событиями потока, т.е. $\mathscr{T}_k = \tau_1 + \ldots + \tau_k$. Тогда в силу теоремы 421 для $\forall\, 0 \leqslant u_1 < \ldots < u_n \leqslant T$

$$P(u_1, \ldots, u_n) := \mathbf{P}\{\mathscr{T}_1 \leqslant u_1, \ldots, \mathscr{T}_n \leqslant u_n,\ \xi(T) = n\} =$$

$$= \mathbf{P}\left\{[\tau_1 \leqslant u_1, \tau_1 + \tau_2 \leqslant u_2, \ldots, \tau_1 + \ldots + \tau_n \leqslant u_n] \cap [\tau_1 + \ldots + \tau_{n+1} > T]\right\} =$$

$$= \int\limits_0^{u_1} \lambda e^{-\lambda t_1} dt_1 \int\limits_0^{u_2 - t_1} \lambda e^{-\lambda t_2}\, dt_2 \cdots \int\limits_0^{u_n - t_1 - \ldots - t_{n-1}} \lambda e^{-\lambda t_n}\, dt_n \int\limits_{T - t_1 - \ldots - t_n}^{\infty} \lambda e^{-\lambda t_{n+1}}\, dt_{n+1}.$$

Интеграл по dt_{n+1}, очевидно, равен $\exp(-\lambda T + \lambda t_1 + \ldots + \lambda t_n)$. Следовательно,

$$P(u_1, \ldots, u_n) = \lambda^n e^{-\lambda T} \int\limits_0^{u_1} dt_1 \int\limits_0^{u_2 - t_1} dt_2 \cdots \int\limits_0^{u_n - t_1 - \ldots - t_{n-1}} dt_n.$$

Многомерный интеграл совпадёт с интегралом в (36), если перейти к новым переменным с помощью взаимно однозначной замены $s_k \dashrightarrow t_1 + \ldots + t_k$, $k = \overline{1,n}$. После деления совместной вероятности $P(u_1, \ldots, u_n)$ на вероятность условия $\mathbf{P}\{\xi(T) = n\} = \lambda^n T^n e^{-\lambda T}(n!)^{-1}$ получаем (36). $\Lleftarrow$

Оказывается, справедливо и обратное. Следующее утверждение позволяет моделировать траектории процесса Пуассона.

426| **Теорема.** *Пусть $\langle \nu, \zeta_1, \zeta_2, \ldots \rangle$ — независимые в совокупности сл.величины, определённые на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$, величина ν имеет распределение Пуассона $\mathcal{P}\mathtt{ois}(\lambda T)$, а сл.в. $\zeta_k \backsim \mathcal{U}\mathtt{n}(0,T)$, $k \geqslant 1$. Обозначим $\mathscr{T}_{1,n} \leqslant \ldots \leqslant \mathscr{T}_{n,n}$ вариационный ряд $\zeta_1, \ldots, \zeta_n$, $n \geqslant 1$; $\mathscr{T}_{0,n} \equiv 0$. Определим $\xi(t) = \max\{k :\ \mathscr{T}_{k,\nu} \leqslant t,\ k = \overline{0,\nu}\}$. Тогда процесс $\langle\, \xi(t),\ t \in [0;T] \,\rangle$ есть процесс Пуассона с интенсивностью λ.*

$\Rrightarrow$ Упрощая запись, положим $T = 1$. Пусть $0 \leqslant s < t \leqslant 1$ и зафиксировано $\nu = n\ (> 0)$, тогда для $k, m \geqslant 0$ событие $\{\xi(s) = k,\ \xi(t) - \xi(s) = m\}$ означает, что из n равномерных $\mathcal{U}\mathtt{n}(0,1)$ сл.величин в интервал $[0; s]$ попало k чисел, а в интервал $(s; t]$ — m чисел. Другими словами, условная вероятность

$$\mathbf{P}\left\{\xi(s)=k,\ \xi(t)-\xi(s)=m \,\middle|\, \nu=n\right\} \;=\; \frac{n!\,s^k(t-s)^m(1-t)^{n-k-m}}{k!m!(n-k-m)!}$$

для $n \geqslant m+k$; для $n < m+k$ эта вероятность равна нулю.

Усредняя по значениям $\nu = n$, получаем вероятность

$$\mathbf{P}\{\xi(s)=k,\ \xi(t)-\xi(s)=m\} \;=$$
$$= \sum_{n=m+k}^{\infty}\left[\frac{n!}{k!m!(n-k-m)!}\ s^k(t-s)^m(1-t)^{n-k-m}\cdot\frac{\lambda^n}{n!}e^{-\lambda}\right] \;=$$
$$= \frac{s^k(t-s)^m\lambda^k\lambda^m}{k!m!}\,e^{-\lambda}\ \sum_{j=0}^{\infty}\frac{\lambda^j}{j!}(1-t)^j = \frac{s^k(t-s)^m\lambda^k\lambda^m}{k!m!}\,e^{-\lambda}\cdot e^{\lambda(1-t)} \;=$$
$$= \frac{(\lambda s)^k}{k!}e^{-\lambda s}\cdot\frac{(\lambda(t-s))^m}{m!}e^{-\lambda(t-s)}$$

в соответствии с известной формулой Тейлора для $e^{\lambda(1-t)}$. Случай, когда вместо двух моментов времени $s < t$ берутся произвольные наборы таких моментов, почти ничем не отличается от рассмотренного. Отсюда можно сделать вывод, что построенный процесс имеет независимые приращения, причём приращение $\xi(t)-\xi(s) \backsim \mathcal{P}\mathtt{ois}(\lambda(t-s))$. $\Lleftarrow$

427| Упр. Покажите, что при фиксированном значении $\xi(T)=n$ условное распределение числа событий потока, появившихся до момента t ($< T$), совпадает с биномиальным распределением $\mathcal{B}\mathtt{in}(n,p)$ с $p = t/T$.

§ 4. Марковские процессы

Для иллюстрации основного свойства, определяющего марковский процесс, рассмотрим процедуру поиска неподвижной точки $f(x) = x$ при сжимающем отображении f. Итерационный процесс состоит в выборе начальной точки x_0 и в последующих приближениях, вычисляемых по формуле $x_{n+1} = f(x_n)$, $n \geqslant 0$. Понятно, что, скажем, x_8 зависит от всех предыдущих значений x_k, $k < 8$, например $x_8 = f(f(f(x_5)))$. Однако ясно, что, если одно и то же значение на шестом шаге x_6 можно было бы получить при нескольких различных значениях x_5, результат на восьмом шаге x_8 всё равно определялся бы полностью x_6 и не зависел от выбора x_5.

Для марковских процессов подобное свойство независимости будущих значений от прошлых при фиксированном настоящем имеет место в стохастическом понимании независимости. Мы начнём с так называемых цепей

Маркова $\boldsymbol{\xi} = \left\langle \xi(t),\ t \in \mathbb{T} \subset \mathbb{R}^1_+ \right\rangle$, определённых на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$, когда фазовое пространство процесса, т.е. пространство $\mathscr{X}$ тех значений, которые могут принимать сл.величины $\xi(t)$, не более чем счётно. Элементы $\mathscr{X}$ называются состояниями процесса. Основной интерес при рассмотрении марковских цепей концентрируется вокруг различных соотношений между вероятностями перехода из одного состояния в другое. В связи с этим природа элементов $\mathscr{X}$ часто не представляет интереса, и можно считать, что $\mathscr{X} = \{e_0, e_1, \ldots, e_N\}$, $N < \infty$, или $\mathscr{X} = \{e_0, e_1, \ldots\}, N = \infty$.

Определение. Процесс $\boldsymbol{\xi}$ называется *марковской цепью*, если для любого «текущего» момента времени $r \in \mathbb{T}$, для любых наборов прошедших $s_1 < \ldots < s_k < r$ и будущих моментов времени $r < t_1 < \ldots < t_m$ и любых значений $\{x_1, \ldots, x_k, y, z_1, \ldots, z_m\} \subset \mathscr{X}$, $k, m \geqslant 1$, условная вероятность

$$\begin{aligned} &\mathbf{P}\left\{ \bigcap\nolimits_{j=1}^{m} \big(\xi(t_j) = z_j\big) \ \Big|\ \bigcap\nolimits_{i=1}^{k} \big(\xi(s_i) = x_i\big) \cap \big(\xi(r) = y\big) \right\} = \\ &= \mathbf{P}\left\{ \bigcap\nolimits_{j=1}^{m} \big(\xi(t_j) = z_j\big) \ \Big|\ \xi(r) = y \right\}. \end{aligned} \tag{37}$$

⚠ Уравнение (37) может быть проверено только тогда, когда вероятности соответствующих условий больше нуля. Обозначим событие, связанное с прошлым, через Π, с настоящим — через R, а с будущим — через F, тогда по определению условной вероятности это уравнение можно переписать в виде

$$\mathbf{P}\{\Pi\mathrm{RF}\}\,\mathbf{P}\{\mathrm{R}\} = \mathbf{P}\{\Pi\mathrm{R}\}\,\mathbf{P}\{\mathrm{RF}\}.$$

Справедливость этого равенства при $\mathbf{P}\{\Pi\mathrm{R}\} = 0$ или $\mathbf{P}\{\mathrm{R}\} = 0$ легко следует из свойств вероятности. Другой аспект применения условных вероятностей связан с формулой умножения $\mathbf{P}\{\Pi\mathrm{R}\} = \mathbf{P}\big\{\Pi \,|\, \mathrm{R}\big\}\,\mathbf{P}\{\mathrm{R}\}$, которая будет справедлива и при $\mathbf{P}\{\mathrm{R}\} = 0$, если договориться считать в этом случае условную вероятность произвольной. В дальнейшем все нюансы применения условных вероятностей будут трактоваться в духе настоящего замечания.

Соотношение (37) называется марковским свойством. Обозначив через $B(t_j) = \{\xi(t_j) = x_j\}$, $j = \overline{1, n}$, события, связанные с реализациями процесса в моменты времени $t_1 < \ldots < t_n$, марковское свойство можно переписать в виде

$$\mathbf{P}\Big\{ \bigcap\nolimits_{m+1}^{n} B(t_j) \ \Big|\ \bigcap\nolimits_{1}^{m} B(t_k) \Big\} = \mathbf{P}\Big\{ \bigcap\nolimits_{m+1}^{n} B(t_j) \ \Big|\ B(t_m) \Big\},$$

причём это равенство справедливо для $\forall\, m < n$. Заметим, что проверку марковского свойства можно ограничить только одним событием из «будущего».

428| **Лемма.** (?!) *Процесс* $\boldsymbol{\xi}$ *образует марковскую цепь т. т. т. когда для* $\forall\, k \geqslant 1$, $\forall\, s_1 < \ldots < s_k < r < t$ *(из* $\mathbb{T}$ *) и любых состояний* $\{x_1, \ldots, x_k, y, z\} \subset \subset \mathscr{X}$ *условная вероятность*

$$\mathbf{P}\Big\{\, \xi(t) = z \;\mid\; \bigcap\nolimits_{j=1}^{k} \big(\xi(s_j) = x_j\big) \cap \big(\xi(r) = y\big) \,\Big\} \;=\; \mathbf{P}\big\{\, \xi(t) = z \mid \xi(r) = y \,\big\}.$$

Марковское свойство, выраженное условием (37) или более простым условием из этой леммы, говорит о «функциональной» независимости вероятностей будущих событий от прошлых при фиксированном настоящем и прошлом. Это свойство можно заменить свойством так называемой условной независимости (стохастической) будущих событий от прошлых относительно настоящего.

429| **Лемма.** (?!) *Процесс* $\boldsymbol{\xi}$ *образует марковскую цепь т. т. т. когда для* $\forall\, k \geqslant 1$, $\forall\, s_1 < \ldots < s_k < r < t$ *(из* $\mathbb{T}$ *) и любых состояний* $\{x_1, \ldots, x_k, y, z\} \subset \subset \mathscr{X}$ *совместная условная вероятность при фиксированном настоящем*

$$\begin{aligned} &\mathbf{P}\Big\{ \big(\xi(t) = z\big) \textstyle\bigcap_1^k \big(\xi(s_j) = x_j\big) \;\Big|\; \xi(r) = y \Big\} = \\ =\;\; &\mathbf{P}\big\{ \xi(t) = z \;\big|\; \xi(r) = y \big\} \cdot \mathbf{P}\Big\{ \textstyle\bigcap_1^k \big(\xi(s_j) = x_j\big) \;\Big|\; \xi(r) = y \Big\}. \end{aligned}$$

430| Пример. Пусть ξ_0 — начальная точка (случайная), $\zeta_0, \zeta_1, \ldots$ — последовательность сл.величин, независимых между собой и от ξ_0. Тогда для любых измеримых функций $f_k(x,z)$ $(\in \mathscr{X})$ последовательность $\xi_{k+1} = = f_k(\xi_k, \zeta_k)$, $k \geqslant 0$, образует марковский процесс с дискретным временем: $\mathbb{T} = \{0, 1, \ldots\}$. ⊚

Для любых событий $B_1, \ldots, B_n$, связанных с реализациями процесса в моменты времени $t_1 < \ldots < t_n$, по формуле условной вероятности и в соответствии с марковским свойством совместная вероятность

$$\begin{aligned} &\mathbf{P}\Big\{ \bigcap_{j=1}^{n} B_j \Big\} \;=\; \mathbf{P}\Big\{ B_n \;\Big|\; \bigcap_{j=1}^{n-1} B_j \Big\} \mathbf{P}\Big\{ \bigcap_{j=1}^{n-1} B_j \Big\} \;= \\ =\;\; &-\mathbf{P}\left\{ B_n \,|\, B_{n-1} \right\} \mathbf{P}\Big\{ \bigcap_{j=1}^{n-1} B_j \Big\} \;=\ldots=\; \mathbf{P}\left\{ B_1 \right\} \prod_{k=0}^{n-2} \mathbf{P}\left\{ B_{n-k} \,|\, B_{n-k-1} \right\}. \end{aligned}$$

Таким образом, если известны *вероятности перехода* из состояния e_k в момент времени s в состояние e_j в момент времени $t > s$, т.е. вероятности

$$p_{k,j}(s,t) := \mathbf{P}\left\{\xi(t) = e_j \mid \xi(s) = e_k\right\}, \tag{38}$$

и вероятности состояний в начальный момент времени:

$$\pi_k := \mathbf{P}\left\{\xi(0) = e_k\right\}, \quad e_k, e_j \in \mathscr{X},$$

то вероятность любого события в фиксированный момент времени и любая конечномерная вероятность могут быть найдены по формулам:

$$\begin{aligned} p_{k_1}(t_1) := \mathbf{P}\left\{\xi(t_1) = e_{k_1}\right\} &= \sum_{k_0=0}^{N} \pi_{k_0} p_{k_0,k_1}(0,t_1), \\ \mathbf{P}\left\{\xi(t_1) = e_{k_1}, \ldots, \xi(t_n) = e_{k_n}\right\} &= p_{k_1}(t_1) \prod_{j=1}^{n-1} p_{k_j,k_{j+1}}(t_j, t_{j+1}), \end{aligned} \tag{39}$$

где $0 \leqslant t_1 < \ldots < t_n$; $k_1, \ldots, k_n \geqslant 0$; $n \geqslant 1$. Напомним, N — «размер» фазового пространства ($N < \infty$ или $N = \infty$). Ясно, что переходные вероятности и начальное распределение $\boldsymbol{\pi} = \left\langle \pi_k, \ k = \overline{1,N} \right\rangle$ удовлетворяют следующим условиям:

(MC$_1$) при каждом k и любых $s < t$ семейство $p_{k,j}(s,t)$ (по j) есть дискретное распределение на $\mathscr{X}$, т.е. для $\forall 0 \leqslant k, j \leqslant N, 0 \leqslant s < t$:

$$(\checkmark) \ \ p_{k,j}(s,t) \geqslant 0, \qquad (\checkmark) \ \ \sum_{j=0}^{N} p_{k,j}(s,t) = 1;$$

(MC$_2$) $p_{k,j}(t,t) = 1$ для $k = j$ и $p_{k,j}(t,t) = 0$ для $k \neq j$;

(MC$_3$) $\boldsymbol{\pi}$ задаёт распределение на $\mathscr{X}$: $\sum_0^N \pi_k = 1$.

Теперь необходимо решить обратный вопрос: при каких условиях на семейство функций $p_{k,j}(s,t)$, $k,j \geqslant 0$; $s < t\,(\in \mathbb{T})$, описывающих переходные вероятности, и семейство $\pi_k, k \geqslant 0$, соотношения (39) будут задавать конечномерное распределение некоторой марковской цепи? По теореме Колмогорова достаточно, чтобы любое маргинальное распределение (39) имело такой же вид. Найдём, для примера, маргинальное распределение $(\xi(t_1), \ldots, \xi(t_{n-2}), \xi(t_n))$:

$$\begin{aligned} &\mathbf{P}\{\xi(t_1) = e_{k_1}, \ldots, \xi(t_{n-2}) = e_{k_{n-2}}, \xi(t_n) = e_{k_n}\} = \\ &= \sum_{j=0}^{N} \mathbf{P}\{\xi(t_1) = e_{k_1}, \ldots, \xi(t_{n-2}) = e_{k_{n-2}}, \xi(t_{n-1}) = e_j, \xi(t_n) = e_{k_n}\} = \end{aligned}$$

$$= p_{k_1}(t_1)\prod_{j=1}^{n-3} p_{k_j,k_{j+1}}(t_j,t_{j+1})\cdot \sum\nolimits_{j=0}^{N} p_{k_{n-2},j}(t_{n-2},t_{n-1})p_{j,k_n}(t_{n-1},t_n).$$

С другой стороны, также из (39) следует, что

$$\mathbf{P}\{\xi(t_1)=e_{k_1},\ldots,\xi(t_{n-2})=e_{k_{n-2}},\xi(t_n)=e_{k_n}\} =$$
$$= p_{k_1}(t_1)\prod_{j=1}^{n-3} p_{k_j,k_{j+1}}(t_j,t_{j+1})\cdot p_{k_{n-2},k_n}(t_{n-2},t_n).$$

Таким образом, можно сформулировать следующее утверждение.

431| **Теорема.** *Набор чисел* $\pi_k \geqslant 0$, $0\leqslant k\leqslant N$, *и семейство функций* $p_{k,j}(s,t)$, $0\leqslant k,j\leqslant N$, $0\leqslant s<t$, *задают марковскую цепь с конечномерными распределениями, определяемыми* (39), *если они удовлетворяют условиям* (MC$_1$) – (MC$_3$) *и, кроме того, для любых* $0\leqslant k,m\leqslant N$, $0\leqslant s<t<u$ *справедливо уравнение Маркова:*

$$\sum_{j=0}^{N} p_{k,j}(s,t)p_{j,m}(t,u) = p_{k,m}(s,u). \tag{40}$$

432| Упр. Покажите, что процесс с конечномерными распределениями (39) удовлетворяет марковскому свойству.

433| Упр. Покажите, что семейство переходных вероятностей (38), определяемых какой-либо цепью Маркова, удовлетворяет уравнению (40).

⚠ Аналог уравнения Маркова в непрерывном случае ($\mathscr{X}=\mathbb{R}^1$) называется уравнением Колмогорова–Чепмена. Для процессов с дискретным временем, т.е. с $\mathbb{T}=\{0,1,\ldots\}$, уравнение (40) при $t=u-1$ называют прямым уравнением Колмогорова–Чепмена, а при $t=s+1$ — обратным.

Уравнение (40) можно переписать в более компактном виде, если воспользоваться операцией матричного умножения. Для каждой пары $s<t$ определим матрицу переходных вероятностей $P(s,t)$ с элементами $p_{k,j}(s,t)$. Если фазовое пространство цепи конечно, то размерность этой матрицы $N\times N$, в противном случае она бесконечномерна. Каждая k-я строка матрицы $P(s,t)$ содержит вероятности перехода из состояния e_k в состояния, номера которых совпадают с номерами столбцов матрицы $P(s,t)$. Сумма всех элементов строки равна 1. Матрица с таким свойством строк называется *стохастической*. Правая часть (40) есть не что иное, как скалярное произведение k-й стро-

ки $P(s,t)$ на m-й столбец $P(t,u)$. Другими словами, уравнение Маркова эквивалентно матричному уравнению

$$P(s,t)P(t,u) = P(s,u), \qquad \forall\, s \leqslant t \leqslant u. \tag{41}$$

Прямое и обратное уравнения Колмогорова–Чепмена имеют вид

$$\begin{aligned} P(s,u) &= P(s,u-1)P(u-1,u), \\ P(s,u) &= P(s,s+1)P(s+1,u). \end{aligned}$$

Обозначим через $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots)$ вектор-строку вероятностей начальных состояний. Тогда вектор-строка вероятностей состояний в каждый фиксированный момент времени (39) находится с помощью аналогичных операций:

$$\boldsymbol{p}(t) = \boldsymbol{\pi}\, P(0,t). \tag{42}$$

434| Пример. Как отмечалось в предыдущем параграфе, процесс Пуассона представляет собой марковскую цепь с начальным распределением, сосредоточенным в нуле: $\pi_0 = 1$. Семейство переходных вероятностей процесса Пуассона, имеющего независимые приращения, совпадает с распределением соответствующих приращений: $\forall\, 0 \leqslant s < t,\ 0 \leqslant k \leqslant j$

$$\begin{aligned} p_{k,j}(s,t) &= \frac{\mathbf{P}\{\xi(s) = k, \xi(t) - \xi(s) = j-k\}}{\mathbf{P}\{\xi(s) = k\}} = \\ &= \mathbf{P}\{\xi(t) - \xi(s) = j-k\} = \frac{(\lambda(t-s))^{j-k}}{(j-k)!} e^{-\lambda(t-s)}. \end{aligned}$$

Так как для $j < k$ и $s < t$ вероятности $p_{k,j}(s,t) \equiv 0$, то уравнение Маркова означает, что должно иметь место равенство

$$\begin{aligned} p_{k,m}(s,u) &= \sum_{j=k}^{m} \frac{(\lambda(t-s))^{j-k}}{(j-k)!} e^{-\lambda(t-s)} \frac{(\lambda(u-t))^{m-j}}{(m-j)!} e^{-\lambda(u-t)} = \\ &= \frac{\lambda^{m-k}}{(m-k)!} e^{-\lambda(u-s)} \sum_{j=k}^{m} \frac{(m-k)!}{(j-k)!(m-j)!} (t-s)^{j-k}(u-t)^{m-j} = \\ &= \frac{\lambda^{m-k}}{(m-k)!} e^{-\lambda(u-s)} \sum_{j=0}^{m-k} \frac{(m-k)!}{j!(m-k-j)!} (t-s)^{j}(u-t)^{m-k-j} = \\ &= \frac{(\lambda(u-s))^{m-k}}{(m-k)!} e^{-\lambda(u-s)}, \end{aligned}$$

где последнее равенство следует из формулы для бинома Ньютона. Таким образом, пуассоновский процесс можно определить как однородную марковскую цепь с пуассоновскими переходными вероятностями. ◎

Напомним, что процесс Пуассона обладает свойством однородности. Для общих марковских процессов это понятие также играет существенную роль.

Определение. Марковская цепь называется *однородной*, если её переходная вероятность для любых моментов $s < t$ зависит только от разности $t - s$, т.е. для $\forall\, 0 \leqslant k, j \leqslant N,\ 0 \leqslant s < t$

$$p_{k,j}(s,t) = p_{k,j}(t-s).$$

✧ Цепи Маркова. Дискретное время. Случай, когда параметр t пробегает счётное подмножество, т.е. $t \in \mathbb{T} = \{0, 1, \ldots, N\}$, где $N \leqslant \infty$, наиболее востребован на практике. Даже при анализе процессов с непрерывным временем, например мониторинг количества голов в футбольном матче, часто бывает удобнее фиксировать ситуацию через какие-то промежутки времени. Такой процесс представляет собой просто последовательность сл.в. Марковскую цепь, образованную последовательностью сл.в., называют *дискретной*. В этом случае семейство переходных вероятностей определяется вероятностями перехода за один шаг. Действительно, в силу (41) (обратного уравнения Колмогорова–Чепмена)

$$P(s,u) = P(s,s+1)P(s+1,u) = \ldots = \prod_{t=s}^{u-1} P(t,t+1).$$

Если дискретная цепь Маркова однородна, то вероятности $p_{k,j}(t,t+1)$ не зависят от $t = 0, 1, \ldots$ Обозначим через $P = P(t,t+1)$ матрицу переходных вероятностей за один шаг однородной дискретной цепи Маркова. Тогда матрица переходных вероятностей для произвольных моментов времени в такой цепи равна соответствующей степени матрицы P:

$$P(s,t) = P^{t-s}.$$

Из (42) следует, что распределение сечения в момент t совпадает с

$$\boldsymbol{p}(t) = \boldsymbol{\pi}\, P^{\,t}. \tag{43}$$

Для описания однородной дискретной цепи Маркова с малым количеством состояний удобно использовать представление матрицы переходов за один шаг в виде стохастического графа. В этом графе каждое состояние обозначается каким-либо символом (кружком), а переходы с положительной вероятностью отмечаются направленными стрелками, над (под) которыми приписываются

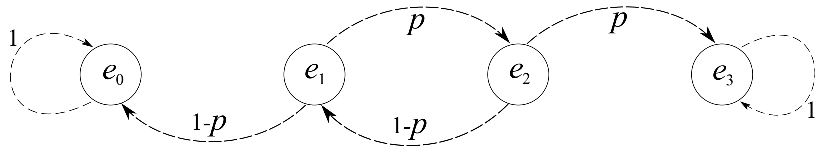


Рис. 8. Стохастический граф наличного капитала игрока

вероятности перехода.

435| Пример. Пусть ξ_n — капитал игрока А после n проведённых игр, в каждой из которых он побеждает (и пополняет свой капитал на 1 у.е.) с вероятностью p; с вероятностью $1-p$ его капитал уменьшается на 1 у.е. На рис. 8 представлен стохастический граф этой дискретной однородной (?!) цепи Маркова (?!) при суммарном капитале игроков $N = a + b = 3$.

Отметим, что для этой цепи состояния $e_0 = 0$ и $e_3 = 3$ — это так называемые *поглощающие состояния*. Они означают разорение или полный успех А. Формально говоря, в этом случае процесс должен закончиться, однако, чтобы соблюсти соответствие между игрой и нашим сл.процессом, мы можем просто считать, что игра как бы продолжается, но с изменёнными (нулевыми) вероятностями победы того или иного игрока.

Чаще всего в такой игре начальные капиталы игроков зафиксированы. В этом случае начальное распределение $\boldsymbol{\pi}$ будет сосредоточено на том состоянии, которое соответствует капиталу А. Например, если игрок А, заходя в игру, имел $a = 1$ у.е. и мечтал выйти из неё с $N = 3$ у.е., то $\boldsymbol{\pi} = (0, 1, 0, 0)$. При «справедливой» игре, когда $p = {}^1\!/_2$, матрица переходных вероятностей за один шаг

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Строка $\boldsymbol{p}_n \stackrel{\text{df}}{=} \boldsymbol{\pi} P^n$ даёт распределение капитала игрока А после n игр.

В табл. 1 приведены примеры такого распределения; чи́сла, меньшие 10^{-6}, полагались равными ($\approx$) нулю.

Как видно, практически уже после 10 игр процесс «остановится» и вероятности, с которыми капитал игрока А достигнет того или иного поглощающего состояния, почти не отличаются от предельных (при $n \to \infty$) значений ${}^1\!/_3, {}^2\!/_3$, которые совпадают (см. дополнение к этой главе) с вероятностями $\frac{b}{a+b}$ — разорения A и $\frac{a}{a+b}$ — разорения B.

Табл. 1. Распределение капитала игрока A

$\boldsymbol{\pi}$, a	$\boldsymbol{p}_n, n=4$	$\boldsymbol{p}_n, n=11$	$\boldsymbol{p}_n, n=20$
$(0,1,0,0)$ $a=1$	(0.625, 0.0625, 0, 0.3125)	(0.6665, 0, 0.0005, 0.3330)	$(\frac{2}{3},\ \approx 0,\ 0,\ \frac{1}{3})$
$(0,0,1,0)$ $a=2$	(0.3125, 0, 0.0625, 0.625)	(0.3330, 0.0005, 0, 0.6665)	$(\frac{1}{3},\ 0,\ \approx 0,\ \frac{2}{3})$

⊚

✧ Эргодическая теорема для цепи Маркова. В примере 435 два состояния были таковы, что при попадании в них сл.точка больше уже никуда не переходит. Такие состояния, т.е. состояния e_k, для которых вероятность перехода $p_{k,k} = 1$, называются *поглощающими*. Чтобы выделить ещё некоторые классы состояний, рассмотрим вероятность

$$p_{k,j}(t) \ = \ \mathbf{P}\left\{\xi(t) = e_j \,|\, \xi(0) = e_k\right\}$$

перехода из состояния e_k в состояние e_j за t шагов, $k \geqslant 0$, $j \geqslant 0$, $t \geqslant 1$. Очевидно, строка $(p_{k,0}(t), p_{k,1}(t), \ldots)$ совпадает с k-й строкой матрицы P^t.

Несущественным называется состояние e_k, если из него можно выйти за конечное число шагов, т.е. для некоторого $t \geqslant 1$ найдётся такое $e_j \in \mathscr{X}$, что $p_{k,j}(t) > 0$, но в него нельзя попасть: $p_{j,k}(t) = 0$ для $\forall t \geqslant 1$, $e_j \in \mathscr{X}$. В противном случае состояние называется *существенным*.

Состояние e_k называется *нулевым*, если $\lim\limits_{t\to\infty} p_{k,k}(t) = 0$.

Сообщающимися называются состояния e_k, e_j, в каждое из которых можно попасть из другого за конечное число шагов: $p_{k,j}(t) > 0$, $p_{j,k}(s) > 0$ для некоторых $t, s \geqslant 1$.

Возвратным называется состояние e_k, для которого вероятность возвращения за конечное число шагов равна единице: $\sum_{t=1}^{\infty} \rho_k(t) = 1$, где

$$\rho_k(t) = \mathbf{P}\left\{\xi(1) \neq e_k, \ldots, \xi(t-1) \neq e_k, \xi(t) = e_k \,|\, \xi(0) = e_k\right\}$$

— вероятность возвращения в состояние e_k в первый раз через t шагов. Имеет место следующее утверждение: *состояние e_k является возвратным т. т. т. когда* $\sum_t p_{j,j}(t) = \infty$.

Если $p_k(2) > 0, p_k(4) > 0, \ldots$, а все вероятности возврата через нечётное число шагов равны нулю, то такое состояние естественно назвать *периодическим* с периодом $M = 2$. В примере 435 все состояния, отличные от

$e_0 = 0, e_N = N$, имеют период $M = 2$. Аналогично вводятся понятия состояний, периодических с периодом $M > 2$. Состояния, не имеющие периода или имеющие период 1 $(\rho_k(t) > 0,\ \forall t \geqslant 1)$, называются *апериодическими*.

436 **Теорема.** [*Эргодическая теорема.*] *Пусть все состояния дискретной однородной цепи Маркова с конечным фазовым пространством и матрицей переходных вероятностей P апериодические, существенные и сообщающиеся. Тогда справедливы следующие утверждения:*

a) *для $\forall\, 0 \leqslant k, j \leqslant N$ существует предел*

$$\lim_{t\to\infty} p_{k,j}(t) = q_j > 0;$$

b) *строка $\boldsymbol{q} = (q_0, q_1, \ldots, q_N)$ определяет вероятностное распределение:*

$$\sum\nolimits_0^N q_j = 1$$

и даёт не содержащее нулей решение уравнения $\boldsymbol{q} = \boldsymbol{q}\, P$, т.е.

$$q_j = \sum_{k=0}^{N} p_{k,j} q_j$$

для $\forall\, j = 0, 1, \ldots, N$. Матрица Q, все строки которой одинаковые и совпадают с $\boldsymbol{q}$, удовлетворяет соотношениям[‡]

$$QP = Q, \qquad \lim_{t\to\infty} \boldsymbol{\pi}\, P^t = \boldsymbol{q} \tag{44}$$

для любого распределения начальных состояний $\boldsymbol{\pi}$;

c) *справедливо обратное: условия* a), b) *гарантируют, что все состояния конечной однородной дискретной цепи Маркова апериодические, существенные и сообщающиеся.*

⚠ Пусть $\cdot^\flat$ — операция транспонирования. Вектор-столбец $\boldsymbol{q}^\flat$ есть собственный вектор с единичным собственным числом матрицы $P^\flat$: $P^\flat \boldsymbol{q}^\flat = \boldsymbol{q}^\flat$, причём все его компоненты ненулевые.

Определение. Дискретная однородная цепь Маркова называется *эргодической*, если для неё выполняются условия a), b) теоремы 436.

⚠ «Качественные» условия эргодической теоремы можно заменить аналитическим условием: утверждения a), b) справедливы т. т. т. когда найдётся такое t_0, что $\min_{k,j} p_{k,j}(t_0) > 0$.

Утверждения b) теоремы, в частности равенство $QP = Q$, есть следствие a). Действительно, требуемое вытекает из прямого уравнения Колмого-

[‡] Здесь сходимость матриц покомпонентная.

рова–Чепмена $P^{t+1} = P^t P$, если в нём перейти к пределу $t \to \infty : P^t \xrightarrow[t\to\infty]{} Q$. Справедливость второго соотношения (44) вытекает, конечно, из того, что все строки Q одинаковые — каждый столбец содержит только одно число.

437 Пример. Рассмотрим матрицу переходных вероятностей

$$P = \begin{pmatrix} 0.2 & 0.4 & 0.4 \\ 0 & 0.6 & 0.4 \\ 0.4 & 0.2 & 0.4 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 1 & 2 & 2 \\ 0 & 3 & 2 \\ 2 & 1 & 2 \end{pmatrix}.$$

Нормированная строка решения уравнения $\boldsymbol{q}P = \boldsymbol{q}$ равна $\boldsymbol{q} = \frac{1}{5}(1, 2, 2)$. По построению $QP = Q$. Легко проверить, что для матрицы

$$E = P - Q = \frac{1}{5}\begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}$$

справедливо равенство $EP = \frac{1}{5}E$ (равенство подобного типа не обязательно будет иметь место для других матриц P). Следовательно, имеем:

$$P = Q + E, \qquad P^2 = (Q+E)P = QP + EP = Q + \frac{1}{5}E,$$

$$P^3 = (Q + \frac{1}{5}E)P = Q + \frac{1}{5^2}E, \quad \ldots, \quad P^s = Q + \frac{1}{5^{s-1}}E \xrightarrow[s\to\infty]{} Q. \qquad \circledcirc$$

Заметим, что однородная марковская цепь с начальным распределением $\boldsymbol{q}$ и матрицей переходных вероятностей Q есть стационарный процесс. Перефразируя, можно сказать, что, каково бы ни было начальное распределение состояний, через некоторый промежуток времени эргодическая цепь Маркова достигнет некой стационарности, когда переходы из состояния в состояние не зависят ни от достигнутого состояния, ни от количества рассматриваемых шагов, — состояние процесса в любой момент в будущем определяется реализацией случайного эксперимента с фиксированным дискретным распределением $\boldsymbol{q}$. Эргодичность означает к тому же, что любые состояния такой цепи достижимы — имеют положительную вероятность осуществления (см. далее закон больших чисел).

438 Примеры. 1) Пусть начальное распределение $\boldsymbol{\pi} = ({}^1\!/_2, {}^1\!/_2, 0, 0)$, т.е. вначале процесс находится в состоянии либо e_0, либо e_1 с равными вероятностями. Пусть, кроме того, матрица переходных вероятностей

$$P = \frac{1}{10}\begin{pmatrix} 4 & 4 & 1 & 1 \\ 4 & 4 & 1 & 1 \\ 0 & 0 & 6 & 4 \\ 0 & 0 & 2 & 8 \end{pmatrix}.$$

Через пять шагов процесс будет находиться в одном из четырех состояний с вероятностями $\boldsymbol{\pi} P^5 \approx (0.16, 0.16, 0.25, 0.43)$. Через 50 шагов распределение состояний будет мало отличаться от предельного распределения $\boldsymbol{q} = (0, 0, {}^1\!/_3, {}^2\!/_3)$, которое также удовлетворяет уравнению $\boldsymbol{q}P = \boldsymbol{q}$. Следует сказать, что, независимо от начального состояния, процесс с такой матрицей переходных вероятностей через некоторое время обязательно выйдет на стационарный режим функционирования, в котором с вероятностями ${}^1\!/_3$ и ${}^2\!/_3$ будут присутствовать только состояния e_2 и e_3. Этот процесс не эргодический, т.е. не все его состояния достижимы с вероятностью 1.

2) Распределение капитала одного из игроков существенно зависит от начального распределения (см. пример 435). Уравнение $\boldsymbol{q}P = \boldsymbol{q}$ имеет здесь два решения $\boldsymbol{q} = (1, 0, \ldots, 0)$ и $\boldsymbol{q} = (0, \ldots, 0, 1)$, которые никак не связаны с предельным распределением процесса $\lim_t \boldsymbol{\pi} P^t$, зависящим к тому же от начального распределения капитала между игроками. ◎

✧ **Закон больших чисел.** Для стационарных процессов были установлены аналоги закона больших чисел. Подобный факт справедлив и для эргодических цепей Маркова.

439| **Теорема.** [*Закон больших чисел.*] *Пусть* $\langle \xi(t) \rangle_{t=1}^{\infty}$ *— однородная дискретная цепь Маркова, для которой существует предельное (стационарное) распределение* $\boldsymbol{q}$, $\mathfrak{N}_A(T) = \sum_{t=1}^{T} \dot{\mathrm{I}}_A(\xi(t))$ *— количество экспериментов (до* T*), в результате которых процесс оказывался в одном из состояний подмножества* A *пространства* $\mathscr{X}$*. Тогда для любого конечного* $A \subset \mathscr{X}$

$$\mathbf{P}\text{-}\lim_{T\to\infty} \frac{1}{T} \mathfrak{N}_A(T) = \boldsymbol{Q}(A),$$

где $\boldsymbol{Q}(A)$ *— вероятностная мера* A *относительно* $\boldsymbol{q}$.

⫤ Пусть сначала множество A содержит только одно состояние: $A = \{e_m\}$. Математическое ожидание $\mathbf{E}\,\dot{\mathrm{I}}_A(\xi(t)) = \mathbf{P}\{\xi(t) = e_m\} = p_m(t) \underset{t\to\infty}{\longrightarrow} q_m = \boldsymbol{Q}(A)$. Поэтому в силу леммы 293, стр. 255,

$$\mathbf{E}\left[\tfrac{1}{T}\mathfrak{N}_A(T)\right] \;=\; \frac{1}{T}\sum_{t=1}^{T} p_m(t) \;\underset{T\to\infty}{\longrightarrow}\; q_m \;=\; \boldsymbol{Q}(A).$$

Найдём математическое ожидание квадрата искомого среднего:

$$\mathbf{E}\left[\frac{1}{T}\mathfrak{N}_A(T)\right]^2 \;=\; \frac{1}{T^2}\sum_{t,u=1}^{T} \mathbf{E}\left[\,\dot{\mathrm{I}}_A(\xi(t))\,\dot{\mathrm{I}}_A(\xi(u))\right] \;=$$

$$=\;\frac{1}{T^2}\sum_{t,u=1}^{T} \mathbf{P}\left\{\xi(t)=e_m, \xi(u)=e_m\right\}.$$

Для $t<u$ имеем

$$\mathbf{P}\big\{\,\xi(t)=e_m, \xi(u)=e_m\,\big\} \;=$$

$$=\;\mathbf{P}\left\{\xi(u)=e_m \,\middle|\, \xi(t)=e_m\right\}\mathbf{P}\big\{\,\xi(u)=e_m\,\big\} \;=\; p_{m,m}(u-t)p_m(t).$$

По условию теоремы $p_{m,m}(s) \underset{s\to\infty}{\longrightarrow} q_m, p_m(s) \underset{s\to\infty}{\longrightarrow} q_m$. Воспользовавшись второй частью леммы 293, стр. 255, получаем, что среднеквадратическое отклонение

$$\mathbf{E}\left[\tfrac{1}{T}\mathfrak{N}_A(T)-q_m\right]^2 \;=\; \mathbf{E}\left[\tfrac{1}{T}\mathfrak{N}_A(T)\right]^2 + q_m^2 - 2q_m\,\mathbf{E}\left[\tfrac{1}{T}\mathfrak{N}_A(T)\right] \underset{T\to\infty}{\longrightarrow} 0.$$

Таким образом, $\dfrac{1}{T}\mathfrak{N}_A(T) \xrightarrow{\mathcal{L}_2} q_m$, следовательно, эта сходимость имеет место и по вероятности.

Для произвольного подмножества $A=\{e_{m_1},\ldots,e_{m_k}\}$, $k<\infty$, утверждение теоремы следует из представления

$$\frac{1}{T}\mathfrak{N}_A(T) \;=\; \textstyle\sum_{j=1}^{k}\frac{1}{T}\mathfrak{N}_{\{e_{m_j}\}}(T) \;\underset{T\to\infty}{\xrightarrow{\mathbf{P}}}\; \sum_{j=1}^{k} q_{m_j} \;=\; \boldsymbol{Q}(A). \qquad ⫤$$

⚠ У эргодической цепи Маркова вероятность любого непустого набора состояний $\boldsymbol{Q}(A)>0$. Как следствие закона больших чисел отсюда получаем, что на «бесконечном» промежутке времени траектория процесса, описывающего эргодическую цепь Маркова, «бесконечно» много раз как посетит любой (непустой и неисчерпывающий) набор состояний A, так и покинет его.

✧ **Марковские случайные функции.** Рассмотрим процесс с фазовым пространством $\mathscr{X}=\mathbb{R}^1$ (или $\mathbb{R}^k$). В этом случае распределение сечений процесса не может быть задано значениями вероятностей в каждой точке $\mathscr{X}$ — требуется знание вероятностей для всех измеримых подмножеств $\mathscr{X}$. Пусть набор измеримых подмножеств $\mathscr{X}$ описывается борелевской σ-алгеброй $\mathscr{B}$.

Определение. Случайная функция $\boldsymbol{\xi}=\left\langle\,\xi(t),\ t\in\mathbb{R}^1_+\,\right\rangle$, заданная на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$, определяет марковский процесс, если для $\forall\, C \in \mathscr{B}$, $\forall\, k \geqslant 1$, любых наборов параметров $s_1 < \ldots < s_k < r < t$ $(\in \mathbb{R}^1_+)$ и почти всех (по распределению сл.вектора $(\xi(t_1), \ldots, \xi(t_k), \xi(r))$) значений $x_1, \ldots, x_k, y$ $(\in \mathscr{X})$ условная вероятность

$$\begin{aligned} &\mathbf{P}\left\{\xi(t) \in C \;\middle|\; \xi(s_1) = x_1, \ldots, \xi(s_k) = x_k, \xi(r) = y\right\} = \\ &= \mathbf{P}\left\{\xi(t) \in C \;\middle|\; \xi(r) = y\right\}. \end{aligned} \tag{45}$$

Запись марковского свойства и его трактовка для сл.функции и цепи Маркова вполне идентичны, если не вдаваться в подробности определения условной вероятности. Для недискретных сл.в. уравнение (45) означает (см. стр. 136, 147), что если функция $\eta : (\mathbb{R}^1, \mathscr{B}) \mapsto (\mathbb{R}^1, \mathscr{B})$ такова, что для $\forall\, A_r \in \mathscr{B}$

$$\mathbf{E}\left[\,\dot{\mathrm{I}}_C(\xi_t)\,\dot{\mathrm{I}}_{A_r}(\xi_r)\right] = \mathbf{E}\left[\eta(\xi_t)\,\dot{\mathrm{I}}_{A_r}(\xi_r)\right],$$

то справедливо $\mathbf{E}\left[\,\dot{\mathrm{I}}_C(\xi_t)\,\dot{\mathrm{I}}_A(\vec{\xi}_1, \xi_r)\right] = \mathbf{E}\left[\eta(\xi_t)\,\dot{\mathrm{I}}_A(\vec{\xi}_1, \xi_r)\right]$ для любого $A \in \mathscr{B}^{k+1}$, где $\vec{\xi}_1 = (\xi(t_1), \ldots, \xi(t_k))$. Математическое ожидания с обеих сторон этого соотношения определяют меры на множествах $A \in \mathscr{B}^{k+1}$. По теореме Каратеодори совпадение таких мер достаточно проверить на множествах $A = \prod_1^k B_j \times A_r$, $B_1, \ldots, B_k, A_r \in \mathscr{B}$. Как и для цепей Маркова, марковское свойство для сл.функций может быть переформулировано в виде свойства условной независимости будущего от прошлого относительно настоящего.

440| **Теорема.** *Случайная функция* $\boldsymbol{\xi}$ *определяет марковский процесс т. т. т. когда для* $\forall\, k \geqslant 1$, $\forall\, B_1, \ldots, B_k, C \in \mathscr{B}$, *любых наборов параметров* $s_1 < \ldots < s_k < r < t$ $(\in \mathbb{R}^1_+)$ *и почти всех (по распределению сл.в.* $\xi(r)$*) значений* y $(\in \mathscr{X})$ *условная вероятность*

$$\begin{aligned} &\mathbf{P}\left\{\bigcap_1^k \big(\xi(s_j) \in B_j\big) \cap \big(\xi(t) \in C\big) \mid \xi(r) = y\right\} = \\ &= \mathbf{P}\left\{\bigcap_1^k \big(\xi(s_j) \in B_j\big) \mid \xi(r) = y\right\} \cdot \mathbf{P}\left\{\xi(t) \in C \mid \xi(r) = y\right\}. \end{aligned}$$

$\Rrightarrow$ Ввиду отсутствия принципиальных отличий рассмотрим только $k = 1$. Как отмечалось выше, необходимо показать, что для любых подмножеств $A_r, B_1, C \in \mathscr{B}$ имеет место равенство

$$\mathbf{E}\left[\,\dot{\mathrm{I}}_A(\xi_{s_1}, \xi_r)\,\dot{\mathrm{I}}_C(\xi_t)\right] = \mathbf{E}\left[\,\dot{\mathrm{I}}_{B_1 \times A_r}(\xi_{s_1}, \xi_r)\;\mathbf{E}\left[\,\dot{\mathrm{I}}_C(\xi_t) \mid \xi_r\right]\right].$$

Условие теоремы эквивалентно тому, что для $\forall\, B_1, A_r \in \mathscr{B}$

$$\begin{aligned}\mathbf{E}\left[\,\dot{\mathrm{I}}_{A_r}(\xi_r)\,\dot{\mathrm{I}}_{B_1}(\xi_{s_1})\,\dot{\mathrm{I}}_C(\xi_t)\right] &= \mathbf{E}\left[\,\dot{\mathrm{I}}_{A_r}(\xi_r)\,\mathbf{E}\big[\,\dot{\mathrm{I}}_{B_1}(\xi_{s_1})\,|\,\xi_r\,\big]\,\mathbf{E}\big[\,\dot{\mathrm{I}}_C(\xi_t)\,|\,\xi_r\,\big]\right] = \\ &= \mathbf{E}\left[\mathbf{E}\big[\,\dot{\mathrm{I}}_{A_r}(\xi_r)\,\dot{\mathrm{I}}_{B_1}(\xi_{s_1})\,|\,\xi_r\,\big]\,\mathbf{E}\big[\,\dot{\mathrm{I}}_C(\xi_t)\,|\,\xi_r\,\big]\right] = \\ &= \mathbf{E}\Big[\mathbf{E}\Big[\Big(\,\dot{\mathrm{I}}_{B_1\times A_r}(\xi_{s_1},\xi_r)\,\mathbf{E}\big[\,\dot{\mathrm{I}}_C(\xi_t)\,|\,\xi_r\,\big]\Big)\,|\,\xi_r\,\Big]\Big] = \\ &= \mathbf{E}\left[\,\dot{\mathrm{I}}_{B_1\times A_r}(\xi_{s_1},\xi_r)\,\mathbf{E}\big[\,\dot{\mathrm{I}}_C(\xi_t)\,|\,\xi_r\,\big]\right],\end{aligned}$$

где равенства в последних трёх строках следуют из свойств усл.м.о. Что и требовалось. Эти же равенства могут быть использованы и для доказательства обратного утверждения. ⫤

Стандартным приёмом перехода от простых функций к измеримым показывается справедливость следующей теоремы.

441 **Теорема.** *Случайная функция $\boldsymbol{\xi}$ определяет марковский процесс т. т. т. когда для любых наборов параметров $s_1 < \ldots < s_k < r < t$ $(\in \mathbb{R}^1_+)$ любого множества $A \in \mathscr{B}^{k+1}$ и любой ограниченной борелевской функции $h : (\mathbb{R}^1, \mathscr{B}) \mapsto (\mathbb{R}^1, \mathscr{B})$*

$$\mathbf{E}\left[h(\xi_t)\,\dot{\mathrm{I}}_A(\xi_{s_1},\ldots,\xi_{s_k},\xi_r)\right] = \mathbf{E}\left[\,\mathbf{E}\big[\,h(\xi_t)\,|\,\xi_r\,\big]\,\dot{\mathrm{I}}_A(\xi_{s_1},\ldots,\xi_{s_k},\xi_r)\right].$$

✧ **Уравнение Колмогорова–Чепмена.** Воспользуемся теоремой 441 для вывода формулы конечномерных распределений марковской сл.функции. Сначала заметим, что при выводе распределения цепи Маркова мы опирались на правую часть уравнения, описывающего марковское свойство, введя понятие вероятности перехода из одного состояния в другое. Для марковских сл.функций мы поступим так же. Правая часть (45) при корректном определении усл.вер. (см. стр. 151 и далее о регулярных условных распределениях) представляет собой при каждом $t, s \geqslant 0$ и $y \in \mathscr{X}$ вероятностную меру на борелевских множествах $C \in \mathscr{B}$. Точнее, для каждой пары $0 \leqslant s < t$ существует функция $\Psi(s,x;t,C)$ (регулярное условное распределение), измеримая по y при каждом фиксированном $C \in \mathscr{B}$, которая при каждом $x \in \mathbb{R}^1$ есть вероятностная мера (по C) на $\mathscr{B}$ и $\Psi(s,\xi_s;t,C) = \mathbf{P}\left\{\xi_t \in C \,\middle|\, \xi_s\right\}$. Функция $\Psi(s,x;t,C)$ называется *переходной вероятностью* марковского процесса из точки x в момент времени s в множество C в момент времени t. Заметим, что усл.м.о. функции $h(\xi_t)$ относительно σ-алгебры, порождённой сл.в. ξ_s, можно вычислить как лебеговский интеграл по мере Ψ :

$$\mathbf{E}\big[\, h(\xi_t) \,|\, \xi_s = x \,\big] \;=\; \int_{\mathbb{R}} h(y)\, \Psi(s,x;t,dy) \;= \int_{\mathbb{R}} h(y)\psi(s,x;t,y)\,dy,$$

где последнее равенство справедливо, если мера $\Psi(s,x;t,\cdot)$ имеет плотность $\psi(s,x;t,y)$ относительно меры Лебега.

Пусть $\Psi_t(C) = \mathbf{P}\{\xi_t \in C\}$, $C \in \mathscr{B}$, — распределение сечения процесса ξ_t в точке t. По определению усл.м.о. для $s < u$

$$\mathbf{P}\{\xi_s \in B_1, \xi_u \in B_3\} \;=\; \mathbf{E}\big[\,\mathbf{E}\big[\,\dot{\mathrm{I}}_{B_3}(\xi_u) \,|\, \xi_s\,\big]\,\dot{\mathrm{I}}_{B_1}(\xi_s)\big] = \int_{B_1} \Psi(s,x_1;u,B_3)\,\Psi_s(dx_1)\,.$$

В частности, полагая здесь $B_1 = \mathbb{R}^1, B_3 = C \subset \mathbb{R}^1$, получаем

$$\Psi_u(C) \;=\; \mathbf{P}\{\xi_u \in C\} \;=\; \int_{\mathbb{R}} \Psi(s,x_1,u,C)\,\Psi_s(dx_1).$$

В силу марковского свойства для $\forall\, s < t < u$ и $\forall\, B_1, B_2, B_3 \in \mathscr{B}$ справедливо представление

$$\begin{aligned}
\mathbf{P}\{\xi_s \in B_1, \xi_t \in B_2, \xi_u \in B_3\} \;&=\; \mathbf{E}\big[\,\mathbf{E}\big[\,\dot{\mathrm{I}}_{B_3}(\xi_u) \,|\, \xi_s, \xi_t\,\big]\,\dot{\mathrm{I}}_{B_1}(\xi_s)\,\dot{\mathrm{I}}_{B_2}(\xi_t)\big] \;= \\
=\; \mathbf{E}\big[\,\mathbf{E}\big[\,\dot{\mathrm{I}}_{B_3}(\xi_u) \,|\, \xi_t\,\big]\,\dot{\mathrm{I}}_{B_1}(\xi_s)\,\dot{\mathrm{I}}_{B_2}(\xi_t)\big] \;&=\; \mathbf{E}\big[\Psi(t,\xi_t;u,B_3)\,\dot{\mathrm{I}}_{B_1}(\xi_s)\,\dot{\mathrm{I}}_{B_2}(\xi_t)\big] \;= \\
=\; \mathbf{E}\Big(\,\mathbf{E}\big[\,\Psi(t,\xi_t;u,B_3)\,\dot{\mathrm{I}}_{B_2}(\xi_t) \,|\, \xi_s\,\big]\,\dot{\mathrm{I}}_{B_1}(\xi_s)\Big) \;&= \\
=\; \int_{B_1}\int_{B_2} \Psi(t,x_2;u,B_3)\,\Psi(s,x_1;t,dx_2)\,\Psi_s(dx_1)\,. &
\end{aligned} \tag{46}$$

Воспользовавшись утверждением 441 , можно доказать аналогичное равенство для любых наборов $t_0 < \ldots < t_k$, $B_0, \ldots, B_k \in \mathscr{B}$:

$$\begin{aligned}
\mathbf{P}\Big\{\bigcap\nolimits_0^k \xi_{t_j} \in B_j\Big\} = \int_{B_0}\!\!\cdots\int_{B_{k-1}} \Psi(t_{k-1},x_{k-1};t_k,B_k)\times \\
\times\prod\nolimits_1^{k-1} \Psi(t_{k-j-1},x_{k-j-1};t_{k-j},dx_{k-j})\,\Psi_{t_0}(dx_0)\,.
\end{aligned} \tag{47}$$

Так как $\mathbf{P}\{\xi_s \in B_1, \xi_u \in B_3\} = \mathbf{P}\big\{\xi_s \in B_1, \xi_t \in \mathbb{R}^1, \xi_u \in B_3\big\}$, то из (46) и предыдущего получаем, что для $\forall\, s < t < u$ переходная вероятность должна удовлетворять уравнению

$$\Psi(s,x;u,B) = \int\limits_{y\in\mathbb{R}^1} \Psi(t,y;u,B)\,\Psi(s,x;t,dy). \tag{48}$$

Это уравнение называется *уравнением Колмогорова–Чепмена*. Понятно, что (48) обеспечивает согласованность конечномерных вероятностей (47). Другими словами, семейство функций $\Psi(s,x;t,B)$, $0 < s < t$, $x \in \mathbb{R}^1$, $B \in \mathscr{B}$, измеримых по x при каждом фиксированном наборе s, t, B и определяющих вероятностную меру на $\mathscr{B}$ при каждом наборе s, t, x, будет задавать переходные вероятности некоторой марковской сл.функции, если это семейство удовлетворяет уравнениям Колмогорова–Чепмена. В этом случае конечномерные распределения находятся по формулам (47), в которых всегда можно взять $t_0 = 0$ и некоторое начальное распределение $\Psi_0(B)$, $B \in \mathscr{B}$. Если переходные вероятности имеют ф.пл. $\psi(s, x; t, y)$, $y \in \mathbb{R}^1$, то любое конечномерное распределение $(\xi(t_1), \ldots, \xi(t_k))$ также имеет плотность:

$$p_{\boldsymbol{\xi}}(x_1, \ldots, x_k; t_1, \ldots, t_k) = \int\limits_{\mathbb{R}} \prod_{j=1}^{k} \psi(t_{j-1}, x_{j-1}; t_j, x_j)\, \Psi_0(dx_0) \tag{49}$$

для $\forall\, 0 = t_0 \leqslant t_1 < \ldots < t_k$. Уравнение Колмогорова–Чепмена эквивалентно

$$\psi(s, x; u, z) = \int_{\mathbb{R}} \psi(t, y; u, z) \psi(s, x; t, y)\, dy. \tag{50}$$

442| П р и м е р. Процесс броуновского движения, как процесс с независимыми приращениями, есть марковский процесс. В качестве иллюстрации покажем, что переходные вероятности этого процесса удовлетворяют уравнению Колмогорова–Чепмена. По теореме 400, стр. 350, условное распределение ξ_t при фиксированном $\xi_s = x$ $(s < t)$ совпадает с нормальным $\mathcal{N}(x, t-s)$ распределением, т.е. плотность переходного распределения (по y)

$$\psi(s, x; t, y) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\Big\{ -\frac{1}{2} \frac{(y-x)^2}{t-s} \Big\}.$$

Правая часть уравнения (50) равна

$$\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi(t-s)}} \exp\Big\{ -\frac{1}{2} \frac{(y-x)^2}{t-s} \Big\} \frac{1}{\sqrt{2\pi(u-t)}} \exp\Big\{ -\frac{1}{2} \frac{(z-y)^2}{u-t} \Big\}\, dy$$

и представляет собой свёртку плотностей нормального $\mathcal{N}(0, u-t)$ распределения и нормального $\mathcal{N}(x, t-s)$ распределения. Хорошо известно, что эта свёртка есть плотность нормального $\mathcal{N}(x, u-t+t-s)$ распределения.

Винеровский процесс выходит из нуля, т.е. его начальное распределение сосредоточено в нуле — $\Psi_0(\{0\}) = 1$. С учётом этого обстоятельства и вида переходной плотности выражение (49) полностью совпадает с плотностью конечномерного распределения винеровского процесса (21), стр. 350. ◎

§ 5. Мартингалы

Для того чтобы проиллюстрировать идею мартингальной теории, рассмотрим снова вводный к этой главе пример, т.е. будем считать, что имеется последовательность независимых сл.в. $\{\xi_k\}_1^\infty$, принимающих значения ± 1. Игрок A в каждой из игр меняет свою ставку в зависимости от результатов предыдущих игр. Другими словами, последовательность его ставок $\{R_k\}_1^\infty$ есть последовательность сл.в., при этом k-я ставка зависит от реализаций $(k-1)$ предыдущих игр: $R_k = R_k(\xi_1, \ldots, \xi_{k-1})$. Самый простой способ формализации этой зависимости состоит в описании сл.в. R_k как функции на измеримом пространстве $(\Omega, \mathscr{F})$, измеримой относительно σ-алгебры, порождённой вектором сл.в. $(\xi_1, \ldots, \xi_{k-1})$; обозначим эту σ-алгебру через $\mathscr{F}_{k-1}$. Доход игрока A после k проведённых игр

$$\eta_k = \textstyle\sum_1^k R_j \xi_j = \eta_{k-1} + R_k \xi_k.$$

Принимая решение о будущей ставке, игрок должен ориентироваться на величину усл.м.о. будущего дохода при фиксированных результатах предыдущих игр. В данном случае — на величину

$$\mathbf{E}\big[\,\eta_k \,\big|\, \mathscr{F}_{k-1}\big] = \eta_{k-1} + R_k\,\mathbf{E}\,\xi_k,$$

где равенство справедливо в силу измеримости η_{k-1} и R_k относительно σ-алгебры $\mathscr{F}_{k-1}$ (т.е. зависимости этих величин только от $\xi_1, \ldots, \xi_{k-1}$) и независимости ξ_k от $\mathscr{F}_{k-1}$. Если игра безобидная, т.е. $\mathbf{E}\,\xi_k = 0$, то мы имеем равенство $\mathbf{E}\big[\,\eta_k \,\big|\, \mathscr{F}_{k-1}\big] = \eta_{k-1}$, что характеризует последовательность η_k как мартингал. Если же $\mathbf{E}\,\xi_k > 0$ (или < 0), то $\{\eta_k\}_1^\infty$ образует субмартингал (супермартингал). Заметим, что определение этой последовательности тесно связано с последовательностью σ-алгебр, т.е. с последовательностью наблюдаемых сл.в., на которых формируются значения изучаемых мартингалов.

Определения. Пусть $(\Omega, \mathscr{F}, \mathbf{P})$ — вероятностное пространство.

a) Возрастающая последовательность σ-алгебр $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \mathscr{F}_2 \subset \ldots \subset \mathscr{F}$ называется *потоком σ-алгебр* или *фильтрацией*.

b) Последовательность сл.в. $\boldsymbol{\xi} = \{\xi_k\}_0^\infty$ образует *мартингал* относительно потока σ-алгебр $\mathcal{F} = \{\mathscr{F}_k\}_0^\infty$, если

i) $\boldsymbol{\xi}$ согласована с фильтрацией $\mathcal{F}$, т.е. для $\forall\, k \geqslant 0$ сл.в. ξ_k измерима относительно σ-алгебры $\mathscr{F}_k$ (обозначим $\xi_k \Subset \mathscr{F}_k$);

ii) математическое ожидание $\mathbf{E}\,|\xi_k| < \infty,\ \forall\, k \geqslant 0$;

iii) для $\forall\, k \geqslant 0$ усл.м.о.

$$\mathbf{E}\bigl[\,\xi_{k+1}\,\big|\,\mathscr{F}_k\,\bigr] = \xi_k \quad (\text{п.н.})\,.$$

Если $\mathbf{E}\bigl[\,\xi_{k+1}\,\big|\,\mathscr{F}_k\,\bigr] \geqslant \xi_k,\ k \geqslant 0$, то $\boldsymbol{\xi}$ называется *субмартингалом*; при обратном неравенстве — *супермартингалом*.

c) *Естественной фильтрацией* (для последовательности сл.в. $\boldsymbol{\eta} = \{\eta_k\}_0^\infty$) называется поток σ-алгебр $\mathcal{F}^{\boldsymbol{\eta}}$, в котором $\mathscr{F}_k = \sigma\bigl(\eta_0, \eta_1, \ldots, \eta_k\bigr)$. Последовательность $\boldsymbol{\xi} = \{\xi_k\}_0^\infty$ образует мартингал относительно этой фильтрации, если $\mathbf{E}\bigl[\,\xi_{k+1}\,\big|\,\eta_0, \ldots, \eta_k\,\bigr] = \xi_k$, с соответствующими изменениями для суб- и супермартингалов.

Поскольку определение мартингалов и субмартингалов опирается на соответствующий поток σ-алгебр, то принято в их обозначении указывать этот поток: $\{\boldsymbol{\xi}, \mathcal{F}\} = \{\xi_k, \mathscr{F}_k\}_0^\infty$. В ситуациях, когда в записи мартингала фильтрация отсутствует, подразумевается, что имеется в виду естественная для $\boldsymbol{\xi}$ фильтрация $\mathcal{F}^{\boldsymbol{\xi}}$.

Ясно, что если ξ_k — субмартингал, то $(-\xi_k)$ — супермартингал, и наоборот, поэтому все теоретические утверждения относительно субмартингалов легко переносятся на супермартингалы.

Поскольку сложных вопросов существования усл.м.о. с тем или иным свойством здесь не возникнет, мы будем опускать напоминание в виде записи (п.н.) о том, что все эти свойства выполняются лишь почти наверное.

Основной рабочий инструмент в теории мартингалов — телескопическое свойство усл.м.о.:

$$\mathbf{E}\bigl[\,\mathbf{E}\bigl[\,\xi\,\big|\,\mathscr{L}\,\bigr]\,\big|\,\mathscr{M}\,\bigr] \;=\; \mathbf{E}\bigl[\,\xi\,\big|\,\mathscr{L}\,\bigr] \;=\; \mathbf{E}\bigl[\,\mathbf{E}\bigl[\,\xi\,\big|\,\mathscr{M}\,\bigr]\,\big|\,\mathscr{L}\,\bigr]$$

для вложенных сигма-подалгебр $\mathscr{L} \subset \mathscr{M} \subset \mathscr{F}$. Напомним также, что i) если $\xi \Subset \mathscr{L}$, то $\mathbf{E}\bigl[\,\xi\,\big|\,\mathscr{L}\,\bigr] = \xi$, и ii) $\mathbf{E}\bigl[\,\xi\,\big|\,\mathscr{L}\,\bigr] = \mathbf{E}\,\xi$, если ξ не зависит от $\mathscr{L}$, т.е. не зависит от индикаторов $\dot{\mathrm{I}}_A,\ \forall\, A \in \mathscr{L}$.

443| Упр. Покажите, что для мартингала $\boldsymbol{\xi}$ справедливо a) $\mathbf{E}\,\xi_k = \mathbf{E}\,\xi_0$, б) $\mathbf{E}\bigl[\,\xi_{k+m}\,\big|\,\mathscr{F}_k\,\bigr] = \xi_k,\ k, m \geqslant 0$.

Кроме того, здесь часто используются классические интегральные неравенства, адаптированные к усл.м.о.

444| **Теорема.** [*Неравенство Йенсена.*] *Если h — выпуклая книзу* *борелевская функция на* $\mathbb{R}^1$, *сл.в.* ξ *на вероятностном пространстве* $(\Omega, \mathscr{F}, \mathbf{P})$ *имеет конечное математическое ожидание* $\mathbf{E}\xi$, *то для* σ*-алгебры* $\mathscr{L} \subset \mathscr{F}$

$$\mathbf{E}\big[\, h(\xi) \,|\, \mathscr{L} \,\big] \geqslant h\big(\mathbf{E}\big[\, \xi \,|\, \mathscr{L} \,\big]\big).$$

Доказательство этого неравенства в общем случае ничем не отличается от доказательства неравенства Йенсена для интегралов (см. следующий пример).

445| Пример. Пусть $\{\xi_k, \mathscr{F}_k\}_0^\infty$ — мартингал, тогда $\{\xi_k^2, \mathscr{F}_k\}_0^\infty$ — субмартингал. Действительно, для квадратической функции $h(x) = x^2$ при $\forall\, x, \mu \in \mathbb{R}^1$ справедливо неравенство $h(x) \geqslant 2\mu(x-\mu) + \mu^2$ (справа здесь стоит касательная к параболе в точке $x_0 = \mu$). Возьмём $\mu = \mathbf{E}\big[\, \xi_{k+1} \,|\, \mathscr{F}_k \,\big] = \xi_k$, тогда

$$\mathbf{E}\big[\, \xi_{k+1}^2 \,|\, \mathscr{F}_k \,\big] \;\geqslant\; \mathbf{E}\big[\, 2\xi_k(\xi_{k+1} - \xi_k) + \xi_k^2 \,|\, \mathscr{F}_k \,\big] \;=\; \xi_k^2$$

в силу $\mathscr{F}_k$-измеримости ξ_k. Аналогично для $|\xi_k|$ и вообще для любой выпуклой книзу функции $h(\xi_k)$. ◎

446| Примеры. 1) Пусть $\mathcal{F} = \{\mathscr{F}_k\}_0^\infty$ — некоторая фильтрация на $(\Omega, \mathscr{F}, \mathbf{P})$, сл.в. ξ интегрируема: $\mathbf{E}\,|\xi| < \infty$. Тогда последовательность $\eta_k = \mathbf{E}\big[\, \xi \,|\, \mathscr{F}_k \,\big]$, $k \geqslant 1$, образует мартингал (мартингал П. Леви) относительно $\mathcal{F}$. Во-первых, по неравенству Йенсена

$$\mathbf{E}\,|\eta_k| \;\leqslant\; \mathbf{E}\big[\, \mathbf{E}\big[\, |\xi| \,|\, \mathscr{F}_k \,\big]\big] \;=\; \mathbf{E}\,|\xi| \;<\; \infty.$$

Во-вторых, по телескопическому свойству (ввиду вложенности $\mathscr{F}_k \subset \mathscr{F}_{k+1}$)

$$\mathbf{E}\big[\, \eta_{k+1} \,|\, \mathscr{F}_k \,\big] \;=\; \mathbf{E}\big[\, \mathbf{E}\big[\, \xi \,|\, \mathscr{F}_{k+1} \,\big] \,|\, \mathscr{F}_k \,\big] \;=\; \mathbf{E}\big[\, \xi \,|\, \mathscr{F}_k \,\big] \;=\; \eta_k\,.$$

2) Суммы сл.в. $S_k = \sum_1^k \xi_j$ с независимыми слагаемыми и нулевыми средними $\mathbf{E}\xi_j = 0$ образуют мартингал относительно естественной фильтрации $\mathcal{F}^{\boldsymbol{\xi}}$. Очевидно, сл.в. ξ_{k+1} не зависит от σ-алгебры $\mathscr{F}_k$ и усл.м.о $\mathbf{E}\big[\, \xi_{k+1} \,|\, \mathscr{F}_k \,\big] = \mathbf{E}\,\xi_{k+1} = 0$. Поэтому

$$\mathbf{E}\big[\, S_{k+1} \,|\, \mathscr{F}_k \,\big] \;=\; \mathbf{E}\big[\, S_k \,|\, \mathscr{F}_k \,\big] + \mathbf{E}\big[\, \xi_{k+1} \,|\, \mathscr{F}_k \,\big] \;=\; S_k$$

в силу свойств усл.м.о. При $\mathbf{E}\xi_j > 0$ последовательность S_k образует субмартингал.

3) Аналогично, если $\mathbf{E}\xi_k = 1$, $k \geqslant 1$, и сл.в. ξ_k, $k \geqslant 1$, независимы в

совокупности, то произведения $P_n = \prod_1^n \xi_k$ образуют мартингал относительно естественной фильтрации:

$$\mathbf{E}\big[\, P_{n+1} \mid \mathscr{F}_n \,\big] \;=\; \mathbf{E}\Big[\, \prod_1^{n+1} \xi_k \mid \mathscr{F}_n \,\Big] \;=\; \Big(\prod_1^{n} \xi_k\Big)\, \mathbf{E}\,\xi_{k+1} \;=\; P_n\,.$$

4) Пусть $\eta_0 = 0$ и $\langle \eta_k \rangle_0^\infty$ — независимые одинаково распределённые сл.в., для которых при некоторых $\lambda \neq 0$ существует производящая функция $\psi(\lambda) = \mathbf{E}\, e^{\lambda\eta_1} < \infty$. Тогда для $S_n = \eta_1 + \ldots + \eta_n$ последовательность

$$\xi_0 = 1, \quad \xi_n = \frac{1}{\psi^n(\lambda)} \exp\{\lambda S_n\}, \quad n \geqslant 1,$$

образует мартингал (мартингал Вальда) относительно потока $\mathscr{F}^{\boldsymbol{\eta}}$.

Например, если $\eta_k \backsim \mathcal{N}(\mu, \sigma^2)$, то для $\forall\, \lambda \in \mathbb{R}^1$ производящая функция $\psi(\lambda) = \exp\{\lambda\mu + \frac{1}{2}\,\lambda^2\sigma^2\}$ и мартингал Вальда

$$\xi_n = \exp\{\lambda\big(S_n - n(\mu + \tfrac{1}{2}\,\lambda\sigma^2)\big)\}.$$

В частности, $\xi_k = \exp\{-2\mu S_k/\sigma^2\}$, $k \geqslant 1$, — мартингал.

5) Предположим, что наблюдается последовательность сл.в. $\boldsymbol{\xi} = \{\xi_k\}_1^\infty$. Сигма-алгебра $\mathscr{F}_n = \sigma\big(\xi_1, \ldots, \xi_n\big)$ описывает набор информации, которой может обладать наблюдатель после n проведённых экспериментов; $\mathscr{F}_0 = \{\varnothing, \Omega\}$. Как известно, наилучший прогноз значения следующего $(n{+}1)$-го наблюдения равен усл.м.о. $\mathbf{E}\big[\,\xi_{n+1} \mid \mathscr{F}_n\,\big]$. Суммарная ошибка прогнозов

$$m_n = \sum_{k=0}^{n-1} \Big(\xi_{k+1} - \mathbf{E}\big[\,\xi_{k+1} \mid \mathscr{F}_k\,\big]\Big), \quad n \geqslant 1,$$

образует мартингал относительно естественной фильтрации $\mathcal{F}^{\boldsymbol{\xi}}$. Действительно, по телескопическому свойству усл.м.о. и в силу $\mathscr{F}_n$-измеримости сл.в. ξ_{k+1} для $k < n$ имеем

$$\begin{aligned}\mathbf{E}\big[\, m_{n+1} \mid \mathscr{F}_n \,\big] \;&=\; \sum_{k=0}^{n} \Big(\mathbf{E}\big[\,\xi_{k+1} \mid \mathscr{F}_n\,\big] - \mathbf{E}\big[\,\xi_{k+1} \mid \mathscr{F}_k\,\big]\Big) \;=\\ &=\; \sum_{k=0}^{n-1} \Big(\xi_{k+1} - \mathbf{E}\big[\,\xi_{k+1} \mid \mathscr{F}_k\,\big]\Big) \;=\; m_n.\end{aligned}$$

6) Пусть $(\mathscr{X}, \mathscr{B})$ — некоторое измеримое пространство, интерпретиру-

емое как множество значений наблюдаемой сл.в. (например, действительная прямая с борелевской σ-алгеброй), $f_0(x), f_1(x)$ — функции плотности относительно некоторой меры μ на $\mathscr{B}$. Образуем пространство $\Omega = \mathscr{X}^{\mathbb{N}}$, $\mathbb{N} = \{1, 2, \ldots\}$, как пространство всех последовательностей из $\mathscr{X}$ с цилиндрической σ-алгеброй $\mathscr{F} = \mathscr{B}^{\mathbb{N}}$ и вероятностным распределением $\mathbf{P}$, определяемым конечномерными вероятностями $(B_1, \ldots, B_n \in \mathscr{B}, n \geqslant 1)$

$$\mathbf{P}\big\{\, B_1 \times \cdots \times B_n \,\big\} = \prod_{k=1}^{n} \int_{B_k} f_0(x)\, d\mu.$$

Вероятностное пространство $(\Omega, \mathscr{F}, \mathbf{P})$ формализует понятие последовательности независимых одинаково распределённых (с плотностью f_0) сл.в. в терминах распределений этих сл.в. Здесь в естественной фильтрации σ-алгебра $\mathscr{F}_k = \mathscr{B}^k \times \mathscr{X}^{\mathbb{N}\setminus\{1,\ldots,k\}}$ и представляет собой минимальный набор подмножеств, описываемых реализациями первых k сл.в. $\xi_k(\boldsymbol{x}) = x_k$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_k, \ldots) \in \Omega$. Рассмотрим сл.в. на $(\Omega, \mathscr{F})$

$$\ell_k(\boldsymbol{x}) = \frac{f_1(x_k)}{f_0(x_k)}, \quad \boldsymbol{x} = (x_1, x_2, \ldots, x_k, \ldots) \in \Omega,$$

если $f_0(x_k) \neq 0$. При $f_0(x_k) = 0$ можно положить $\ell_k(\boldsymbol{x}) = 0$, т.к. вероятность $\mathbf{P}\{\boldsymbol{x} : f_0(x_k) = 0\} = \int_{\{f_0(x_k)=0\}} f_0(x_k)\, d\mu = 0$. Очевидно, сл.в. ℓ_k, $k \geqslant 1$, независимы в совокупности, и математическое ожидание

$$\mathbf{E}\,\ell_k = \int\limits_{\{f_0(x_k)\neq 0\}} \frac{f_1(x_k)}{f_0(x_k)} f_0(x_k)\, d\mu = \int\limits_{\{f_0(x_k)\neq 0\}} f_1(x_k)\, d\mu.$$

Если распределение, определяемое плотностью f_1, абсолютно непрерывно относительно распределения с плотностью f_0, то последний интеграл равен единице. И этом случае (как в примере 3) в этом блоке) последовательность

$$\mathscr{L}_n = \prod_{k=1}^{n} \frac{f_1(\xi_k)}{f_0(\xi_k)}, \qquad n \geqslant 1,$$

образует мартингал относительно естественной фильтрации и вероятностной меры, определяемой плотностью f_0. В математической статистике сл.в. $\mathscr{L}_n$, называемая отношением правдоподобия, участвует в формировании оптимального критерия, предназначенного для различения двух простых гипотез $H_0 : \xi \backsim f_0$ и $H_1 : \xi \backsim f_1$ о распределении наблюдаемой сл.в. ◎

✧ **Разложение Дуба для субмартингалов**. Пусть $\{\xi_n\}_0^\infty$ — некоторый субмартингал относительно потока σ-алгебр $\mathcal{F} = \{\mathscr{F}_n\}_0^\infty$. Образуем мартингал (см. пример 4, 446)

$$m_n = \xi_0 + \sum\nolimits_0^{n-1} \Big(\xi_{k+1} - \mathbf{E}\big[\,\xi_{k+1} \mid \mathscr{F}_k\,\big]\Big), \quad n \geqslant 1,$$

$m_0 = \xi_0$. Разность между субмартингалом ξ_n и мартингалом m_n

$$A_n = \sum\nolimits_0^{n-1} \Big(\mathbf{E}\big[\,\xi_{k+1} \mid \mathscr{F}_k\,\big] - \xi_k\Big)$$

будет $\mathscr{F}_{n-1}$-измеримой сл.в., т.е. последовательность $\{A_n\}_1^\infty$ предсказуема относительно потока σ-алгебр $\mathcal{F}$. Кроме того, из определения субмартингала следует неотрицательность и монотонность A_n: $A_n \leqslant A_{n+1}$, $n \geqslant 0$. Представление субмартингала $\boldsymbol{\xi}$ в виде

$$\begin{aligned} \xi_n &= m_n + A_n, \\ &m_n \ \text{ — мартингал}, \\ &0 \leqslant A_n \leqslant A_{n+1}, \qquad A_n \Subset \mathscr{F}_{n-1},\ n \geqslant 1, \quad A_0 = 0, \end{aligned}$$

называется *разложением Дуба* для субмартингалов; последовательность A_n, $n \geqslant 0$, называется *компенсатором субмартингала*. Несложно показать, что подобное разложение единственно с точностью до нулевой вероятности.

Компенсатор субмартингала ξ_n^2, образованного из квадратично-интегрируемого мартингала $\boldsymbol{\xi} = \{\xi_n\}_0^\infty$, называется *квадратической характеристикой* мартингала $\boldsymbol{\xi}$ и обозначается $\big\langle \boldsymbol{\xi} \big\rangle$. Таким образом, для квадратично-интегрируемого мартингала $\boldsymbol{\xi}$ справедливо представление

$$\boldsymbol{\xi}^2 = \boldsymbol{m} + \big\langle \boldsymbol{\xi} \big\rangle.$$

447| П р и м е р. Пусть $\{\xi_1, \xi_2 \ldots\}$ — последовательность независимых сл.в. с нулевыми математическими ожиданиями $\mathbf{E}\,\xi_k = 0$, $k \geqslant 1$, и конечными дисперсиями $\sigma_k^2 = \mathbf{E}\,\xi_k^2 < \infty$. Как было показано ранее, последовательность сумм $S_n = \sum_1^n \xi_k$ образует мартингал относительно естественной фильтрации $\mathscr{F}_n = \sigma\big(\xi_1, \ldots, \xi_n\big)$. Заметим, что т.к. $\mathbf{E}\big[\,S_{n+1} \mid \mathscr{F}_n\,\big] = S_n$, то

$$\mathbf{E}\big[\,S_{n+1}^2 \mid \mathscr{F}_n\,\big] - S_n^2 \ = \ \mathbf{E}\big[\,(S_{n+1} - S_n)^2 \mid \mathscr{F}_n\,\big]$$

(ср. с двумя способами вычисления дисперсии). Поэтому квадратическая характеристика этого мартингала

$$\langle \boldsymbol{S}\rangle_n = \sum\nolimits_{k=0}^{n-1}\Big(\mathbf{E}\big[S_{k+1}^2 \mid \mathscr{F}_k\big] - S_k^2\Big) = \sum\nolimits_{k=0}^{n-1}\big(\mathbf{E}\big[\xi_{k+1}^2 \mid \mathscr{F}_k\big]\big) =$$
$$= \sum\nolimits_{k=0}^{n-1}\sigma_{k+1}^2 = \sum\nolimits_{k=1}^{n}\sigma_k^2 = \mathbf{D}S_n$$

ввиду независимости ξ_{k+1} от σ-алгебры $\mathscr{F}_k$. ◎

448| **Следствие.** *Пусть $\boldsymbol{\xi} = \{\xi_k,\ k \geqslant 1\}$ — последовательность независимых одинаково распределённых сл.в. с $\mathbf{E}\xi_1 = 0,\ \sigma^2 = \mathbf{D}\xi_1 < \infty$. Положим $S_n = \sum_1^n \xi_k$, тогда последовательность $\{S_n^2 - n\sigma^2,\ n \geqslant 1\}$ образует мартингал относительно естественной фильтрации $\mathcal{F}^{\boldsymbol{\xi}}$.*

При больших n поведение квадратично-интегрируемого мартингала в значительной степени определяется поведением его квадратической характеристики. Точнее, справедлива

449| **Теорема.** *Пусть $\boldsymbol{\xi} = \big(\xi_n, \mathscr{F}_n\big)$ — квадратично-интегрируемый мартингал, $\big(\big\langle \boldsymbol{\xi} \big\rangle_n, \mathscr{F}_n\big)$ — его квадратическая характеристика. Если $\mathbf{P}\big\{\lim_n \big\langle\, \xi\, \big\rangle_n = \infty\big\} = 1$, то почти наверное*

$$\lim_n \frac{\xi_n}{\big\langle \boldsymbol{\xi} \big\rangle_n} = 0.$$

✧ Случайная замена времени. Тождество Вальда. При наблюдениях за сл.проц. часто возникает вопрос о моменте прекращения наблюдений. Естественно, решение о прекращении наблюдений в какой-то момент не может зависеть от последующих наблюдений. При рассмотрении мартингалов естественно считать, что момент остановки (прекращения наблюдений) согласован с фильтрацией, определяющей этот мартингал. Таким образом, значение сл.процесса $\xi_n,\ n \geqslant 0$, в момент остановки τ можно представить в виде ряда:

$$\xi_\tau = \sum\nolimits_0^\infty \xi_n\, \dot{\mathrm{I}}\,(\tau = n).$$

Очевидно (?!), ξ_τ — случайная величина.

Представляют интерес вопросы о сохранении свойств мартингалами (субмартингалами) при случайной остановке. В частности, имеет ли место равенство $\mathbf{E}\xi_\tau = \mathbf{E}\xi_0$? Этот факт, называемый часто теоремой о свободной остановке (optional stopping theorem), всегда выполняется, если момент остановки ограничен. Обозначим $a \wedge b = \min\{a, b\}$.

<u>450</u>| **Лемма.** *Пусть* $\{\xi_n, \mathscr{F}_n\}_0^\infty$ *— мартингал (субмартингал) и* ν *— марковский момент относительно* $\{\mathscr{F}_n\}_0^\infty$. *Тогда* $\{\xi_{\nu\wedge n}, \mathscr{F}_n\}_0^\infty$ *— мартингал (субмартингал) и* $\mathbf{E}\,\xi_n =_{(\geqslant)} \mathbf{E}\left[\xi_{\nu\wedge n}\right] =_{(\geqslant)} \mathbf{E}\,\xi_0$ *для* $\forall\, n \geqslant 0$.

⤇ Заметим, что $\xi_{\nu\wedge(n+1)} = \sum_0^n \xi_k\,\dot{\mathrm{I}}\,(\nu = k) + \xi_{n+1}\,\dot{\mathrm{I}}\,(\nu > n)$. Так как события $\{\nu = k\} \in \mathscr{F}_k \subset \mathscr{F}_n$, $k \leqslant n$, и событие $\{\nu > n\} = \{\nu \leqslant n\}^c \in \mathscr{F}_n$, то

$$\mathbf{E}\big[\,\xi_{\nu\wedge(n+1)} \,|\, \mathscr{F}_n\big] \;=\; \sum_0^n \mathbf{E}\big[\,\xi_k\,\dot{\mathrm{I}}\,(\nu = k) \,|\, \mathscr{F}_n\big] + \mathbf{E}\big[\,\xi_{n+1}\,\dot{\mathrm{I}}\,(\nu > n) \,|\, \mathscr{F}_n\big] \;=$$
$$=\; \sum_0^n \xi_k\,\dot{\mathrm{I}}\,(\nu = k) + \mathbf{E}\big[\,\xi_{n+1} \,|\, \mathscr{F}_n\big]\,\dot{\mathrm{I}}\,(\nu > n) \;\underset{(\geqslant)}{=}\; \xi_{\nu\wedge n},$$

т.к. $\mathbf{E}\big[\,\xi_{n+1} \,|\, \mathscr{F}_n\big] \underset{(\geqslant)}{=} \xi_n$. Следовательно, $\mathbf{E}\,\xi_{\nu\wedge n} \underset{(\geqslant)}{=} \mathbf{E}\,\xi_{\nu\wedge 0} = \mathbf{E}\,\xi_0$. Докажем теперь неравенство $\mathbf{E}\,\xi_n \geqslant \mathbf{E}\left[\xi_{\nu\wedge n}\right]$ (актуальное для субмартингалов). Заметим, что для $k < n$

$$\mathbf{E}\left[\xi_n\,\dot{\mathrm{I}}\,(\nu = k)\right] \;=\; \mathbf{E}\left[\,\dot{\mathrm{I}}\,(\nu = k)\,\mathbf{E}\big[\,\xi_n \,|\, \mathscr{F}_k\,\big]\right] \;\geqslant\; \mathbf{E}\left[\,\dot{\mathrm{I}}\,(\nu = k)\xi_k\right].$$

Поэтому

$$\mathbf{E}\,\xi_n \;=\; \sum_{k=0}^{n-1} \mathbf{E}\left[\xi_n\,\dot{\mathrm{I}}\,(\nu = k)\right] + \mathbf{E}\left[\xi_n\,\dot{\mathrm{I}}\,(\nu \geqslant n)\right] \;\geqslant$$
$$\geqslant\; \sum_{k=0}^{n-1} \mathbf{E}\left[\xi_k\,\dot{\mathrm{I}}\,(\nu = k)\right] + \mathbf{E}\left[\xi_n\,\dot{\mathrm{I}}\,(\nu \geqslant n)\right] \;=\; \mathbf{E}\left[\xi_{\nu\wedge n}\right]. \qquad ⤆$$

В следующей теореме рассмотрен общий момент остановки.

<u>451</u>| **Теорема.** *Пусть* $\{\xi_n, \mathscr{F}_n\}_0^\infty$ *— мартингал (субмартингал) и* ν *— момент остановки относительно* $\{\mathscr{F}_n\}_0^\infty$. *Если:*

$$\text{i)}\ \mathbf{E}\,|\xi_\nu| < \infty, \qquad \text{ii)}\ \lim_n \mathbf{E}\left[\xi_n\,\dot{\mathrm{I}}_{\{\nu > n\}}\right] = 0,$$

то $\mathbf{E}\,\xi_\nu =_{(\geqslant)} \mathbf{E}\,\xi_0$.

⤇ Для $\forall\, n \geqslant 0$ имеем

$$\mathbf{E}\,\xi_\nu \;=\; \mathbf{E}\left[\xi_\nu\,\dot{\mathrm{I}}\,(\nu \leqslant n)\right] + \mathbf{E}\left[\xi_\nu\,\dot{\mathrm{I}}\,(\nu > n)\right] =$$
$$=\; \mathbf{E}\left[\xi_{\nu\wedge n}\right] - \mathbf{E}\left[\xi_{\nu\wedge n}\,\dot{\mathrm{I}}\,(\nu > n)\right] + \mathbf{E}\left[\xi_\nu\,\dot{\mathrm{I}}\,(\nu > n)\right].$$

По лемме <u>450</u> первое слагаемое равно (не меньше) $\mathbf{E}\,\xi_0$. Второе слагаемое $\mathbf{E}\left[\xi_{\nu\wedge n}\,\dot{\mathrm{I}}\,(\nu > n)\right] = \mathbf{E}\left[\xi_n\,\dot{\mathrm{I}}\,(\nu > n)\right] \underset{n}{\to} 0$ по условию.

Так как для моментов остановки вероятность $\mathbf{P}\{\nu > n\} \underset{n}{\to} 0$, то по теореме Лебега об ограниченной сходимости (напомним, что по условию $\mathbf{E}\,|\xi_\nu| < \infty$)

математическое ожидание $\mathbf{E}\left[\xi_\nu\,\dot{\mathrm{I}}(\nu > n)\right] \underset{n}{\to} 0$. ⫤

Приведём ещё два варианта условий, обеспечивающих это свойство.

452| **Лемма.** *Пусть $\{\xi_n, \mathscr{F}_n\}_0^\infty$ — мартингал и ν — момент остановки относительно $\{\mathscr{F}_n\}_0^\infty$. Если $|\xi_{\nu\wedge n}| \leqslant \eta$* (п.н.), *где $\mathbf{E}\eta < \infty$, то $\mathbf{E}\xi_\nu = \mathbf{E}\xi_0$.*

⫢ Легко видеть, что $|\xi_\nu| \leqslant \eta$ (п.н.), поэтому $\mathbf{E}|\xi_\nu| < \infty$ и

$$\left|\mathbf{E}\xi_{\nu\wedge n} - \mathbf{E}\xi_\nu\right| \;\leqslant\; \mathbf{E}\left[\left|\xi_n - \xi_\nu\right|\dot{\mathrm{I}}(\nu > n)\right] \;\leqslant\; 2\,\mathbf{E}\left[\eta\,\dot{\mathrm{I}}(\nu > n)\right] \underset{n}{\to} 0$$

по теореме Лебега об ограниченной сходимости. Равенство $\mathbf{E}\xi_{\nu\wedge n} = \mathbf{E}\xi_0$ следует из леммы 450. ⫤

453| **Лемма.** *Пусть $\{\xi_n, \mathscr{F}_n\}_0^\infty$ — мартингал и ν — момент остановки относительно $\{\mathscr{F}_n\}_0^\infty$. Кроме того,*

a) $\mathbf{E}\nu < \infty$,

b) $\mathbf{E}\big[\,|\xi_{n+1} - \xi_n|\,\big|\,\mathscr{F}_n\big] \leqslant C$ (п.н.), $\forall n \geqslant 0$,

с некоторой константой $C < \infty$. Тогда $\mathbf{E}|\xi_\nu| < \infty$ и $\mathbf{E}\xi_\nu = \mathbf{E}\xi_0$.

⫢ Определим $\eta = \sum_0^\nu \zeta_k$, $\zeta_k = |\xi_k - \xi_{k-1}|$, $\zeta_0 = 0$. Тогда, ввиду возможности замены порядка суммирования в ряду с положительными членами,

$$\begin{aligned}\mathbf{E}\eta = &\sum_{n=0}^{\infty}\sum_{k=0}^{n}\mathbf{E}\left[\zeta_k\,\dot{\mathrm{I}}(\nu = n)\right] \;=\; \sum_{k=0}^{\infty}\sum_{n=k}^{\infty}\mathbf{E}\left[\zeta_k\,\dot{\mathrm{I}}(\nu = n)\right] \;= \\ = &\sum_{k=0}^{\infty}\mathbf{E}\left[\zeta_k\,\dot{\mathrm{I}}(\nu \geqslant k)\right] \;=\; \sum_{k=0}^{\infty}\mathbf{E}\Big[\,\dot{\mathrm{I}}(\nu \geqslant k)\,\mathbf{E}\big[\,|\xi_k - \xi_{k-1}|\,\big|\,\mathscr{F}_{k-1}\big]\Big],\end{aligned}$$

где последнее равенство следует из того, что событие $\{\nu \geqslant k\} = \{\nu \leqslant k-1\}^c \in$ $\in \mathscr{F}_{k-1}$. Таким образом, в силу условия b)

$$\mathbf{E}\eta \;\leqslant\; C\sum\nolimits_0^\infty \mathbf{P}\{\nu \geqslant k\} \;=\; C(1 + \mathbf{E}\nu) \;<\; \infty$$

по известному свойству математического ожидания неотрицательных целочисленных сл.в. Так как по построению $|\xi_{\nu\wedge n}| \leqslant \eta$, то требуемое утверждение следует из предыдущей леммы. ⫤

454| **Теорема.** [*Тождество Вальда.*] *Пусть $\theta_1, \theta_2, \ldots$ — независимые одинаково распределённые сл.в. с $\mathbf{E}\theta_1 = \mu$ и $\mathbf{D}\theta_1 = \sigma^2 < \infty$. Если марковский момент имеет конечное математическое ожидание $\mathbf{E}\nu < \infty$, то для $S_n =$ $= \theta_1 + \ldots + \theta_n$ выполняются следующие равенства:*

i) $\mathbf{E}S_\nu = \mu\,\mathbf{E}\nu$, ii) $\mathbf{E}(S_\nu - \nu\mu)^2 = \sigma^2\,\mathbf{E}\nu$.

$\Rrightarrow$ i) Применим 453 к мартингалу $\xi_n = S_n - n\mu$ (относительно фильтрации $\mathscr{F}_n = \sigma(\theta_1, \ldots, \theta_n)$). Условие b) здесь выполняется, т.к. θ_{n+1} не зависит от $\mathscr{F}_n$, и потому

$$\mathbf{E}\big[\,|\xi_{n+1} - \xi_n|\,\big|\,\mathscr{F}_n\big] \;=\; \mathbf{E}\big[\,|\theta_{n+1} - \mu|\,\big|\,\mathscr{F}_n\big] \;\leqslant\; 2\,\mathbf{E}\,|\theta_1| \quad (\text{п.н.}).$$

По лемме 453 имеем $0 = \mathbf{E}\,\xi_1 = \mathbf{E}\,S_\nu - \mu\,\mathbf{E}\,\nu$.

ii) Можно считать $\mu = \mathbf{E}\,\theta_1 = 0$. В силу следствия 448 последовательность $S_k^2 - k\sigma^2$ образует мартингал, поэтому (с учётом леммы 450) $\mathbf{E}\left[S^2_{\nu\wedge n} - (\nu \wedge n)\sigma^2\right] = 0$. Отсюда по лемме Фату

$$\mathbf{E}\,S_\nu^2 \;=\; \mathbf{E}\left[\lim_n S^2_{\nu\wedge n}\right] \;\leqslant\; \varliminf_n \mathbf{E}\left[S^2_{\nu\wedge n}\right] \;=\; \sigma^2 \varliminf_n \mathbf{E}\left[\nu \wedge n\right] \;=\; \sigma^2\,\mathbf{E}\,\nu.$$

Обратно,

$$\begin{aligned}
\mathbf{E}\,S_\nu^2 \;&=\; \mathbf{E}\left[S_\nu^2\,\dot{\mathrm{I}}(\nu < n)\right] + \mathbf{E}\left[\,\mathbf{E}\big[S_\nu^2\,\dot{\mathrm{I}}(\nu \geqslant n)\,\big|\,\mathscr{F}_n\big]\right] \;\geqslant^{(1)} \\
&\geqslant\; \mathbf{E}\left[S^2_{\nu\wedge n}\,\dot{\mathrm{I}}(\nu < n)\right] + \mathbf{E}\left[\big(\mathbf{E}\big[\,\dot{\mathrm{I}}(\nu \geqslant n)|S_\nu|\,\big|\,\mathscr{F}_n\big]\big)^2\right] \;\geqslant^{(2)} \\
&\geqslant\; \mathbf{E}\left[S^2_{\nu\wedge n}\,\dot{\mathrm{I}}(\nu < n)\right] + \mathbf{E}\left[\,\dot{\mathrm{I}}(\nu \geqslant n) S_n^2\right] \;=\; \mathbf{E}\,S^2_{\nu\wedge n},
\end{aligned}$$

где неравенство «$\geqslant^{(1)}$» следует из неравенства Коши–Буняковского (для усл.м.о.), а неравенство «$\geqslant^{(2)}$» — из неравенства

$$(*) \qquad \mathbf{E}\big[\,\dot{\mathrm{I}}(\nu \geqslant n)|S_\nu|\,\big|\,\mathscr{F}_n\big] \;\geqslant\; |S_n|\,\dot{\mathrm{I}}(\nu \geqslant n), \quad n \geqslant 1.$$

Для установления последнего заметим сначала, что в силу неравенства треугольника $S_\nu \leqslant \sum_1^\nu |\theta_k| =: R_\nu$, поэтому

$$\mathbf{E}\,|S_\nu| \leqslant \mathbf{E}\,R_\nu \;=\; \mathbf{E}\,|\theta_1|\,\mathbf{E}\,\nu \;<\; \infty$$

по тождеству Вальда i) из настоящей теоремы. Отсюда

$$\mathbf{E}\left[|S_m|\,\dot{\mathrm{I}}(\nu > m)\right] \;\leqslant\; \mathbf{E}\left[R_m\,\dot{\mathrm{I}}(\nu > m)\right] \;\leqslant\; \mathbf{E}\left[R_\nu\,\dot{\mathrm{I}}(\nu > m)\right] \;\underset{m\to\infty}{\longrightarrow}\; 0.$$

Возьмём произвольное событие $B \in \mathscr{F}_n$. Для $\forall\, k \geqslant n$ событие $\{\nu > k\} \cap$ $\cap\, B \in \mathscr{F}_k$, поэтому усл.м.о.

$$\begin{aligned}
\mathbf{E}\big[\,|S_{k+1}|\,\dot{\mathrm{I}}(\nu > k)\,\dot{\mathrm{I}}(B)\,\big|\,\mathscr{F}_k\big] \;&=\; \dot{\mathrm{I}}(\nu > k)\,\dot{\mathrm{I}}(B)\,\mathbf{E}\big[\,|S_{k+1}|\,\big|\,\mathscr{F}_k\big] \;\geqslant \\
&\geqslant\; \dot{\mathrm{I}}(\nu > k)\,\dot{\mathrm{I}}(B)|S_k|,
\end{aligned}$$

т.к. $|S_n|$ есть субмартингал, поскольку при нулевом среднем $\mu = \mathbf{E}\,\theta_1$ суммы

S_n образуют мартингал. Поэтому для $\forall\, m \geqslant n$

$$\begin{aligned}
&\mathbf{E}\left[|S_n|\,\dot{\mathrm{I}}(\nu \geqslant n)\,\dot{\mathrm{I}}(B)\right] \;= \\
&=\; \mathbf{E}\left[|S_\nu|\,\dot{\mathrm{I}}(\nu = n)\,\dot{\mathrm{I}}(B)\right] + \mathbf{E}\left[|S_n|\,\dot{\mathrm{I}}(\nu > n)\,\dot{\mathrm{I}}(B)\right] \;\leqslant \\
&\leqslant\; \mathbf{E}\left[|S_\nu|\,\dot{\mathrm{I}}(\nu = n)\,\dot{\mathrm{I}}(B)\right] + \mathbf{E}\left[|S_{n+1}|\,\dot{\mathrm{I}}(\nu \geqslant n+1)\,\dot{\mathrm{I}}(B)\right] \;= \\
&=\; \mathbf{E}\left[|S_\nu|\,\dot{\mathrm{I}}(n \leqslant \nu \leqslant n+1)\,\dot{\mathrm{I}}(B)\right] + \mathbf{E}\left[|S_{n+1}|\,\dot{\mathrm{I}}(\nu > n+1)\,\dot{\mathrm{I}}(B)\right] \;\leqslant \\
&\leqslant \cdots \leqslant\; \mathbf{E}\left[|S_\nu|\,\dot{\mathrm{I}}(n \leqslant \nu \leqslant m)\,\dot{\mathrm{I}}(B)\right] + \mathbf{E}\left[|S_m|\,\dot{\mathrm{I}}(\nu > m)\,\dot{\mathrm{I}}(B)\right].
\end{aligned}$$

Устремляя $m \to \infty$, получаем неравенство $\mathbf{E}\left[|S_n|\,\dot{\mathrm{I}}(\nu \geqslant n)\,\dot{\mathrm{I}}(B)\right] \leqslant \mathbf{E}\left[|S_\nu|\,\dot{\mathrm{I}}(\nu \geqslant n)\,\dot{\mathrm{I}}(B)\right]$, которое эквивалентно (*) ввиду произвольности выбора $B \in \mathscr{F}_n$ и измеримости $|S_n|\,\dot{\mathrm{I}}(\nu \geqslant n)$ относительно $\mathscr{F}_n$. ⫤

455| Примеры. 1) Пусть $\left\langle \xi_j \backsim \mathcal{B}\mathtt{ern}(p) \right\rangle_1^\infty$ — последовательность независимых бернуллиевских сл.в. Рассмотрим момент достижения частной суммой $S_n = \sum_1^n \xi_j$ заданного (целого) уровня M: $\nu = \min\{n : \; S_n = M\}$. Сл.в. ν равна номеру того эксперимента в схеме испытаний Бернулли, при котором впервые произошло ровно M успехов. Другими словами, ν имеет распределение Паскаля с параметрами (M, p). Легко понять, что ν можно представить в виде $\nu = \eta_1 + \ldots + \eta_M$ с независимыми геометрическими сл.в. $\eta_i \backsim \mathcal{G}\mathtt{eo}(p)$. Поскольку $\mathbf{P}\{\eta_1 = k\} = p(1-p)^{k-1}$, то математическое ожидание

$$\mathbf{E}\,\eta_1 \;=\; \sum_{k=1}^{\infty} k\,p(1-p)^{k-1} \;=\; -p\Big(\sum_{k=1}^{\infty}(1-p)^k\Big)'_p \;=\; -p\Big(\frac{1}{p}\Big)' \;=\; \frac{1}{p}.$$

Следовательно, $\mathbf{E}\,\nu = M/p$, или в форме, эквивалентной тождеству Вальда: $M = \mathbf{E}\,S_\nu = \mathbf{E}\,\xi_1 \cdot \mathbf{E}\,\nu$. Легко видеть, что дисперсия геометрической сл.в. $\mathbf{D}\eta_1 = (1-p)/p^2$, поэтому $\mathbf{E}\,\nu^2 = M(1-p)/p^2 + M^2/p^2$. Стало быть, как и следует из тождества Вальда,

$$\mathbf{E}(S_\nu - \nu p)^2 \;=\; M^2 - 2Mp\frac{M}{p} + p^2\frac{M + M^2 - Mp}{p^2} \;=\; M(1-p) \;=\; \mathbf{D}\xi_1 \cdot \mathbf{E}\,\nu.$$

2) Рассмотрим игру, описанную в вводном примере на стр. 313. Суммарный доход игрока A после n проведённых игр равен $S_n = \theta_1 + \ldots + \theta_n$, где $\{\theta_k\}_1^\infty$ — последовательность независимых одинаково распределённых сл.в. с $\mathbf{P}\{\theta_1 = 1\} = p = 1 - \mathbf{P}\{\theta_1 = -1\}$. Определим момент остановки $\nu = \min\{n : S_n = b$ или $S_n = -a\}$, где a — начальный капитал A, b — началь-

ный капитал его соперника. Понятно, что если $S_\nu = b$, то последовательность игр заканчивается победой А (разорением Б). Обозначим вероятность этого через $p_{_A}$. В приложении к этой главе приведён вывод формулы для $p_{_A}$ с помощью рекуррентных соотношений. Сейчас же мы воспользуемся тождеством Вальда, предполагая, что $\mathbf{P}\{\nu<\infty\}=1$ и $\mathbf{E}\,\nu<\infty$, в частности $p_{_A}+p_{_Б}=1$ (доказательство см. в Приложении А).

Обсудим сначала вариант безобидной игры: $p={}^1\!/_2$. В этом случае $\mathbf{E}\,\theta_1 = 0$ и ввиду сохранения мартингального свойства

$$0 \;=\; \mathbf{E}\,S_\nu \;=\; bp_{_A} - a(1-p_{_A}),$$

т.е. $p_{_A} = \dfrac{a}{a+b}$. Поскольку дисперсия $\sigma^2=1$, то из тех же соображений, применённых к мартингалу S_n^2-n,

$$0 \;=\; \mathbf{E}(S_\nu^2-\nu) \;=\; b^2p_{_A}+a^2(1-p_{_A})-\mathbf{E}\,\nu \;=\; ab-\mathbf{E}\,\nu.$$

Стало быть, средняя продолжительность игры до разорения одного из игроков равна ab. Пусть теперь вероятность выигрыша А в одной игре $p<{}^1\!/_2$, т.е. отношение шансов $\lambda=(1-p)/p>1$. В этом случае $\mu=\mathbf{E}\,\theta_1=2p-1<0$. Рассмотрим неотрицательные сл.в. $\zeta_k=\lambda^{\theta_k}$. Очевидно, $\mathbf{E}\,\zeta_k=1$. Поэтому для мартингала $\xi_n=\prod_1^n\zeta_k=\lambda^{S_n}$ справедливы равенства

$$1 \;=\; \mathbf{E}\,\zeta_1 \;=\; \mathbf{E}\,\xi_\nu \;=\; p_{_A}\lambda^b+(1-p_{_A})\,\lambda^{-a}.$$

Отсюда вероятность победы А равна $p_{_A}=(\lambda^a-1)/(\lambda^{a+b}-1)$.

Среднее время игры найдём из тождества Вальда:

$$(2p-1)\,\mathbf{E}\,\nu \;=\; \mathbf{E}\,S_\nu \;=\; bp_{_A}-a(1-p_{_A}).$$

Поэтому $\mathbf{E}\,\nu=(a-(a+b)p_{_A})/(1-2p)$. ◎

✧ **Максимальные неравенства для субмартингалов.** Неравенство Колмогорова для сумм независимых сл.в. с нулевым средним легко переносится на случай субмартингалов.

456 **Теорема.** *Пусть* $\{\xi_n,\mathscr{F}_n\}_0^\infty$ *— неотрицательный субмартингал,* $\overline{\xi}_n=\max\limits_{0\leqslant k\leqslant n}\xi_k$, *тогда для* $\forall\,A>0$, $p>1$ *и* $\forall\,n=0,1,\ldots$

i) $$\mathbf{P}\Big\{\max_{0\leqslant k\leqslant n}\xi_k>A\Big\} \;\leqslant\; \frac{1}{A}\,\mathbf{E}\left[\xi_n\,\mathrm{I}(\overline{\xi}_n>A)\right] \;\leqslant\; \frac{1}{A}\,\mathbf{E}\,\xi_n;$$

ii) $$\|\overline{\xi}_n\|_p \;\leqslant\; \frac{p}{p-1}\|\xi_n\|_p,$$

где $\mathcal{L}_p$-норма $\|\eta\|_p = \sqrt[p]{\mathbf{E}\,|\eta|^p}$.

$\Rrightarrow$ i) Пусть марковский момент $\nu = \min\{k \geqslant 0:\ \xi_k > A\}$. Тогда по лемме 450 в силу неотрицательности ξ_n

$$\mathbf{E}\,\xi_n \ \geqslant\ \mathbf{E}\left[\xi_{\nu\wedge n}\right] \ \geqslant\ \mathbf{E}\left[\xi_{\nu\wedge n}\,\dot{\mathrm{I}}\,(\overline{\xi}_n > A)\right] + \mathbf{E}\left[\xi_{\nu\wedge n}\,\dot{\mathrm{I}}\,(\overline{\xi}_n \leqslant A)\right] \ \geqslant$$
$$\geqslant\ A\,\mathbf{P}\big\{\,\overline{\xi}_n > A\,\big\} + \mathbf{E}\left[\xi_n\,\dot{\mathrm{I}}\,(\overline{\xi}_n \leqslant A)\right],$$

т.к. $\xi_{\nu\wedge n} = \xi_\nu > A$ в области $\{\overline{\xi}_n > A\}$. Требуемое неравенство следует отсюда простыми арифметическими преобразованиями.

ii) Напомним, что для неотрицательной сл.в. ввиду i)

$$\mathbf{E}\,\overline{\xi}_n^{\,p} \ =\ p\int_0^\infty x^{p-1}\,\mathbf{P}\big\{\,\overline{\xi}_n > x\,\big\}\,dx \ \overset{\text{i)}}{\leqslant}\ p\int_0^\infty x^{p-2}\,\mathbf{E}\left[\xi_n\,\dot{\mathrm{I}}\,(\overline{\xi}_n > x)\right]dx.$$

Для неотрицательных функций по теореме Фубини можно всегда произвести замену порядка интегрирования. Так как $\int_0^\infty x^{p-2}\,\dot{\mathrm{I}}\,(a > x)\,dx = a^{p-1}/(p-1)$, то после применения неравенства Гёльдера (551, стр. 487) с $q = p/(p-1)$ получаем

$$\mathbf{E}\,\overline{\xi}_n^{\,p} \ \leqslant\ \frac{p}{p-1}\,\mathbf{E}\left[\xi_n\,\overline{\xi}_n^{\,p-1}\right] \ \leqslant\ \frac{p}{p-1}\,\|\xi_n\|_p\,\|\overline{\xi}_n^{\,p-1}\|_q \ =\ \frac{p}{p-1}\,\|\xi_n\|_p\,\sqrt[q]{\mathbf{E}\,\overline{\xi}_n^{\,p}}.$$

Если $\mathbf{E}\,\overline{\xi}_n^{\,p} < \infty$, то отсюда вытекает требуемое неравенство. Если же данное условие не выполняется, то, применяя метод усечения (к сл.в. $\overline{\xi}_n$), можно показать, что и $\mathbf{E}\,\xi_n^p = \infty$. $\Lleftarrow$

✧ **Теоремы сходимости для субмартингалов**. Приведём без доказательства несколько результатов, касающихся вопросов сходимости последовательностей, образующих мартингалы (субмартингалы).

457| Теорема. [*Дуб.*] *Пусть $\{\xi_n, \mathscr{F}_n\}_0^\infty$ — субмартингал с ограниченным математическим ожиданием:* $\sup_n \mathbf{E}\,|\xi_n| < \infty$. *Тогда существует сл.в. ξ_∞ такая, что $\mathbf{E}\,|\xi_\infty| < \infty$ и $\xi_n \xrightarrow{\text{п.н.}} \xi_\infty$.*

Если субмартингал $\{\xi_n\}_0^\infty$ равномерно интегрируем, то имеет место сходимость в $\mathcal{L}_1$: $\xi_n \xrightarrow{\mathcal{L}_1} \xi_\infty$. Более того, $\mathbf{E}\big[\,\xi_\infty \,\big|\, \mathscr{F}_n\,\big] \geqslant \xi_n$ со знаком равенства для мартингалов.

[*П. Леви.*] *Пусть $\mathbf{E}\,|\eta| < \infty$, тогда для любого потока σ-алгебр $\mathcal{F} = \{\mathscr{F}_n\}_0^\infty$ имеет место сходимость (п.н. и в смысле $\mathcal{L}_1$) мартингала Леви:*

$$\mathbf{E}\big[\,\eta \,\big|\, \mathscr{F}_n\,\big] \to \mathbf{E}\big[\,\eta \,\big|\, \mathscr{F}_\infty\,\big],$$

где $\mathscr{F}_\infty = \sigma\big(\cup_n \mathscr{F}_n\big)$ *— минимальная* σ*-алгебра, содержащая все составляющие потока* $\mathcal{F}$*.*

✧ **Мартингалы с непрерывным временем.** Самым ярким представителем мартингальных процессов с непрерывным временем $t \in \mathbb{T} = [0;\infty)$ является стандартный винеровский процесс. Как и в дискретном случае, для определения мартингала (субмартингала) сначала на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbf{P})$ вводится семейство σ-алгебр $\mathcal{F} = \big\langle \mathscr{F}_t,\ t \in \mathbb{T} \big\rangle$, образующих фильтрацию:

$$\mathscr{F}_s \subset \mathscr{F}_t \subset \mathscr{F} \quad \text{при } s < t, \quad s, t \in \mathbb{T}.$$

Фильтрация здесь должна быть расширенной, т.е. все входящие в $\mathcal{F}$ σ-алгебры должны быть полными и, более того, должны содержать все нулевые подмножества $\mathscr{F}$. Кроме того, поток σ-алгебр должен быть непрерывным справа: $\mathscr{F}_{t+} = \bigcap_{\varepsilon>0} \mathscr{F}_{t+\varepsilon} = \mathscr{F}_t$. Последнее условие обеспечивает измеримость событий вида $\{\nu < t\}$ относительно σ-алгебры $\mathscr{F}_t$ для моментов ν выхода за границы множеств достаточно общего вида.

Определение. Процесс $\boldsymbol{\xi} = \big\langle \xi_t,\ t \in [0;\infty) \big\rangle$, согласованный с фильтрацией $\mathcal{F}$, называется мартингалом (субмартингалом) относительно $\mathcal{F}$ (обозначается $(\boldsymbol{\xi}, \mathcal{F})$), если

i) $\mathbf{E}\,|\xi_t| < \infty$;

ii) $\mathbf{E}\big[\,\xi_t \,\big|\, \mathscr{F}_s\big] =_{(\geqslant)} \xi_s$ (п.н.), $\forall\, s < t,\ s, t \in \mathbb{T}$.

Теорема о свободной остановке для мартингалов формулируется следующим образом.

458| **Теорема.** *Пусть* $(\boldsymbol{\xi}, \mathcal{F})$ *— мартингал с непрерывными справа траекториями,* $\nu_1 \leqslant \nu_2$ *— опциональные моменты относительно* $\mathcal{F}$ *(т.е. события* $\{\nu_j < t\} \in \mathscr{F}_t, \forall\, t \geqslant 0,\ j = 1, 2$*). Если* $\nu_2 \leqslant c$*, где* c *— некоторая константа, то* $\mathbf{E}\,\xi_{\nu_1} \leqslant \mathbf{E}\,\xi_{\nu_2}$*.*

459| Пример. Наиболее часто теорема о свободной остановке применяется к винеровскому процессу. В частности, с помощью этой техники (теорема 403, стр. 361) мы нашли распределение и математическое ожидание для момента достижения траекторией процесса заданного уровня. ◎

Дополнения

А. Задача о разорении игрока

Найдём вероятность разорения игрока А из вводного примера (стр. 313). Пусть $p \in (0;1)$ — вероятность выигрыша игрока А в одной игре. Воспользуемся следующими рассуждениями. Если при $a, b > 0$

$$\{-a < X_1 < b, \quad -a < X_1 + X_2 < b\}, \quad \text{то}$$

$$-a - b \ < \ -a - X_1 \ < \ X_2 \ < \ b - X_1 \ < \ b + a,$$

т.е. $|X_j| < a + b, \ \forall\, j = 1, 2.$

Пусть $n = km$, определим $S_t^u = \sum_{j=t}^{u} \xi_j$, где $\mathbf{P}\{\xi_j = 1\} = p = 1 - \mathbf{P}\{\xi_j = -1\}$. Тогда в силу аналогичных соображений

$$\gamma_n := \mathbf{P}\Big\{-a < S_1^j < b, \quad \forall\, j \leqslant n\Big\} \ \leqslant \ \mathbf{P}\Big\{\bigcap_{l=1}^{k} \big(\big|S_{(l-1)m+1}^{lm}\big| < a + b\big)\Big\}.$$

Ввиду независимости и одинаковой распределённости $\xi_1, \xi_2, \ldots$

$$\gamma_n \leqslant \big(\mathbf{P}\big\{\big|S_1^m\big| < a + b\big\}\big)^k.$$

Всегда можно подобрать m так, чтобы $\mathbf{P}\big\{\big|S_1^m\big| < a + b\big\} = \varepsilon < 1$, поскольку в противном случае дисперсия $\mathbf{D} S_1^m = mp(1-p) < (a+b)^2$ для $\forall\, m > 1$, что невозможно для $p \in (0;1)$. Таким образом, для $\varepsilon_1 = \sqrt[m]{\varepsilon}$ при достаточно больших n вероятность того, что все частичные суммы до n включительно находятся в интервале $(-a; b)$

$$\gamma_n \leqslant \varepsilon_1^n \underset{n}{\to} 0.$$

Другими словами, с вероятностью единица игра завершится разорением одного из игроков. Более того, если обозначить через ν момент наступления такого события, то математическое ожидание $\mathbf{E}\,\nu = \sum_0^\infty \mathbf{P}\{\nu > n\} = \sum_0^\infty \gamma_n < \infty$.

Пусть D — событие пространства Ω_∞, означающее разорение игрока А. Ясно, что на каждом этапе бесконечной игры вероятность D зависит только от накопленного игроком А капитала k и не зависит от конкретных результатов уже проведённых игр (если, конечно, к этому этапу один из игроков не разорится), т.е. каждую следующую игру можно рассматривать как первую с новым состоянием кошелька А. Пусть $\pi(k)$ — вероятность разорения А, если его капитал равен k ($\in [0; R]$). Очевидно, $\pi(0) = 1$, $\pi(R) = 0$. В случае, когда $0 < k < R$, следующая игра может закончиться либо победой А, и тогда его капитал станет равным $k+1$, либо его поражением — капитал А уменьшится до $k-1$. Из этих соображений по формуле полной вероятности следует, что вероятность $\pi(k)$ удовлетворяет разностному уравнению

$$\pi(k) = \pi(k+1)p + \pi(k-1)(1-p). \tag{51}$$

При $p = {}^1\!/_2$ это уравнение имеет два линейно независимых частных решения $\pi(k) \equiv 1$ и $\pi(k) = k$. Отсюда общее решение находим в виде их линейной комбинации:

$$\pi(k) = c_1 + c_2 k.$$

Подставляя сюда «краевые» условия $\pi(0) = 1$, $\pi(R) = 0$, получаем, что вероятность разорения игрока А при начальном его капитале a и капитале противника b

$$\pi(a) = \frac{b}{a+b}\,. \tag{52}$$

В случае $p \neq {}^1\!/_2$ будем искать частные решения (51) в виде $\pi(k) = x^k$. После подстановки получаем квадратное уравнение: $px^2 - x + (1-p) = 0$. Таким образом, общее решение (51) есть линейная комбинация двух линейно независимых частных решений с $x = 1$ и $x = (1-p)/p$:

$$\pi(k) = c_1 + c_2\lambda^k, \qquad \lambda = \frac{(1-p)}{p}.$$

С учётом краевых условий получаем вероятность разорения А:

$$\pi(a) \;=\; \pi(a \,|\, b) \;=\; \frac{\lambda^{a+b} - \lambda^a}{\lambda^{a+b} - 1}.$$

При игре А против казино капитал игрока В следует воспринимать как количество условных единиц (количество ставок), которыми игрок А планирует пополнить свой кошелёк; по достижении этой цели игрок А прекращает игру. Приведём некоторые значения вероятности $1 - \pi(a \,|\, b)$ окончательного успеха А при игре в «красное–чёрное» в рулетку с 18 красными, 18 чёрными ячейками и одной ячейкой «зеро». Другими словами, вероятность победы А в одной игре равна ${}^{18}\!/_{37}$, т.е. отношение шансов казино против А равно

$\lambda = {}^{19}/_{18}\ (> 1)$.

(✓) Если ставка в одной игре составляет малую часть капитала A, т.е. a достаточно велико, то вероятность $1 - \pi(a \mid b)$ в большей степени зависит от уровня претензий b, чем от начального капитала a. Для $a > 40$ вероятность $1-\pi(a \mid b)$ при фиксированном b почти не изменяется относительно a и равна приблизительно λ^{-b}.

(✓) Уменьшение ставки (т.е. увеличение a и пропорциональное увеличение b) приводит к увеличению времени «наслаждения» игрой, но, увы, к уменьшению вероятности окончательной победы A. Для увеличения вероятности выигрыша следует выбрать максимальную ставку, сообразную с претензиями A. Если игрок мечтает удвоить свой капитал, то не надо растягивать удовольствие — следует сразу поставить весь капитал на выбранный цвет; если A мечтает учетверить свой капитал, то лучше всего два раза сыграть «ва-банк», если, конечно, повезёт при первой игре. Например, если игрок A, заходя в игру с 10 руб., мечтает выйти, имея 40 руб., то при игре «ва-банк» вероятность осуществления мечты равна $({}^{18}/_{37})^2 \approx 0.237$, при постоянной ставке в 10 руб. эта вероятность равна 0.230, а при ставке в 1 руб. — 0.093.

(✓) Приблизительно с одинаковой вероятностью можно увеличить свой капитал в 36 раз, сыграв один раз «ва-банк» на какое-то конкретное число ($\mathbf{P} = 0.027$), или увеличить в 32 раза капитал, сыграв 5 раз «ва-банк» на какой-либо цвет ($\mathbf{P} = 0.031$).

(✓) Стратегия игры «Мартингейл» формализует высказанное выше утверждение о том, что наибольшая вероятность осуществления цели достигается при ставке, максимально приближенной к претензиям игрока A. Если он поставил целью прибавить к своему капиталу 1 у.е., то при первом подходе к столу ставка должна равняться 1 у.е. В случае неудачи претензии A увеличиваются до 2 у.е., поэтому при втором подходе к столу ставка должна составлять 2 у.е. Таким образом, если капитал игрока A равен $2^k - 1$ у.е., то следует, начиная с 1 у.е., последовательно удваивая ставку до 2 у.е., 4 у.е., 8 у.е. ..., дождаться первой победной игры и, получив дополнительную 1 у.е. к своему первоначальному капиталу, закончить игру. Разорение A наступит только тогда, когда произойдут подряд k неудач — вероятность этого, очевидно, равна $(1-p)^k$. Интересно, что при игре с равными шансами ($p = {}^1/_2$) эта стратегия имеет ту же вероятность (?!) выигрыша одной у.е., что и стратегия, при которой ставка всегда равна 1 у.е.

В ситуации со ставками на «красное–чёрное» в рулетку ($p = {}^{18}/_{37}$) при начальном капитале 4095 у.е. ($= 2^{12} - 1$) вероятность выигрыша 1 у.е. с помощью «Мартингейла» равна 0.99966. Стратегия, при которой каждая ставка равна 1 у.е. с тем же начальным капиталом, как отмечалось выше, имеет вероятность выигрыша одной у.е., равную $1 - \pi(4095 \,|\, 1) \approx {}^{1}/_{\lambda^1} \approx 0.9474$. В связи с этим отметим, во-первых, что, конечно, стратегия «Мартингейл» повышает шансы на успех, однако только большой пессимист будет считать приведённые выше вероятности сильно различающимися. Во-вторых, кажущаяся «бесконечная близость» к единице вероятности победы при использовании стратегии «Мартингейл» таит в себе небольшое, если не сказать большое, лукавство. Трудно представить себе игрока, пришедшего в казино и ставящего целью добавить к наличному капиталу в 4095 руб. всего один рубль. Понятно, что амбиции любого игрока «чуть» выше. Количество у.е. уменьшается пропорционально этим амбициям. Так, если при тех же условиях игрок мечтает заполучить 585 руб., то его капитал равняется 7 у.е. В этом случае стратегия «Мартингейл» имеет 86.5% на успех. Стратегия, при которой каждая ставка равна 585, приводит к успеху с вероятностью 0.85. Что называется, оцените разницу. Если казино раздаёт бесплатные напитки игрокам, то вторая стратегия представляется предпочтительнее, поскольку имеет большее среднее значение объёма выпитого сока ☺. При повышении ставки доходы казино возрастают. Так, средний доход казино с игроков, удовлетворяющихся 1 рублём сверх начальных 4095 руб., составляет всего $4095 \cdot 0.000336 - 1 \cdot 0.999664 \approx 0.38$ руб., а с игроков, мечтающих в этой ситуации разбогатеть на 585 руб., — почти 50 рублей. Чем выше амбиции игрока, тем больше на нём заработает казино. Аналогичные выводы справедливы и для $p = {}^{1}/_{2}$.

В. Упражнения

Упр. В.1. Доказать стационарность в широком смысле сл.проц. $\xi_k = \cos\big(k\eta + \upsilon\big)$, $k = 0, 1, \ldots$, где сл.в. η, υ независимы и $\upsilon \backsim \mathcal{U}\mathrm{n}(0, 2\pi)$, η — произвольная сл.в.

Упр. В.2. Найти математическое ожидание стационарной последовательности авторегрессии второго порядка $\xi_n = 0.7\xi_{n-1} - 0.1\xi_{n-2} + \varepsilon_n$, если $\mathbf{E}\varepsilon_n = 2$.

Упр. В.3. Найти представление в виде скользящего среднего для стационарной последовательности авторегрессии $\xi_n = \xi_{n-1} - \frac{1}{4}\,\xi_{n-2} + \varepsilon_n$.

Упр. B.4. Показать, что последовательность $\xi_n = \frac{1}{2}(\varepsilon_{n-1}+\varepsilon_{n+1})$, образованная с помощью стандартного белого шума $\boldsymbol{\varepsilon}$, стационарна и её спектральная плотность $g(\theta) = \frac{1}{2\pi}\cos^2(\theta)$.

Упр. B.5. Показать, что пуассоновский процесс $\boldsymbol{\xi} = \left\langle \xi(t) \right\rangle_{t\geqslant 0}$ является субмартингалом относительно естественной фильтрации $\mathscr{F}^{\boldsymbol{\xi}}$; найти представление $\xi(t) = \eta(t) + A(t)$, где $\left\langle \eta_t, \mathscr{F}^{\boldsymbol{\xi}} \right\rangle$ — мартингал, $A(t)$ — детерминированная функция.

Упр. B.6. Пусть $w(t)$ — стандартный винеровский процесс. Показать, что процесс $\xi(t) = \exp\big(w(t) - \frac{1}{2}t\big) - 1$ имеет ортогональные приращения и, кроме того, $\mathbf{E}\xi(t) = 0$, $\mathbf{D}\xi(t) = 1$ для $\forall t$.

Упр. B.7. Пусть $\boldsymbol{\xi} = \big(\xi(t),\, t \geqslant 0\big)$ — процесс с независимыми приращениями, $\mathbf{E}\xi(t) = 0$, $t \geqslant 0$. Показать, что $\boldsymbol{\xi}$ есть мартингал относительно естественной фильтрации $\mathscr{F}^{\boldsymbol{\xi}}$. Установить справедливость разложения Дуба (в случае квадратичной интегрируемости $\boldsymbol{\xi}$) для субмартингала $\big(\xi^2(t),\, t \geqslant 0\big)$:

$$\xi^2(t) = m(t) + \left\langle \boldsymbol{\eta} \right\rangle_t,$$

где $\big(m(t), \mathscr{F}^{\boldsymbol{\xi}}\big)$ — мартингал, $\left\langle \boldsymbol{\eta} \right\rangle_t = \mathbf{D}\xi(t)$ — компенсатор ξ^2 (квадратическая характеристика $\boldsymbol{\xi}$).

C. Указания к решению задач

✦ 362, стр. 316. Пусть $B\times\mathscr{X}^{\mathbb{T}\setminus\tau}$, $B\in\mathscr{B}^k$, — произвольный цилиндр с набором $\tau = (t_1,\ldots,t_k)\subset\mathbb{T}$. Так как по условию сл.вектор $\vec{\zeta} = (\xi_{t_1},\ldots,\xi_{t_k})$ измерим относительно $\mathscr{B}^k$, то прообраз $\vec{\zeta}^{-1}(B)\in\mathscr{F}$, поэтому и $\boldsymbol{\xi}^{-1}(B\times\mathscr{X}^{\mathbb{T}\setminus\tau})\in\mathscr{F}$. Другими словами, прообраз любого конечномерного цилиндрического множества $\mathscr{F}$-измерим, что по теореме 507, стр. 464, гарантирует $\boldsymbol{\xi}^{-1}(\mathscr{B}^{\mathbb{T}})\subset\mathscr{F}$.

✦ 363, стр. 318. Так как ϑ измеримо, то и $\eta = (\omega,\vartheta(\omega)) : (\Omega,\mathscr{F}) \mapsto (\Omega\times\mathbb{T}, \mathscr{F}\otimes\mathscr{B})$ измеримо; сл.в. ξ_ϑ есть суперпозиция измеримых отображений $\xi_\vartheta(\omega) = \xi(\omega,t)\circ\eta$.

✦ 367, стр. 322. Распределение разности независимых сл.в. $\xi_u - \xi_t$ одинаково для любых u, t; эта разность может равняться нулю т. т. т. когда $\xi_u = \xi_t$ (п.н.).

✦ 368, стр. 322. Закрашенная полоса между $\pm 1 + \cos(2\pi t)$, $t\in[0;1]$.

✦ 375, стр. 327. Если $U_t = \{\xi(t)\neq\eta(t)\}$, то $\mathbf{P}\{U_t\} = 0$ и в силу счётности $\mathbb{T}$ вероятность неотличимости $\mathbf{P}\{\xi(t) = \eta(t),\ \forall t\in\mathbb{T}\} = 1 - \mathbf{P}\big\{\bigcup_{t\in\mathbb{T}} U_t\big\} = 1$.

✦ 376, стр. 327. Пусть T — счётное всюду плотное в $\mathbb{T}$ подмножество. По условию найдётся подмножество $U\subset\Omega$ такое, что $\mathbf{P}\{U\} = 1$ и $\xi_t(\omega) = \eta_t(\omega)$ для $\forall t\in T$, $\omega\in U$. В силу непрерывности траекторий и плотности подмножества T равенство $\xi_t(\omega) = \eta_t(\omega)$

будет выполняться для $\forall t \in \mathbb{T}$.

✦ 389 , стр. 344. $\mathbf{E}\left[\xi(s)-\xi(t)\right]^2 = \mathbf{D}\xi(s)+\mathbf{D}\xi(t)-2\,\mathbf{Cov}(\xi(s),\xi(t)) = 2\big(\mathbf{D}\xi(s)-\mathcal{R}(t-s)\big) \to$ $\to 0$, т.к. в силу непрерывности в нуле $\lim_{s\to t}\mathcal{R}(t-s) = \mathcal{R}(0) = \mathbf{D}\xi(s)$.

✦ 393 , стр. 347. $(\Rightarrow)$ Пусть $\forall 0 \leqslant s \leqslant t \leqslant u \leqslant v$, тогда $\mathbf{Cov}(\xi(v)-\xi(u),\xi(t)-$ $-\ \xi(s)) = \mathbb{V}(t)-\mathbb{V}(t)-\mathbb{V}(s)+\mathbb{V}(s) = 0$. Обратно, пусть $\mathbb{V}(t) = \mathbf{D}\xi(t)$, тогда $0 < \mathbf{E}\big[\xi(s)-$ $-\ \xi(t)\big]^2 = \mathbb{V}(s)+\mathbb{V}(t)-2\,\mathbf{E}\big[\xi(s)\big(\xi(t)-\xi(s)\big)\big]-2\mathbb{V}(s) = \mathbb{V}(t)-\mathbb{V}(s)$. Ковариация для $s \leqslant t$ $\mathbf{Cov}\,\big(\xi(s),\xi(t)\big) = \mathbf{E}\big[\xi(s)\big(\xi(t)-\xi(s)\big)\big]+\mathbb{V}(s) = \mathbb{V}(s)$.

✦ (?!) к (22), стр. 355. Если $\eta \backsim \mathcal{N}(0,1)$, то $\eta^2 \backsim \mathcal{K}\mathtt{hi}(1)$ и (см. стр. 167) $\mathbf{D}\eta^2 = 2$.

✦ (?!), стр. 356. Это гауссовский процесс с непрерывными траекториями, выходящий из нуля; ковариация $\mathbf{Cov}(\sqrt{c}w(s/c),\sqrt{c}w(t/c)) = c\min(s/c,t/c) = \min(s,t)$.

✦ 405 , стр. 363. Сечения, как линейные комбинации нормальных сл.в., имеют нормальное распределение. Математическое ожидание, очевидно, равно нулю. Ковариация для $s \leqslant t$ равна $\mathbf{E}\big[(w(t)-tw(1))(w(s)-sw(1)\big] = s-ts-st+ts1 = s(1-t)$.

✦ 428 , стр. 387. Путь F — событие из прошлого, R — событие из настоящего, B_1, B_2 — два события из будущего (B_1 — по времени раньше). Тогда $\mathbf{P}\big\{B_1B_2 \,\big|\, RF\big\} =$ $= \mathbf{P}\big\{B_2 \,\big|\, B_1RF\big\}\,\mathbf{P}\big\{B_1 \,\big|\, RF\big\} = \mathbf{P}\big\{B_2 \,\big|\, B_1\big\}\,\mathbf{P}\big\{B_1 \,\big|\, R\big\} = \mathbf{P}\big\{B_1B_2 \,\big|\, R\big\}$ по условию.

✦ (?!) к 435 , стр. 392. Для $j < k < n$ и совпадающей чётности разностей $f-r, n-$ $-\ k$ вероятность $\mathbf{P}\big\{\xi_n = f \,\big|\, \xi_k = r, \xi_j = p\big\} = \mathrm{C}_{n-k}^{f-r}\, p^{f-r}(1-p)^{n-k-f+r}$, если $f-r > 0$ (при отрицательной разности изменения очевидны, при различающейся чётности указанных разностей вероятность равна нулю). Таким образом, переходная вероятность зависит только от последнего момента времени (марковость) через разность значений (однородность).

✦ (?!), стр. 409. Если $\tau(\omega) = n < \infty$, то $\xi_\tau = \xi_n$, если же $\tau(\omega) = \infty$, то $\xi_\tau = 0$. Поэтому ряд, определяющий сл.в. ξ_τ, сходится при всех $\omega \in \Omega$, стало быть, этот ряд измерим, т.к. является пределом измеримых сл.в.

✦ (?!), стр. 418. Если игрок мечтает выиграть 1 у.е. и имеет в наличии 2^k-1 у.е., то по формуле (52), стр. 417, с $b = 1, a = 2^k-1$ вероятность разорения равна 2^{-k} .

✦ B1. Заметить, что для любых фиксированных k, n, y математическое ожидание (относительно $v \backsim \mathcal{U}\mathtt{n}(0,2\pi)$) $\mathbf{E}\big[\cos\big(ky+v\big)\big] = 0$ и $\mathbf{E}\big[\cos\big(ky+v\big)\cos\big(ny+v\big)\big] =$ $= \frac{1}{2}\cos\big((k-n)y\big)+\frac{1}{2}\,\mathbf{E}\big[\cos\big(ky+ny+2v\big)\big] = \frac{1}{2}\cos\big((k-n)y\big)$. То есть математическое ожидание процесса равно нулю, а ковариация зависит от разности моментов времени.

✦ B2. Стационарная последовательность существует, т.к. уравнение $1-0.7\gamma+0.1\gamma^2 = 0$ имеет корни $2, 5\ (> 1)$. Поэтому для $\mu = \mathbf{E}\xi_k : \mu = 0.7\mu-0.1\mu+2$, т.е. $\mu = 5$.

✦ B3. Уравнение $1-\gamma+\frac{1}{4}\gamma^2 = 0$ имеет два совпадающих корня $\gamma = 2$. Поэтому

$$\frac{1}{1-\gamma+\frac{1}{4}\gamma^2} = \frac{1}{(1-\frac{1}{2}\gamma)^2} = \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\frac{\gamma^{k+m}}{2^{k+m}} = \sum_{j=0}^{\infty}\frac{j+1}{2^j}\gamma^j.$$

Следовательно, стационарный процесс авторегрессии $\xi_n = \sum_{k=0}^{\infty}(k+1)2^{-k}\varepsilon_{n-k}$.

✦ B4. ξ_n и ξ_k независимы, если $|n-k| \neq 2$. Поэтому $\mathcal{R}(0) = \frac{1}{2}$, $\mathcal{R}(\pm 2) = \frac{1}{4}$, $\mathcal{R}(j) = 0$ для $j \neq 0, \neq \pm 2$. По формуле (8), стр. 334, спектральная плотность

$$g(\theta) = \frac{1}{2\pi}\left(e^{\mathrm{i}2\theta}\,\frac{1}{4} + e^{-\mathrm{i}2\theta}\,\frac{1}{4} + e^{\mathrm{i}0\theta}\,\frac{1}{2}\right) = \frac{1}{4\pi}\big(1+\cos(2\theta)\big) = \frac{1}{2\pi}\cos^2(\theta).$$

✦ B5. В силу независимости $\xi(t) - \xi(s)$ от σ-алгебры $\mathscr{F}_s^{\xi}$ усл.м.о. $\mathbf{E}\big[\,\xi(t)\,\big|\,\mathscr{F}_s^{\xi}\,\big] = \mathbf{E}\big[\xi(t) - \xi(s)\big] + \xi(s) = \lambda(t-s) + \xi(s) > \xi(s)$ для $t > s$. Отсюда понятно, что $\eta(t) = \xi(t) - \lambda t$ есть мартингал и $\xi(t) = \eta(t) + \lambda t$.

✦ B6. Заметить, что для $\eta \backsim \mathcal{N}(0, \sigma^2)$ математическое ожидание $\mathbf{E}\, e^{\eta} = e^{\sigma^2/2}$.

✦ B7. В силу независимости приращений для $s < t$ имеем: i) $\mathbf{E}\big[\,\xi_t\,\big|\,\mathscr{F}_s\,\big] = \mathbf{E}\big[\,\xi_t - \xi_s + \xi_s\,\big|\,\mathscr{F}_s\,\big] = \mathbf{E}\,\big(\xi_t - \xi_s\big) + \xi_s = \xi_s$; ii) $\mathbf{E}\,\big(\xi_t - \xi_s\big)^2 = \mathbf{E}\,\big[\xi_t^2 - \xi_s^2 - 2(\xi_t - \xi_s)\xi_s\big] = \mathbf{D}\xi_t - \mathbf{D}\xi_s$; iii) $\mathbf{E}\big[\,\xi_t^2 - \xi_s^2\,\big|\,\mathscr{F}_s\,\big] = \mathbf{E}\big[\,(\xi_t - \xi_s)(\xi_t - \xi_s + 2\xi_s)\,\big|\,\mathscr{F}_s\,\big] = \mathbf{E}\,\big(\xi_t - \xi_s\big)^2$.

IX Математические основания. Пространства с мерой. Интеграл Лебега

Аппарат теории меры и конструкция измеримых пространств являются тем фундаментом, на котором покоится всё здание теории вероятностей. Мы начнём с описания естественного метода построения мер на прямой и плоскости. Возникающие при этом объекты и понятия применяются при построении мер на произвольных пространствах Ω, не обязательно совпадающих с $\mathbb{R}^k$. Основные факты теории множеств собраны в отдельном приложении (стр. 520).

§ 1. Функции множеств

Сначала очертим круг естественных требований, предъявляемых к любой мере. Во-первых, мера должна приписывать неотрицательные значения не отдельным точкам, а подмножествам Ω. Подмножества, для которых возможно вычисление меры, естественно назвать измеримыми. Таким образом, мера есть функция, заданная на некотором классе измеримых подмножеств. Ясно, что пустое множество должно быть измеримо и иметь меру нуль. Во-вторых, мера должна удовлетворять естественному свойству аддитивности — мера измеримого множества, сложенного из измеримых частей, должна равняться сумме мер этих частей.

Для введения первоначальных понятий нам понадобятся сведения из теории множеств (см. стр. 520), в частности следующие утверждения.

460| Упр. Пусть $\langle Q_n \rangle_1^\infty$ — набор подмножеств Ω. Докажите, что

а) $\dot{\mathrm{I}}\left(\bigcap\limits_{n=1}^{\infty} Q_n\right) = \prod\limits_{n=1}^{\infty} \dot{\mathrm{I}}(Q_n), \quad \dot{\mathrm{I}}\left(\biguplus\limits_{n=1}^{\infty} Q_n\right) = \sum\limits_{n=1}^{\infty} \dot{\mathrm{I}}(Q_n);$

б) $\bigcup\limits_{n=1}^{\infty} Q_n = \biguplus\limits_{k=1}^{\infty} B_k,$ где $B_1 = Q_1,\ B_k = Q_k \bigcap\limits_{n=1}^{k-1} Q_n^c, \quad k \geqslant 2;$

в) $\dot{\mathrm{I}}(\varliminf\limits_n Q_n) = \varliminf\limits_n \dot{\mathrm{I}}(Q_n), \quad \dot{\mathrm{I}}(\varlimsup\limits_n Q_n) = \varlimsup\limits_n \dot{\mathrm{I}}(Q_n);$

г) для монотонной последовательности $\varliminf\limits_n Q_n = \varlimsup\limits_n Q_n$.

Напомним, что знак $\biguplus$ вместо привычного знака объединения применяется нами в ситуациях, когда предполагается или подразумевается, что множества, входящие в объединение, не пересекаются.

Обозначим через $\overline{\mathbb{R}}_+^1$ положительную часть числовой прямой $[0; +\infty]$ с добавленной точкой $\{+\infty\}$.

Определения. Неотрицательная функция $\mu : \mathcal{F} \mapsto \overline{\mathbb{R}}_+^1$, заданная на классе $\mathcal{F}$ подмножеств пространства Ω, называется

— *аддитивной*, если значение $\mu(B)$ на любом элементе $B \in \mathcal{F}$, который можно представить как конечное объединение непересекающихся подмножеств $S_i \in \mathcal{F},\ i = \overline{1, N}$, равно сумме значений μ на этих подмножествах:

$$B,\ \{S_i\}_{i=1}^N \in \mathcal{F}: \quad B = \biguplus\nolimits_{i=1}^N S_i \quad \Rightarrow \quad \mu(B) \ = \ \sum\nolimits_{i=1}^N \mu(S_i); \tag{1}$$

— *счётно-аддитивной (сигма-аддитивной, σ-аддитивной)*, если в (1) допустимы счётные $\mathcal{F}$-разбиения $B \in \mathcal{F}$ (т.е. допустимо $N = \infty$);

— *непрерывной*, если значение $\mu(B)$ на любом элементе $B \in \mathcal{F}$, представимом в виде предела монотонной последовательности измеримых подмножеств $S_n \in \mathcal{F}$, равно пределу значений μ на этих подмножествах:

$$B,\ S_1, S_2, \ldots \in \mathcal{F}: \ B = \lim_n{}_{\Downarrow(\Uparrow)} S_n \quad \Rightarrow \quad \mu(B) \ = \ \lim_n \mu(S_n);$$

— *непрерывной в нуле*, если $\lim_n \mu(S_n) = 0$ при $\forall\,(S_n \searrow \varnothing)$.

Требования σ-аддитивности и непрерывности тоже вполне объяснимы, т.к. чаще всего нам придётся иметь дело с некоторыми асимптотическими утверждениями, для которых соответствующие множества точек будут представлять собой пределы или бесконечные суммы измеримых подмножеств.

461| Упр. Докажите, что

✦ для аддитивной или σ-аддитивной функции $\mu : \mathcal{F} \mapsto \overline{\mathbb{R}}^1_+$ значение $\mu(\varnothing) = 0$, если только $\varnothing \in \mathcal{F}$ и $\exists B \in \mathcal{F}: \quad \mu(B) < \infty$;

✦ σ-аддитивная функция аддитивна, если $\mu(\varnothing) = 0$.

Случай $\mu(B) = \infty$ для $\forall B \in \mathcal{F}$ по понятным причинам рассматриваться не будет.

Определение. Неотрицательная функция множеств $\mu : \mathcal{F} \mapsto \overline{\mathbb{R}}^1_+$ называется *мерой на* $\mathcal{F}$, если она σ-аддитивная и $\mu(\varnothing) = 0$.

462| Пример. В любом пространстве Ω на классе $\mathcal{F} = \mathscr{P}(\Omega)$ всех подмножеств можно задать так называемую считающую меру. Пусть $\mathcal{S}$ — произвольное подмножество Ω. Положим $\mathcal{Z}(A) = \#(A \cap \mathcal{S})$ — количество элементов $\mathcal{S}$, вошедших в A. Легко показать (?!), что $\mathcal{Z}$ — мера. Если $\mathcal{S}$ конечно, то мера любого подмножества конечна. С другой стороны, если, скажем, $\Omega = \mathbb{R}^1$ и $\mathcal{S}$ — достаточно плотное бесконечное подмножество, то мера большинства более или менее интересных подмножеств будет бесконечной.

Очень популярная, особенно у физиков, мера Дира́ка δ_{ω_0}, сосредоточенная в точке ω_0, есть частный случай считающей меры, когда множество $S = \{\omega_0\}$ состоит из одной точки. Мера Дира́ка любого множества равна 1 или 0 в зависимости от того, содержит или нет это множество точку ω_0. ⊚

⚠ Любое конструирование нетривиальной меры может начинаться только с некоторого узкого класса множеств. Кроме того, необходимо указать процедуру продолжения этой конструкции на всё более расширяющееся семейство подмножеств. Окончательный класс подмножеств $\mathscr{F}$, на котором определяется мера, с одной стороны, должен быть достаточно широк, чтобы содержать все привычные множества вроде интервала, круга и т.п. С другой стороны, он не должен быть излишне широк, чтобы любой не слишком сложный процесс построения меры позволял включить этот класс в совокупность измеримых подмножеств.

Естественное понятие длины и площади первоначально было введено для простых объектов типа отрезка и прямоугольника:

$$\lambda\big(\lfloor a; b\rceil\big) \;=\; (b-a), \qquad \lambda\big(\lfloor a_1; b_1\rceil \times \lfloor a_2; b_2\rceil\big) \;=\; (b_1-a_1)(b_2-a_2).$$

Здесь символы $\lfloor$ и $\rceil$ означают, что для этих мер не важно, какого рода, открытого или замкнутого, выбираются соответствующие интервалы. Доказательство непрерывности меры λ на классе всех таких интервалов не представляет труда. Можно доказать и её сигма-аддитивность. Однако нас больше интересуют некие общие черты, которые могут пригодиться при построении других мер. Здесь существенным является то, что класс интервалов образует полукольцо (см. стр. 521).

463| Упр. Докажите, что класс $\mathcal{H}$ подмножеств Ω образует полукольцо, если

а) он замкнут относительно попарных пересечений;

б) для $\forall Q, B \in \mathcal{H}$, $Q \subset B$, найдётся конечный набор подмножеств $\langle Q_1, \ldots, Q_k \rangle \subset \mathcal{H}$ такой, что $B = Q + \biguplus_{j=1}^{k} Q_j$.

Определение. Функция $\mu : \mathcal{F} \mapsto \overline{\mathbb{R}}^1_+$, заданная на классе $\mathcal{F}$ подмножеств Ω, удовлетворяет свойству *полуаддитивности* (σ*-полуаддитивности*), если для любого множества $B \in \mathcal{F}$ и любого его покрытия конечным (счётным) объединением $N < \infty$ (соответственно $N = \infty$) элементов класса $\mathcal{F}$ справедливо

$$B, \langle S_i \rangle_{i=1}^{N} \in \mathcal{F} : \quad B \subset \bigcup\nolimits_1^N S_i \quad \Rightarrow \quad \mu(B) \leqslant \sum\nolimits_1^N \mu(S_i).$$

464| **Лемма.** *Пусть $\gamma : \mathcal{H} \mapsto \overline{\mathbb{R}}^1_+$ — аддитивная неотрицательная функция, заданная на полукольце $\mathcal{H}$ подмножеств Ω. Тогда*

(i$_1$) *γ монотонна: $Q \subset W \ (\in \mathcal{H}) \ \Rightarrow \ \gamma(Q) \leqslant \gamma(W)$;*

(i$_2$) *γ удовлетворяет свойству полуаддитивности;*

(i$_3$) *если γ сигма-аддитивна (мера), то она σ-полуаддитивна;*

(i$_4$) *если γ сигма-полуаддитивна, то она σ-аддитивна (мера).*

$\Rrightarrow$ (i$_1$) Если $Q \subset W$, то ввиду 463 $W = Q + \biguplus_1^k q_j$, где $q_j \in \mathcal{H}$. В силу аддитивности $\gamma(W) = \gamma(Q) + \sum_1^k \gamma(q_j) \geqslant \gamma(Q)$.

(i$_2$) Пусть множество $W = \bigcup_1^N Q_i \ (\in \mathcal{H})$ представлено в виде конечного объединения элементов $\mathcal{H}$. Тогда $W \supset Q_i$ и, снова в силу 463, $W = \biguplus_{j=0}^{k_i} q_{ij}$, где q_{ij} — элементы $\mathcal{H}$, причём $q_{i0} = Q_i$. Таким образом,

$$W = \bigcap_{i=1}^{N} W = \bigcap_{i=1}^{N} \biguplus_{j=0}^{k_i} q_{ij} = \biguplus_{j_1=0}^{k_1} \cdots \biguplus_{j_N=0}^{k_N} (q_{1j_1} \cap \ldots \cap q_{Nj_N}),$$

где элементы $(q_{1j_1} \cap \ldots \cap q_{Nj_N})$ не пересекаются между собой и по свойству полукольца принадлежат этому полукольцу. В силу условия аддитивности

$$\gamma(W) \;=\; \sum_{j_1=0}^{k_1}\cdots\sum_{j_N=0}^{k_N}\gamma(q_{1j_1}\cap\ldots\cap q_{Nj_N}). \tag{2}$$

Заметим, что аналогично предыдущему

$$Q_1 \;=\; q_{10}\cap W \;=\; q_{10}\bigcap_{i=2}^{N}\biguplus_{j=0}^{k_i} q_{ij} \;=\; \biguplus_{j_2=0}^{k_2}\cdots\biguplus_{j_N=0}^{k_N}(q_{10}\cap q_{2j_2}\cap\ldots\cap q_{Nj_N}).$$

Стало быть, в (2) можно выделить слагаемые, сумма которых даст значение $\gamma(Q_1)$ (как и значения $\gamma(Q_i)$, причём каждое из слагаемых (2) ровно один раз входит в представление хотя бы для одного $\gamma(Q_i)$). Следовательно, если $W=\bigcup_1^N Q_i$, то $\gamma(W)\leqslant\sum_1^N\gamma(Q_i)$.

Случай $W\subset\bigcup_1^N Q_i$, ввиду равенства $W=\bigcup_1^N\big(W\cap Q_i\big)$, сводится к предыдущему, поскольку $W\cap Q_i\in\mathcal{H}$ и функция γ монотонна:

$$\gamma(W) \;\leqslant\; \sum\nolimits_1^N\gamma\big(W\cap Q_i\big) \;\leqslant\; \sum\nolimits_1^N\gamma(Q_i).$$

(i_3) Как и при доказательстве полуаддитивности, достаточно рассмотреть случай $W=\bigcup_1^\infty Q_n$, где W и все множества Q_n входят в полукольцо.

Легко проверить, что $\bigcup_{n=1}^\infty Q_n=\biguplus_{n=1}^\infty\Big(Q_n\setminus\bigcup_{i=1}^{n-1}Q_i\Big)$ (см. 460). Применяя два раза лемму 584, стр. 521, получаем, что для $\forall\, n\geqslant 1$ существуют элементы полукольца $Q_{ni},G_{nj}\in\mathcal{H},\ i=\overline{1,L_n},j=\overline{1,Z_n}$, такие, что

$$Q_n\setminus\bigcup_{i=1}^{n-1}Q_i \;=\; \biguplus_{i=1}^{L_n}Q_{ni},\qquad Q_n\setminus\biguplus_{i=1}^{L_n}Q_{ni} \;=\; \biguplus_{j=1}^{Z_n}G_{nj}.$$

Поэтому $Q_n=\biguplus_{j=1}^{Z_n}G_{nj}+\biguplus_{i=1}^{L_n}Q_{ni}$, и в силу аддитивности γ

$$\gamma(Q_n) \;=\; \sum\nolimits_{j=1}^{Z_n}\gamma(G_{nj})+\sum\nolimits_{i=1}^{L_n}\gamma(Q_{ni}) \;\geqslant\; \sum\nolimits_{i=1}^{L_n}\gamma(Q_{ni}).$$

Теперь пришло время воспользоваться σ-аддитивностью:

$$\gamma(W) \;=\; \gamma\Big(\biguplus_{n=1}^{\infty}\biguplus_{i=1}^{L_n}Q_{ni}\Big) \;=\; \sum_{n=1}^{\infty}\sum_{i=1}^{L_n}\gamma(Q_{ni}) \;\leqslant\; \sum_{n=1}^{\infty}\gamma(Q_n).$$

(i_4) Пусть $W=\biguplus_1^\infty Q_i$, где $W,\big\{Q_i\big\}_1^\infty$ — элементы полукольца. По условию $\gamma(W) \leqslant \sum_1^\infty \gamma(Q_i)$, поэтому для доказательства σ-аддитивности необходимо только доказать противоположное неравенство.

Для произвольного $N < \infty$ конечное объединение $\biguplus_1^N Q_i \subset W$. Снова применяя лемму 584, стр. 521, представим W в виде конечной суммы:

$$W \;=\; \biguplus_{i=1}^{N} Q_i \;+\; \biguplus_{j=1}^{L} q_j, \qquad q_j \in \mathcal{H},\ L < \infty.$$

Следовательно, $\gamma(W) \geqslant \sum_1^N \gamma(Q_i)$. Полагая здесь $N \to \infty$, получаем требуемое неравенство $\gamma(W) \geqslant \sum_1^\infty \gamma(Q_i)$. $\lhd$

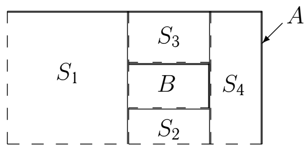
Как показывают следующие примеры, полукольцо интервалов (прямоугольников) может содержать меньшее число элементов, чем те, что были описаны нами при определении меры λ на прямой и плоскости.

465| Примеры. (?!) 1) На числовой прямой $\Omega = \mathbb{R}^1$ система интервалов $\mathcal{K} = \big\langle\, (a;b],\ -\infty < a \leqslant b < \infty \,\big\rangle$ образует полукольцо.

Аналогичная система $\big\langle\, (a;b],\ U \leqslant a \leqslant b \leqslant V \,\big\rangle$ на интервале $\Omega = (U;V]$ образует уже полуалгебру.

2) На плоскости $\mathbb{R}^2$ полукольцо образует класс прямоугольников

$$\mathcal{K} = \big\langle\, (a_1;b_1] \times (a_2;b_2],\quad a_1 < b_1,\ a_2 < b_2 \,\big\rangle.$$



На рисунке справа $A = B + \biguplus_1^4 S_j$.

Ограничение только полуоткрытыми интервалами продиктовано техническими потребностями. Такое полукольцо образует в некотором смысле минимальный набор, на котором можно определить достаточно широкий класс мер. Заметим, что сложность доказательства σ-аддитивности конкретной меры на классе подмножеств $\mathcal{H}$ возрастает с увеличением числа элементов $\mathcal{H}$. ◎

Пример полукольца на плоскости указывает способ образования полуколец на прямом произведении пространств.

466| Упр. Докажите, что если в пространствах Ω_j выделены полукольца множеств $\mathcal{H}_j$, $j = \overline{1,k}$, то на прямом произведении $\Omega = \Omega_1 \times \ldots \times \Omega_k$ система $\mathcal{H} = \mathcal{H}_1 \times \ldots \times \mathcal{H}_k$ образует полукольцо.

Доказательство сигма-аддитивности меры λ мы приведём позднее при построении более общей меры Лебега–Стилтьеса.

Следующий шаг в определении меры на прямой (плоскости) также естествен: для любого конечного объединения непересекающихся интервалов (прямоугольников) полагается по определению $\lambda\big(\biguplus_1^n(a_j;b_j]\big) = \sum_1^n \lambda\big((a_j;b_j]\big)$.

Система всевозможных конечных объединений непересекающихся элементов полукольца образует (минимальное) кольцо множеств, включающее все элементы полукольца (см. 586, стр. 522). При этом мера, доопределённая с полукольца на порождённое кольцо указанным здесь способом, сохраняет свойства σ-аддитивности и непрерывности, т.е. действительно задаёт меру на кольце.

467 **Теорема.** *Пусть* $\mu : \mathcal{H} \mapsto \overline{\mathbb{R}}_+^1$ *— мера, заданная на полукольце. Существует единственная мера* $\tilde{\mu} : \mathfrak{C}(\mathcal{H}) \mapsto \overline{\mathbb{R}}_+^1$, *определённая на кольце* $\mathfrak{C}(\mathcal{H})$, *порождённом* $\mathcal{H}$, *совпадающая с* μ *на* $\mathcal{H}$: $\tilde{\mu}(B) = \mu(B),\ B \in \mathcal{H}$. *При этом для* $\forall\, B \in \mathfrak{C}(\mathcal{H})$ *и любого конечного* $\mathcal{H}$*-разбиения* $B = \biguplus_1^n S_j,\quad n < \infty$,

$$\tilde{\mu}(B) \;=\; \sum\nolimits_1^n \mu(S_j). \tag{3}$$

⊨ Поскольку любой элемент кольца $\mathfrak{C}(\mathcal{H})$ есть конечное объединение непересекающихся элементов полукольца, то соотношение (3) можно взять в качестве определения меры. Совпадение её с мерой μ на полукольце очевидно.

✓ Покажем сначала, что такое определение корректно. Пусть множество B представлено двумя способами в виде объединения элементов полукольца:

$$B \;=\; \biguplus_1^n S_i \;=\; \biguplus_1^k D_j.$$

Тогда $S_i = S_i\, B = \biguplus_{j=1}^k D_j\, S_i$ и в силу аддитивности μ на полукольце $\mu(S_i) = \sum_{j=1}^k \mu(D_j\, S_i),\ i = 1,\dots,n$. Аналогично $\mu(D_j) = \sum_{i=1}^n \mu(D_j\, S_i)$, следовательно,

$$\sum_{i=1}^n \mu(S_i) \;=\; \sum_{i=1}^n \sum_{j=1}^k \mu(D_j\, S_i) \;=\; \sum_{j=1}^k \sum_{i=1}^n \mu(S_i\, D_j) \;=\; \sum_{j=1}^k \mu(D_j).$$

✓ Аддитивность функции $\tilde{\mu}$ очевидна (?!).

✓ Пусть множество B из $\mathfrak{C}(\mathcal{H})$ есть счётное объединение непересекающихся элементов кольца: $B = \biguplus_1^\infty S_k$, причём

$$B \;=\; \biguplus_{j=1}^n B_j, \quad S_k \;=\; \biguplus_{i=1}^{N_k} D_{ki}, \qquad n, N_k < \infty, \quad B_j, D_{ki} \in \mathcal{H},$$

где все множества D_{ki} попарно не пересекаются.

Так как мера μ σ-аддитивна на полукольце и $B_j = B_j B$, то

$$\mu(B_j) \;=\; \mu\Bigl(\biguplus_{k=1}^{\infty}\biguplus_{i=1}^{N_k} B_j\, D_{ki}\Bigr) \;=\; \sum_{k=1}^{\infty}\sum_{i=1}^{N_k}\mu(B_j\, D_{ki}) \;=\; \sum_{k=1}^{\infty}\tilde{\mu}(B_j\, S_k).$$

Поскольку члены бесконечного ряда с положительными слагаемыми можно произвольно переставлять, ввиду аддитивности $\tilde{\mu}$

$$\tilde{\mu}(B) \stackrel{\text{df}}{=} \sum_{j=1}^{n}\mu(B_j) \;=\; \sum_{k=1}^{\infty}\sum_{j=1}^{n}\tilde{\mu}(B_j\, S_k) \;=\; \sum_{k=1}^{\infty}\tilde{\mu}(S_k\, B) \;=\; \sum_{k=1}^{\infty}\tilde{\mu}(S_k). \qquad \Lleftarrow$$

⚠ Продолжение меры с полукольца на кольцо единственно, поэтому для обозначения этого продолжения можно оставить тот же символ, что использовался для меры на полукольце: $\tilde{\mu} = \mu$.

✧ Эквивалентность σ-аддитивности и непрерывности мер

468| **Теорема.** *Пусть функция множеств $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}^1_+$ задана на кольце $\mathcal{A}$ подмножеств Ω.*

I) *Если μ непрерывна в нуле и аддитивна, то она сигма-аддитивна.*

II) *Если μ сигма-аддитивна, то для любой монотонной последовательности $\langle S_n \rangle_1^{\infty} \subset \mathcal{A}$ (при условии $\mu(S_k) < \infty$ для некоторого $k \geqslant 1$)*

$$\lim\nolimits_n \mu(S_n) \;=\; \mu(\lim\nolimits_n S_n).$$

$\Rrightarrow$ I) Пусть $B = \biguplus_1^{\infty} S_j$, тогда $\bigcap_{j=1}^{n}\bigl[B \setminus S_j\bigr] \searrow \varnothing$. Если $\sum_1^n \mu(S_j) < \infty$ при $\forall\, n \geqslant 1$, то в силу непрерывности и аддитивности меры

$$\mu(B) \;=\; \mu\Bigl(\biguplus\nolimits_1^{n} S_j\Bigr) + \mu\Bigl(\bigcap\nolimits_1^{n}\bigl[B \setminus S_j\bigr]\Bigr) \;\underset{n}{\to}\; \sum\nolimits_1^{\infty}\mu(S_j).$$

Если же $\sum_1^n \mu(S_j) = \infty$ при всех $n > n_0$, то очевидно и $\mu(B) = \infty$.

II) Пусть $B = \lim\limits_{n}{}_{\uparrow} S_n = \bigcup_{n=1}^{\infty} S_n = \bigcup_{n=k}^{\infty} S_n$ и $\mu(S_n) < \infty, \forall\, n$. Случай $\mu(S_k) = \infty$ при некотором k очевиден. Так как $S_1 \subset S_2 \subset \ldots$, то в силу 460, стр. 424,

$$\bigcup\nolimits_k^{\infty} S_n \;=\; S_k + \biguplus\nolimits_k^{\infty}(S_{n+1} \setminus S_n)$$

и мера $\mu(S_{n+1} \setminus S_n) = \mu(S_{n+1}) - \mu(S_n)$, ибо $\mu(S_n) < \infty$. Следовательно, в силу σ-аддитивности μ в этом случае

$$\begin{aligned}\mu(B) \;&=\; \mu(S_k) + \sum\nolimits_{n=k}^{\infty}(\mu(S_{n+1}) - \mu(S_n)) = \\ &=\; \lim_{N}\Bigl(\mu(S_k) + \sum\nolimits_{n=k}^{N}(\mu(S_{n+1}) - \mu(S_n))\Bigr) \;=\; \lim_{N}\mu(S_N).\end{aligned}$$

Пусть $B = \lim\limits_{n}{}_{\downarrow} S_n = \bigcap_1^\infty S_n = \bigcap_k^\infty S_n$ для $\forall k \geqslant 1$. По правилу двойственности d'Morgána

$$S_k \setminus \bigcap\nolimits_1^\infty S_n \;=\; \bigcup\nolimits_{n=k+1}^\infty (S_k \setminus S_n) \;=\; \lim_n{}_{\uparrow} (S_k \setminus S_n),$$

где учтено то, что $S_k \setminus S_{k+1} \subset S_k \setminus S_{k+2}$. По уже доказанному свойству непрерывности меры $(\mu(S_k) < \infty)$

$$\mu(S_k) - \mu\Big(\bigcap\nolimits_1^\infty S_n\Big) \;=\; \mu(S_k) - \lim_n \mu(S_n) \;\Rightarrow\; \mu(B) \;=\; \lim_n \mu(S_n). \qquad \Lleftarrow$$

⚠ Для функций множеств на кольце условия σ-аддитивности и непрерывности эквивалентны. Кроме того, любое кольцо есть полукольцо, следовательно, для таких функций справедливы утверждения леммы <u>464</u>, стр. 426. Поэтому определение меры иногда удобнее начинать с кольца подмножеств.

469| Упр. Если $\mu(S_n) = \infty,\ \forall n \geqslant 1$, то $\mu(\lim_\uparrow S_n) = \lim \mu(S_n) = \infty$. Приведите пример, когда в этом случае может быть $\mu(\lim_\downarrow S_n) = 0$.

✧ **Внешняя мера. Теорема Каратеодори.** Определение меры для более сложных подмножеств и расширение класса измеримых подмножеств, естественно, невозможно без некоего предельного перехода. Результат будет зависеть от способа такого перехода. Так, мера Жордана в $\mathbb{R}^k$ определяется для тех подмножеств G, для которых

$$\inf_{G^* \in \overline{\mathcal{G}}} \sum_{A \in G^*} \lambda\big(A\big) \;=\; \sup_{G_* \in \underline{\mathcal{G}}} \sum_{A \in G_*} \lambda\big(A\big),$$

где $\overline{\mathcal{G}} = \Big\langle\, G^* = \big\{A_i\big\}_1^n \subset \mathcal{H} : \ \ \biguplus_1^n A_i \supset G, \ \ n < \infty \,\Big\rangle$ — совокупность «приближений» множества G извне *конечными* объединениями элементов полукольца $\mathcal{H}$ (интервалов, прямоугольников и т.п.). Аналогично, $\underline{\mathcal{G}} = \Big\langle\, G_* = \big\{A_i\big\}_1^n \subset \subset \mathcal{H} : \ \ \biguplus_1^n A_i \subset G, \ \ n < \infty \,\Big\rangle$ задаёт приближения G изнутри. К сожалению, класс измеримых по Жордану множеств не слишком широк (см. задачу ниже). Этот класс не замкнут относительно счётных объединений и пересечений, что неудобно: каждое утверждение асимптотического характера будет требовать обязательной проверки измеримости соответствующего множества.

470| Упр. Докажите, что множество всех рациональных точек на отрезке [0;1] неизмеримо по Жордану.

Продолжение Лебега строится чуть иначе.

Определение. Пусть $\mathcal{A}$ — кольцо подмножеств пространства Ω и мера $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}^1_+$ задана на $\mathcal{A}$. *Внешней мерой* (относительно μ) называется функция μ^*, определённая для любого подмножества $G \subset \Omega$:

$$\mu^*(G) \;=\; \inf\Big\{ \sum_{B_j \in G^\sigma} \mu(B_j) : \quad G^\sigma = \{B_j\}_1^\infty \subset \mathcal{A}, \;\; \bigcup\nolimits_1^\infty B_j \supset G \Big\}.$$

При этом полагается по определению $\mu^*(\varnothing) = 0$ и $\mu^*(G) = \infty$, если не существует ни одного $\mathcal{A}$-покрытия множества G.

471| Упр. Воспользовавшись утверждением 6) задачи 460, стр. 424, покажите, что в определении внешней меры можно ограничиться покрытиями $\biguplus B_i \supset G$ с непересекающимися B_i. Более того, эти подмножества можно выбирать только из полукольца, порождающего кольцо $\mathcal{A}$.

472| **Лемма.** (?!) *Внешняя мера μ^* обладает следующими свойствами:*

(i$_1$) *μ^* — монотонная мера, т.е. $A \subset B \;\;\Rightarrow\;\; \mu^*(A) \leqslant \mu^*(B)$;*

(i$_2$) *μ^* — полная мера, т.е. подмножества множеств нулевой внешней меры имеют внешнюю меру нуль: $\mu^*(B) = 0, A \subset B \;\;\Rightarrow\;\; \mu^*(A) = 0$;*

(i$_3$) *$\mu^*(\Omega) = \mu(\Omega)$, если в кольцо $\mathcal{A}$ входит Ω (т.е. если $\mathcal{A}$ — алгебра).*

Можно было бы на этом и остановиться, однако внешняя мера не обязательно обладает необходимыми качествами непрерывности и σ-аддитивности — она монотонна и в общем случае лишь счётно-полуаддитивна.

Определение. Множество $G \subset \Omega$ *измеримо по мере μ^**, если

$$\mu^*(B) = \mu^*(BG) + \mu^*(B\,G^c), \qquad \forall\, B \subset \Omega.$$

⚠ 1) Если $B = \Omega$, то для любого покрытия $\bigcup_j^\infty A_j \supset G^c$ подмножество $\big(\bigcup_j^\infty A_j\big)^c$ $(\subset G)$ можно интерпретировать как приближение G изнутри. Таким образом, в случае $\mu^*(\Omega) < \infty$ измеримость означает совпадение приближений меры G изнутри и снаружи. Для неизмеримых множеств G в силу полуаддитивности μ^* справедливо только неравенство $\mu^*(B) \leqslant \mu^*(B\,G) + \mu^*(B\,G^c)$.

2) Измеримость множества не означает, что его внешняя мера конечна.

473| Упр. Докажите, что пустое множество $\varnothing$ и всё пространство Ω измеримы и, кроме того, измеримы все множества нулевой меры (внешней).

Здесь также можно было бы остановиться, выбрав, например в $\mathbb{R}^1$, в качестве общего класс подмножеств, измеримых по Лебегу, однако этот класс излишне широк — не всякий процесс построения меры сможет «дотянуться» до него. Желательно найти класс, в который входили бы измеримые подмножества любой достаточно «естественной» меры. Как будет видно из следующей теоремы, различные части которой принадлежат Каратеодори и Хану ([9], стр. 150, 152), на эту роль может претендовать σ-алгебра, порождённая полукольцом или кольцом (см. стр. 522 дополнения).

474| Упр. Докажите, что если $\mathcal{A} = \mathfrak{C}(\mathcal{H})$ — кольцо, порождённое полукольцом $\mathcal{H}$, то минимальная σ-алгебра $\sigma(\mathcal{H}) = \sigma(\mathcal{A})$.

Определение. Мера μ, заданная на классе $\mathcal{F}$ подмножеств Ω, называется *σ-конечной*, если выполнено одно из двух (эквивалентных) условий:

— найдётся такое $\mathcal{F}$-покрытие $\Omega = \bigcup_1^\infty F_n$, что $\mu(F_n) < \infty$ при $\forall n \geqslant 1$;

— найдётся возрастающая последовательность подмножеств $F_n \subset F_{n+1} \in \in \mathcal{F}$ такая, что $F_n \nearrow \Omega$ и $\mu(F_n) < \infty$ при $\forall n \geqslant 1$.

475| **Теорема.** [*К. Каратеодори.*] *Пусть на кольце $\mathcal{A}$ подмножеств пространства Ω задана мера $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}^1_+$, μ^* — соответствующая ей внешняя мера на $\mathscr{P}(\Omega)$. Тогда*

(I_1) *класс $\Psi = \Psi_\mu$ подмножеств Ω, измеримых относительно μ^*, образует σ-алгебру;*

(I_2) *кольцо $\mathcal{A} \subset \Psi$ и, следовательно, Ψ содержит минимальную σ-алгебру $\sigma(\mathcal{A})$, порождённую кольцом $\mathcal{A}$;*

(I_3) *μ^* — мера на Ψ, причём $\mu^*(B) = \mu(B)$ для $\forall B \in \mathcal{A}$;*

(I_4) *если μ — σ-конечная мера на $\mathcal{A}$, то μ^* — единственная мера на σ-алгебре $\sigma(\mathcal{A})$, совпадающая с μ на кольце $\mathcal{A}$.*

$\Rrightarrow$ Стр. 513–516. $\Lleftarrow$

476| **Следствие.** *Для любой сигма-конечной меры μ, заданной на полукольце $\mathcal{H}$, существует единственное продолжение μ^* до минимальной σ-алгебры $\sigma(\mathcal{H})$, порождённой $\mathcal{H}$.*

► Итак, можно выделить следующие этапы процесса построения меры на подмножествах произвольного абстрактного пространства Ω:

— задаётся полукольцо $\mathcal{H}$ подмножеств Ω;

— на элементах $\mathcal{H}$ определяется функция $\mu : \mathcal{H} \mapsto \overline{\mathbb{R}}^1_+$;

— устанавливается σ-аддитивность (или аддитивность и σ-полуаддитивность) и σ-конечность μ на $\mathcal{H}$;

— мера μ продолжается на кольцо $\mathfrak{C}$ всех конечных объединений непересекающихся элементов полукольца $\mathcal{H}$;

— делается вывод (со ссылкой на теорему Каратеодори) о существовании единственной меры μ^*, служащей продолжением меры μ на σ-алгебру $\sigma(\mathcal{H})$, порождённую полукольцом $\mathcal{H}$ (или, что эквивалентно, кольцом $\mathfrak{C}$).

⚠ Можно определить меру сразу на кольце или даже алгебре подмножеств. При этом, с учетом теоремы эквивалентности 468, вместо доказательства σ-аддитивности можно попытаться доказать её непрерывность.

Таким образом, задача свелась к выбору полукольца (кольца) $\mathcal{H}$, для которого порождённая σ-алгебра содержит достаточно широкий класс подмножеств. Любую меру можно считать сразу заданной на этой σ-алгебре.

Определения. Пространство $(\Omega, \mathscr{F})$ с выделенной σ-алгеброй $\mathscr{F}$ подмножеств Ω называется *измеримым пространством*, а подмножества Ω, входящие в σ-алгебру $\mathscr{F}$, — *измеримыми* (или $\mathscr{F}$-*измеримыми*).

Если на измеримом пространстве определена мера μ, то такое пространство $(\Omega, \mathscr{F}, \mu)$ называется *пространством с мерой*.

Пространство с *нормированной* мерой $\mu(\Omega) = 1$ называется *вероятностным пространством*, а множества из $\mathscr{F}$ — *событиями*.

Интересно, что элементы σ-алгебры могут быть приближены элементами кольца. Напомним, что операция $A \triangle B = (A \setminus B) + (B \setminus A) = AB^c + BA^c$ называется симметрической разностью и представляет собой как бы теоретико-множественную характеристику расхождения двух множеств.

477| **Лемма.** *Пусть* $\mu : \sigma(\mathcal{A}) \mapsto \overline{\mathbb{R}}^1_+$ *— мера, заданная на* σ*-алгебре, порождённой кольцом* $\mathcal{A}$*, причём на* $\mathcal{A}$ *мера* μ σ*-конечна. Тогда для* $\forall \varepsilon > 0$ *и* $\forall G \in \sigma(\mathcal{A})$, $\mu(G) < \infty$*, существует такое* $B \in \mathcal{A}$*, что* $\mu(B \triangle G) < \varepsilon$.

⤃ Пусть множество $Q \in \mathcal{A}$ таково, что $\mu(Q) < \infty$. Обозначим через $\mathcal{G}$ класс подмножеств $G \in \sigma(\mathcal{A})$, для которых при $\forall \varepsilon > 0$ найдётся подмножество $B \in \mathcal{A}$ такое, что $\mu(B \triangle (QG)) < \varepsilon$. Этот класс, очевидно, содержит все подмножества из кольца $\mathcal{A}$ (с $B = QG$). Покажем, что он замкнут относительно операции пересечения. Пусть $G_1, G_2 \in \mathcal{G}$, т.е. для $\forall \varepsilon > 0$ найдутся подмножества $B_1, B_2 \in \mathcal{A}$ такие, что

$$\mu(B_j \triangle (QG_j)) \;=\; \mu(B_j Q^c \cup B_j G_j^c) + \mu(QG_j B_j^c) \;<\; \tfrac{1}{2}\varepsilon, \qquad i = 1, 2.$$

Тогда

$$\big(Q(G_1G_2)\big) \vartriangle (B_1B_2) \;=\; QG_1G_2(B_1^c \cup B_2^c) + B_1B_2(Q^c \cup G_1^c \cup G_2^c) \;\subset$$
$$\subset\; \big(B_1Q^c \cup B_1G_1^c + QG_1B_1^c\big) \bigcup \big(B_2Q^c \cup B_2G_2^c + QG_2B_1^c\big).$$

Откуда $\mu(Q(G_1G_2) \vartriangle (B_1B_2)) < \varepsilon$. Аналогичным образом устанавливается замкнутость $\mathcal{G}$ относительно разности двух множеств.

Покажем, что $\mathcal{G}$ есть монотонный класс. Пусть имеется возрастающая последовательность элементов $\mathcal{G} : G_k \nearrow G$. По свойству непрерывности меры для $\forall \varepsilon > 0$ найдётся $k:$ $\mu(QG \setminus (QG_k)) < \frac{1}{2}\varepsilon$. Так как $G_k \in \mathcal{G}$, то найдётся $B \in \mathcal{A}$ такое, что $\mu(B \vartriangle QG_k) < \frac{1}{2}\varepsilon$. Тогда

$$QGB^c + B(QG)^c \;\subset\; QG_kB^c \cup B(Q^c \cup G_k^c) \cup QGG_k^c$$

ввиду включения $G \supset G_k$. Следовательно, $\mu(B \vartriangle QG) \leqslant \mu(B \vartriangle QG_k) + \mu(QG \setminus (QG_k)) < \varepsilon$.

Таким образом, класс $\mathcal{G}$ есть λ-система, замкнутая относительно конечных пересечений. Так как, очевидно, $\Omega \in \mathcal{G}$, то по лемме 591, стр. 523, класс $\mathcal{G}$ образует σ-алгебру. Кольцо $\mathcal{A} \subset \mathcal{G}$, поэтому $\mathcal{G} = \sigma\big(\mathcal{A}\big)$, т.е. для любого подмножества $G \in \sigma\big(\mathcal{A}\big)$ и $\forall \varepsilon > 0$ найдётся подмножество из кольца $B \in \mathcal{A}$ такое, что $\mu(B \vartriangle (QG)) < \varepsilon$.

Пусть теперь $G \in \sigma\big(\mathcal{A}\big)$ и $\mu(G) < \infty$. Поскольку мера μ сигма-конечна на кольце, найдётся множество $Q \in \mathcal{A}$ такое, что $\mu(G \setminus (QG)) < \frac{1}{2}\varepsilon$. По предыдущему можно указать множество $B \in \mathcal{A}$, для которого $\mu(B \vartriangle (QG)) < \frac{1}{2}\varepsilon$, откуда $\mu(B \vartriangle G) < \varepsilon$. ⫤

«Аппроксимирующее свойство» меры может быть расширено до всех подмножеств исходного пространства.

478| **Лемма.** *Пусть $(\Omega, \mathscr{F}, \mu)$ — пространство с мерой на σ-алгебре $\mathscr{F}$, μ^* – внешняя мера μ. Тогда для $\forall\, A \subset \Omega$ существует подмножество $G \in \mathscr{F}$ такое, что $G \supset A$ и $\mu^*(A) = \mu(G)$.*

⊨ По определению внешней меры найдётся последовательность $\big\langle\, F_{jn} \,\big\rangle_{j=1}^{\infty}$, $n \geqslant 1$, счётных $\mathscr{F}$-покрытий множества A таких, что $\mu^*(A) = \lim_n \sum_{j=1}^{\infty} \mu(F_{jn})$. По лемме 460, стр. 424, каждое из этих покрытий можно выбрать так, чтобы все его элементы не пересекались. Поэтому для множеств $B_n = \biguplus_j F_{jn} \in \mathscr{F}$ последовательность $\mu(B_n) \to \mu^*(A)$. Легко понять, что любое конечное пересечение $\cap_1^n B_j$ также образует покрытие A и, кроме того,

$$\min_{j\leqslant n}\mu(B_j) \;\geqslant\; \mu\Big(\bigcap_1^n B_j\Big) \geqslant \mu^*(A).$$

Так как $\mu(B_j) \geqslant \mu^*(A)$ при всех $j \geqslant 1$, то $\min\limits_{j\leqslant n}\mu(B_j) \underset{n}{\to} \mu^*(A)$. Поэтому по свойству непрерывности меры $\mu(\cap_1^\infty B_j) = \mu^*(A)$. ⫤

Для множеств нулевой меры, опираясь только на определение внешней меры, можно несколько уточнить предыдущие утверждения.

479| **Лемма.** (?!) *Пусть* $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}_+^1$ *— мера на кольце подмножеств. Внешняя мера* $\mu^*(N) = 0$ *т. т. т. когда для* $\forall\,\varepsilon > 0$ *найдётся семейство непересекающихся подмножеств* $\langle\, S_n \,\rangle_1^\infty \subset \mathcal{A}$*, что* $\biguplus_n S_n \supset N$ *и* $\sum_n \mu(S_n) < \varepsilon$ *.*

✧ Пополнение по мере. Представляется весьма удобным, когда, доказав равенство нулю меры какого-либо множества, мы можем перенести это свойство на более мелкие подмножества: $B \in \mathscr{F},\ \mu(B) = 0,\ A \subset B \ \Rightarrow\ A \in \in \mathscr{F},\ \mu(A) = 0$. Мера (или σ-алгебра) с таким свойством называется *полной*. Любую σ-алгебру можно пополнить по мере. Рассмотрим класс

$$\mathscr{F}_\mu \;=\; \Big\langle\, B \subset \Omega : \ \ \exists A, C \in \mathscr{F},\ A \subset B \subset C,\ \mu(C \setminus A) = 0 \,\Big\rangle. \qquad (4)$$

Этот класс образует уже σ-алгебру (?!) относительно естественного продолжения μ: $\tilde{\mu}(B) \;=\; \mu(C)$, если $A \subset B \subset C$, $\mu(C \setminus A) = 0$.

Как отмечалось выше, класс Ψ_μ множеств, измеримых относительно внешней меры μ^*, полон. Покажем, что $\Psi_\mu = \mathscr{F}_\mu$, если μ — σ-конечная мера на σ-алгебре $\mathscr{F}$. В этом случае, ввиду утверждения 6) леммы 460, стр. 424, существует разбиение $\Omega = \biguplus_n \Omega_n$ такое, что $\mu(\Omega_n) < \infty$ при $\forall\, n \geqslant 1$. Если множество $B \in \Psi_\mu$, т.е. B измеримо относительно внешней меры μ^*, тогда

$$\mu^*(\Omega_n) \;=\; \mu^*(B\Omega_n) + \mu^*(B^c\Omega_n).$$

По лемме 478 найдутся множества $C_n, G_n \in \mathscr{F}$ такие, что:

$$\begin{aligned}\mu^*(B\Omega_n) \;&=\; \mu(C_n), \;\; B\Omega_n \;\subset\; C_n,\\ \mu^*(B^c\Omega_n) \;&=\; \mu(G_n), \;\; B^c\Omega_n \;\subset\; G_n.\end{aligned}$$

Легко понять, что множества $C_n\Omega_n$ и $G_n\Omega_n$ также удовлетворяют этим свойствам, поэтому можно считать, что $C_n, G_n \subset \Omega_n$. Следовательно, т.к. $\mu^*(\Omega_n) = = \mu(\Omega_n) < \infty$, то $\mu(\Omega_n) \;=\; \mu(C_n) + \mu(G_n)$, т.е. $\mu(\Omega_n \setminus G_n) \;=\; \mu(C_n)$.

Ясно, что $A_n = \Omega_n \setminus G_n \subset B\Omega_n$, стало быть, для $\forall n \geqslant 1$ найдутся множества $A_n, C_n \in \mathscr{F}$ такие, что $A_n \subset B\Omega_n \subset C_n \subset \Omega_n$ и $\mu(A_n) = \mu(C_n) < \infty$.

Положим $A = \biguplus_n A_n, C = \biguplus_n C_n$. Так как $B = \biguplus_n B\Omega_n$, то $A \subset B \subset C$ и

$$\mu(C \setminus A) \;=\; \sum_n \mu(C_n \setminus A_n) \;=\; 0.$$

Таким образом, пополнение σ-алгебры $\mathscr{F}$ по σ-конечной мере μ совпадет с классом измеримых множеств Ψ_μ.

§ 2. Борелевская σ-алгебра и мера в евклидовом пространстве

✧ **Борелевская σ-алгебра в $\mathbb{R}^1$.** Применим описанный выше процесс сначала к построению меры Лебега–Стилтьеса на числовой прямой. Как уже отмечалось, минимальный естественный класс подмножеств $\mathbb{R}^1$, с которого удобно начинать построение меры, есть полукольцо $\mathcal{K}$ всех интервалов вида $(a;b]$, $-\infty < a \leqslant b < \infty$.

Определение. Минимальная σ-алгебра $\sigma(\mathcal{K})$, порождённая полукольцом $\mathcal{K}$, называется *борелевской* и обозначается $\mathscr{B}(\mathbb{R}^1)$ или $\mathscr{B}$.

480| Упр. Докажите $\mathscr{B}$-измеримость

- одноточечных множеств и любых счётных подмножеств $\mathbb{R}^1$;
- интервалов вида $(a;b)$, $(a;b]$, $[a;b)$, $[a;b]$ (конечных или бесконечных);
- открытых или замкнутых множеств $\mathbb{R}^1$.

Возможны и другие варианты определения борелевской σ-алгебры. Доказательство их идентичности зиждется на утверждениях леммы 595, стр. 525.

481| **Лемма.** (?!) *Борелевская σ-алгебра*

$$\mathscr{B} = \sigma(\mathcal{K}) = \sigma(\mathscr{O}) = \sigma(\mathcal{K}_{()}) = \sigma(\mathcal{K}_{[)}) = \sigma(\mathcal{U}) = \sigma(\mathcal{C}),$$

где $\mathcal{K}_{()}$ — *класс открытых интервалов* $(a;b)$, $a,b \in \mathbb{R}^1$,
$\mathcal{K}_{[)}$ — *класс полуоткрытых интервалов* $[a;b)$, $a,b \in \mathbb{R}^1$,
$\mathscr{O}$ — *класс всех открытых множеств на прямой,*
$\mathcal{U}$ — *кольцо всех конечных объединений элементов* $\mathcal{K}$,
$\mathcal{C}$ — *система интервалов вида* $(-\infty;b]$, $b \in \mathbb{R}^1$.

⤃ $\left[\sigma(\mathscr{O}) = \sigma(\mathcal{K}_{()})\right]$ Так как любое открытое множество $\mathbb{R}^1$ есть счётное объединение открытых интервалов, то ввиду 480 класс $\mathscr{O} \subset \mathscr{B}$. С другой стороны, класс интервалов $\mathcal{K}_{()}$ порождает $\mathscr{B}$ и $\mathcal{K}_{()} \subset \mathscr{O} \subset \mathscr{B}$. Поэтому $\mathscr{B} = \sigma\big(\mathcal{K}_{()}\big) \subset \sigma\big(\mathscr{O}\big) \subset \mathscr{B}$. ⇚

⚠ Э. Борель (1898) ввёл определение меры открытых множеств $\mu(B) = \sum_n (c_n - a_n)$, если $B = \biguplus_n (a_n; c_n)$, и продолжил её на σ-алгебру, порождённую всеми открытыми множествами. Поэтому в топологическом пространстве Ω с классом $\mathscr{O}$ всех открытых множеств (см. стр. 526) σ-алгебру $\mathscr{B}(\Omega) := \sigma\big(\mathscr{O}\big)$ естественно называть борелевской. Отметим, что эта σ-алгебра будет совпадать с σ-алгеброй, порождённой базой топологии, только если эта база счётная, что верно для польских (например, евклидовых) пространств.

Из 476 непосредственно вытекает следующая важная

482| **Теорема.** *Любая σ-конечная мера, заданная на полукольце интервалов $\mathcal{K}$, может быть единственным образом продолжена до борелевской σ-алгебры $\mathscr{B}$. Неформально говоря, элементы борелевской σ-алгебры $\mathscr{B}$ измеримы относительно любой меры, заданной на полукольце $\mathcal{K}$; в частности борелевские множества измеримы по Лебегу.*

⚠ Класс (σ-алгебра) измеримых по Лебегу множеств шире борелевской σ-алгебры. Из дальнейшего будет понятно, что этот класс может быть получен из σ-алгебры $\mathscr{B}$ с помощью процесса пополнения по мере Лебега.

Ещё одну характеризацию борелевской σ-алгебры даёт следующая теорема, являющаяся следствием теоремы 597, стр. 525. Эта характеризация объясняет причину, по которой чаще всего задание какой-либо меры начинают не с полукольца, а сразу с порождённого кольца.

483| **Теорема.** *Борелевская σ-алгебра $\mathscr{B}$ совпадает с минимальным монотонным классом $\mathfrak{M}(\mathcal{U})$, порождённым кольцом $\mathcal{U} = \mathfrak{C}(\mathcal{K})$ всех конечных объединений непересекающихся интервалов полукольца $\mathcal{K}$.*

✧ **Индуцированная σ-алгебра. Отрезок $[0;1]$.** В общем случае σ-алгебра $\mathscr{F}$, определённая на пространстве Ω, сужается (индуцируется) на любое подмножество $Q \subset \Omega$ следующим образом.

484| **Лемма.** (?!) *Пусть $\mathscr{F}$ — σ-алгебра в Ω и $Q \subset \Omega$. Тогда совокупность пересечений Q со всеми множествами из $\mathscr{F}$*

$$Q \Cap \mathscr{F} := \big\langle Q \cap B : \ B \in \mathscr{F} \big\rangle$$

образует σ-алгебру подмножеств Q.

Подсказка. *Роль основного пространства выполняет не Ω, а Q.*

Здесь существенна связь индуцированной (на Q (!)) σ-алгебры с порождающим исходную σ-алгебру семейством.

485 **Теорема.** *Пусть $\sigma(Q \Cap \mathcal{H}; Q)$ — минимальная σ-алгебра на Q, порождённая $Q \Cap \mathcal{H}$. Тогда индуцированная на Q σ-алгебра*

$$Q \Cap \sigma(\mathcal{H}) \;=\; \sigma(Q \Cap \mathcal{H}; Q).$$

⊫ Так как $Q \Cap \sigma(\mathcal{H}) \supset Q \Cap \mathcal{H}$, то из предыдущей леммы следует, что $Q \Cap \sigma(\mathcal{H}) \supset \sigma(Q \Cap \mathcal{H}; Q)$.

Для доказательства обратного рассмотрим на Ω класс подмножеств $\mathcal{H}_Q = \left\langle B \in \sigma(\mathcal{H}) : \; Q\,B \in \sigma(Q \Cap \mathcal{H}; Q) \right\rangle$, который, очевидно, замкнут относительно счётных объединений и содержит Ω. Так как $\sigma(Q \Cap \mathcal{H}; Q)$ есть σ-алгебра на Q, то из включения $Q\,B \in \sigma(Q \Cap \mathcal{H}; Q)$ следует, что $Q\,B^c = Q \setminus (Q\,B) \in \sigma(Q \Cap \mathcal{H}; Q)$. Следовательно, $\mathcal{H}_Q$ — σ-алгебра на Ω.

С другой стороны, по построению $\mathcal{H} \subset \mathcal{H}_Q \subset \sigma(\mathcal{H})$, стало быть, $\mathcal{H}_Q = \sigma(\mathcal{H})$, т.е. для $\forall\, B \in \sigma(\mathcal{H}) \;\Rightarrow\; Q\,B \in \sigma(Q \Cap \mathcal{H}; Q)$. ⫤

⚠ Если $Q \in \sigma(\mathcal{H})$ и подмножество $B \subset Q \subset \Omega$, то, установив его измеримость в более узком пространстве — $B \in \sigma(Q \Cap \mathcal{H}; Q)$, мы можем утверждать, что оно измеримо и в широком пространстве — $B \in \sigma(\mathcal{H})$. И обратно, измеримость в широком пространстве влечёт измеримость в узком.

Особую роль среди подпространств $\mathbb{R}^1$ играет отрезок $[0;1]$. В соответствии с предыдущей леммой борелевская σ-алгебра на $[0;1]$, индуцированная борелевской σ-алгеброй $\mathbb{R}^1$, может быть определена как минимальная σ-алгебра (на $[0;1]$), порождённая семейством интервалов $\left\langle [0;b], 0 \leqslant b \leqslant 1 \right\rangle$, или семейством $\left\langle (a;b], 0 \leqslant a \leqslant b \leqslant 1 \right\rangle$, или … (см. 481).

► Иногда бывает удобно считать достижимыми бесконечные точки, т.е. вместо прямой $\mathbb{R}^1 = (-\infty;+\infty)$ рассматривать прямую $\overline{\mathbb{R}}^1 = [-\infty;+\infty]$ с добавленными «точками» $\{-\infty\}$, $\{+\infty\}$. Борелевскую σ-алгебру $\mathscr{B}(\overline{\mathbb{R}}^1)$ в этом пространстве можно построить исходя из полуалгебры

$$\tilde{\mathcal{K}} = \left\langle\, [-\infty;b], \quad (a;+\infty], \quad [-\infty;+\infty], \quad (c;d] \quad a,b,c,d \in \mathbb{R}^1 \right\rangle$$

или класса интервалов вида $\widehat{\mathcal{K}} = \big\langle\, [-\infty; b],\quad [-\infty; +\infty],\quad b \in \mathbb{R}^1 \,\big\rangle$.

Заметим, что $\lim\limits_{b\to+\infty} (a;b] = (a;+\infty) \neq (a;+\infty]$.

Несмотря на то что пересечение $\mathbb{R}^1 \Cap \tilde{\mathcal{K}}$ содержит больше элементов, чем полукольцо $\mathcal{K}$ конечных интервалов, очевидно, индуцированная (на $\mathbb{R}^1$) σ-алгебра $\mathbb{R}^1 \Cap \mathscr{B}(\overline{\mathbb{R}}^1)$ совпадает с борелевской σ-алгеброй $\mathscr{B}(\mathbb{R}^1)$.

486| **Лемма.** (?!) *Множество $B \in \mathscr{B}(\overline{\mathbb{R}}^1)$ т. т. т. когда подмножество его конечных точек $B \setminus \{-\infty, +\infty\} \in \mathscr{B}(\mathbb{R}^1)$.*

✧ Мера Лебега–Стилтьеса на борелевской σ-алгебре $\mathbb{R}^1$. Пусть задана конечная, непрерывная справа, неубывающая функция $F(x)$ действительного аргумента $x \in \mathbb{R}^1$. В соответствии с конструкцией меры на σ-алгебре, определим сначала функцию множеств на полукольце $\mathcal{K}$ интервалов вида $(a;b],\quad a \leqslant b,\quad a,b \in \mathbb{R}^1$, по формуле

$$\mu_F(a;b] \;=\; F(b) - F(a).$$

Неотрицательность μ_F очевидна. Кроме того, т.к. прямую $\mathbb{R}^1$ можно представить в виде счётного объединения интервалов с целыми границами, то очевидно также, что μ_F σ-конечна. Для доказательства σ-аддитивности μ_F по лемме 464, стр. 426, необходимо установить её аддитивность и σ-полуаддитивность.

✓ Аддитивность покажем, используя геометрические свойства прямой. Рассмотрим конечное $\mathcal{K}$-разбиение $(A;B] = \biguplus_1^N (a_j;b_j]$. Перегруппируем интервалы так, чтобы $A = a_1 < b_1 = a_2 < \ldots < b_{N-1} = a_N < b_N = B$. Тогда

$$\mu_F(A;B] \;=\; F(B) - F(A) \;=\; F(b_N) - F(a_N) + F(b_{N-1}) - \ldots$$
$$- F(a_2) + F(b_1) - F(a_1) \;=\; \sum\nolimits_1^N \mu_F(a_j;b_j].$$

✓ Для доказательства σ-полуаддитивности μ_F воспользуемся топологическими свойствами $\mathbb{R}^1$ и непрерывностью справа функции F. Рассмотрим счётное $\mathcal{K}$-разбиение интервала $(A;B] = \biguplus_1^\infty (a_j;b_j]$. Выберем (пока произвольно) точки $\widetilde{A} > A$ $(< B)$ и $\tilde{b}_j > b_j$. Для этих точек замкнутый отрезок

$$[\widetilde{A};B] \;\subset\; (A;B] \;=\; \biguplus\nolimits_1^\infty (a_j;b_j] \;\subset\; \bigcup\nolimits_1^\infty (a_j;\tilde{b}_j).$$

Таким образом, компакт (ограниченное и замкнутое множество) на числовой прямой покрывается открытыми интервалами. По свойству компактов (лемма

Гейне–Бореля) здесь можно выделить конечный набор N интервалов, покрывающих $[\widetilde{A}\,;B]$ (и тем более $(\widetilde{A}\,;B])$:

$$(\widetilde{A}\,;B] \;\subset\; [\widetilde{A}\,;B] \;\subset\; \bigcup\nolimits_1^N (a_j;\tilde{b}_j) \;\subset\; \bigcup\nolimits_1^N (a_j;\tilde{b}_j].$$

По доказанному в лемме 464, стр. 426, свойству полуаддитивности

$$\mu_F(\widetilde{A}\,;B] \;\leqslant\; \sum\nolimits_1^N \mu_F(a_j;\tilde{b}_j].$$

Воспользуемся теперь непрерывностью справа функции F и выберем для фиксированного $\varepsilon > 0$ точки $\widetilde{A}\,,\tilde{b}_j$ так, чтобы

$$0 \leqslant F(\widetilde{A}\,) - F(A) < \varepsilon \quad \text{и} \quad 0 \leqslant F(\tilde{b}_j) - F(b_j) < \frac{1}{2^j}\,\varepsilon, \quad j \geqslant 1.$$

Тогда $\mu_F(A\,;B] - \mu_F(\widetilde{A}\,;B] = F(\widetilde{A}\,) - F(A) \;< \varepsilon$, и при $\forall\, j \geqslant 1$ разность мер $\mu(a_j;\tilde{b}_j] - \mu_F(a_j;b_j] \;=\; F(\tilde{b}_j) - F(b_j) \;< \varepsilon\, 2^{-j}$. Следовательно,

$$\mu_F(A;B] \;<\; \mu_F(\widetilde{A}\,;B] + \varepsilon \;\leqslant\; \sum\nolimits_1^N \mu_F(a_j;\tilde{b}_j] + \varepsilon \;<$$
$$<\; \sum\nolimits_1^\infty \mu_F(a_j;b_j] + \sum\nolimits_1^\infty \frac{\varepsilon}{2^j} + \varepsilon \;=\; \sum\nolimits_1^\infty \mu_F(a_j;b_j] + 2\varepsilon.$$

Полагая здесь $\varepsilon \to 0$, получаем искомое неравенство, а вместе с ним доказательство σ-полуаддитивности и σ-аддитивности μ_F.

Борелевская σ-алгебра $\mathscr{B}$ порождается полукольцом интервалов $\mathcal{K}$; из следствия 476, стр. 433, к теореме Каратеодори следует, что на σ-алгебре $\mathscr{B}$ существует единственная мера μ, для которой $\mu(a;b] = \mu_F(a;b] = F(b) - F(a)$.

Определения. Пусть функция $F : \mathbb{R}^1 \mapsto \mathbb{R}^1$ всюду непрерывна справа и не убывает. Мера μ_F на борелевской σ-алгебре $\mathscr{B}(\mathbb{R}^1)$, для которой $\mu_F(a;b] = F(b) - F(a)$, называется мерой *Лебега–Стилтьеса.*

Мера Лебега–Стилтьеса λ с $\lambda(a;b] = b - a$ называется *мерой Лебега.* Множества, измеримые относительно внешней меры Лебега λ^*, называются *измеримыми по Лебегу* (лебеговскими).

Как следует из предыдущего (см. стр. 436), класс измеримых по Лебегу множеств совпадает с пополнением борелевской σ-алгебры по мере Лебега, т.е. для любого такого множества B найдутся борелевские подмножества $A \subset B \subset C$ такие, что $\lambda(C \setminus A) = 0$.

⚠ Иногда эти названия закрепляют за полными мерами на подмножествах, измеримых относительно соответствующих внешних мер. Меры, определённые на борелевской σ-алгебре, называют мерами Бореля–Стилтьеса и Бореля–Лебега, а сужение меры Лебега на класс измеримых по Жордану множеств называют мерой Жордана. Мы не будем столь педантичными и, отдавая дань вкладу Бореля и Жордана, остановимся всё же на указанных здесь наименованиях.

⚠ Ничто не мешает нам определить меру Лебега–Стилтьеса с помощью непрерывной *слева* функции F на полукольце интервалов вида $[a;b)$ по той же формуле $\mu_F[a;b) = F(b) - F(a)$. Следует сказать, что ни по богатству научных результатов, ни по степени трудоемкости их доказательства эти два способа определения меры ничем друг от друга не отличаются. Приверженность тому или иному способу объясняется исключительно привычками авторов публикаций, например принадлежностью к той или иной научной школе. Заметим, что деление на два мирных клана никак не связано с их географическим местоположением.

Указанный здесь способ выбран по нескольким причинам. Первая, лежащая в плоскости филологии (а посему не бесспорная), проистекает из того, что большинство людей фразу «с часу до двух» воспринимает скорее как интервал $(1;2]$ или, в крайнем случае, как $[1;2]$, но никак не интервал $[1;2)$. Вторая причина, вытекающая, по-видимому, из первой, объясняется тем, что почти во всех современных компьютерных программах математического характера (Excel, Mathematica и т.п.) реализован способ именно с непрерывной справа функцией F. Наконец, третья, более существенная, причина связана с тем, что в многомерном случае краткая запись $\vec{x} \leqslant \vec{y}$, под которой подразумевается выполнение неравенства $x_j \leqslant y_j$ для каждой пары компонент векторов, полностью соответствует геометрической интерпретации такого векторного неравенства. В то же время неравенство $\vec{x} < \vec{y}$ (в геометрическом смысле) может иметь место и при некоторых совпадающих компонентах: $(0,0) < (1,0)$.

Кстати, можно вовсе не требовать непрерывности функции F и определить меру интервалов в соответствии с равенствами следующего упражнения.

487| Упр. Покажите, что в любом случае определения меры посредством функции F всюду непрерывной справа или слева справедливо:

$$\mu_F(a;b] = F(b+) - F(a+), \qquad \mu_F[a;b) = F(b-) - F(a-),$$
$$\mu_F[a;b] = F(b+) - F(a-), \qquad \mu_F(a;b) = F(b-) - F(a+),$$
$$\mu_F(\{b\}) = F(b+) - F(b-).$$

Мера μ, принимающая конечные значения на конечных интервалах, является мерой Лебега–Стилтьеса. В качестве одного из вариантов определяющей её непрерывной справа функции можно взять функцию $F(x) = \mu(0; x]$, если $x \geqslant 0$, и $F(x) = -\mu(x; 0]$, если $x \leqslant 0$. Для конечной меры μ можно положить $F(x) = \mu(-\infty; x]$. Такой приём реализован в теории вероятностей, где F называется функцией распределения. Легко понять, что все генерирующие функции одной меры Лебега–Стилтьеса на числовой прямой отличаются друг от друга на постоянное слагаемое.

✧ Считающие меры. Пусть функция $F(x) = \llbracket x \rrbracket$ равна целой части (с округлением вниз) числа x. Эта функция непрерывна справа и постоянна на участках между любыми двумя соседними целыми числами. График $F(x)$ представляет собой ступенчатую линию со ступеньками единичной высоты в целых точках. Мера Лебега–Стилтьеса, порождённая этой функцией, очевидно, может быть определена для *любого* подмножества $B \subset \mathbb{R}^1$ и равна количеству точек из множества целых чисел $\mathbb{Z}$, попавших в это подмножество: $\mu(B) = \#(B \cap \mathbb{Z})$. Другими словами, эта мера есть мера, считающая целые числа (ср. с 462, стр. 425).

Мера Дира́ка δ_{x_0}, сосредоточенная в т. x_0, порождается индикаторной функцией $F(x) = \dot{\mathrm{I}}\,(x; [x_0; \infty))$.

✧ Меры на индуцированном подпространстве. Очень часто вместо всего пространства $\mathbb{R}^1$ рассматривают его конечную часть, например отрезок $[0; 1]$. Можно предложить два способа задания меры на подмножествах $[0; 1]$. Во-первых, заметим, что борелевская σ-алгебра $\mathscr{B}([0; 1])$ совпадает с индуцированной на $[0; 1]$ борелевской σ-алгеброй $\mathscr{B}(\mathbb{R}^1)$. Поскольку этот отрезок измерим относительно борелевской σ-алгебры $\mathscr{B}(\mathbb{R}^1)$, то σ-алгебра $\mathscr{B}([0; 1]) \subset \mathscr{B}(\mathbb{R}^1)$. Поэтому мера, заданная на $\mathscr{B}(\mathbb{R}^1)$, естественным образом индуцирует (суженную) меру на борелевских подмножествах отрезка $[0; 1]$.

С другой стороны, как следует из леммы 485, стр. 439, σ-алгебра $\mathscr{B}([0; 1])$ есть минимальная σ-алгебра на $[0; 1]$, порождённая полукольцом интервалов $\mathcal{K} = \left\langle\, [0; a], (b; c], 0 \leqslant a \leqslant 1, 0 \leqslant b \leqslant c \leqslant 1 \,\right\rangle$. Как и в $\mathbb{R}^1$, на $[0; 1]$ можно построить меру Лебега–Стилтьеса с помощью конечной, непрерывной справа и неубывающей функции $F : [0; 1] \mapsto \mathbb{R}^1$, определив вначале её на элементах $\mathcal{K}$ по формулам $\mu_F[0; a] = F(a) - F(0)$, $\mu_F(b; c] = F(c) - F(b)$ и продолжив затем по теореме Каратеодори на все подмножества $\mathscr{B}([0; 1])$. Легко понять, что эта мера совпадает с сужением на $\mathscr{B}([0;1])$ меры Лебега–Стилтьеса на $\mathbb{R}^1$, порождённой функцией $\widetilde{F}(x) = F(x)$ при $x \in [0;1]$, $\widetilde{F}(x) = F(0)$ при $x \leqslant 0$, $\widetilde{F}(x) = F(1)$ при $x \geqslant 1$. Заметим, что функция $\widetilde{F}$ непрерывна в точке $x = 0$.

Мера Лебега–Стилтьеса, порождённая функцией F на $\mathbb{R}^1$ и суженная до $\mathscr{B}([0;1])$, может не совпадать с мерой Лебега–Стилтьеса, порождённой сужением функции F, если эта функция имеет разрывы на краях отрезка. Например, мера, считающая целые точки, индуцированная на отрезок $[0;1]$, так и будет считать количество целых чисел $\{0,1\}$, попавших в множества из этого отрезка. В то же время функция $F(x) = \llbracket x \rrbracket$, $x \in [0;1]$, порождает меру Лебега–Стилтьеса на $[0;1]$, учитывающую только наличие в измеряемом подмножестве точки $\{1\}$. Заметим, кстати, что индуцированная на отрезок $[0;1]$ считающая мера не может быть описана как мера Лебега–Стилтьеса в стиле предыдущего абзаца.

Чтобы избежать подобных трудностей, вместо замкнутого отрезка $[0;1]$ рассматривают иногда открытый слева интервал $(0;1]$ или открытый справа интервал $[0;1)$, если мера Лебега–Стилтьеса задаётся через непрерывную справа или соответственно непрерывную слева генерирующую функцию.

✧ **Борелевская σ-алгебра в $\mathbb{R}^k$.** Борелевскую σ-алгебру в евклидовом пространстве $\mathbb{R}^k$ (как и в $\mathbb{R}^1$) можно построить несколькими способами. Первый способ связан с тем, что евклидово пространство есть метрическое пространство (с метрикой $\rho(\vec{x},\vec{y})$) с базой топологии $\mathscr{O} = \left\langle U_r(\vec{c}),\ \ \vec{c} \in \mathbb{R}^k, 0 < r < \infty \right\rangle$ из открытых шаров $U_r(\vec{c}) = \{\vec{x} \in \mathbb{R}^k : \rho(\vec{x},\vec{c}) < r\}$. Тогда σ-алгебра $\mathscr{B}(\mathbb{R}^k) := \sigma\big(\mathscr{O}\big)$ — естественный претендент на звание борелевской. Легко понять, что эта σ-алгебра совпадает с σ-алгеброй, порождённой всеми открытыми множествами $\mathbb{R}^k$.

Второй способ восходит к ситуациям, когда рассматривается прямое произведение измеримых пространств. Пусть $(\Omega_1, \mathscr{F}_1), \ldots, (\Omega_k, \mathscr{F}_k)$ — набор измеримых пространств. Рассмотрим прямое произведение $\Omega_\times = \Omega_1 \times \ldots \times \Omega_k$ и класс (полуалгебру в $\Omega_\times$) $\mathscr{F}_\times = \mathscr{F}_1 \times \ldots \times \mathscr{F}_k$, состоящий из всевозможных произведений $B_1 \times \ldots \times B_k$ соответствующих измеримых подмножеств. Множества вида $B_1 \times \ldots \times B_k$, где $B_i \in \mathscr{F}_i$, $i = \overline{1,k}$, и хотя бы одно $B_i = \Omega_i$, принято называть цилиндрами. Сигма-алгебра

$$\mathscr{F}_1 \otimes \ldots \otimes \mathscr{F}_k \ := \ \sigma\big(\mathscr{F}_1 \times \ldots \times \mathscr{F}_k\big),$$

порождённая $\mathscr{F}_\times$, называется *прямым произведением* σ-алгебр или *цилиндрической* σ-алгеброй. Если $(\Omega_i, \mathscr{F}_i)$ — копии одного и того же пространства $(\Omega, \mathscr{F})$, то обозначают $\Omega_\times =: \Omega^k$, $\mathscr{F}_1 \otimes \ldots \otimes \mathscr{F}_k =: \mathscr{F}^{\otimes k}$ (или для краткости[‡] $\mathscr{F}^k$).

Определять меру, начиная с классов $\mathcal{O}$ и $\mathscr{F}_\times$, не всегда удобно. Рассмотрим класс $\mathcal{K}^k$ подмножеств, полученных прямым произведением элементов полукольца $\mathcal{K}$ на прямой. Этот класс есть полукольцо в $\mathbb{R}^k$, образованное параллелепипедами вида $(\vec{a}; \vec{c}] :=(a_1; c_1] \times \ldots \times (a_k; c_k]$, $-\infty < a_j \leqslant c_j < \infty$, $j = \overline{1, k}$, и σ-алгебра $\sigma(\mathcal{K}^k)$ — минимальная σ-алгебра, порождённая этим полукольцом.

488| **Теорема.** *Борелевская σ-алгебра в евклидовом пространстве $\mathbb{R}^k$ есть σ-алгебра*[‡]

$$\mathscr{B}^k := \mathscr{B}(\mathbb{R}^k) = \mathscr{B}^{\otimes k} = \sigma(\mathcal{K}^k) = \sigma(\mathcal{K}^\urcorner),$$

где $\mathcal{K}^\urcorner$ — класс «угловых» множеств вида $\langle (-\infty; c_1] \times \ldots \times (-\infty; c_k] \rangle$.

$\Rrightarrow$ Докажем сначала совпадение σ-алгебр $\mathscr{B}^{\otimes k} = \sigma(\mathcal{K}^k)$. Из включения $\mathcal{K} \times \ldots \times \mathcal{K} \subset \mathscr{B} \times \ldots \times \mathscr{B}$ следует, что

$$\sigma(\mathcal{K} \times \ldots \times \mathcal{K}) \subset \sigma(\mathscr{B} \times \ldots \times \mathscr{B}) = \mathscr{B}^{\otimes k}.$$

Покажем теперь, что для любых двух (в общем случае k) борелевских подмножеств B_1, B_2 прямое произведение $B_1 \times B_2 \in \sigma(\mathcal{K} \times \mathcal{K})$, что повлечёт $\mathscr{B}^{\otimes 2} = \sigma(B_1 \times B_2 : B_1, B_2 \in \mathscr{B}) \subset \sigma(\mathcal{K} \times \mathcal{K})$.

Если $B_1 \in \mathcal{K}$, то совокупность $\mathcal{A}$ борелевских подмножеств $A \subset \mathbb{R}^1$, для которых прямое произведение $B_1 \times A \in \sigma(\mathcal{K} \times \mathcal{K})$, совпадает с σ-алгеброй $\mathscr{B}$. Действительно, эта совокупность содержит все интервалы $A \in \mathcal{K}$. Кроме того, очевидно $\mathbb{R}^1 \in \mathcal{A}$. Поэтому если $A \in \mathcal{A}$, то $B_1 \times A^c = B_1 \times \mathbb{R}^1 \setminus B_1 \times A \in \sigma(\mathcal{K} \times \mathcal{K})$. Подобным же образом устанавливается, что $\mathcal{A}$ замкнута относительно счётных объединений. Таким образом, класс $\mathcal{A}$ есть σ-алгебра, содержащая полукольцо $\mathcal{K}$, т.е. $\mathcal{A} = \mathscr{B}$.

Аналогично показывается равенство $\mathcal{A} = \mathscr{B}$ для совокупности $\mathcal{A}$ борелевских подмножеств A, для которых прямое произведение $A \times B_2 \in \sigma(\mathcal{K} \times \mathcal{K})$. Следовательно, $B_1 \times B_2 \in \sigma(\mathcal{K} \times \mathcal{K})$.

Остальные утверждения теоремы следуют из утверждений 595, стр. 525, и следующего упражнения. $\Lleftarrow$

[‡] Это только обозначение, но никак не k-кратное декартово произведение σ-алгебр.

489| Упр. Докажите $\sigma\big(\mathcal{K}^k\big)$ -измеримость следующих множеств:

✦ одноточечных, конечных и счётных подмножеств $\mathbb{R}^k$;

✦ параллелепипедов $\mathbb{R}^k$ вида $\prod_{i=1}^{k}\lfloor a_i; c_i\rceil$ с любыми сочетаниями «открытости» ($[\cdot;\cdot]$, $[\cdot;\cdot)$, $(\cdot;\cdot)$, $(\cdot;\cdot]$) и конечности или бесконечности рёбер;

✦ открытых и замкнутых шаров;

✦ открытых и замкнутых подмножеств $\mathbb{R}^k$.

Покажите, что при $\forall\,\vec{a} < \vec{c}$ параллелепипед $(\vec{a};\vec{c}\,] \in \sigma\big(\mathcal{K}^{\urcorner}\big)$.

⚠ Различные характеризации борелевской σ-алгебры используются для описания различных объектов. Так, соотношение $\mathscr{B}^k = \sigma\big(\mathcal{K}^k\big)$ кладётся в основу построения меры Лебега–Стилтьеса, а характеризация $\mathscr{B}^k = \mathscr{B}^{\otimes k}$ позволяет описывать случайные векторы как упорядоченные наборы случайных величин.

✧ **О произведениях измеримых пространств.** При рассмотрении произведений пространств возникает вопрос корректности определения этого произведения, поскольку, формально говоря, пространство $\Omega_1\times\Omega_2\times\Omega_3$ отличается от пространства $\Omega_1\times(\Omega_2\times\Omega_3)$, т.к. элементы первого суть векторы с тремя компонентами $(\omega_1,\omega_2,\omega_3)$, а элементы второго $(\omega_1,(\omega_2,\omega_3))$ — векторы с двумя компонентами (ω_1, y_2), в которых вторая компонента y_2 $(\in \Omega_2\times\Omega_3)$ есть двухкомпонентный вектор. Понятно, что между элементами этих пространств можно установить взаимно однозначное соответствие, поэтому можно считать, что $\Omega_1\times\Omega_2\times\Omega_3 = \Omega_1\times(\Omega_2\times\Omega_3) = (\Omega_1\times\Omega_2)\times\Omega_3$. Осталось показать совпадение прямых произведений соответствующих σ-алгебр. Напомним, что прямое произведение σ-алгебр $\mathscr{F}_1\otimes\cdots\otimes\mathscr{F}_k$ есть наименьшая σ-алгебра в пространстве $\Omega_1\times\cdots\times\Omega_k$, содержащая всевозможные произведения измеримых подмножеств:

$$\mathscr{F}_1\otimes\cdots\otimes\mathscr{F}_k \overset{\text{df}}{=} \sigma\Big(\,B_1\times\cdots\times B_k :\ B_j\in\mathscr{F}_j,\ j=\overline{1,k}\ \Big) \;=\; \sigma\Big(\,\mathscr{F}_1\times\cdots\times\mathscr{F}_k\Big).$$

490| **Лемма.** *Пусть* $(\Omega_n,\mathscr{F}_n)$, $n=\overline{1,k}$, *— измеримые пространства. Тогда*

$$\bigotimes_{n=1}^{k}\mathscr{F}_n \;=\; \Big(\bigotimes_{i=1}^{m}\mathscr{F}_i\Big)\otimes\Big(\bigotimes_{j=m+1}^{k}\mathscr{F}_j\Big), \qquad 1\leqslant m<k.$$

⤃ Поскольку $\mathscr{F}_\alpha \times \mathscr{F}_\beta \subset \mathscr{F}_\alpha \otimes \mathscr{F}_\beta = \sigma\big(\mathscr{F}_\alpha \times \mathscr{F}_\beta\big)$, стало быть,

$$\mathop{\times}_{n=1}^{k} \mathscr{F}_n \;=\; \big(\mathop{\times}_{i=1}^{m} \mathscr{F}_i\big) \times \big(\mathop{\times}_{j=m+1}^{k} \mathscr{F}_j\big) \;\subset\; \big(\bigotimes_{i=1}^{m} \mathscr{F}_i\big) \times \big(\bigotimes_{j=m+1}^{k} \mathscr{F}_j\big) \;\subset$$
$$\subset\; \big(\bigotimes_{i=1}^{m} \mathscr{F}_i\big) \otimes \big(\bigotimes_{j=m+1}^{k} \mathscr{F}_j\big).$$

Обратно, произведение $\bigotimes_{i=1}^{m} \mathscr{F}_i \times \big(\mathop{\times}_{j=m+1}^{k} \Omega_i\big)$, т.е. класс подмножеств вида $A \times \big(\mathop{\times}_{j=m+1}^{k} \Omega_j\big)$, $A \in \bigotimes_{i=1}^{m} \mathscr{F}_i$, есть σ-подалгебра σ-алгебры $\bigotimes_{i=1}^{k} \mathscr{F}_i$. Поэтому

$$\bigotimes_{i=1}^{m} \mathscr{F}_i \times \bigotimes_{j=m+1}^{k} \mathscr{F}_j \;=\; \Big(\bigotimes_{i=1}^{m} \mathscr{F}_i \times \big(\mathop{\times}_{j=m+1}^{k} \Omega_j\big)\Big) \cap \Big(\big(\mathop{\times}_{i=1}^{m} \Omega_i\big) \times \bigotimes_{j=m+1}^{k} \mathscr{F}_j\Big)$$
$$\subset\; \bigotimes_{i=1}^{k} \mathscr{F}_i,$$

что доказывает лемму. ⫤

✧ **Мера Лебега–Стилтьеса в пространстве $\mathbb{R}^k$.** Как и в одномерном случае, рассмотрим непрерывную справа функцию $F(\vec{x})$, $\vec{x} \in \mathbb{R}^k$, т.е. $F(\vec{x} + \vec{\varepsilon}) \to F(\vec{x})$, $\vec{\varepsilon} \geqslant \vec{0}$, $\vec{\varepsilon} \to \vec{0}$, где неравенство $\vec{\varepsilon} \geqslant \vec{0}$ понимается покомпонентно: $\varepsilon_j \geqslant 0$, $j = \overline{1,k}$. Условие возрастания здесь должно быть усилено. Рассмотрим параллелепипед $(\vec{a};\vec{c}] = (a_1;c_1] \times \ldots \times (a_k;c_k]$ и связанные с ним разностные операторы

$$\Delta^{(j)}_{a,c} F(\vec{x}) = F(x_1,\ldots,x_{j-1},c,x_{j+1},\ldots,x_k) - F(x_1,\ldots,x_{j-1},a,x_{j+1},\ldots,x_k),$$
$$\Delta_{\vec{a},\vec{c}} F = \Delta^{(1)}_{a_1,c_1} \ldots \Delta^{(k)}_{a_k,c_k} F(\vec{x}). \tag{5}$$

Определения. Пусть функция $F : \mathbb{R}^k \mapsto \mathbb{R}^1$ непрерывна справа и $\Delta_{\vec{a},\vec{c}} F \geqslant 0$ для $\forall\, \vec{a} \leqslant \vec{c}$. Мера μ_F на борелевской σ-алгебре $\mathscr{B}^k$, для которой $\mu_F(\vec{a};\vec{c}] = \Delta_{\vec{a},\vec{c}} F$, называется *мерой Лебега–Стилтьеса.*

Мера Лебега–Стилтьеса с $F(\vec{x}) = x_1 \cdot \ldots \cdot x_k$, так что мера параллелепипеда равна его объёму: $\mu_F(\vec{a};\vec{c}] = \prod_1^k (c_j - a_j)$, называется *мерой Лебега* в $\mathbb{R}^k$.

Доказательство корректности такого определения повторяет одномерный случай (см. также стр. 38).

⚠ Функция F не обязана возрастать в «обычном смысле». Например, для $F(x_1,x_2) = x_1 x_2$ значение $F(-1,-1) > F(0,0)$, хотя $(-1,-1) < (0,0)$.

491| **Лемма.** (?!) *Если* $F(\vec{x}) = \prod_1^k F_j(x_j)$, *то мера параллелепипеда* $\Delta_{\vec{a},\vec{c}} F = \prod_1^k (F_j(c_j) - F_j(a_j))$.

492| Упр. Любая конечная мера на борелевском пространстве $(\mathbb{R}^k, \mathscr{B}^k)$ будет мерой Лебега–Стилтьеса (с функцией $F(\vec{x})$, равной мере множества $(-\infty; x_1] \times \ldots \times (-\infty; x_k])$. Покажите, что любая мера, принимающая конечные значения на конечных параллелепипедах, будет мерой Лебега–Стилтьеса.

Конечные меры, в частности меры Лебега–Стилтьеса с ограниченной порождающей функцией, обладают аппроксимирующим свойством, подобным 478, стр. 435. Напомним, что в конечномерном евклидовом пространстве $\mathbb{R}^k$ подмножество K компактно т. т. т. когда оно замкнуто и ограничено. Поэтому, кстати, все компакты $\mathbb{R}^k$ измеримы по Борелю.

493| **Лемма.** *Если мера* μ *на* $(\mathbb{R}^k, \mathscr{B}^k)$ *конечна, т.е.* $\mu(\mathbb{R}^k) < \infty$, *то для любого борелевского подмножества* $B \subset \mathbb{R}^k$ *и любого* $\varepsilon > 0$ *найдутся компакты* K, U *такие, что* $K \subset B$, $U \subset B^c$ *и* $\mu(U^c) - \varepsilon \leqslant \mu(B) \leqslant \mu(K) + \varepsilon$.

⤃ Из всех возможных определений борелевской σ-алгебры здесь нам понадобится определение её как минимальной σ-алгебры, содержащей все параллелепипеды вида $[\vec{a}; \vec{c}]$. Определим класс борелевских подмножеств $\mathcal{G} = \langle B \in \mathscr{B}^k \rangle$, удовлетворяющих утверждению леммы. В этот класс входят все замкнутые параллелепипеды $B = [\vec{a}; \vec{c}]$. Действительно, можно выбрать компакт $K = B$, и в силу непрерывности меры для $\forall \varepsilon > 0$ найдётся целое $n \geqslant 1$ такое, что открытый параллелепипед $O = (\vec{a} - \frac{1}{n}; \vec{c} + \frac{1}{n}) \supset B$ и $\mu(O) < \mu(B) + \frac{1}{2}\varepsilon$. Кроме того, т.к. мера μ конечна, то найдётся замкнутый параллелепипед $K_0 \supset O$ такой, что $\mu(\mathbb{R}^k) < \mu(K_0) + \frac{1}{2}\varepsilon$. Множество $U = K_0 \setminus O$ есть искомый компакт, поскольку U замкнуто и ограничено, $U \subset B^c$ и

$$\mu(B^c) - \mu(U) \;=\; \mu(\mathbb{R}^k) - \mu(B) - \mu(K_0) + \mu(O) \;<\; \varepsilon.$$

Покажем, что $\mathcal{G}$ есть алгебра и монотонный класс; откуда в силу утверждения I) леммы 591, стр. 523, будет следовать, что $\mathcal{G}$ – σ-алгебра, содержащая класс параллелепипедов, порождающий $\mathscr{B}^k$, и потому анонсированное в лемме свойство выполняется для всех борелевских множеств.

Из предыдущих рассуждений следует, что $\mathbb{R}^k \in \mathcal{G}$ (достаточно выбрать $K = K_0, U = \varnothing$). В силу симметрии относительно операции дополнения класс $\mathcal{G}$ вместе с каждым подмножеством B содержит и его дополнение.

Пусть теперь $B = B_1 \cup B_2$, где $B_1, B_2 \in \mathcal{G}$, $\varepsilon > 0$ и K_1, U_1, K_2, U_2 — соответствующие компакты: $\mu(U_j^c) - \varepsilon \leqslant \mu(B_j) \leqslant \mu(K_j) + \varepsilon$, $j = 1, 2$. Рассмотрим компакты $K = K_1 \cup K_2, U = U_1 \cap U_2$. Тогда, очевидно, $B_1 \cup B_2 \subset K$ и

$$\mu([B_1 \cup B_2] \setminus K) \ \leqslant \ \mu(B_1 \setminus K_1) + \mu(B_2 \setminus K_2) \ \leqslant \ 2\varepsilon.$$

Аналогично для $[B_1 \cup B_2]^c \setminus U \subset \left[B_1^c \setminus U_1\right] \bigcup \left[B_2^c \setminus U_2\right]$.

Рассмотрим монотонную последовательность $B_n \in \mathcal{G}, B_n \nearrow B = \bigcup_n B_n$. В силу непрерывности меры найдётся номер $n \geqslant 1$ такой, что $\mu(B) \leqslant \mu(B_n) + \frac{1}{2}\varepsilon$. По построению существует компакт $K \subset B_n \subset B$, удовлетворяющий условию $\mu(B_n) \leqslant \mu(K) + \frac{1}{2}\varepsilon.$ Следовательно,

$$\mu(B) \ \leqslant \ \mu(B_n) + \frac{\varepsilon}{2} \ \leqslant \ \mu(K) + \varepsilon.$$

Кроме того, для любого $n \geqslant 1$ существует компакт $U_n \subset B_n^c$ такой, что $\mu(U_n^c \setminus B_n) \leqslant \varepsilon/2^n$. Множество $U = \bigcap_n U_n$ компактно, т.к. есть пересечение компактных множеств. Очевидно, $U^c = \bigcup_n U_n^c \supset \bigcup_n B_n = B$, и

$$U^c \setminus B \ = \ \Big[\bigcup\nolimits_n U_n^c\Big] \bigcap \Big[\bigcap\nolimits_m B_m^c\Big] \ \subset \ \bigcup\nolimits_n \left[U_n^c \setminus B_n\right].$$

Поэтому $\mu(U^c \setminus B) \leqslant \sum_n \varepsilon/2^n = \varepsilon$. ⫤

✧ Пространство числовых последовательностей $\mathbb{R}^{\mathbb{N}}$. Формализация понятия последовательности функций с заданными свойствами требует введения не только конечномерных пространств, но и бесконечных произведений измеримых пространств. Здесь при построении меры удобнее не оперировать сразу всем набором пространств, а рассмотреть сначала класс так называемых конечномерных цилиндров.

Пусть $\mathbb{R}^\infty = \{\boldsymbol{x} = (x_1, x_2, \ldots) : \ x_n \in \mathbb{R}^1, n \geqslant 1\}$ — пространство всех числовых последовательностей (иногда будем использовать обозначение $\mathbb{R}^{\mathbb{N}}$). Наиболее естественную топологию открытых множеств здесь определяет так называемая тихоновская топология (см. стр. 528), база окрестностей которой состоит из всех цилиндрических множеств вида $\prod_k^\infty B_k$, где конечное число множеств $B_{k_j} = (a_j; c_j)$, $a_j, c_j \in \mathbb{R}^1$, $j = \overline{1, n}$, $n < \infty$, остальные $B_k = \mathbb{R}^1$. Тогда борелевской естественно назвать минимальную σ-алгебру $\mathscr{B}(\mathbb{R}^\infty)$, порождённую открытыми множествами в этой топологии.

Другие способы определения σ-алгебры подмножеств $\mathbb{R}^\infty$ опираются на системы множеств, полученных как прямые произведения классов множеств, определяющих борелевские σ-алгебры на координатных прямых. Пусть $B \subset \subset \mathbb{R}^k$ — подмножество k-мерного пространства. Цилиндром в $\mathbb{R}^\infty$ с основанием B будем называть множество

$$B \times \mathbb{R}^\infty_{k+1} = \{\boldsymbol{x} \in \mathbb{R}^\infty : (x_1, \ldots, x_k) \in B\},$$

т.е. набор всех последовательностей, у которых k-мерные векторы, образованные первыми k элементами, принадлежат подмножеству B, а остальные элементы произвольны. Заметим, что любой цилиндр с k-мерным основанием можно рассматривать как цилиндр с m-мерным основанием при $\forall m > k$. Действительно, если $B \subset \mathbb{R}^k$, то $B' = B \times \mathbb{R}^{m-k} \subset \mathbb{R}^m$ и $B \times \mathbb{R}^\infty_{k+1} = B' \times \mathbb{R}^\infty_{m+1}$.

Рассмотрим три класса цилиндрических множеств:

$\mathcal{B} = \left\langle B \times \mathbb{R}^\infty_{n+1} : \ B \in \mathscr{B}^n, n \geqslant 1 \right\rangle$ — алгебру (?!) цилиндров с борелевскими основаниями;

$\mathcal{C} = \left\langle (B_1 \times \ldots \times B_n) \times \mathbb{R}^\infty_{n+1} : \ \{B_j\}_1^n \subset \mathscr{B}^1, n \geqslant 1 \right\rangle$ — полуалгебру (?!) цилиндров с основаниями в виде «борелевских прямоугольников» — прямых произведений одномерных борелевских подмножеств;

$\mathcal{K} = \left\langle (\vec{a}; \vec{c}] \times \mathbb{R}^\infty_{n+1} : \ \vec{a}, \vec{c} \in \mathbb{R}^n, n \geqslant 1 \right\rangle$ — класс цилиндров с параллелепипедами в основании; этот класс будет полукольцом, если к нему добавить цилиндры, в основании которых есть интервалы вида $(-\infty; c], (a; +\infty)$.

494| **Теорема.** *Борелевская σ-алгебра в $\mathbb{R}^\infty$*

$$\mathscr{B}^\infty := \mathscr{B}(\mathbb{R}^\infty) = \sigma(\mathcal{B}) = \sigma(\mathcal{C}) = \sigma(\mathcal{K}).$$

⤇ Так как, очевидно, $\sigma(\mathcal{B}) \supset \sigma(\mathcal{C}) \supset \sigma(\mathcal{K})$, то для доказательства совпадения этих трех σ-алгебр достаточно показать, что $\sigma(\mathcal{B}) \subset \sigma(\mathcal{K})$. Пусть $\mathcal{K}^{(n)}$ — класс n-мерных параллелепипедов. Так как борелевская σ-алгебра $\mathscr{B}^n = \sigma\big(\mathcal{K}^{(n)}\big)$, то минимальная σ-алгебра, порождённая классом цилиндров с n-мерными параллелепипедами в основании, $\sigma\big(\mathcal{K}^{(n)} \times \mathbb{R}^\infty_{n+1}\big) = \mathscr{B}^n \times \mathbb{R}^\infty_{n+1}$. Следовательно, для любого n-мерного подмножества $B \in \mathscr{B}^n$ справедливо включение $B \times \mathbb{R}^\infty_{n+1} \in \sigma\big(\mathcal{K}^{(n)} \times \mathbb{R}^\infty_{n+1}\big) \subset \sigma(\mathcal{K})$, что и требовалось. ⤆

495| Упр. Покажите, что $\mathscr{B}(\mathbb{R}^\infty) = \sigma(\mathcal{K})$.

✧ **Вероятностная мера в пространстве $\mathbb{R}^\infty$.** В предыдущем разделе σ-алгебра в пространстве $\mathbb{R}^\infty$ всех числовых последовательностей $\boldsymbol{x} = (x_1, x_2, \ldots)$ определена на основе классов цилиндрических подмножеств с «конечномерными» основаниями. Связано это не только с тем, что эта σ-алгебра совпадает с борелевской σ-алгеброй $\mathscr{B}^\infty$, порождённой открытыми множествами в тихоновской топологии. В большей степени это обусловлено процессом построения меры на подмножествах $\mathbb{R}^\infty$. Любой такой процесс начинается с определения меры так называемых конечномерных подмножеств, т.е. подмножеств $\boldsymbol{B} \subset \mathbb{R}^\infty$, для которых найдётся конечномерное (k-мерное) борелевское подмножество $B \in \mathscr{B}^k$ такое, что $\boldsymbol{B} = B \times \mathbb{R}^\infty_{k+1} := \{\boldsymbol{x} \in \mathbb{R}^\infty : (x_1, \ldots, x_k) \in B\}$. Дабы избежать ненужных здесь сложностей и построить меру, которая не равна тождественно $+\infty$ (или нулю), будем осуществлять это построение для вероятностных мер $\mathbf{P}$, т.е. мер, удовлетворяющих условию $\mathbf{P}\{\mathbb{R}^\infty\} = 1$.

Если какая-то мера $\mathbf{P}$ задана на всех подмножествах $\mathscr{B}^\infty$, то, выбирая подмножества вида $B \times \mathbb{R}^\infty_{k+1}$, где $B \in \mathscr{B}^k$, можно определить меру для всех борелевских подмножеств $\mathbb{R}^k$ по формуле

$$\mathrm{P}_k\{B\} \;=\; \mathbf{P}\left\{B \times \mathbb{R}^\infty_{k+1}\right\}. \tag{6}$$

Поскольку $B \times \mathbb{R}^\infty_{k+1} = (B \times \mathbb{R}^1) \times \mathbb{R}^\infty_{k+2}$, то семейство (по k) таких мер, очевидно, удовлетворяет так называемому условию *согласованности*:

$$\mathrm{P}_{k+1}\{B \times \mathbb{R}^1\} \;=\; \mathrm{P}_k\{B\}, \qquad \forall\, k \in \mathbb{N}, B \in \mathscr{B}^k. \tag{7}$$

Это условие — единственное условие к семейству конечномерных вероятностных мер, при котором существует вероятностная мера на борелевских подмножествах $\mathbb{R}^\infty$, ограничения которой на конечномерные пространства совпадают с соответствующими элементами семейства.

496| **Теорема.** [*О продолжении согласованных мер. Колмогоров.*] *Пусть семейство вероятностных мер* $\left\langle \mathrm{P}_k\{B\}, B \in \mathscr{B}^k \right\rangle_{k=1}^\infty$ *на борелевских подмножествах конечномерных евклидовых пространств* $\mathbb{R}^k$, $k \geqslant 1$, *удовлетворяет условию согласованности* (7). *Тогда на измеримом пространстве* $(\mathbb{R}^\infty, \mathscr{B}^\infty)$ *найдётся единственная вероятностная мера* $\mathbf{P}$, *для которой справедливо* (6).

⊫ Определим меру $\mathbf{P}$ на цилиндрических подмножествах пространства $\mathbb{R}^\infty$, в основании которых лежат борелевские подмножества конечномерных пространств $\mathbb{R}^k$, $k \geqslant 1$, по формуле (6). Условие согласованности гарантирует корректность такого определения: при любых способах описания одного и того же множества с основаниями из k-мерного или $(k+n)$-мерного ($n \geqslant 1$) борелевского пространства мера этого множества будет одинакова.

Класс указанных цилиндрических множеств образует кольцо (алгебру) подмножеств $\mathbb{R}^\infty$, поэтому сигма-аддитивность этой меры будет гарантирована, если показать её аддитивность и непрерывность в «нуле» (см. 468, стр. 430). Пусть $B_1^{(k)}\times\mathbb{R}^\infty_{k+1}$, $B_2^{(m)}\times\mathbb{R}^\infty_{m+1}$ — два непересекающихся цилиндрических подмножества с $B_1^{(k)}\in\mathscr{B}^k, B_2^{(m)}\in\mathscr{B}^m, k<m$ (случай $k=m$ разбирается аналогично). Очевидно, подмножество $B_1^{(k)}\times\mathbb{R}^{m-k}\in\in\mathscr{B}^m$ и не пересекается с $B_2^{(m)}$. Поэтому $\big[B_1^{(k)}\times\mathbb{R}^\infty_{k+1}\big]\uplus\big[B_2^{(m)}\times\mathbb{R}^\infty_{m+1}\big]==\big[\big(B_1^{(k)}\times\mathbb{R}^{m-k}\big)\uplus B_2^{(m)}\big]\times\mathbb{R}^\infty_{m+1}$ и

$$\mathbf{P}\Big\{\big[B_1^{(k)}\times\mathbb{R}^\infty_{k+1}\big]\biguplus\big[B_2^{(m)}\times\mathbb{R}^\infty_{m+1}\big]\Big\} = \mathrm{P}_m\Big\{\big(B_1^{(k)}\times\mathbb{R}^{m-k}\big)\biguplus B_2^{(m)}\Big\} =$$

$$= \mathrm{P}_m\big\{B_1^{(k)}\times\mathbb{R}^{m-k}\big\}+\mathrm{P}_m\big\{B_2^{(m)}\big\} = \mathbf{P}\big\{B_1^{(k)}\times\mathbb{R}^\infty_{k+1}\big\}+\mathbf{P}\big\{B_2^{(m)}\times\mathbb{R}^\infty_{m+1}\big\}$$

в силу аддитивности меры P_m. Следовательно, мера $\mathbf{P}$ аддитивна и потому монотонна: $\boldsymbol{B}\subset\boldsymbol{A}\Rightarrow\mathbf{P}\{\boldsymbol{B}\}\leqslant\mathbf{P}\{\boldsymbol{A}\}$.

Для доказательства непрерывности меры рассмотрим произвольную монотонно убывающую последовательность $\big\langle\boldsymbol{B}_n=B_n^{(d_n)}\times\mathbb{R}^\infty_{d_n+1}\big\rangle_{n=1}^\infty$ цилиндрических множеств, пересечение которых пусто. Будем считать размерность основания $B_n^{(d_n)}$ цилиндрического множества $d_n=n$ (см. по этому поводу замечание 582, стр. 510), в связи с чем можно положить $B_n^{(d_n)}=B_n\ (\in\mathbb{R}^n)$. Предположим, от противного, что $\mathbf{P}\big\{\boldsymbol{B}_n\big\}>2\varepsilon$ для некоторого $\varepsilon>0$ при $\forall\, n\geqslant1$.

По лемме 493 для любого борелевского множества B_n в пространстве $\mathbb{R}^n$ найдётся такое компактное подмножество $C_n\subset B_n$, что $\mathrm{P}_n\{B_n\setminus C_n\}\leqslant\leqslant\varepsilon/2^n$. Определим множество $K_n=C_n\bigcap_1^{n-1}\big(C_j\times\mathbb{R}^{n-j}\big)\subset\mathbb{R}^n$. Это множество компактно, т.к. оно ограничено ($K_n\subset C_n$) и замкнуто, поскольку этими свойствами обладают компакты $C_j,\ j=\overline{1,n}$. Для множества $\boldsymbol{K}_n==K_n\times\mathbb{R}^\infty_{n+1}=\bigcap_1^n\big(C_j\times\mathbb{R}^\infty_{j+1}\big)\subset\boldsymbol{B}_n$ справедливо

$$\mathbf{P}\{\boldsymbol{B}_n\}-\mathbf{P}\{\boldsymbol{K}_n\} = \mathbf{P}\{\boldsymbol{B}_n\setminus\boldsymbol{K}_n\} \leqslant \sum_{j=1}^n\mathbf{P}\Big\{\boldsymbol{B}_n\setminus\big(C_j\times\mathbb{R}^\infty_{j+1}\big)\Big\} \leqslant$$

$$\leqslant \sum_{j=1}^n\mathbf{P}\Big\{\boldsymbol{B}_j\setminus\big(C_j\times\mathbb{R}^\infty_{j+1}\big)\Big\} = \sum_{j=1}^n\mathrm{P}_n\{B_j\setminus C_j\} \leqslant \sum_1^n\frac{\varepsilon}{2^j} < \varepsilon.$$

Поскольку $\mathbf{P}\big\{\boldsymbol{B}_n\big\}>2\varepsilon$, получаем, что $\mathbf{P}\{\boldsymbol{K}_n\}>\varepsilon$ при $\forall\, n\geqslant1$.

Таким образом, множества $\boldsymbol{K}_n\neq\varnothing,\ \forall\, n\geqslant1$, и последовательность этих множеств монотонно убывает. В соответствии с леммой 497 пересечение $\bigcap_1^\infty\boldsymbol{K}_n\neq\varnothing$, что противоречит условию $\bigcap_1^\infty\boldsymbol{B}_n=\varnothing$. Следовательно, мера $\mathbf{P}$ на кольце цилиндрических множеств $\mathbb{R}^\infty$ сигма-аддитивна. По теореме Каратеодори её можно единственным образом продолжить на σ-алгебру, порождённую этим кольцом, т.е. на σ-алгебру $\mathscr{B}^\infty$. ⊲

В приведённом доказательстве теоремы Колмогорова существенно использовалось следующее утверждение.

497| **Лемма.** *Пусть* $\left\langle K_n \times \mathbb{R}^\infty_{n+1} \right\rangle_1^\infty$ *— монотонное семейство цилиндрических множеств, в основании которых лежат непустые компактные подмножества* n*-мерных евклидовых пространств:* $K_n \subset \mathbb{R}^n,\ K_n \times \mathbb{R}^1 \supset K_{n+1},\ n \geqslant 1$. *Тогда* $\bigcap_1^\infty (K_n \times \mathbb{R}^\infty_{n+1}) \neq \varnothing$.

$\Rrightarrow$ Так как множества $K_n \times \mathbb{R}^\infty_{n+1}$ не пусты, то в каждом таком множестве можно выбрать по одной точке (последовательности) $\boldsymbol{x}^{(n)} = (x_{1n}, x_{2n}, \ldots)$. В силу монотонности указанного семейства все числа $x_{1n} \in K_1$. Поскольку K_1 — компакт, из последовательности x_{1n} можно выделить сходящуюся подпоследовательность $x_{1n_1} \to X_1 \in K_1$. Аналогично, из последовательности $(x_{1n_1}, x_{2n_1}) \in K_2$ можно выделить сходящуюся подпоследовательность $(x_{1n_2}, x_{2n_2}) \to (X_1, X_2) \in K_2$. Продолжая подобным образом, получаем после применения стандартного приёма выделения диагональной подпоследовательности, что существует элемент $\boldsymbol{x}^{(0)} = (X_1, X_2, \ldots)$ такой, что $(X_1, \ldots, X_n) \in K_n$ при $\forall\, n \geqslant 1$, стало быть, $\boldsymbol{x}^{(0)} \in \bigcap_1^\infty (K_n \times \mathbb{R}^\infty_{n+1})$. $\Lleftarrow$

При доказательстве теоремы Колмогорова никак не использовалась конкретная структура оснований цилиндрических множеств $B \times \mathbb{R}^\infty_{k+1}$, важно было лишь то, что класс таких множеств образует кольцо, порождающее цилиндрическую σ-алгебру. Это позволило, с одной стороны, воспользоваться эквивалентностью свойств сигма-аддитивности и непрерывности (в нуле) для функции множеств, а с другой стороны — применить теорему Каратеодори о продолжении меры с кольца на порождённую σ-алгебру. Из этих рассуждений следует, что семейство конечномерных вероятностных мер может быть определено на более узком, чем борелевская σ-алгебра, классе подмножеств $\mathbb{R}^k$, лишь бы класс соответствующих цилиндирческих множеств образовывал кольцо или полукольцо, порождающее цилиндрическую σ-алгебру (см. теорему 494). Стало быть, справедлива

498| **Теорема.** [*О продолжении согласованных мер.*] *Пусть задано семейство вероятностных мер* $\left\langle \mathrm{P}_k\{B\}, B \in \mathcal{C}_k\ (\mathcal{K}^k) \right\rangle_{k=1}^\infty$, *каждая из которых определена либо на полуалгебре* $\mathcal{C}_k$ *борелевских прямоугольников, либо на полукольце* $\mathcal{K}^k$ *параллелепипедов с конечными или бесконечными границами в конечномерном евклидовом пространстве* $\mathbb{R}^k,\ k \geqslant 1$. *Если это семейство удовлетворяет условию согласованности* (7), *то на измеримом пространстве*

$(\mathbb{R}^\infty, \mathscr{B}^\infty)$ *найдётся единственная вероятностная мера* $\mathbf{P}$, *для которой справедливо* (6).

Конечная мера в конечномерном пространстве была описана как мера Лебега–Стилтьеса μ_F (см. стр. 447), порождённая непрерывной справа функцией F, удовлетворяющей при $\forall\,\vec{a} \lessdot \vec{c}$ условию $\Delta_{\vec{a},\vec{c}}F(\vec{x}) \geqslant 0$ на разностный оператор (5). У вероятностных мер функция F называется функцией распределения. Теорему Колмогорова можно переформулировать в терминах этих функций.

499| **Теорема.** *Пусть* $\left\langle F_k(\vec{x}),\ \vec{x} \in \mathbb{R}^k \right\rangle_{k=1}^{\infty}$ — *семейство функций распределения в конечномерных евклидовых пространствах, удовлетворяющее условию согласованности*

$$F_{k+1}(x_1, \ldots, x_k, +\infty) \;=\; F_k(x_1, \ldots, x_k).$$

Тогда на измеримом пространстве $(\mathbb{R}^\infty, \mathscr{B}^\infty)$ *существует единственная вероятностная мера* $\mathbf{P}$, *для которой при* $\forall\, k \geqslant 1, \vec{a}, \vec{c} \in \mathbb{R}^k$

$$\mathbf{P}\left\{ (\vec{a}; \vec{c}] \times \mathbb{R}^\infty_{k+1} \right\} \;=\; \Delta_{\vec{a},\vec{c}} F(\vec{x}). \tag{8}$$

⊨ Заметим, что каждая функция распределения F_k задаёт соответствующую меру Лебега–Стилтьеса μ_k $(= \mu_{F_k})$ в k-мерном пространстве $\mathbb{R}^k$. Рассмотрим подмножества $\mathbb{R}^{k+1}$ вида $B \times \mathbb{R}^1$, где $B \in \mathscr{B}^k$. В силу условия согласованности мера $\mu_{k+1}(B \times \mathbb{R}^1) = \mu_k(B)$ для всех множеств вида $B = \prod_{j=1}^{k}(-\infty; x_j]$. Класс таких множеств замкнут относительно пересечений, поэтому это равенство верно для $\forall\, B \in \mathscr{B}^k$ (см. стр. 35). Другими словами, семейство мер Лебега–Стилтьеса $\left\langle \mu_k(B),\ B \in \mathscr{B}^k \right\rangle$ удовлетворяет условию согласованности, и, следовательно, по теореме Колмогорова это семейство определяет единственную вероятностную меру на $(\mathbb{R}^\infty, \mathscr{B}^\infty)$, для которой выполняется равенство (8). ⫤

✧ **Цилиндрическая σ-алгебра в пространстве функций $\mathbb{R}^{\mathbb{T}}$.** Потребности изучения изменчивости величин во времени (или в пространстве) диктуют необходимость рассмотрения элементов, принимающих значения в пространстве функций. Пусть $\mathbb{T}$ — некоторое множество параметров t. Чаще всего t интерпретируется как время (момент времени), прошедшее от начала наблюдения над процессом $(\mathbb{T} \subset \mathbb{R}^1_+)$, или координата пространства $(\mathbb{T} \subset \mathbb{R}^k)$, в зависимости от которой изменяется исследуемая характеристика (например, высота снежного покрова). Пространство всех функций $X(=X_t) : \mathbb{T} \mapsto \mathbb{R}^1$, заданных на $\mathbb{T}$ и принимающих значения в $\mathbb{R}^1$, обозначается как $\mathbb{R}^{\mathbb{T}} = \prod_{t\in\mathbb{T}} \mathbb{R}^1$ (сравните с пространством $\mathbb{R}^k$, когда $\mathbb{T} = \{1, 2, \dots, k\}$, и пространством последовательностей $\mathbb{R}^\infty$ — здесь $\mathbb{T} = \mathbb{N}$). Удобно считать сомножители, входящие в $\mathbb{R}^{\mathbb{T}}$, различными, для чего снабдим их соответствующим индексом: $\mathbb{R}^{\mathbb{T}} = \prod\limits_{t\in\mathbb{T}} \mathbb{R}^1_t$.

Определение борелевской σ-алгебры опирается на соответствующую топологию открытых множеств. Базис окрестностей тихоновской топологии $\mathbb{R}^{\mathbb{T}}$ (стр. 528) образуют все цилиндрические множества вида $B_{t'} \times \prod_{t\neq t'} \mathbb{R}^1_t = \{X \in \mathbb{R}^{\mathbb{T}} : X_{t'} \in B_{t'}\}$, $t' \in \mathbb{T}, B_{t'} \subset \mathbb{R}^1$, т.е. множества функций, у которых значения в одной точке t' ограничены открытым подмножеством $B_{t'} \subset \mathbb{R}^1$, а в остальном произвольны. Визуально любой базисный элемент тихоновской топологии можно описать как совокупность всех функций, проходящих в момент времени $t = t'$ через «ворота», задаваемые открытым подмножеством $B_{t'}$. Отметим, что непрерывность отображений, принимающих значения в $\mathbb{R}^{\mathbb{T}}$, в тихоновской топологии эквивалентна покоординатной непрерывности, т.е. отображение $\xi : \Omega \mapsto \mathbb{R}^{\mathbb{T}}$ непрерывно т. т. т. когда для $\forall t \in \mathbb{T}$ непрерывно отображение $\xi(\omega)_t : \Omega \mapsto \mathbb{R}^1$. Минимальную σ-алгебру, порождённую тихоновской топологией, хотелось бы назвать борелевской в $\mathbb{R}^{\mathbb{T}}$. К сожалению, эта σ-алгебра для наших целей слишком широка. Это связано с тем, что обычный способ построения меры начинается с некоторого класса подмножеств и продолжается затем на минимальную σ-алгебру, порождённую этим классом. В пространстве $\mathbb{R}^{\mathbb{T}}$ в качестве такого класса мы можем взять только базу топологии. При несчётном $\mathbb{T}$ σ-алгебра, порождённая базой топологии $\mathbb{R}^{\mathbb{T}}$, не будет совпадать с σ-алгеброй, порождённой всей топологией. Указанное совпадение σ-алгебр имеет место для полных сепарабельных метрических пространств (так называемых польских пространств), в которых топология может быть задана счётной базой. При несчётном множестве индексов $\mathbb{T}$ для тихоновской топологии в $\mathbb{R}^{\mathbb{T}}$ это несправедливо.

Описанный процесс построения меры приводит к определению так называемой цилиндрической σ-алгебры. Пусть $\tau = (t_1, \dots, t_k)$ — набор различных индексов из параметрического множества $\mathbb{T}$ с фиксированным порядком записи (т.е., к примеру, $(1, 2) \neq (2, 1)$). С этим набором свяжем k-мерное борелевское пространство $(\mathbb{R}^\tau, \mathscr{B}^\tau) = (\prod\limits_{t\in\tau} \mathbb{R}_t, \bigotimes\limits_{t\in\tau} \mathscr{B}_t)$ (с учётом договорённости, конечно, $(\mathbb{R}^\tau, \mathscr{B}^\tau) = (\mathbb{R}^k, \mathscr{B}^k)$). Назовём цилиндром (цилиндрическим множеством) с основанием $B \in \mathscr{B}^\tau$ подмножество пространства $\mathbb{R}^{\mathbb{T}}$ вида

$$\boldsymbol{B}_\tau \;=\; B \times \prod\nolimits_{t \notin \tau} \mathbb{R}_t \;=\; B \times \mathbb{R}^{\mathbb{T} \setminus \tau} \;=\; \{X \in \mathbb{R}^{\mathbb{T}} :\; (X_{t_1}, \ldots, X_{t_k}) \in B\},$$

т.е. совокупность всех функций, у которых k-мерный вектор значений в моменты $t_1, \ldots, t_k$ попадает в множество B; в остальном значения функций произвольны.

500| ⚠ Из определения видно, что при различных наборах параметров $\tau \neq \sigma$ цилиндры $\boldsymbol{B}_\tau$ и $\boldsymbol{C}_\sigma$ могут не совпадать, даже если они имеют одинаковые основания $B = C$. Например, цилиндры (с $\tau = (4, 2)$, $B = (a; c] \times (b; d]$)

$$\begin{aligned} B \times \mathbb{R}^{\mathbb{T} \setminus \tau} &= \{X \in \mathbb{R}^{\mathbb{T}} :\; a < X_4 \leqslant c, b < X_2 \leqslant d\}, \\ B \times \mathbb{R}^{\mathbb{T} \setminus \sigma} &= \{X \in \mathbb{R}^{\mathbb{T}} :\; a < X_{s_1} \leqslant c, b < X_{s_2} \leqslant d\}, \end{aligned}$$

вообще говоря, различны, если $s_1 \neq 4$ или $s_2 \neq 2$. В частности, будут различаться цилиндры $B \times \mathbb{R}^{\mathbb{T} \setminus \tau}$ и $\big[\pi B\big] \times \mathbb{R}^{\mathbb{T} \setminus \tau}$, если πB получено какой-либо перестановкой π сомножителей B: $\pi B = (b; d] \times (a; c]$. Однако подобные два цилиндра будут совпадать, если произвести одновременно аналогичную перестановку компонент набора параметров τ (например, $\pi\tau = (2, 4)$):

$$B \times \mathbb{R}^{\mathbb{T} \setminus \tau} \;=\; \big[\pi B\big] \times \mathbb{R}^{\mathbb{T} \setminus \pi\tau}.$$

Также из определения ясно, что цилиндр не изменится, если в его записи к вектору параметров добавить (в любом месте) любое количество параметров с одновременным добавлением на соответствующих местах в записи основания сомножителей, совпадающих со всем пространством $\mathbb{R}^1$. Например,

$$\big[(a; c] \times (b; d]\big] \times \mathbb{R}^{\mathbb{T} \setminus \tau} \;=\; \widehat{B} \times \mathbb{R}^{\mathbb{T} \setminus \hat{\tau}},$$

где $\tau = (4, 2)$, $\hat{\tau} = (0, 4, 7, 2, 1.5)$, $\widehat{B} = \mathbb{R} \times (a; c] \times \mathbb{R} \times (b; d] \times \mathbb{R}$.

Как всегда, множество $B = (\vec{a}; \vec{c}] \subset \mathbb{R}^k$ будем называть (k-мерным) параллелепипедом, а множество $B = B_1 \times \cdots \times B_k$, где $B_j \in \mathscr{B}^1$, $j = \overline{1, k}$, — (k-мерным) борелевским прямоугольником.

Легко видеть, что класс всех цилиндров $\mathcal{B}_{\mathbb{T}} = \big\langle \boldsymbol{B}_\tau,\; \tau = (t_1, \ldots, t_k) \subset \subset \mathbb{T},\; B \in \mathscr{B}^\tau,\; k \geqslant 1 \big\rangle$ образует алгебру подмножеств $\mathbb{R}^{\mathbb{T}}$.

Определение. *Цилиндрической σ-алгеброй* $\mathscr{B}^{\mathbb{T}} = \bigotimes\limits_{t \in \mathbb{T}} \mathscr{B}_t$ в пространстве функций $\mathbb{R}^{\mathbb{T}}$ называется минимальная σ-алгебра, порождённая классом цилиндров $\mathcal{B}_{\mathbb{T}}$: $\mathscr{B}^{\mathbb{T}} = \sigma\big(\mathcal{B}_{\mathbb{T}}\big)$.

Рассмотрим произвольное подмножество параметров $\mathbb{S} \subset \mathbb{T}$ и соответствующее ему измеримое пространство $(\mathbb{R}^{\mathbb{S}}, \mathscr{B}^{\mathbb{S}})$ с цилиндрической σ-алгеброй. Легко понять, что класс

$$\mathscr{B}^{\mathbb{S}} \times \mathbb{R}^{\mathbb{T}\setminus\mathbb{S}} := \left\langle B \times \mathbb{R}^{\mathbb{T}\setminus\mathbb{S}} : \; B \in \mathscr{B}^{\mathbb{S}} \right\rangle$$

образует σ-подалгебру σ-алгебры $\mathscr{B}^{\mathbb{T}}$. Следующая теорема говорит, что цилиндрическая σ-алгебра $\mathscr{B}^{\mathbb{T}}$ исчерпывается, в некотором смысле, σ-подалгебрами $\mathscr{B}^{\mathbb{S}}$ со счётным набором $\mathbb{S}$.

501| **Теорема.** *Пусть $\mathbb{T}$ — несчётное множество. Цилиндрическая σ-алгебра $\mathscr{B}^{\mathbb{T}}$ совпадает с объединением всех σ-подалгебр $\mathscr{B}^{\mathbb{S}} \times \mathbb{R}^{\mathbb{T}\setminus\mathbb{S}}$ со счётным набором параметров $\mathbb{S} \subset \mathbb{T}$:* $\mathscr{B}^{\mathbb{T}} = \bigcup_{\mathbb{S}\subset\mathbb{T}} \left[\mathscr{B}^{\mathbb{S}} \times \mathbb{R}^{\mathbb{T}\setminus\mathbb{S}}\right]$.

$\Rrightarrow$ Обозначим правую часть доказываемого равенства через $\mathscr{U}$. Покажем, что $\mathscr{U}$ образует σ-алгебру. Выберем произвольные подмножества $A_j \in \mathscr{U}$, $j = 1, 2 \ldots$ Для каждого A_j найдётся счётный набор индексов $\mathbb{S}_j$ такой, что $A_j \in \mathscr{B}^{\mathbb{S}_j} \times \mathbb{R}^{\mathbb{T}\setminus\mathbb{S}_j}$. Поскольку счётное объединение счётных классов также счётно, то набор $\mathbb{S} = \bigcup_j \mathbb{S}_j$ содержит счётное число индексов. Кроме того, как уже отмечалось выше, каждое подмножество $A_j \in \in \mathscr{B}^{\mathbb{S}} \times \mathbb{R}^{\mathbb{T}\setminus\mathbb{S}}$. Следовательно, $\bigcup_j A_j \in \mathscr{B}^{\mathbb{S}} \times \mathbb{R}^{\mathbb{T}\setminus\mathbb{S}} \subset \mathscr{U}$. Принадлежность к $\mathscr{U}$ пустого множества и всего пространства $\mathbb{R}^{\mathbb{T}}$, а также замкнутость $\mathscr{U}$ относительно операции дополнения очевидны.

Далее, любой конечномерный цилиндр есть элемент одной из σ-подалгебр $\mathscr{B}^{\mathbb{S}} \times \mathbb{R}^{\mathbb{T}\setminus\mathbb{S}}$, поэтому $\mathscr{U}$ содержит все цилиндрические множества $\mathbb{R}^{\mathbb{T}}$. Так как $\mathscr{B}^{\mathbb{T}}$ есть минимальная σ-алгебра, содержащая такие множества, и, очевидно, $\mathscr{U} \subset \mathscr{B}^{\mathbb{T}}$, то $\mathscr{U} = \mathscr{B}^{\mathbb{T}}$. $\Lleftarrow$

Таким образом, множество, измеримое относительно цилиндрической σ-алгебры, представляет собой семейство функций на $\mathbb{T}$, свойства которых описываются не более чем счётным набором параметров $(t_1, t_2, \ldots) \subset \mathbb{T}$. Поэтому подмножества функций $X \in \mathbb{R}^{\mathbb{T}}$, свойства которых могут быть описаны лишь посредством значений аргумента t, пробегающего несчётное подмножество $T_0 \subset \mathbb{T}$, скорее всего, будут неизмеримы. К таковым относятся, например, множества, порождаемые утверждениями о значении максимума типа $\left\langle X : \sup_t X_t \leqslant 1 \right\rangle$ или утверждениями о непрерывности, дифференцируемости X_t и т.п.

По теореме 494 борелевская σ-алгебра в пространстве $\mathbb{R}^{\mathbb{N}}$ может быть определена не только через класс цилиндров, в основании которых лежат конечномерные борелевские множества, но и через класс цилиндров с основаниями в виде параллелепипедов или в виде борелевских прямоугольников. Это утверждение, а также утверждение теоремы 501 позволяют сформулировать

502| **Следствие.** *Цилиндрическая σ-алгебра $\mathscr{B}^{\mathbb{T}}$ в пространстве функций $\mathbb{R}^{\mathbb{T}}$ совпадает с минимальной σ-алгеброй, порождённой цилиндрами с конечномерными параллелепипедами в основании, а также с минимальной σ-алгеброй, порождённой цилиндрами с конечномерными борелевскими прямоугольниками в основании.*

✧ **Вероятность в пространстве $\mathbb{R}^{\mathbb{T}}$.** Как и в пространстве последовательностей, для описания вероятностной меры в пространстве функций $\mathbb{R}^{\mathbb{T}}$ необходимо согласованное семейство конечномерных вероятностных мер. В данном случае условие согласованности удобнее формулировать для мер, заданных на цилиндрах с борелевскими прямоугольниками в основании, т.е. с основаниями вида $B^{(k)} := B_1 \times \cdots \times B_k$, где борелевские множества $B_j \subset \mathbb{R}^1$, $j = \overline{1,k}$. С каждой перестановкой k натуральных чисел $\pi(1,\ldots,k) = (j_1,\ldots,j_k)$ можно связать «переставленный» прямоугольник $\pi B^{(n)} = B_{j_1} \times \cdots \times B_{j_k}$.

Обозначим через $\mathbb{T}^{*k}$ совокупность всех k-векторов с фиксированным порядком записи $\tau = (t_1,\ldots,t_k)$, $t_j \in \mathbb{T}$, и неповторяющимися координатами: $t_j \neq t_n$, $j \neq n$. Предположим, что для каждого $k \geqslant 1$ и каждого $\tau \in \mathbb{T}^{*k}$ задана вероятностная мера $\mathrm{P}(B;\tau)$, $B \in \mathscr{B}^k$, в k-мерном пространстве $\mathbb{R}^k$. Будем говорить, что семейство $\mathcal{P} = \left\langle \mathrm{P}(\cdot\,;\tau),\ \tau \in \bigcup_k \mathbb{T}^{*k} \right\rangle$ удовлетворяет условиям:

— *перестановочности* (*симметрии*), если вероятность любого цилиндрического множества $B^{(k)} \times \mathbb{R}^{\mathbb{T}\setminus\tau}$, $\tau = (t_1,\ldots,t_k)$, не зависит от порядка записи прямоугольника $B^{(k)} = B_1 \times \cdots \times B_k$, при условии сохранения информации о том, к какому параметру относится то или иное множество B_j:

$$\mathrm{P}\{B^{(k)};\tau\} \ = \ \mathrm{P}\{\pi B^{(k)};\pi\tau\} \tag{9}$$

при любых $\tau \in \bigcup_k \mathbb{T}^{*k}$ и любой перестановке π;

— *согласованности*, если для $\forall \tau \in \bigcup_k \mathbb{T}^{*k}$ и $\forall s \in \mathbb{T} \setminus \tau$ мера $\mathrm{P}(\cdot\,;\ \tau)$

есть частное распределение $\mathrm{P}(\cdot;\ (\tau, s))$:

$$\mathrm{P}\{B^{(k)} \times \mathbb{R}^1_s;\ (\tau, s)\} \ = \ \mathrm{P}\{B^{(k)};\ \tau\}, \quad \forall\, B^{(k)} \in \mathscr{B}^k, \tag{10}$$

где $(\tau, s) = (t_1, \ldots, t_k, s)$. Очевидно, условие согласованности эквивалентно условию, при котором вместо одного параметра s берётся любой набор параметров $(s_1, \ldots, s_m) \subset \mathbb{T} \setminus \tau$.

Если семейство $\mathcal{P}$ удовлетворяет условиям симметрии и согласованности, то соотношение (10) допускает усиление: при $\forall\, k \geqslant 1,\ \tau \in \mathbb{T}^{*k},\ 1 \leqslant n \leqslant k$

$$\mathrm{P}\Big\{ \prod_{j=1}^{k} B_j;\ \tau \Big\} \ = \ \mathrm{P}\Big\{ \prod_{j \neq n} B_j;\ \tau \setminus t_n \Big\}, \quad \forall\, B_1 \times \cdots \times B_k \in \mathscr{B}^k,\ B_n = \mathbb{R}^1_n. \tag{11}$$

503 **Теорема.** [*О продолжении согласованных мер.*] *Пусть семейство конечномерных вероятностных мер* $\big\langle\, \mathrm{P}\{B;\ \tau\},\ \tau \in \bigcup_n \mathbb{T}^{*k},\ B \in \mathscr{B}^\tau \,\big\rangle$ *на борелевских подмножествах конечномерных евклидовых пространств удовлетворяет условиям перестановочности* (9) *и согласованности* (10)*. Тогда на измеримом пространстве* $(\mathbb{R}^{\mathbb{T}}, \mathscr{B}^{\mathbb{T}})$ *с цилиндрической* σ*-алгеброй найдётся единственная вероятностная мера* $\mathbf{P}$*, для которой равенство*

$$\mathbf{P}\big\{ B \times \mathbb{R}^{\mathbb{T} \setminus \tau} \big\} \ = \ \mathrm{P}\{B;\ \tau\} \tag{12}$$

справедливо при $\forall\, B \in \mathscr{B}^k,\ \tau \in \mathbb{T}^{*k},\ k \geqslant 1$.

⇛ Определим функцию множеств $\mathbf{P}$ на цилиндрах по формуле (12). Как отмечено в 500, одинаковые цилиндры могут иметь различные описания, связанные либо с различными перестановками в записи сомножителей основания, либо с добавлением в эту запись сомножителей, совпадающих со всем пространством $\mathbb{R}^1$. Условия перестановочности и согласованности гарантируют, что мера $\mathbf{P}$ не зависит от способа записи, если произвести соответствующие изменения в наборе параметров τ.

Далее, поскольку борелевские прямоугольники образуют алгебру и порождают борелевскую σ-алгебру, то совпадение двух мер $\mathrm{P}\{B;\tau\}, \mathrm{P}\{B \times \mathbb{R}^1; (\tau, s)\}$ на борелевских прямоугольниках $B = B^{(k)}$ обеспечивает совпадение этих мер на любых борелевских подмножествах $B \subset \mathbb{R}^k$. Другими словами, условие согласованности выполняется для $\forall\, B \in \mathscr{B}^k,\ k \geqslant 1$.

Покажем, что функция множеств $\mathbf{P}$, определённая на цилиндрах, сигма-аддитивна. Выберем непересекающиеся цилиндрические множества

$B_j \times \mathbb{R}^{\mathbb{T}\setminus\tau_j}$, $B_j \in \mathbb{R}^{\tau_j}$, $j = 1, 2, \ldots$, такие, что их объединение тоже есть цилиндр:

$$B_0 \times \mathbb{R}^{\mathbb{T}\setminus\tau_0} = \biguplus_j B_j \times \mathbb{R}^{\mathbb{T}\setminus\tau_j}, \quad B_0 \in \mathbb{R}^{\tau_0}.$$

Рассмотрим счётный набор параметров $\mathbb{S} = \bigcup_{j=0}^{\infty} \tau_j$. Тогда множества $B_j \times \mathbb{R}^{\mathbb{S}\setminus\tau_j}$, $j = 0, 1, 2, \ldots$, представляют собой цилиндры в пространстве $\mathbb{R}^{\mathbb{S}}$. По теореме Колмогорова 496 семейство мер $\left\langle \mathrm{P}\{\cdot\,;\tau\},\ \tau \in \bigcup_k \mathbb{S}^{*k} \right\rangle$ допускает продолжение $\mathbf{P}_{\mathbb{S}}$ на измеримое пространство $(\mathbb{R}^{\mathbb{S}}, \mathscr{B}^{\mathbb{S}})$. Следовательно, в силу σ-аддитивности $\mathbf{P}_{\mathbb{S}}$

$$\mathbf{P}\left\{B_0 \times \mathbb{R}^{\mathbb{T}\setminus\tau_0}\right\} = \mathrm{P}\{B_0;\tau_0\} = \mathbf{P}_{\mathbb{S}}\left\{\biguplus_j B_j \times \mathbb{R}^{\mathbb{S}\setminus\tau_j}\right\} =$$
$$= \sum_j \mathbf{P}_{\mathbb{S}}\left\{B_j \times \mathbb{R}^{\mathbb{S}\setminus\tau_j}\right\} = \sum_j \mathrm{P}\{B_j;\tau_j\} = \sum_j \mathbf{P}\left\{B_j \times \mathbb{R}^{\mathbb{T}\setminus\tau_j}\right\}.$$

Таким образом, функция множеств $\mathbf{P}$, определённая на кольце цилиндрических множеств, порождающем цилиндрическую σ-алгебру $\mathscr{B}^{\mathbb{T}}$, обладает свойством σ-аддитивности. По теореме Каратеодори существует единственное продолжение этой меры на все подмножества $\mathscr{B}^{\mathbb{T}}$. ⋞

Если параметрическое множество $\mathbb{T}$ имеет порядок, то часто семейство конечномерных вероятностных мер бывает удобнее задавать только для упорядоченных наборов $\tau = \left\langle t_1 < \ldots < t_k \right\rangle$. В этом случае условие перестановочности не проверяется, а, наоборот, используется для определения меры для всех возможных перестановок. Это очевидное преимущество несколько портит то обстоятельство, что тогда условие согласованности должно быть проверено в усиленном варианте (11), когда множество $B = \mathbb{R}^1$ добавляется не на последнее место в записи борелевского прямоугольника, а в любое возможное место внутри этой записи. К примеру, для того чтобы по аналогии с теоремой 499 задать меру в $(\mathbb{R}^{\mathbb{T}}, \mathscr{B}^{\mathbb{T}})$ с помощью функций распределения, можно описать семейство

$$\left\langle F(\vec{x};\tau),\ \tau = (t_1, \ldots, t_k) \in \mathbb{T}, t_1 < \ldots < t_k,\ \vec{x} \in \mathbb{R}^k,\ k \geqslant 1 \right\rangle,$$

которое должно удовлетворять условию согласованности

$$F(\vec{x};\tau) = F(\vec{x}^{(j)};\tau \setminus t_j) \quad \text{при} \quad \begin{cases} \forall\, \vec{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k,\ x_j = +\infty, \\ \vec{x}^{(j)} = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k) \in \mathbb{R}^{k-1}, \\ \forall\, \tau = (t_1, \ldots, t_k) \subset \mathbb{T}, t_1 < \ldots < t_k. \end{cases}$$

⚠ Существуют варианты теоремы Колмогорова, в которых мера строится на цилиндрической σ-алгебре прямого произведения пространств $\prod_{t\in\mathbb{T}}\Omega_t$ с Ω_t, отличными от числовой прямой.

i) Для счётного $\mathbb{T}$ и особого способа задания семейства конечномерных вероятностей через так называемые переходные вероятности по теореме Ионеску Тулчи (стр. 508) структура Ω_t может быть полностью произвольна.

ii) Также ничем не ограничен выбор Ω_t и при несчётном $\mathbb{T}$, если конечномерные вероятности задаются через произведения соответствующих одномерных вероятностей (см. теорему Ломницкого–Улама на стр. 511).

iii) Для произвольного семейства согласованных конечномерных вероятностей теорема о продолжении будет справедлива для так называемых борелевских измеримых пространств $(\Omega_t, \mathscr{F}_t)$, которые алгебраически изоморфны некоторому борелевскому подпространству $(\mathbb{R}^1, \mathscr{B}^1)$, т.е. для них существует взаимно однозначное отображение $h:\Omega_t\mapsto X$, где $X\in\mathscr{B}[0;1]$, причём прообраз $h^{-1}(B)\in\mathscr{F}_t$ для любого борелевского подмножества $B\in X\bigcap\mathscr{B}^1$ (см., например, [6], [5]).

► **О пространствах $\mathcal{C}(\mathbb{T})$ и $\mathcal{D}(\mathbb{T})$.** Как отмечено выше, многие семейства функций, представляющие для нас интерес, будут неизмеримы относительно цилиндрической σ-алгебры. К счастью, чаще всего можно предположить, что наблюдаемые реализации функции X_t принадлежат более узкому пространству функций, например, что эти реализации непрерывны всюду или без разрывов второго рода. Рассмотрим совокупность $\mathcal{C}(\mathbb{T})$ всех непрерывных функций, заданных на компакте $\mathbb{T}$ и принимающих значения в $\mathbb{R}^1$. Если $\mathbb{T}\subset\mathbb{R}^1$, то, не слишком сильно ограничивая себя, можно считать $\mathbb{T}=[0;1]$. Снабжённое супремум-метрикой

$$\rho_U(X,Y) \;=\; \sup_{t\in\mathbb{T}}|X_t-Y_t|,$$

это пространство становится польским [13]. Можно показать, что цилиндрическая σ-алгебра на $\mathcal{C}[0;1]$ совпадает с σ-алгеброй, порождённой всеми открытыми шарами, и с σ-алгеброй, порождённой открытыми множествами.

Согласно построению цилиндрической σ-алгебры, при каждом фиксированном $t\in\mathbb{T}$ множества непрерывных функций вида $\{X\in\mathcal{C}(\mathbb{T}):\ X_t\leqslant \leqslant M\}$ измеримы. Поэтому для любого подмножества $T_0\subset\mathbb{T}$ пересечение $\cap_{t\in QT_0}\{X\in\mathcal{C}(\mathbb{T}):\ X_t\leqslant M\}\in\mathscr{B}(\mathcal{C}(\mathbb{T}))$, где Q — множество рациональных чисел. Однако ясно, что для непрерывной функции справедливость неравенства $X_t \leqslant M$ при всех рациональных t влечёт его справедливость при всех t. Таким образом, $\{X \in \mathcal{C}(\mathbb{T}) : \ \sup_{t\in T_0} X_t \leqslant M\} \in \mathscr{B}(\mathcal{C}(\mathbb{T}))$.

Аналогичные факты имеют место и для пространства $\mathcal{D}(\mathbb{T})$ всех функций на $\mathbb{T} = [0;1]$, непрерывных справа (при $t < 1$) и имеющих пределы слева (при $t > 0$). Снабжённое метрикой Скорохода, это пространство становится сепарабельным, что весьма важно для нас, поскольку в этом случае цилиндрическая σ-алгебра в $\mathcal{D}(\mathbb{T})$ совпадает с борелевской. Пусть Λ — множество всех строго возрастающих непрерывных функций $\lambda(t)$, $t \in [0;1]$, $\lambda(0) = 0, \lambda(1) = 1$, тогда метрика Скорохода определяется как

$$\rho_S(X,Y) \ = \ \inf\{\varepsilon > 0 : \ \exists\lambda \in \Lambda : \ \sup_t |X_t - Y_{\lambda(t)}| + \sup_t |t - \lambda(t)| < \varepsilon\}.$$

Отметим, что сходимость $X^{(n)} \to X$ в этой метрике влечёт сходимость $X_t^{(n)} \to X_t$ во всех точках t непрерывности функции X. В случае, когда X непрерывна всюду, сходимость в метрике Скорохода ρ_S влечёт сходимость в равномерной метрике ρ_U. Другими словами, в пространстве $\mathcal{C}([0;1])$ метрики ρ_U и ρ_S порождают одну топологию.

§3. Измеримые функции

Пусть задана некоторая функция $\xi : \Omega_1 \mapsto \Omega_2$. При изучении свойств этой функции нас всегда будет интересовать вопрос измеримости множеств вида $\{\omega \in \Omega_1 : \xi(\omega) \in B\}$ с некоторым $B \subset \Omega_2$.

Определения. *Прообразом множества* $B \subset \Omega_2$ относительно функции $\xi : \Omega_1 \mapsto \Omega_2$ (обозначается $\xi^{-1}(B)$ или $\{\xi \in B\}$) называется совокупность всех точек Ω_1, образы (значения) которых лежат в B:

$$\xi^{-1}(B) \ := \ \{\omega \in \Omega_1 : \ \ \xi(\omega) \in B\}.$$

Прообразом класса $\mathcal{H}$ подмножеств Ω_2 называется совокупность всех прообразов элементов $\mathcal{H}$: $\xi^{-1}(\mathcal{H}) = \left\langle \xi^{-1}(B) : B \in \mathcal{H} \right\rangle$.

504| **Лемма.** (?!) *Для любых наборов подмножеств* $B_\alpha \subset \Omega_2, \alpha \in \mathfrak{A}$:

а) $\xi^{-1}(\Omega_2) = \Omega_1, \qquad \xi^{-1}(\varnothing) = \varnothing;$

б) $\xi^{-1}\Big(\bigcap_{\alpha\in\mathfrak{A}} B_\alpha\Big) = \bigcap_{\alpha\in\mathfrak{A}} \xi^{-1}(B_\alpha), \qquad \xi^{-1}\Big(\bigcup_{\alpha\in\mathfrak{A}} B_\alpha\Big) = \bigcup_{\alpha\in\mathfrak{A}} \xi^{-1}(B_\alpha);$

в) $\xi^{-1}(B_1 \setminus B_2) = \xi^{-1}(B_1) \setminus \xi^{-1}(B_2), \qquad \xi^{-1}(B^c) = (\xi^{-1}(B))^c$.

⊫ Доказем только первое соотношение `в)`:

$$\omega \in \xi^{-1}(B_1 \setminus B_2) \;\Leftrightarrow\; \xi(\omega) \in B_1 \setminus B_2 \;\Leftrightarrow\; \xi(\omega) \in B_1,\; \xi(\omega) \notin B_2 \;\Leftrightarrow$$
$$\Leftrightarrow\; \omega \in \xi^{-1}(B_1),\; \omega \notin \xi^{-1}(B_2) \;\Leftrightarrow\; \omega \in \xi^{-1}(B_1) \setminus \xi^{-1}(B_2).$$

Остальные соотношения доказываются аналогично. ⫤

505| **Лемма.** (?!) *Пусть* $\mathscr{F}_k$ — σ*-алгебра подмножеств* Ω_k, $k = 1, 2$. *Тогда для любой функции* $\xi : \Omega_1 \mapsto \Omega_2$ *классы подмножеств*

$$\begin{aligned} \xi^{-1}(\mathscr{F}_2) &= \left\langle \xi^{-1}(B) : \; B \in \mathscr{F}_2 \right\rangle, \\ \tilde{\xi}(\mathscr{F}_1) &= \left\langle A \subset \Omega_2 : \; \xi^{-1}(A) \in \mathscr{F}_1 \right\rangle \end{aligned}$$

образуют σ*-алгебры на* Ω_1 *и* Ω_2 *соответственно.*

⊫ Покажем, к примеру, замкнутость $\tilde{\xi}(\mathscr{F}_1)$ относительно счётных объединений. Выберем произвольные подмножества $A_i \in \tilde{\xi}(\mathscr{F}_1), i = 1, 2, \ldots$ (т.е. для них $\xi^{-1}(A_i) \in \mathscr{F}_1$). В силу 504 $\xi^{-1}\Big(\bigcup_1^\infty A_i\Big) = \bigcup_1^\infty \xi^{-1}(A_i) \in \mathscr{F}_1$. ⫤

Другими словами, если на пространстве Ω_2 выделена σ-алгебра подмножеств $\mathscr{F}_2$, то прообраз этой σ-алгебры $\xi^{-1}(\mathscr{F}_2)$ будет определять σ-алгебру на Ω_1. Обратно, если на Ω_1 задана σ-алгебра подмножеств $\mathscr{F}_1$, то совокупность $\xi(\mathscr{F}_1)$ всех образов элементов $\mathscr{F}_1$ может и не быть σ-алгеброй. Однако совокупность всех тех подмножеств Ω_2, прообразы которых входят в σ-алгебру $\mathscr{F}_1$, будет уже σ-алгеброй подмножеств Ω_2.

Определение. Сигма-алгебру $\sigma\big(\xi\big) := \xi^{-1}(\mathscr{F}_2)$ называют *сигма-алгеброй, порождённой (индуцированной) функцией* ξ.

Определение. Функция $\xi(\omega_1) : (\Omega_1, \mathscr{F}_1) \mapsto (\Omega_2, \mathscr{F}_2)$, заданная на измеримом пространстве[‡] и принимающая значения в измеримом пространстве, называется *измеримой* (или $(\mathscr{F}_1|\mathscr{F}_2)$*-измеримой*), если прообраз любого $\mathscr{F}_2$-измеримого множества $\mathscr{F}_1$-измерим: $\xi^{-1}(B) \in \mathscr{F}_1$ для $\forall\, B \in \mathscr{F}_2$.

Иначе говоря, функция ξ измерима, если порождённая этой функцией σ-алгебра полностью входит в σ-алгебру $\mathscr{F}_1$: $\xi^{-1}(\mathscr{F}_2) \subset \mathscr{F}_1$. Очевидно, что $\mathscr{F} = \sigma\big(\xi\big)$ есть минимальная σ-алгебра в Ω_1 такая, что ξ $(\mathscr{F}|\mathscr{F}_2)$-измерима, и, аналогично, σ-алгебра $\mathscr{F} = \tilde{\xi}(\mathscr{F}_1)$ из леммы 505 есть максимальная σ-алгебра в Ω_2 такая, что функция ξ $(\mathscr{F}_1|\mathscr{F})$-измерима.

[‡] Запись $\xi : (\Omega_1, \mathscr{F}_1) \mapsto (\Omega_2, \mathscr{F}_2)$ применяется с единственной целью – подчеркнуть наличие в пространствах Ω_1, Ω_2 некоторых выделенных классов подмножеств; функция ξ определена на элементах $x \in \Omega_1$, а её значениями служат элементы $y \in \Omega_2$.

► Важно, что суперпозиция измеримых функций также измерима.

506| Лемма. *Если функция* $\xi(\omega_1) : (\Omega_1, \mathscr{F}_1) \mapsto (\Omega_2, \mathscr{F}_2)$ *и функция* $\eta(\omega_2) : (\Omega_2, \mathscr{F}_2) \mapsto (\Omega_3, \mathscr{F}_3)$ *измеримы, то измерима суперпозиция* $(\eta \circ \xi)(\omega_1) = \eta(\xi(\omega_1)) : (\Omega_1, \mathscr{F}_1) \mapsto (\Omega_3, \mathscr{F}_3)$.

$$\models\!\!\!\Rightarrow \quad (\eta \circ \xi)^{-1}(\mathscr{F}_3) \; = \; \xi^{-1}(\eta^{-1}(\mathscr{F}_3)) \; \subset \; \xi^{-1}(\mathscr{F}_2) \; \subset \; \mathscr{F}_1. \qquad \Leftarrow\!\!\!\dashv$$

Обратно, измеримость действительной функции η относительно σ-алгебры, порождённой действительной (векторной) функцией ξ, можно интерпретировать как зависимость η от ξ (см. теорема 516, стр. 469).

✧ Измеримость функции и порождающий класс. Чтобы установить измеримость какой-либо функции, не обязательно проверять измеримость прообразов всех $\mathscr{F}_2$-измеримых множеств — достаточно это проделать для всех подмножеств из семейства, порождающего $\mathscr{F}_2$.

507| Лемма. 1) *Прообраз* $\xi^{-1}(\sigma(\mathcal{X}))$ σ*-алгебры, порождённой семейством подмножеств* $\mathcal{X}$, *совпадает с* σ*-алгеброй* $\sigma(\xi^{-1}(\mathcal{X}))$, *порождённой прообразами* $\mathcal{X}$:

$$\xi^{-1}(\sigma(\mathcal{X})) \; = \; \sigma(\xi^{-1}(\mathcal{X})).$$

2) *Пусть функция* $\xi(\omega_1) : (\Omega_1, \mathscr{F}_1) \mapsto (\Omega_2, \mathscr{F}_2)$, где $\mathscr{F}_2 = \sigma(\mathcal{X})$. *Тогда*

$$\xi^{-1}(\mathscr{F}_2) \subset \mathscr{F}_1 \quad \Leftrightarrow \quad \xi^{-1}(\mathcal{X}) \subset \mathscr{F}_1.$$

$\models\!\!\!\Rightarrow$ 1) Так как σ-алгебра $\xi^{-1}(\sigma(\mathcal{X}))$ содержит $\xi^{-1}(\mathcal{X})$, то она должна содержать и порождаемую σ-алгебру: $\xi^{-1}(\sigma(\mathcal{X})) \supset \sigma(\xi^{-1}(\mathcal{X}))$.

Пусть $\mathscr{A}_1 = \sigma(\xi^{-1}(\mathcal{X}))$. Класс $\tilde{\xi}(\mathscr{A}_1)$ есть набор подмножеств $B \subset \Omega_2$, прообразы которых $\xi^{-1}(B) \in \mathscr{A}_1$ (см. 505). Поскольку для $\forall B \in \mathcal{X} \;\Rightarrow\; \xi^{-1}(B) \in \xi^{-1}(\mathcal{X}) \subset \mathscr{A}_1$, то имеет место включение $\mathcal{X} \subset \tilde{\xi}(\mathscr{A}_1)$. Так как класс $\tilde{\xi}(\mathscr{A}_1)$ есть σ-алгебра (см. 505), то $\sigma(\mathcal{X}) \subset \tilde{\xi}(\mathscr{A}_1)$. Другими словами, прообраз любого множества $\sigma(\mathcal{X})$ принадлежит $\sigma(\xi^{-1}(\mathcal{X}))$, т.е. $\xi^{-1}(\sigma(\mathcal{X})) \subset \sigma(\xi^{-1}(\mathcal{X}))$.

2) Необходимость ($\Rightarrow$) очевидна. Достаточность ($\Leftarrow$) вытекает из первого утверждения данной леммы и того факта, что минимальная σ-алгебра $\sigma(\xi^{-1}(\mathcal{X})) \subset \mathscr{F}_1$, раз $\xi^{-1}(\mathcal{X}) \subset \mathscr{F}_1$:

$$\xi^{-1}(\mathscr{F}_2) \; = \; \xi^{-1}(\sigma(\mathcal{X})) \; = \; \sigma(\xi^{-1}(\mathcal{X})) \; \subset \; \mathscr{F}_1. \qquad \Leftarrow\!\!\!\dashv$$

Из этой леммы и теоремы 488, стр. 445, следует корректность классического определения борелевской функции.

Определение. Функция $\vec{\xi}(\omega) : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^k, \mathscr{B}^k)$, принимающая значения в евклидовом пространстве $\mathbb{R}^k$, называется *измеримой по Борелю* (или просто *измеримой*), если для $\forall\, \vec{x} \in \mathbb{R}^k$

$$\{\vec{\xi} \leqslant \vec{x}\} \;:=\; \{\omega \in \Omega:\; \vec{\xi}(\omega) \leqslant \vec{x}\} \quad \in \mathscr{F}.$$

Если исходное пространство $(\Omega, \mathscr{F}) = (\mathbb{R}^m, \mathscr{B}^m)$ также борелевское, то измеримая функция называется *борелевской*.

508| Лемма. *Каждое из следующих условий необходимо и достаточно для измеримости векторной функции* $\vec{\xi} = (\xi_1, \ldots, \xi_k) : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^k, \mathscr{B}^k)$:

a) *измерима каждая компонента* $\xi_i, \quad i = \overline{1, k}$;

b) $\{\vec{\xi} \in (\vec{a}; \vec{b}]\} \in \mathscr{F}$ *для* $\forall\, \vec{a} < \vec{b}$;

c) $\{\vec{\xi} \leqslant \vec{a}\} \in \mathscr{F}$ *для* $\forall\, \vec{a} \in \mathbb{R}^k$;

d) $\{\|\vec{\xi} - \vec{x}\| < r\} \in \mathscr{F}$ *для* $\forall\, \vec{x} \in \mathbb{R}^k,\ r > 0$.

⇛ Все утверждения следуют из того, что соответствующие классы подмножеств порождают борелевскую σ-алгебру в $\mathbb{R}^k$ (теорема 488, стр. 445). Например, если все компоненты векторной функции измеримы, то для $\forall\, B_1, \ldots, B_k \in \mathscr{B}$ прообраз $\vec{\xi}^{-1}(B_1 \times \ldots \times B_k) = \bigcap_i \xi_i^{-1}(B_i) \in \mathscr{F}$. Другими словами, для класса $\mathcal{X} = \mathscr{B} \times \ldots \times \mathscr{B}$, порождающего борелевскую σ-алгебру в $\mathbb{R}^k$, имеет место включение $\vec{\xi}^{-1}(\mathcal{X}) \subset \mathscr{F}$, что по предыдущей лемме гарантирует измеримость $\vec{\xi}$. Обратное утверждение в a) следует из соотношений, аналогичных $\xi_1^{-1}(B) = \vec{\xi}^{-1}(B \times \mathbb{R}^1 \times \ldots \times \mathbb{R}^1) \in \mathscr{F}$. ⇚

509| Лемма. 1) *Если* $\vec{\xi} : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^k, \mathscr{B}^k)$ *— непрерывная функция и* $\mathscr{F}$ *содержит все открытые множества* Ω, *то* $\vec{\xi}$ *измерима по Борелю.*

2) *Монотонная функция* $\xi : (\Omega, \mathscr{B}(\Omega)) \mapsto (\mathbb{R}^1, \mathscr{B})$, *заданная на подмножестве* $\Omega \subset \mathbb{R}^1$, *измерима относительно индуцированной на* Ω *борелевской* σ*-алгебры* $\mathscr{B}(\Omega)$.

⇛ 1) Из курса анализа известно, что функция непрерывна т. т. т. когда прообраз любого открытого множества открыт. Таким образом, если σ-алгебра $\mathscr{F}$ в Ω содержит все открытые множества, то любая непрерывная функция $\vec{\xi} : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^k, \mathscr{B}^k)$ будет измеримой по Борелю, поскольку борелевская σ-алгебра порождается классом открытых множеств.

2) Если функция не убывает, то прообраз $\xi^{-1}(-\infty; x]$ можно представить

в виде $(-\infty; y] \cap \Omega$ или $(-\infty; y) \cap \Omega$ с некоторым $y \in \mathbb{R}^1$. По определению индуцированной σ-алгебры эти множества принадлежат $\mathscr{B}(\Omega)$. $\Lleftarrow$

510| Примеры. 1) Индикаторная функция $\dot{\mathrm{I}}(\omega; A)$ подмножества $A \subset \Omega$ измерима т. т. т. когда измеримо множество A.

2) Если пространство $\Omega = \biguplus_1^\infty A_j$ разбито на счётное число непересекающихся измеримых подмножеств, то измерима функция

$$\xi(\omega) \ = \ \sum_{j=1}^{\infty} c_j \, \dot{\mathrm{I}}(\omega; A_j), \qquad c_j \in \mathbb{R}^1, \ j = 1, 2, \ldots$$

3) Функции x^k $(k = 1, 2, \ldots)$, e^x, $|x|$, $\min(x, c)$, $\max(x, c), c \in \mathbb{R}^1$, определённые на всей борелевской прямой, непрерывны, а посему измеримы. Функция $\frac{1}{x} : (\Omega, \mathscr{B}(\Omega)) \mapsto (\mathbb{R}^1, \mathscr{B})$, определённая на множестве $\Omega = (-\infty; 0) \cup (0; \infty)$ с борелевской σ-алгеброй $\mathscr{B}(\Omega)$, порождённой всеми открытыми подмножествами Ω, непрерывна всюду на Ω. Аналогично, измерима функция $\sqrt[b]{x^a} : (\mathbb{R}^1, \mathscr{B}) \mapsto (\mathbb{R}^1, \mathscr{B})$ при $a, b \in \mathbb{N}$, b нечётно, а также функции x^c, $\ln x : ((0; \infty), \mathscr{B}(0; \infty)) \mapsto (\mathbb{R}^1, \mathscr{B})$ при произвольном $c \in \mathbb{R}^1$.

4) Функции $x + y$, $x - y$, xy, $\max\{x, y\}$, $\min\{x, y\}$ непрерывны на всем пространстве $\mathbb{R}^2$ и поэтому являются борелевскими. Функция $\frac{x}{y} : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B})$, определённая на $\Omega = \mathbb{R}^1 \times ((-\infty; 0) \cup (0; \infty))$ с индуцированной борелевской σ-алгеброй, также будет непрерывной и измеримой.

5) Измеримость рассмотренных здесь функций можно проверить прямыми методами. Например, если обозначить через $\mathbb{Q}$ множество всех рациональных чисел на прямой, то множество

$$\{(x, y) : x + y < a\} \ = \ \bigcup_{q \in \mathbb{Q}} \{(x, y) : x < q, y < a - q\} \in \mathscr{B}^2,$$

т.к. счётное объединение множеств вида $\{x < b, y < c\} \in \mathscr{B}^2$. Напомним, что класс интервалов вида $(-\infty; a)$ также порождает борелевскую σ-алгебру. ◎

При проверке измеримости одномерных (действительных) функций полезны утверждения следующей теоремы.

511| **Теорема.** *Если функции $\xi_n : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B}), n \geqslant 1$, измеримы и для $\omega \in \Omega$ существует предел $\xi(\omega) = \lim_n \xi_n(\omega)$, то ξ измерима. Кроме того,*

$$\varliminf_n \xi_n, \quad \varlimsup_n \xi_n, \quad \sup_n \xi_n, \quad \inf_n \xi_n$$

— также измеримые функции.

$\Rrightarrow$ Условие ξ $(= \lim_n \xi_n)$ $\leqslant x$ имеет место, только когда «для $\forall k > 0$ $\exists N \geqslant 1$, что для $\forall n \geqslant N$ $\xi_n \leqslant x + \dfrac{1}{k}$ ». Заменив здесь все логические операции на соответствующие теоретико-множественные операции, получаем, что множество $\{\xi \leqslant x\}$ измеримо, т.к. допускает представление посредством объединения и пересечения счётного числа измеримых множеств:

$$\{\omega : \xi(\omega) \leqslant x\} \;=\; \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \{\omega : \xi_n(\omega) \leqslant x + \tfrac{1}{k}\}.$$

Измеримость остальных функций следует из равенств:

$$\{\sup_n \xi_n \leqslant x\} \;=\; \bigcap_{n=1}^{\infty} \{\xi_n \leqslant x\}, \qquad \{\inf_n \xi_n < x\} \;=\; \bigcup_{n=1}^{\infty} \{\xi_n < x\},$$

$$\underline{\lim}_n\, \xi_n \;=\; \lim_n \Big(\inf_{k \geqslant n} \xi_k\Big), \qquad \overline{\lim}_n\, \xi_n \;=\; \lim_n \Big(\sup_{k \geqslant n} \xi_k\Big)$$

с учётом утверждений леммы 508. $\Lleftarrow$

✧ Простые функции.

Определение. Пусть $\Omega = \biguplus_1^K A_j$ — конечное $\mathscr{F}$-разбиение Ω. Действительная функция $h(\omega)$, $\omega \in \Omega$,

$$h(\omega) \;=\; \sum\nolimits_{j=1}^{K} c_j \, \dot{\mathrm{I}}\,(\omega; A_j), \qquad c_j \in \mathbb{R}^1, \quad A_j \in \mathscr{F}, \quad j = \overline{1, K},$$

неизменная на множествах из разбиения, называется *простой функцией*.

512| **Теорема.** *Неотрицательная функция* $\xi(\omega) : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1_+, \mathscr{B})$ *измерима т. т. т. когда существует неубывающая последовательность неотрицательных простых функций* $h_n(\omega)$, *сходящихся в каждой точке* $\omega \in \Omega$ *к* $\xi(\omega)$:

$$0 \;\leqslant\; h_{n-1}(\omega) \;\leqslant\; h_n(\omega) \;\underset{n \to \infty}{\longrightarrow}\; \xi(\omega), \qquad \omega \in \Omega.$$

$\Rrightarrow$ Если ξ измерима, то для любых целых $k, n \geqslant 1$ измеримы множества

$$A_{nk} \;=\; \Big\{\omega : \frac{(k-1)}{2^n} \leqslant \xi(\omega) < \frac{k}{2^n}\Big\}, \qquad B_n = \{\omega : \xi(\omega) \geqslant n\}.$$

Заметим, что при $\forall n > 1$ множества A_{nk}, $1 \leqslant k \leqslant n2^n$, не пересекаются и вместе с B_n образуют разбиение Ω. Легко проверить, что при увеличении n на единицу множества A_{nk} делятся пополам: $A_{(n+1)(2k-1)} + A_{(n+1)(2k)} = A_{nk}$.

Определим простую функцию

$$h_n(\omega) = \sum\nolimits_{k=1}^{n2^n} c_{nk}\,\dot{\mathrm{I}}\,(\omega; A_{nk}) + n\,\dot{\mathrm{I}}\,(\omega; B_n),$$

где $c_{nk} = (k-1)/2^n$ — «наименьшее» значение ξ на множестве A_{nk}. По построению $h_n(\omega) \leqslant h_{n+1}(\omega) \leqslant \xi(\omega)$ при $\forall\,\omega \in \Omega$, и, кроме того, любая точка ω, начиная с некоторого $n > 1$, попадёт в одно из множеств A_{nk}, поэтому $0 \leqslant \xi(\omega) - h_n(\omega) < 1/2^n \to 0$ при $n \to \infty$.

Достаточность следует из предыдущей теоремы. ⫷

Функцию, принимающую любые значения из $(-\infty; \infty)$, представим в виде разности двух положительных функций:

$$\xi(\omega) \;=\; \xi^+(\omega) - \xi^-(\omega), \qquad \xi^{\pm}(\omega) \geqslant 0,$$

где $\xi^+(\omega) = \xi(\omega)\,\dot{\mathrm{I}}\,(\omega; \xi(\omega) > 0)$ — «положительная часть» ξ, равная $\xi(\omega)$, если $\xi(\omega) > 0$, и равная 0, если $\xi(\omega) \leqslant 0$; функция $\xi^-(\omega) = (-\xi(\omega))^+ = -\xi(\omega)\,\dot{\mathrm{I}}\,(\omega; \xi(\omega) < 0)$ — «отрицательная часть» ξ.

513| Упр. а) Докажите измеримость положительной и отрицательной частей любой измеримой функции.

б) Проверьте, что $|\xi| = \xi^+ + \xi^-$, $\quad (-2\xi)^{\pm} = 2\xi^{\mp}$.

в) Докажите, что если $\xi \leqslant \eta$, то $\xi^+ \leqslant \eta^+$, $\xi^- \geqslant \eta^-$.

Из предыдущей теоремы непосредственно вытекает

514| **Следствие.** *Для любой измеримой по Борелю функции ξ найдётся такая последовательность простых функций $\langle\, h_n \,\rangle_1^{\infty}$, что $|h_n(\omega| \leqslant |\xi(\omega)|$ и $h_n(\omega) \underset{n}{\to} \xi(\omega)$ для $\forall\,\omega \in \Omega$; причём эта сходимость равномерная на области $\{\omega : -M < \xi(\omega) < M\}$ при $\forall\,M > 0$.*

Часто свойства, связанные с измеримыми функциями, могут быть легко установлены (или следуют из определения) для некоторого набора индикаторных функций. Перенос этого свойства на все измеримые функции может быть осуществлён с помощью следующей леммы. Семейство функций $\mathcal{H}$ будем называть *монотонным*, если оно замкнуто относительно монотонно возрастающих пределов: $\langle\, h_n \,\rangle_1^{\infty} \subset \mathcal{H}$, $h_n \nearrow h \;\Rightarrow\; h \in \mathcal{H}$.

515| **Лемма.** *Пусть $\mathcal{H}$ — монотонное семейство неотрицательных функций на Ω, содержащее индикаторные функции $\dot{\mathrm{I}}_B$ для всех B из некоторого класса $\mathcal{D}$ подмножеств Ω. Тогда семейство $\mathcal{H}$ содержит все неотрицательные $\sigma\big(\mathcal{D}\big)$-измеримые функции, если выполнено какое-либо из* *следующих условий:*

a) *класс $\mathcal{D}$ есть σ-алгебра, и семейство $\mathcal{H}$ замкнуто относительно линейных комбинаций с неотрицательными коэффициентами:*

$$h_1, h_2 \in \mathcal{H},\ c_1, c_2 \in \mathbb{R}^1_+ \ \Rightarrow\ c_1 h_1 + c_2 h_2 \in \mathcal{H};$$

b) *класс $\mathcal{D}$ замкнут относительно пересечений (образует π-систему) и $\mathcal{D} \ni \Omega$, семейство $\mathcal{H}$ замкнуто относительно неотрицательных линейных комбинаций:*

$$h_1, h_2 \in \mathcal{H},\ c_1, c_2 \in \mathbb{R}^1,\ c_1 h_1 + c_2 h_2 \geqslant 0 \ \Rightarrow\ c_1 h_1 + c_2 h_2 \in \mathcal{H}.$$

⊫ a) Любая неотрицательная $\sigma(\mathcal{D})$-измеримая функция принадлежит $\mathcal{H}$, т.к. она может быть приближена монотонной последовательностью простых $\mathcal{D}$-измеримых функций, каждая из которых по условию принадлежит $\mathcal{H}$.

b) Пусть $\mathcal{G} = \langle A \subset \Omega :\ \dot{\mathrm{I}}_A \in \mathcal{H} \rangle$. По условию к семейству $\mathcal{H}$ класс $\mathcal{G}$ есть монотонный класс. Кроме того, если $A, B \in \mathcal{G}$ и $A \subset B$, то $\dot{\mathrm{I}}_B - \dot{\mathrm{I}}_A \geqslant 0$, следовательно, $\dot{\mathrm{I}}_{B \setminus A} = \dot{\mathrm{I}}_B - \dot{\mathrm{I}}_A \in \mathcal{H}$, т.е. $B \setminus A \in \mathcal{G}$. Таким образом, класс $\mathcal{G}$ образует λ-систему, содержащую $\mathcal{D}$. По теореме 598, стр. 526, $\mathcal{G} \supset \sigma(\mathcal{D})$. Дальнейшие рассуждения идентичны пункту a). ⫤

В качестве примера докажем теорему о суперпозиции двух функций. Пусть $\xi : \Omega \mapsto \mathscr{X}$ — некоторая функция из Ω в измеримое пространство $(\mathscr{X}, \mathscr{F})$. Рассмотрим семейство $\mathcal{H}$ неотрицательных действительных функций $\eta : \Omega \mapsto \mathbb{R}^1$, для которых существует измеримая функция $h : (\mathscr{X}, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B}^1)$ такая, что $\eta(\omega) = h(\xi(\omega))$, т.е. η есть суперпозиция $h \circ \xi$. Очевидно, семейство $\mathcal{H}$ замкнуто относительно линейных комбинаций с неотрицательными коэффициентами. Для любого измеримого подмножества $B \in \mathscr{F}$ индикаторная функция $\dot{\mathrm{I}}_D(\omega) = \dot{\mathrm{I}}_B(\xi(\omega))$ относительно прообраза $D = \xi^{-1}(B)$ принадлежит $\mathcal{H}$. То есть $\mathcal{H}$ содержит все индикаторные функции подмножеств из класса $\xi^{-1}(\mathscr{F})$, образующего σ-алгебру $\sigma(\xi)$, порождённую отображением ξ. Пусть $\eta_n = h_n \circ \xi \nearrow \eta_0$. Определим последовательность борелевских функций $f_n = \max_{1 \leqslant k \leqslant n} h_j$. Очевидно, последовательность $f_n,\ n \geqslant 1$, не убывает; положим $f_0 = \widetilde{f}\ \dot{\mathrm{I}}(\widetilde{f} < \infty)$, где $\widetilde{f} = \lim_n f_n$. По теореме 511 функция f_0 борелевская. Далее, поскольку $f_n(\xi) = \eta_n$, то $\eta_0 = f_0(\xi) \in \mathcal{H}$. Следовательно, по лемме 515 $\mathcal{H}$ содержит все неотрицательные функции, измеримые относительно σ-алгебры, порождённой функцией ξ. Таким образом, справедлива

516| Теорема. *Пусть $\sigma(\xi) = \xi^{-1}(\mathscr{F})$ — σ-алгебра на Ω, порождённая отображением ξ из Ω в измеримое пространство $(\mathscr{X}, \mathscr{F})$. Функция*

$\eta : (\Omega, \sigma(\xi)) \mapsto (\mathbb{R}^1, \mathscr{B}^1)$ *измерима т. т. т. когда найдётся измеримая функция* $h : (\mathscr{X}, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B}^1)$ *такая, что* $\eta = h \circ \xi$, *т.е.* $\eta(\omega) = h(\xi(\omega))$.

▶ Изучим некоторые свойства измеримых функций. Утверждение следующей леммы основано на факте измеримости суперпозиции измеримых функций (лемма 506, стр. 464).

517| **Лемма.** *Если функция* $\xi : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B})$ *измерима, а функция* $\eta : Q \mapsto \mathbb{R}^1$, *заданная на борелевском подмножестве* $Q \subset \mathbb{R}^1$, *монотонна или непрерывна, причём* $\xi(\Omega) \subset Q$, *то суперпозиция* $h \circ \xi$ *измерима.*

⤇ Следует из лемм 509 и 506. ⤆

518| Примеры. Постоянно используются суперпозиции непрерывных функций $\sin(\xi)$, $\cos(\xi)$, e^{ξ}, $|\xi|$, ξ^k $(k = 1, 2, \ldots)$, $\xi^+ = \max(0, \xi)$, определённых на всей прямой $(Q = \mathbb{R}^1)$. Для функции $\sqrt{\xi}$ необходимо предполагать, что $\xi \geqslant 0$ $(Q = \mathbb{R}^1_+)$. Аналогично, функции $\ln(\xi)$ и ξ^c $(c > 0)$ измеримы, если $\xi > 0$ $(Q = (0; \infty))$, а функция $1/\xi$ — если $\xi \neq 0$ $(Q = (-\infty; 0) \cup (0; \infty))$. ⊚

519| Упр. Покажите измеримость функции $\mathrm{tg}(\xi)$, если измеримая функция $\xi \neq \frac{1}{2}\pi + \pi k$, $k = 0, \pm 1, \pm 2, \ldots$

520| **Теорема.** *Пусть функции* $\xi, \eta : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B})$ *измеримы. Тогда измеримы функции*

$$\xi + \eta, \quad \xi - \eta, \quad \xi \cdot \eta, \quad \frac{\xi}{\eta} \ (\eta \neq 0), \quad \max\{\xi, \eta\}, \quad \min\{\xi, \eta\}.$$

⤇ Функция $h(x, y) = x + y : (\mathbb{R}^2, \mathscr{B}^2) \mapsto (\mathbb{R}^1, \mathscr{B})$ непрерывна, а посему измерима. Двумерная функция $\zeta = (\xi, \eta) : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^2, \mathscr{B}^2)$ измерима, т.к. по условию компоненты этой функции измеримы (лемма 508). Следовательно, сумма $\xi + \eta : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B})$ измерима как суперпозиция измеримых функций. Аналогично для $\xi - \eta$, $\xi\eta$, $\max\{\xi, \eta\}$, $\min\{\xi, \eta\}$, ξ/η.

Ещё один способ доказательства использует теоремы 512 и 514 о представлении измеримых функций в виде предела простых. К примеру, пусть

$$h(\omega) = \sum_{i=1}^{K} h_i \, \dot{\mathrm{I}}(\omega; A_i), \qquad q(\omega) = \sum_{j=1}^{N} q_j \, \dot{\mathrm{I}}(\omega; B_j), \tag{13}$$
$$\Omega = \biguplus_{i=1}^{K} A_i = \biguplus_{j=1}^{N} B_j, \qquad h_i, q_j \in \mathbb{R}^1, \quad i = \overline{1, K}, \ j = \overline{1, N}.$$

Тогда $A_i = A_i \Omega = \biguplus_{j=1}^{N} A_i B_j$, $B_j = \biguplus_{i=1}^{K} B_j A_i$ и пространство $\Omega =$

$= \biguplus_{i=1}^{K} \biguplus_{j=1}^{N} A_i\, B_j$. Таким образом, функции h, q можно записать в виде

$$h(\omega) \;=\; \sum_{i=1}^{K}\sum_{j=1}^{N} h_i\, \dot{\mathrm{I}}\,(\omega; A_i\, B_j), \quad q(\omega) \;=\; \sum_{i=1}^{K}\sum_{j=1}^{N} q_j\, \dot{\mathrm{I}}\,(\omega; A_i\, B_j). \tag{14}$$

Следовательно, сумма функций $[h+q](\omega) \;=\; \sum_{i=1}^{K}\sum_{j=1}^{N}(h_i+q_j)\, \dot{\mathrm{I}}\,(\omega; A_i\, B_j)$ есть простая функция, т.к. пересечения измеримых множеств (и объединения этих пересечений, если они понадобятся) измеримы.

Для функций ξ и η по теореме 514 найдутся последовательности простых функций таких, что $\xi(\omega) = \lim_n h_n(\omega)$, $\eta(\omega) = \lim_n q_n(\omega)$. Стало быть,

$$\big(\xi + \eta\big) = \lim_n h_n + \lim_n q_n = \lim_n \big(h_n + q_n\big) \quad —$$

измеримая функция, как поточечный предел простых функций. ⫤

521| Упр. Обобщите на счётное число слагаемых теорему 520 : докажите, что если ряд $\eta(\omega) = \sum_i \xi_i(\omega)$ из измеримых функций при $\forall\omega \in \Omega$ сходится, то функция η измерима.

§4. Интеграл Лебега

Схема построения интеграла Лебега состоит из трёх блоков:

— определяется интеграл от неотрицательных простых функций;

— интеграл от неотрицательных измеримых функций определяется как монотонный предел интегралов простых функций;

— интеграл от измеримой функции задаётся как разность интегралов положительной и отрицательной частей функции, если такая разность арифметически допустима, т.е. нет неопределённости вида $+\infty - \infty$.

Доказательство свойств интеграла Лебега также может следовать этой схеме. Сначала свойство проверяется на простых неотрицательных функциях (индикаторах), затем устанавливается факт сохранения свойства при предельном переходе, тем самым свойство доказывается для неотрицательных функций (в связи с этим см. также лемму 515). После чего делается заключение о возможности переноса этого свойства на произвольные измеримые функции.

Рассмотрим абстрактное пространство с σ-конечной мерой $(\Omega, \mathscr{F}, \mu)$. В дальнейшем будем всегда полагать $0 \cdot \infty = 0$.

Пусть $\mathfrak{P}_+(\Omega, \mathscr{F})$ — класс всех простых неотрицательных функций на измеримом пространстве $(\Omega, \mathscr{F})$.

► *Интегралом Лебега* от простой неотрицательной функции

$$h(\omega) \;=\; \sum\nolimits_{j=1}^{K} c_j \,\dot{\mathrm{I}}\,(\omega; A_j) \in \mathfrak{P}_+(\Omega, \mathscr{F}), \qquad c_j \geqslant 0,\; j = \overline{1, K},\; K < \infty,$$

по Ω относительно меры μ называется величина

$$\int\limits_{\Omega} h\, d\mu \;:=\; \sum_{j=1}^{K} c_j \mu(A_j),$$

где $c_j\mu(A_j) = 0$, если $c_j = 0$, $\mu(A_j) = \infty$. Иногда мы будем использовать расширенные варианты обозначений $\int_\Omega h(\omega)\, d\mu$ или $\int_\Omega h(\omega)\, \mu(d\omega)$. Или, наоборот, сокращённые варианты $\int h\, d\mu$, $\int h$, когда из контекста понятны мера и пространство интегрирования.

Приведём некоторые полезные свойства интеграла Лебега.

522| **Лемма.** *Пусть* $g, h \in \mathfrak{P}_+(\Omega, \mathscr{F})$, *тогда:*

($\mathrm{L_{S1}}$) $\int_\Omega h\, d\mu$ *не зависит от способа представления функции* h;

($\mathrm{L_{S2}}$) $\int_\Omega c\,\dot{\mathrm{I}}_A\, d\mu = c\,\mu(A), \quad c \geqslant 0$;

($\mathrm{L_{S3}}$) *если* $\mu(A) = 0$, *то* $\int_\Omega h\,\dot{\mathrm{I}}_A\, d\mu = 0$;

($\mathrm{L_{S4}}$) *(линейность) если константы* $a, b > 0$, *то*

$$\int_\Omega (a\,h + b\,g)\, d\mu \;=\; a\int_\Omega h\, d\mu + b\int_\Omega g\, d\mu;$$

($\mathrm{L_{S5}}$) *(монотонность) если* $h \leqslant g$, *то* $\int_\Omega h\, d\mu \leqslant \int_\Omega g\, d\mu.$

⊨ ($\mathrm{L_{S1}}$) Пусть $h(\omega) = \sum_{i=1}^{K} c_i \,\dot{\mathrm{I}}\,(\omega; A_i) = \sum_{j=1}^{N} q_j \,\dot{\mathrm{I}}\,(\omega; B_j)$ с $\Omega = \biguplus_{i=1}^{K} A_i = \biguplus_{j=1}^{N} B_j$. Произведение $c_i\mu(A_iB_j) = q_j\mu(A_iB_j)$ при всех i, j. Действительно, если найдётся $\omega \in A_iB_j$, то $c_i = h(\omega) = q_j$. В противном случае $\mu(A_iB_j) = 0$. Следовательно, в силу аддитивности меры

$$\sum_{i=1}^{K} c_i\,\mu(A_i) \;=\; \sum_{i=1}^{K}\sum_{j=1}^{N} c_i\,\mu(A_iB_j) \;=\; \sum_{i=1}^{K}\sum_{j=1}^{N} q_j\,\mu(A_iB_j) \;=\; \sum_{j=1}^{N} q_j\,\mu(B_j)\,.$$

($\mathrm{L_{S4}}$) Воспользуемся соотношениями (13, 14). По определению

$$\textstyle\int_\Omega (ah + bg)\, d\mu \;=\; \displaystyle\sum_{i=1}^{K}\sum_{j=1}^{N} (ac_i + bz_j)\mu(A_iB_j)\,,$$

что после очевидных преобразований доказывает утверждение ($\mathrm{L_{S4}}$).

(L_{S5}) Снова используем представления (13, 14). Как и при доказательстве (L_{S1}), легко устанавливается неравенство $c_i\,\mu(A_iB_j) \leqslant z_j\,\mu(A_iB_j)$ при всех i, j, что доказывает теорему. ⩤

► Пусть теперь $\xi : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1_+, \mathscr{B})$ — произвольная неотрицательная измеримая по Борелю функция на Ω. Интегралом от функции ξ по Ω называется величина

$$\int_\Omega \xi\, d\mu \;:=\; \lim_n \int_\Omega h_n\, d\mu,$$

где $h_n \in \mathfrak{P}_+(\Omega, \mathscr{F}),\ \ n \geqslant 1$, — возрастающая последовательность простых неотрицательных функций такая, что $\xi(\omega) = \lim_n\!\uparrow h_n(\omega)$ для $\forall\,\omega \in \Omega$.

Покажем корректность такого определения. По свойству монотонности интеграла (L_{S5}) последовательность интегралов, рассмотренных в определении, возрастает и, следовательно, предел всегда существует или равен $+\infty$.

523| **Лемма.** *Интеграл* $\int_\Omega \xi\, d\mu$ *от неотрицательной измеримой функции не зависит от способа её приближения монотонной последовательностью простых функций:*

$$\xi(\omega) \;=\; \lim_n\!\uparrow x_n(\omega) \;=\; \lim_n\!\uparrow y_n(\omega), \quad \forall\,\omega \in \Omega; \quad \big\langle\, x_n,\ y_n \big\rangle_1^\infty \subset \mathfrak{P}_+(\Omega, \mathscr{F}),$$
$$\Downarrow$$
$$\lim_n \int\limits_\Omega x_n(\omega)\, d\mu \;=\; \lim_n \int\limits_\Omega y_n(\omega)\, d\mu.$$

⊫ Здесь нам понадобится неравенство Маркова — лемма 524. Прежде всего отметим, что по этому неравенству мера $\mu(\xi > \varepsilon) < \infty$ при $\forall\,\varepsilon > 0$, если $\lim_n \int_\Omega x_n\, d\mu < \infty$.

Зафиксируем $m \geqslant 1\ \ (y_m \not\equiv 0)$, и пусть $\lim_n \int_\Omega x_n\, d\mu < \infty$. Так как y_m принимает конечное число ненулевых значений, то существуют числа

$$y^* = \max_\omega y_m(\omega) < \infty \quad \text{и} \quad y_* = \min_{\omega : y_m(\omega) > 0} y_m(\omega) > 0.$$

Выберем $0 < \varepsilon < y_*$. Тогда для $Q_n = \{\omega \in \Omega :\ x_n(\omega) + \varepsilon > y_m(\omega)\},\ n \geqslant 1$,

$$\begin{aligned} y_m \;&=\; y_m\, \dot{\mathrm{I}}_{Q_n\{y_m>0\}} + y_m\, \dot{\mathrm{I}}_{Q_n^c} \;\leqslant\; (x_n + \varepsilon)\, \dot{\mathrm{I}}_{Q_n\{y_m>0\}} + y^*\ \dot{\mathrm{I}}_{Q_n^c} \;\leqslant \\ &\leqslant\; x_n + \varepsilon\, \dot{\mathrm{I}}_{Q_n\{y_m>0\}} + y^*\ \dot{\mathrm{I}}_{Q_n^c}\,. \end{aligned}$$

Так как $x_n \nearrow \xi \geqslant y_m$, то $\lim_n\!\uparrow Q_n = \Omega$. Если $\omega \in Q_n^c$, тогда $\xi \geqslant y_m \geqslant \geqslant x_n + \varepsilon \geqslant \varepsilon$, поэтому $\mu(Q_n^c) \leqslant \mu(\xi \geqslant \varepsilon) < \infty$, следовательно, по свойству непрерывности меры $\lim\limits_{n}{}_{\downarrow}\,\mu(Q_n^c) = 0$.

Если $\omega \in Q_n \cap \{y_m > 0\} = Q_n \cap \{y_m \geqslant y_*\}$, то $x_n > y_* - \varepsilon > 0$. Следовательно, при каждом $n \geqslant 1$

$$\mu(Q_n\{y_m > 0\}) \;\leqslant\; \lim\limits_{n}{}_{\uparrow}\,\mu(Q_n\{y_m > 0\}) \;\leqslant$$
$$\leqslant\; \lim\limits_{n}{}_{\uparrow}\,\mu(x_n > y_* - \varepsilon) \;=\; \mu(\xi > y_* - \varepsilon) < \infty.$$

Поэтому для $\forall\, \varepsilon < y_*$

$$\int_\Omega y_m\,d\mu \;\leqslant\; \int_\Omega x_n\,d\mu + \varepsilon\,\mu(\xi > y_* - \varepsilon) + y^*\mu(Q_n^c)$$
$$\underset{n\to\infty}{\longrightarrow}\; \lim_n \int_\Omega x_n\,d\mu + \varepsilon\mu(\xi > y_* - \varepsilon) \;\underset{\varepsilon\to 0}{\longrightarrow}\; \lim_n \int_\Omega x_n\,d\mu.$$

Таким образом, если $\lim_n \int x_n\,d\mu < \infty$, то $\lim_m \int y_m\,d\mu \leqslant \lim_n \int x_n\,d\mu < \infty$. Произведя перестановку $x \leftrightarrow y$, получаем требуемое. ⇚

524 **Лемма.** [*Неравенство Маркова.*] *Пусть* $\xi \geqslant 0$ *и последовательность* $\langle\, x_n \,\rangle_1^\infty \subset \mathfrak{P}_+(\Omega, \mathscr{F})$ *такова, что* $\lim\limits_{n}{}_{\uparrow}\, x_n = \xi$. *Обозначим предел* $\lim_n \int_\Omega x_n\,d\mu = \int_\Omega \xi\,d\mu$[‡]. *Тогда для* $\forall\, A > 0$

$$\mu(\xi > A) \;\leqslant\; \frac{1}{A}\int\limits_\Omega \xi\,d\mu\,. \tag{15}$$

⇛ По свойствам линейности и монотонности интеграла от неотрицательных простых функций и по свойству непрерывности меры

$$\int_\Omega \xi\,d\mu \;=\; \lim_n \Big(\int_\Omega x_n\,\dot{\mathrm{I}}_{\{x_n > A\}}\,d\mu + \int_\Omega x_n\,\dot{\mathrm{I}}_{\{x_n \leqslant A\}}\,d\mu \Big) \;\geqslant$$
$$\geqslant\; A \lim\limits_{n}{}_{\uparrow}\,\mu(x_n > A) \;=\; A\,\mu(\xi > A).$$
⇚

⚠ Однозначность определения интеграла Лебега посредством произвольно взятой возрастающей последовательности простых функций выгодно отличает его от интеграла Римана. Интеграл Лебега можно определить, воспользовавшись конкретным приближением из теоремы 512, стр. 467:

$$\int\limits_\Omega \xi\,d\mu \;=\; \lim_n \left[\sum_{i=1}^{n2^n} \frac{i-1}{2^n}\mu\left(\frac{i-1}{2^n} \leqslant \xi < \frac{i}{2^n}\right) + n\mu\,(\xi \geqslant n) \right],$$

где предел всегда существует или равен $+\infty$.

[‡] Здесь $\int_\Omega \xi\,d\mu$, понимаемое пока как обозначение для $\lim_n \int_\Omega x_n\,d\mu$, после леммы 523 становится полновесным интегралом (см. также стр. 486).

Другое определение интеграла от положительной функции, не требующее проверки корректности, содержит следующая теорема. Обозначим через $L(\xi) = \langle\, h \in \mathfrak{P}_+(\Omega, \mathscr{F}) : \ 0 \leqslant h \leqslant \xi \,\rangle$ класс всех неотрицательных простых функций, «подпирающих» неотрицательную функцию ξ снизу.

525 **Теорема.** *Пусть $\xi : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1_+, \mathscr{B})$ — измеримая функция. Тогда*

$$\int_\Omega \xi \, d\mu \ = \ \sup_{h \in L(\xi)} \int_\Omega h \, d\mu.$$

⤃ Обозначим правую часть доказываемого равенства через J. По свойству супремума всегда найдётся такая последовательность $\langle\, h_n \,\rangle_1^\infty \subset L(\xi)$, что $\int h_n \, d\mu \to J$. Пусть $\langle\, q_n \,\rangle_1^\infty \subset L(\xi)$ — произвольная монотонная последовательность, для которой $q_n \nearrow \xi$. Образуем новую последовательность простых функций $z_n = \max\{h_1, \ldots, h_n, q_n\}$, $n \geqslant 1$, которая возрастает и принадлежит $L(\xi)$. Так как $h_n \leqslant z_n$, то $\int z_n \, d\mu \to J$. С другой стороны, $q_n \leqslant z_n \leqslant \xi$, следовательно, $z_n \nearrow \xi$ и по лемме 523 $\int \xi \, d\mu = \lim_n \int z_n \, d\mu = J$. ⤂

Приведём свойства интеграла от неотрицательных функций.

526 **Теорема.** *Пусть $\xi, \eta \geqslant 0$ — измеримые функции. Тогда:*

($\mathrm{L_{p1}}$) *если $\mu(\xi \neq \eta) = 0$, то $\int_\Omega \xi \, d\mu \ = \ \int_\Omega \eta \, d\mu$;*

($\mathrm{L_{p2}}$) $\int_\Omega \xi \, d\mu \geqslant 0$, *с равенством т. т. т. когда $\mu(\xi > 0) = 0$;*

($\mathrm{L_{p3}}$) *(монотонность) если $\mu(\xi > \eta) = 0$, то*

$$\int_\Omega \xi \, d\mu \ \leqslant \ \int_\Omega \eta \, d\mu$$

со знаком равенства, если (и только если) $\mu(\xi \neq \eta) = 0$ (при $\int_\Omega \eta \, d\mu < \infty$);

($\mathrm{L_{p4}}$) *(линейность) для любых констант $a, b > 0$*

$$\int_\Omega (a\xi + b\eta) \, d\mu \ = \ a \int_\Omega \xi \, d\mu + b \int_\Omega \eta \, d\mu;$$

если $a\xi(\omega) - b\eta(\omega) \geqslant 0$ для $\forall\, \omega \in \Omega$, то при условии $\int_\Omega \eta \, d\mu < \infty$

$$\int_\Omega (a\xi - b\eta) \, d\mu \ = \ a \int_\Omega \xi \, d\mu - b \int_\Omega \eta \, d\mu.$$

⤃ ($\mathrm{L_{p1}}$) Пусть $W = \{\omega : \xi(\omega) = \eta(\omega)\}$, тогда для любой функции $h \in L(\xi)$ простая функция $q = h\, \dot{\mathrm{I}}_W \leqslant \eta$, т.е. $q \in L(\eta)$, причём по свойству

(L_{S3}) для простых функций интегралы от h и q равны:

$$\int_\Omega h\,d\mu \;=\; \int_\Omega h\,\dot{\mathrm{I}}_W\;d\mu + \int_\Omega h\,\dot{\mathrm{I}}_{W^c}\;d\mu \;=\; \int_\Omega q\,d\mu.$$

Другими словами, если $\mu(\xi \neq \eta) = 0$, то для любой простой функции из $L(\xi)$ найдётся простая функция из $L(\eta)$, интегралы от которых равны. Очевидно, что справедливо и обратное утверждение. Следовательно, супремумы интегралов по этим двум классам простых функций совпадают.

(L_{p2}) По неравенству Маркова (лемма 524) для $\forall\, k > 0$

$$\mu\Bigl(\xi > \frac{1}{k}\Bigr) \;\leqslant\; k\int_\Omega \xi\,d\mu.$$

Поэтому если интеграл равен нулю, то $\mu\bigl(\xi > \frac{1}{k}\bigr) = 0$ и $\mu(\xi > 0) = \lim\limits_{k}{}_{\uparrow}\,\mu\bigl(\xi > \frac{1}{k}\bigr) = 0$. Обратно, если $\mu(\xi > 0) = 0$, то по свойству (L_{p1}) $\int \xi\,d\mu = 0$.

(L_{p3}) В силу свойства (L_{p1}) достаточно рассмотреть случай $\xi(\omega) \leqslant \eta(\omega)$ при $\forall\,\omega \in \Omega$. В соответствии с ($L_{p4}$)

$$\int_\Omega \eta\,d\mu \;=\; \int_\Omega \xi\,d\mu + \int_\Omega (\eta - \xi)\,d\mu \;\geqslant\; \int_\Omega \xi\,d\mu;$$

причём ввиду (L_{p2}) равенство здесь будет достигаться, лишь когда $\mu(\xi < \eta) = 0$, если, конечно, $\int \eta\,d\mu \neq \infty$.

(L_{p4}) Если $h_n, q_n \in \mathfrak{P}_+(\Omega, \mathscr{F})$ и $\lim\limits_{n}{}_{\uparrow}\, h_n = \xi$, $\lim\limits_{n}{}_{\uparrow}\, q_n = \eta$, то $\lim\limits_{n}{}_{\uparrow}\,(ah_n + bq_n) = (a\xi + b\eta)$ и по определению интеграла

$$\int_\Omega (a\xi + b\eta)\,d\mu \;=\; \lim_{n}{}_{\uparrow} \int_\Omega (ah_n + bz_n)\,d\mu \;=\; a\int_\Omega \xi\,d\mu + b\int_\Omega \eta\,d\mu,$$

в силу линейности интеграла от простых функций.

Второе утверждение этого пункта следует из первого:

$$a\int_\Omega \xi\,d\mu \;=\; \int_\Omega \bigl((a\xi - b\eta) + b\eta\bigr)\,d\mu \;=\; \int_\Omega (a\xi - b\eta)\,d\mu + b\int_\Omega \eta\,d\mu.$$

Заметим, что справедливость заявленного соотношения следует из приведённого равенства только в случае $\int_\Omega \eta\,d\mu < \infty$. ⟽

► Пусть $\xi : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B})$ — измеримая по Борелю функция, ξ^+, ξ^- — её положительная и отрицательная части. Интегралом от ξ на Ω называется величина

$$\int\limits_\Omega \xi\,d\mu \;:=\; \int\limits_\Omega \xi^+\,d\mu - \int\limits_\Omega \xi^-\,d\mu,$$

если хотя бы один из интегралов правой части конечен (говорят, что интеграл на Ω *существует*). Если оба интеграла правой части конечны, то $\int_\Omega \xi\,d\mu$ конечен и функция ξ называется *интегрируемой на* Ω по Лебегу.

Интегрируемость по Лебегу носит абсолютный характер — функция интегрируема только вместе со своим модулем.

527| **Теорема.** [*Неравенство «треугольника».*] *Измеримая функция ξ интегрируема т. т. т. когда интегрируема функция $|\xi|$, при этом*

$$\left|\int_\Omega \xi\,d\mu\right| \leqslant \int_\Omega |\xi|\,d\mu.$$

⊫ В силу линейности интеграла от положительных функций

$$\int_\Omega |\xi|\,d\mu = \int_\Omega (\xi^+ + \xi^-)\,d\mu = \int_\Omega \xi^+\,d\mu + \int_\Omega \xi^-\,d\mu,$$

что доказывает первую часть теоремы. Вторая часть следует из неравенства треугольника для действительных чисел и предыдущего равенства:

$$\left|\int_\Omega \xi^+\,d\mu - \int_\Omega \xi^-\,d\mu\right| \leqslant \int_\Omega \xi^+\,d\mu + \int_\Omega \xi^-\,d\mu\,. \qquad ⫤$$

⚠ Если мера μ — считающая мера какого-либо не более чем счётного подмножества Ω, то приведённое в теореме неравенство действительно представляет собой неравенство треугольника для сумм действительных чисел.

528| Упр. Докажите, что если $\int \eta\,d\mu < \infty$ и $|\xi| \leqslant \eta$, то ξ интегрируема и $\left|\int \xi\,d\mu\right| \leqslant \int \eta\,d\mu$.

✧ Свойства интеграла Лебега

529| **Теорема.** *Пусть ξ, η — измеримые функции, тогда:*

(L_1) (?!) *если* $\mu(\xi \neq \eta) = 0$, *то* $\displaystyle\int_\Omega \xi\,d\mu = \int_\Omega \eta\,d\mu$,

когда существует хотя бы один из интегралов;

(L_2) *(однородность)* $\displaystyle\int_\Omega c\xi\,d\mu = c\int_\Omega \xi\,d\mu$ *для любых* $c \in \mathbb{R}^1$;

(L_3) *(монотонность) если* $\xi \leqslant \eta$, *то* $\displaystyle\int_\Omega \xi\,d\mu \leqslant \int_\Omega \eta\,d\mu$,

когда существуют оба интеграла;

(L_4) *(аддитивность)* $\displaystyle\int_\Omega (\xi+\eta)\,d\mu = \int_\Omega \xi\,d\mu + \int_\Omega \eta\,d\mu$,

когда выражение в правой части корректно определено.

$\Rrightarrow$ Так как мера μ и пространство Ω здесь неизменны, будем использовать сокращённый вариант записи $\int_\Omega \zeta\, d\mu = \int \zeta$.

(L_2) При $c<0$ имеем $(c\xi)^{\pm} = -c\xi^{\mp}$. Поэтому

$$\int (c\xi)^+ - \int (c\xi)^- \;=\; c\,(-\int \xi^- + \int \xi^+) = c\int \xi.$$

(L_3) Легко понять, что из условия теоремы следуют неравенства $\xi^+ \leqslant \eta^+$, $\xi^- \geqslant \eta^-$. Поэтому если $\int \xi = +\infty$, т.е. $\int \xi^+ = \infty$, $\int \xi^- < \infty$, то по свойству монотонности интеграла от положительных функций $\int \eta^+ = \infty$, $\int \eta^- < \infty$, т.е. $\int \eta = +\infty$. Аналогично для случая $\int \eta = -\infty$.

Пусть теперь функции ξ, η интегрируемы: $\int \xi^{\pm} < +\infty$, $\int \eta^{\pm} < +\infty$. Условие теоремы приводит к неравенству для неотрицательных функций:

$$\xi^+ + \eta^- \;\leqslant\; \eta^+ + \xi^-.$$

Воспользовавшись последовательно свойствами монотонности и аддитивности для интеграла от неотрицательных функций, а также интегрируемостью всех слагаемых, получаем утверждение теоремы.

(L_4) Нам понадобятся следующие соотношения:

$$(*) \qquad (\xi+\eta)^+ \;\leqslant\; \xi^+ + \eta^+, \qquad (\xi+\eta)^- \;\leqslant\; \xi^- + \eta^-,$$

$$(^*_*) \qquad (\xi+\eta)^- \;=\; \left[\xi^- + \eta^-\right] - \left[(\xi^+ + \eta^+) - (\xi+\eta)^+\right] \quad (\geqslant 0).$$

Неравенства $(*)$ проверяются простым перебором, а равенство $(^*_*)$ следует из равенств $\xi^+ - \xi^- + \eta^+ - \eta^- \;=\; \xi + \eta \;=\; (\xi+\eta)^+ - (\xi+\eta)^-$. Заметим, что в $(^*_*)$ оба выражения, стоящие в квадратных скобках, неотрицательны.

Из всех ситуаций, когда выражение $\int \xi + \int \eta$ имеет смысл, достаточно рассмотреть три.

($\checkmark$) В случае $\int \xi^+ < \infty$, $\int \eta^+ < \infty$, $\int \xi^- = \int \eta^- = \infty$ из $(*)$ следует, что $\int (\xi+\eta)^+ < \infty$, а из $(^*_*)$, ввиду линейности интеграла от положительных функций (свойство (L_{p4})), что $\int (\xi+\eta)^- = \infty$. Поэтому

$$\int (\xi+\eta)^+ - \int (\xi+\eta)^- \;=\; -\infty \;=\; \int \xi + \int \eta.$$

($\checkmark$) Ситуация $\int \xi^{\pm} < \infty$, $\int \eta^+ < \infty$, $\int \eta^- = \infty$ разбирается аналогично.

($\checkmark$) Если $\int \xi^{\pm} < \infty$, $\int \eta^{\pm} < \infty$, то, как и раньше, $\int (\xi+\eta)^{\pm} < \infty$ и из $(^*_*)$ имеем

$$\int (\xi+\eta)^+ \;=\; \int \xi^+ + \int \eta^+ - \int \xi^- - \int \eta^- + \int (\xi+\eta)^-.$$

Откуда (L_4) следует простыми рассуждениями. $\Lleftarrow$

530| Упр. Покажите, что для любой конечной функции h интеграл относительно меры Дира́ка δ_{ω_0}, сосредоточенной в точке ω_0, равен

$$\int_\Omega h\,d\delta_{\omega_0} \;=\; h(\omega_0).$$

► Интеграл Лебега может быть распространён на функции, принимающие значения $\pm\infty$. Необходимость в этом возникает, например, при рассмотрении ряда из положительных функций, если этот ряд при некоторых $\omega \in \Omega$ не сходится $(= +\infty)$. Обозначим через $\phi_\xi = \{\xi(\omega) < \infty\}$ множество точек Ω, в которых неотрицательная функция ξ принимает конечные значения. По неравенству Маркова (15), если функция ξ интегрируема, то

$$\mu(\phi_\xi^c) \;=\; \lim_{N}{}_{\downarrow}\,\mu(\xi > N) \;\leqslant\; \lim_{N} \frac{1}{N}\int_\Omega \xi\,d\mu \;=\; 0.$$

531| **Лемма.** (?!) *Пусть для расширенной неотрицательной функции* $\xi \geqslant 0$ *мера* $\mu(\xi = +\infty) > 0$. *Если монотонная последовательность простых функций* $\langle h_n \rangle \subset \mathfrak{P}_+(\Omega, \mathscr{F})$ *такова, что* $h_n \nearrow \xi$, *то* $\lim_{n}{}_{\uparrow} \int_\Omega h_n\,d\mu = +\infty$.

Итак, можно положить по определению $\int_\Omega \xi\,d\mu = +\infty$, если $\mu(\xi = +\infty) > 0$, и $\int_\Omega \xi\,d\mu = \int_\Omega \xi\,\dot{\mathrm{I}}_{\phi_\xi}\,d\mu$ в противном случае.

► **Интеграл по сумме мер.** Иногда мера, по которой производится интегрирование, может быть представлена в виде конечной суммы мер.

532| **Лемма.** (?!) *Пусть* μ_1, μ_2 *— сигма-конечные меры на измеримом пространстве* $(\Omega, \mathscr{F})$. *Тогда для любых* $a_1, a_2 \geqslant 0$:

а) *функция множеств* $a_1\mu_1 + a_2\mu_2$ *также задаёт меру;*

б) *для любой измеримой функции* $\xi : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1_+, \mathscr{B})$

$$\int_\Omega \xi\,d(a_1\mu_1 + a_2\mu_2) \;=\; a_1\int_\Omega \xi\,d\mu_1 + a_2\int_\Omega \xi\,d\mu_2.$$

533| Упр. Используя 530, покажите, что для меры μ, считающей точки множества $\mathcal{S} = \{\omega_1, \ldots, \omega_n\}$, $n < \infty$, интеграл $\int \xi\,d\mu = \sum_1^n \xi(\omega_i)$.

✧ **Теоремы Леви, Фату и Лебега.** В редких случаях, как при доказательстве линейности, можно обойтись чисто алгебраическими методами — чаще приходится прибегать к методам, использующим асимптотические свойства интеграла.

534| **Теорема.** [*О монотонной сходимости. Беппо Леви.*] *Если монотонно возрастающая последовательность функций* $\xi_{n-1} \leqslant \xi_n \nearrow \eta,\ n \to \infty$, *причём* $\int \xi_m^- \, d\mu < \infty$ *для некоторого* $m \geqslant 1$, *то*

$$\lim_n{}_{\uparrow} \int_\Omega \xi_n \, d\mu \ = \ \int_\Omega \eta \, d\mu.$$

⊫ По условию теоремы, начиная с $n = m$, все интегралы $\int \xi_n \, d\mu$ существуют. Если $\int \xi_k \, d\mu = +\infty$ для некоторого $k > m$, то при $\forall n \geqslant k$ в силу свойства монотонности интеграла

$$\int_\Omega \eta \, d\mu \ \geqslant \ \int_\Omega \xi_n \, d\mu \ \geqslant \ \int_\Omega \xi_k \, d\mu \ = \ \infty.$$

Поэтому достаточно рассмотреть случай $\xi_n \geqslant 0$ и $\int \xi_n \, d\mu < \infty,\ n \geqslant 1$, а в общем случае перейти к последовательности $\xi_n - \xi_m,\ n \geqslant m$.

Для каждой функции ξ_n построим возрастающую последовательность простых функций $h_{nk} \in \mathfrak{P}_+(\Omega, \mathscr{F})$: $\lim_k{}_{\uparrow} h_{nk} = \xi_n$. Зададим новую последовательность $z_k = \max\{h_{1k}, \ldots, h_{kk}\}$, которая, очевидно, не убывает, а посему существует измеримая функция $\zeta = \lim_k{}_{\uparrow} z_k$. Кроме того, при $\forall\, k \geqslant n \geqslant 1$

$$h_{nk} \leqslant z_k \leqslant \xi_k \leqslant \eta.$$

Переходя здесь к пределу по $k \to \infty$, получаем, что $\xi_n \leqslant \zeta \leqslant \eta$ при $\forall n \geqslant 1$. Откуда, полагая $n \to \infty$, получаем $\zeta = \lim_k{}_{\uparrow} z_k = \eta$ и

$$\int_\Omega \eta \, d\mu \ = \ \lim_k{}_{\uparrow} \int_\Omega z_k \, d\mu \ \leqslant \ \lim_k{}_{\uparrow} \int_\Omega \xi_k \, d\mu \ \leqslant \ \int_\Omega \eta \, d\mu. \qquad ⫤$$

⚠ Теорема о монотонной сходимости верна и для последовательности, предельные значения которой могут равняться $+\infty$.

535| **Лемма.** (?!) *Пусть* $\xi_n \geqslant 0,\ n \geqslant 1$, *тогда*

$$\int_\Omega \sum\nolimits_1^\infty \xi_n \, d\mu \ = \ \sum\nolimits_1^\infty \int_\Omega \xi_n \, d\mu.$$

536| **Следствие.** *Пусть* $\xi(\omega) = \sum_1^\infty x_j \, \dot{\mathrm{I}}_{B_j}(\omega)$ *— счётнозначная функция, принимающая постоянные значения на множествах разбиения* $\Omega = \biguplus_j B_j$, *тогда*

$$\int_\Omega \xi \, d\mu \ = \ \sum\nolimits_1^\infty x_j \, \mu(B_j),$$

причём $\left|\int_\Omega \xi\, d\mu\right| < \infty$ *т. т. т. когда указанный ряд сходится абсолютно.*

⊫ В силу предыдущей леммы

$$\int_\Omega |\xi|\, d\mu \;=\; \sum\nolimits_1^\infty \int_\Omega |x_j|\, \dot{\mathrm{I}}_{B_j}\, d\mu \;=\; \sum\nolimits_1^\infty |x_j|\, \mu(B_j).$$

Поэтому $\int_\Omega |\xi|\, d\mu < \infty$ т. т. т. когда $\sum_1^\infty |x_j|\, \mu(B_j) < \infty$. В этом случае ряд $\sum_1^\infty x_j\, \mu(B_j)$ суммируем при любой перестановке слагаемых, следовательно,

$$\sum\nolimits_1^\infty x_j\, \mu(B_j) \;=\; \sum_{j\in J^+} x_j\, \mu(B_j) + \sum_{j\in J^-} x_j\, \mu(B_j) \;=\; \int_\Omega \xi^+\, d\mu - \int_\Omega \xi^-\, d\mu,$$

где $J^+ = \{j:\; x_j > 0\}$, $J^- = \{j:\; x_j < 0\}$.

Пусть теперь $\int_\Omega \xi^+\, d\mu = \infty$, $\int_\Omega \xi^-\, d\mu < \infty$; положим $x_j^+ = \max(0, x_j)$, $x_j^- = -\min(0, x_j)$. Снова по предыдущей лемме ряд $\sum_j x_j^- \mu(B_j) < \infty$, тогда как ряд $\sum_j x_j^+ \mu(B_j) = \infty$. Предположив, что $|\sum_j x_j \mu(B_j)| < \infty$, в силу возможности почленного сложения сходящихся рядов приходим к противоречию:

$$\sum\nolimits_j x_j^+ \mu(B_j) \;=\; \sum\nolimits_j x_j \mu(B_j) + \sum\nolimits_j x_j^- \mu(B_j) \;<\; \infty.$$

Стало быть, $\sum_j x_j \mu(B_j) = +\infty = \int_\Omega \xi\, d\mu$. ⫤

Следующее утверждение даёт одно из наиболее действенных средств, наряду с теоремой о монотонной сходимости и теоремой Лебега для анализа последовательностей интегралов.

537 **Лемма.** [*П. Фату.*] *Если* $\xi_n \geqslant 0$ — *последовательность измеримых функций* ($n \geqslant 1$), *то*

$$\varliminf_n \int_\Omega \xi_n\, d\mu \;\geqslant\; \int_\Omega \varliminf_n \xi_n\, d\mu.$$

⊫ Напомним, что $\varliminf_n \xi_n = \lim_{n\uparrow} (\inf_{k\geqslant n} \xi_k)$. Таким образом, для последовательности $\gamma_n = \inf_{k\geqslant n} \xi_k$ выполняются условия теоремы о монотонной сходимости и, кроме того, $\gamma_n \leqslant \xi_n$, поэтому

$$\int_\Omega \varliminf_n \xi_n\, d\mu \;=\; \int_\Omega \lim_{n\uparrow} \gamma_n\, d\mu \;=\; \lim_{n\uparrow} \int_\Omega \gamma_n\, d\mu \;\leqslant\; \varliminf_n \int_\Omega \xi_n\, d\mu. \qquad ⫤$$

538 **Следствие.**

a) *Если* $\xi_n \geqslant \eta$, $n \geqslant 1$, *причём* $\int \eta\, d\mu > -\infty$, *то*

$$\varliminf_n \int_\Omega \xi_n\, d\mu \;\geqslant\; \int_\Omega \varliminf_n \xi_n\, d\mu.$$

б) *Если* $\xi_n \leqslant \eta,\ n \geqslant 1$, *причём* $\int \eta\, d\mu < +\infty$, *то*

$$\varlimsup_n \int_\Omega \xi_n\, d\mu \ \leqslant\ \int_\Omega \varlimsup_n \xi_n\, d\mu.$$

в) *Если* $|\xi_n| \leqslant \eta,\ n \geqslant 1$, *причём* $\int \eta\, d\mu < +\infty$, *то*

$$\int_\Omega \varliminf_n \xi_n\, d\mu \ \leqslant\ \varliminf_n \int_\Omega \xi_n\, d\mu \ \leqslant\ \varlimsup_n \int_\Omega \xi_n\, d\mu \ \leqslant\ \int_\Omega \varlimsup_n \xi_n\, d\mu.$$

⊨⇒ б) Если $\int \eta\, d\mu = -\infty$, то по свойству монотонности $\int \xi_n\, d\mu \leqslant \int \eta\, d\mu = -\infty$ и, аналогично, $\int \varlimsup_n \xi_n\, d\mu \leqslant \int \eta\, d\mu = -\infty$. Поэтому достаточно рассмотреть случай $\left|\int \eta\, d\mu\right| < \infty$. Рассмотрим последовательность функций $\zeta_n = \eta - \xi_n \geqslant 0,\ n \geqslant 1$. Заметим сначала, что $\int \zeta_n\, d\mu = \int \eta\, d\mu - \int \xi_n\, d\mu$ и $\varliminf_n \zeta_n = \eta - \varlimsup_n \xi_n$. Следовательно, по лемме Фату

$$\int_\Omega \eta\, d\mu - \varlimsup_n \int_\Omega \xi_n\, d\mu \ \geqslant\ \int_\Omega \eta\, d\mu - \int_\Omega \varlimsup_n \xi_n\, d\mu,$$

что доказывает б), т.к. $\left|\int \eta\, d\mu\right| < \infty$. Утверждение а) доказывается переходом к $-\xi_n$, в) следует из а), б). ⇐⫤

Утверждение следующей теоремы полезно при анализе возможности предельного перехода под знаком интеграла. При этом зачастую требуется расширенный вариант сходимости последовательности функций. Пусть $\mathfrak{A}$ — некоторое высказывание об элементах ω пространства с мерой $(\Omega, \mathscr{F}, \mu)$, которое выделяет в Ω измеримое подмножество $A \in \mathscr{F}$ тех ω, на которых $\mathfrak{A}$ истинно.

Определение. Говорят, что высказывание $\mathfrak{A}$ выполняется *почти всюду* относительно меры μ или μ*-почти всюду*, если подмножество, противоположное $\mathfrak{A}$, имеет меру $\mu(A^c) = 0$. Кратко записывается как $\mathfrak{A}$ (п.в.), или $\mathfrak{A}$ (μ-п.в.), или $\mathfrak{A}$ (a.e.) (almost everywhere), или $\mathfrak{A}\,[\mu]$.

539| Примеры.

1) $\xi_n \underset{n\to\infty}{\to} \xi_0$ [п.в.] $\Leftrightarrow \mu\big(\omega : \{\nexists \lim_n \xi_n(\omega)\} \cup \{\exists \lim_n \xi_n(\omega) \neq \xi_0(\omega)\}\big) = 0$.

2) $\xi = \eta\,[\mu] \quad\Leftrightarrow\quad \mu(\omega : \xi(\omega) \neq \eta(\omega)) = 0$. ◎

540| **Теорема.** [*О мажорируемой сходимости. А.Л. Лебег.*] *Если измеримые функции* $\eta, \xi_0, \xi_1, \ldots$ *таковы, что* μ*-почти всюду*

a) $\lim_n \xi_n = \xi_0$, b) $|\xi_n| \leqslant \eta, n \geqslant 1$, *и* c) $0 \leqslant \int_\Omega \eta\, d\mu < \infty$,

то ξ_0 *интегрируема и*

$$\text{i)}\quad \lim_n \int_\Omega \xi_n\, d\mu \;=\; \int_\Omega \xi_0\, d\mu, \qquad \text{ii)}\quad \lim_n \int_\Omega |\xi_n - \xi_0|\, d\mu \;=\; 0.$$

$\Rrightarrow$ Предположим сначала, что условия теоремы верны при $\forall\, \omega \in \Omega$. Тогда по предыдущему следствию

$$\int_\Omega \lim_n \xi_n\, d\mu \;\leqslant\; \underline{\lim_n} \int_\Omega \xi_n\, d\mu \;\leqslant\; \overline{\lim_n} \int_\Omega \xi_n\, d\mu \;\leqslant\; \int_\Omega \lim_n \xi_n\, d\mu,$$

что эквивалентно утверждению i) теоремы. Утверждение ii) следует из i), если заметить, что $|\xi_n - \xi_0| \leqslant 2\eta$.

Пусть измеримое множество $A \in \mathscr{F}$ состоит из всех тех $\omega \in \Omega$, в которых выполняются условия теоремы относительно сходимости $\xi_n(\omega) \to \xi_0(\omega)$ и мажорируемости $|\xi_n(\omega)| \leqslant \eta(\omega)$. Так как для σ-полуаддитивной меры счётное объединение множеств меры нуль имеет меру нуль — $\mu(A^c) = 0$, то по свойству интеграла от «почти совпадающих» функций для $\forall\, n \geqslant 0$

$$\int_\Omega \xi_n\, \dot{\mathrm{I}}_A\, d\mu \;=\; \int_\Omega \xi_n\, d\mu \quad \text{и} \quad \int_\Omega \eta\, \dot{\mathrm{I}}_A\, d\mu \;=\; \int_\Omega \eta\, d\mu.$$

Кроме того, $\lim_n \xi_n(\omega)\, \dot{\mathrm{I}}_A(\omega) = \xi_0(\omega)\, \dot{\mathrm{I}}_A(\omega)$ и $\xi_n(\omega)\, \dot{\mathrm{I}}_A(\omega) \leqslant \eta(\omega)\, \dot{\mathrm{I}}(\omega)$ при всех $\omega \in \Omega$, что доказывает теорему. $\Lleftarrow$

✧ **Замена переменной в интеграле Лебега.** Измеримая функция, заданная на пространстве с мерой $\xi : (\Omega, \mathscr{F}, \mu) \mapsto (\mathfrak{X}, \mathscr{Q})$, порождает (индуцирует) на $(\mathfrak{X}, \mathscr{Q})$ меру по формуле

$$\mu_\xi(B) \;=\; \mu(\xi^{-1}(B)), \quad B \in \mathscr{Q}.$$

541| **Лемма.** (?!) *Функция μ_ξ есть мера на $(\mathfrak{X}, \mathscr{Q})$.*

Интеграл от суперпозиции двух функций $\eta \circ \xi(\omega) = \eta(\xi(\omega))$ можно вычислить двумя способами. Первый — относительно меры μ на $(\Omega, \mathscr{F}, \mu)$, второй — относительно меры μ_ξ, индуцированной отображением ξ на пространстве её значений. Следующая теорема показывает, что оба способа совпадают.

542| **Теорема.** [*О замене переменных.*] *Пусть измеримая функция $\xi : (\Omega, \mathscr{F}, \mu) \mapsto (\mathfrak{X}, \mathscr{Q})$ задана на пространстве с мерой, $\eta : (\mathfrak{X}, \mathscr{Q}) \mapsto (\mathbb{R}^1, \mathscr{B})$ — измеримая по Борелю функция. Тогда*

$$\int_{\mathfrak{X}} \eta(x) \mu_\xi(dx) \;=\; \int_\Omega \eta(\xi(\omega))\, \mu(d\omega).$$

$\Rrightarrow$ Воспользуемся схемой построения интеграла Лебега.

✓ Пусть сначала $\eta = \dot{\mathrm{I}}_B$, с $B \in \mathscr{Q}$. Тогда по определению

$$\int_{\mathfrak{X}} \eta(x)\mu_\xi(dx) \;=\; \mu_\xi(B) \;=\; \mu(\xi^{-1}(B)).$$

С другой стороны, $\eta(\xi(\omega)) = \dot{\mathrm{I}}(\omega; \xi(\omega) \in B) = \dot{\mathrm{I}}(\omega; \xi^{-1}(B))$. Следовательно,

$$\int_\Omega \eta(\xi(\omega))\, \mu(d\omega) \;=\; \mu(\xi^{-1}(B)),$$

т.е. требуемое равенство выполняется для индикаторных функций, а потому и для простых неотрицательных функций.

✓ В силу теоремы о монотонной сходимости равенство будет справедливо для монотонных пределов простых неотрицательных функций, т.е. для любых измеримых неотрицательных функций.

✓ Для произвольной функции η по предыдущему $\int_{\mathfrak{X}} \eta^{\pm}\, d\mu_\xi = \int \eta^{\pm} \circ \xi\, d\mu$. Если существует $\int_{\mathfrak{X}} \eta\, d\mu_\xi$, то, например, $\int_{\mathfrak{X}} \eta^{+}\, d\mu_\xi < \infty$ и, следовательно, $\int \eta^{+} \circ \xi\, d\mu < \infty$. То есть если существует один интеграл, то существует другой интеграл и они равны. ⫤

В качестве примера приведём здесь формулу для простой линейной замены в интеграле относительно меры Лебега.

Определения. Интеграл Лебега от $\xi : \mathbb{R}^k \mapsto \mathbb{R}^1$ относительно меры Лебега–Стилтьеса μ_F, порождённой функцией F, называется *интегралом Лебега–Стилтьеса* и обозначается $\int_{\mathbb{R}^k} \xi(\vec{x})\, dF(\vec{x})$. Интеграл Лебега–Стилтьеса от функции $\xi \cdot \dot{\mathrm{I}}_A$ обозначается $\int_A \xi(\vec{x})\, dF(\vec{x})$, если такая запись не несёт двусмысленности (см. ниже 569, стр. 498).

Интеграл относительно меры Лебега будем обозначать $\int_{\mathbb{R}^k} \xi(\vec{x})\, d\vec{x}$. Чтобы отделить интеграл Лебега по отрезку $[a; b]$ на прямой от аналогичного интеграла Римана, последний будем обозначать через $\int_a^b \xi(x)\, dx$.

543| Лемма. *Для* $\forall\, b \neq 0, \vec{c} \in \mathbb{R}^k$

$$\int\limits_{\mathbb{R}^k} \eta(\vec{x})\, d\vec{x} \;=\; |b|^k \int\limits_{\mathbb{R}^k} \eta(b\,\vec{t} + \vec{c})\, d\vec{t},$$

если хотя бы один из интегралов существует.

⫥ Рассмотрим только одномерный случай, и пусть $b > 0$. Пусть λ — мера Лебега на прямой, т.е. $\lambda(u; y] = y - u$. Измеримая функция $\xi(t) = bt + c$

индуцирует меру, для которой

$$b\,\mu_\xi(u;y] \;=\; b\,\lambda\big(\xi^{-1}(u;y]\big) \;=\; b\,\lambda\Big(\frac{u-c}{b};\frac{y-c}{b}\Big] \;=\; y-u \;=\; \lambda(u;y].$$

По теореме Каратеодори, если меры совпадают на интервалах, то они совпадают на всех борелевских подмножествах: $b\,\mu_\xi(A)=\lambda(A), A\in\mathscr{B}$. По конструкции интеграла Лебега отсюда следует, что $b\int_{\mathbb{R}}\eta\,d\mu_\xi=\int_{\mathbb{R}}\eta\,d\lambda$ для измеримой функции η. В соответствии с формулой замены переменных получаем

$$b\int_{\mathbb{R}}\eta(bt+c)\,dt \;=\; b\int_{\mathbb{R}}\eta(x)\,\mu_\xi(dx) \;=\; \int_{\mathbb{R}}\eta(x)\,dx.$$

Многомерный случай и случай с $b<0$ оставляем читателю. ⫤

544| Упр. Докажите, что мера Лебега инвариантна относительно сдвигов и однородна: $\lambda(bA+\vec{c})=|b|^k\lambda(A),\ A\in\mathscr{B}^k,\ \vec{c}\in\mathbb{R}^k,\ b\in\mathbb{R}^1$.

▸ Рассмотрим теперь произвольное взаимно однозначное преобразование ξ. Обозначим через $\lfloor a;b\rceil$ любой возможный интервал на числовой прямой — $[a;b]$, или $(a;b]$, или $(-\infty;b)$ и т.п.

545| **Теорема.** *Если функция $\xi:\lfloor a;b\rceil\mapsto\lfloor u;v\rceil$ строго возрастает и обратная функция $\tilde{\xi}(x)=\inf\{t\in\lfloor a;b\rceil:\xi(t)\geqslant x\}$, $x\in\lfloor u;v\rceil$, абсолютно непрерывна, то*[‡] *для любой борелевской функции $\eta:\lfloor u;v\rceil\to\mathbb{R}^1$*

$$\int\limits_{\lfloor a;b\rceil}\eta(\xi(t))\,dt \;=\; \int\limits_{\lfloor u;v\rceil}\eta(x)\,|\tilde{\xi}\,'(x)|\,dx. \tag{16}$$

⫢ По теореме 542, стр. 483, необходимо найти сначала меру λ_ξ, индуцированную функцией ξ и мерой Лебега λ. Заметим, что прообраз $\xi^{-1}(-\infty;x]=$ $=\lfloor a;\tilde{\xi}(x)]$ или $=\lfloor a;\tilde{\xi}(x)\big)$ в зависимости от того, $\xi(\tilde{\xi}(x))\leqslant x$ или $\xi(\tilde{\xi}(x))>x$. В любом случае индуцированная мера интервала $\lambda_\xi(x;y]=$ $=\tilde{\xi}(y)-\tilde{\xi}(x)$. Стало быть, λ_ξ есть мера Лебега–Стилтьеса на $\lfloor u;v\rceil$, порождённая абсолютно непрерывной функцией $\tilde{\xi}$, что вместе с утверждением следствия 564 и равенством (20), стр. 493, доказывает теорему. ⫤

Доказательство многомерного варианта этой теоремы требует несколько больших усилий, поэтому мы приведём только формулировку результата (см., например, [12, стр. 78], [11, стр. 162], [23, т. 2, стр. 388]).

(‡) Здесь, конечно, $\tilde{\xi}\,'(x)\geqslant 0$; формула содержит модуль $|\tilde{\xi}\,'(x)|$, дабы подчеркнуть возможность её применения для убывающих функций (с $\tilde{\xi}(x)=\inf\{t:\xi(t)\leqslant x\}$).

546| **Теорема.** *Пусть функция* $\vec{\xi} = (\xi_1, \ldots, \xi_k) : \Omega \mapsto E$ *взаимно однозначно отображает борелевскую область* $\Omega \subseteq \mathbb{R}^k$ *в борелевскую область* $E \subseteq \mathbb{R}^k$, $\vec{\xi}^{-1} = (\tilde{\xi}_1, \ldots, \tilde{\xi}_k)$ *— обратная функция. Если компоненты функции* $\vec{\xi}^{-1}$ *непрерывно дифференцируемы и всюду на* E *якобиан*

$$J(x_1, \ldots, x_k) \;=\; \left\| \left(\frac{\partial \tilde{\xi}_j(x_1, \ldots, x_k)}{\partial x_n} \right)_{nj} \right\| > 0,$$

то для борелевской функции $\eta : E \mapsto \mathbb{R}^1$

$$\int\limits_{\Omega} \eta(\vec{\xi}(\vec{\omega}))\, d\vec{\omega} \;=\; \int\limits_{E} \eta(\vec{x}) J(\vec{x})\, d\vec{x},$$

если хотя бы один из интегралов сходится абсолютно.

§ 5. Некоторые интегральные неравенства

Самое востребованное в теории вероятностей интегральное неравенство Маркова почти не требует доказательства.

547| **Лемма.** (?!) [*Неравенство Маркова.*] *Пусть измеримая функция* $\eta \geqslant 0$, *тогда при* $\forall Y > 0$

$$\mu(\eta > Y) \;\leqslant\; \frac{1}{Y} \int\limits_{\Omega} \eta\, \dot{\mathrm{I}}\,(\eta > Y)\, d\mu \;\leqslant\; \frac{1}{Y} \int\limits_{\Omega} \eta\, d\mu. \tag{17}$$

Обсуждение других популярных неравенств, потребность в которых возникает при изучении интегральных выражений, начнём с неравенства Йенсена для выпуклых функций (см. определения и свойства на стр. 530). Это неравенство удобнее формулировать для вероятностной меры, т.е. меры с $\mu(\Omega) = 1$. Для такой меры можно использовать привычное теоретико-вероятностное обозначение для интеграла от функции ξ — $\int_\Omega \xi\, d\mu =: \mathbf{E}\,\xi$.

Учитывая потребности многомерного случая, определим

$$\mathbf{E}\,\vec{\xi} \;=\; \mathbf{E}(\xi_1, \ldots, \xi_k) \;:=\; (\mathbf{E}\,\xi_1, \ldots, \mathbf{E}\,\xi_k)\,.$$

548| [*Неравенство Йенсена.*] *Пусть* $(\Omega, \mathscr{F}, \mu)$ *— пространство с вероятностной мерой,* $h : (\vec{a}; \vec{c}) \mapsto \mathbb{R}^1$ *— выпуклая (книзу) функция на открытом, возможно бесконечном, параллелепипеде* $(\vec{a}; \vec{c})$ *евклидова пространства* $\mathbb{R}^k$. *Если измеримая по Борелю функция* $\vec{\xi} : \Omega \mapsto (\vec{a}; \vec{c})$ *такова, что* $\left|\mathbf{E}\,\xi_j\right| < \infty$, $j = \overline{1, k}$, *то справедливо соотношение*

$$\mathbf{E}\left[h \circ \vec{\xi}\,\right] \;\geqslant\; h\big(\mathbf{E}\,\vec{\xi}\,\big) \tag{18}$$

со строгим неравенством, когда функция h *строго выпуклая и* $\mu(\vec{\xi} = \vec{x}_0) < 1$ *при* $\forall\, \vec{x}_0 \in (\vec{a}; \vec{c})$.

Для вогнутой функции h *справедливо противоположное неравенство.*

⊫ Выберем $\vec{x}_0 = \mathbf{E}\,\vec{\xi}$; очевидно, $\vec{a} < \vec{x}_0 < \vec{c}$. Построим опорную («касательную») плоскость (⊛), стр. 531, к графику функции $y = h(\vec{x})$ в т. $\vec{x}_0$:

$$h(\vec{x}) \;\geqslant\; \vec{b}^{\flat}(\vec{x} - \vec{x}_0) + h(\vec{x}_0), \quad \forall\, \vec{x} \in (\vec{a}; \vec{c}),$$

с некоторым $\vec{b} \in \mathbb{R}^k$. Проинтегрировав обе части этого неравенства (с заменой $\vec{x} \dashrightarrow \vec{\xi}$) и учтя, что $\mathbf{E}(\vec{\xi} - \vec{x}_0) = \vec{0}$, получим требуемое неравенство. ⫤

Функция $h(x) = x^a, x > 0$, выпуклая при $a > 1$. Поэтому для любой вероятностной меры μ и измеримой функции ξ справедливо неравенство $\left(\int_\Omega |\xi|\, d\mu\right)^a \leqslant \int_\Omega |\xi|^a\, d\mu$. Применяя это неравенство к $a = \frac{r}{s}$, $r > s > 0$, и заменяя функцию $\xi \dashrightarrow \xi^s$, получаем следующее утверждение.

549| [*Неравенство Ляпунова.*] *Пусть* $\xi : (\Omega, \mathscr{F}, \mu) \mapsto (\mathbb{R}^1, \mathscr{B})$ *— измеримая функция на пространстве с вероятностной мерой. Тогда для любых* $r > s > 0$

$$\left(\int_\Omega |\xi|^s\, d\mu\right)^{1/s} \leqslant \left(\int_\Omega |\xi|^r\, d\mu\right)^{1/r}$$

со строгим неравенством, если $\mu(\xi = x_0) \neq 1$ *при* $\forall\, x_0 \in \mathbb{R}^1$. *В частности, если* $\int_\Omega |\xi|^r\, d\mu < \infty$, *то* $\int_\Omega |\xi|^s\, d\mu < \infty$ *для* $\forall\, s \in (0; r]$.

550| Упр. Приведите пример с неограниченной мерой, когда неравенство Ляпунова нарушается.

Рассмотрим функцию $h(x_1, x_2) = x_1^a x_2^{1-a}$, $x_1 > 0, x_2 > 0$, с $0 < a < 1$. Матрица вторых производных функции $-h$ в точке $x_1 = u, x_2 = v$ равна $a(1-a)u^{a-2}v^{-1-a}\begin{pmatrix} v^2 & -uv \\ -uv & u^2 \end{pmatrix}$. Очевидно, эта матрица неотрицательно определена, т.е. функция h вогнута, поэтому, согласно неравенству Йенсена, $\mathbf{E}\left[|\xi_1|^a |\xi_2|^{1-a}\right] \leqslant \big(\mathbf{E}\,|\xi_1|\big)^a \big(\mathbf{E}\,|\xi_2|\big)^{1-a}$. Принято использовать это неравенство с $a = \dfrac{1}{p}, 1 - a = \dfrac{1}{q}, \dfrac{1}{p} + \dfrac{1}{q} = 1$, и $|\xi_1|^a \dashrightarrow |\xi_1|$, $|\xi_2|^{1-a} \dashrightarrow |\xi_2|$.

551| [*Неравенство Гёльдера.*] *Пусть* $\xi_1, \xi_2 : (\Omega, \mathscr{F}, \mu) \mapsto (\mathbb{R}^1, \mathscr{B})$ *—*

измеримые функции на пространстве с мерой. Тогда при $p, q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$,

$$\int_\Omega |\xi_1 \xi_2|\, d\mu \ \leqslant\ \Big(\int_\Omega |\xi_1|^p\, d\mu\Big)^{1/p}\Big(\int_\Omega |\xi_2|^q\, d\mu\Big)^{1/q}.$$

⤃ Для вероятностной меры μ неравенство было доказано в абзаце, предшествующем теореме. Для произвольной меры μ неравенство выполняется тривиальным образом, если хотя бы один интеграл в правой части равен $+\infty$. Если оба эти интеграла конечны, достаточно применить неравенство Гёльдера к функциям $\xi_1 = 1, \xi_2 = \xi_2/|\xi_1|^{p/q}$ с вероятностной мерой (см. 556 ниже) $\tilde{\mu}(A) = \int_A |\xi_1|^p\, d\mu/C$, где $C = \int_\Omega |\xi_1|^p\, d\mu$, и

$$\int_\Omega 1 \frac{|\xi_2|}{|\xi_1|^{p/q}}\, \frac{|\xi_1|^p}{C} d\mu \ \leqslant\ 1 \cdot \Big(\int_\Omega \frac{|\xi_2|^q}{|\xi_1|^p}\, \frac{|\xi_1|^p}{C} d\mu\Big)^{1/q}.$$

Осталось только заметить, что $p - \frac{p}{q} = 1$, $C^{1-1/q} = C^{1/p}$, и если $C = 0$, то неравенство Гёльдера, очевидно, справедливо по свойству интеграла Лебега от неотрицательной функции. ⤂

⚠ Если один из интегралов правой части неравенства Гёльдера равен нулю, а второй — $+\infty$, то это неравенство также остаётся верным (?!), если договориться считать $0 \cdot \infty = 0$.

552| Упр. [*Ляпунов.*] Пусть $\int_\Omega |\xi|^a\, d\mu < \infty$ при $\forall\, a > 0$. Докажите, что функция $G(a) = \ln\big(\int_\Omega |\xi|^a\, d\mu\big)$, $a \in (0; \infty)$, выпуклая.

Выбирая $p = q = 2$, немедленно получаем неравенство, которое в разных вариантах было доказано Коши (для рядов, 1821 г.), Буняковским (для интеграла Римана, 1859 г.) и Шварцем (для интеграла Римана, 1884 г.).

553| [*Неравенство Коши–Буняковского.*] *Пусть измеримые функции* $\xi_1, \xi_2 : (\Omega, \mathscr{F}, \mu) \mapsto (\mathbb{R}^1, \mathscr{B})$ *заданы на пространстве с мерой. Тогда*

I) $$\Big(\int_\Omega \xi_1 \xi_2\, d\mu\Big)^2 \ \leqslant\ \int_\Omega \xi_1^2\, d\mu \int_\Omega \xi_2^2\, d\mu,$$

при условии что определён интеграл в левой части неравенства, например когда оба интеграла в правой части конечны.

II) *Если интегралы в правой части* I) *конечны, то знак равенства в* I) *достигается т. т. т. когда одна из функций, скажем* $\xi_1 = a\,\xi_2$ [μ-п.в.], *с некоторой константой* a *(возможно, равной нулю).*

⊫ II) Примем сокращённую форму записи $\int_\Omega \eta\, d\mu = \int \eta$. В одну сторону («когда») утверждение теоремы очевидно. Предположим теперь, что в I) имеет место равенство, причём $\int \xi_2^2 \neq 0$ (в противном случае утверждение верно с $\xi_2 = 0 \cdot \xi_1$), и пусть для определённости $z = \int \xi_1 \xi_2 < 0$.

Положим $c = -\big[\int \xi_1^2 / \int \xi_2^2\big]^{1/2}$. После возведения в квадрат правой части равенства $0 = \Big(\sqrt{\int \xi_1^2} + c\sqrt{\int \xi_2^2}\Big)^2$ и применения равенства I) (с учётом знака z) получим $\int (\xi_1 - c\,\xi_2)^2 = 0$. По свойству интеграла Лебега отсюда следует, что $\xi_1 = c\xi_2$ [μ-п.в.]. ⫤

554| Упр. Докажите 553, воспользовавшись неравенством $\int (\xi_1 - a\xi_2)^2 \geqslant 0$, справедливым при $\forall\, a \in \mathbb{R}^1$.

555| [*Неравенство Минковского.*] *Пусть* $\xi_1, \xi_2 : (\Omega, \mathscr{F}, \mu) \mapsto (\mathbb{R}^1, \mathscr{B})$ — *измеримые функции на пространстве с мерой. Тогда для* $\forall\, p \geqslant 1$

$$\Big(\int_\Omega |\xi_1 + \xi_2|^p\, d\mu\Big)^{1/p} \leqslant \Big(\int_\Omega |\xi_1|^p\, d\mu\Big)^{1/p} + \Big(\int_\Omega |\xi_2|^p\, d\mu\Big)^{1/p}. \tag{19}$$

⊫ Понятно, что доказательства требует только случай $p > 1$, причём когда оба интеграла в правой части конечны. В этом случае интеграл в левой части также конечен, поскольку ввиду неравенства (25), стр. 532, для выпуклых функций $|\xi_1 + \xi_2|^p \leqslant 2^{p-1}(|\xi_1|^p + |\xi_2|^p),\ p > 1$. Выберем $q = p/(p-1)$, т.е. $\dfrac{1}{q} + \dfrac{1}{p} = 1$, тогда в силу неравенства треугольника и неравенства Гёльдера имеем (используя сокращённую запись интеграла по Ω относительно μ):

$$\begin{aligned}
\int |\xi_1 + \xi_2|^p \ &= \int |\xi_1 + \xi_2||\xi_1 + \xi_2|^{p-1} \ \leqslant \\
&\leqslant \int |\xi_1||\xi_1 + \xi_2|^{p-1} + \int |\xi_2||\xi_1 + \xi_2|^{p-1} \ \leqslant \\
\leqslant \Big(\int |\xi_1|^p\Big)^{1/p} \Big(\int |\xi_1 + \xi_2|^{(p-1)q}\Big)^{1/q} &+ \Big(\int |\xi_2|^p\Big)^{1/p} \Big(\int |\xi_1 + \xi_2|^{(p-1)q}\Big)^{1/q} \ = \\
&= \Big(\int |\xi_1 + \xi_2|^p\Big)^{1-1/p} \Big[\Big(\int |\xi_1|^p\Big)^{1/p} + \Big(\int |\xi_2|^p\Big)^{1/p}\Big],
\end{aligned}$$

что эквивалентно (19) (при конечной левой части). ⫤

⚠ В анализе пространство функций, интегрируемых со степенью p (относительно меры μ), обозначают (читается «эль-пе»)

$$\mathcal{L}_p = \mathcal{L}_p(\mu) = \big\langle\, \xi : (\Omega, \mathscr{F}, \mu) \mapsto (\mathbb{R}^1, \mathscr{B}) : \ \int_\Omega |\xi|^p\, d\mu < \infty \,\big\rangle.$$

Неравенство Минковского показывает, что при $p \geqslant 1$ это пространство линейно, и для функции $\|\xi\|_p := (\int_\Omega |\xi|^p \, d\mu)^{1/p}$ справедливо неравенство треугольника: $\|\xi_1+\xi_2\|_p \leqslant \|\xi_1\|_p + \|\xi_2\|_p$. Поэтому если считать равными функции, совпадающие почти всюду (т.е. перейти к так называемому фактор-пространству), то можно сказать, что функция $\|\cdot\|_p$ задаёт в $\mathcal{L}_p$ норму (L_p-норму). Если мера μ конечна, то из неравенства Ляпунова следует включение $\mathcal{L}_p \subset \mathcal{L}_m$, $1 \leqslant p < m$.

Из неравенства Коши–Буняковского следует, что L_2-норма может быть задана скалярным произведением $(\xi,\eta) := \int_\Omega \xi\eta \, d\mu$, каковое превращает $\mathcal{L}_2$ в гильбертово пространство (снова с оговоркой об отождествлении функций, совпадающих почти всюду).

§ 6. Разложение Лебега для мер. Теорема Радона-Никодима

Для дальнейшего нам понадобятся следующие

Определения. Пусть ψ, μ — меры на измеримом пространстве $(\Omega, \mathscr{F})$. Мера ψ называется:

— *абсолютно непрерывной* относительно μ : $\psi \ll \mu$, если для любого μ-нулевого множества $N \in \mathscr{F}$ мера $\psi(N) = 0$: $\mu(N) = 0 \Rightarrow \psi(N) = 0$,

— *сингулярной* относительно μ (μ-сингулярной): $\psi \perp \mu$, если найдётся множество $N \in \mathscr{F}$ такое, что $\psi(N^c) = \mu(N) = 0$.

Пусть $\xi(\omega) : (\Omega, \mathscr{F}, \mu) \mapsto (\overline{\mathbb{R}}^1_+, \mathscr{B})$ — измеримая неотрицательная функция, заданная на пространстве с мерой. Определим функцию множеств

$$M_\xi(B) = \int\limits_B \xi \, d\mu := \int\limits_\Omega \xi \, \dot{\mathrm{I}}_B \, d\mu, \qquad B \in \mathscr{F}.$$

Эту функцию часто называют *неопределённым интегралом* ξ. Заметим, что если ξ интегрируема, то, очевидно, $M_\xi(B) \leqslant \int_\Omega \xi \, d\mu < \infty$ при $\forall B \in \mathscr{F}$.

556| **Теорема.** *Для любой $\mathscr{F}$-измеримой функции $\xi \geqslant 0$:*

a) *функция множеств M_ξ есть мера на $(\Omega, \mathscr{F})$;*

b) *мера M_ξ абсолютно непрерывна относительно меры μ;*

c) *если μ — σ-конечная мера, то мера M_ξ σ-конечна т. т. т. когда функция ξ п.в. конечна: $\mu(\xi = \infty) = 0$;*

d) *для $\mathscr{F}$-измеримой функции η интеграл $\int_\Omega \eta \, dM_\xi = \int_\Omega \eta \, \xi \, d\mu.$*

$\Rrightarrow$ a) Пусть измеримое множество $B = \biguplus_n A_n$, тогда $\dot{\mathrm{I}}_B = \sum_n \dot{\mathrm{I}}_{A_n}$. Из леммы 535 следует, что M_ξ — неотрицательная σ-аддитивная функция:

$$M_\xi(B) \;=\; \int_\Omega \sum\nolimits_1^\infty \xi\, \dot{\mathrm{I}}_{A_n}\, d\mu \;=\; \sum\nolimits_1^\infty \int_\Omega \xi\, \dot{\mathrm{I}}_{A_n}\, d\mu \;=\; \sum\nolimits_1^\infty M_\xi(A_n).$$

b) Абсолютная непрерывность есть свойство интеграла Лебега: если $\mu(B)=0$, то $\xi\,\dot{\mathrm{I}}_B = 0\,[\mu$-п.в.], поэтому $\int \xi\,\dot{\mathrm{I}}_B\, d\mu = 0$.

c) Для σ-конечной меры μ найдётся семейство множеств $\Omega_n \nearrow \Omega$, $n \to \infty$, с $\mu(\Omega_n) < \infty$. Если $\mu(\xi=\infty)=0$, то $M_\xi(\xi=\infty)=0$ и для подмножества $\widehat{\Omega}_n = \Omega_n \cap \{\xi < n\}$ мера

$$M_\xi(\widehat{\Omega}_n \cup \{\xi=\infty\}) \;=\; M_\xi(\widehat{\Omega}_n) \;=\; \int_{\widehat{\Omega}_n} \xi\, d\mu \;\leqslant\; n\mu(\widehat{\Omega}_n) \;\leqslant\; n\mu(\Omega_n) \;<\; \infty.$$

Очевидно, $\widehat{\Omega}_n \cup \{\xi=\infty\} \nearrow \Omega,\ n \to \infty$.

Обратно, если найдётся последовательность подмножеств $\Omega_n \nearrow \Omega$, $n \to \infty$, с $M_\xi(\Omega_n) < \infty$, то и $M_\xi(\Omega_n \bigcap \{\xi=\infty\}) < \infty$, что в силу определения интеграла от расширенной функции возможно, только если $\mu(\Omega_n \bigcap \{\xi=\infty\}) = 0$. Поэтому $\mu(\xi=\infty) = \lim_{n}{}_{\uparrow}\, \mu(\Omega_n \bigcap \{\xi=\infty\}) = 0$.

Для индикаторных функций $\eta = \dot{\mathrm{I}}_B$ равенство d) выполняется по определению меры M_ξ; для произвольных измеримых функций — по конструкции интеграла Лебега. ⫤

557| Упр. Пусть $\int |\xi|\, d\mu < \infty$. Докажите, что если для последовательности множеств $\langle A_n \rangle_{n=1}^\infty$ предел $\lim_n \mu(A_n) = 0$, то $\lim_n \int_{A_n} \xi\, d\mu = 0$.

Такой класс мер важен, особенно в теории вероятностей. Справедлива

558| Теорема. [*И. Радон, О. Никодим.*] *Если σ-конечная мера γ абсолютно непрерывна относительно σ-конечной меры μ, то найдётся измеримая функция $\xi \geqslant 0$ такая, что*

$$\gamma(B) \;=\; \int\limits_B \xi\, d\mu, \qquad B \in \mathscr{F}.$$

Функция ξ определяется однозначно с точностью до μ-меры нуль, называется производной Радона–Никодима или плотностью меры γ относительно меры μ и обозначается $\frac{d\gamma}{d\mu} = \frac{d\gamma}{d\mu}(\omega)$.

Доказательство теоремы основано на следующем замечательном утверждении, называемом разложением Лебега для мер.

559| Лемма. [*О разложении мер. А. Лебег.*] *Пусть γ, μ — сигма-конечные меры на измеримом пространстве $(\Omega, \mathscr{F})$, тогда:*

a) *найдётся единственная (с точностью до μ-меры нуль) измеримая функция $\xi \geqslant 0$ и μ-сингулярная мера ψ такие, что*

$$\gamma(A) \;=\; \psi(A) + \int\limits_A \xi\, d\mu, \qquad A \in \mathscr{F};$$

b) *существует единственное разложение Лебега меры γ в виде суммы двух мер, одна из которых μ-непрерывная, а другая μ-сингулярная.*

⊨ Стр. 516–518. ⫤

Докажем теорему Радона–Никодима.

⊨ Пусть для множества $N \in \mathscr{F}$ мера $\mu(N) = 0$ и $\psi(N^c) = 0$. Если мера $\gamma \ll \mu$, то и $\gamma(N) = 0$. Из разложения Лебега 559 и равенства $\psi(A) = \psi(A \cap N)$ получаем, что при $\forall A \in \mathscr{F}$ мера $\psi(A) = \gamma(AN) - \int_{AN} \xi\, d\mu = 0$, т.е. $\gamma(A) = \int_A \xi\, d\mu$, что и требовалось. ⫤

560| ⚠ Можно доказать вариант теоремы Радона–Никодима для произвольной меры γ и σ-конечной меры $\mu \gg \gamma$. Мы приведём здесь доказательство теоремы Радона–Никодима с произвольной мерой γ и конечной мерой μ — вариант, востребованный при определении условного математического ожидания в теории вероятностей (это доказательство позаимствовано нами из монографии Ж. Невё [16]).

⊨ Определим класс множеств конечной γ-меры: $\mathfrak{S} = \langle\, S : \gamma(S) < \infty \,\rangle$. Выберем последовательность $S_n \in \mathfrak{S}$: $\lim_n \mu(S_n) = \sup \langle\, \mu(S), S \in \mathfrak{S} \,\rangle$. Так как класс $\mathfrak{S}$ замкнут относительно конечных объединений: $\gamma(S_1 \cup S_2) \leqslant \gamma(S_1) + \gamma(S_2) < \infty$, то можно выбрать монотонно неубывающую последовательность $S_n \subset S_{n+1}, n \geqslant 1$.

Определим меру $\gamma_n(A) = \gamma(A(S_n \setminus S_{n-1})), A \in \mathscr{F}$, которая, очевидно, ограничена, и $\gamma_n \ll \mu$. По теореме Радона–Никодима существует $\xi_n \geqslant 0$:

$$\gamma_n(A) \;=\; \int_A \xi_n\, d\mu, \qquad A \in \mathscr{F},$$

причём если выбрать $A = \{\omega \notin (S_n \setminus S_{n-1}) : \xi_n(\omega) > 0\}$, то по определению $\gamma_n(A) = 0$, и, следовательно, $\mu(A) = 0$ (в противном случае получим противоречие с представлением Радона–Никодима). Таким образом, можно считать $\xi_n(\omega) = 0$, если $\omega \notin (S_n \setminus S_{n-1})$.

Легко видеть, что т.к. $\gamma \ll \mu$, то

$$\gamma(A) \;=\; \textstyle\sum_{n=1}^\infty \gamma_n(A), \quad \text{если} \quad \mu\Big(A \cap \big(\bigcup_{n=1}^\infty S_n\big)^c\Big) = 0,$$

и $\gamma(A) = \infty$ в противном случае. Положим $\xi(\omega) = \sum_n \xi_n(\omega)$, $\omega \in \bigcup_n S_n = \bigcup_n (S_n \setminus S_{n-1})$ и $\xi(\omega) = \infty$, если $\omega \notin \bigcup_n S_n$. Из предыдущих построений ясно, что $\gamma(A) = \int_A \xi\, d\mu$, $A \in \mathscr{F}$. ⊲

Теорема Радона–Никодима и утверждение d) теоремы 556 позволяют сформулировать

561| **Следствие.** *Если мера $\gamma \ll \mu$, то для $\forall$ функции $\eta \geqslant 0$*

$$\int\limits_{\Omega} \eta\, d\gamma \;=\; \int\limits_{\Omega} \eta \frac{d\gamma}{d\mu}\, d\mu. \tag{20}$$

562| Упр. Докажите, что если меры на измеримом пространстве $(\Omega, \mathscr{F})$ связаны соотношениями $\gamma \ll \rho \ll \mu$, то для измеримой функции ξ

$$\int_{\Omega} \xi\, d\gamma \;=\; \int_{\Omega} \xi\, \frac{d\gamma}{d\rho}\, \frac{d\rho}{d\mu}\, d\mu\,.$$

В анализе также имеется понятие абсолютной непрерывности, относящееся к обычным функциям на евклидовом пространстве. Чтобы пояснить связь между этими звучащими одинаково понятиями, докажем следующее утверждение.

563| **Теорема.** *Пусть $\gamma, \mu : (\Omega, \mathscr{F}) \mapsto \overline{\mathbb{R}}^1_+$ — произвольные меры, причём мера γ конечна: $\gamma(\Omega) < \infty$, тогда следующие условия эквивалентны:*

а) $\gamma \ll \mu$, *т.е.* $\mu(N) = 0 \;\Rightarrow\; \gamma(N) = 0$;

б) *для $\forall \varepsilon > 0$ найдётся такое $\delta > 0$, что*

$$\text{для} \;\; \forall\, A \in \mathscr{F},\; \mu(A) < \delta \quad \Rightarrow \quad \gamma(A) < \varepsilon;$$

в) *если σ-алгебра $\mathscr{F} = \sigma\big(\mathcal{K}\big)$ порождена полукольцом $\mathcal{K}$, то для $\forall \varepsilon > 0\; \exists \delta > 0$, что для любого счётного семейства элементов полукольца*

$$\forall \left\langle\, A_i \,\right\rangle_1^J \subset \mathcal{K}, A_i A_j = \varnothing,\, i \neq j,\;\, J = \infty,\;\, \sum_{i=1}^{J} \mu(A_i) < \delta \quad \Rightarrow \;\; \sum_{i=1}^{J} \gamma(A_i) < \varepsilon.$$

⚠ *Условие $J = \infty$ можно заменить на условие $\forall\, J < \infty$.*

⊳ а) ⇒ б) Предположим от противного, что для некоторого $\varepsilon > 0$ и любого $n \geqslant 1$ найдётся $A_n \in \mathscr{F}$, что $\mu(A_n) < 2^{-n}$ и $\gamma(A_n) > \varepsilon$. Положим $A = \lim\limits_{n}{}_{\downarrow} \bigcup_{i=n}^{\infty} A_i$. Так как конечные меры непрерывны относительно убывающих последовательностей, то

$$\gamma(A) \;=\; \lim_n \gamma\Big(\bigcup_{i=n}^{\infty} A_i\Big) \;\geqslant\; \overline{\lim_n}\, \gamma(A_n) \;>\; \varepsilon.$$

С другой стороны, мера $\mu(A) \leqslant \sum_{i=n}^{\infty} \mu(A_i) < 2 \cdot 2^{-n}$ при любом $n \geqslant 1$, т.е. $\mu(A) = 0$, что противоречит $\gamma \ll \mu$.

б) $\Rightarrow$ а) Если $\mu(N) = 0$, то в силу б) $\gamma(N) < \varepsilon$ для $\forall \varepsilon > 0$.

б) $\Rightarrow$ в) Следует из σ-аддитивности мер.

в) $\Rightarrow$ а) Пусть $\mu(N) = 0$. Для $\varepsilon > 0$ выберем $\delta > 0$ в соответствии с условием в). В силу 479, стр. 436, найдётся такое семейство $\langle S_n \rangle_1^{\infty}$ непересекающихся подмножеств из кольца $\mathfrak{S}(\mathcal{K})$, порождённого полукольцом $\mathcal{K}$, что $S = \biguplus_n S_n \supset N$ и $\mu(S) = \sum_n \mu(S_n) < \delta$. Так как подмножества из порождённого кольца представимы в виде конечных объединений непересекающихся элементов из полукольца, то можно считать все $S_n \in \mathcal{K}$. По условию в) отсюда следует, что $\gamma(N) \leqslant \gamma(S) = \sum_n \gamma(S_n) < \varepsilon$, т.е. $\gamma(N) = 0$. ⇚

Рассмотрим теперь некоторую меру Лебега–Стилтьеса на борелевской прямой $\mathbb{R}^1$, порождённую конечной, неубывающей, непрерывной функцией F. В качестве полукольца $\mathcal{K}$ можно выбрать класс интервалов вида $(a; b]$, для которых мера Лебега–Стилтьеса (с непрерывной F) равна $\mu_F(a; b] = F(b) - F(a)$, а мера Лебега — $\lambda(a; b] = (b - a)$. Таким образом, условие в) (с $\mu = \lambda$ и $\gamma = \mu_F$) есть не что иное, как условие абсолютной непрерывности F.

Из курса анализа известно, что любая монотонная функция почти всюду (по мере Лебега) дифференцируема. Абсолютная непрерывность такой функции гарантирует возможность восстановления функции по её производной:

$$F(b) - F(a) \;=\; \int_a^b F'(x)\, dx,$$

где для определённости полагается $F'(x) = 0$ в тех точках, в которых производная не существует (см., например, [13], гл. VI, §4).

564| **Лемма.** *Мера Лебега–Стилтьеса μ_F на прямой, порождённая неубывающей, непрерывной, ограниченной функцией F, абсолютно непрерывна относительно меры Лебега λ : $\mu_F \ll \lambda$ т. т. т. когда функция F абсолютно-непрерывна. При этом*

$$\mu_F(A) \;=\; \int\limits_A F'(x)\, dx, \qquad \forall\, A \in \mathscr{B}.$$

$\Rrightarrow$ Справедливость этого равенства для любых измеримых множеств следует из теоремы Каратеодори. $\Lleftarrow$

Таким образом, для того чтобы найти плотность меры Лебега–Стилтьеса μ_F, необходимо найти производную F' во всех точках, где она существует, установив предварительно абсолютную непрерывность меры μ_F или функции F. Очень часто, не только на практике, но и в теории обязательное условие абсолютной непрерывности «по забывчивости» не проверяется. Приведём некоторые факты, касающиеся абсолютной непрерывности функций.

✦ Липшицевы функции, т.е. такие функции F, что

$$\left|F(x) - F(y)\right| \leqslant L\left|x - y\right|, \qquad \forall\, x, y \in Q \subset \mathbb{R}^1,$$

абсолютно непрерывны на Q;

✦ непрерывно дифференцируемые функции — Липшицевы (значит, абсолютно непрерывные) на любом конечном замкнутом интервале;

✦ существуют равномерно-непрерывные на конечном отрезке, ограниченные функции, не являющиеся абсолютно непрерывными на этом отрезке (например, канторова функция 52, стр. 58).

В следующей теореме, которую можно рассматривать как вариант теоремы Радона–Никодима в евклидовом пространстве, приводится условие, позволяющее в некоторой степени обезопасить себя от возможных ошибок.

Определим $F'^{+}(x) = \varliminf\limits_{h\to+0} (F(x+h)-F(x))/h$, т.е. в тех точках x, где функция F дифференцируема, $F'^{+}(x)$ равна производной функции F. По теореме Лебега этот факт имеет место для почти всех x. Поэтому всюду в дальнейшем будем писать F', подразумевая, где это необходимо, F'^{+}. Понятно, что для неубывающей функции $F'^{+}(x) \geqslant 0$. Кроме того, отметим, что F'^{+} расширенная, с возможностью принятия значения $+\infty$, борелевская функция, т.к. она есть нижний предел борелевских функций, измеримость которых, в свою очередь, следует из монотонности функции F.

565 **Теорема.** 1) *Конечная мера Лебега–Стилтьеса μ_F, заданная на борелевской прямой $(\mathbb{R}^1, \mathscr{B})$ с помощью ограниченной функции F, абсолютно непрерывна относительно меры Лебега λ т. т. т. когда функция F абсолютно непрерывна.*

2) *Пусть $F : \mathbb{R}^1 \mapsto \mathbb{R}^1$ — неубывающая непрерывная функция с*

$|F(\pm\infty)| < \infty$[‡]. *Мера Лебега–Стилтьеса* μ_F *абсолютно непрерывна относительно меры Лебега* λ *т. т. т. когда*

$$\int_{\mathbb{R}^1} F'(x)\,dx \;\geqslant\; \mu_F(\mathbb{R}^1) \;=\; F(+\infty) - F(-\infty).$$

В обоих случаях плотность $d\mu_F/d\lambda = F'$, *и для* $\forall\, A \in \mathscr{B}$ *и любой неотрицательной борелевской функции* ξ:

$$\text{✦}\quad \mu_F(A) \;=\; \int\limits_A F'(x)\,dx, \qquad\qquad \text{✦}\quad \int\limits_{\mathbb{R}^1} \xi(x)\,dF(x) \;=\; \int\limits_{\mathbb{R}^1} \xi(x)F'(x)\,dx,$$

в частности $\int_{\mathbb{R}^1} F'(x)\,dx = \mu_F(\mathbb{R}^1) = F(+\infty) - F(-\infty)$.

⚠ *Существуют меры Лебега–Стилтьеса с* $\mu_F(\mathbb{R}^1) = \infty$, *для которых утверждение теоремы неверно.*

⤃ 1) Повторяет утверждение 564. Доказательство 2) приведено в главе I, теорема 60, стр. 63, для многомерной функции распределения. ⤂

§ 7. Связь интеграла Лебега-Стилтьеса с интегралом Римана-Стилтьеса

Теорема Радона–Никодима даёт прямой путь вычисления интегралов Лебега относительно мер Лебега–Стилтьеса, если воспользоваться для этого хорошо разработанной в анализе техникой римановского интегрирования.

Если вспомнить известный признак Дарбу существования интеграла по Риману $\int_a^b h(x)\,dx$ от функции h на конечном замкнутом отрезке, то можно заметить, что нижние суммы Дарбу представляют собой интеграл Лебега от простых функций, последовательность которых можно подобрать так, что они будут, монотонно возрастая, сходиться к функции h. Поэтому если интеграл Римана существует (и автоматически конечен), то этот интеграл будет совпадать с интегралом Лебега.

Обратно, по известному критерию Лебега интегрируемости функции на отрезке, функция $h : [a;b] \mapsto \mathbb{R}^1$ интегрируема по Риману т. т. т. когда она ограничена и множество её точек разрыва имеет лебегову меру нуль. В частности, интегрируемы по Риману на отрезке функции, имеющие разрывы первого рода в конечном или счётном числе точек. Поэтому при формулировке обрат-

[‡] Как всегда, $F(\pm\infty) = \lim\limits_{x\to\pm\infty} F(x)$.

ных утверждений необходимо вводить дополнительные условия, связанные с непрерывностью функции.

Если рассмотреть интеграл Римана в несобственном смысле по открытому или бесконечному интервалу (как предел интегралов по замкнутым отрезкам), то аналогичные рассуждения показывают, что из существования конечного интеграла Римана от модуля функции $|h|$ следует существование конечного интеграла Лебега от $|h|$, что, в свою очередь, влечёт существование конечного интеграла Лебега от h. Те же рассуждения справедливы и при вычислениях интегралов по многомерным областям.

566| **Лемма.** *Пусть* $\xi : (\mathbb{R}^k, \mathscr{B}^k) \mapsto (\mathbb{R}^1, \mathscr{B})$ *— борелевская функция.*

а) *Если интеграл Римана* $\int_Q \xi(\vec{x})\, d\vec{x}$ *по замкнутой ограниченной области* $Q \subset \mathbb{R}^k$ *существует (конечен), то* ξ *интегрируема на* Q *по Лебегу.*

б) *Если конечны несобственные интегралы Римана* $\int_Q \xi(\vec{x})\, d\vec{x}$ *и* $\int_Q |\xi(\vec{x})|\, d\vec{x}$, *то функция* ξ *интегрируема на* Q *по Лебегу.*

в) *Если* ξ *интегрируема по Лебегу на* $Q \subset \mathbb{R}^k$ *и кусочно-непрерывна, то она интегрируема (в собственном или несобственном смысле) по Риману.*

Во всех случаях интегралы Римана и Лебега от ξ *совпадают.*

г) *Если функция* ξ *условно интегрируема по Риману (т.е. интеграл от* ξ *сходится, а от* $|\xi|$ *расходится), то функция* ξ *неинтегрируема по Лебегу.*

⚠ Учитывая утверждения леммы, всегда при определении интегрируемости какой-либо функции ξ подчеркивают, что интеграл от ξ должен сходиться абсолютно, т.е. $\int_{\mathbb{R}^k} \xi(\vec{x})\, d\vec{x}$ существует и $\int_{\mathbb{R}^k} |\xi(\vec{x})|\, d\vec{x} < \infty$.

Для теории вероятностей важна именно интегрируемость по Лебегу. Можно привести пример плотности f, для которой интеграл $\int_{\mathbb{R}} x f(x)\, dx$ условно конечен, однако закон больших чисел Хинчина не выполняется.

Поскольку вопрос существования интеграла Римана–Стилтьеса $\int_{-\infty}^{\infty} \xi(x)\, dF(x)$ тесно связан с взаимным расположением точек разрывов подынтегральной функции ξ и функции F, аналогичное утверждение о связи этого интеграла с интегралом Лебега–Стилтьеса удобнее формулировать для непрерывных всюду функций ξ. Приведём (без доказательства) следующее утверждение, где под функцией распределения понимается непрерывная справа неубывающая конечная функция на $\mathbb{R}^1$.

567| **Теорема.** *Пусть* $\xi : (\mathbb{R}^1, \mathscr{B}) \mapsto (\mathbb{R}^1, \mathscr{B})$ *непрерывная (стало быть, борелевская) функция,* F *— функция распределения.*

I) *Если F непрерывна в точках $a, b \in \mathbb{R}^1$, то интеграл Лебега–Стилтьеса $\int_{(a;b]} \xi(x)\, dF(x)$ совпадает с интегралом Римана–Стилтьеса $\int_a^b \xi(x)\, dF(x)$.*

II) *Если F непрерывна в точке $a \in \mathbb{R}^1$ и интеграл Римана–Стилтьеса $\int_a^\infty |\xi(x)|\, dF(x) < \infty$, то интеграл Лебега–Стилтьеса $\int_{(a;\infty)} \xi(x)\, dF(x)$ совпадает с интегралом Римана–Стилтьеса $\int_a^\infty \xi(x)\, dF(x)$.*

Для интервалов $(-\infty; c], (-\infty; +\infty)$ справедлив аналогичный результат.

568 Пример. Пусть мера Лебега–Стилтьеса определяется функцией

$$F(x) = e^x\, \dot{\mathrm{I}}\,(x < 0) + (4 - e^{-x})\, \dot{\mathrm{I}}\,(x \geqslant 0)$$

с разрывом в $x = 0$. Эту функцию можно записать в виде суммы двух функций, одна из которых (F_a) абсолютно непрерывна, а другая (F_d) кусочно-постоянна с единственным скачком в $x = 0$:

$$F(x) = F_a(x) + F_d(x) := \Big[e^x\, \dot{\mathrm{I}}\,(x < 0) + (2 - e^{-x})\, \dot{\mathrm{I}}\,(x \geqslant 0)\Big] + \Big[2\, \dot{\mathrm{I}}\,(x \geqslant 0)\Big].$$

Плотность меры, определяемой функцией F_a, очевидно, равна $(F_a(x))' = e^{-|x|}$. Вторая мера есть не что иное, как удвоенная мера Дира́ка в нуле; в полном соответствии с разложением Лебега эта мера сингулярна относительно меры Лебега. Таким образом, если существует $\int_{-\infty}^{\infty} |\xi(x)|\, e^{-|x|}\, dx < \infty$, то интеграл от функции ξ равен (см. 532, стр. 479)

$$\int_{\mathbb{R}} \xi(x)\, dF(x) = \int_{-\infty}^{\infty} \xi(x) e^{-|x|}\, dx + 2\xi(0). \qquad \circledcirc$$

569 ⚠ Запись интеграла Римана в виде $\int_a^b \xi(x)\, dx$ не несёт двусмысленности относительно принадлежности крайних точек a, b области интегрирования, поскольку мера Лебега этих точек равна нулю — присоединение их произвольным образом к области интегрирования не повлияет как на значение интеграла, так и на сам факт интегрируемости или неинтегрируемости функции ξ. Аналогичные соображения верны и для интеграла Римана–Стилтьеса $\int_a^b \xi(x)\, dF(x)$, если функция F непрерывна в крайних точках a, b. Если же F разрывна на краю, то интеграл Римана–Стилтьеса может и не существовать.

Что же касается интеграла Лебега–Стилтьеса $\int_{[a;b]} \xi(x)\, dF(x)$, то здесь вопрос существования не стоит так остро. Однако имеется проблема, связанная с интерпретацией его записи в таком виде. Напомним, что такое обозначение было введено для сокращения записи $\int_{\mathbb{R}} \dot{\mathrm{I}}_{[a;b]}\, \xi(x)\, \mu_F(dx)$, понимаемой как интеграл Лебега относительно меры Лебега–Стилтьеса μ_F, порождённой непрерывной справа (или слева) функцией F на $\mathbb{R}^1$. С другой стороны, ничто не мешает нам использовать обозначение $\int_{[a;b]} \xi(x)\,dF(x)$ для интеграла Лебега–Стилтьеса по отрезку $[a;b]$ (что обычно и делается). Легко понять, что если функция $F:\mathbb{R}^1 \mapsto \mathbb{R}^1$ непрерывна в точках a, b, то оба интеграла будут совпадать между собой.

Неоднозначность трактовки проявляется, когда сужение меры Лебега–Стилтьеса с $\mathbb{R}^1$ до области интегрирования A не совпадает с мерой, построенной на функции F, определённой на A. Например, непрерывная справа функция $F(x) = \dot{\mathrm{I}}(x;[0;\infty))$ порождает на $\Omega = \mathbb{R}^1$ меру Дира́ка δ_0, сосредоточенную в $x = 0$. Сужение δ_0 на пространство $\widetilde{\Omega} = [0;1]$ также сосредоточено в $x = 0$. С другой стороны, на $\widetilde{\Omega}$ функция $F \equiv 1$, а мера Лебега–Стилтьеса $\tilde{\mu}_F$, порождённая на $\widetilde{\Omega}$ этой функцией, тождественно равна нулю. Связано это с тем, что на $\mathbb{R}^1$ мера отрезка $[0;x]$, генерируемая непрерывной справа функцией F, равна $\mu_F[0;x] = F(x) - F(0-)$. В то же время, на отрезке $[0;1]$ генерируемая той же функцией мера равна $\tilde{\mu}_F[0;x] = F(x) - F(0)$. Поэтому запись $\int_{[0;1]} \xi\,dF$, интерпретируемая как $\int_{\mathbb{R}} \dot{\mathrm{I}}_{[0;1]}\,\xi\,dF$, даёт значение $\xi(0)$. Если же её трактовать как интеграл Лебега–Стилтьеса относительно меры $\tilde{\mu}_F$, то получим 0 при любых ξ.

Можно, конечно, ввести специальное обозначение для одного из этих интегралов, например, $\int_{[0;1]} \xi\,F(d\mathbb{R})$, $\int_{[0;1]} \xi\,d_{\mathbb{R}}F$ и тому подобное. Однако исторический опыт показывает, что такие обозначения не приживаются. Поэтому единственный надёжный способ избежать двусмысленности — всегда пояснять способ вычисления такого интеграла. Если не оговорено противное, мы будем под интегралом понимать интеграл Лебега–Стилтьеса на $\mathbb{R}^k$ с индикаторной функцией соответствующего подмножества $\mathbb{R}^k$, что обеспечит возможность применения свойства аддитивности интеграла.

570 П р и м е р. Пусть $F(x) = \llbracket x \rrbracket$ — функция, генерирующая меру, «считающую» целые точки (см. стр. 443), $\xi(x) = 1 + x$. Тогда

$$\int_{(0;1)} \xi\,dF = 0, \quad \int_{[0;1)} \xi\,dF = 1, \quad \int_{(0;1]} \xi\,dF = 2, \quad \int_{[0;1]} \xi\,dF = 3.$$

Это не противоречит непрерывности неопределённого интеграла Лебега, ибо, например, предел $\lim\limits_{a\to+0}(a;1] = (0;1] \neq [0;1]$.

Отметим, что здесь не играет роли направление непрерывности (слева или справа) функции F, важно только наличие скачка у функции F в точке из области интегрирования, что обусловливает положительность меры Лебега–Стилтьеса в этой точке. Ту же самую считающую меру порождает непрерывная слева функция $\widetilde{F}(x) = [\![x]\!]$, округляющая числа вверх. Ситуация изменится, если рассматриваются интегралы по мере Лебега–Стилтьеса на соответствующем отрезке (обозначим сейчас такие интегралы со звёздочкой $\int^*$). Например,

$$\int_{[0;1]}^{*} (1+x)\, dF = 2, \quad \int_{[0;1]}^{*} (1+x)\, d\widetilde{F} = 1 \quad (?!)\,. \qquad \circledcirc$$

§ 8. Теорема Фубини–Тонелли

Практически всегда при вычислениях многомерных интегралов используется приём изменения порядка следования кратных интегралов. Возможность его применения возникает в ситуациях, когда мера на всём пространстве представима в виде произведения мер на «координатных» пространствах. Соответствующее утверждение называется теоремой Фубини.

Определение. Пространство $(\Omega_1 \times \Omega_2,\, \mathscr{F}_1 \otimes \mathscr{F}_2,\, \mu_1 \times \mu_2)$ есть *прямое произведение пространств* $(\Omega_j, \mathscr{F}_j, \mu_j)$, $j = 1, 2$, с σ-конечными мерами, если σ-алгебра $\mathscr{F}_1 \otimes \mathscr{F}_2 = \sigma\big(\mathscr{F}_1 \times \mathscr{F}_2\big)$ — минимальная σ-алгебра, содержащая все подмножества $\Omega_1 \times \Omega_2$ вида $B_1 \times B_2$, $B_j \in \mathscr{F}_j, j = 1, 2$, и мера (также σ-конечная) на подмножествах такого вида равна $\mu_1 \times \mu_2(B_1 \times B_2) = \mu_1(B_1)\mu_2(B_2)$.

Для мер Лебега–Стилтьеса факт существования меры с требуемым свойством на произведении пространств $\mathbb{R}^k \times \mathbb{R}^m$ вытекает из того, что функция $F_1(\vec{x})F_2(\vec{y})$ удовлетворяет всем свойствам, необходимым для построения меры Лебега–Стилтьеса в пространстве $\mathbb{R}^{k+m}$. Для общего случая нам понадобятся некоторые свойства измеримых множеств и функций на прямом произведении пространств. Оговоримся: обе координаты в произведении равнозначны, поэтому все утверждения, приводимые для одной из координат, могут (и даже должны) быть переформулированы и для другой координаты.

571| **Лемма.** 1) *Для любого множества* $B \in \mathscr{F}_1 \otimes \mathscr{F}_2$ *и любого* $\omega_1 \in \Omega_1$ *сечение* B *вдоль* ω_1 *измеримо относительно* $\mathscr{F}_2$:

$$B_{\omega_1} := \{\omega_2 : (\omega_1, \omega_2) \in B\} \in \mathscr{F}_2.$$

2) *Пусть* $\xi(\omega_1, \omega_2) : (\Omega_1 \times \Omega_2, \mathscr{F}_1 \otimes \mathscr{F}_2) \mapsto (\mathfrak{O}, \mathscr{A})$ — *измеримая функция, тогда для* $\forall\, \omega_2 \in \Omega_2$ *функция* $\xi_{\omega_2}(\omega_1) : (\Omega_1, \mathscr{F}_1) \mapsto (\mathfrak{O}, \mathscr{A})$, *определяемая равенством* $\xi_{\omega_2}(\omega_1) = \xi(\omega_1, \omega_2)$, *измерима.*

⊨ 1) Покажем, что совокупность всех множеств, удовлетворяющих утверждению теоремы:

$$\mathcal{U} \;=\; \Big\langle B:\; B\in\mathscr{F}_1\otimes\mathscr{F}_2,\quad B_{\omega_1}\in\mathscr{F}_2\Big\rangle,$$

образует σ-алгебру.

- Сечение пустого множества и сечение $\Omega_1\times\Omega_2$, очевидно, входят в $\mathscr{F}_2$.
- Если $B\in\mathcal{U}$, т.е. $\{\omega_2:\ (\omega_1,\omega_2)\in B\}\in\mathscr{F}_2$, то

  $$(B^c)_{\omega_1} \;=\; \{\omega_2:\ (\omega_1,\omega_2)\notin B\} \;=\; \Omega_2\setminus\{\omega_2:\ (\omega_1,\omega_2)\in B\} \;\in\; saF_2.$$

- Если $\big\langle B_j, j\geqslant 1\big\rangle\subset\mathcal{U}$, то

  $$\Big(\bigcup\nolimits_1^\infty B_j\Big)_{\omega_1} \;=\; \Big\{\omega_2:\ (\omega_1,\omega_2)\in\bigcup\nolimits_1^\infty B_j\Big\} \;=\; \bigcup\nolimits_1^\infty (B_j)_{\omega_1} \;\in\; \mathscr{F}_2.$$

Заметим теперь, что сечение B_{ω_1} любого прямоугольника $B=A_1\times A_2$, где $A_1\in\mathscr{F}_1, A_2\in\mathscr{F}_2$, совпадает с A_2, если $\omega_1\in A_1$, и $B_{\omega_1}=\varnothing$, если $\omega_1\notin A_1$, т.е. в любом случае $(A_1\times A_2)_{\omega_1}\in\mathscr{F}$. Поэтому σ-алгебра $\mathcal{U}$ содержит произведение $\mathscr{F}_1\times\mathscr{F}_2$ и, значит, совпадает с σ-алгеброй $\mathscr{F}_1\otimes\mathscr{F}_2$.

2) Прообраз любого множества $A\in\mathscr{A}$ при отображении $\xi(\omega_1,\omega_2)$ с фиксированным ω_2

$$\{\omega_1:\xi(\omega_1,\omega_2)\in A\} \;=\; \{(\omega_1,\omega_2):\xi(\omega_1,\omega_2)\in A\}_{\omega_2} \;\in\mathscr{F}_1$$

ввиду измеримости ξ и по утверждению 1) леммы. ⫤

⚠ Утверждение 1) леммы необратимо. Существует пример неизмеримого множества, все сечения которого по обеим координатам измеримы.

Приём, использованный при доказательстве леммы, когда показывается, что совокупность множеств, удовлетворяющих требуемому свойству, образует σ-алгебру и содержит класс подмножеств, порождающих основную σ-алгебру, — этот приём применяется и при доказательстве существования прямого произведения мер.

572| **Лемма.** *Пусть* $(\Omega_i,\mathscr{F}_i,\mu_i)$, $i=1,2$, *— пространства с σ-конечными мерами. Существует единственная σ-конечная мера μ на $\mathscr{F}_1\otimes\mathscr{F}_2$ такая, что $\mu(B_1\times B_2)=\mu_1(B_1)\mu_2(B_2)$, $B_1\in\mathscr{F}_1, B_2\in\mathscr{F}_2$. При этом для $\forall\, B\in\mathscr{F}_1\otimes\mathscr{F}_2$*

$$\mu(B) \;=\; \int_{\Omega_2}\mu_1(B_{\omega_2})\mu_2(d\omega_2) \;=\; \int_{\Omega_1}\mu_2(B_{\omega_1})\mu_1(d\omega_1).$$

⊨ См. стр. 518–520. ⫤

573| **Теорема.** [*Г. Фубини.*] *Пусть* $(\Omega_i, \mathscr{F}_i, \mu_i)$, $i = 1, 2$, *— пространства с* σ*-конечными мерами. Если функция* $\xi(\omega_1, \omega_2) : (\Omega_1 \times \Omega_2, \mathscr{F}_1 \otimes \mathscr{F}_2) \mapsto (\mathbb{R}^1, \mathscr{B})$ *интегрируема по мере* $\mu_1 \times \mu_2$, *то:*

a) *для* μ_1*-почти всех* ω_1 *интеграл* $\left|\int_{\Omega_2} \xi(\omega_1, \omega_2)\, \mu_2(d\omega_2)\right| < \infty$ *(аналогично для интеграла по* ω_1 *при фиксированном* ω_2*);*

b) *интеграл* $\int_{\Omega_2} \xi(\omega_1, \omega_2)\, \mu_2(d\omega_2)$, *как функция* ω_1 *на* $(\Omega_1, \mathscr{F}_1)$, *измерим по Борелю (аналогично для интеграла по* ω_1*);*

c) *имеют место равенства*

$$\begin{aligned} \int\limits_{\Omega_1 \times \Omega_2} \xi \, d\mu_1 \times \mu_2 \;&=\; \int\limits_{\Omega_1} \Bigg(\int\limits_{\Omega_2} \xi(\omega_1, \omega_2)\, \mu_2(d\omega_2) \Bigg) \mu_1(d\omega_1) \;= \\ &=\; \int\limits_{\Omega_2} \Bigg(\int\limits_{\Omega_1} \xi(\omega_1, \omega_2)\, \mu_1(d\omega_1) \Bigg) \mu_2(d\omega_2). \end{aligned} \tag{21}$$

Часто легче проанализировать конечность повторного интеграла; отсюда

574| **Теорема.** [*Л. Тонелли.*] *Если*

$$\int\limits_{\Omega_1} \Bigg(\int\limits_{\Omega_2} \big|\xi(\omega_1, \omega_2)\big|\, \mu_2(d\omega_2) \Bigg) \mu_1(d\omega_1) \;<\; \infty,$$

то $\int\limits_{\Omega_1 \times \Omega_2} |\xi| \, d\mu_1 \times \mu_2 < \infty$ *и справедливы равенства (21).*

Для неотрицательных функций справедлива

575| **Теорема.** *Если измеримая функция* $\xi(\omega_1, \omega_2) \geqslant 0$ *для* $\mu_1 \times \mu_2$*-почти всех* (ω_1, ω_2), *то справедливы равенства (21) в том смысле, что если конечна (бесконечна) какая-либо из частей этих соотношений, то конечны (бесконечны) и две другие его части и они совпадают.*

$\Rrightarrow$ Докажем все три теоремы одновременно. Заметим, что по лемме 571 для $(\mathscr{F}_1 \otimes \mathscr{F}_2)$-измеримой функции g сечение $g_{\omega_1}(\omega_2) = g(\omega_1, \omega_2)$ при любом фиксированном ω_1 будет $\mathscr{F}_2$-измеримой функцией $\omega_2 \in \Omega_2$.

Рассмотрим сначала $\xi = \dot{\mathrm{I}}_B$, где $B \in \mathscr{F}_1 \otimes \mathscr{F}_2$, тогда при фиксированном ω_1 функция $\xi_{\omega_1}(\omega_2) \equiv \dot{\mathrm{I}}_{B_{\omega_1}}(\omega_2)$, $\omega_2 \in \Omega_2$, и интеграл $\int_{\Omega_2} \xi(\omega_1, \omega_2)\, \mu_2(d\omega_2) = \mu_2(B_{\omega_1})$. Равенство (21) верно, поскольку совпадает с определением меры $\mu_1 \times \mu_2$ (см. доказательство леммы 572, где была также показана измеримость $\mu_2(B_{\omega_1})$). Отсюда следует (21) для неотрицательных простых функций.

Пусть теперь $(\mathscr{F}_1 \otimes \mathscr{F}_2)$-измеримая функция $\xi \geqslant 0$ и монотонная последовательность неотрицательных простых функций h_n на $\Omega_1 \times \Omega_2$ такова, что $\lim\limits_{n}{\uparrow}\, h_n = \xi$. По определению интеграла Лебега и в силу предыдущего

$$\int_{\Omega_1\times\Omega_2} \xi\, d\mu_1 \times d\mu_2 \;=\; \lim_{n}{\uparrow} \int_{\Omega_1\times\Omega_2} h_n\, d\mu_1 \times d\mu_2 \;=\; \lim_{n}{\uparrow} \int_{\Omega_1} \left(\int_{\Omega_2} (h_n)_{\omega_1}\, d\mu_2 \right) d\mu_1.$$

Заметим, что при $\forall\, \omega_1 \in \Omega_1$ сечение $(h_{n-1})_{\omega_1} \leqslant (h_n)_{\omega_1} \nearrow \xi_{\omega_1},\quad n \to \infty$, где функции $(h_n)_{\omega_1}$ простые. Поэтому снова по определению интеграла Лебега $\lim\limits_{n}{\uparrow} \int_{\Omega_2} (h_n)_{\omega_1}\, d\mu_2 = \int_{\Omega_2} \xi_{\omega_1}\, d\mu_2$. В свою очередь, по теореме о монотонной сходимости для интегралов отсюда следует справедливость (21) $(\mu = \mu_1 \times \mu_2)$:

$$\int_{\Omega_1\times\Omega_2} \xi\, d\mu \;=\; \lim_{n}{\uparrow} \int_{\Omega_1} \left(\int_{\Omega_2} (h_n)_{\omega_1}\, d\mu_2 \right) d\mu_1 \;=\; \int_{\Omega_1} \left(\int_{\Omega_2} \xi_{\omega_1}(\omega_2)\, d\mu_2 \right) d\mu_1.$$

Важно отметить, что равенство здесь симметрично. Другими словами, если интеграл справа конечен, то и интеграл слева также конечен, и наоборот. То есть можно считать доказанным вариант теоремы Фубини в редакции Тонелли. Кроме того, из этого равенства следует справедливость теоремы Фубини для неотрицательных функций (даже если интегралы равны $+\infty$).

Любая измеримая функция есть разность двух неотрицательных $\xi = \xi^+ - \xi^-$, интегралы от которых по условию теоремы Фубини конечны, т.е. к ним можно применить доказанные утверждения. По свойству линейности интеграла отсюда следует справедливость теоремы Фубини. ⩤

576| Пример. Пусть $\Omega_1 = (0;1]$, $\Omega_2 = [1;\infty)$ с мерами Лебега μ_1, μ_2. Элементарными вычислениями показывается, что для функции $\xi(x,y) = e^{-xy} - 2e^{-2xy}$

$$\int_{\Omega_2} \left[\int_{\Omega_1} \xi(x,y)\, dx \right] dy \;<0\; < \int_{\Omega_1} \left[\int_{\Omega_2} \xi(x,y)\, dy \right] dx,$$

т.е. теорема Фубини «не работает». Причина, конечно, в том, что функция ξ не является абсолютно интегрируемой на $\Omega_1 \times \Omega_2$. Например, $\int_{\Omega_1} |\xi(x,y)|\, dx = \frac{1}{2y} + (e^{-2y} - e^{-y})/y$, т.е. как функция y не интегрируется на Ω_2, стало быть, условие теоремы Тонелли не выполнено. ◎

§ 9. Формула интегрирования по частям

Сначала докажем простейший вариант формулы интегрирования по частям. При выводе этой формулы используется простой алгебраический факт:

$$(*)\qquad \sum\nolimits_0^M (b_{k+1} - b_k) \;=\; \sum\nolimits_0^M b_{k+1} - \sum\nolimits_0^M b_k \;=\; b_{M+1} - b_0.$$

Отсюда легко получается так называемая формула суммирования по частям (преобразование Абеля):

$$\sum\nolimits_0^M U_k(V_{k+1} - V_k) = \sum\nolimits_0^M U_k V_{k+1} - \sum\nolimits_0^M U_k V_k =$$
$$= \sum\nolimits_0^M U_{k+1}V_{k+1} - \sum\nolimits_0^M U_k V_k - \sum\nolimits_0^M (U_{k+1} - U_k)V_{k+1} =$$
$$= -\sum\nolimits_0^M (U_{k+1} - U_k)V_{k+1} + U_{M+1}V_{M+1} - U_0 V_0.$$

Отсюда и из $(*)$ получаем равенство

$$\binom{*}{*} \qquad \sum\nolimits_0^M (U_k - U_0)(V_{k+1} - V_k) = \sum\nolimits_0^M (U_{k+1} - U_k)(V_{M+1} - V_{k+1}).$$

577| **Теорема.** *Пусть* $G(x)$ *— неубывающая и непрерывная, а* $F(x)$ *— неубывающая и непрерывная справа функции, принимающие конечные значения при* $\forall\, x \in (A; B)$, *тогда для* $\forall\, A < a < b < B$

$$\int\limits_{(a;b]} G(x)\, dF(x) + \int\limits_{(a;b]} F(x)\, dG(x) = G(b)F(b) - G(a)F(a).$$

$\Rrightarrow$ Для сокращения записи рассмотрим интервал $(a; b] = (0; 1]$. Пусть $A_{nk} = (k/2^n; (k+1)/2^n]$, $k = 0, 1, \ldots, 2^n - 1$, тогда для $\forall\, x \in (0; 1]$ при $n \to \infty$ последовательности простых неотрицательных функций

$$G_n(x) - G(0) = \sum_{k=0}^{2^n-1} \Big(G\Big(\frac{k}{2^n}\Big) - G(0)\Big)\, \dot{\mathrm{I}}\,(x; A_{nk}) \nearrow G(x) - G(0),$$
$$F(1) - F_n(x) = \sum_{k=0}^{2^n-1} \Big(F(1) - F\Big(\frac{k+1}{2^n}\Big)\Big)\, \dot{\mathrm{I}}\,(x; A_{nk}) \nearrow F(1) - F(x).$$

Действительно, если $x \in A_{nk}$ при некотором k, то $x \leqslant \dfrac{(k+1)}{2^n} \leqslant x + \dfrac{1}{2^n}$ и ввиду монотонности и непрерывности справа функции F

$$0 \leqslant \big(F(1) - F(x)\big) - \big(F(1) - F_n(x)\big) = F\Big(\frac{k+1}{2^n}\Big) - F(x) \leqslant$$
$$\leqslant F\Big(x + \frac{1}{2^n}\Big) - F(x) \to 0 \qquad (n \to \infty).$$

Аналогично для функции G.

Интегралы от $G_n(x) - G(0)$ относительно F и от $F(1) - F_n(x)$ относительно G совпадают ввиду равенства $(^*_*)$:

$$\int\limits_{(0;1]} \big(G_n(x) - G(0)\big)\, dF(x) \;=\; \sum_{k=0}^{2^n-1} \Big(G\Big(\frac{k}{2^n}\Big) - G(0)\Big)\Big(F\Big(\frac{k+1}{2^n}\Big) - F\Big(\frac{k}{2^n}\Big)\Big) \;=$$

$$= \sum_{k=0}^{2^n-1} \Big(F(1) - F\Big(\frac{k+1}{2^n}\Big)\Big)\Big(G\Big(\frac{k+1}{2^n}\Big) - G\Big(\frac{k}{2^n}\Big)\Big) = \int\limits_{(0;1]} \big(F(1) - F_n(x)\big)\, dG(x).$$

В соответствии с определением интеграла Лебега отсюда следует равенство интегралов от функций $G(x) - G(0)$ и $F(1) - F(x)$. ⫤

⚠ Функции F и G входят в формулу интегрирования по частям симметричным образом. Поэтому в условиях теоремы можно потребовать непрерывности любой из них; причём отказаться от этого условия, вообще говоря, нельзя. Рассмотрим в качестве примера две одинаковые функции $F(x) = G(x) = -\dot{\mathrm{I}}\,\big(x; (-\infty; 0)\big) + \dot{\mathrm{I}}\,\big(x; [0; \infty)\big)$, порождающие меру Лебега–Стилтьеса, сосредоточенную в точке $x = 0$: $\mu_F(\{0\}) = \mu_G(\{0\}) = 2$. Поэтому для этих функций

$$\int\limits_{(-1;1]} F(x)\, dG(x) \;=\; \int\limits_{(-1;1]} G(x)\, dF(x) \;=\; G(0)\mu_F(\{0\}) \;=\; 2,$$

в то же время $F(1)G(1) - F(-1)G(-1) = 0$. Здесь существенным оказалось то, что функции терпят разрыв в общей точке.

578| Упр. Докажите, что для непрерывной монотонной функции F

$$\int\limits_{[a;b]} F(x)\, dF(x) \;=\; \frac{1}{2}\,(F^2(b) - F^2(a)).$$

► **Интеграл по функции с ограниченной вариацией.** Самый естественный способ обобщения формулы интегрирования по частям состоит в рассмотрении наряду с неубывающими функциями их линейных комбинаций. Действительное расширение произойдёт, если мы рассмотрим класс функций, представимых в виде разности двух неубывающих функций. Из анализа известно, что эквивалентное описание такого класса состоит в требовании ограниченности полной вариации функции на отрезке $[a; b]$: $V_a^b[F] < \infty$,

$$V_a^b[F] \;:=\; \sup\Big\langle\; \sum_{k=1}^{n} \big|F(x_k) - F(x_{k-1})\big| : a = x_0 < x_1 < \ldots < x_n = b \;\Big\rangle.$$

Если функция F имеет ограниченную вариацию на $[a;b]$, то $F(x) = F_1(x) - F_2(x)$, где $F_1(x) = V_a^x[F] + F(a)$, причём функции F_1, F_2 сохраняют свойства непрерывности (слева или справа) функции F.

Итак, для функции $F = F_1 - F_2$, где F_1, F_2 — неубывающие, непрерывные справа функции на $[a;b]$, можно определить интеграл Лебега–Стилтьеса от любой борелевской функции ξ по формуле

$$\int\limits_{(a;b]} \xi(x)\,dF(x) \;=\; \int\limits_{(a;b]} \xi(x)\,dF_1(x) - \int\limits_{(a;b]} \xi(x)\,dF_2(x),$$

если выражение справа имеет смысл, т.е. оба интеграла существуют и нет неопределённости вида $+\infty - \infty$.

579| Упр. Докажите формулу интегрирования по частям для функций F, G с ограниченной вариацией, одна из которых всюду непрерывна, а вторая всюду непрерывна справа.

Наиболее востребован случай, когда одна из функций абсолютно непрерывна. Как известно, условие абсолютной непрерывности на конечном отрезке влечёт условие ограниченности вариации.

580| **Теорема.** *Пусть функция F имеет ограниченное изменение на интервале $(a;b]$ и непрерывна справа на нём, а функция G абсолютно непрерывна на $(a;b]$. Тогда*

$$\int\limits_{(a;b]} G(x)\,dF(x) + \int\limits_{(a;b]} F(x)G'(x)\,dx \;=\; G(b)F(b) - G(a)F(a).$$

Дополнения и доказательства

A. Варианты теоремы Колмогорова о продолжении семейств мер

По теореме Колмогорова 496, стр. 451, семейство согласованных конечномерных вероятностных мер определяет вероятностную меру на борелевской σ-алгебре пространства числовых последовательностей, порождённой цилиндрическими подмножествами. Этот результат может быть распространён на произвольные измеримые пространства. Введём следующее

Определение. Пусть $(X_j, \mathscr{F}_j)$, $j = 1, 2$, — измеримые пространства. Функция $\mu : \mathscr{F}_2 \times X_1 \mapsto [0;1]$ называется *переходной мерой* (с X_1 на $\mathscr{F}_2$), если при $\forall\, x \in X_1$ функция множеств $\mu(B \,|\, x)$, $B \in \mathscr{F}_2$, есть мера на σ-алгебре $\mathscr{F}_2$, а при $\forall\, B \in \mathscr{F}_2$ функция $\mu(B \,|\, x)$ переменной x измерима по Борелю как функция из $(X_1, \mathscr{F}_1)$ в $[0;1]$.

Частный случай следующей леммы с переходными мерами $\mu(\cdot \,|\, x)$, не зависящими от x (лемма 572, стр. 501, доказательство см. стр. 518), используется при доказательстве теоремы Фубини.

581 **Лемма.** *Пусть $\mu : \mathscr{F}_2 \times X_1 \mapsto [0;1]$ — переходная мера, определённая на измеримых пространствах $(X_j, \mathscr{F}_j)$, $j = 1, 2$. Тогда для любой неотрицательной измеримой по Борелю функции $h : (X_1 \times X_2, \mathscr{F}_1 \otimes \mathscr{F}_2) \mapsto (\mathbb{R}^1_+, \mathscr{B})$ при $\forall\, x_1 \in X_1$ определён интеграл Лебега*

$$\mathcal{E}(h \,|\, x_1) \;=\; \int_{X_2} h(x_1, x_2)\, \mu(dx_2 \,|\, x_1),$$

и функция $\mathcal{E}(h \,|\, x_1)$, $x_1 \in X_1$, измерима по Борелю относительно $\mathscr{F}_1$.

⊫ Рассмотрим сначала индикаторную функцию $h = \dot{\mathrm{I}}_B(x_1, x_2)$ с $B = B_1 \times B_2$, $B_j \in \mathscr{F}_j$, $j = 1, 2$. По определению интеграла Лебега $\mathcal{E}(h \,|\, x_1) = \dot{\mathrm{I}}_{B_1}(x_1)\mu(B_2 \,|\, x_1)$, что представляет собой измеримую функцию x_1 по определению переходной меры. Обозначим через $\mathcal{W}$ семейство измеримых подмножеств $B \in \mathscr{F}_1 \otimes \mathscr{F}_2$, для которых справедливо утверждение об измеримости интеграла $\mathcal{E}(\dot{\mathrm{I}}_B \mid x_1)$. Из предыдущего полуалгебра $\mathscr{F}_1 \times \mathscr{F}_2 \subset \mathcal{W}$. Кроме того, если два непересекающихся подмножества $A, B \in \mathcal{W}$, то индикаторная функция $\dot{\mathrm{I}}_{A+B}(x_1, x_2) = \dot{\mathrm{I}}_A(x_1, x_2) + \dot{\mathrm{I}}_B(x_1, x_2)$, поэтому в силу линейности интеграла $\mathcal{E}(\dot{\mathrm{I}}_{A+B} \mid x_1) = \mathcal{E}(\dot{\mathrm{I}}_A \mid x_1) + \mathcal{E}(\dot{\mathrm{I}}_B \mid x_1)$. Стало быть, семейство $\mathcal{W}$

содержит алгебру, порождённую полуалгеброй $\mathscr{F}_1 \times \mathscr{F}_2$. Если показать, что $\mathcal{W}$ есть монотонный класс, то в силу 591, стр. 523, это будет означать, что $\mathcal{W}$ — σ-алгебра.

Пусть $B_n \searrow B$ для некоторой убывающей последовательности подмножеств из семейства $\mathcal{W}$. Тогда последовательность индикаторов $\dot{\mathrm{I}}_{B_n} \searrow \dot{\mathrm{I}}_B$ и для $\forall\, x_1 \in X_1$ по теореме о монотонной сходимости $\mathcal{E}(\dot{\mathrm{I}}_{B_n} \mid x_1) \searrow \mathcal{E}(\dot{\mathrm{I}}_B \mid x_1)$. Таким образом, интеграл $\mathcal{E}(\dot{\mathrm{I}}_B \mid x_1)$ измерим как предел измеримых функций. Поскольку σ-алгебра $\mathscr{F}_1 \otimes \mathscr{F}_2 \stackrel{\mathrm{df}}{=} \sigma\big(\mathscr{F}_1 \times \mathscr{F}_2\big)$, то $\mathcal{W} = \mathscr{F}_1 \otimes \mathscr{F}_2$, т.е. для $\forall\, B \in \mathscr{F}_1 \otimes \mathscr{F}_2$ интеграл $\mathcal{E}(\dot{\mathrm{I}}_B \mid x_1)$ измерим по Борелю. В силу линейности интеграла Лебега отсюда следует измеримость $\mathcal{E}(h \,|\, x_1)$ для любой простой функции h, представимой в виде конечной линейной комбинации индикаторов. Так как любая неотрицательная измеримая функция h есть монотонный предел простых функций: $h_n \nearrow h$, то в силу теоремы о монотонной сходимости $\mathcal{E}(h \,|\, x_1) = \lim_n \mathcal{E}(h_n \,|\, x_1)$ при $\forall\, x_1 \in X_1$, что и доказывает измеримость $\mathcal{E}(h \,|\, x_1)$. ⩤

Для дальнейшего удобно ввести следующие обозначения, связанные с семейством измеримых пространств $(\Omega_j, \mathscr{F}_j)$, $j = 1, 2, \ldots$:

$$\Omega_l^m = \Omega_l \times \cdots \times \Omega_m, \qquad \vec{\omega}^{(m)} = (\omega_1, \ldots, \omega_m) \in \Omega_1^m, \quad 1 \leqslant l \leqslant m \leqslant \infty,$$
$$\mathscr{F}_l^m = \bigotimes_{n=l}^{m} \mathscr{F}_n \stackrel{\mathrm{df}}{=} \sigma\big(\mathscr{F}_l \times \cdots \times \mathscr{F}_m\big), \quad 1 \leqslant l \leqslant m < \infty.$$

Здесь знак $\stackrel{\mathrm{df}}{=}$ напоминает, что σ-алгебра $\mathscr{F}_l^m$ на прямом произведении измеримых пространств есть минимальная σ-алгебра, содержащая всевозможные «прямоугольники» вида $B_l \times \cdots \times B_m$ с $B_n \in \mathscr{F}_n$, $n = \overline{l, m}$. В бесконечномерном пространстве Ω_1^∞ σ-алгебра $\mathscr{F}_1^\infty$ есть цилиндрическая σ-алгебра, порождённая всеми конечномерными цилиндрами $B^{(k)} \times \Omega_{k+1}^\infty$ с $B^{(k)} \in \mathscr{F}_1^k$ (или $B^{(k)} = B_1 \times \cdots \times B_k$ с $B_n \in \mathscr{F}_n$, $n = \overline{1, k}$), $k \geqslant 1$.

[*Ионеску Тулчи.*] *Пусть* $(\Omega_j, \mathscr{F}_j)$, $j = 1, 2, \ldots$, *— измеримые пространства,* $P_k(B \,|\, \vec{\omega}^{(k-1)})$, $\vec{\omega}^{(k-1)} \in \Omega_1^{k-1}$, $B \in \mathscr{F}_k$, *— семейство переходных вероятностей,* P_1 *— некоторая вероятностная мера на* $(\Omega_1, \mathscr{F}_1)$. *Тогда на пространстве* $(\Omega_1^\infty, \mathscr{F}_1^\infty)$ *существует единственная мера* $\mathbf{P}$, *для которой*

$$\mathbf{P}\left\{B^{(k)} \times \Omega_{k+1}^\infty\right\} = \tag{22}$$
$$= \int_{B_1} P_1(d\omega_1) \int_{B_2} P_2(d\omega_2 \,|\, \omega_1) \ldots \int_{B_k} P_k(d\omega_k \,|\, \omega_1, \ldots, \omega_{k-1})$$

при $\forall\, B^{(k)} = B_1 \times \ldots \times B_k$, $B_n \in \mathscr{F}_n$, $n = \overline{1, k}$, $k \geqslant 1$.

⊫ Для любого $m \geqslant 1$ и любой функции h на Ω^{m+1}, принимающей значения в отрезке $[0;1]$ и измеримой по Борелю относительно σ-алгебры $\mathscr{F}_1^{m+1}$, в силу леммы 581 можно определить измеримую функцию

$$\mathcal{E}_m(h\,|\,\vec{\omega}^{(m)}) \;=\; \int_{\Omega_m} h(\vec{\omega}^{(m+1)}) P_m(d\omega_{m+1}\,|\,\vec{\omega}^{(m)}).$$

Для каждого фиксированного $k > 1$, выбирая в качестве h индикаторную функцию множества $B \in \mathscr{F}_1^k$, можно последовательно вычислить

$$\begin{aligned} \widetilde{P}_{k-1,k}\big(B\,|\,\vec{\omega}^{(k-1)}\big) \;&:= \mathcal{E}_{k-1}\big(\,\dot{\mathrm{I}}_B\,|\,\vec{\omega}^{(k-1)}\big),\\ \widetilde{P}_{m,k}\big(B\,|\,\vec{\omega}^{(m)}\big) \;&:= \mathcal{E}_m\big(\widetilde{P}_{m+1,k}\big(B\,|\,\cdot\,\big)\,|\,\vec{\omega}^{(m)}\big), \quad m = k-2,\ldots,1,\\ \mathbf{P}_k\{B\} \;&=\; \int_{\Omega_1} \widetilde{P}_{1,k}\big(B\,|\,\omega_1\big)\,P_1(d\omega_1). \end{aligned}$$

Очевидно, правая часть (22) равна $\mathbf{P}_k\{B\}$ с «прямоугольным» множеством $B = B_1\times\cdots\times B_k,\ B_n \in \mathscr{F}_n,\ n \leqslant k$. Кроме того, ясно, что $\mathbf{P}_k\left\{\Omega_1^k\right\} = 1$.

Покажем, что функция множеств $\mathbf{P}_k$ на σ-алгебре $\mathscr{F}_1^k$ сигма-аддитивна, т.е. определяет меру. Пусть $B = \biguplus_1^\infty A_n,\ A_n \in \mathscr{F}_1^k$, тогда $\dot{\mathrm{I}}_B = \sum_1^\infty \dot{\mathrm{I}}_{A_n}$ и в силу свойства аддитивности интеграла и теоремы о монотонной сходимости

$$\begin{aligned} \widetilde{P}_{k-1,k}\big(B\,|\,\vec{\omega}^{(k-1)}\big) \;&=\; \mathcal{E}_{k-1}\big(\,\dot{\mathrm{I}}_B\,|\,\vec{\omega}^{(k-1)}\big) \;=\; \sum\nolimits_1^\infty \mathcal{E}_{k-1}\big(\,\dot{\mathrm{I}}_{A_n}\,|\,\vec{\omega}^{(k-1)}\big) \;=\\ &=\; \sum\nolimits_1^\infty \widetilde{P}_{k-1,k}\big(A_n\,|\,\vec{\omega}^{(k-1)}\big), \quad \forall\,\vec{\omega}^{(k-1)} \in \Omega_1^{k-1}. \end{aligned}$$

Аналогичным образом, по индукции доказывается возможность суммирования счётных рядов с положительными слагаемыми для остальных функций $\widetilde{P}_{m,k}, 1 \leqslant m < k-1$, а также для функции $\mathbf{P}_k$.

Определим теперь функцию множеств $\mathbf{P}$ на цилиндрах $B^{(k)}\times\Omega_{k+1}^\infty,\ B^{(k)} \in \mathscr{F}_1^k,\ k \geqslant 1$, как $\mathbf{P}\{B^{(k)}\times\Omega_{k+1}^\infty\} := \mathbf{P}_k\{B^{(k)}\}$. Поскольку, очевидно, для $\forall\,B^{(k)} \in \mathscr{F}_1^k$, функция $\mathcal{E}_k\big(\,\dot{\mathrm{I}}\,\big(\vec{\omega}^{(k+1)};\,B^{(k)}\times\Omega_{k+1}\big)\,|\,\vec{\omega}^{(k)}\big) = \dot{\mathrm{I}}\,\big(\vec{\omega}^{(k)};\,B^{(k)}\big)$, то значение $\mathbf{P}$ не зависит от способа представления основания цилиндрического множества. Из этого замечания и аддитивности функций $\mathbf{P}_k$ сразу следует аддитивность функции $\mathbf{P}$ и, следовательно, её монотонность.

Осталось установить непрерывность в нуле функции $\mathbf{P}$. Пусть $B_n^{(d_n)} \in \mathscr{F}_1^{d_n}$ — последовательность (конечномерных) множеств, для которых последовательность цилиндрических множеств $B_n^{(d_n)}\times\Omega_{d_n+1}^\infty \searrow \varnothing$, т.е. эта последовательность убывает и $\bigcap B_n^{(d_n)}\times\Omega_{d_n+1}^\infty = \varnothing$. Необходимо показать, что (монотонный) предел $\lim_{n}{}_{\downarrow}\, \mathbf{P}\big\{\,B_n^{(d_n)}\times\Omega_{d_n+1}^\infty\,\big\} = 0$. Ограничимся случаем, когда размерность $d_n = n$ (см. по этому поводу замечание 582 ниже), и обозначим для краткости $B_n^{(d_n)} = B_n$.

Предположим, от противного, что $\mathbf{P}\big\{\, B_n \times \Omega_{n+1}^{\infty} \,\big\} > \varepsilon$ при $\forall\, n \geqslant 1$ для некоторого $\varepsilon > 0$. По определению при $\forall\, n \geqslant 1$

$$\mathbf{P}\big\{\, B_n \times \Omega_{n+1}^{\infty} \,\big\} \;=\; \int_{\Omega_1} \widetilde{P}_{1,n}\big(B_n \,|\, \omega_1\big).\, P_1(d\omega_1).$$

Поскольку $B_n \times \Omega_{n+1} \supset B_{n+1}$, то по конструкции оператора $\mathcal{E}$ получаем, что

$$\mathcal{E}_{n+1}\big(\,\dot{\mathrm{I}}_{B_{n+1}}(\vec{\omega}^{(n)},\,\cdot\,) \,|\, \vec{\omega}^{(n)}\big) \;\leqslant\; \dot{\mathrm{I}}_{B_n}(\vec{\omega}^{(n)}).$$

Отсюда следует, что при $\forall\, \omega_1 \in \Omega_1$

$$\widetilde{P}_{1,n+1}\big(B_{n+1} \,|\, \omega_1\big) \;\leqslant\; \widetilde{P}_{1,n}\big(B_n \,|\, \omega_1\big),$$

т.е. последовательность $\widetilde{P}_{1,n}\big(B_n \,|\, \omega_1\big)$, $n \geqslant 1$, убывает. Следовательно, существует ω_{01} такое, что $\widetilde{P}_{1,n}\big(B_n \,|\, \omega_{01}\big) > \varepsilon$ при $\forall\, n \geqslant 1$, кроме того, $\omega_{01} \in B_1$ и для $\forall\, n \geqslant 2$ цилиндр $(\omega_1) \times \Omega_2^n \in B_n$, т.е. ω_{01} принадлежит проекции любого B_n на Ω_1 (в противном случае $\dot{\mathrm{I}}_{B_n}(\omega_{01},\,\cdot\,) \equiv 0$).

Предположим теперь, что для некоторого $k \geqslant 1$ мы смогли построить вектор $\vec{\omega}_0^{(k)} = (\omega_{01}, \ldots, \omega_{0k}) \in B_k$ такой, что $\widetilde{P}_{k,n}\big(B_n \,|\, \vec{\omega}_0^{(k)}\big) > \varepsilon$ при $\forall\, n \geqslant k$. Из определения функций $\widetilde{P}$ имеем для $n > k$

$$\varepsilon < \widetilde{P}_{k,n}\big(B_n \,|\, \vec{\omega}_0^{(k)}\big) \;=\; \int_{\Omega_{k+1}} \widetilde{P}_{k+1,n}\big(B_n \,|\, (\vec{\omega}_0^{(k)}, \omega_{k+1})\big)\, P_k(d\omega_{k+1} \,|\, \vec{\omega}_0^{(k)}).$$

Аналогично предыдущему отсюда заключаем, что найдётся $\omega_{0(k+1)} \in \Omega_{k+1}$ такое, что вектор $\vec{\omega}_0^{(k+1)} = (\omega_{01}, \ldots, \omega_{0(k+1)}) \in B_{k+1}$ будет иметь свойства, аналогичные $\vec{\omega}_0^{(k)}$. Следовательно, существует $(\omega_{01}, \omega_{02}, \ldots) \in B_n \times \Omega_{n+1}^{\infty}$ при $\forall\, n \geqslant 1$, что противоречит условию $B_n \times \Omega_{n+1}^{\infty} \searrow \varnothing$. ⫤

582| ⚠ При рассмотрении последовательности цилиндрических множеств $B_n^{(d_n)} \times \Omega_{d_n+1}^{\infty} \searrow \varnothing$ можно считать, во-первых, что размерность d_n в представлении основания $B_n^{(d_n)}$ выбрана наименьшей возможной, т.е. проекция этого основания на пространство Ω_{d_n} не совпадает с Ω_{d_n}. Во-вторых, отсюда следует, что последовательность d_n, $n \geqslant 1$, не убывает. Действительно, если $d_{n-1} > d_n$, то ввиду монотонности $B_n^{(d_n)} \times \Omega_{d_n+1}^{(d_{n-1})} \subset B_{n-1}^{(d_{n-1})}$ и, следовательно, проекция $B_{n-1}^{(d_{n-1})}$ на пространство $\Omega_{d_{n-1}}$ совпадает с $\Omega_{d_{n-1}}$.

В-третьих, случай $\sup_n d_n = k < \infty$ анализируется просто, поскольку тогда мера $\mathbf{P}\big\{\, B_n^{(d_n)} \times \Omega_{d_n+1}^{\infty} \,\big\} = \mathbf{P}_k\,\big\{B_n^{(d_n)} \times \Omega_{d_n+1}^{k}\big\}$ и свойство σ-аддитивности этой меры следует из аналогичного свойства для меры $\mathbf{P}_k$.

В-четвёртых, можно считать, что $d_n < d_{n+1}$, поскольку если $d_{n-1} < d_n =$

$= \dots = d_{n+m} < d_{n+m+1},\ m > 1$, то в (монотонной) последовательности цилиндрических множеств можно исключить все множества от n до $n+m-1$, оставив только $B_{n+m}^{(d_{n+m})} \times \Omega_{d_{n+m+1}}^{\infty}$.

Наконец, в-пятых, при $d_n < d_{n+1} - 1$ выбор очередного элемента ω_{0n} следует производить не из «одномерного» пространства $\Omega_{d_{n+1}}$, а из произведения пространств $\Omega_{d_n+1} \times \dots \times \Omega_{d_{n+1}}$.

Из теоремы Ионеску Тулчи следует результат, который позволяет допускать существование последовательностей независимых случайных элементов со значениями в произвольных пространствах.

$\big[$*Ломницкий–Улам.*$\big]$ *Пусть* $(\Omega_t, \mathscr{F}_t, P_t),\ t \in \mathbb{T}$, *— семейство вероятностных пространств. Тогда на измеримом пространстве функций* $(\prod_{t\in\mathbb{T}} \Omega_t, \mathscr{F}_{\mathbb{T}})$ *с цилиндрической* σ*-алгеброй* $\mathscr{F}_{\mathbb{T}}$ *существует единственная вероятностная мера* $\mathbf{P}$, *для которой при любых наборах параметров* $\tau = (t_1, \dots, t_k) \subset \mathbb{T}$

$$\mathbf{P}\Big\{ \prod_{j=1}^{k} B_j \times \prod_{t \notin \tau} \Omega_t \Big\} = \prod_{j=1}^{k} P_{t_j}(B_j)$$

при $\forall\, B_j \in \mathscr{F}_{t_j},\ j = \overline{1,k},\ k \geqslant 1.$

$\Rrightarrow$ Заметим, что при счётном $\mathbb{T}$ приведённое утверждение есть частный случай теоремы Ионеску Тулчи, т.к. вероятность P_t на $(\Omega_t, \mathscr{F}_t)$ удовлетворяет условиям переходной вероятности с Ω_s на $\mathscr{F}_t$ при любых $s \neq t \in \mathbb{T}$. Поэтому для любого счётного набора параметров $\mathbb{S} \subset \mathbb{T}$ в измеримом пространстве $(\prod_{s\in\mathbb{S}} \Omega_s, \mathscr{F}_{\mathbb{S}})$ с соответствующей цилиндрической σ-алгеброй $\mathscr{F}_{\mathbb{S}}$ существует единственная вероятностная мера $\mathbf{P}_{\mathbb{S}}$, для которой (в соответствии с (22)) при $\forall\, k \geqslant 1,\ \tau = (s_1, \dots, s_k) \subset \mathbb{S},\ B_j \in \mathscr{F}_{s_j}, j = \overline{1,k}$, вероятность цилиндра

$$\mathbf{P}_{\mathbb{S}} \Big\{ \prod_{j=1}^{k} B_j \times \prod_{s\in \mathbb{S}\setminus\tau} \Omega_s \Big\} = \prod_{j=1}^{k} P_{s_j}(B_j).$$

Важно отметить, что как сам прямоугольник, так и его вероятность не зависят от количества и мест расположения рёбер прямоугольника с $B_j = \Omega_{s_j}$.

Рассмотрим два счётных набора $\mathbb{S}_1 \subset \mathbb{S}_2 \subset \mathbb{T}$ и соответствующие им меры $\mathbf{P}_{\mathbb{S}_1}, \mathbf{P}_{\mathbb{S}_2}$ на цилиндрических σ-алгебрах $\mathscr{F}_{\mathbb{S}_1}, \mathscr{F}_{\mathbb{S}_2}$ соответственно. Ясно, что если $k \geqslant 1, \tau = (s_1, \dots, s_k) \subset \mathbb{S}_1,\ B_j \in \mathscr{F}_{s_j},\ j = \overline{1,k}$, то для цилиндрического множества $C = \prod_{j=1}^{k} B_{s_j} \times \prod_{s \in \mathbb{S}_1 \setminus \tau} \Omega_s \in \mathscr{F}_{\mathbb{S}_1}$ справедливо

$$(\sharp) \qquad \mathbf{P}_{\mathbb{S}_1}\{C\} \;=\; \mathbf{P}_{\mathbb{S}_2}\Big\{C \times \prod_{t \in \mathbb{S}_2 \setminus \mathbb{S}_1} \Omega_t\Big\}.$$

Поскольку семейство таких цилиндров образует полуалгебру, порождающую $\mathscr{F}_{\mathbb{S}_1}$, то это равенство будет справедливо для $\forall\, C \in \mathscr{F}_{\mathbb{S}_1}$.

По теореме 501, стр. 457, для любого подмножества $\boldsymbol{C} \in \mathscr{F}_{\mathbb{T}}$ из цилиндрической σ-алгебры на $\prod_{\mathbb{T}} \Omega_t$ найдётся счётный набор параметров $\mathbb{S} \subset \mathbb{T}$ и подмножество $C \in \mathscr{F}_{\mathbb{S}}$ такие, что $\boldsymbol{C} = C \times \prod_{t \notin \mathbb{S}} \Omega_t$. Положим

$$\mathbf{P}\{\boldsymbol{C}\} \;=\; \mathbf{P}_{\mathbb{S}}\{C\}$$

и покажем, что такое определение корректно. Пусть $\boldsymbol{C} = C_1 \times \prod_{t \notin \mathbb{S}_1} \Omega_t = C_2 \times \prod_{t \notin \mathbb{S}_2} \Omega_t$, где $C_j \in \mathscr{F}_{\mathbb{S}_j}$, $j = 1, 2$. Определим новый счётный набор параметров $\mathbb{U} = \mathbb{S}_1 \bigcup \mathbb{S}_2$, тогда, очевидно, $C_j \times \prod_{t \in (\mathbb{U} \setminus \mathbb{S}_j)} \Omega_t \in \mathscr{F}_{\mathbb{U}}$, $j = 1, 2$. Ввиду $(\sharp)$ имеем

$$\mathbf{P}_{\mathbb{S}_j}\{C\} \;=\; \mathbf{P}_{\mathbb{U}}\Big\{C_j \times \prod_{t \in \mathbb{U} \setminus \mathbb{S}_j} \Omega_t\Big\} \;=\; \mathbf{P}_{\mathbb{U}}\{\boldsymbol{C}\}, \qquad j = 1, 2.$$

Докажем сигма-аддитивность такой меры. Снова в силу 501, стр. 457, можно сразу рассмотреть счётное семейство непересекающихся подмножеств $\mathscr{F}_{\mathbb{T}}$ вида $\boldsymbol{B}_j = B_j \times \prod_{t \notin \mathbb{S}_j} \Omega_t$, $B_j \in \mathscr{F}_{\mathbb{S}_j}$, со счётными наборами параметров $\mathbb{S}_j$, $j \geqslant 1$. Для счётного набора $\mathbb{U} = \bigcup_j \mathbb{S}_j$ подмножества $B_j \times \prod_{t \in (\mathbb{U} \setminus \mathbb{S}_j)} \Omega_t \in \mathscr{F}_{\mathbb{U}}$, $j \geqslant 1$, и не пересекаются. Следовательно,

$$\mathbf{P}\Big\{\biguplus_j \boldsymbol{B}_j\Big\} \;=\; \mathbf{P}_{\mathbb{U}}\Big\{\biguplus_j \Big[B_j \times \prod_{t \in \mathbb{U} \setminus \mathbb{S}_j} \Omega_t\Big]\Big\} \;=$$
$$= \sum_j \mathbf{P}_{\mathbb{U}}\Big\{\Big[B_j \times \prod_{t \in \mathbb{U} \setminus \mathbb{S}_j} \Omega_t\Big]\Big\} = \sum_j \mathbf{P}\Big\{B_j \times \prod_{t \in \mathbb{U} \setminus \mathbb{S}_j} \Omega_t\Big\} \;=\; \mathbf{P}\{\boldsymbol{B}_j\}$$

в силу σ-аддитивности меры $\mathbf{P}_{\mathbb{U}}$. ⇚

B. Доказательства

✧ Доказательство теоремы Каратеодори (стр. 433)

$\Rrightarrow$ Для доказательства сигма-полуаддитивности внешней меры возьмём $G \subset \bigcup_1^\infty S_n$. Если $\mu^*(S_n) = \infty$ для некоторого $n \geqslant 1$, то свойство σ-полуаддитивности выполняется в силу монотонности внешней меры. Итак, пусть $\mu^*(S_n) < \infty$ для $\forall\, n \geqslant 1$. По определению точной нижней грани

$$\forall\, \varepsilon > 0 \;\; \forall\, n \geqslant 1 \;\; \exists \bigcup\nolimits_{j=1}^{\infty} B_{nj} \supset S_n : \qquad \mu^*(S_n) > \sum\nolimits_{j=1}^{\infty} \mu(B_{nj}) - \frac{\varepsilon}{2^n},$$

где все множества B_{nj} принадлежат кольцу $\mathcal{A}$.

Так как $G \subset \bigcup_{n=1}^\infty S_n \subset \bigcup_{n=1}^\infty \bigcup_{j=1}^\infty B_{nj}$, т.е. имеем счётное $\mathcal{A}$-покрытие G, то по определению внешней меры

$$\mu^*(G) \;\leqslant\; \sum\nolimits_n \sum\nolimits_j \mu(B_{nj}) \;<\; \sum\nolimits_n \mu^*(S_n) + \varepsilon.$$

Устремляя $\varepsilon \to 0$, получаем доказательство σ-полуаддитивности.

(I_1) Пусть $G \in \Psi$, т.е. $\mu^*(B) = \mu^*(B\,G) + \mu^*(B\,G^c)$ для $\forall\, B \subset \Omega$. Отсюда очевидным образом следует измеримость дополнительного множества G^c.

Пусть $S, G \in \Psi$, тогда для $\forall\, B \subset \Omega$

$$\begin{aligned} \mu^*(B) \;&=\; \mu^*(B\,S) + \mu^*(B\,S^c) \;=\; (\text{т.к. } \; G \in \Psi) \\ &=\; \mu^*(B\,S) + \mu^*((B\,S^c)\,G) + \mu^*((B\,S^c)\,G^c). \end{aligned}$$

Так как $S = S\,(S \cup G), \quad G\,S^c = (S \cup G)S^c$ и $S \in \Psi$, то

$$\begin{aligned} \mu^*(B\,S) \,+\, \mu^*((B\,S^c)\,G) \;&=\; \mu^*(B\,(S \cup G)\,S) \,+\, \mu^*(B\,(S \cup G)S^c) \;= \\ &=\; \mu^*(B\,(S \cup G)). \end{aligned}$$

Таким образом, $\mu^*(B) = \mu^*(B\,(S \cup G)) + \mu^*(B\,(S \cup G)^c)$, т.е. конечное объединение $S \cup G \in \Psi$.

Из двух предыдущих фактов следует, что класс Ψ замкнут и относительно операций пересечения $S\,G = (S^c \cup G^c)^c$ и разности $S \setminus G = S\,G^c$.

Пусть $S_n \in \Psi, \;\; S_n\,S_k = \varnothing, k, n \geqslant 1$. Тогда для $\forall\, B \subset \Omega$

$$\mu^*\Big(B \biguplus\nolimits_1^n S_i\Big) \;=\; \sum\nolimits_1^n \mu^*(B\,S_i).$$

Действительно, это равенство верно при $n = 1$. Если же оно верно для какого-то $n \geqslant 1$, то в силу измеримости S_{n+1} имеем

$$\mu^*\Big(B \biguplus_{i=1}^{n+1} S_i\Big) = \mu^*\Big(\Big(B \biguplus_{i=1}^{n+1} S_i\Big) S_{n+1}\Big) + \mu^*\Big(\Big(B \biguplus_{i=1}^{n+1} S_i\Big) S_{n+1}^c\Big) =$$

$$= \mu^*(B\, S_{n+1}) + \mu^*\Big(B \biguplus_{i=1}^{n} S_i\Big) = \sum_{i=1}^{n+1} \mu^*(B\, S_i).$$

Выбрав $B = \Omega$, получаем из предыдущего доказательство аддитивности внешней меры на классе Ψ.

Как отмечалось, $\biguplus_{i=1}^n S_i \in \Psi$ и μ^* монотонна (см. 472, стр. 432), поэтому

$$\mu^*(B) = \mu^*\Big(B \biguplus_{i=1}^{n} S_i\Big) + \mu^*\Big(B \Big(\biguplus_{i=1}^{n} S_i\Big)^c\Big) \geqslant$$

$$\geqslant \sum_{i=1}^{n} \mu^*(B\, S_i) + \mu^*\Big(B \Big(\biguplus_{i=1}^{\infty} S_i\Big)^c\Big)$$

при $\forall\, n \geqslant 1$. Полагая здесь $n \to \infty$ и воспользовавшись доказанной выше σ-полуаддитивностью μ^*, получаем

$$\mu^*(B) \geqslant \sum_{i=1}^{\infty} \mu^*(B\, S_i) + \mu^*\Big(B \Big(\biguplus_{i=1}^{\infty} S_i\Big)^c\Big) \geqslant \tag{23}$$

$$\geqslant \mu^*\Big(B \biguplus_{i=1}^{\infty} S_i\Big) + \mu^*\Big(B \Big(\biguplus_{i=1}^{\infty} S_i\Big)^c\Big) \geqslant \mu^*(B), \tag{24}$$

где последнее неравенство следует из полуаддитивности μ^* : $\mu^*(B\,A) + \mu^*(B\,A^c) \geqslant \mu^*(B)$. Следовательно, все знаки неравенства ($\geqslant$) здесь должны быть заменены на знаки равенства. Поэтому, во-первых, из соотношения (24) (теперь уже равенства) следует, что $\biguplus_1^\infty S_i \in \Psi$. Во-вторых, полагая в (23) $B = \biguplus_1^\infty S_i$ $(\in \Psi)$, получаем σ-аддитивность μ^*.

Для произвольных множеств $S_n,\ n \geqslant 1$, из класса Ψ

$$\bigcup\nolimits_1^\infty S_n = \biguplus\nolimits_{n=1}^\infty \Big(S_n \setminus \bigcup\nolimits_{i=1}^{n-1} S_i\Big) \in \Psi,$$

т.к. по доказанному класс Ψ замкнут относительно разностей, любых конечных объединений и счётных объединений непересекающихся множеств.

(I_2) Пусть $G \in \mathcal{A}$, тогда в силу полуаддитивности $\mu^*(B) = \mu^*(B\,G + B\,G^c) \leqslant \mu^*(B\,G) + \mu^*(B\,G^c)$. Для доказательства обратного воспользуемся определением внешней меры как инфимума, взятого по всем счётным покрытиям. Тогда для $\forall\varepsilon>0$ существует $\mathcal{A}$-покрытие $B\subset\bigcup_1^\infty S_n$ такое, что

$$\mu^*(B) \;>\; \sum\nolimits_1^\infty \mu(S_n)-\varepsilon.$$

Заметим, что $S_n=G\,S_n+G^c\,S_n$ (поэтому $\mu(S_n)=\mu(G\,S_n)+\mu(G^c\,S_n)$) и

$$B\,G \;\subset\; \bigcup\nolimits_1^\infty G\,S_n, \qquad B\,G^c \;\subset\; \bigcup\nolimits_1^\infty G^c\,S_n.$$

Следовательно,

$$\begin{aligned}\mu^*(B) \;&>\; \sum\nolimits_1^\infty \mu(S_n)-\varepsilon \;=\; \sum\nolimits_1^\infty \mu(G\,S_n)+\sum\nolimits_1^\infty \mu(G^c\,S_n)-\varepsilon \;\geqslant\\ &\geqslant\; \mu^*(B\,G)+\mu^*(B\,G^c)-\varepsilon.\end{aligned}$$

Полагая $\varepsilon\to 0$, получаем требуемое неравенство.

(I$_3$) Сигма-аддитивность μ^* была доказана в п. (I$_1$).

Пусть множество $G\in\mathcal{A}$ принадлежит кольцу. Тогда по определению внешней меры как минимума суммы мер $\mathcal{A}$-покрытий следует, что $\mu^*(G)\leqslant\leqslant\mu(G)$. С другой стороны, для любого $\mathcal{A}$-покрытия $\cup_i B_i\supset G$ по свойству σ-полуаддитивности $\mu(G)\leqslant\sum_i\mu(B_i)$, т.е. $\mu^*(G)\geqslant\mu(G)$.

(I$_4$) Предположим, что мера $\gamma:\sigma\big(\mathcal{A}\big)\mapsto\mathbb{R}^1_+$ совпадает с мерой μ на элементах кольца $\mathcal{A}$. По условию теоремы существует такое $\mathcal{A}$-покрытие $\Omega==\bigcup\limits_{n=1}^{\infty}S_n$, что $\mu(S_n)=\gamma(S_n)<\infty$ для $\forall\,n\geqslant 1$. При этом, как было показано выше, можно выбрать $S_n\cap S_j=\varnothing,\quad k\neq j$.

Зафиксируем $n\geqslant 1$ и определим класс

$$\mathscr{Q} \;=\; \big\langle\, G\in\sigma\big(\mathcal{A}\big):\; \mu(G\,S_n)=\gamma(G\,S_n)\,\big\rangle.$$

По условию $\mathcal{A}\subset\mathscr{Q}\subset\sigma\big(\mathcal{A}\big)$. Покажем, что $\mathscr{Q}$ — сигма-алгебра.

Очевидно, что $\varnothing,\Omega\in\mathscr{Q}$. Так как $S_n=G\,S_n+G^c\,S_n$, то

$$\gamma(G\,S_n)+\gamma(G^c\,S_n) \;=\; \mu(G\,S_n)+\mu(G^c\,S_n).$$

Поэтому если $G\in\mathscr{Q}$, то и $G^c\in\mathscr{Q}$, т.к. значения мер в последнем равенстве конечны.

Докажем монотонность $\mathscr{Q}$. Пусть $B=\lim\limits_{k}{}_{\uparrow(\downarrow)}G_k,\quad G_k\in\mathscr{Q}$. Тогда в силу непрерывности меры (снова учитываем, что $\gamma(G_k\,S_n)=\mu(G_k\,S_n)<\infty$)

$$\mu(B\,S_n) \;=\; \lim_k\mu(G_k\,S_n) \;=\; \lim_k\gamma(G_k\,S_n) \;=\; \gamma(B\,S_n).$$

По теореме 597, стр. 525, получаем, что $\mathscr{Q}=\mathfrak{M}\,(\mathcal{A})=\sigma\big(\mathcal{A}\big)$.

Таким образом, $\mu(B) = \sum_n^\infty \mu(B\, S_n) = \sum_n^\infty \gamma(B\, S_n) = \gamma(B)$ для $\forall\, B \in \sigma\big(\mathcal{A}\big)$ с $\mu(B) < \infty$ или $\gamma(B) < \infty$. ⫤

✧ Доказательство теоремы Лебега о разложении мер (стр. 491)

⊫ Предположим сначала, что меры γ, μ конечны.

a) Рассмотрим совокупность измеримых функций $H = \big\langle\, h \geqslant 0 : \int_A h\, d\mu \leqslant \gamma(A), \quad \forall\, A \in \mathscr{F} \big\rangle$, содержащую как минимум все индикаторные функции μ-нулевых множеств. В этот класс наряду с любыми двумя функциями h_1, h_2 входит и функция $h(\omega) = \max(h_1(\omega), h_2(\omega))$:

$$\int\limits_A h\, d\mu \;=\; \int\limits_{A\{h_1 \leqslant h_2\}} h_2\, d\mu + \int\limits_{A\{h_1 > h_2\}} h_1\, d\mu \;\leqslant$$
$$\leqslant\; \gamma(A\{h_1 \leqslant h_2\}) + \gamma(A\{h_1 > h_2\}) \;=\; \gamma(A).$$

По свойству супремума найдётся последовательность $h_n \in H$ такая, что

$$\lim_{n\to\infty} \int_\Omega h_n\, d\mu \;=\; J \;:=\; \sup_{h\in H} \int_\Omega h\, d\mu \;\leqslant\; \gamma(\Omega) \;<\; \infty.$$

Определим монотонную последовательность $g_n = \max\big\langle\, h_i, i = \overline{1, n}\,\big\rangle$. По установленному выше свойству $g_n \in H, n \geqslant 1$, и, кроме того, в силу монотонности интеграла Лебега $\int_\Omega h_n\, d\mu \leqslant \int_\Omega g_n\, d\mu \nearrow J, \quad n \to \infty$. По теореме 534 о монотонной сходимости отсюда следует, что для функции $\xi = \lim\limits_n{}_\uparrow\, g_n$

$$\int_A \xi\, d\mu \;=\; \lim_n \int_A g_n\, d\mu \;\leqslant\; \gamma(A), \qquad A \in \mathscr{F},$$

и $\int_\Omega \xi\, d\mu = J$. Таким образом, функция множеств $\psi(A) = \gamma(A) - \int_A \xi\, d\mu$ есть мера, поскольку она неотрицательная и, очевидно, сигма-аддитивная. Осталось показать её сингулярность, т.е. построить μ-нулевое множество, дополнение которого имеет ψ-меру нуль.

Зафиксируем $n \geqslant 1$ и для $\forall\, A \in \mathscr{F}$ с $\mu(A) > 0$ определим класс

$$D(A) \;=\; \Big\langle\, B \in \mathscr{F} : \quad B \subset A, \quad \psi(B) < \frac{1}{n}\mu(B)\,\Big\rangle.$$

Этот класс, очевидно, замкнут относительно счётных объединений непересекающихся подмножеств и $D(A_1) \subset D(A_2)$ при $A_1 \subset A_2$. Покажем теперь, что этот класс всегда не пуст. Действительно, если бы это было не так, то для функции $h_0 = \frac{1}{n}\,\dot{\mathrm{I}}_A$ при $\forall\, B \in \mathscr{F}$

$$\int_B h_0\,d\mu \;=\; \frac{1}{n}\mu(AB) \;\leqslant\; \psi(AB) \;\leqslant\; \psi(B) \;=\; \gamma(B) - \int_B \xi\,d\mu.$$

Таким образом, функция $\xi + h_0 \in H$, но по построению $J \geqslant \int_\Omega (\xi + h_0)\,d\mu = = \int_\Omega \xi\,d\mu + \mu(A) = J + \mu(A)$, что противоречит $\mu(A) > 0$ (при $J < \infty$).

Полагая $B_0 = \varnothing$, определим $\forall\, k \geqslant 1$ множества $\tilde{B}_k = \bigcap\limits_{i=0}^{k-1} B_i^c$ и

$$B_k \in D(\tilde{B}_k): \qquad \mu(B_k) \;\geqslant\; \varepsilon_k = \frac{1}{2}\sup\left\langle\, \mu(B): \quad B \in D(\tilde{B}_k)\,\right\rangle,$$

если $\mu(\tilde{B}_k) > 0$, иначе $B_k = \varnothing$, $\varepsilon_k = 0$. Из монотонности семейства $D(A)$ следует, что $B_k \in D(\Omega)$, если $B_k \neq \varnothing$, и, кроме того, т.к. множества $B_k, k \geqslant \geqslant 0$, не пересекаются, то и $Z = \biguplus_k B_k \in D(\Omega)$. Поэтому

$$\infty \;>\; \mu(Z) \;=\; \sum_{k=1}^{\infty} \mu(B_k) \;\geqslant\; \sum_{k=1}^{\infty} \varepsilon_k$$

ввиду конечности μ. Следовательно, $\varepsilon_k \to 0$, $k \to \infty$.

Если $\mu(Z^c) > 0$, то по определению D найдётся множество $Q \in D(Z^c)$ с $\mu(Q) > n\psi(Q) \geqslant 0$. Однако, т.к. $Z^c \subset \tilde{B}_k$, то

$$2\varepsilon_k \;=\; \sup\left\langle\, \mu(B): B \in D(\tilde{B}_k)\,\right\rangle \;\geqslant\; \sup\left\langle\, \mu(B): B \in D(Z^c)\,\right\rangle \;\geqslant\; \mu(Q) \;>\; 0.$$

Следовательно, для $Z_n = Z$ ($\in D(\Omega)$) мера $\mu(Z_n^c) = 0 = \mu(\Omega) - \mu(Z_n)$ и $\psi(Z_n) < \frac{1}{n}\mu(Z_n) = \frac{1}{n}\mu(\Omega)$. Поэтому

$$\psi\Big(\bigcap_{n=1}^{\infty} Z_n\Big) \;\leqslant\; \varliminf_{n\to\infty} \psi(Z_n) \;=\; 0,$$

$$\mu\Big(\Big(\bigcap_{n=1}^{\infty} Z_n\Big)^c\Big) \;=\; \mu\Big(\bigcup_{n=1}^{\infty} Z_n^c\Big) \;\leqslant\; \sum_{n=1}^{\infty} \mu(Z_n^c) \;=\; 0,$$

что доказывает a) (единственность ξ легко следует из b).

b) Пусть имеются два разложения Лебега $\gamma = \chi_1^\perp + \chi_1^a = \chi_2^\perp + \chi_2^a$ на μ-сингулярную ($\chi^\perp$) и μ-непрерывную (χ^a) части, а множества $N_1, N_2 \in \mathscr{F}$ таковы, что $\mu(N_1^c) = \mu(N_2^c) = \chi_1^\perp(N_1) = \chi_2^\perp(N_2) = 0$.

Очевидно, для $\forall\, B \in \mathscr{F}$ меры $\chi_i^\perp(B) = \chi_i^\perp(BN^c)$, где $N = N_1 \cap N_2$; при этом $\mu(BN^c) = 0$, и, следовательно, $\chi_i^a(BN^c) = 0$. Таким образом,

$$\chi_1^\perp(B) \;=\; \chi_1^\perp(BN^c) + \chi_1^a(BN^c) \;=\; \chi_2^\perp(BN^c) + \chi_2^a(BN^c) \;=\; \chi_2^\perp(B)$$

и $\chi_1^a(B)=\chi_2^a(B)$ при всех $B\in\mathscr{F}$.

Осталось избавиться от предположения конечности мер. В силу сигма-конечности μ,γ существует такое $\mathscr{F}$-разбиение $\Omega=\biguplus_{k=1}^{\infty}\Omega_k$, что $\mu(\Omega_k),\gamma(\Omega_k)<\infty$. Рассмотрим измеримое пространство $(\Omega_k,\mathscr{F}_k)$ с σ-алгеброй $\mathscr{F}_k=\Omega_k\Cap\mathscr{F}=\left\langle B\Omega_k : B\in\mathscr{F}\right\rangle$. Меры μ_k,γ_k на $\mathscr{F}_k$ совпадают с мерами μ,γ на множествах вида $B\cap\Omega_k, B\in\mathscr{F}$. По доказанному, найдутся мера $\psi_k\perp\mu_k$ и $\mathscr{F}_k$-измеримая функция $\xi_k\geqslant 0$ такие, что для $\forall\, B\in\mathscr{F}$

$$\gamma_k(B\Omega_k)\;=\;\psi_k(B\Omega_k)+\int\limits_{B\Omega_k}\xi_k\,d\mu_k$$

и для некоторого множества $N_k\in\mathscr{F}_k$ мера $\mu_k(\Omega_k\setminus N_k)=0$ и $\psi_k(N_k)=0$.

Определим на Ω функцию $\tilde{\xi}_k(\omega)=\xi_k(\omega)$, если $\omega\in\Omega_k$, и $\tilde{\xi}_k(\omega)=0$ в противном случае. Очевидно, эта функция $\mathscr{F}$-измерима и по построению интеграла Лебега $\int_B\tilde{\xi}_k\,d\mu=\int_{B\Omega_k}\tilde{\xi}_k\,d\mu=\int_{B\Omega_k}\xi_k\,d\mu_k$. Поэтому для функции $\tilde{\xi}=\sum_k\tilde{\xi}_k\,\dot{\mathrm{I}}_{\Omega_k}$ в силу теоремы о монотонном пределе под знаком интеграла

$$\int\limits_B\tilde{\xi}\,d\mu\;=\;\sum_{k=1}^{\infty}\int\limits_{B\Omega_k}\xi_k\,d\mu_k.$$

Зададим меру $\psi(B)=\sum_k\psi_k(B\Omega_k), B\in\mathscr{F}$. Поскольку множества $N_k,\Omega_k\setminus N_k\subset\Omega_k$, то при разных k они не пересекаются, и потому

$$\psi\Big(\biguplus_{k=1}^{\infty}N_k\Big)\;=\;\sum_{k=1}^{\infty}\psi_k(N_k)\;=\;0,\qquad \mu\Big(\Big(\biguplus_{k=1}^{\infty}N_k\Big)^c\Big)\;=\;\sum_{k=1}^{\infty}\mu_k(\Omega_k\setminus N_k)\;=\;0.$$

То есть мера $\psi\perp\mu$ и пара $\tilde{\xi},\psi$ удовлетворяет 559, а), стр. 491. Единственность разложения мер по Лебегу доказывается аналогично. $\Lleftarrow$

✧ Доказательство леммы о прямом произведении мер (стр. 501)

$\Rrightarrow$ Предположим сначала, что обе меры конечны. Везде индекс ω_1 у множеств $B\in\Omega_1\times\Omega_2$ будет означать взятие сечения вдоль $\omega_1\in\Omega_1$.

✓ Пусть $\mathcal{W}$ — совокупность множеств $B\in\mathscr{F}_1\otimes\mathscr{F}_2$, для которых функция $\mu_2(B_{\omega_1}),\omega_1\in\Omega_1$, $\mathscr{F}_1$-измерима и интеграл $\int_{\Omega_1}\mu_2(B_{\omega_1})\,d\mu_1<\infty$. Покажем, что $\mathcal{W}$ — σ-алгебра, совпадающая с $\mathscr{F}_1\otimes\mathscr{F}_2$.

Класс $\mathscr{F}_1\times\mathscr{F}_2\subset\mathcal{W}$, т.к. при $\forall\, B=B_1\times B_2$ мера сечения $\mu_2(B_{\omega_1})=$ $=\mu_2(B_2)$, если $\omega_1\in B_1$, или $\mu_2(B_{\omega_1})=0$, если $\omega_1\notin B_1$. Поэтому

$$\int_{\Omega_1} \mu_2(B_{\omega_1})\mu_1(d\omega_1) \;=\; \mu_2(B_2)\int_{B_1} \mu_1(d\omega_1) \;=\; \mu_2(B_2)\mu_1(B_1) \;<\; \infty.$$

Если $A, B \in \mathcal{W}$ и $AB = \varnothing$, то $A_{\omega_1} \cap B_{\omega_1} = \varnothing$ и, кроме того, $(A+B)_{\omega_1} = A_{\omega_1} + B_{\omega_1}$, поэтому $\mu_2((A+B)_{\omega_1}) = \mu_2(A_{\omega_1}) + \mu_2(B_{\omega_1})$ и $A + B \in \mathcal{W}$.

Таким образом, $\mathcal{W}$ содержит как полуалгебру $\mathscr{F}_1 \times \mathscr{F}_2$, так и все конечные объединения непересекающихся элементов этой полуалгебры, т.е. содержит алгебру, порождённую $\mathscr{F}_1 \times \mathscr{F}_2$. Для установления равенства $\mathcal{W} = \mathscr{F}_1 \otimes \mathscr{F}_2$ в силу теоремы 597, стр. 525, достаточно показать монотонность класса $\mathcal{W}$.

Пусть $B_n \subset B_{n+1} \in \mathcal{W}, n \geqslant 1$ и $B = \lim\limits_{n}{}_{\uparrow} B_n = \bigcup_n B_n$. Тогда из очевидных соотношений $(B_n)_{\omega_1} \subset (B_{n+1})_{\omega_1}, B_{\omega_1} = \bigcup_n (B_n)_{\omega_1}$ следует, что $\mu_2((B_n)_{\omega_1}) \leqslant \mu_2((B_{n+1})_{\omega_1}) \leqslant \mu_2(\Omega_2) < \infty$, и в силу непрерывности меры

$$\lim_{n}{}_{\uparrow} \mu_2((B_n)_{\omega_1}) \;=\; \mu_2(B_{\omega_1}) \;<\; \infty.$$

Поскольку предел измеримых функций измерим, то $\lim\limits_{n}{}_{\uparrow} B_n \in \mathcal{W}$.

✓ Покажем теперь, что неотрицательная функция множеств $\mu(B) = \int_{\Omega_1} \mu_2(B_{\omega_1})\, d\mu_1$, $B \in \mathscr{F}_1 \otimes \mathscr{F}_2$, есть мера. Пусть $B = \biguplus_n B_n$, $B_n \in \mathscr{F}_1 \otimes \mathscr{F}_2$, тогда, очевидно, $B_{\omega_1} = \biguplus_n (B_n)_{\omega_1}$, т.к. $(B_n)_{\omega_1}(B_m)_{\omega_1} = \varnothing, k \neq m$. Применяя теорему о монотонной сходимости для интеграла Лебега, отсюда получаем

$$\mu(B) \;=\; \int_{\Omega_1} \left[\sum\nolimits_1^\infty \mu_2((B_n)_{\omega_1})\right] d\mu_1 \;=\; \sum\nolimits_1^\infty \int_{\Omega_1} \mu_2((B_n)_{\omega_1})\, d\mu_1 \;=\; \sum\nolimits_1^\infty \mu(B_n).$$

Эта мера, очевидно, конечна, поэтому она в силу теоремы Каратеодори задаёт единственную меру на $\mathscr{F}_1 \otimes \mathscr{F}_2$, для которой $\mu(B_1 \times B_2) = \mu_1(B_1)\mu_2(B_2)$.

✓ Пусть теперь меры μ_1, μ_2 сигма-конечны. Тогда найдутся разбиения $\Omega_1 = \biguplus_n Q_{1n}, \Omega_2 = \biguplus_n Q_{2n}$ на непересекающиеся $\mathscr{F}_1$-, $\mathscr{F}_2$-измеримые подмножества, для которых $\mu_1(Q_{1n}) < \infty, \mu_2(Q_{2n}) < \infty, n \geqslant 1$. Тогда $\Omega_1 \times \Omega_2$ также разбивается на μ-конечные подмножества:

$$\Omega_1 \times \Omega_2 \;=\; \left(\biguplus\nolimits_1^\infty Q_{1n} \times \Omega_2\right) \bigcap \left(\biguplus\nolimits_1^\infty \Omega_1 \times Q_{2m}\right) \;=\; \biguplus\nolimits_{n,m=1}^\infty Q_{1n} \times Q_{2m}.$$

Для $B \in \mathscr{F}_1 \otimes \mathscr{F}_2$ положим $B^{(nm)} = B \cap (Q_{1n} \times Q_{2m})$, тогда, очевидно, $B_{\omega_1} = \biguplus_{nm} B^{(nm)}_{\omega_1}$ и в силу σ-аддитивности μ_2

$$\mu_2(B_{\omega_1}) \;=\; \sum\nolimits_{nm} \mu_2(B^{(nm)}_{\omega_1}).$$

Так как $\mu_2(B^{(nm)}_{\omega_1})$ измеримы (по ω_1) при каждых n, m, то и функция $\mu_2(B_{\omega_1})$

также измерима. Поэтому можно определить функцию множеств

$$\mu(B) \ := \ \sum\nolimits_{nm} \ \int_{\Omega_1} \mu_2(B^{(nm)}_{\omega_1})\, d\mu_1,$$

для которой σ-аддитивность следует из σ-аддитивности μ_2, теоремы о монотонной сходимости для интеграла Лебега, а также из возможности перестановки слагаемых ряда с неотрицательными членами. Равенство $\mu(B_1 \times B_2) = \mu_1(B_1)\mu_2(B_2)$ проверяется простыми преобразованиями. ⫤

C. Основные понятия теории множеств. Кольца. Алгебры

Будем придерживаться следующих обозначений и сокращений:

- $A \cap B = A \cdot B = A\, B$ — пересечение множеств;
- $\{\omega \in \Omega : Q(\omega)\} = \{Q\}$ — совокупность (класс, набор) элементов множества Ω, удовлетворяющих свойству Q;
- $\langle\, A_1, A_2, \dots \,\rangle = \langle\, A_i \,\rangle_i = \langle\, A_i \,\rangle_1^n$ — совокупность выбранных элементов;
- $\mathscr{P}(\Omega) = \langle\, A \subseteq \Omega \,\rangle$ — класс всех подмножеств Ω, включая $\varnothing, \Omega$;
- $\uplus$ (или $+$) — знак объединения непересекающихся множеств:

$$A_i\, A_j = \varnothing,\ i \neq j, \quad \Leftrightarrow \quad A_1 \cup A_2 = A_1 + A_2, \ \ \bigcup\nolimits_{j=1}^{N} A_j = \biguplus\nolimits_{j=1}^{N} A_j;$$

- $B \subset \bigcup_j S_j, \ \ S_j \in \mathcal{H}$, — *покрытие* множества B элементами класса $\mathcal{H}$ или $\mathcal{H}$*-покрытие*;
- $B = \biguplus_j S_j, \ \ S_j \in \mathcal{H}$, — *разбиение* множества B на непересекающиеся элементы класса $\mathcal{H}$ или $\mathcal{H}$*-разбиение*;
- индикаторная функция множества $A \subset \Omega$:

$$\dot{\mathrm{I}}_A = \dot{\mathrm{I}}\,(\omega; A) = \dot{\mathrm{I}}\,(A) = \begin{cases} 1, & \text{если} \ \ \omega \in A, \\ 0, & \text{если} \ \ \omega \notin A; \end{cases}$$

- *нижний предел* счётного семейства $\langle\, A_n \,\rangle_{n=1}^{\infty}$ подмножеств Ω есть набор всех точек $\omega \in \Omega$, которые, начиная с некоторого n, входят во все A_n:

$$\underline{\lim_n}\, A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k;$$

- *верхний предел* есть набор тех точек, которые входят в бесконечное число (б.ч.) множеств последовательности: $\{A_n \text{ б.ч.}\} = \overline{\lim_n}\, A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k;$

✦ *предел убывающей последовательности* $A_n \supset A_{n+1},\ n \geqslant 1$:

$$\lim_n {\downarrow} A_n = \underline{\lim}_n\, A_n = \overline{\lim_n}\, A_n = \bigcap\nolimits_{n=1}^{\infty} A_n;$$

✦ *предел возрастающей последовательности* $A_n \subset A_{n+1}$:

$$\lim_n {\uparrow} A_n = \underline{\lim}_n\, A_n = \overline{\lim_n}\, A_n = \bigcup\nolimits_{n=1}^{\infty} A_n.$$

► Полукольцо.

Определение. Класс $\mathcal{H}$ подмножеств Ω образует *полукольцо*, если:

C1) $\mathcal{H}$ замкнут относительно конечных пересечений:

$$A, B \in \mathcal{H} \quad \Rightarrow \quad AB \in \mathcal{H};$$

C2) $A, B \in \mathcal{H} \quad \Rightarrow \quad \exists$ конечное $\mathcal{H}$-разбиение $A \setminus B$:

$$A \setminus B \;=\; \biguplus\nolimits_{j=1}^{n} S_j, \qquad \big\langle\, S_j \,\big\rangle_{j=1}^{n} \;\subset\; \mathcal{H}, \quad n < \infty.$$

Полукольцо $\mathcal{H}$, содержащее основное пространство Ω, т.е. полукольцо с «единицей», называется *полуалгеброй*.

583| Упр. а) Докажите, что системы множеств, описанные в примере 465, стр. 428, удовлетворяют свойствам C1), C2).

б) Докажите, что система интервалов $[a; b)$, $a \leqslant b$, образует полукольцо. Если, кроме того, в систему включить все интервалы вида $[a; +\infty), (-\infty; a)$, а также прямую $\mathbb{R}^1$, то такая система будет полуалгеброй.

в) Приведите примеры, когда разность $A \setminus B \notin \mathcal{H}$ и объединение $A \bigcup B \notin$ $\notin \mathcal{H}$ для подмножеств A, B из полукольца $\mathcal{H}$.

584| **Лемма.** *Пусть* $Q_1, \ldots, Q_n \in \mathcal{H}$ *— элементы полукольца. Тогда существует* $\mathcal{H}$ *-разбиение* $(B_1, \ldots, B_L \in \mathcal{H})$

$$Q_n \setminus \bigcup\nolimits_{i=1}^{n-1} Q_i \;= \biguplus\nolimits_{j=1}^{L} B_j.$$

⊨ Так как $Q_3 \setminus (Q_1 \cup Q_2) \;=\; (Q_3 \setminus Q_2) \setminus Q_1$, то по свойству полукольца, применённому к разностям $Q_3 \setminus Q_2$ и $S_k^{(3)} \setminus Q_1$, справедливы равенства

$$Q_3 \setminus (Q_1 \cup Q_2) \;=\; \biguplus\nolimits_{k=1}^{N_3} S_k^{(3)} \setminus Q_1 \;=\; \biguplus\nolimits_{k=1}^{N_3} \biguplus\nolimits_{i=1}^{M_k} S_{ki}^{(3)} \;=\; \biguplus\nolimits_{j=1}^{L_3} B_j$$

с некоторыми наборами $\{S_{ki}^{(3)}, i = \overline{1, M_k}\}$, $\{S_k^{(3)}, k = \overline{1, N_3}\}$ непересекающихся элементов полукольца, где последнее равенство получено путем простого переобозначения: $\{S_{ki}^{(3)}, i=\overline{1,M_k}, k=\overline{1,N_3}\} = \{B_j, j=\overline{1,L_3}\}$. Общий случай доказывается аналогичным образом по индукции. К примеру:

$$Q_4 \setminus \bigcup_1^3 Q_i \;=\; (Q_4 \setminus Q_3) \setminus \bigcup_1^2 Q_i \;=\; \biguplus_{k=1}^{N_4} \Big(S_k^{(4)} \setminus \bigcup_1^2 Q_i \Big). \qquad \Lleftarrow$$

► Кольцо.

Определение. Совокупность $\mathcal{D}$ подмножеств пространства Ω называется *кольцом*, если для $\forall\, n<\infty$ выполняются следующие условия:

R1) $A, B \in \mathcal{D} \quad \Rightarrow \quad B \setminus A \in \mathcal{D}$ (замкнутость относительно разности);

R2) $\langle A_i \rangle_1^n \subset \mathcal{D} \;\Rightarrow\; \bigcup\limits_1^n A_i \in \mathcal{D}$ и $\bigcap\limits_1^n A_i \in \mathcal{D}$ (замкнутость относительно конечных объединений и пересечений).

Кольцо, содержащее Ω, т.е. кольцо с «единицей», называется *алгеброй*.

585| Упр. Докажите, что а) условие R2) можно заменить условием R2'): $A, B \in \mathcal{D} \Rightarrow A \bigcup B \in \mathcal{D}$, б) класс $\mathcal{D}$ является кольцом т. т. т. когда $\mathcal{D}$ полукольцо, удовлетворяющее R2').

Из определения видно, что если кольцо содержит все элементы некоторого полукольца, то оно содержит и все конечные объединения этих элементов.

586| **Лемма.** (?!) *Класс множеств, состоящий из всевозможных конечных объединений непересекающихся элементов полукольца, образует кольцо (так называемое кольцо, порождённое полукольцом). Это кольцо совпадает с классом всех конечных объединений (возможно, пересекающихся) элементов полукольца. Кроме того, это кольцо есть минимальное кольцо, содержащее все элементы полукольца.*

✧ Сигма-алгебра и монотонный класс

Определение. Совокупность $\mathscr{F}$ подмножеств пространства Ω образует σ*-алгебру* (сигма-алгебру), если:

S1) $\Omega \in \mathscr{F}, \quad \varnothing \in \mathscr{F}$ – в $\mathscr{F}$ входит как всё Ω, так и пустое подмножество;

S2) $A \in \mathscr{F} \quad \Rightarrow \quad A^c = \Omega \setminus A \in \mathscr{F}$ – $\mathscr{F}$ замкнута относительно операции дополнения;

S3) $\langle A_j \rangle_1^\infty \subset \mathscr{F} \quad \Rightarrow \quad \bigcup_1^\infty A_j \in \mathscr{F}$ – $\mathscr{F}$ замкнута относительно счётного объединения своих элементов.

587| Упр. Покажите, что для любой σ-алгебры $\mathscr{F}$:

а) S3) эквивалентно замкнутости относительно счётных пересечений;

б) $\mathscr{F}$ замкнута относительно конечных объединений и пересечений;

в) если $A, B \in \mathscr{F}$, то $A \setminus B \in \mathscr{F}$, $A \Delta B \in \mathscr{F}$;

г) любая конечная алгебра, т.е. алгебра, состоящая из конечного числа подмножеств, является также и σ-алгеброй.

588| Примеры. Следующие системы подмножеств образуют σ-алгебры:

(✦) $\mathscr{F} = \langle \varnothing, \Omega \rangle$; (✦) $\mathscr{F} = \langle \varnothing, \Omega, A, A^c \rangle$, где $A \subset \Omega$;

(✦) $\mathscr{F} = \langle A \subseteq [0;1], \ A$ — измеримо по Лебегу $\rangle$;

(✦) $\mathscr{F} = \mathscr{P}(\Omega) = \langle A \subseteq \Omega \rangle$ — совокупность всех подмножеств Ω.

Последнюю σ-алгебру часто обозначают как 2^{Ω}, поскольку для пространства Ω, состоящего из конечного числа k элементов, количество элементов этой σ-алгебры $\#(\mathscr{P}(\Omega))$ равно 2^k. ◎

589| Упр. Докажите, что $\#(\mathscr{P}(\Omega)) = 2^k$ при конечном $k = \#(\Omega) < \infty$.

⚠ Интересно, что σ-алгебра не может содержать бесконечное, но счётное число подмножеств.

Наряду с σ-алгеброй удобно ввести ещё классы множеств, замкнутых относительно монотонных пределов.

Определение. Семейство множеств $\mathcal{M}$ пространства Ω называется *монотонным классом,* если предел любой монотонной последовательности элементов $\mathcal{M}$ принадлежит $\mathcal{M}$. Монотонный класс $\mathcal{M}$ называется *λ-системой,* если он замкнут относительно монотонных разностей: $B_1, B_2 \in \mathcal{M}, B_1 \subset \subset B_2 \Rightarrow B_2 \setminus B_1 \in \mathcal{M}$.

590| Примеры. 1) Очевидно, любая σ-алгебра является монотонным классом и λ-системой.

2) Класс интервалов на прямой с конечными или бесконечными границами есть монотонная полуалгебра (но не λ-система). ◎

591| **Лемма.** I) *Алгебра $\mathcal{A}$ является σ-алгеброй т. т. т. когда она образует монотонный класс.*

II) *Класс множеств $\mathcal{M}$ является σ-алгеброй т. т. т. когда $\mathcal{M}$ есть λ-система, замкнутая относительно конечных пересечений, и $\Omega \in \mathcal{M}$.*

$\Rrightarrow$ I) Так как $\mathcal{A}$ — алгебра, то для $\forall \left\langle S_i \right\rangle_1^\infty \subset \mathcal{A}$ множества $B_n = \bigcup_1^n S_i \in \mathcal{A}$ и $B_n \subset B_{n+1}$. Поэтому $\bigcup_1^\infty S_i = \lim\limits_n \uparrow B_n \in \mathcal{A}$ в силу монотонности $\mathcal{A}$. Аналогично для счётных пересечений.

II) Следует из утверждения I) и хорошо известного правила де Моргàна: $A \cup B = \Omega \setminus \big((\Omega \setminus A) \cap (\Omega \setminus B)\big)$. $\Lleftarrow$

592| Пример. Рассмотрим какое-либо счётное разбиение $\Sigma = \left\langle S_i \right\rangle_1^\infty$ основного пространства: $\Omega = \biguplus_1^\infty S_i$. Свяжем с этим разбиением класс множеств, полученных путём конечного или счётного объединения элементов Σ:

$$\mathfrak{S} \;=\; \left\langle\, \biguplus_{i=1}^\infty S_i^{(j_i)} : j_i = \pm 1, i \geqslant 1 \,\right\rangle,$$

где $S^{(-1)} = \varnothing, S^{(+1)} = S$. Ясно, что $\varnothing, \Omega \in \mathfrak{S}$. Кроме того, этот класс, очевидно, замкнут относительно операций вычитания и конечных объединений и пересечений. Проверка монотонности $\mathfrak{S}$ также не вызывает затруднений. Таким образом, $\mathfrak{S}$ есть σ-алгебра, которая содержит все элементы разбиения Σ, и любая другая σ-алгебра, удовлетворяющая этому свойству, будет содержать в себе класс $\mathfrak{S}$, т.е. $\mathfrak{S}$ есть минимальная σ-алгебра, содержащая все элементы разбиения Σ. ◎

593| **Теорема.** *Для любой совокупности $\mathcal{Q}$ подмножеств Ω здесь существуют единственные минимальные σ-алгебра $\sigma\big(\mathcal{Q}\big)$, монотонный класс $\mathfrak{M}(\mathcal{Q})$ и λ-система $\Lambda(\mathcal{Q})$, содержащие все элементы $\mathcal{Q}$.*

$\Rrightarrow$ При доказательстве всех утверждений используется тот факт, что пересечение любого набора классов данного типа (набора σ-алгебр, монотонных классов, λ-систем) будет классом того же типа (см. упражнение ниже). К примеру, система $\mathscr{P}(\Omega)$ всех подмножеств Ω есть σ-алгебра, поэтому совокупность $\left\langle \mathscr{A}_t \right\rangle_{t \in \Upsilon}$ всех σ-алгебр, каждая из которых содержит все элементы $\mathcal{Q}$, не пуста. Тогда класс, состоящий из множеств, входящих во все σ-алгебры $\mathscr{A}_t,\ t \in \Upsilon$,

$$\sigma\big(\mathcal{Q}\big) \;=\; \bigcap\nolimits_{t \in \Upsilon} \mathscr{A}_t,$$

удовлетворяет утверждению теоремы. Аналогично для λ-систем и монотонных классов. $\Lleftarrow$

594| Упр. Докажите, что пересечение любого набора σ-алгебр (монотонных классов, λ-систем) есть σ-алгебра (монотонный класс, λ-система).

При доказательстве совпадения σ-алгебр, определённых двумя различными способами, весьма востребованы утверждения следующего упражнения.

595| Упр. Докажите, что:

а) $\mathcal{Q}_1 \subset \mathcal{Q}_2 \quad\Rightarrow\quad \sigma(\mathcal{Q}_1) \subset \sigma(\mathcal{Q}_2)$;

б) $\mathcal{Q}_1 \subset \mathcal{Q}_2$ и $\mathcal{Q}_2 \subset \sigma(\mathcal{Q}_1) \quad\Rightarrow\quad \sigma(\mathcal{Q}_1) = \sigma(\mathcal{Q}_2)$,

т.е. две σ-алгебры совпадают, если σ-алгебра, порождённая более узким классом, содержит все элементы более широкого класса;

в) $\mathcal{Q}_1 \subset \sigma(\mathcal{Q}_2)$ и $\mathcal{Q}_2 \subset \sigma(\mathcal{Q}_1) \quad\Rightarrow\quad \sigma(\mathcal{Q}_1) = \sigma(\mathcal{Q}_2)$,

т.е. две σ-алгебры совпадают, если σ-алгебра, порождённая одним классом, содержит все элементы другого класса, и наоборот;

г) операции $\mathfrak{M}(\mathcal{Q})$, $\Lambda(\mathcal{Q})$ также удовлетворяют а) – в).

596| Упр. Опишите минимальный монотонный класс $\mathfrak{M}(\mathcal{K})$, порождённый полукольцом интервалов на прямой вида $(a;b]$.

597| **Теорема.** *Если $\mathcal{H}$ — кольцо и минимальный монотонный класс $\mathfrak{M}(\mathcal{H}) \ni \Omega$, то $\mathfrak{M}(\mathcal{H}) = \sigma(\mathcal{H})$.*

⤇ По предыдущей лемме необходимо установить, что $\mathfrak{M}(\mathcal{H})$ — кольцо, т.е. $A_1 \cup A_2,\ A_1 \setminus A_2 \in \mathfrak{M}(\mathcal{H})$ для $\forall\, A_1, A_2 \in \mathfrak{M}(\mathcal{H})$.

Покажем, что для любого $S \subset \Omega$ класс множеств

$$\mathfrak{T}(S) \;=\; \big\langle\, A \subset \Omega :\quad S \cup A,\ S \setminus A,\ A \setminus S \in \mathfrak{M}(\mathcal{H}) \,\big\rangle$$

образует монотонный класс. Пусть $\big\langle\, C_n \,\big\rangle_1^\infty \subset \mathfrak{T}(S)$ — возрастающая последовательность и $C = \lim\limits_{n}{}_{\uparrow} C_n$. Тогда по определению $\mathfrak{T}(S)$

$$S \cup C_n,\quad C_n \setminus S,\quad S \setminus C_n \in \mathfrak{M}(\mathcal{H}),$$

причём последовательности $\big\langle\, S \cup C_n \,\big\rangle_1^\infty, \big\langle\, C_n \setminus S \,\big\rangle_1^\infty$ — возрастающие, а последовательность $\big\langle\, S \setminus C_n \,\big\rangle_1^\infty$ — убывающая. Так как $\mathfrak{M}(\mathcal{H})$ — монотонный класс, то, например, $(S \cup C) = \lim\limits_{n}{}_{\uparrow} (S \cup C_n) \in \mathfrak{M}(\mathcal{H})$. Следовательно, $\lim\limits_{n}{}_{\uparrow} C_n \in \mathfrak{T}(S)$ и, аналогично, $\lim\limits_{n}{}_{\downarrow} C_n \in \mathfrak{T}(S)$.

Если $S \in \mathcal{H}$ и $A \in \mathcal{H}$, то из определения кольца следует, что $A \in \mathfrak{T}(S)$, т.е. $\mathcal{H} \subset \mathfrak{T}(S)$. Поэтому и минимальный монотонный класс $\mathfrak{M}(\mathcal{H}) \subset \mathfrak{T}(S)$. Другими словами, каковы бы ни были $S \in \mathcal{H}$ и $A_1 \in \mathfrak{M}(\mathcal{H})$,

$$S \cup A_1,\ S \setminus A_1,\ A_1 \setminus S \in \mathfrak{M}(\mathcal{H}),$$

т.е. $S \in \mathfrak{T}(A_1)$, а значит, $\mathcal{H} \subset \mathfrak{T}(A_1)$ и минимальный монотонный класс $\mathfrak{M}(\mathcal{H}) \subset \mathfrak{T}(A_1)$. Следовательно, $A_1 \cup A_2,\ A_1 \setminus A_2,\ A_2 \setminus A_1 \in \mathfrak{M}(\mathcal{H})$ для $\forall\, A_1, A_2 \in \in \mathfrak{M}(\mathcal{H})$, что и требовалось. ⤆

Кольцо в этом утверждении замкнуто относительно операций пересечения и объединения. В следующей теореме порождающий класс замкнут только относительно операции пересечения.

598| **Теорема.** *Пусть $\mathcal{H}$ — класс подмножеств Ω, замкнутый относительно конечных пересечений. Если минимальная λ-система $\Lambda(\mathcal{H}) \ni \Omega$, то $\sigma(\mathcal{H}) \subset \Lambda(\mathcal{H})$.*

$\Rrightarrow$ Достаточно показать (с учетом утверждения II) предыдущей леммы), что λ-система $\Lambda(\mathcal{H})$ замкнута относительно конечных пересечений.

Зафиксируем произвольное $F \in \mathcal{H}$ и покажем, что класс множеств

$$\mathrm{T}(F) := \left\langle B \subset \Omega : \ B \cap F \in \Lambda(\mathcal{H}) \right\rangle \supset \Lambda(\mathcal{H}).$$

Так как $\Lambda(\mathcal{H})$ — минимальная λ-система, содержащая $\mathcal{H}$, и $\mathcal{H} \subset \mathrm{T}(F)$ (последний факт следует из того, что $B \cap F \in \mathcal{H} \subset \Lambda(\mathcal{H})$ при $B \in \mathcal{H}$), то для доказательства включения $\mathrm{T}(F) \supset \Lambda(\mathcal{H})$ достаточно показать, что и $\mathrm{T}(F)$ есть λ-система.

Монотонность класса $\mathrm{T}(F)$ следует из монотонности λ-системы $\Lambda(\mathcal{H})$. Если теперь $B_1 \subset B_2, B_1, B_2 \in \mathrm{T}(F)$, то $(B_2 \setminus B_1) \cap F = (B_2 \cap F) \setminus (B_1 \cap F) \in \Lambda(\mathcal{H})$ ввиду замкнутости $\Lambda(\mathcal{H})$ относительно монотонных разностей.

Таким образом, для $\forall F \in \mathcal{H}$ и $B \in \Lambda(\mathcal{H})$ пересечение $B \cap F \in \Lambda(\mathcal{H})$. Отсюда следует, что для $\forall B \in \Lambda(\mathcal{H})$ класс $\mathrm{T}(B) \supset \mathcal{H}$. Аналогично предыдущему получаем, что λ-система $\mathrm{T}(B) \supset \Lambda(\mathcal{H})$, т.е. $\forall F, B \in \Lambda(\mathcal{H})$ пересечение $B \cap F \in \Lambda(\mathcal{H})$. $\Lleftarrow$

D. Топологические пространства

Обсудим вкратце некоторые вопросы общей топологии, потребность в которых возникает при построении вероятностных пространств.

▸ Задать *топологию* в каком-либо пространстве Ω — значит описать класс $\mathcal{T}$ его подмножеств, удовлетворяющий трём аксиомам:

a) $\Omega, \varnothing \in \mathcal{T}$;

b) конечное пересечение множеств из $\mathcal{T}$ принадлежит $\mathcal{T}$;

c) любое объединение множеств из $\mathcal{T}$ принадлежит $\mathcal{T}$.

✦ Множества из топологии называются *открытыми*; дополнения открытых множеств суть *замкнутые* множества. Любое открытое множество, содержащее заданную точку $x \in \Omega$, называется *окрестностью* x (иногда чуть шире — любое множество, содержащее открытое множество, содержащее x).

П р и м е р ы. а) Естественная топология на числовой прямой задаётся открытыми множествами, которые с каждой своей точкой x содержат и некоторый накрывающий эту точку открытый интервал $(a;c) \ni x$.

б) В любом пространстве можно задать тривиальную (самую бедную) топологию $\mathcal{T} = \langle \varnothing, \Omega \rangle$ и дискретную (самую богатую) топологию $\mathcal{T} = \langle B : B \subset \Omega \rangle$, в которой все подмножества открыты (и замкнуты).

✦ Пусть $B \subset \Omega$. Объединение $B^{o} := \bigcup_{O \subset B} O$ всех открытых множеств, лежащих внутри B, т.е. наибольшее открытое множество с этим свойством, называется *внутренней частью* (*внутренностью*) B. Множество $[B] := ((B^c)^o)^c$ называется *замыканием* B; $\partial B := [B] \setminus B^o$ — *граница* B.

► Способы описания топологии (класса открытых множеств):

— задать *базу* $\mathfrak{B}_{\mathcal{T}}$ топологии, т.е. описать класс подмножеств такой, что любое открытое множество есть объединение элементов базы. Семейство $\mathfrak{B} \subset \mathcal{T}$ есть база топологии $\mathcal{T}$ т. т. т. когда для любой точки $x \in \Omega$ и любой окрестности $U \ni x$ найдется $V \in \mathfrak{B}$, что $x \in V \subset U$;

— задать *базис* (*предбазу*) $\mathfrak{b}_{\mathcal{T}}$ топологии, т.е. описать такой класс подмножеств, для которого система всех конечных пересечений образует базу топологии;

— описать топологию как минимальную, удовлетворяющую заданному требованию, например как минимальную топологию, в которой непрерывны функции из некоторого класса функций.

► В метрическом пространстве с метрикой (расстоянием) $\rho(x,y)$ между любыми точками $x, y \in \Omega$ базу топологии образуют все шары —

$$\mathfrak{B}_{\mathcal{T}} = \left\langle U_r(x) := \{\omega \in \Omega : \rho(x,\omega) < r\} : \ x \in \Omega, r > 0 \right\rangle.$$

П р и м е р ы. а) В евклидовом пространстве $\Omega = \mathbb{R}^k$ с обычной метрикой $\rho(\vec{x},\vec{y}) = \|\vec{x} - \vec{y}\|$ существует счётная база, состоящая из шаров с центрами в точках с рациональными координатами и рациональными радиусами (или из открытых параллелепипедов с рациональными вершинами).

б) Базу дискретной топологии задают все одноточечные множества.

► Подмножество $K \subseteq \Omega$ называется *компактом* (*компактным множеством*) т. т. т. когда из любого его покрытия открытыми множествами можно выделить конечное покрытие:

$$\Big[\bigcup\nolimits_{\beta} B_\beta \supset K, \quad \mathcal{B} = \langle B_\beta \rangle \subset \mathcal{T} \Big] \Rightarrow \Big[\bigcup\nolimits_1^n B_{\beta_j} \supset K, \quad \langle B_{\beta_j} \rangle_1^n \subset \mathcal{B}, \ n < \infty. \Big]$$

✦ Множество K компактно т. т. т. когда из любой системы замкнутых подмножеств $Z_\beta \subset K$, $\beta \in \mathbb{B}$, имеющей пустое пересечение $\bigcap_{\beta\in\mathbb{B}} Z_\beta = \varnothing$, можно выделить конечную подсистему с пустым пересечением.

✦ Если пространство $(\Omega, \mathcal{T})$ *хаусдорфово* (т.е. любые две точки Ω имеют непересекающиеся окрестности), то любой его компакт замкнут.

✦ Если топология определяется некоторой метрикой, то множество K компактно т. т. т. когда из любой его последовательности можно выделить сходящуюся (к элементу K) подпоследовательность. (На заре становления топологии как раздела математики именно так определялись компактные множества, а компактные множества в теперешнем понимании назывались бикомпактными.)

✦ В евклидовом пространстве $\mathbb{R}^k$ множество компактно т. т. т. когда оно ограничено и замкнуто.

► Последовательность $\left\langle x_n \right\rangle_1^\infty \subset \Omega$ сходится к точке $x_0 \in \Omega$ в топологии $\mathcal{T}$, если для любой окрестности $U(x_0) \in \mathcal{T}$ точки x_0, начиная с некоторого номера N (возможно, зависящего от выбранной окрестности), весь «хвост» последовательности попадает в эту окрестность: $x_n \in U(x_0)$ для $\forall\, n \geqslant N$.

► Отображение $h : (\Omega_1, \mathcal{T}_1) \mapsto (\Omega_2, \mathcal{T}_2)$ между двумя топологическими пространствами называется *непрерывным* в точке $x \in \Omega_1$, если прообраз любой окрестности точки $h(x)$ содержит окрестность точки x. Отображение h непрерывно всюду т. т. т. когда $h^{-1}(\mathcal{T}_2) \subset \mathcal{T}_1$, т.е. прообраз любого открытого множества (в топологии $\mathcal{T}_2$) открыт (в топологии $\mathcal{T}_1$). Если $\mathfrak{B}_2$ — база топологии $\mathcal{T}_2$, то отображение h непрерывно всюду т. т. т. когда $h^{-1}(\mathfrak{B}_2) \subset \mathcal{T}_1$.

► Пусть $(X_\alpha, \mathcal{T}_\alpha)$, $\alpha \in \aleph$, — семейство топологических пространств. *Прямое произведение* $\Omega = \prod_\alpha X_\alpha$ определяется как пространство всех правил $\omega(\alpha)$ $(= \omega_\alpha)$, которые каждому $\alpha \in \aleph$ ставят в соответствие некоторый элемент X_α. Если $X_\alpha = X$ при $\forall\, \alpha \in \aleph$, то $\Omega = X^\aleph$ — класс всех функций из $\aleph$ в X. *Тихоновской* топологией на прямом произведении называется минимальная топология, относительно которой непрерывны все проекции на координатные пространства, т.е. функции $h_\alpha : \Omega \mapsto X_\alpha$ такие, что $h_\alpha(\omega) = \omega_\alpha$.

✦ Базис тихоновской топологии образуют все множества вида $\{\omega \in \Omega : \omega_\alpha \in B_\alpha\}$, где $B_\alpha \in \mathfrak{B}_\alpha, \alpha \in \aleph$, $\mathfrak{B}_\alpha$ — база топологии $\mathcal{T}_\alpha$.

✦ Функция $h : (Y, \mathcal{T}) \mapsto \Omega$ $(= \prod_\alpha X_\alpha)$ непрерывна в тихоновской топологии т. т. т. когда для $\forall\, \alpha \in \aleph$ непрерывна координатная функция $h(y)_\alpha$ $(\in X_\alpha)$.

✦ Последовательность $\left\langle \omega^{(n)}, n = 1, 2, \dots \right\rangle \subset \Omega$ сходится к ω' в тихоновской топологии т. т. т. когда она сходится поточечно (покоординатно):

$$\omega^{(n)} \to \omega' \quad \Leftrightarrow \quad \omega_\alpha^{(n)} \to \omega'_\alpha \ \text{ для } \ \forall\, \alpha \in \aleph\,.$$

✦ При $\aleph = \{1, 2, \ldots, k\}$ прямое произведение $X_1 \times \ldots \times X_k$ может быть описано как набор всех упорядоченных (т.е. для которых важен порядок расположения) векторов $(x_1, \ldots, x_k)$, где i-й элемент $x_i \in X_i$. В пространстве $\mathbb{R}^k$ (т.е. при $X_i = \mathbb{R}^1, \forall\, i = 1, \ldots, k$) с топологией каждого координатного пространства, определяемой метрикой $\rho(x, y) = |x - y|$, тихоновская топология совпадает с топологией евклидового пространства.

✦ При $\aleph = \mathbb{N} = \{1, 2, \ldots\}$ и $(X_i, \mathfrak{B}_i) = (X, \mathfrak{B}), \forall\, i \in \mathbb{N}$, прямое произведение $\prod_i X_i = X^{\mathbb{N}} = X^\infty$ может быть описано как набор всех последовательностей элементов из X. Базу тихоновской топологии задают все цилиндрические множества вида $\prod_i B_i$, где $B_i \in \mathfrak{B}$, $i \leqslant n$, и $B_i = X$ для $\forall\, i > n$. Если топология в X определяется метрикой ρ, то тихоновская топология совпадает с топологией, задаваемой метрикой, например:

$$\rho_\infty(\boldsymbol{x}, \boldsymbol{y}) \ = \ \sum_1^\infty \frac{\rho(x_k, y_k)}{2^k(1 + \rho(x_k, y_k))}, \quad \boldsymbol{x} = (x_1, \ldots), \, \boldsymbol{y} = (y_1, \ldots) \in X^\infty.$$

► Полное сепарабельное метрическое пространство называется *польским* пространством.

✦ Метрическое пространство (Ω, ρ) с метрикой ρ *полно*, если любая его фундаментальная последовательность сходится:

$$\forall\, \varepsilon > 0 \ \exists N \ \forall\, n, m > N : \rho(x_n, x_m) < \varepsilon \quad \Rightarrow \quad \exists x \in \Omega : \ \rho(x_n, x) \to 0.$$

✦ Топологическое пространство $(\Omega, \mathcal{T})$ *сепарабельно*, если найдётся счётное всюду плотное подмножество $\mathfrak{S} \subset \Omega$, т.е. любое открытое множество из Ω имеет непустое пересечение с $\mathfrak{S}$.

✦ В сепарабельном метрическом пространстве борелевская σ-алгебра, порождённая топологией, совпадает с σ-алгеброй, порождённой системой счётных шаров с центрами из всюду плотного множества и рациональными радиусами.

► Действительное линейное пространство E называется евклидовым, если на нём задано скалярное произведение, т.е. задана функция, обозначаемая (x, y), определённая для каждой пары элементов $x, y \in E$ и удовлетворяющая аксиомам:

1) $(x, y) = (y, x)$, 2) $(x, x) \geqslant 0$, $(x, x) = 0 \ \Leftrightarrow \ x = 0$,

3) $(x_1+x_2,y)=(x_1,y)+(x_2,y)$, 4) $(cx,y)=c(x,y),\ c\in\mathbb{R}^1$.

✦ В евклидовом пространстве норма определяется как $\|x\|=\sqrt{(x,x)}$.

✦ Если евклидово пространство полно относительно нормы, определяемой скалярным произведением, то такое пространство называется *гильбертовым* (см. [13, гл. III, §4], [9, гл. IV, 26]).

✦ Пусть $(\Omega,\mathscr{F},\mu)$ — пространство с σ-конечной мерой. Совокупность действительных функций $\mathcal{L}_2(\Omega,\mathscr{F},\mu)=\left\langle \xi:\Omega\mapsto\mathbb{R}^1:\ \int_\Omega \xi^2\,d\mu<\infty\right\rangle$, интегрируемых с квадратом, образует гильбертово пространство относительно скалярного произведения $(\xi,\eta)=\int_\Omega \xi\eta\,d\mu$, если считать неразличимыми функции, совпадающие почти всюду по мере μ.

✦ Если H' — замкнутое линейное подпространство гильбертова пространства H, то для любого элемента $x\in H$ найдётся единственный элемент $x'\in H'$ такой, что $\|x-x'\|=\min_{y\in H'}\|x-y\|$, причём разность $x-x'$ ортогональна к H'. Последнее означает, что $(x-x',y)=0$ для $\forall\, y\in H'$. Например, совокупность $\mathcal{L}_2(\Omega,\mathscr{A},\mathbf{P})$ всех интегрируемых с квадратом функций на вероятностном пространстве $(\Omega,\mathscr{F},\mathbf{P})$, измеримых относительно некоторой σ-подалгебры $\mathscr{A}\subset\mathscr{F}$, есть замкнутое линейное подпространство $\mathcal{L}_2(\Omega,\mathscr{F},\mathbf{P})$. Поэтому для любой сл.в. ξ, у которой конечен второй момент, найдётся единственная (п.н.) сл.в. η, измеримая относительно $\mathscr{A}$, доставляющая минимум среднеквадратической ошибки $\mathbf{E}\left[(\xi-\eta)^2\right]$. Эту сл.в. можно назвать условным математическим ожиданием $\mathbf{E}\left[\,\xi\,\middle|\,\mathscr{A}\,\right]$ относительно σ-алгебры $\mathscr{A}$.

E. Выпуклые функции

Определения. Область $\mathcal{X}\subset\mathbb{R}^k$ называется *выпуклой*, если вместе с любыми двумя точками $\vec{x}_1,\vec{x}_2\in\mathcal{X}$ она содержит и весь отрезок, соединяющий эти точки: $\alpha\vec{x}_1+(1-\alpha)\vec{x}_2\in\mathcal{X},\quad \forall\,\alpha\in[0;1]$.

Функция $h:\mathcal{X}\mapsto\mathbb{R}^1$, заданная на выпуклой области $\mathcal{X}\subset\mathbb{R}^k$, называется:

— *выпуклой* (книзу), если для $\forall\,\vec{x}_1\neq\vec{x}_2\ (\in\mathcal{X})$ и $\forall\,\alpha\in[0;1]$

$$h(\alpha\vec{x}_1+(1-\alpha)\vec{x}_2)\leqslant\alpha h(\vec{x}_1)+(1-\alpha)h(\vec{x}_2); \tag{$*$}$$

— *строго выпуклой*, если неравенство $(*)$ строгое для $\forall\,\alpha\in(0;1)$;

— *вогнутой*, если выпукла функция $-h$.

Другими словами, любая дуга графика выпуклой функции лежит ниже хорды, соединяющей края этой дуги. Рисунок 9 иллюстрирует случай $k=1$; здесь

$$x_\alpha = \alpha x_1 + (1-\alpha)x_2,$$
$$y_\alpha = \alpha h(x_1) + (1-\alpha)h(x_2).$$

На рисунке приведена ещё касательная, иллюстрирующая ключевое свойство выпуклых функций.

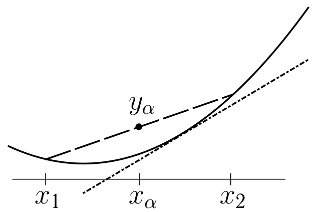


Рис. 9. Свойство выпуклости функции

✓ Функция h выпукла т. т. т. когда для любой внутренней точки $\vec{x}_0 \in \mathcal{X}^o$ можно построить так называемую опорную плоскость («касательную» линию), проходящую через точку графика $(\vec{x}_0, h(\vec{x}_0))$ и целиком лежащую не выше графика функции (ниже графика при $\vec{x} \neq \vec{x}_0$ для строго выпуклых функций):

$$(\circledast) \qquad \forall\, \vec{x}_0 \in \mathcal{X}^o\ \exists \vec{b}: \quad \vec{b}^\flat(\vec{x}-\vec{x}_0) + h(\vec{x}_0) \leqslant h(\vec{x}),\ \forall\, \vec{x} \in \mathcal{X}.$$

⤇ Докажем этот факт для случая $k=1$ и области $\mathcal{X} = (a;c)$. Пусть $x_0 \in (a;c)$. Покажем, что функция

$$q(x) \;=\; \frac{h(x)-h(x_0)}{x-x_0}, \qquad x \neq x_0,$$

не убывает. Путь $x' < x_0 < x$, тогда можно представить $x_0 = \alpha x' + (1-\alpha)x$, где $\alpha = (x-x_0)/(x-x') \in (0;1)$. В силу выпуклости h имеем $h(x_0) \leqslant \alpha h(x') +$ $+\,(1-\alpha)h(x)$, т.е. $\alpha(h(x')-h(x_0)) \geqslant (1-\alpha)(h(x_0)-h(x))$. Таким образом,

$$\big(h(x')-h(x_0)\big)\,\frac{x-x_0}{x-x'} \;\geqslant\; \big(h(x_0)-h(x)\big)\,\frac{x_0-x'}{x-x'},$$

т.е. $-q(x') \geqslant -q(x)$. Остальные случаи расположения x', x, x_0 разбираются аналогичным образом. Положим $b = \inf_{x>x_0} q(x)$. Тогда для любых $x' < x_0 < x$ справедливо $q(x') \leqslant b \leqslant q(x)$. Отсюда следует, во-первых, что $-\infty < b < \infty$, и, во-вторых, что заявленное неравенство справедливо. ⤆

✓ Функция h выпукла т. т. т. когда выпукла область, лежащая над графиком функции, — так называемый надграфик (или epigráph) функции: $\mathrm{epi}(h) =$ $= \{(\vec{x}, y):\ y \geqslant h(\vec{x}),\ \vec{x} \in \mathcal{X}\}$.

✓ Из предыдущего свойства (как, впрочем, и из определения) следует неравенство Йенсена, названное в честь первого исследователя выпуклых функций:

$$h\Big(\sum_1^n \alpha_j \vec{x}_j\Big) \leqslant \sum_1^n \alpha_j h(\vec{x}_j),$$

$$\forall\, n \geqslant 1,\ \{\vec{x}_j\}_1^n \subset \mathcal{X},\ 0 \leqslant \alpha_1, \ldots, \alpha_n \leqslant 1,\ a_1 + \ldots + \alpha_n = 1.$$

✧ Пример. Функция $h(\vec{x}) = \|\vec{x}\| = \sqrt{\vec{x}^\flat \vec{x}}$ на $\mathbb{R}^k$ выпукла, т.к. по неравенству треугольника $\|\alpha\vec{x}_1 + (1-\alpha)\vec{x}_2\| \leqslant \alpha\|\vec{x}_1\| + (1-\alpha)\|\vec{x}_2\|$.

599| Упр. Докажите, что ограниченная выпуклая функция непрерывна всюду во внутренней части области определения.

600| Упр. Приведите пример: а) разрывной ограниченной выпуклой функции; б) непрерывной неограниченной выпуклой функции.

✓ Если функция $h(x)$, $x \in \mathbb{R}^1$, всюду дифференцируема, то она выпукла т. т. т. когда её производная h' всюду возрастает.

✓ Дважды дифференцируемая всюду функция h выпукла т. т. т. когда второй дифференциал $d^2h = d\vec{x}^\flat h''(\vec{x}_0)d\vec{x}$ в любой точке $\vec{x}_0$ есть неотрицательная функция $d\vec{x}$; в одномерном случае — т. т. т. когда $h''(x_0) \geqslant 0,\ \forall\, x_0$.

Пример. Функция $h(x) = e^x$ строго выпукла, т.к. $h''(x) > 0$. Уравнение касательной к графику этой функции в точке $x = 0$, очевидно, имеет вид $y = x+1$. Это доказывает часто используемое неравенство $e^x \geqslant x+1$. Заметим, что ввиду строгой выпуклости равенство здесь имеет место только при $x = 0$.

Аналогично, функция $h(x) = \ln x,\ x > 0$, строго вогнута и её касательная в т. $x = 1$ равна $y = x - 1$. Поэтому при любых $x > 0, x \neq 1$, справедливо неравенство $\ln x < x - 1$.

Применяя к этой функции неравенство Йенсена с $\alpha_j = \dfrac{1}{n},\ j = \overline{1,n}$, получаем неравенство $\ln(\sum_1^n x_j/n) \geqslant \frac{1}{n}\sum_1^n \ln(x_j)$ или, после потенцирования, $\frac{1}{n}\sum_1^n x_j \geqslant \big(\prod_1^n x_j\big)^{1/n}$, т.е. среднее арифметическое неотрицательных чисел не меньше среднего геометрического.

Примеры. 1. Функция $h(x) = x^a,\ x > 0$, строго выпуклая при $a > 1$, т.к. $h''(x) = a(a-1)x^{a-2} > 0$. Поэтому для $\forall\, x \neq y\ (> 0)$

$$(x+y)^a = 2^a\Big(\frac{1}{2}x + \frac{1}{2}y\Big)^a < 2^a\Big(\frac{1}{2}x^a + \frac{1}{2}y^a\Big) = 2^{a-1}\big(x^a + y^a\big). \tag{25}$$

При $0 < a < 1$ функция $h(x) = x^a,\ x > 0$, строго вогнута, поэтому

$$(x+y)^a > \frac{x^a + y^a}{2^{1-a}}, \qquad x, y > 0,\ x \neq y,\ 0 < a < 1.$$

Например, $\sqrt{2}\sqrt{x+y} > \sqrt{x} + \sqrt{y}$.

2. Как уже отмечалось, функция $h(x_1, x_2) = \sqrt{x_1^2 + x_2^2} = \|(x_1, x_2)\|$ выпуклая. Вектор первых производных в точке $(x_1, x_2) = (a, b)$, $a^2 + b^2 \neq 0$, равен $h' = (a, b)/h(a, b)$. Построив уравнение касательной плоскости в точке (a, b), приходим (после элементарных преобразований) к неравенству

$$\sqrt{x_1^2 + x_2^2} \ \geqslant \ \frac{a x_1 + b x_2}{\sqrt{a^2 + b^2}} \, .$$

Впрочем, это неравенство, как и предыдущее для $x^{1/2}$, легко доказывается возведением в квадрат обеих его частей. Продемонстрированный здесь способ, помимо всего прочего, даёт инструмент для получения подобных неравенств.

F. Некоторые факты из теории матриц

Если A — какая-либо матрица, то через $A^\flat$ будем обозначать транспонированную матрицу.

✓ Симметричная квадратная матрица $A = A^{(k \cdot k)}$ размера $k \cdot k$ называется *неотрицательно (положительно) определённой*, обозначается $A \geqslant 0$ $(A > 0)$, если $\vec{x}^\flat A \vec{x} \geqslant_{(>)} 0$ для любого ненулевого вектора $\vec{x} \in \mathbb{R}^k$. По критерию Сильвестра матрица $A \geqslant 0$ т. т. т. когда все её главные миноры неотрицательны; $A > 0$ т. т. т. когда положительны k главных миноров, построенных от левого верхнего угла матрицы.

✓ Любая квадратная матрица $A = A^{(k \cdot k)}$ имеет не более k собственных чисел λ, для которых определитель $|A - \lambda \dot{\mathbb{I}}| = 0$. Каждому собственному числу λ отвечает собственное подпространство векторов $\vec{x}$, удовлетворяющих характеристическому уравнению $A\vec{x} = \lambda\vec{x}$.

✓ Матрица $A \geqslant 0$ т. т. т. когда все её собственные числа неотрицательны (положительны т. т. т. когда $A > 0$).

✓ Квадратная матрица Q называется ортогональной, если $Q^\flat Q = Q Q^\flat = \dot{\mathbb{I}}$. Преобразование $\vec{y} = Q^\flat \vec{x}$ с ортогональной матрицей геометрически эквивалентно повороту системы координат без растяжения.

✓ Для любой матрицы $A \geqslant 0$ найдётся ортогональная матрица Q и диагональная матрица $\Gamma = \mathrm{diag}(\gamma_1^2, \ldots, \gamma_k^2)$ с неотрицательными элементами такие, что $A = Q \Gamma Q^\flat$. Количество ненулевых диагональных элементов Γ совпадает с рангом матрицы A. Квадратичная форма $\vec{x}^\flat A\,\vec{x}$ преобразованием $\vec{y} = Q^\flat\vec{x}$ приводится к каноническому виду $\vec{y}^\flat\Gamma\,\vec{y} = \sum_1^k \gamma_j^2 y_j$.

✓ Если $A = Q\Gamma Q^\flat \geqslant 0$ с ортогональной матрицей Q и диагональной матрицей $\Gamma = \operatorname{diag}(\gamma_1^2, \dots, \gamma_k^2)$, то матрица $\widetilde{A} = Q\operatorname{diag}(\gamma_1, \dots, \gamma_k)\,Q^\flat$ удовлетворяет уравнению $\widetilde{A}^\flat\widetilde{A} = A$, т.е. $\widetilde{A} = \sqrt{A}$.

G. Указания к решению задач

✦ 460, стр. 424. б) $B_kB_m = \varnothing$. Пусть ω принадлежит левой части, k — такое минимальное число n, что $Q_n \ni \omega$. Тогда $\omega \in Q_k, \omega \notin \bigcup_1^{k-1} Q_j$, т.е. $\omega \in B_k$. в) $\dot{\mathrm{I}}(\bigcap_k B_k) = \inf_k \dot{\mathrm{I}}(B_k)$. Индикатор $\dot{\mathrm{I}}\left(\,\underline{\lim}_n Q_n\right) = \dot{\mathrm{I}}\left(\lim_\uparrow \bigcap_{k=n}^\infty Q_k\right) = \lim_\uparrow \dot{\mathrm{I}}\left(\bigcap_{k=n}^\infty Q_k\right) = \lim_n \inf_{k\geqslant n} \dot{\mathrm{I}}(Q_k) = \underline{\lim}_n\ \dot{\mathrm{I}}(Q_n)$.

✦ 461, стр. 425. $B = B \biguplus \varnothing$, $\quad A \biguplus B = A \biguplus B \biguplus \varnothing \biguplus \varnothing \dots$.

✦ 462, стр. 425. $\#((A \uplus B)\mathcal{S}) = \#(A\mathcal{S}) + \#(B\mathcal{S})$ — очевидно. Рассмотреть случаи $\mu\big(\biguplus_1^\infty B_k\big) < \infty$ и $= \infty$. Во втором случае $\forall M\ \exists K:\ \mu\big(\biguplus_1^K B_k\big) > M$.

✦ 463, стр. 426. Если $Q, B \in \mathcal{H}$, то $B \supset QB \in \mathcal{H} \ \Rightarrow\ B = QB + \biguplus_1^k Q_j = QB + (B \setminus Q)$.

✦ 465, стр. 428. Пересечение $(a;c] \cap (b;d]$ или пусто, или представляет собой интервал того же вида. Если $(a;c] \subsetneq (b;d]$, то $(b;d] = (a;c] + \big[(b;a] \uplus (c;d]\big]$. Относительно прямоугольников см. 466, стр. 428.

✦ 466, стр. 428. В соответствии с 463, стр. 426, пусть $A_1 \times A_2 \supset B_1 \times B_2 \ \Rightarrow\ A_1 = B_1 + \uplus_j Q_j,\ A_2 = B_2 + \uplus_n R_n \ \Rightarrow\ A_1 \times A_2 = B_1 \times B_2 + \uplus_j Q_j \times A_2 + \uplus_n A_2 \times R_n + \uplus_{jn} Q_j \times R_n$.

✦ 469, стр. 431. $S_n = [n; \infty)$.

✦ 470, стр. 431. Если интервалы покрывают все рациональные точки, то они будут покрывать и весь отрезок $[0;1]$. Нет ни одного интервала, полностью лежащего в множестве рациональных чисел.

✦ 471, стр. 432. Если $\bigcup_i B_i \supset G$, то $\biguplus_k Q_k = \bigcup_i B_i \supset G$ и $\mu(\biguplus_k Q_k) \leqslant \sum_i \mu(B_i)$. Поэтому инфимум по более узкому классу будет совпадать с инфимумом по широкому классу покрытий. Любой элемент кольца $Q_k = \biguplus_{j=1}^{J_k} Z_j$, Z_j из полукольца, и по свойству рядов из положительных чисел $\mu(\biguplus_k \biguplus_j Z_j) = \sum_k \sum_j \mu(Z_j)$.

✦ 472, стр. 432. (i_3) — $\mu(\Omega) \leqslant \sum_i \mu(B_i)$ в силу σ-полуаддитивности.

✦ 473, стр. 432. В силу монотонности и полуаддитивности внешней меры: $\mu^*(B) \leqslant \mu^*(BG) + \mu^*(BG^c) = \mu^*(BG^c) \leqslant \mu^*(B)$.

✦ 474, стр. 433. $\mathcal{H} \subset \mathcal{A}$ и $\sigma\big(H\big) \supset \mathcal{A}$.

✦ 479, стр. 436. Воспользоваться 472, стр. 432, и тем, что инфимум равен нулю т. т. т. когда найдётся последовательность элементов, сходящаяся к нулю.

✦ (4), стр. 436. $\mu(C) = \mu(A) + \mu(C \setminus A) = \mu(A) \leqslant \mu(D)$ для $\forall D \supset B \supset E \ \Rightarrow\ \mu(C) =$

$\mu(D)$, если $\mu(C \setminus A) = 0 = \mu(D \setminus E)$. Если $A \subset B \subset C$, то $C^c \subset B^c \subset A^c$ и $\mu(A^c \setminus C^c) = \mu(C \setminus A)$. Если $A_k \subset B_k \subset C_k$, то $\bigcap A_k \subset \bigcap B_k \subset \bigcap C_k$ и $\mu(\bigcap C_k \setminus \bigcap A_k) \leqslant \sum_k \mu(C_k \setminus A_k)$.

✦ 480, стр. 437. $(a;b) = \bigcup_k (a; b + \frac{1}{k}\,]$. Любое открытое множество есть счётное объединение открытых интервалов; замкнутое множество есть дополнение какого-то открытого множества до $\mathbb{R}^1$.

✦ 481, стр. 437. «$\underset{4}{=}$» $[a;b) = \lim_k (a - \frac{1}{k}\,; b - \frac{1}{k}\,] \in \sigma(\mathcal{K})$, $(a;b] = \lim_k [a + \frac{1}{k}\,; b + \frac{1}{k}\,) \in \sigma(\mathcal{K}_{[)})$. «$\underset{5}{=}$» $\mathcal{K} \subset \mathcal{U} \subset \sigma(\mathcal{K})$. «$\underset{6}{=}$» $(a;b] = (-\infty; b] \setminus (-\infty; a]$.

✦ 484, стр. 438. $A_k \in Q \Cap \mathscr{F} \;\Rightarrow\; A_k = QB_k,\ B_k \in \mathscr{F} \;\Rightarrow\; \bigcup A_k = Q \bigcup B_k,\ \bigcup B_k \in \mathscr{F} \;\Rightarrow\; \bigcup A_k \in Q \Cap \mathscr{F}$.

✦ 486, стр. 440. Класс множеств $B \in \mathscr{B}(\overline{\mathbb{R}}^1)$, удовлетворяющих условию леммы, есть σ-алгебра, содержащая интервалы, определяющие σ-алгебру $\mathscr{B}(\overline{\mathbb{R}}^1)$.

✦ 489, стр. 446. Ср. с 480, стр. 437. Так как $(\vec{a};\vec{c}] = \{\vec{x} :\ x_1 \leqslant c_1, a_2 < x_1 \leqslant c_2, \ldots\} - \{\vec{x} :\ x_1 \leqslant a_1, a_2 < x_1 \leqslant c_2, \ldots\}$, то для принадлежности к σ-алгебре $\sigma(\mathcal{K}^{\urcorner})$ достаточно установить сей факт для каждого из этих подмножеств; далее по индукции.

✦ (?!), стр. 450. Любые конечные наборы множеств из $\mathcal{B}$ суть цилиндры с общим конечномерным пространством.

✦ 487, стр. 442. Пусть F непрерывна слева, тогда $\mu[a;b] = \lim_k \mu[a; b + \frac{1}{k}\,) = \lim_k (F(b + \frac{1}{k}\,) - F(a)) = F(b+) - F(a) = F(b+) - F(a-)$.

✦ 492, стр. 448. Покажите, что функция $F(x_1, x_2) = \mathrm{sign}(x_1 x_2)\mu(\Pi)$, где Π — параллелепипед от $(\min(x_1, 0), \min(x_2, 0))$ до $(\max(x_1, 0), \max(x_2, 0))$, генерирует меру μ.

✦ 504, стр. 462. $\omega \in \xi^{-1}(\bigcup B_\alpha) \;\Leftrightarrow\; \xi(\omega) \in \bigcup B_\alpha \;\Leftrightarrow\; \big[\exists \alpha :\ \xi(\omega) \in B_\alpha\big] \;\Leftrightarrow\; \omega \in \bigcup_\alpha \xi^{-1}(B_\alpha)$.

✦ 513, стр. 468. а) $\{\xi^+ \leqslant x\}$ равно $\{\xi \leqslant x\}$ при $x \geqslant 0$ и пусто в противном случае. в) Если $0 < \xi \leqslant \eta$, то $\xi^- = 0 = \eta^-$; если $\xi \leqslant 0 < \eta$, то $\xi^- = -\xi \geqslant 0 = \eta^-$; если $\xi \leqslant \eta \leqslant 0$, то $\xi^- = -\xi \geqslant -\eta = \eta^-$.

✦ 521, стр. 471. Бесконечная сумма есть предел измеримых конечных сумм.

✦ 529, стр. 477, (L_1). $\mu(\xi^{\pm} = \eta^{\pm}) = 0$;

✦ 530, стр. 479. $\delta(h \neq \tilde{h}) = 0$, где $\tilde{h} = h(\omega_0)\,\dot{\mathrm{I}}(\omega = \omega_0)$.

✦ 531, стр. 479. По неравенству Маркова $\lim\limits_n \int x_n \, d\mu \geqslant a\mu(\xi > a),\ a > 0$.

✦ 532, стр. 479. б) Для простых очевидно.

✦ 533, стр. 479. $\mu(\xi \neq \tilde{\xi}) = 0$, где $\tilde{\xi} = \sum_i \xi(\omega_i)\,\dot{\mathrm{I}}(\omega = \omega_i)$.

✦ 535, стр. 480. В силу 531, стр. 479, можно считать $\mu(\eta = \infty) = 0$. Применить теорему Б. Леви к последовательности $\xi_n\,\dot{\mathrm{I}}(\eta < \infty) \nearrow \eta\,\dot{\mathrm{I}}(\eta < \infty)$.

✦ 541, стр. 483. $\mu_\xi(\uplus B_k) = \mu(\uplus\, \xi^{-1}(B_k)) = \sum \mu(\xi^{-1}(B_k)) = \sum \mu_\xi(B_k)$.

✦ 544, стр. 485. Меры $\lambda(aB + c)$ и $|a|\lambda(B)$ совпадают на параллелепипедах. Другой способ: $x \in aB + c \;\Leftrightarrow\; (x - c)/a \in B$; по предшествующей лемме $|a|\lambda(B) = |a| \int \dot{\mathrm{I}}_B(t)dt = \int \dot{\mathrm{I}}_B\big((x - c)/a\big)dx = \lambda(aB + c)$.

✦ 550, стр. 487. Использовать неравенство $\sqrt[2]{2} > \sqrt[3]{2}$.

✦ Замечание, стр. 488. $\int |\xi_1|^p = 0 \Rightarrow \xi_1 = 0 \Rightarrow \xi_1\xi_2 = 0$ [п.в.] .

✦ 552 , стр. 488. Рассмотреть функцию от выпуклой комбинации $G(c\,a_1 + (1-c)\,a_2), 0 < < c < 1$; применить неравенство Гёльдера к $\int_\Omega |\xi|^{ca_1}|\xi|^{(1-c)a_2}\,d\mu$ с $p = 1/c, q = 1/(1-c)$.

✦ 554 , стр. 489. Квадратичная форма не меняет знак т. т. т. когда её дискриминант неположителен.

✦ 557 , стр. 491. Применить теорему Лебега об ограниченной сходимости к последовательности $\xi\,\dot{\mathrm{I}}_{A_n}$.

✦ 562 , стр. 493. Применить лемму 561 , стр. 493, к парам $\gamma \ll \rho$ и $\rho \ll \mu$.

✦ 578 , стр. 505. Применить формулу интегрирования по частям.

✦ 583 , стр. 521. а) Пересечение прямоугольников — прямоугольник. Любые два прямоугольника задают сетку разбиения плоскости, их разность описывается ячейками сетки. Открытость (замкнутость) соответствующих сторон нужна для того, чтобы ячейки сетки не пересекались по границам. в) Рассмотреть разность вложенных интервалов и объединение непересекающихся интервалов.

✦ 585 , стр. 522. а) $A \cap B = (A \cup B) \setminus (A \setminus B)(B \setminus A)$. б) Из определения полукольца $B \setminus A = \bigcup_i S_i \overset{R2'}{\in} \mathcal{D}$.

✦ 586 , стр. 522. Если A, B_k из полукольца, то $B \setminus A = \left(\biguplus_k B_k\right) \setminus A = \biguplus_k \left(B_k \setminus A\right) = \biguplus_k \biguplus_j Q_{kj}$. Далее применить равенство $(B_1 \uplus B_2) \setminus (A_1 \uplus A_2) = (B_1 \setminus A_1) \setminus A_2 \ \uplus\ (B_2 \setminus A_1) \setminus A_2$.

✦ 587 , стр. 523. а) Правило d'Morgána. б) $A \cup B = A \cup B \cup \varnothing \cup \ldots$ в) $A \setminus B = = AB^c$. г) Любой счётный набор подмножеств содержит только конечное число различных.

✦ 589 , стр. 523. Если $\Omega = \{\omega_1, \ldots, \omega_k\}$, то любое подмножество $A \subset \Omega$ описывается вектором $(b_1, \ldots, b_k)$, элементы которого $b_j\ (= 0, 1)$ равны 1, если $\omega_j \in A,\ j = \overline{1, k}$.

✦ 594 , стр. 524. а) $\sigma\big(\mathcal{Q}_1\big), \sigma\big(\mathcal{Q}_2\big)$ — σ-алгебры, содержащие $\mathcal{Q}_1$, причём первая из них — минимальная. б, в) Применить а) и $\sigma\big(\sigma\big(\mathcal{Q}\big)\big) = \sigma\big(\mathcal{Q}\big)$.

✦ 596 , стр. 525. $\lfloor a; b\rceil,\ \infty \leqslant a < b \leqslant \infty,\ \lfloor = [$ или $(,\ \rceil =]$ или $)$.

✦ 599 , стр. 532. При доказательстве свойства $(\circledast)$ было показано, что функция $q(x)$ не убывает, поэтому существуют конечные пределы $\lim_{x\to x_0+0} q(x)$ и $\lim_{x\to x_0-0} q(x)$, что возможно, только если $\lim_{x\to x_0} h(x) = h(x_0)$.

✦ 600 , стр. 532. а) $f(x) = x^2,\ |x| < 1,\ \ f(\pm 1) = 2$, б) $f(x) = \dfrac{1}{x},\ x > 0$.

Греческий алфавит

α : альфа	β : бета	γ : гамма	δ : дельта
ε, ϵ : эпсилон	ζ : зета	η : эта	λ : лямбда
$\vartheta,\ \theta$: тета	$\varkappa$: каппа	μ : мю	ν : ню
ξ : кси	π : пи	ρ : ро	σ : сигма
τ : тау	φ : фи	χ : хи	ψ : пси
Γ : Гамма	Δ : Дельта	Θ : Тета	Λ : Лямбда
Σ : Сигма	Φ : Фи	Ψ : Пси	Ω : Омега

Готический шрифт

$\mathfrak{A}$: А	$\mathfrak{B}$: Б	$\mathfrak{C}$: Ц	$\mathfrak{D}$: Д	$\mathfrak{E}$: Е
$\mathfrak{F}$: эФ	$\mathfrak{G}$: Ж	$\mathfrak{H}$: аШ	$\mathfrak{I}$: И	$\mathfrak{J}$: Йот
$\mathfrak{K}$: Ка	$\mathfrak{L}$: эЛь	$\mathfrak{M}$: эМ	$\mathfrak{N}$: эН	$\mathfrak{O}$: О
$\mathfrak{P}$: П	$\mathfrak{Q}$: Ку	$\mathfrak{R}$: эР	$\mathfrak{S}$: эС	$\mathfrak{T}$: Т
$\mathfrak{U}$: У	$\mathfrak{W}$: дубль В	$\mathfrak{X}$: Икс	$\mathfrak{Y}$: Игрик	$\mathfrak{Z}$: Зет

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ЛИТЕРАТУРА

[1] Андерсон Т. *Введение в многомерный статистический анализ* / Т. Андерсон. – М.: Физматлит, 1963. – 500 с.

[2] Биллингсли П. *Сходимость вероятностных мер* / П. Биллингсли. – М.: Наука, 1977. – 352 с.

[3] Боровков А. А. *Теория вероятностей* / А. А. Боровков. – М.: Наука, 1976. – 472 с.

[4] Бородин А. Н. *Случайные процессы*: учебник / А. Н. Бородин. – СПб.: Лань, 2013. – 640 с.

[5] Булинский А. В. *Теория случайных процессов* / А. В. Булинский, А. Н. Ширяев. – М.: Физматлит, 2003. – 400 с.

[6] Вентцель А. Д. *Курс теории случайных процессов* / А. Д. Вентцель. – М.: Физматлит, 1996. – 402 с.

[7] Гнеденко Б. В. *Курс теории вероятностей: учебник* / Б. В. Гнеденко. – 8-е изд., испр. и доп. – М.: Едиториал УРСС, 2005. – 408 с.

[8] Гнеденко Б. В. *Предельные распределения для сумм независимых случайных величин* / Б. В. Гнеденко, А. Н. Колмогоров. – М.: ГИТТЛ, 1949. – 264 с.

[9] Данфорд Н. *Линейные операторы. Общая теория* / Н. Данфорд, Дж. Шварц. – М.: ИЛ, 1962. – 896 с.

[10] Дороговцев А. Я. *Элементы общей теории меры и интеграла:* учеб. пособие / А. Я. Дороговцев. – Киев: Вища школа, 1988. – 152 с.

[11] Зорич В. А. *Математический анализ* Ч. II/ В. А. Зорич. – 4-е изд. – М.: МЦНМО, 2002. – 794 с.

[12] Ильин В. А. *Основы математического анализа.* Ч. II: учеб.пособие / В. А. Ильин, Э. Г. Позняк. – М.: Физматлит, 2002. – 464 с.

[13] Колмогоров А. Н. *Элементы теории функций и функционального анализа* / А. Н. Колмогоров, С. В. Фомин. – 6-е изд. – М.: Наука, 1989. – 543 с.

[14] Лукач Е. *Характеристические функции* / Е. Лукач. – М.: Наука, 1979. – 424 с.

[15] Миллер Б. М. *Теория случайных процессов в примерах и задачах* / Б. М. Миллер, А. Р. Панков. – М.: Физматлит, 2002. – 320 с.

[16] Невё Ж. *Математические основы теории вероятностей* / Ж. Невё. – М.: Мир, 1969. – 310 с.

[17] Петров В. В. *Суммы независимых случайных величин* / В. В. Петров. – М.: Наука, 1972. – 416 с.

[18] Прохоров А. В. *Задачи по теории вероятностей* / А. В. Прохоров, В. Г. Ушаков, Н. Г. Ушаков. – М.: Наука, 1986. – 328 с.

[19] Секей Г. *Парадоксы в теории вероятностей и математической статистике:* пер. с англ. / Г. Секей. – М.: Мир, 1990. – 240 с., ил.

[20] Стоянов Й. *Контрпримеры в теории вероятностей.* Электронное издание / Й. Стоянов. – М.: МЦНМО, 2014. – 294 с.

[21] Феллер В. *Введение в теорию вероятностей и её приложения.* Т. 1 / В. Феллер. – 2-е изд. – М.: Мир, 1984. – 528 с.

[22] Феллер В. *Введение в теорию вероятностей и её приложения.* Т. 2 / В. Феллер. – 2-е изд. – М.: Мир, 1984. – 752 с.

[23] Фихтенгольц П. *Курс дифференциального и интегрального исчисления.* Т. 1 / П. Фихтенгольц. – М.: Физматлит, 2001. – 607 с.

[24] Фихтенгольц П. *Курс дифференциального и интегрального исчисления.* Т. 3 / П. Фихтенгольц. – М.: Физматлит, 2001. – 662 с.

[25] Халмош П. *Теория меры* / П. Халмош. – М.: ИЛ, 1953. – 291 с.

[26] Шилов Г. Е. *Интеграл, мера и производная. Общая теория*/ Г. Е. Шилов, Б. Л. Гуревич. – М.: Наука, 1967. – 220 с.

[27] Ширяев А. Н. *Вероятность* — 1 / А. Н. Ширяев. – М.: Изд-во МЦНМО, 2004. – 520 с.

[28] Ширяев А. Н. *Вероятность* — 2 / А. Н. Ширяев. – М.: Изд-во МЦНМО, 2004. – 408 с.

[29] Ширяев А. Н. *Задачи по теории вероятностей*: учеб. пособие / А. Н. Ширяев. – М.: Изд-во МЦНМО, 2006. – 416 с.

[30] Acosta A. de. A new proof of the Hartman–Wintner law of the iterated logarithm// *The Ann.of Probab.* – 1983. Vol. **11**, № 2. – p. 270–276.

[31] Dudley R.M. *Real analysis and probability* / R. M. Dudley. – Cambridge: University Press, 2004. – 555 p.

[32] Chow Y. S. *Probability Theory* / Y. S. Chow, H. Teiher. – 3rd ed. – N.Y.: Springer-Verlag, 1997. – 488 p.

[33] Sklar A. Fonctions de répartion á *n* dimensions et leurs marges // *Publ. Inst. Statist. Univ. P.*, 1959. Vol. 8. – p. 229–231.

Сергей Владимирович СИМУШКИН

МЕТОДЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Учебное пособие

Редакция
естественнонаучной литературы