КАЗАНСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ

ИНСТИТУТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ
И ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Кафедра математической статистики

С.В. СИМУШКИН

МЕТОДЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Часть 1

Многомерные модели
Математические основания

Учебное пособие

Казань — 2016

Принято на заседании
кафедры математической статистики

Протокол № 8 от 6 апреля 2016 года

Рецензенты:

доктор физ.-мат. наук, профессор **И.Н. Володин**,
кандидат физ.-мат. наук, доцент **О.Е. Тихонов**

С и м у ш к и н С. В.

Методы теории вероятностей. Часть 1. Многомерные модели. Математические основания: Учебное пособие / С.В. Симушкин. — Казань: Казан.ун-т, 2016. — 255 с.

В пособии разобраны методы анализа вероятностных распределений систем случайных величин (случайных векторов). 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 физико-математических специальностей университетов.

УДК 519.2 (075.8)

СОДЕРЖАНИЕ

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

(?!)	—	факт, доказательство которого доверено читателю;
⚠	—	замечание;
$\mathfrak{Ynp}$	—	упражнение для самостоятельного решения;
$\Rrightarrow$ $\Lleftarrow$	—	знаки начала и завершения доказательства;
$\circledcirc$	—	знак окончания примера (серии примеров);
$\{x : \mathfrak{A}(x)\}$	—	совокупность элементов x, удовлетворяющих $\mathfrak{A}$;
$\langle\, O_\alpha,\ \alpha \in \aleph\,\rangle$	—	совокупность объектов O_α, индексированных $\alpha \in \aleph$, $\aleph$ («алеф») — произвольный набор индексов;
$\uplus$	—	объединение непересекающихся событий;
$\forall$	—	любой, любого, любых и т.п.;
$S := D\ \ (=: S)$	—	S определяется (впервые) посредством D;
$S \overset{\mathrm{df}}{=} D$	—	напоминание, что S определялось (когда-то) через D;
$\llbracket a \rrbracket$	—	целая часть a — наибольшее целое, не больше a;
$\#(A)$	—	количество элементов A;
$\lvert M\rvert$	—	определитель матрицы M;
$M^\flat,\ \vec{x}^\flat$	—	операция транспонирования;
$\vec{x}$	—	вектор-столбец;
$\vec{x} + c$	—	$(x_1 + c, \ldots, x_k + c)$ — сдвиг координат вектор-столбца;
$\vec{x} \lessdot \vec{y}$	—	$x_j < y_j$ для $\forall\, j$;
$\xi \overset{d}{\backsim} F\ \ (\xi \overset{d}{\backsim} \eta)$	—	ξ имеет распределение F (такое, как η);
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;
Ω	—	пространство исходов ω;
$\mathcal{K}$	—	полукольцо (интервалов);
$\mathfrak{C}$	—	кольцо, порождённое классом подмножеств;
σ	—	сокращение, связанное со счётным суммированием 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	—	произвольные переменные;
s, u, v, t	—	произвольные переменные;
ξ, η, ζ	—	случайные величины, функции на Ω,
ϵ, ε	—	малые величины;
$\mathbb{N}$ $(\mathbb{N}_0)$	—	множество натуральных (целых неотрицательных) чисел;
$\mathcal{N}(\mu, \sigma^2)$	—	нормальный (гауссовский) закон (список см. стр. 249);
μ	—	мера или значение математического ожидания сл.в.;
λ	—	мера Лебега (гл. III) или параметр вероятностных законов;
φ	—	характеристическая функция (преобразование Фурье);
ϕ, Φ	—	плотность и функция распределения нормального закона;
f, F (g, G)	—	функция плотности, функция распределения сл.в., сл.вектора;
h, g	—	функции действительного (векторного) аргумента;
$\dot{\mathbb{I}}$	—	единичная матрица (размерность по контексту);
$\mathrm{I}_A = \mathrm{I}(A)$	—	индикаторная функция множества A;
$\mathbf{P}$, $\mathbf{E}$, $\mathbf{D}$	—	вероятность, матем.ожидание, дисперсия;
Γ, B	—	гамма-функция, бета-функция;
$\boldsymbol{\Sigma}$	—	матрица ковариаций;
$\sum, \displaystyle\sum$	—	знаки суммирования; индекс суммирования опускается, если он очевиден из контекста.

*В*ведение. В пособии представлены методы анализа вероятностных распределений систем случайных величин (случайных векторов). В стандартных 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 «физическая» концепция плотности как производной функции распределения здесь редко расходится с истиной. В многомерном случае такая концепция в очень неожиданных ситуациях приводит к ложным выводам.

Кроме того, в литературе по теории вероятностей слабо описаны методы исследования преобразований случайных векторов, хотя это и есть основная задача теории вероятностей (как её анонсируют авторы многих учебников). Вторая глава данного пособия в некоторой степени восполняет этот пробел.

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 III. Этот раздел пособия 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 III можно рекомендовать в качестве самостоятельного дополнительного пособия к курсу теории меры.

Структура пособия не соответствует классическому линейному построению учебников, когда каждый новый факт подкреплён фактами, приведёнными на предыдущих страницах учебника. Ввиду того, что основания теории вероятностей, описанные в главе III, занимают очень большое пространство как по объёму, так и по времени, в первых главах приводятся факты, относящиеся к многомерному анализу, опирающиеся на теоремы из последней главы. Тем самым как-бы подчёркивается, что без теории меры и без интеграла Лебега теорию вероятностей не построить, но, с другой стороны, если читатель уже знаком, хотя бы вкратце, с основами этих теорий, то он вполне способен освоить материал первых двух глав, используя последнюю главу как дополнительный источник информации.

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

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, 219, стр. 233 — теорема 219 на стр. 233; в противном случае ссылка содержит только номер, например, (17) — формула (17) на одной из ближайших страниц выше или ниже по тексту. Если необходима ссылка на близлежащую формулу, обращение к которой более нигде не потребуется, то эта ссылка оформляется в виде символов типа $(*)$ или $(\natural)$ и т.п.

Краткое введение в методологию теории вероятностей. 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%-ой уверенностью), но не относится к области интересов теории вероятностей. В цифровую форму ответы на подобные вопросы облекать смысла нет никакого, и правильные ответы на них должны состоять из фраз типа «ну, небольшая» или «значительная» и т.п.

Вероятностная модель включает в себя описание всех мыслимых исходов эксперимента ω, т.е. описание так называемого пространства элементарных исходов Ω $(\ni \omega)$. Элементарными они называются потому, что их дробление на более мелкие части не предполагается. Появление того или иного элементарного исхода описывается с помощью вероятностной меры $\mathbf{P}$. Расцвет теории вероятностей начался с того момента, когда было понято, что задание вероятности удобнее всего проводить в рамках теории меры. Другими словами, необходимо сначала определить класс $\mathscr{F}$ измеримых подмножеств Ω, каковые здесь уже называются событиями, и на этих событиях задать нормированную (т.е. с единичной мерой всего Ω) сигма-аддитивную меру $\mathbf{P}$. Тройка объектов $(\Omega, \mathscr{F}, \mathbf{P})$ и представляет собой вероятностное пространство — вероятностную модель эксперимента.

Чаще всего наблюдается не сам исход ω, а его преобразование с помощью некоторой действительной функции. Для того чтобы можно было отвечать на вопросы о вероятностях значений этой функции, необходимо (подобно Ω) описать класс событий в $\mathbb{R}^1$. Удобнее всего здесь оказался класс, называемый борелевской сигма-алгеброй (σ-алгеброй) $\mathscr{B}$. По определению борелевская σ-алгебра есть минимальная σ-алгебра, содержащая все открытые подмножества $\mathbb{R}^1$. Легко показать (см. главу III), что эту σ-алгебру можно охарактеризовать как минимальную σ-алгебру, содержащую все интервалы определённого вида, например, только интервалы вида $(a;c]$ или интервалы вида $(-\infty;c]$, $a, c \in \mathbb{R}^1$. Таким образом, нас интересуют только те функции $\xi(\omega)$, $\omega \in \Omega$, у которых прообразы $\xi^{-1}(B) = \{\xi \in B\} := \{\omega : \xi(\omega) \in B\}$ любых борелевских подмножеств $B \in \mathscr{B}$ измеримы. В теории вероятностей такие функции называются случайными величинами (в дальнейшем будем писать кратко сл.в. — формы прочтения этой записи в зависимости от падежа и множественного или единственного числа будут зависеть, конечно, от контекста). Как отмечено выше, сл.в. ξ можно определить как функцию, для которой множество элементарных исходов $\{\xi \leqslant x\} := \{\omega \in \Omega : \xi(\omega) \leqslant x\} \in \mathscr{F}$, то есть измеримо при $\forall\, x \in \mathbb{R}^1$.

Сл.в. ξ порождает (индуцирует) вероятностную меру $\mathbb{P}$

на измеримом борелевском пространстве $(\mathbb{R}^1, \mathscr{B})$ по формуле $\mathbb{P}\{B\} = \mathbf{P}\{\xi \in B\}$, $B \in \mathscr{B}$. Эта мера называется распределением сл.в. Её можно описать с помощью чуть меньшего класса событий из $\mathbb{R}^1$, а именно класса интервалов вида $(-\infty; x]$, $x \in \mathbb{R}^1$. Функция $F(x) = \mathbb{P}\{(-\infty; x]\} = \mathbf{P}\{\xi \leqslant x\}$, $x \in \mathbb{R}^1$, называется функцией распределения (ф.р.). Таким образом, сл.в. можно определить как случайную реализацию эксперимента с действительным исходом, значения (реализации) которой появляются в соответствии с вероятностным распределением $\mathbb{P}$ (или ф.р. F). Подобная интерпретация возможна ещё и потому, что любая числовая характеристика сл.в., как-то: вероятность некоторого события или среднее значение — может быть вычислена либо как интеграл Лебега от некоторой функции $h(x)$ относительно меры $\mathbb{P}$, либо как интеграл Лебега от функции $h(\xi(\omega))$ относительно меры $\mathbf{P}$:

$$\mathbf{E}h(\xi) := \int_\Omega h(\xi(\omega))\, d\,\mathbf{P} = \int_\mathbb{R} h(x)\, d\mathbb{P} \quad \Big(= \int_\mathbb{R} h(x)\, \mathbb{P}(dx)\Big).$$

Этот интеграл называют математическим ожиданием функции $h(\xi)$. Поскольку чаще всего описание сл.в. осуществляют с помощью ф.р., для обозначения интеграла относительно меры $\mathbb{P}$ используют запись в виде интеграла Лебега–Стилтьеса:

$$\mathbf{E}h(\xi) = \int_\mathbb{R} h(x)\, dF(x).$$

Такая запись намекает на возможность вычисления матем.ожидания с помощью техники интегрирования по Риману–Стилтьесу. Например, если ф.р. F всюду непрерывна и её производная $f(x) = F'(x)$ существует и непрерывна всюду, кроме, быть может, конечного числа точек, то

$$\mathbf{E}h(\xi) = \int_{-\infty}^{\infty} h(x) f(x)\, dx,$$

где интеграл понимается уже как несобственный интеграл Римана, и для того чтобы он был конечен, необходимо требовать его абсолютную сходимость.

Обращение к ф.р. сл.в. удобно, если нас интересуют характеристики только одной этой сл.в. Если же появляется ещё одна сл.в., то необходимо рассматривать аналогичные конструкции в пространстве $\mathbb{R}^2$ (для трёх сл.в. в $\mathbb{R}^3, \ldots,$ для последовательности сл.в. в $\mathbb{R}^\infty$). Подобное описание не только возможно, но часто и необходимо. Однако с терминологической точки зрения удобно считать, что все сл.в. заданы на некотором общем вероятностном пространстве, и любые числовые характеристики этих систем сл.в. вычислять как интеграл Лебега от некоторой суперпозиции функций:

$$\int_\Omega h(\xi_1(\omega), \ldots, \xi_k(\omega))\ \mathbf{P}(d\omega).$$

Реализации нескольких сл.в. можно понимать теперь не только как набор реализаций этих сл.в., но и как значение некоторой функции (случайного вектора) от единственной реализации ω случайного эксперимента, описываемого вероятностной моделью $(\Omega, \mathscr{F}, \mathbf{P})$.

I Многомерные случайные величины

*С*лучайный вектор удобнее всего определять 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. 137), определение измеримости функций, принимающих значения в $\mathbb{R}^k$, также может быть, в зависимости от ситуации, связано с различными классами подмножеств.

Определение. *Случайным вектором* $\vec{\xi}$ называется

a) упорядоченный набор сл.в. $\vec{\xi} = (\xi_1, \ldots, \xi_k)$ или отображение $\vec{\xi} : (\Omega, \mathscr{F}, \mathbf{P}) \mapsto (\mathbb{R}^k, \mathscr{B}^k)$ [‡] такое, что

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

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

Из леммы 130, стр. 157, и теоремы 117, стр. 138, следует

[‡] $\vec{\xi}$ задано на Ω и принимает значения в $\mathbb{R}^k$; здесь подчёркивается наличие в этих пространствах соответствующих σ-алгебр и вероятности. При отсутствии $\mathbf{P}$ нет причин называть это отображение «случайным» (отображение просто измеримо).

1| Лемма. (?!) *Условия определений* a-c) *эквивалентны между собой и, кроме того, эквивалентны каждому из следующих условий:*

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

e) *прообраз* $\forall$ *открытого множества* $B \subset \mathbb{R}^k$ *измерим:*

$$\vec{\xi}^{\,-1}(B) \in \mathscr{F}.$$

§ 1. Распределение случайных векторов

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

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

где в последнем равенстве мера $\mathbb{P}_{\vec{\xi}}$ на борелевском пространстве $(\mathbb{R}^k, \mathscr{B}^k)$ индуцирована (порождена) отображением $\vec{\xi}$. Справедливость этого равенства следует из теоремы 163, стр. 180, о замене переменных в интеграле Лебега.

Определение. Вероятностная мера $\mathbb{P}_{\vec{\xi}}$:

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

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

(‡) 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 $\vec{x} \lessdot \vec{y}$.

Так как распределение $\mathbb{P}_{\vec{\xi}}$ позволяет вычислить 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 $\mathbb{P}_{\vec{\xi}}$ на $\mathscr{B}^k$. Другими словами, задание сл.вектора эквивалентно заданию некоторого вероятностного закона (распределения) на борелевских подмножествах евклидова пространства $\mathbb{R}^k$. Обращение к единому источнику в виде пространства $(\Omega, \mathscr{F}, \mathbf{P})$ упрощает формулировку и интерпретацию свойств сл.в. 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 использовать понятие *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 $\mathbb{P}_{\vec{\xi}}\{\mathscr{X}\} = \mathbf{P}\{\, \vec{\xi} \in \mathscr{X} \,\} = 1$, часто не совсем удобно с описательной точки зрения. Поэтому мы будем, по мере надобности, изменять вид носителя, сохраняя за ним только свойство единичной вероятности.

Дискретные распределения. Проще всего задаются распределения 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

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

Функция $p(\vec{x}) = \mathbf{P}\{\, \vec{\xi} = \vec{x} \,\}$, $\vec{x} \in \mathscr{X}$, называется *функцией* *вероятностей* $\vec{\xi}$ (ф.вер.).

Запись дискретного распределения упрощается, если 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: $\mathscr{X}_{\vec{\xi}} = \mathscr{X}_{\xi_1} \times \mathscr{X}_{\xi_2}$. 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	$1/3$.

В одномерном случае, наоборот, точки с нулевой вероятностью исключаются из носителя.

По конструкции интеграла Лебега матем.ожидание любой функции от дискретного сл.вектора находится по формуле

$$\mathbf{E}h(\vec{\xi}) \overset{\text{df}}{=} \int_{\mathbb{R}^k} h(\vec{x})\, \mathbb{P}_{\vec{\xi}}(d\vec{x}) = \sum\nolimits_1^N h(\vec{x}_j)p(\vec{x}_j).$$

Если ряд сходится абсолютно, то говорят, что матем.ожидание конечно; если при $N = \infty$ ряд сходится условно, то матем.ожидание не существует. В случаях, когда для положительной части $h^+(\vec{x}) = \max\{h(\vec{x}), 0\}$

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

а матем.ожидание отрицательной части $\mathbf{E}\left[h^-(\vec{\xi})\right] < +\infty$ (или наоборот), говорят, что матем.ожидание $\mathbf{E}h(\vec{\xi})$ существует и равно $+\infty$ $(-\infty)$.

2| П р и м е р ы. Наиболее востребованные в приложениях дискретные многомерные 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 $k > 2$ 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 выборки, принадлежащих j-ой категории. Если генеральная совокупность состоит из N_j объектов j-ой категории, $N = N_1 + \ldots + N_k$ — общее число объектов в совокупности и производится выбор без возвращения n объектов, тогда функция вероятностей $\vec{\xi} = (\xi_1, \ldots, \xi_k)$, т.е. набор вероятностей того, что в выборке будет n_1 объектов 1-ой категории, $\ldots$, n_k объектов k-ой категории равна[‡]

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

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

Компоненты $\vec{\xi}$ линейно связаны: $\xi_1 + \ldots + \xi_k = n$, поэтому, в принципе, его распределение нужно описывать в пространстве меньшей размерности ($\mathbb{R}^{k-1}$), как это обычно и делается для $k = 2$, когда в эксперименте фиксируется только количество объектов одной категории, при этом к второй категории (в общем случае к k-ой) 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

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

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. $\vec{\xi}$ равна

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

для всех тех $(n_1, \ldots, n_k)$, для которых $n_1 + \ldots + n_k = n$. Косвенно справедливость этой формулы подтверждает 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 $k = 2$ модель $\mathcal{M}\mathtt{ult}_2(n; q, 1-q)$, по-существу, 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-ой 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. ◎

3| $\mathfrak{Упр}$. а) Докажите формулы для ф.вер. многомерной гипергеометрической $p_{\text{МН}}$ и мультиномиальной $p_{\text{М}}$ моделей.

б) Покажите, что для фиксированных $q_j = {}^{N_j}/_{N}$, $j = \overline{1,k}$, вероятность $p_{\text{МН}}(n_1, \ldots, n_k) \to p_{\text{М}}(n_1, \ldots, n_k)$ при $N \to \infty$ для $\forall\,(n_1, \ldots, n_k)$.

4| 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 (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 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, то можно усомниться в истинности предположений о распределении голосов избирателей или в беспристрастности (равновероятности) отбора при опросе. ⊚

Функция распределения случайного вектора. Распределение сл.вектора, как и распределение одномерной сл.величины, может быть описано посредством его функции распределения.

О п р е д е л е н и е. Функция (от $\vec{x} \in \mathbb{R}^k$)

$$F(\vec{x}) \ := \ \mathbf{P}\big\{ \, \omega : \vec{\xi}(\omega) \leqslant \vec{x} \, \big\} = \mathbf{P}\big\{ \, \xi_1 \leqslant x_1 \, , \, \dots \, , \, \xi_k \leqslant x_k \, \big\}$$

называется *функцией распределения* (ф.р.) сл.вектора $\vec{\xi}$.

Напомним, что запятая между событиями под знаком вероятности заменяет знак пересечения (союз И).

Свойства ф.р. сл.векторов повторяют в основном свойства ф.р. сл.в. Нам понадобится разностный оператор $\Delta_{\vec{a},\vec{c}}$, связанный с параллелепипедом $(\vec{a};\vec{c}] = (a_1;c_1] \times \ldots \times (a_k;c_k]$ (см. (5), стр. 152).

5| Т е о р е м а. *Пусть* $F(\vec{x})$ *ф.р. сл.вектора* $\vec{\xi}$. *Тогда:*

(F_1) F *не убывает:* $F(\vec{x}) \leqslant F(\vec{y})$, *если* $\vec{x} < \vec{y}$;

(F_2) $\mathbf{P}\big\{ \, \vec{\xi} \in (\vec{a};\vec{c}] \, \big\} \ = \ \Delta_{\vec{a},\vec{c}} F(\vec{x}), \quad \vec{a} \leqslant \vec{c}$;

(F_3) F *непрерывна справа по совокупности переменных:*

$$\lim_{\vec{\epsilon} \to +0} F(\vec{x} + \vec{\epsilon}) = F(\vec{x}) \ ^{(\ddagger)};$$

(F_4) $\lim\limits_{\vec{x} \to \vec{y}} F(\vec{x}) = 1$, *если все координаты* $\vec{y}$ *равны* $+\infty$,

$\lim\limits_{\vec{x} \to \vec{y}} F(\vec{x}) = 0$, *если хотя бы одна координата* $y_j = -\infty$:

$$F(+\infty, \ldots, +\infty) = 1, \qquad F(y_1, \ldots, -\infty, \ldots, y_k) = 0.$$

⊫ Аналитически свойство (F_2) (к примеру, в трехмерном случае) можно доказать так: искомая вероятность

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

где $G_1(x_1) \qquad = \mathbf{P}\{\xi_1 \leqslant x_1, \ a_2 < \xi_2 \leqslant c_2, \ a_3 < \xi_3 \leqslant c_3\} =$

$$= \Delta^{(2)}_{a_2,c_2} G_2(x_1, x_2),$$

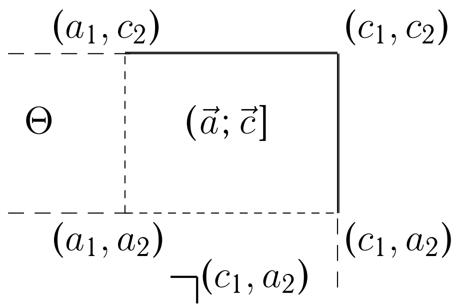
где $G_2(x_1, x_2) \ = \mathbf{P}\{\xi_1 \leqslant x_1, \ \xi_2 \leqslant x_2, \ a_3 < \xi_3 \leqslant c_3\} =$

(‡) $\vec{\epsilon} \to +0 \ \Leftrightarrow \ \vec{\epsilon} \to \vec{0}, \ \vec{\epsilon} \geqslant \vec{0}$.

$$= \Delta^{(3)}_{a_3,c_3} F(x_1, x_2, x_3).$$

В двумерном случае доказательство свойства (F_2) можно представить геометрически. Пусть $\neg^{(u,v)} = \{(x,y) : x \leqslant u, y \leqslant v\}$ — «угловая» область на плоскости $\mathbb{R}^2$. Тогда

$$\mathbf{P}\{ (\xi_1, \xi_2) \in \neg^{(u,v)} \} = F(u, v).$$



Как видно из рисунка, прямоугольник

$$(\vec{a}; \vec{c}] = \neg^{(c_1, c_2)} - \neg^{(c_1, a_2)} - \Theta,$$

где, в свою очередь, область

$$\Theta = \neg^{(a_1, c_2)} - \neg^{(a_1, a_2)}.$$

Таким образом, ввиду вложенности соответствующих областей

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

Доказательство остальных свойств аналогично $\mathbb{R}^1$. ⟸

В многомерном случае, в отличие от одномерного, условий непрерывности справа (F_3), монотонности (F_1) и нормированности (F_4) недостаточно, чтобы функция F определяла некоторую вероятностную меру. Следующая теорема является прямым следствием утверждения, приведённого в главе III, стр. 152, о существовании в пространстве $\mathbb{R}^k$ меры Лебега–Стилтьеса.

6| Т е о р е м а. *Если функция* $F : \mathbb{R}^k \mapsto [0; 1]$ *обладает свойствами* (F_3), (F_4) *и для* $\forall\, \vec{a} < \vec{c}$ $(\in \mathbb{R}^k)$ *оператор* $\Delta_{\vec{a},\vec{c}} F \geqslant 0$, *тогда на борелевском пространстве* $(\mathbb{R}^k, \mathscr{B}^k)$ *существует единственная вероятностная мера* $\mathbf{P}$ *такая, что* $\forall\, \vec{a} < \vec{c}$ $(\in \mathbb{R}^k)$

$$\mathbf{P}\{ \vec{x} \in \mathbb{R}^k : \vec{x} \leqslant \vec{c} \} = F(\vec{c}), \qquad \mathbf{P}\{ (\vec{a}, \vec{c}] \} = \Delta_{\vec{a},\vec{c}} F. \tag{2}$$

Теорема Каратеодори, гарантирующая существование меры Лебега–Стилтьеса, устанавливает факт существования меры, удовлетворяющей только второму равенству (2).

7| Л е м м а. (?!) *Мера* **P**, *построенная в теореме* 6, *действительно есть вероятность с ф.р.* F, *т.е.*

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

b) *для неё справедливы оба соотношения* (2).

В частности, функция F *удовлетворяет свойству* (F_1).

8| $\mathfrak{Y}\mathfrak{np}$. Функция $G(x_1, x_2) = \mathbf{I}(x_1 + x_2 > 0)$ удовлетворяет свойствам (F_1), (F_3), (F_4), но не может быть ф.р.

9| П р и м е р ы. 1) Если $F = F_1(x_1)F_2(x_2)$, где F_1, F_2 — некоторые одномерные ф.р., то оператор $\Delta_{\vec{a},\vec{c}}F = (F_1(c_1) - F_1(a_1))(F_2(c_2) - F_2(a_2))$. В частности, если все ф.р. $F_i(x) = x$, $0 \leqslant x \leqslant 1$, то ф.р. $F_U(x_1, x_2) = x_1 x_2$ определяет сл.вектор $\vec{\xi} \overset{d}{\backsim} \mathcal{U}\mathrm{n}([0;1]^2)$ 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 прямой $y = x$.

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

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

$$\mathbf{E}h(\vec{\xi}\,) \overset{\mathrm{df}}{=} \int_{\mathbb{R}^k} h(\vec{x})\,\mathbb{P}_{\vec{\xi}}(d\vec{x}) =: \int_{\mathbb{R}^k} h(\vec{x})\,dF_{\vec{\xi}}(\vec{x}). \tag{3}$$

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. 190, стр. 203), средства вычисления которого разработаны в анализе. Для этого, правда, каждый раз приходится делать оговорку об абсолютной интегрируемости функции $h(\vec{x})$, если матем.ожидание конечно:

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

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. 210), 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 вектора задаётся посредством ф.р. $F_{\vec{\xi}}$, вероятности любых событий $B \subset \mathbb{R}^k$ можно найти по формуле

$$\mathbf{P}\big\{\, \vec{\xi} \in B \,\big\} \;=\; \int_{\mathbb{R}^k} \mathbf{I}_B(\vec{x})\, dF_{\vec{\xi}}(\vec{x}) \;=:\; \int_B dF_{\vec{\xi}}(\vec{x}). \tag{4}$$

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

$$F_{\vec{\xi}}(\vec{x}) \;\stackrel{\text{df}}{=}\; \mathbf{P}\big\{\, \vec{\xi} \leqslant \vec{x} \,\big\} \;=\; \int_\Omega \mathbf{I}(\omega; \vec{\xi} \leqslant \vec{x})\; \mathbf{P}(d\omega).$$

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 $F_{\vec{\xi}}(x_1, \ldots, x_k)$ называется *совместной* функцией распределения $\vec{\xi}$.

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 $\vec{\xi}$ называется *маргинальной (частной)* функцией распределения.

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

Слово Marginal переводится как «предельный, крайний». Следующая лемма даёт 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.

10| Лемма. (?!) *Если* $(\xi_1,\ldots,\xi_k)\stackrel{d}{\sim}F(x_1,\ldots,x_k)$, *то*

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

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

Другое объяснение термина «маргинальный» восходит к естественному способу отыскания одномерных распределений по таблице двумерного дискретного распределения.

$\xi_2\backslash^{\xi_1}$	-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	

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. 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

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

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 ξ_1 (?!)

$$p_{\xi_1}(x_i) = \sum\nolimits_{j=1}^{n_2} p(x_i, y_j), \qquad x_i \in \mathscr{X}_1. \tag{5}$$

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

11| Примеры. 1) Легко понять из определения, что маргинальное распределение многомерной гипергеометрической модели (1), стр. 17, снова будет иметь тот же тип:

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

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 $n_j^* = n - (n_1 + \ldots + n_{j-1})$

$$\begin{aligned} &\mathbf{P}\{\xi_1 = n_1, \ldots, \xi_{j-1} = n_{j-1}, \xi_j^* = n_j^*\} = \\ &\qquad = \sum_{n_j + \ldots + n_k = n_j^*} p_{\mathrm{MH}}(n_1, \ldots, n_j, \ldots, n_k) = \\ &= \Big(\prod_{i=1}^{j-1} \mathrm{C}_{N_i}^{n_i}\Big) \frac{\mathrm{C}_{N_j^*}^{n_j^*}}{\mathrm{C}_N^n} \left[\sum_{n_j + \ldots + n_k = n_j^*} \Big(\prod_{l=j}^{k} \mathrm{C}_{N_l}^{n_l}\Big) \frac{1}{\mathrm{C}_{N_j^*}^{n_j^*}} \right] = \\ &\qquad = \Big(\prod_{i=1}^{j-1} \mathrm{C}_{N_i}^{n_i}\Big) \mathrm{C}_{N_j^*}^{n_j^*} \frac{1}{\mathrm{C}_N^n}, \end{aligned}$$

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 $(\xi_1, n-\xi_1) \overset{d}{\backsim} \mathcal{H}\mathrm{g}_2(n; N_1, N-N_1)$, что эквивалентно гипергеометрическому распределению 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

$$
\begin{aligned}
(\xi_1, \ldots, \xi_k) \quad &\overset{d}{\backsim} \quad \mathcal{M}\mathtt{ult}_k(n; q_1, \ldots, q_k) \qquad \Rightarrow \\
(\xi_1, \ldots, \xi_{j-1}, \xi_j^*) \quad &\overset{d}{\backsim} \quad \mathcal{M}\mathtt{ult}_j(n; q_1, \ldots, q_{j-1}, q_j^*)\,.
\end{aligned}
$$

В частности, $(\xi_1, n-\xi_1) \overset{d}{\backsim} \mathcal{M}\mathtt{ult}_2(n; q_1, 1-q_2)$, что эквивалентно биномиальной модели: $\xi_1 \overset{d}{\backsim} \mathcal{B}\mathtt{in}(n; q_1)$ $(\xi_2 = n - \xi_1)$. ◎

О точках непрерывности функции распределения. Напомним, что одномерная ф.р. терпит разрыв в точке x т. т. т. когда эта точка имеет положительную вероятность: $\mathbf{P}\{\xi = x\} = F(x+) - F(x-) > 0$. Поскольку количество точек x, для которых вероятность $\mathbf{P}\{\xi = x\} \geqslant {}^1/_n$, не более n, то для одномерной ф.р. имеет место

12| Лемма. (?!) *Множество точек разрыва любой одномерной функции распределения разве лишь 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.

13 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:

$$\begin{array}{ll} a)\ x = 0,\ y \geqslant 0, & b)\ x = 1,\ y \geqslant 0, \\ c)\ y = 0,\ x \geqslant 0, & d)\ y = 1,\ x \geqslant 0. \end{array}$$

Точка $(0, -1)$ — точка непрерывности, хотя маргинальное распределение ξ терпит разрыв при $x = 0$. С другой стороны, скажем, точка $(0, 2)$ будет точкой 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, пересекающих 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 $\min\{x, y\} \leqslant 0$. Эта функция непрерывна в точке $(0, 0)$, хотя обе частные ф.р. $F_\xi(t) = F_\eta(t) = \big(\frac{1}{2}t + {}$ $+ \frac{1}{4}\big)\mathbf{I}_{[0;1)}(t) + \mathbf{I}_{[1;\infty)}(t)$ терпят разрыв в нуле.

3) Пример 28, стр. 41, демонстрирует ещё один аспект отличий одномерного и многомерного случаев. В этом примере, непрерывная многомерная ф.р. имеет целую линию сосредоточения — вероятность осуществления казалось бы невозможного при непрерывной ф.р. события $\xi = \eta$ равна $^1/_2$. ◎

14 Лемма. (?!) *Пусть* $F(x_1, \ldots, x_k)$ *— ф.р. сл.вектора* $\vec{\xi}$, $F_j(x_j)$, $j = \overline{1, k}$, *— 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 «углового» множества* $\{\vec{x} : \vec{x} \leqslant \vec{x}_0\}$ *равна нулю. В частности, если* $\mathbf{P}\{\vec{\xi} = \vec{x}_0\} > 0$, *то в точке* $\vec{x}_0$ *ф.р.* $\vec{\xi}$ *терпит разрыв.*

II.a) *Ф.р.* F *непрерывна в точке* $(x_{01}, \ldots, x_{0k})$, *если каждая 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. Справедлива

15 Лемма. *Множество* $A \subset \mathbb{R}^k$, *в котором любые два вектора* $\vec{x}, \vec{y} \in A$ *имеют различные координаты:* $x_j \neq y_j$, $\forall\, j = \overline{1, k}$, *содержит не более чем счётное число точек разрыва функции распределения* F.

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

⊫ Пусть A — несчётное подмножество D 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 14 (пункт II.a) гарантирует непрерывность ф.р. в этих точках.

Во втором утверждении важна только выпуклость параллелепипеда и то, что при $\vec{a} \lessdot \vec{c}$ найдутся два вектора $\vec{x}, \vec{y}$ из параллелепипеда, координаты которых различаются. Поэтому отрезок $\lambda\vec{x} + (1-\lambda)\vec{y},\ \lambda \in [0;1]$, соединяющий эти два вектора, полностью лежит внутри параллелепипеда и каждая пара точек этого отрезка имеет различные координаты. В силу первого утверждения леммы, на этом отрезке имеется несчётное число точек непрерывности F. ⫤

Из лемм 12, 15 следует важная для дальнейшего

16 Теорема. *Если ф.р. $F_1(\vec{x}) = F_2(\vec{x})$ во всех $\vec{x}$, являющихся точками непрерывности F_1 и F_2, 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 (по условию $F_1(\vec{x}_n) = F_2(\vec{x}_n)$). Отсюда, ввиду непрерывности справа ф.р., следует, что и $F_1(\vec{x}) = F_2(\vec{x})$. ⫤

Обобщённая функция распределения. 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 6 это свойство обеспечивают условия $(\mathtt{F}_4)$. 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

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

$x \geqslant 0$. Этой мере соответствует ф.р. $(G(x_1) + 1)(G(x_2) + 1)$, которая не может быть получена с помощью линейного преобразования F.

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

Задание распределения через копулу. Если ф.р. F одномерной сл.в. ξ непрерывна, то сл.в. $F(\xi) \stackrel{d}{\backsim} \mathcal{U}\mathrm{n}([0;1])$. Это соотношение основывается на том, что $F(F^{-1}(x)) = x$, $x \in \in [0;1]$. В многомерном случае в принципе нельзя определить обратную ф.р. Однако очевидно, что если каждая одномерная маргинальная ф.р. непрерывна, то ф.р. $\mathcal{C}(y_1, \ldots, y_k)$ сл.вектора $(\eta_1, \ldots, \eta_k) = (F_{\xi_1}(\xi_1), \ldots, F_{\xi_k}(\xi_k)) \in [0;1]^k$, имеет равномерные на $[0;1]$ одномерные маргинальные распределения, а ф.р. сл.вектора $\vec{\xi}$ равна

$$F(\vec{x}) = \mathcal{C}(F_{\xi_1}(x_1), \ldots, F_{\xi_k}(x_k)). \tag{6}$$

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

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

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

Теорема Sklar'а утверждает, что функция $F(\vec{x})$ есть многомерная ф.р. т. т. т. когда она может быть представлена в виде (6) 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.

17| Примеры. 1) Ф.р. равномерного распределения на кубе задаёт копулу с независимыми компонентами.

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

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

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 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. (сл.векторы) заданы на одном вероятностном пространстве. В дальнейшем это условие подразумевается без специальной оговорки.

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

$$\mathbf{P}\{\xi_1 \in B_1, \ldots, \xi_k \in B_k\} = \prod_1^k \mathbf{P}\{\xi_j \in B_j\}$$

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

Сл.векторы $\vec{\xi}_j \in \mathbb{R}^{k_j},\ j = \overline{1, m}$, *независимы в совокупности*, если $\mathbf{P}\{\,\vec{\xi}_1 \in B_1, \ldots, \vec{\xi}_m \in B_m\,\} = \prod_1^m \mathbf{P}\Big\{\,\vec{\xi}_j \in B_j\Big\}$ для любых борелевских 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 сл.вектора, у которого любой подвектор (не обязательно двумерный) будет состоять из независимых в совокупности сл.в., однако весь вектор не будет удовлетворять условию совместной независимости.

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

$$F_{\vec{\xi}}(\vec{x}) = \prod\nolimits_1^k \mathbf{P}\{\xi_j \leqslant x_j\} = \prod\nolimits_1^k F_{\xi_j}(x_j). \tag{7}$$

Справедливо и обратное утверждение.

18 Т е о р е м а. *Компоненты вектора* $\vec{\xi} = (\xi_1, \ldots, \xi_k)$ *независимы в совокупности т. т. т. когда совместная ф.р.* $F_{\vec{\xi}}$ *удовлетворяет* (7), *т.е. равна произведению одномерных частных ф.р.*

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

$$\mathbf{P}\left\{\xi_1 \in B \,|\, \xi_2 \leqslant x_2\right\} =: \mu_1\{B\} \tag{8}$$

задаёт (?!) вероятностную меру на $B \in \mathscr{B}$. В силу (7) 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 совпадают и определяемые ими вероятностные меры: для $\forall\, B \in \mathscr{B}$ $\mu_1\{B\} = \mathbf{P}\left\{\xi_1 \in\right.$

$\in B \,|\, \xi_2 \leqslant x_2\} = \mathbf{P}\{\xi_1 \in B\}$, поэтому $\mathbf{P}\{\xi_1 \in B, \xi_2 \leqslant x_2\} = \mathbf{P}\{\xi_1 \in B\}\,\mathbf{P}\{\xi_2 \leqslant x_2\}$. 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 $\mu_2\{A\} := \mathbf{P}\left\{\xi_2 \in A \,|\, \xi_1 \in B_1\right\}$ с фиксированным множеством B_1, получим доказательство теоремы. $\Lleftarrow$

Аналогичное справедливо и для независимых сл.векторов.

19| 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:*

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

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.

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

21| Пример. При сравнении теоретических (истинных) дисперсий по критерию Фишера, основанному на двух независимых выборках $(\xi_1, \ldots, \xi_n)$ и $(\eta_1, \ldots, \eta_m)$ из нормального распределения с параметрами (μ_1, σ_1^2) и (μ_2, σ_2^2) соответственно, существенно используется тот факт, что выборочные дисперсии $n^{-1} \sum_{i=1}^{n} (\xi_i - \sum_{j=1}^{n} \xi_j / n)^2$ и $m^{-1} \sum_{i=1}^{m} (\eta_i - \sum_{j=1}^{m} \eta_j / m)^2$ независимы. ⊚

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

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 $A_j \in \mathscr{L}_j$, $j = \overline{1, n}$.

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

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 $\sigma(\mathcal{F})$, порождённой $\mathcal{F}$. Способ доказательства этой теоремы сродни способу доказательства теоремы 18.

23| Теорема. *Для независимости σ-алгебр $\mathscr{L}_1 = \sigma(\mathcal{F}_1), \ldots, \mathscr{L}_n = \sigma(\mathcal{F}_n)$, порождённых классами подмножеств $\mathcal{F}_1, \ldots, \mathcal{F}_n$ σ-алгебры $\mathscr{F}$, достаточно, чтобы классы*

$\mathcal{F}_1, \ldots, \mathcal{F}_n$ *были независимы и определяли меру.*

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

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

По условию теоремы эти 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. Аналогично по индукции показывается независимость σ-алгебр $\mathscr{L}_1, \ldots, \mathscr{L}_n$. $\Lleftarrow$

Утверждения теорем 18 и 19 являются, по-существу, следствиями теоремы 23. Так, если рассмотреть полукольцо $\mathcal{B}$ всех интервалов вида $\big\langle (a; b],\ -\infty < a < b < \infty \big\rangle$, то по условию теоремы 18 независимы (?!) полукольца $\xi_1^{-1}(\mathcal{B}), \ldots, \xi_k^{-1}(\mathcal{B})$, индуцированные отображениями $\xi_1, \ldots, \xi_k$. Согласно теореме Каратеодори эти классы множеств определяют меру. В силу леммы 130, стр. 157, прообразы борелевской σ-алгебры $\xi_j^{-1}(\mathscr{B}) = \sigma(\xi_j^{-1}(\mathcal{B})),\ j = 1, \ldots k$. По теореме 23 σ-алгебры $\sigma(\xi_j^{-1}(\mathcal{B})),\ j = 1, \ldots k,$ независимы. Независимость сл.в. $\xi_1, \ldots, \xi_k$ следует теперь из леммы 22 .

Следующий факт также вытекает из теоремы 23 . Для краткости теорема сформулирована только для четырёх сл.в., хотя вполне очевидно, что она справедлива для любого количества сл.в. и даже для сл.векторов. Доказательство теоремы почти дословно повторяет рассуждения предыдущего абзаца.

24| Т е о р е м а. *Пусть сл.в.* $\xi_1, \xi_2, \xi_3, \xi_4$ *независимы в совокупности, тогда независимы сл.векторы* (ξ_1, ξ_2) *и* (ξ_3, ξ_4). *Следовательно, 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* $\vec{h}_2(\xi_3, \xi_4)$.

Математическое ожидание произведения сл.в. Теорема 18 вместе с теоремой Фубини (стр. 210) позволяют упростить доказательство свойства матем.ожидания произведения функций независимых сл.векторов.

25| Теорема. *Если векторы* $\left\langle\, \vec{\xi}_j,\ j=\overline{1,J}\, \right\rangle$ *независимы в совокупности и матем.ожидания* $\mathbf{E}\big|h_j(\vec{\xi}_j)\big| < \infty,\ \ j=\overline{1,J}$, *то*

$$\mathbf{E}\Big[\prod\nolimits_1^J h_j(\vec{\xi}_j)\Big] \ = \ \prod\nolimits_1^J \mathbf{E}\Big[h_j(\vec{\xi}_j)\Big].$$

$\Rrightarrow$ Пусть вектор $\vec{\xi} \overset{d}{\backsim} F_1(\vec{x}),\ \vec{x} \in \mathbb{R}^k$, и вектор $\vec{\eta} \overset{d}{\backsim} F_2(\vec{y}),\ \vec{y} \in \mathbb{R}^m$. Очевидно, функция $F(\vec{x},\vec{y}) = F_1(\vec{x})F_2(\vec{y})$ задаёт ф.р. $(\vec{\xi},\vec{\eta})$ в пространстве $\mathbb{R}^{k+m}$, причём соответствующая ей мера Лебега–Стилтьеса есть прямое произведение мер $\mu_1 \times \mu_2$, порождённых F_1, F_2. По условию теоремы

$$\int_{\mathbb{R}^k}\bigg(\int_{\mathbb{R}^m} |h_1(\vec{x})h_2(\vec{y})|\, dF_2(\vec{y})\bigg)\, dF_1(\vec{x}) =$$

$$= \int_{\mathbb{R}^k} |h_1(\vec{x})|\, dF_1(\vec{x}) \int_{\mathbb{R}^m} |h_2(\vec{y})|\, dF_2(\vec{y}) < \infty.$$

Отсюда, применяя теорему Фубини 199, стр. 211, получаем

$$\mathbf{E}\big[h_1(\,\vec{\xi}\,)h_2(\vec{\eta})\big] = \int\limits_{\mathbb{R}^{k+m}} h_1(\vec{x})h_2(\vec{y})\, dF(\vec{x},\vec{y}) =$$

$$= \int\limits_{\mathbb{R}^k} h_1(\vec{x})\, dF_1(\vec{x}) \int\limits_{\mathbb{R}^m} h_2(\vec{y})\, dF_2(\vec{y}) = \mathbf{E}\big[h_1(\,\vec{\xi}\,)\big]\, \mathbf{E}\big[h_2(\vec{\eta})\big].$$

Общий случай $J > 2$ доказывается по индукции. $\Lleftarrow$

§ 3. Абсолютно-непрерывные распределения

Если распределение $\mathbb{P}_{\vec{\xi}}$ абсолютно-непрерывно относительно меры Лебега (см. определение на стр. 187; допуская некоторую вольность речи, говорят просто абсолютно-непрерывно), то по теореме Радона–Никодима (174 , стр. 188)

$$\mathbf{P}\{\,\vec{\xi}\in B\,\} \;=\; \int\limits_B f(\vec{x})\,d\vec{x}, \qquad B\in\mathscr{B}^k, \tag{9}$$

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

$$(\checkmark)\;\; f(\vec{x})\geqslant 0, \quad \forall\,\vec{x}\in\mathbb{R}^k, \qquad\qquad (\checkmark)\;\; \int\limits_{\mathbb{R}^k} f(\vec{x})\,d\vec{x} \;=\; 1. \tag{10}$$

Часто плотность задают только на носителе $\mathscr{X} = \{\vec{x} : f(\vec{x}) > 0\}$ в виде $f = f(\vec{x})$, $\vec{x}\in\mathscr{X}$, подразумевая, что $f = 0$ при $\vec{x}\notin\mathscr{X}$. Если сл.вектор $\vec{\xi} \overset{d}{\backsim} f(\vec{x})$, $\vec{x}\in\mathscr{X}$, то матем.ожидание любой измеримой функции $h(\vec{\xi}\,)$ может быть найдено по формуле (10), стр. 190:

$$\mathbf{E}h(\vec{\xi}\,) \;=\; \int\limits_{\mathscr{X}} h(\vec{x})f(\vec{x})\,d\vec{x} \;=\; \int\limits_{\mathbb{R}^k} \mathbf{I}_{\mathscr{X}}(\vec{x})h(\vec{x})f(\vec{x})\,d\vec{x}. \tag{11}$$

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

Условие абсолютной непрерывности можно задать через представление для ф.р. Определим «правую» производную[(‡)]

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

положив $F^{(k)}(\vec{x}) := 0$ там, где предел равен $+\infty$. Эта производная измерима по Борелю, как предел измеримых функций. 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. [9], т. 1, стр. 244-245).

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

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

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

$$F(\vec{x}) = \int\limits_{\vec{u}\leqslant\vec{x}} f(\vec{u})\,d\vec{u} = \int\limits_{\mathbb{R}^k} \mathbf{I}_{\{\vec{u}\leqslant\vec{x}\}}(\vec{u}) f(\vec{u})\,d\vec{u}. \tag{12}$$

II) *Если $\int_{\mathbb{R}^k} F^{(k)}(\vec{x})\,d\vec{x} \geqslant 1$, то мера $\mathbb{P}$ 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 (10), то она есть плотность распределения с ф.р., определяемой (12).*

⊫ I) В одну сторону ($\Rightarrow$) следует из теоремы Радона–Никодима 174, стр. 188. Обратно, пусть для ф.р. справедливо представление (12). 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 $\mathbb{P}$ на множествах вида $B = \{\vec{u} : \vec{u} \leqslant \vec{x}\}$. Другими словами, ф.р. этих мер совпадают, а следовательно, совпадают и сами меры.

II) Как и выше, 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

$$\mathbb{P}\{(\vec{a}; \vec{c}]\} = \Delta_{\vec{a};\vec{c}} F \geqslant \mathsf{M}(\vec{a}; \vec{c}].$$

Заметим, что это гарантирует конечность меры M.

Так как ф.р. монотонна, то

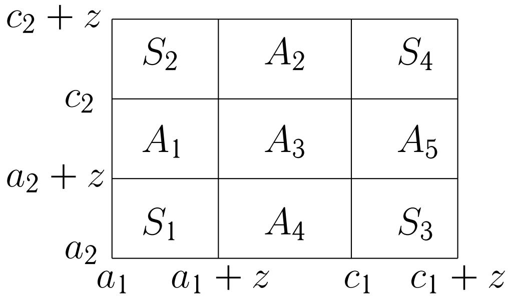
$$F(\vec{x}) \leqslant \frac{1}{z^k} \int_{(\vec{x},\vec{x}+z]} F(\vec{u})\, d\vec{u} \leqslant F(\vec{x}+z)$$

при $\forall\, \vec{x} \in \mathbb{R}^k$, $z > 0$. В силу непрерывности F отсюда получаем, что $1/z^k \int_{(\vec{x},\vec{x}+z]} F(\vec{u})\, d\vec{u} \to F(\vec{x})$ при $z \to +0$. В целях сокращения записи дальнейшие построения будем производить в двумерном случае. Таким образом, имеем

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

с областями (см. рисунок)

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



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

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

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

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

Θ_1. Например,

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

Следовательно, в силу леммы Фату 158, стр. 177,

$$\mathbb{P}\{(\vec{a}; \vec{c}]\} = \lim_{z\to+0} \int_{\Theta_1} \frac{\Delta_{\vec{u};\vec{u}+z}F}{z^k}\, d\vec{u} \;\geqslant\; \int_{\Theta_1} \varliminf_{z\to+0} \frac{\Delta_{\vec{u};\vec{u}+z}F}{z^k}\, d\vec{u} \;=$$
$$= \int_{\Theta_1} F^{(k)}(\vec{u})\, d\vec{u} = \mathsf{M}(\vec{a}; \vec{c}].$$

Ввиду непрерывности мер $\mathbb{P}, \mathsf{M}$ 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

$$\big[(-\infty; a_1]\uplus(a_1; c_1] \uplus (c_1; \infty)\big]\times\big[(-\infty; a_2] \uplus (a_2; c_2] \uplus (c_2; \infty)\big]$$
$$= \big[(-\infty; a_1]\times(-\infty; a_2]\big] \;\uplus\; \big[(-\infty; a_1]\times(a_2; c_2]\big] \;\uplus\; \ldots$$
$$\uplus\; \big[(c_1; \infty]\times(c_2; \infty]\big]$$

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

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

Следовательно, для любого конечного параллелепипеда $(\vec{a};\vec{c}]$ мера $\mathsf{M}(\vec{a};\vec{c}] = \mathbb{P}\{(\vec{a};\vec{c}\,]\}$. Отсюда по теореме Каратеодори 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. Свойства (F_1), (F_3), (F_4) (см. 6, стр. 21) следуют из монотонности интеграла и теоремы Лебега о предельном переходе под знаком интеграла. Например, если $\vec{x}_n \to \vec{y}$ и одна из координат $\vec{y}$ равна $-\infty$, то функция $\mathbf{I}(\vec{u};\vec{u} \leqslant \vec{x}_n) \to 0$. Поэтому $F(\vec{x}_n) \to 0$. Также показывается, что F непрерывна всюду. Свойство (F_2) «монотонности» функции следует из тождества следующей леммы, в силу которого по свойству линейности интеграла

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

27| Лемма. (?!) *Для индикатора* $\mathbf{I}(\vec{u};\vec{u} \leqslant \vec{x})$, *как функции* $\vec{x}$ *при фиксированном* $\vec{u}$, *имеет место равенство* $\Delta_{\vec{a},\vec{c}}\,\mathbf{I}(\vec{u};\vec{u} \leqslant \vec{x}) = \mathbf{I}(\vec{u};\vec{u} \in (\vec{a};\vec{c}\,])$.

28| Пример. 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 $y = x$ $(x \leqslant y)$, а три другие — ниже диагонали $(x \geqslant y)$, имеем

$$\Delta_A F = F(b_1,b_2) - F(b_1,a_2) - F(a_1,b_2) + F(a_1,a_2) = G^2(b_2) - G^2(a_2) - G(a_1)G(b_2) + G^2(a_2) = G(b_2)(G(b_2) - G(a_1)) \geqslant 0.$$

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: $\mathbf{P}\{A\} = \Delta_A F = G^2(b_2) - G^2(a_2) - G^2(b_2) + G^2(a_2) = 0$. 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 $\mathcal{U}\mathtt{n}(0,1)$ закона. Ясно, что ф.р. $F(x,y)$ непрерывна всюду и дифференцируема всюду, кроме границ единичного квадрата и его диагонали. Если всё же положить производную $F''(x,y) = 1 \cdot \mathbf{I}(0 < x < y < 1)$, то $\int_{\mathbb{R}^2} F''(x,y)\,dxdy = {}^1\!/_2 < 1$. Условие предыдущей теоремы не выполнено. Поэтому мера Лебега–Стилтьеса μ_F, генерируемая ф.р. F, не является абсолютно-непрерывной.

Из выражения для ф.р. F не сразу видно, куда пропала вероятностная масса ${}^1\!/_2$ в результате вычисления производной. На самом деле это распределение задаёт меру μ_F, значение которой на диагональном множестве $y = x$ отлично от нуля, а мера Лебега этой области равна нулю. Зафиксируем $N > 0$ и накроем диагональ квадратами $A_k = (a_{k-1}, a_k] \times (a_{k-1}, a_k]$, где $a_k = {}^k\!/_{2^N}$, $k = 1, \ldots, 2^N$. Мера $\mu_F(A_k) = {}^k\!/_{2^{2N}}$, поэтому

$$\mu_F(y = x) = \lim_N \sum\nolimits_{k=1}^{2^N} \mu_F(A_k) = \lim_N \frac{2^N(2^N+1)}{2 \cdot 2^{2N}} = \frac{1}{2}.$$

В действительности этот факт справедлив для любой непрерывной ф.р. G. Легко понять, что ф.р. F есть ф.р. вектора $(\xi, \max\{\xi,\eta\})$, где ξ, η — независимые с.в. с общей ф.р. G. Поэтому мера μ_F тех точек, где $y = x$, равна

$$\mu_F(y = x) = \mathbf{P}\{\xi = \max\{\xi,\eta\}\} = \mathbf{P}\{\xi \geqslant \eta\}\,. \tag{13}$$

Для независимых одинаково распределённых сл.в. с непрерывной ф.р. эта вероятность равна (?!) $^1\!/_2$. ⊚

Частная плотность. Если вектор $\vec{\xi}$ имеет абсолютно-непрерывное распределение, то это справедливо и для любых частных распределений. Понятие абсолютной непрерывности связано с мерой Лебега в пространстве значений сл.вектора. Поскольку, строго говоря, в пространствах $\mathbb{R}^1, \mathbb{R}^2, \ldots$ это разные меры (длина, площадь, …), то, переходя к маргинальному распределению, необходимо указывать соответствующую меру Лебега.

29| Теорема. *Если вектор* $\vec{\xi} = (\xi_1, \ldots, \xi_k)$ *имеет абсолютно-непрерывное распределение относительно меры Лебега в* $\mathbb{R}^k$ *с плотностью* $f(\vec{x})$, *то сл.вектор* (ξ_1, ξ_2) *и сл.величина* ξ_1 *также имеют абсолютно-непрерывное распределение относительно мер Лебега в* $\mathbb{R}^2$ *и* $\mathbb{R}^1$ *соответственно с плотностями*

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

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

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

⊫ Пусть $B \subset \mathbb{R}^2$ — 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 $\vec{\xi}$

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

Индикаторная функция $\mathbf{I}(\vec{x}; B \times \mathbb{R}^{k-2}) = \mathbf{I}((x_1, x_2); B)$ для $\forall\, \vec{x} \in \mathbb{R}^k$. Отсюда по теореме Фубини получаем

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

Второе утверждение теоремы доказывается аналогично. ⫷

Формулы плотности других подвекторов имеют схожую структуру — совместная плотность интегрируется по всем переменным, не относящимся к этим подвекторам.

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

30| Т е о р е м а. I) *Если сл.в.* $\xi_1, \ldots, \xi_k$ *независимы в совокупности и распределение каждой из сл.в.* ξ_j *абсолютно-непрерывно с плотностью* $f_j(x_j)$, *то распределение сл.вектора* $\vec{\xi} = (\xi_1, \ldots, \xi_k)$ *имеет плотность*

$$f_{\vec{\xi}}(\vec{x}) \;=\; \prod_{j=1}^{k} f_j(x_j), \qquad \vec{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k. \tag{14}$$

II) *Если плотность распределения сл.вектора* $\vec{\xi}$ *может быть представлена в виде* (14), *то сл.в.* $\xi_1, \ldots, \xi_k$ *независимы в совокупности и распределение* ξ_j *имеет плотность* f_j.

⫸ Схема доказательства теоремы весьма проста. Например, для I) необходимо: а) учтя независимость, 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 (12). ⫷

31| П р и м е р. Легко видеть, что для сл.вектора $\vec{\xi} = (\xi_1, \xi_2)$ 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 $A = (a_1; c_1] \times (a_2; c_2]$, 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

$$\mathbf{P}\{\, \vec{\xi} \in A \,\} = F(c_1, c_2) - F(c_1, a_2) - F(a_1, c_2) + F(a_1, a_2) = 0.$$

Иначе говоря, вектор (ξ_1, ξ_2) сосредоточен на диагонали. Распределение (ξ_1, ξ_2) не абсолютно-непрерывно, т.к. вероятность попадания на диагональ $\mathbf{P}\{\,\xi_1 = \xi_2\,\} = 1$, в то время как плоская мера Лебега диагонали равна нулю. Однако оба маргинальных распределения вектора $\vec{\xi}$ совпадают с абсолютно-непрерывным равномерным распределением $\mathcal{U}\mathtt{n}(0,1)$ (например, $F_1(x_1) = F(x_1, 1) = x_1$). ⊚

32| Примеры. 1) Сл.вектор $\vec{\xi}$ имеет *равномерное* распределение в области $\mathscr{X} \subset \mathbb{R}^k$: $\vec{\xi} \overset{d}{\backsim} \mathcal{U}\mathtt{n}(\mathscr{X})$, если вероятность попадания вектора в любое подмножество B этой области пропорциональна объёму B :

$$\mathbf{P}\{\,\vec{\xi} \in B\,\} = \frac{V(B \cap \mathscr{X})}{V(\mathscr{X})} = \frac{1}{V(\mathscr{X})}\int\limits_B \mathbf{I}_{\mathscr{X}}(\vec{x})\,d\vec{x},\ B \in \mathscr{B}^k, \tag{15}$$

где объём (мера Лебега) $V(\mathscr{X}) = \int_{\mathscr{X}} d\vec{x} < \infty$. В силу (9) плотность равномерного распределения можно положить равной $f(\vec{x}) = {}^1/_{V(\mathscr{X})}$, $\vec{x} \in \mathscr{X}$.

Иногда при описании равномерного распределения приходится учитывать, что этот вектор всегда принадлежит $\mathscr{X}$, и вычислять объём подмножеств посредством меры Лебега, определённой на $\mathscr{X}$. Например, двумерный 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. Понятно, что эта конкретная модель эквивалентна модели одномерного равномерного распределения в отрезке $[0; 2\pi]$. См. также пример 13, стр. 27

2) Другому популярному распределению — нормальному, будет посвящён целый раздел. Здесь же приведём один частный

вариант двумерной плотности с коэффициентом корреляции ρ :

$$f(x_1, x_2) = \frac{(2\pi)^{-1}}{\sqrt{1-\rho^2}} \exp\Big\{ -\frac{1}{2(1-\rho^2)}\big(x_1^2 - 2\rho x_1 x_2 + x_2^2\big)\Big\}, \quad (16)$$

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.

3) Ещё одна популярная модель — *модель Дирихле* или *многомерное бета-распределение*. Бета-распределение применяется для описания сл.величин, принимающих значения в отрезке $[0;1]$, например 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 $(k+1)$-ую категорию «всё остальное», и задавать его плотность распределения в пространстве $\mathbb{R}^k$, точнее, в части гиперкуба $[0;1]^k$:
$\mathscr{X}_{\mathfrak{D}} = \big\langle\, \vec{x} \in \mathbb{R}^k : \;\; 0 \leqslant x_1, \ldots, x_k \leqslant 1, \quad x_1 + \ldots + x_k \leqslant 1 \,\big\rangle,$

$$f_{\mathfrak{D}}(\vec{x}) = \frac{1}{\mathrm{B}(a_1, \ldots, a_{k+1})} \prod_{j=1}^{k} x_j^{a_j - 1} \Big(1 - \Big(\sum_{j=1}^{k} x_j\Big)\Big)^{a_{k+1}-1}, \quad (17)$$

$f_{\mathfrak{D}}(\vec{x}) = 0$ вне носителя $\mathscr{X}_{\mathfrak{D}}$; параметры модели $a_i > 0$ для $\forall\, i = 1, \ldots, k+1$. Кратко пишем $\vec{\xi} \overset{d}{\backsim} \mathtt{Dir}_k(a_1, \ldots, a_{k+1})$. Функция

$$\mathrm{B}(a_1, \ldots, a_{k+1}) = \int_{\mathscr{X}_{\mathfrak{D}}} f_{\mathfrak{D}}(\vec{x})\, d\vec{x} = \frac{\Gamma(a_1)\cdots\Gamma(a_{k+1})}{\Gamma(a_1 + \ldots + a_{k+1})}$$

есть многомерный вариант бета-функции Эйлера.

При $k = 1$ имеем бета-распределение в отрезке $[0;1]$: $\mathcal{D}\mathtt{ir}_1(a_1, a_2) \overset{d}{\backsim} \mathcal{B}\mathtt{et}(a_1, a_2)$. При $k = 2$ и фиксированном x_1

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

Подстановка $x_2 \dashrightarrow (1 - x_1)t,\ t \in [0;1]$, даёт в правой части

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

Интеграл равен бета-функции $\mathrm{B}(a_2, a_3) = \dfrac{\Gamma(a_2)\Gamma(a_3)}{\Gamma(a_2 + a_3)}$, т.е.

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

что совпадает с плотностью бета-распределения $\mathcal{B}\mathtt{et}(a_1, a_2{+}a_3)$. Аналогичные выкладки справедливы и для $k > 2$, т.е. интеграл от функции (17) по какой-либо переменной по всей области возможных значений даёт функцию того же вида.

Таким образом, во-первых, функция $f_{\mathcal{D}}$ действительно есть ф.пл. ($\int f_{\mathcal{D}}\, d\vec{x} = 1$). Во-вторых, в соответствии с теоремой 29 сл.вектор $(\xi_1, \ldots, \xi_{k-1}) \overset{d}{\backsim} \mathcal{D}\mathtt{ir}_{(k-1)}(a_1,\ \ldots, a_{k-1}, a_k + a_{k+1})$. И аналогично для других подвекторов: все исключённые сл.в. присоединяются вместе со своими параметрами к категории «всё остальное»; в частности, $\xi_1 \overset{d}{\backsim} \mathcal{B}\mathtt{et}(a_1, a_2 + \ldots + a_{k+1})$. Позднее мы покажем, что если часть сл.в. объединить (путем сложения) в новую сл.в., то новый вектор снова будет иметь распределение Дирихле. Например, $(\xi_1, \ldots, \xi_{k-2}, (\xi_{k-1} + \xi_k)) \overset{d}{\backsim} \mathcal{D}\mathtt{ir}_{(k-1)}(a_1, \ldots, (a_{k-1} + a_k), a_{k+1})$. ◎

§ 4. Числовые характеристики. Корреляция

Для каждой компоненты ξ_j сл.вектора $\vec{\xi} \overset{d}{\backsim} F$ можно определить математическое ожидание 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 кусочно-непрерывную плотность f_{ξ_j} и интеграл сходится абсолютно. Ранее (3), стр. 22, матем.ожидание любой функции $h(\vec{\xi}\,)$ определялось через совместную ф.р. F.

33| Лемма. *Пусть $h : \mathbb{R}^m \mapsto \mathbb{R}^1$ — борелевская функция, $\widetilde{h} : \mathbb{R}^k \mapsto \mathbb{R}^1$ — её естественное продолжение на пространство размерности $k = m+n$: $\widetilde{h}(x_1,\ldots,x_m,y_1,\ldots,y_n) = h(x_1,\ldots,x_m)$, $\forall\,(x_1,\ldots,x_m) \in \mathbb{R}^m, (y_1,\ldots,y_n) \in \mathbb{R}^n$. Тогда для любого k-мерного сл.вектора $\vec{\zeta} = (\xi_1,\ldots,\xi_m,\eta_1,\ldots,\eta_n)$ с ф.р. F матем.ожидание*

$$\mathbf{E}\widetilde{h}(\zeta) = \int_{\mathbb{R}^k} \widetilde{h}(\vec{z})\,dF(\vec{z}) = \int_{\mathbb{R}^m} h(\vec{x})\,dF_{\vec{\xi}}(\vec{x}) = \mathbf{E}h(\vec{\xi}), \qquad (18)$$

где $F_{\vec{\xi}}(\vec{x})$, $\vec{x} \in \mathbb{R}^m$, — (маргинальная) функция распределения $(\xi_1,\ldots,\xi_m)$ и интегралы существуют или не существуют, конечны или бесконечны одновременно.

В частности, $\mathbf{E}\xi_j = \int_{\mathbb{R}} x\,dF_{\xi_j}(x) = \int_{\mathbb{R}^k} x_j\,dF_{\vec{\xi}}(\vec{x})$, $j = \overline{1,k}$.

$\Rrightarrow$ Воспользуемся схемой построения интеграла Лебега. Напомним, что интеграл Лебега–Стилтьеса относительно ф.р. (матем.ожидание) есть просто обозначение для интеграла Лебега относительно меры Лебега–Стилтьеса, порождённой этой ф.р. Поэтому достаточно доказать только равенство двух крайних частей (18). Пусть сначала $h(\vec{x}) = c\,\mathbf{I}_B(\vec{x})$, $\vec{x} \in \mathbb{R}^m$, — (масштабированная) индикаторная функция множества $B \in \mathscr{B}^m$, $c \in \mathbb{R}^1$. Тогда $\widetilde{h}(\vec{x},\vec{y}) = c\,\mathbf{I}_{\widetilde{B}}(\vec{x},\vec{y})$ с k-мерным множеством («полосой») $\widetilde{B} = B \times \mathbb{R}^n$ и по определению интеграла

$$\mathbf{E}\widetilde{h}(\vec{\xi},\vec{\eta}) = c\,\mathbf{P}\big\{\,\vec{\xi} \in B, \vec{\eta} \in \mathbb{R}^n\,\big\} = c\,\mathbf{P}\big\{\,\vec{\xi} \in B\,\big\} = \mathbf{E}h(\vec{\xi}).$$

Отсюда следует, что (18) справедливо для любых простых функций h, представимых в виде конечной линейной комбинации индикаторных функций. Пусть теперь $h \geqslant 0$ — неотрицательная измеримая функция и $h_n \nearrow h$ — последовательность простых функций, монотонно сходящихся к h. Тогда из предыдущего $\mathbf{E}\widetilde{h}_n(\vec{\xi}, \vec{\eta}) = \mathbf{E}h_n(\vec{\xi})$. По определению интеграла Лебега правая часть здесь сходится к $\mathbf{E}h(\vec{\xi})$. Так как, очевидно, $\widetilde{h}_n \nearrow \widetilde{h}$, то по теореме о монотонной сходимости Б. Леви левая часть будет сходиться к $\mathbf{E}\widetilde{h}$.

Последний шаг схемы лебеговского интегрирования (переход от положительных функций к измеримым) очевиден. $\Lleftarrow$

Линейная регрессия. Аналогично матем.ожиданию сл.в. ξ_j можно определить её дисперсию $\sigma_j^2 = \mathbf{D}\xi_j = \mathbf{E}(\xi_j - \mu_j)^2$. Напомним одно из свойств дисперсии: если дисперсия конечна, то она доставляет минимум средне-квадратической ошибки прогноза сл.в. посредством константы, т.е. $\sigma_j^2 = \min_c \mathbf{E}(\xi_j - c)^2$. Рассмотрим две сл.в. ξ_1, ξ_2 и найдём минимум средне-квадратической ошибки прогноза сл.в. ξ_1 посредством линейной функции от ξ_2.

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

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

Задача минимизации здесь эквивалентна (?!) поиску минимума функции

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

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

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. 170).

Определения. *Коэффициентом ковариации* называется $\sigma_{12} = \mathbf{Cov}(\xi_1,\xi_2) := \mathbf{E}\big[(\xi_1-\mu_1)(\xi_2-\mu_2)\big]$, 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. В ситуации, когда коэф.корреляции равен нулю, говорят, что сл.в. *не коррелируют*.

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

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

35| $\mathfrak{Y}\mathfrak{n}\mathfrak{p}$. 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 $\mathbf{D}(\xi_1-\xi_2) = \mathbf{D}\xi_1 - \mathbf{D}\xi_2$.

Таким образом, справедлива

36| Теорема. *Уравнение* $x_1 = \mu_1 + \beta(x_2 - \mu_2)$ *с* $\beta = \rho\frac{\sigma_1}{\sigma_2}$ *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. Из вида остаточной дисперсии (каковая всегда неотрицательна) легко выводятся свойства коэфф.корреляции.

37| Теорема. *Пусть* ρ *— коэффициент корреляции между сл.в.* 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.],

т.е. ξ_1, ξ_2 *линейно связаны, причём знак* $\mathrm{sign}(\beta) = \mathrm{sign}(\rho)$.

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

(R_4) $\mathbf{Corr}(v\,\xi_1 + u,\ b\,\xi_2 + a) = \mathrm{sign}(vb)\,\rho$

для $\forall\, a, b, u, v \in \mathbb{R}^1$, *т.е. коэф.корр. не изменяется при линейных преобразованиях сл.в., если угловые коэффициенты имеют один знак.*

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

25 , стр. 36:

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

Свойство (R_4) легко проверить (?!) непосредственной подстановкой в определение коэф.корр. ⫷

⚠ Свойство (R_3), вообще говоря, необратимо. Можно привести примеры, в которых сл.в. зависимы, но не коррелируют (см. также пример 59 ниже).

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

$$\mathbf{Corr}(\xi_2, \xi_1 - \alpha - \beta\xi_2) = 0. \tag{19}$$

Из (19), следует, что дисперсия $\mathbf{D}\xi_1 = \sigma^2_{1\cdot 2} + \sigma^2_{1/2}$, где $\sigma^2_{1\cdot 2} = \sigma^2_1(1 - \rho^2)$ — остаточная дисперсия, $\sigma^2_{1/2} = \mathbf{D}[\alpha + \beta\xi_2]$ — дисперсия регрессии. Другими словами, изменчивость сл.в. ξ_1 на $\rho^2 \cdot 100\%$ может быть объяснена влиянием на неё сл.в. ξ_2.

Матрица ковариаций и матрица корреляций.

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

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

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

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

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

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

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

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

В частности, $\mathbf{E}\vec{\xi} = \vec{\mu} = (\mathbf{E}\xi_1, \ldots, \mathbf{E}\xi_k)^\flat$ (— вектор-столбец). Легко устанавливаются простейшие алгебраические свойства такого матем.ожидания.

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

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

Отсюда следует $\mathbf{E}\,[\vec{u}^\flat \vec{\xi}\,] = \vec{u}^\flat \mathbf{E}\vec{\xi} = (\mathbf{E}\vec{\xi})^\flat \vec{u}$ *для* $\forall\, \vec{u} \in \mathbb{R}^k,$ *если все компоненты вектора матем.ожиданий* $\mathbf{E}\vec{\xi}$ *конечны.*

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

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

b) *матрица ковариаций неотрицательно определена:*

$$\boldsymbol{\Sigma} \geqslant 0;$$

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

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

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

$\Rrightarrow$ a) Произведение $\vec{\xi}\,\vec{\xi}^\flat$ вектор-столбца $\vec{\xi}$ на строку $\vec{\xi}^\flat$ есть матрица всех пар произведений компонент вектора $\vec{\xi}.$ Далее можно воспользоваться свойством 34 .

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

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

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

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

40| С л е д с т в и е. *Матрица ковариаций вырождена, то есть* $|\,\mathbf{Cov}(\vec{\xi})| = 0$ *т. т. т. когда компоненты вектора* $\vec{\xi}$ *линейно связаны:* $\vec{u}^\flat\vec{\xi} - c = 0$ [п.н.] *с некоторым детерминированным вектором* $\vec{u}$ *и скалярной константой* $c \in \mathbb{R}^1$.

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

О п р е д е л е н и е. Определитель матрицы ковариаций $|\boldsymbol{\Sigma}|$ называется *обобщённой дисперсией* случайного вектора.

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

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

Утверждение b) леммы 39 может быть обращено.

41| С л е д с т в и е. *Симметричная матрица* $\boldsymbol{\Sigma}$ *является матрицей ковариаций некоторого сл.вектора т. т. т. когда она неотрицательно определена.*

⊫ Здесь нам понадобится факт из общей теории матриц, по которому для любой матрицы $\boldsymbol{\Sigma} \geqslant 0$ найдётся матрица A $(= \sqrt{\boldsymbol{\Sigma}}\,)$ такая, что $\boldsymbol{\Sigma} = AA^\flat$. Пусть $\vec{\xi}$ — произвольный сл.вектор с независимыми компонентами и единичными дисперсиями. Тогда матрица ковариаций $\mathbf{Cov}(\vec{\xi}) = \dot{\mathbb{I}}$ и в силу леммы 39, c) ковариация $\mathbf{Cov}(A\vec{\xi}) = A\,\dot{\mathbb{I}}\,A^\flat = \mathbf{\Sigma}$. $\Lleftarrow$

Регрессия сл.величины на вектор. Решим теперь задачу построения наилучшего прогноза значений одной сл.в. с помощью линейной функции от вектора сл.в. Обозначим через ζ прогнозируемую сл.в., через $\vec{\xi} = (\xi_1, \ldots, \xi_k)^\flat$ — наблюдаемый вектор, по значениям которого мы хотим предсказать значение ζ. Необходимо найти вектор $\vec{c}$, для которого достигается минимум средне-квадратического расхождения $H(\vec{c}) = \mathbf{E}(\zeta - \vec{c}^{\,\flat}\vec{\xi}\,)^2$. Так как минимум средне-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, 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 39, a) матрица ковариаций $\vec{\xi}_*$

$$\mathbf{Cov}(\vec{\xi}_*) = \mathbf{E}\left[\begin{pmatrix}\zeta\\ \vec{\xi}\end{pmatrix}(\zeta, \vec{\xi}^{\,\flat})\right] = \begin{pmatrix}\mathbf{E}\zeta^2 & \mathbf{E}\zeta\vec{\xi}^{\,\flat}\\ \mathbf{E}\vec{\xi}\zeta & \mathbf{E}\vec{\xi}\vec{\xi}^{\,\flat}\end{pmatrix} = \begin{pmatrix}\sigma^2_\zeta & \vec{\sigma}^{\,\flat}_{\zeta\xi}\\ \vec{\sigma}_{\zeta\xi} & \Sigma_\xi\end{pmatrix}, \quad (20)$$

где $\sigma^2_\zeta = \mathbf{D}\zeta$ — дисперсия ζ, $\vec{\sigma}_{\zeta\xi}$ — вектор (k-мерный) ковариаций ζ со всеми компонентами $\vec{\xi}$, $\Sigma_\xi = \mathbf{Cov}(\vec{\xi})$ — матрица ковариаций $\vec{\xi}$. В этих обозначениях, с учётом утверждения леммы 39, c) и правил обращения с блочными матрицами

$$H(\vec{c}) = \mathbf{D}\vec{c}^{\,\flat}_*\vec{\xi}_* = \vec{c}^{\,\flat}_*\,\mathbf{Cov}(\vec{\xi}_*)\vec{c}_* = \sigma^2_\zeta - 2\vec{c}^{\,\flat}\vec{\sigma}_{\zeta\xi} + \vec{c}^{\,\flat}\Sigma_\xi\vec{c}.$$

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 $H = \sigma_\zeta^2 + \vec{c}^\flat \Sigma_\xi \vec{c} \geqslant \sigma_\zeta^2$, ввиду неотрицательной определённости Σ_ξ; 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

$$\mathbf{E}\big[\vec{\xi}\eta\big] = \mathbf{E}\big[\vec{\xi}\zeta\big] + \mathbf{E}\big[\vec{\xi}\vec{\xi}^\flat \vec{e}\big] = \vec{\sigma}_{\zeta\xi} + \Sigma_\xi \vec{e}.$$

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

42| Т е о р е м а. I) *Если $\Sigma_\xi > 0$, то линейная функция*

$$Z(\vec{x}) = \mu_\zeta + \vec{\sigma}^\flat_{\zeta\xi}\, \Sigma_\xi^{-1}\big(\vec{x} - \vec{\mu}_\xi\big) \tag{21}$$

с вектором $\vec{\sigma}_{\zeta\xi} = \mathbf{Cov}(\zeta, \vec{\xi})$ доставляет наилучший в средне-квадратическом линейный прогноз значения сл.в ζ по реализациям сл.вектора $\vec{\xi} = \vec{x}$.

II) *Остаточная дисперсия прогноза (дисперсия ошибки)*

$$\sigma^2_{\zeta\boldsymbol{\cdot}\xi} = \sigma_\zeta^2 - \vec{\sigma}^\flat_{\zeta\xi}\, \Sigma_\xi^{-1}\vec{\sigma}_{\zeta\xi} = \mathbf{D}\zeta - \mathbf{D}Z(\vec{\xi}). \tag{22}$$

⚠ Применяя известные операции к блочным матрицам, замечаем, что остаточная дисперсия с точностью до множителя $|\Sigma_\xi|$ совпадает с определителем матрицы ковариаций (20). Таким образом, остаточную дисперсию можно записать в виде $|\,\mathbf{Cov}(\zeta, \vec{\xi}^\flat)|\,/\,|\Sigma_\xi| = 1/\tilde{\sigma}_\zeta^2$, где $\tilde{\sigma}_\zeta^2$ — первый диагональный элемент обратной матрицы ковариаций $\big(\mathbf{Cov}(\zeta, \vec{\xi}^\flat)\big)^{-1}$.

Как было отмечено, 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 несколько вольной интерпретации, что сл.функция $Z(\vec{\xi})$ несёт в себе «всю» информацию о влиянии $\vec{\xi}$ на ζ. Это даёт основание ввести

Определение. Пусть $Z(\vec{x})$ — линейная средне-квадратическая регрессия сл.в. ζ на сл.вектор $\vec{\xi}$. Коэф.корр.

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

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

В качестве альтернативного определения множественного коэф.корр. может быть взято любое из свойств (R_2), (R_5) следующей теоремы, в которой сохранены предыдущие обозначения.

43| Теорема. *Пусть $|\Sigma_\xi| > 0$. Множественный коэффициент корреляции $\rho_{\zeta\bullet\xi}$ удовлетворяет следующим свойствам:*

(R_1) $0 \leqslant \rho_{\zeta\bullet\xi} \leqslant 1$;

(R_2) $$\rho_{\zeta\bullet\xi} = \frac{\sqrt{\vec{\sigma}^{\flat}_{\zeta\xi}\,\Sigma_\xi^{-1}\vec{\sigma}_{\zeta\xi}}}{\sigma_\zeta} = \left(1 - \frac{\sigma^2_{\zeta\bullet\xi}}{\sigma^2_\zeta}\right)^{1/2};$$

(R_3) $\rho_{\zeta\bullet\xi} = 0$ *т. т. т. когда сл.в. ζ не коррелирует со всеми сл.в. ξ_j, $j = \overline{1,k}$;*

(R_4) $\rho_{\zeta\bullet\xi} = 1$ *т. т. т. когда сл.в. ζ линейно зависит от компонент вектора $\vec{\xi}$, т.е. почти наверное $\zeta = Z(\vec{\xi})$;*

(R_5) $\rho_{\zeta\bullet\xi}$ *есть максимальный коэф.корреляции между сл.в. ζ и линейными функциями компонент вектора $\vec{\xi}$:*

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

⊫ (R$_5$) Пусть $\eta = \vec{c}^{\flat}\vec{\xi}$ — некоторая линейная комбинация $\vec{\xi}$. Заметим сначала, что по свойству обычного коэфф.корреляции этот коэффициент не изменяется при линейных преобразованиях сл.величин. Поэтому, не умаляя общности, можно предположить, что $\mathbf{E}\zeta = 0, \mathbf{E}\vec{\xi} = \vec{0}, \mathbf{E}\eta = 0$, дисперсии $\mathbf{D}\zeta = \mathbf{D}\eta = 1$. В этих предположениях искомый коэфф.корреляции равен коэффициенту ковариации:

$$\mathbf{Corr}(\zeta, \eta) = \mathbf{E}\left[\eta\zeta\right] = \mathbf{E}\left[\vec{c}^{\flat}\vec{\xi}\zeta\right] = \vec{c}^{\flat}\mathbf{E}\left[\vec{\xi}\zeta\right] = \vec{c}^{\flat}\vec{\sigma}_{\zeta\xi}. \tag{23}$$

Функцию линейной регрессии с этой же целью представим в виде $Z(\vec{\xi}) = \sigma_Z \vec{\beta}^{\flat}\vec{\xi}$ с соответствующим вектором $\vec{\beta}$ и $\sigma_Z^2 = \mathbf{D}Z = \vec{\sigma}^{\flat}_{\zeta\xi}\Sigma_{\xi}^{-1}\vec{\sigma}_{\zeta\xi}$. Из предыдущих выкладок $\mathbf{Corr}(\zeta, Z) = \vec{\beta}^{\flat}\vec{\sigma}_{\zeta\xi}$.

По определению регрессии $\mathbf{E}\left(\zeta - \sigma_Z\vec{\beta}^{\flat}\vec{\xi}\right)^2 \leqslant \mathbf{E}\left(\zeta - \sigma_Z\vec{c}^{\flat}\vec{\xi}\right)^2$. Откуда, учитывая предположение о дисперсиях линейных комбинаций, имеем $1 + \sigma_Z^2 - 2\sigma_Z\vec{\beta}^{\flat}\vec{\sigma}_{\zeta\xi} \leqslant 1 + \sigma_Z^2 - 2\sigma_Z\vec{c}^{\flat}\vec{\sigma}_{\zeta\xi}$, что в сочетании с (23) доказывает (R$_5$). ⫤

44| $\mathfrak{Y}\mathfrak{np}$. Докажите свойства (R$_1$)–(R$_4$).

45| Примеры. 1) Пусть $\vec{\xi} \stackrel{d}{\backsim} \mathcal{H}\mathrm{g}_k(n; N_1, \ldots, N_k)$. Как отмечалось в примере 11, стр. 25, вектор $(\xi_1, n - \xi_1) \stackrel{d}{\backsim} \mathcal{H}\mathrm{g}_2(n; N_1, N - N_1)$, т.е. ξ_1 имеет гипергеометрическое (одномерное) распределение. Поэтому матем.ожидание и дисперсия компонент многомерного гипергеометрического распределения равны:

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

где $q_j = N_j/N$ — относительная доля j-ой 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 коэфф.корреляции

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

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

2) Если $(\xi_1, \ldots, \xi_k) \overset{d}{\backsim} \mathcal{M}\mathtt{ult}_k(n; q_1, \ldots, q_k)$, то, как и в предыдущем примере, $(\xi_1, n - \xi_1) \overset{d}{\backsim} \mathcal{M}\mathtt{ult}_2(n; q_1, 1 - q_2)$, т.е. $\xi_1 \overset{d}{\backsim} \mathcal{B}\mathtt{in}(n; q_1)$. 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 $(\xi_1, \xi_2, \xi_3) \overset{d}{\backsim} \mathcal{M}\mathtt{ult}_3(n; q_1, q_2, 1 - q_1 - q_2)$:

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

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

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

где последнее равенство справедливо ввиду равенства 1 суммы всех вероятностей модели $\mathcal{M}\mathtt{ult}_3(n'; q_1, q_2, 1-q_1-q_2)$. Таким образом, коэфф.ковариации $\mathbf{Cov}(\xi_1, \xi_2) = q_1 q_2 n(n-1) - nq_1\, nq_2 =$

$= -nq_1q_2,$ а коэфф.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 $\vec{\xi} \overset{d}{\backsim} \mathtt{Mult}_k(n; q_1, \ldots, q_k)$ можно записать в виде

$$\mathbf{Cov}(\,\vec{\xi}\,) = n\Big(Q - \vec{q}\,\vec{q}^{\,\flat}\Big),$$

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 $\xi_1,$ по значению $\xi_2 = x_2$ другой компоненты мультиномиального (и многомерного гипергеометрического) сл.вектора $(\xi_1, \xi_2, \ldots, \xi_k)$ равен

$$x_1^* = \frac{q_1}{1-q_2}(n - x_2), \tag{24}$$

даже в случае, если сумма вероятностей (частот) $q_1 + q_2 = 1.$

Здесь есть небольшая проблема применимости, связанная с дискретностью $\xi_1.$ 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 анализа остальных выстрелов потребуется время. Во время ожидания результатов этого анализа может возникнуть желание заключить пари на количество попаданий в «яблочко». Линейный прогноз ожидаемого результат даёт $x_1 = 8.18$. Такой ответ, естественно, никого не устроит. Чтобы дать приемлемый ответ в виде целого числа, не надо торопиться округлять, прогнозируя $x_1 = 8$.

В данной ситуации штраф за ошибочный прогноз редко представляет собой квадрат разности $(d - r)^2$, где d — выдвинутый прогноз, а r — действительный результат. Обычно 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 $0.424 < 0.5$.

3) Для модели Дирихле $(\xi_1, \xi_2) \overset{d}{\backsim} \mathcal{D}\mathtt{ir}_2(a_1, a_2, a_3)$ (см. (17), стр. 46), 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 $m = 1, r = 1$)

$$\mathbf{E}\xi_1^m\xi_2^r = \int\limits_0^1 dx_1 \int\limits_0^{1-x_1} x_1^m x_2^r \frac{x_1^{a_1-1}x_2^{a_2-1}}{\mathrm{B}(a_1,a_2,a_3)}(1-(x_1+x_2))^{a_3-1}\,dx_2 =$$

$$= \frac{\mathrm{B}(m+a_1, r+a_2, a_3)}{\mathrm{B}(a_1,a_2,a_3)}\Bigg|_{\substack{m=1,\\ r=1}} = \frac{a_1a_2}{(a_1+a_2+a_3+1)(a_1+a_2+a_3)}\,.$$

Полагая $\boldsymbol{a} = a_1 + \ldots + a_{k+1}$, имеем $\mathbf{E}\,[\xi_1\xi_2] = a_1a_2/(\boldsymbol{a}(\boldsymbol{a} + 1))$ для $\vec{\xi} \overset{d}{\backsim} \mathcal{D}\mathtt{ir}_k(a_1, \ldots, a_{k+1})$. В этих обозначениях матем.ожидание $\mathbf{E}\xi_1 = \mathrm{B}(1 + a_1, a_2, a_3)/\mathrm{B}(a_1, a_2, a_3) = a_1/\boldsymbol{a}$. Поэтому коэф.ковариации равен

$$\mathbf{E}\,[\xi_1\xi_2] - \mathbf{E}\xi_1\,\mathbf{E}\xi_2 = -\frac{a_1a_2}{\boldsymbol{a}^2(\boldsymbol{a}+1)}\,,$$

а коэф.корреляции

$$\mathbf{Corr}(\xi_1,\xi_2) = -\sqrt{\frac{a_1a_2}{(\boldsymbol{a}-a_1)(\boldsymbol{a}-a_2)}} = -\sqrt{\frac{q_1q_2}{(1-q_1)(1-q_2)}}\,,$$

где $q_j = a_j/\boldsymbol{a}$. Сравните с дискретными моделями. ⊚

§ 5. Эллипсоид рассеяния

Наряду с числовыми характеристиками положения $\vec{\mu} = \mathbf{E}\vec{\xi}$ и разброса $\boldsymbol{\Sigma} = \mathbf{Cov}(\vec{\xi})$ сл.вектора $\vec{\xi}$ часто рассматривают «производную» от них геометрическую характеристику — так называемый эллипсоид рассеяния. Кроме центра сосредоточения и области изменения (в основном) сл.вектора, эллипсоид рассеяния показывает характер зависимости сл.в. Из всего разнообразия геометрических фигур эллипсоид выбран по следующим причинам. Во-первых, эллипсоид имеет наиболее удобное аналитическое представление (ср. с аналитическим представлением параллелограмма, 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 матем.ожиданий k-мерного сл.вектора $\vec{\xi}$, $\boldsymbol{\Sigma} = \mathbf{Cov}(\vec{\xi})$ — его матрица ковариаций, $\boldsymbol{\Sigma} > 0$. Эллипсоид (относительно вектора $\vec{x} \in \mathbb{R}^k$)

$$(\vec{x} - \vec{\mu})^{\flat} \boldsymbol{\Sigma}^{-1} (\vec{x} - \vec{\mu}) \leqslant k + 2 \tag{25}$$

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

Причины выбора такой формы эллипсоида объясняет

46| Теорема. *Пусть* $\boldsymbol{\Sigma} > 0$, *тогда эллипсоид* (25) — *единственный* k*-мерный эллипсоид, равномерное распределение внутри которого имеет идентичные с* $\vec{\xi}$ *вектор матем.ожиданий* $\vec{\mu}$ *и матрицу ковариаций* $\boldsymbol{\Sigma}$.

$\Rrightarrow$ Пусть $\mathfrak{E} : (\vec{x}-\vec{a})^\flat B\,(\vec{x}-\vec{a}) \leqslant c^2$ определяет некоторый эллипсоид с положительно определённой матрицей $B > 0$. Плотность равномерного распределения $f(\vec{x}) = {}^1\!/_V$ при $\vec{x} \in \mathfrak{E}$, где V — объём эллипсоида. Как известно из курса алгебры, 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 $(\vec{x}-\vec{a}) = Q^\flat\Lambda^{-1}\vec{y}$ переводит взаимно-однозначно эллипсоид $\mathfrak{E}$ в шар $\mathfrak{O} : \vec{y}^{\,\flat}\vec{y} \leqslant c^2$. Поскольку якобиан этого преобразования равен $||Q^\flat\Lambda^{-1}|| = 1/\sqrt{|B|}$, 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, т.е. для $r > 0$ и $(\theta, \alpha_1, \ldots, \alpha_{k-2}) \in \Theta$, где

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

положим

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

Как известно, 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

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

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 $\mathfrak{O}$. Т.е. матем.ожидание вектора с равномерным распределением внутри эллипсоида $\mathfrak{E}$ равно $\vec{a}$.

Ковариация равна (в матричной форме записи интегралов)

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

Не углубляясь в процесс вычисления, скажем, 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

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

Таким образом, матрица ковариаций равномерного распределения равна

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

Теорема доказана. ⫤

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

О п р е д е л е н и е. *Обобщённой дисперсией* сл.вектора $\vec{\xi}$ называется определитель матрицы ковариаций $\mathfrak{S}^2(\vec{\xi}) = |\,\mathbf{Cov}(\vec{\xi})|$.

В связи с этим см. также стр. 54.

В одномерном случае обобщённая дисперсия совпадает с дисперсией сл.в.: $\mathfrak{S}^2(\xi) = \sigma^2$. Эллипсоид рассеяния представляет собой отрезок $[\mu - \sqrt{3}\sigma; \mu + \sqrt{3}\sigma]$.

В двумерном случае эллипсоид рассеяния имеет вид

$$X_1^2 - 2\rho X_1 X_2 + X_2^2 \leqslant 4(1 - \rho^2),$$

где ρ — коэфф.корреляции между сл.в. ξ_1, ξ_2, переменные $X_j = (x_j - \mu_j)/\sigma_j$, $j = 1, 2$. Обобщённая дисперсия $\mathfrak{S}^2(\xi_1, \xi_2) = \sigma_1^2 \sigma_2^2 (1 - \rho^2)$. Ясно, что главные оси эллипса рассеяния параллельны осям координат т. т. т. когда коэфф.корреляции $\rho = 0$.

Нетрудно показать, что главная ось этого эллипса совпадает с так называемой линией ортогональной регрессии. Другими словами, на линии $X_2 = \mathrm{sign}(\rho) X_1$ (которая совпадает с главной осью эллипса) достигается минимум средне-квадратического расстояния от реализаций сл.вектора (ξ_1, ξ_2) до линии регрессии:

$$\min_{b,c} \frac{\mathbf{E}(\xi_2 - b\xi_1 - c)^2}{1 + b^2}.$$

§ 6. 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-ой компоненты ζ_1 максимальна среди всех нормированных линейных комбинаций $\vec{\xi}$;

(✓) дисперсия j-ой 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 $|\mathbf{\Sigma}| = 0$. Собственные векторы, отвечающие различным собственным числам, ортогональны. Если j-ый столбец матрицы Q образован нормированным собственным вектором, отвечающим j-ому собственному числу γ_j, то

$$\mathbf{\Sigma}Q = Q\Gamma$$

с диагональной матрицей $\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_k)$. Кроме того, ввиду ортонормированности столбцов матрица Q будет ортогональной: $QQ^\flat = Q^\flat Q = \dot{\mathbb{I}}$.

Некоторая проблема при доказательстве следующей теоремы возникает в случае, когда часть собственных чисел совпадает. Чтобы обойти это неудобство будем считать, что все собственные числа матрицы ковариаций $\mathbf{\Sigma}$ различны.

47| Т е о р е м а. *Пусть* $\mathbf{\Sigma}$ *– матрица ковариаций* $\vec{\xi}$. *Тогда вектор-столбцы* $\vec{c}_j$, *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 $\gamma_1 > \cdots > \gamma_k > 0$.

$\Rrightarrow$ Во-первых, как всегда, когда дело касается дисперсий сл.в., можно предположить, что матем.ожидания равны нулю. Пусть матрица Q образована нормированными вектор-столбцами решений характеристического уравнения с упорядоченными по убыванию собственными числами. Тогда матрица ковариаций сл.вектора $\vec{\zeta} = Q^\flat\vec{\xi}$

$$\mathbf{Cov}(\vec{\zeta}) = Q^\flat\,\mathbf{Cov}(\vec{\xi})Q = Q^\flat\mathbf{\Sigma}Q = Q^\flat Q\Gamma = \Gamma$$

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 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 $\vec{b} = Q^\flat\vec{a}$. Поэтому дисперсия сл.в. $\vec{a}^\flat\vec{\xi}$

$$\mathbf{D}(\vec{a}^\flat\vec{\xi}) = \vec{a}^\flat\mathbf{\Sigma}\vec{a} = \vec{b}^\flat Q^\flat\mathbf{\Sigma}Q\vec{b} = \vec{b}^\flat\Gamma\vec{b} = b_1^2\gamma_1 + \ldots + b_k^2\gamma_k \leqslant \gamma_1,$$

поскольку все собственные числа $\gamma_2, \ldots, \gamma_k < \gamma_1$ и норма вектора $\vec{b}$ равна единице: $\vec{b}^\flat\vec{b} = \vec{a}^\flat QQ^\flat\vec{a} = \vec{a}^\flat\mathbb{I}\vec{a} = 1$, т.к. вектор

$\vec{a}$ должен быть нормированным. Последнее неравенство будет строгим т. т. т. когда коэффициент $b_1 \neq 1$. Это доказывает, что сл.в. $\zeta_1 = \vec{c}_1^{\,\flat}\vec{\xi}$ есть первая главная компонента.

Пусть 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 линейных комбинаций вектора $\vec{\xi}$, не коррелирующих с первыми $j-1$ главными компонентами. $\Lleftarrow$

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:

$$\xi_j^* = \frac{\xi_j - \mu_j}{\sigma_j}.$$

Легко понять, что в этом случае ковариационная матрица $\mathbf{\Sigma}$ совпадает с корреляционной матрицей. Таким образом, метод главных компонент, применённый к нормированным показателям, отличается от этого же метода, но применённого к ненормированным показателям, тем, что в этом случае решается характеристическое уравнение относительно матрицы корреляций, а не ковариаций. Заметим, кстати, что тогда сумма дисперсий главных компонент равна k, так как на диагонали матрицы корреляций стоят единицы.

Свойства главных компонент.

Пусть $\vec{\zeta} = Q^\flat \vec{\xi}$ — вектор главных компонент $\vec{\xi}$.

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-ой главной компоненты равна j-ому собственному числу матрицы $\mathbf{\Sigma}$:

$$\mathbf{D}\zeta_j = \gamma_j \,,\ j = 1, \ldots, k.$$

4) Переход к главным компонентам сохраняет обобщённую дисперсию

$$\mathfrak{S}^2(\vec{\xi}) = |\mathbf{\Sigma}| = \mathfrak{S}^2(\vec{\zeta}) = \prod\nolimits_1^k \gamma_j.$$

5) Переход к главным компонентам сохраняет сумму дисперсий признаков

$$\sum\nolimits_1^k \mathbf{D}\xi_j = \sum\nolimits_1^k \mathbf{D}\zeta_j = \sum\nolimits_1^k \gamma_j.$$

6) Переход к главным компонентам представляет собой поворот осей системы координат параллельно осям эллипсоида рассеяния сл.вектора $\vec{\xi}$.

Доказательства этих свойств сразу следуют либо из утверждения предыдущей теоремы, либо из её доказательства. Например,

$$\sum\nolimits_1^k \mathbf{D}\xi_j = \mathbf{E}\left[\vec{\xi}^\flat \vec{\xi}\right] = \mathbf{E}\left[\vec{\zeta}^\flat Q^\flat Q \vec{\zeta}\right] = \mathbf{E}\left[\vec{\zeta}^\flat \vec{\zeta}\right] = \sum\nolimits_1^k \gamma_j.$$

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. Для каждого $1 \leqslant q \leqslant k$ рассмотрим отношение

$$\mathfrak{m}_q = \frac{\gamma_1 + \ldots + \gamma_q}{\gamma_1 + \ldots + \gamma_k}.$$

Величина $\mathfrak{m}_q \cdot 100\%$ показывает ту долю разброса наблюдений, которая может быть объяснена факторами, определяющими первые q главных компонент. Если $\mathfrak{m}_q \cdot 100\%$ достаточно велико, то оставшиеся $k-q$ главных компонент могут быть исключены из рассмотрения. Точное значение понятия «достаточно велико» конкретизируется, исходя непосредственно из рассматриваемой исследовательской задачи. Можно предложить следующую градацию для описания информативности главных компонент.

Если $\mathfrak{m}_q \cdot 100\% > 75\%$, то говорят, что первые q главных компонент содержат *основную массу* информации о разбросе наблюдаемых признаков, если $\mathfrak{m}_q \cdot 100\% > 90\%$, то *почти всю* информацию, и наконец, если $\mathfrak{m}_q \cdot 100\% > 95\%$, то *всю* информацию. Приведённые цифры, 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.

48| Лемма. (?!) 1) *Наилучшее линейное приближение вектора $\vec{\xi}$ по значениям первых q главных компонент $\vec{\zeta}^{(q)}$ доставляет преобразование $\vec{\xi}^* = Q^{(q)}\vec{\zeta}^{(q)}$ с матрицей $Q^{(q)}$ размера $k \cdot q$, построенной на первых q столбцах матрицы Q.*

2) *Относительная суммарная ошибка этого прогноза равна*

$$\frac{\sum_1^k \mathbf{E}(\xi_j - \xi_j^*)^2}{\sum_1^k \sigma_j^2} = 1 - \mathfrak{m}_q.$$

§ 7. Условное распределение

Строгому определению условного матем.ожидания и связанному с ним понятию условного распределения для общей вероятностной модели в дальнейшем мы посвятим специальный раздел. Здесь же введём эти понятия для дискретных и абсолютно-непрерывных моделей. В дальнейшем знаки $\mathbf{P}, \mathbf{E}$ будут использоваться для обозначения вероятности (матем.ожидания) относительно распределения на исходном вероятностном пространстве, знаки $\mathbb{P}, \mathbb{E}$ будут обозначать соответствующие операции в пространстве значений сл.величин ($\mathbb{R}^k$). Такое разграничение применяется только для того, чтобы обратить внимание на способ вычисления соответствующей характеристики, например: матем.ожидание любой функции от сл.в. ξ равно

$$\mathbf{E}\big[h(\xi)\big] = \int_\Omega h(\xi(\omega))\ \mathbf{P}(d\omega) = \int_\mathbb{R} h(x)\,\mathbb{P}_\xi(dx) = \mathbb{E}\big[h(\xi)\big].$$

Кроме того, договоримся в этом разделе понимать интегралы как лебеговские, независимо от их записи.

Пусть $p(x,y)$, $x \in \mathscr{X}$, $y \in \mathscr{Y}$, — ф.вер. 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 $\mathbb{P}_{\xi|\eta}(x\,|\,y) := \mathbf{P}\left\{\xi = x\,|\,\eta = y\right\} = p(x,y)/p_\eta(y)$. Легко понять из (5), стр. 25, что сумма вероятностей $\mathbb{P}_{\xi|\eta}(x\,|\,y)$ по всем $x \in \mathscr{X}$ равна 1. Таким образом, набор вероятностей

$$\mathbb{P}_{\xi|\eta}(x\,|\,y) = \frac{p(x,y)}{p_\eta(y)}, \quad x \in \mathscr{X}, \tag{26}$$

определяет дискретное распределение, которое естественно назвать *условным распределением ξ при фиксированном значении сл.в. $\eta = y$*. 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 ф.вер. $\mathbb{P}_{\xi|\eta}(x\,|\,y)$ и при $y \notin \mathscr{Y}$. Поскольку вероятность совокупности всех точек $y \notin \mathscr{Y}$ (подчеркнём, что именно всех точек сразу, а не каждой точки в отдельности) равна нулю, такое доопределение можно осуществить произвольным образом, учитывая, например, удобство записи, лишь бы только функция $\mathbb{P}_{\xi|\eta}(x\,|\,y)$ задавала некоторое распределение. Например, можно положить $\mathbb{P}_{\xi|\eta}(x_0\,|\,y) = 1, \mathbb{P}_{\xi|\eta}(x\,|\,y) = 0,\ x \neq x_0,$ с некоторым произвольным $x_0 \in \mathscr{X}$, возможно зависящим от y.

Относительно условного распределения можно определить любые числовые характеристики. Например, для измеримой функции $h(x, y)$ величина

$$\mathbb{E}\left[h(\xi,\eta)\,|\,y\right] := \sum_{x\in\mathscr{X}} h(x,y)\mathbb{P}_{\xi|\eta}(x\,|\,y) \tag{27}$$

задаёт условное матем.ожидание (усл.м.о.) $h(\xi, \eta)$ при фиксированном значении $\eta = y$. В частности, для индикаторной функции $h(x, y) = \mathbf{I}_A(x)$ получим классическую условную вероятность $\mathbb{P}_\xi\{A\,|\,y\} = \mathbf{P}\left\{\xi \in A\,|\,\eta = y\right\}$ 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. отличается от него только на множестве $\mathscr{Y}^c$ нулевой меры, т.е. почти наверное совпадает с ним.

Легко видеть, что усл.м.о. $\mathbb{E}\left[h(\xi,\eta)\,|\,y\right]$ как функция y измеримо, поэтому суперпозиция $\mathbb{E}\left[h(\xi,\eta)\,|\,\eta\right]$ представляет собой сл.в. Её матем.ожидание

$$\mathbb{E}\Big[\mathbb{E}\left[h(\xi,\eta)\,|\,\eta\right]\Big] = \sum_{y\in\mathscr{Y}}\Big[\sum_{x\in\mathscr{X}} h(x,y)\mathbb{P}_{\xi|\eta}(x\,|\,y)\Big]p_\eta(y) = \\ = \sum_{(x,y)\in\mathscr{X}\times\mathscr{Y}} h(x,y)p(x,y) = \mathbb{E}[h(\xi,\eta)]. \tag{28}$$

Для индикаторной функции $h(x,y) = \mathbf{I}_A(x)$ это соотношение даёт хорошо известную формулу полной вероятности:

$$\mathbf{P}\{\xi\in A\} = \mathbb{E}[\,\mathbf{I}_A(\xi)] = \sum_{y\in\mathscr{Y}} \mathbb{P}_\xi\{A\,|\,y\}p_\eta(y).$$

Пусть теперь сл.вектор (ξ,η) имеет абсолютно-непрерывную ф.р. $F.$ В этой ситуации мы не можем вычислить условную вероятность $\mathbf{P}\left\{\xi\in A\,|\,\eta=y\right\}$ стандартным способом, поскольку вероятность любого одноточечного события $\{\eta=y\}$ равна нулю. Приведём некоторые соображения, позволяющие всё-таки ввести понятие условного распределения и в этом случае.

Первая идея исходит из принципа «по аналогии». Условная вероятность равна отношению совместной вероятности событий на вероятность условия. В непрерывном случае аналогом вероятности одноточечного события выступает ф.пл. Пусть $f(x,y)$ — совместная ф.пл. (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

$$f_{\xi|\eta}(x\,|\,y) = \frac{f(x,y)}{f_\eta(y)}, \qquad x\in\mathbb{R}^1. \tag{29}$$

49 Лемма. (?!) *При $f_\eta(y)>0$ функция (29) задаёт плотность некоторого распределения, т.е. $f_{\xi|\eta}(x\,|\,y)\geqslant 0$ и $\int_{\mathbb{R}} f_{\xi|\eta}(x\,|\,y)\,dx = 1$.*

К той же самой ф.пл. мы придём, если определим условную плотность в точке «естественным» образом как предел условных вероятностей в «малой» окрестности условия:

$$\lim_{\epsilon\to+0}\frac{1}{\epsilon}\,\mathbf{P}\bigl\{\xi\in(x-\epsilon;x+\epsilon]\,|\,\eta\in(y-\epsilon;y+\epsilon]\bigr\} =$$

$$=\lim_{\epsilon\to+0}\frac{\bigl(\Delta_{\vec{u}-\epsilon,\vec{u}+\epsilon}F\bigr)/\epsilon^2}{\bigl(F_\eta(y+\epsilon)-F_\eta(y-\epsilon)\bigr)/\epsilon}=\frac{f(x,y)}{f_\eta(y)},$$

где $\vec{u}\pm\epsilon=(x\pm\epsilon,y\pm\epsilon)$ и последнее равенство справедливо для почти всех (по распределению F) точек (x,y) с $f_\eta(y)>0$.

Отметим, что приведённые здесь соображения носят чисто эвристический характер и не претендуют на строгость обоснования. Важнее не то, как мы определили условное распределение, а те свойства, которые отсюда могут быть получены. Ключевым является свойство (28) (или его более широкий вариант (31)), позволяющее вычислять любые характеристики сл.вектора (ξ,η) последовательно — сначала при фиксированном значении η, затем усредняя по распределению η. В этих целях, доопределим сначала условную плотность (29) для точек с $f_\eta(y)=0$. Поскольку вероятность совокупности всех таких точек, очевидно, равна нулю, можно определить $f_{\xi|\eta}(x\,|\,y)$ произвольным образом, лишь бы только эта функция задавала ф.пл., например, положить $f_{\xi|\eta}(x\,|\,y)=f_0(x)$ с некоторой фиксированной ф.пл. $f_0(x)$ на $\mathbb{R}^1$ относительно меры Лебега.

Математическое ожидание относительно условной плотности

$$\mathbb{E}\{h(\xi,\eta)\,|\,y\}=\mathbb{E}\{h(\xi,\eta)\,|\,\eta=y\}:=\int\limits_{\mathbb{R}}h(x,y)f_{\xi|\eta}(x\,|\,y)\,dx \qquad (30)$$

назовём условным матем.ожиданием $h(\xi,\eta)$ при фиксированном $\eta=y$. Выбирая здесь $h(x,y)=\mathbf{I}_A(x)$, т.е. индикаторную функцию некоторого множества $A\subset\mathbb{R}^1$ или множества вида $A=\{x\in\mathbb{R}^1: x\leqslant z\}$, $z\in\mathbb{R}^1$, получаем условное распределение или, соответственно, условную ф.р. сл.в. ξ :

$$\mathbb{P}_{\xi|\eta}\{A\,|\,y\} := \int_A f_{\xi|\eta}(x\,|\,y)\,dx, \qquad F_{\xi|\eta}(z\,|\,y) := \int_{-\infty}^{z} f_{\xi|\eta}(x\,|\,y)\,dx.$$

Введённые выше понятия условного матем.ожидания удовлетворяют ключевому соотношению, которое берётся за основу при определении усл.м.о. для общих вероятностных моделей.

50| 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*

$$\begin{aligned} \mathbf{E}\left[g(\eta)h(\xi,\eta)\right] &= \mathbf{E}\Big(g(\eta)\mathbb{E}\left[h(\xi,\eta)\,|\,\eta\right]\Big) = \\ &= \int\limits_{\mathbb{R}} g(y)\Big[\int\limits_{\mathbb{R}} h(x,y) f_{\xi|\eta}(x\,|\,y)\,dx\Big] f_\eta(y)\,dy. \end{aligned} \tag{31}$$

Выбирая в качестве g и h индикаторные функции $g(y) = \mathbf{I}_A(y)$, $h(x,y) = \mathbf{I}_B(x,y)$ множеств $A = \{y : y \leqslant z\}$, $B = \{(x,y) : x \leqslant u, -\infty < y < \infty\}$ и применяя запись интеграла через ф.р., получаем следующее утверждение.

51| Следствие. *Пусть* F_η *— функция распределения* η. *Тогда совместная ф.р.* (ξ,η) *равна*

$$F(u,z) \overset{\mathrm{df}}{=} \mathbf{P}\{\xi \leqslant u, \eta \leqslant z\} = \int\limits_{-\infty}^{z} F_{\xi|\eta}(u\,|\,y)\,dF_\eta(y). \tag{32}$$

Если в соотношениях (31) выбрать $g \equiv 1$, $h(x,y) = \mathbf{I}_A(x)$ с некоторым борелевским множеством $A \subset \mathbb{R}^1$, то получим формулу полной вероятности в непрерывном случае:

$$\mathbf{P}\{\xi \in A\} = \mathbf{E}\left[\mathbf{P}\left\{\xi \in A\,|\,\eta\right\}\right] = \int_{\mathbb{R}} \mathbb{P}_{\xi|\eta}\{A\,|\,y\} f_\eta(y)\,dy.$$

⚠ 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.

52| 𝔜𝔭𝔭. Пусть $\mathbf{D}\big[\xi \,|\, \eta\big]$ — условная дисперсия сл.в. ξ при фиксированной сл.в. η. Покажите, что $\mathbf{D}\xi = \mathbf{E}\big(\mathbf{D}\big[\xi \,|\, \eta\big]\big) + \mathbf{D}\big(\mathbf{E}\big[\xi \,|\, \eta\big]\big)$.

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 соотношение, аналогичное равенствам (28), (31):

$$\mathbf{E}\big[\,\mathbf{I}_B(\eta)h(\xi,\eta)\big] \;=\; \mathbf{E}\big[\,\mathbf{I}_B(\eta)\mathbf{E}\big[h(\xi,\eta) \,|\, \eta\big]\big], \quad \forall\, B \in \mathscr{B}. \tag{33}$$

В соответствии с этим определением, условная ф.р. ξ при фиксированном $\eta = y$ есть измеримая (по y) функция распределения (по x) $F_{\xi|\eta}(x \,|\, y)$, удовлетворяющая (32). Условные распределения, определённые выше, являются только частными вариантами общей конструкции, в том смысле, что любой другой вариант будет отличаться от них лишь почти наверное.

⚠ Отметим одно обстоятельство, связанное с усл.м.о. Исходя из (27) и (30), можно заметить, что при всех y

$$\mathbf{E}\big[h(\xi,\eta) \,|\, \eta = y\big] = \mathbf{E}\big[h(\xi,y) \,|\, \eta = y\big]. \tag{34}$$

В общем случае такое равенство может быть обосновано либо при дополнительных условиях на вектор (ξ,η), например, когда сл.в. независимы, либо когда указан способ вычисления соответствующих усл.м.о. (например, через ф.плотности или ф.вероятности, как это сделано выше, или через условное распределение, если конкретизирован способ его построения).

Из соотношения (32) в случае, когда распределение η абсолютно-непрерывно с плотностью f_η, 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 как

$$F(x\,|\,y) = \frac{1}{f_\eta(y)} \frac{\partial F(x,y)}{\partial y}\,.$$

53| Пример. Рассмотрим функцию вида $F(x,y) = G(x)G(y)\,\mathbf{I}(x \leqslant y) + G^2(y)\,\mathbf{I}(x \geqslant y)$ с некоторой непрерывной ф.р. G (см. пример 28, стр. 41). Маргинальная ф.р. второй компоненты $F_\eta(y) = F(+\infty, y) = G^2(y)$. Если бы ф.р. G имела плотность g, то плотность $f_\eta(y) = 2g(y)G(y)$. В этом случае предыдущая формула даёт усл.ф.р.

$$F(x\,|\,y) = \tfrac{1}{2}\,\tfrac{G(x)}{G(y)}\,\mathbf{I}(x < y) + 1 \cdot \mathbf{I}(x > y).$$

При $x = y$ эту усл.ф.р. мы вольны выбрать произвольно, но с сохранением свойств ф.р. (непрерывность справа), т.е. положить $F(y\,|\,y) = 1$. В записи $F(x\,|\,y)$ пропала плотность g. Можно предположить, что этот результат справедлив и при отсутствии плотности. Проверим справедливость равенства (32).

Если рассмотреть случай $u < z$, то ф.р. $F(u,z) = G(u)G(z)$. Для вычисления интеграла в (32) его необходимо разбить на две части $y \leqslant u$ и $u < y \leqslant z$:

$$\int\limits_{-\infty}^{u} + \int\limits_{u}^{z} F(u\,|\,y)\,dG^2(y) = \int\limits_{-\infty}^{u} 1\,dG^2(y) + \frac{1}{2}G(x)\int\limits_{u}^{z} \frac{dG^2(y)}{G(y)}\,.$$

Первый интеграл, очевидно, равен $G^2(u)$. Во втором интеграле можно сначала осуществить замену $G(y) \dashrightarrow x$, $G(u) \leqslant \leqslant x \leqslant G(z)$, см. 86, стр. 105. В результате, после простых вычислений, получим подтверждение (32). В случае $u > z$ ф.р. $F(u,z) = G^2(z)$, а в предыдущем уравнении останется только первый интеграл (здесь уже до z), что снова приведёт к равенству в (32). ⊚

Совместное распределение через условное. Равенство (32) можно использовать и в обратном направлении. Часто ф.р. $F_\xi(x\,|\,y)$ наблюдаемой в эксперименте сл.в. ξ зависит от некоторых параметров y. Если, в свою очередь, значения y представляют собой реализации некоторого сл.вектора η, то (32) даёт способ определения совместного распределения (ξ,η) (аналог формулы Фубини).

54| Т е о р е м а. *Пусть функция $G(x\,|\,y)$ при каждом $y \in \in \mathbb{R}$ есть ф.р., а при каждом фиксированном x она измерима как функция переменной y; F_0 — ф.р. некоторой сл.в. η. Тогда*

i) *функция*

$$F(x,y) = \int_{-\infty}^{y} G(x\,|\,u)\,dF_0(u)$$

задаёт ф.р. двумерного сл.вектора (ξ,η);

ii) *для любой неотрицательной функции $h(x,y)$* [‡]

$$\mathbf{E}\big[h(\xi,\eta)\big] = \int\limits_{y\in\mathbb{R}} \Big[\int\limits_{x\in\mathbb{R}} h(x,y)\,dG(x\,|\,y)\Big]\,dF_0(y),$$

(‡) Запись вида $\int_{u\in A}$ используется для указания на u как переменную интегрирования.

в частности, при $\forall\, A, B \in \mathscr{B}$

$$\mathbf{P}\{\xi \in A, \eta \in B\} = \int\limits_{y\in B} \Big[\int\limits_{x\in A} dG(x\,|\,y)\Big]\, dF_0(y).$$

⊫ i) Свойства нормированности и непрерывности ф.р. легко проверяются с привлечением теоремы Лебега о мажорируемой сходимости. Свойство монотонности следует из представления

$$\Delta_{\vec{a},\vec{c}} F(x,y) = \int_{a_1}^{c_1} \big(G(c_2\,|\,t) - G(a_2\,|\,t)\big)\, dF_0(t) \geqslant 0.$$

Доказательство части ii) теоремы почти полностью повторяет доказательство теоремы Фубини (см., например, монографию Ж. Невё [7]). ⫤

55| 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*. Тогда для любой ф.плотности* $f_\eta(y), y \in \mathbb{R}^1$*, относительно меры* μ_2 *на* $(\mathbb{R}^1, \mathscr{B}^1)$ *функция*

$$f(x,y) = f(x\,|\,y) f_\eta(y)$$

есть ф.плотности распределения некоторого сл.вектора (ξ, η) *относительно произведения мер* $\mu_1 \times \mu_2$ *на* $(\mathbb{R}^2, \mathscr{B}^2)$*, при этом*

$$\mathbf{E}\big[h(\xi,\eta)\big] = \int\limits_{y\in\mathbb{R}} \Big[\int\limits_{x\in\mathbb{R}} h(x,y) f(x\,|\,y)\, d\mu_1(x)\Big] f_\eta(y)\, d\mu_2(y).$$

56| 𝔜𝔭𝔭. Докажите следствие через теорему 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 $f_{\vec{\eta}}(y_1, y_2, y_3) =$

$= \int_{\mathbb{R}^2} f(x_1, x_2, y_1, y_2, y_3)\, dx_1 dx_2$ — частная плотность $\vec{\eta} = (\eta_1, \eta_2, \eta_3)$ и условное матем.ожидание при фиксированном значении $\vec{\eta} = \vec{y}$ любой интегрируемой (относительно совместной плотности f) функции $h(x_1, x_2, y_1, y_2, y_3)$ равно

$$\mathbf{E}\big[\, h(\vec{\xi}, \vec{\eta}) \,|\, \vec{y}\,\big] = \frac{1}{f_{\vec{\eta}}(\vec{y})} \int_{\mathbb{R}^2} h(\vec{x}, \vec{y}) f(\vec{x}, \vec{y})\, d\vec{x}.$$

57| Пример. В статистическом эксперименте наблюдается последовательность независимых сл.в. $\xi_1, \ldots, \xi_n$ с общей ф.р. $F(x\,|\,\theta)$, зависящей от параметра θ; цель исследования — принятие решения о параметре 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 совместное распределение вектора $\vec{\xi} = (\xi_1, \ldots, \xi_n)$ и параметра ϑ описывается ф.р.

$$F(\vec{x}, \theta) = \int\limits_{-\infty}^{\theta} \prod_{i=1}^{n} F(x_i \,|\, u)\, dG(u).$$

Оптимальное в байесовском подходе решение всегда основывается на так называемом *апостериорном распределении* — условном распределении ϑ при фиксированном значении наблюдаемого вектора $\vec{\xi} = \vec{x}$. ⊚

§ 8. Наилучший нелинейный прогноз

Ранее мы нашли линейную функцию, наилучшим образом прогнозирующую реализации одной сл.в. по другой сл.в. Решим аналогичную задачу в классе всех функций, воспользовавшись теорией условных распределений. Рассмотрим две сл.в. ξ, η и найдём функцию h, для которой значения $h(\eta)$ наилучшим образом прогнозируют значения ξ в смысле достижения минимума средне-квадратической ошибки прогноза $\mathbf{E}\big(\xi - h(\eta)\big)^2$. Пусть $F_{\xi|\eta}(x \,|\, y)$ — какой-либо из вариантов усл.ф.р. ξ при фиксированном η. Определим усл.м.о. относительно этой усл.ф.р.

$$\mu^*(y) = \mathbb{E}\left[\xi \,|\, \eta = y\right] = \int\limits_{x \in \mathbb{R}} x \, dF_{\xi|\eta}(x \,|\, y).$$

58 Теорема. *Если* $\mathbf{E}\xi^2 < \infty$, *то функция* $\mathbb{E}\left[\xi \,|\, \eta = y\right]$ *минимизирует средне-квадратическую ошибку прогноза значений сл.в.* ξ *по значениям* η.

⊫ Ввиду (31) $\mathbb{E}\left[\xi^2 \,|\, \eta = y\right] < \infty$, а значит, и $|\mu^*(y)| < \infty$, при почти всех $y \in \mathbb{R}^1$. По известному свойству матем.ожидания $\mathbf{E}\big[\,(\xi - \mu^*(y))^2 \,|\, \eta = y\,\big] \leqslant \mathbf{E}\big[\,(\xi - C)^2 \,|\, \eta = y\,\big]$ для любой константы C (относительно значений сл.в. ξ). В частности $\mathbf{E}\big[\,(\xi - \mu^*(y))^2 \,|\, \eta = y\,\big] \leqslant \mathbf{E}\big[\,(\xi - h(y))^2 \,|\, \eta = y\,\big]$ для любого $h(y)$. Таким образом, снова в силу (31)

$$\mathbf{E}\big(\xi - \mu^*(\eta)\big)^2 = \mathbf{E}\Big[\mathbb{E}\left[(\xi - \mu^*(\eta))^2 \,|\, \eta\right]\Big] \leqslant \mathbf{E}\big(\xi - h(\eta)\big)^2. \qquad ⫤$$

Определение. Условное матем.ожидание $\mathbb{E}\left[\xi \,|\, \eta = y\right]$ называется *(средне-квадратической) регрессией* ξ *на* η.

⚠ Теорема 58 даёт ещё один способ определения усл.м.о. сл.величины ξ при фиксированном значении $\eta = y$ как такой функции $h^*(y)$, для которой достигается минимум (по всем измеримым функциям h) среднего $\mathbf{E}\big(\xi - h(\eta)\big)^2$. Существование минимума легко устанавливается с привлечением теории гильбертовых пространств к пространству интегрируемых с квадратом функций. Единственное неудобство этого определения — необходимость предположения конечности второго момента у ξ, что, конечно, не существенно для индикаторных функция $\xi = \mathbf{I}_B$.

59| Примеры. 1) Так как для независимых сл.в. совместная плотность равна произведению частных плотностей, то при любом фиксированном $\xi_2 = x_2$ условная плотность ξ_1 равна частной (безусловной) плотности ξ_1; условное среднее $\mathbb{E}\left[\xi_1 \,|\, \xi_2\right]$ совпадает с матем.ожиданием $\mathbf{E}\xi_1$.

2) Плотность двумерной нормальной модели (16), стр. 46, после элементарных преобразований можно записать в виде

$$\begin{aligned} f(x_1, x_2) &= \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left\{-\frac{1}{2(1-\rho^2)}(x_1 - \rho x_2)^2 - \frac{1}{2}x_2^2\right\} = \\ &= \phi(x_2 \,|\, 0, 1)\ \phi(x_1 \,|\, \rho x_2, (1-\rho^2)), \end{aligned}$$

где $\phi(x \,|\, \mu, \sigma^2) = \exp\{-(x-\mu)^2/2\sigma^2\}/(\sqrt{2\pi}\sigma)$ — одномерная нормальная плотность со средним μ и дисперсией σ^2.

Отсюда ясно, что, т.к. $\int_{\mathbb{R}^1} \phi(x_1 \,|\, \rho x_2, (1-\rho^2))\, dx_1 = 1$, частная плотность $f_2(x_2) = \phi(x_2 \,|\, 0, 1)$. Поэтому при фиксированном $\xi_2 = x_2$ условное распределение $\xi_1\big|_{\xi_2 = x_2} \overset{d}{\backsim} \phi(\cdot \,|\, \rho x_2, (1 - \rho^2))$ — нормально с соответствующими параметрами. Таким образом, регрессия $\mu_{1|2}(x_2) = \rho x_2$, т.е. совпадает с линейной.

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 всегда равно 0. Снова мы получили регрессию, совпадающую с линейной регрессией: $\mu_{1|2}(x_2) = 0 + 0x_2$. Если вспомнить общий вид линии регрессии (теорема 36, стр. 51), то отсюда можно сделать вывод, что коэф.корр. $\rho = 0$. При этом сл.в. ξ_1, ξ_2 не являются независимыми, поскольку условная плотность одной сл.в. зависит от зафиксиро-

ванного значения другой сл.в.

4) Из конструкции многомерного гипергеометрического распределения $\mathcal{H}\mathrm{g}_k(n; N_1, \ldots, N_k)$ (стр. 17) ясно, что если в выборке зафиксировать значение числа 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 при фиксированных значениях $\xi_2 = n_2, \ldots, \xi_J = n_J$ $(J < k)$

$$\mathbb{E}\big[\xi_1 \,|\, \xi_2 = n_2, \ldots, \xi_J = n_J\big] = \frac{N_1\big(n - (n_2 + \ldots + n_J)\big)}{N - (N_2 + \ldots + N_J)} \,.$$

Таким образом, регрессия линейно зависит от условий. Заметим, что это выражение совпадает с линейной регрессией (24), если обозначить $q_r = N_r / N$.

Аналогичное верно и для мультиномиальной модели (стр. 18).

5) Если у вектора $\vec{\xi} \overset{d}{\backsim} \mathcal{D}\mathtt{ir}_k(a_1, \ldots, a_{k+1})$, имеющего распределение Дирихле (17), зафиксирована одна компонента $\xi_k = p_k$ $(\in (0; 1))$, то область возможных значений остальных компонент сужается: $\xi_1 + \ldots + \xi_{k-1} \leqslant 1 - p_k$. Чтобы привести условное распределение к стандартному виду, следует рассмотреть нормированный вектор $(\xi_1, \ldots, \xi_{k-1})/(1 - p_k)$. Так как маргинальное распределение $\xi_k \overset{d}{\backsim} \mathcal{B}\mathtt{et}(a_k, \boldsymbol{a} - a_k)$ $(\boldsymbol{a} = a_1 + \ldots + a_{k+1})$, то простыми арифметическими выкладками легко показывается,

что условное распределение

$$\frac{1}{(1-p_k)}(\xi_1,\ldots,\xi_{k-1})\Big|_{\xi_k=p_k} \overset{d}{\backsim} \mathtt{Dir}_{(k-1)}(a_1,\ldots,a_{k-1},a_{k+1}).$$

Таким образом, условное матем.ожидание $\mathbb{E}\left[\xi_1 \,|\, \xi_k\right] = (1 - \xi_k)a_1/(\boldsymbol{a} - a_k)$.

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. $\xi \overset{d}{\backsim} \mathtt{Bin}(n,\theta)$ 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. В иллюстративных целях, не вдаваясь в подробности механизма образования, предположим, что доля θ в каждой партии априори (до опыта) представляет собой реализацию бета-сл.в. $\vartheta \overset{d}{\backsim} \mathtt{Bet}(a_1,a_2)$. Таким образом, совместное распределение вектора (ξ,ϑ) принадлежит к составному типу (одна компонента дискретна, вторая непрерывна). Вместо ф.р. такое распределение удобнее описывать с помощью функции

$$\mathbf{P}\left\{\xi=k,\vartheta\leqslant\theta\right\} = \frac{1}{\mathrm{B}(a_1,a_2)}\int_0^\theta \mathrm{C}_n^k p^{k+a_1-1}(1-p)^{n-k+a_2-1}\,dp =$$
$$= \frac{\mathrm{C}_n^k\,\mathrm{B}(k+a_1,n-k+a_2)}{\mathrm{B}(a_1,a_2)}\,\mathcal{B}(\,\theta\,;\,k+a_1,n-k+a_2),$$

где $\theta\in[0;1]$, $k=0,1,\ldots n$, функция $\mathcal{B}(\cdot\,;\,r,s)$ есть ф.р. бета-закона с параметрами r,s. Отсюда легко понять, что условное распределение $\vartheta\big|_{\xi=k} \overset{d}{\backsim} \mathtt{Bet}(k+a_1, n-k+a_2)$. Следовательно, регрессия ϑ на ξ, доставляющая по теореме 58 наилучшую в средне-квадратическом оценку θ по результатам контрольных испытаний, равна

$$\hat{\theta} = \mathbb{E}\left[\vartheta \,|\, \xi = k\right] = \frac{k+a_1}{n+a_1+a_2} = \frac{k}{n}\,q_n + \frac{a_1}{a_1+a_2}(1-q_n),$$

где $q_n = n/(n+a_1+a_2)$. Последнее равенство показывает, что оценка θ есть выпуклая комбинация выборочной относительной частоты k/n и априорного среднего $a_1/(a_1+a_2)$. ◎

А. Указания к решению задач

✦ 1, стр. 14. Функция ξ измерима относительно σ-алгебры $\mathscr{F} = \sigma(\mathcal{F})$, порождённой некоторым классом подмножеств $\mathcal{F}$, т. т. т. когда прообраз любого борелевского множества $\xi^{-1}(B) \in \mathcal{F}$. Далее для каждого из пунктов леммы воспользоваться указанными ссылками со специфическим классом $\mathcal{F}$.

✦ 3, стр. 19. а) Гипергеометрическая модель: общее число исходов — C_N^n (из N объектов выбирается n), в «благоприятные» попадают любые n_1 из первой группы ($\mathrm{C}_{N_1}^{n_1}$ вариантов), …, n_k из k-ой группы ($\mathrm{C}_{N_k}^{n_k}$ 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 при $M \to \infty$ и фиксированном $m > 0$ отношение $M^{-m} M!/_{(M-m)!} \to 1$.

✦ 7, стр. 22. Воспользоваться (F_4); (F_1) справедливо для любой ф.р.

✦ 8, стр. 22. Найти прямоугольник $(\vec{a}; \vec{c}]$, для которого

$\Delta_{\vec{a};\vec{c}} G < 0$.

✦ 10, стр. 24. $\mathbf{P}\{\,\xi_1 \leqslant x_1\,\} = \mathbf{P}\{\,\xi_1 \leqslant x_1, \xi_2 \in \mathbb{R}^1\,\}$.

✦ (5), стр. 25. $\mathbf{P}\{\xi_1 = n\} = \sum_{m \in \mathscr{X}_2} \mathbf{P}\{\xi_1 = n, \xi_2 = m\}$.

✦ 12, стр. 26. Оценить количество точек $\bigcup_n \langle\, x \in \mathbb{R}^1 : \mathbf{P}\{\,\xi = x\,\} \geqslant {}^1/_n \rangle$.

✦ 14, стр. 28. i) Границу углового множества представить в виде монотонного предела $\lim\limits_{z\downarrow 0}{}_{\downarrow} \{\vec{x} : \vec{x} \leqslant \vec{x}_0\} \setminus \{\vec{x} : \vec{x} \leqslant \leqslant \vec{x}_0 - z\}$. 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 $\exists b :\ F(a, b) = \mathbf{P}\{\xi \leqslant a, \eta \leqslant b\} > \mathbf{P}\{\xi \leqslant a\} - {}^q/_2$.

✦ (8), стр. 32. Проверить свойства вероятности.

✦ 20, стр. 33. 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 $\vec{\eta}_1, \ldots, \vec{\eta}_n$.

✦ 22, стр. 34. $\forall\, A_j \in \sigma(\vec{\xi}_j)\ \exists B_j \in \mathscr{B}^{m_j} :\ A_j = \vec{\xi}_j^{\,-1}(B_j)$.

✦ (?!), стр. 35. Показать: $\mathbf{P}\{(A_1 \uplus A_2)B\} = \mathbf{P}\{A_1 B\} + + \mathbf{P}\{A_2 B\} =$ (независимость) $= \mathbf{P}\{A_1\}\,\mathbf{P}\{B\} + \mathbf{P}\{A_2\}\,\mathbf{P}\{B\} = \mathbf{P}\{A_1 \uplus A_2\}\,\mathbf{P}\{B\}$.

✦ 27, стр. 41. Воспользоваться доказательством свойства (F_2) для ф.р.

✦ 13, стр. 42. (1) $\mathbf{P}\{\xi = \eta\} = 0$; (2) $(\xi, \eta) \overset{d}{\backsim} (\eta, \xi) \Rightarrow \mathbf{P}\{\xi < \eta\} = \mathbf{P}\{\xi > \eta\}$.

✦ 34, стр. 50. Применить свойство линейности матем.ожидания.

✦ 35, стр. 51. б) $\mathbf{D}(\xi_2 + (\xi_1 - \xi_2)) = \mathbf{D}\xi_2 + \mathbf{D}(\xi_1 - \xi_2)$.

✦ (19), стр. 52. $\mathbf{Cov}(\xi_2, \xi_1 - \rho\sigma_1 {}^{\xi_2}/_{\sigma_2}) = \sigma_{12} - \rho\sigma_1\sigma_2 = 0$.

✦ 38, стр. 53. В матрице $\mathbf{E}\big[A\Xi\big]$ элемент на пересечении i-ой строки j-ого столбца 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. совпадает с (i, j)-ым элементом матрицы $A\,\mathbf{E}\big[\Xi\big]$.

✦ 44, стр. 58. Свойство (R_4) и неравенство $\rho_{\zeta\boldsymbol{.}\xi} \leqslant 1$ 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). (R_2) Для вычисления коэф.корреляции надо найти дисперсии компонент и их ковариацию. Вывести из некоррелированности остатка $\zeta - Z(\vec{\xi})$ и компонент $\vec{\xi}$, что ковариация $\mathbf{Cov}(\zeta, Z(\vec{\xi})) = \mathbf{D}Z(\vec{\xi})$. Найти вид этой дисперсии из предшествующей теоремы.

✦ 49, стр. 73. Применить формулу для маргинальной плотности.

✦ 52, стр. 76. Применить формулу условной плотности.

✦ (31), стр. 75. В записи $\mathbf{E}\big[g(\eta)h(\xi, \eta)\big]$ через совместную плотность воспользоваться условной плотностью; применить теорему Фубини–Тонелли.

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. В этой главе мы изучим способы отыскания распределений некоторых функций от случайных величин с известным распределением.

§ 1. Индуцированное распределение

Пусть $\vec{\xi} = (\xi_1, \ldots, \xi_k)$ — сл.вектор, для которого известен закон распределения, т.е. известны вероятности $\mathbb{P}_{\vec{\xi}}\{A\} \overset{\mathrm{df}}{=} \mathbf{P}\{\vec{\xi} \in \in A\}$ 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

$$\mathbb{P}_{\vec{\eta}}\{B\} = \mathbb{P}_{\vec{\xi}}\{h^{-1}(B)\}, \qquad B \in \mathscr{B}^m.$$

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.

60| Пример. Пусть $F_\xi(x)$ — ф.р. сл.величины ξ и $h(x) = b\,x + a$ с некоторыми $a \in \mathbb{R}^1$ и $b > 0$. Заметим, что обратная функция $h^{-1}(y) = \frac{1}{b}\,(y-a)$ и неравенство $h(x) \leqslant y$ эквивалентно $x \leqslant \frac{1}{b}\,(y-a)$. Поэтому ф.р. сл.величины $\eta = b\,\xi + a$ равна

$$F_\eta(y) \overset{\mathrm{df}}{=} \mathbf{P}\{\eta \leqslant y\} = \mathbf{P}\Big\{\xi \leqslant \frac{1}{b}(y-a)\Big\} = F_\xi\left(\frac{1}{b}(y-a)\right).$$

Если распределение ξ имеет плотность $f_\xi(x)$ с носителем $\mathscr{X}_\xi = (L; R)$, с конечными или бесконечными границами, то

$$F_\eta(y) \;=\; F_\xi\left(\frac{1}{b}(y-a)\right) \;=\; \int_{-\infty}^{\frac{y-a}{b}} f_\xi(x)\,\mathbf{I}_{(L;R)}(x)\,dx.$$

Замена переменной $x = h^{-1}(t) = \frac{1}{b}\,(t-a)$, $-\infty < t < y$, в этом интеграле переводит индикаторную функцию $\mathbf{I}_{(L;R)}(x)$ в индикаторную функцию $\mathbf{I}_{(U;V)}(x)$, где $U = b\,L + a$ $(= \lim_{x\to L}(b\,x + a))$ и $V = b\,R + a$. В силу возможности линейной замены в интеграле Лебега

$$F_\eta(y) \;=\; \int\limits_{-\infty}^{y} \underbrace{\frac{1}{b}\,f_\xi\left(\frac{1}{b}(t-a)\right)}_{f_\eta(t)}\,\mathbf{I}_{(U;V)}(t)\,dt.$$

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 $\xi \overset{d}{\backsim} \mathcal{N}(0,1)$, то $\sigma\xi + \mu \overset{d}{\backsim} \mathcal{N}(\mu, \sigma^2)$. ◎

61| $\mathfrak{Y}\mathfrak{np}$. Как изменятся ф.р. и ф.пл. $b\xi + a$, если $b < 0$?

Плотность монотонного преобразования. Плотность монотонного (возрастающего) преобразования $\eta = h(\xi)$ можно найти по аналогичной схеме — сначала ф.р. η представляется через интеграл от плотности: $F_\eta(y) = \mathbf{P}\{\eta \leqslant y\} = \mathbf{P}\{\xi \leqslant h^{-1}(y)\} = \int_{-\infty}^{h^{-1}(y)} f_\xi(t)\,dt$; затем производится замена переменной $t = \widetilde{h}(u) = h^{-1}(u)$, $u \leqslant y$, какова́я с учётом якобиана $dt = \widetilde{h}'(u)du$ приводит ф.р. к виду $F_\eta(y) = \int_{-\infty}^{y} f_\xi(\widetilde{h}(u))\widetilde{h}'(u)\,du$. Подынтегральное выражение даст ф.плотности η.

Следующая теорема содержит формальное обоснование этой схемы. Чтобы избежать излишних подробностей будем понимать здесь под носителем распределения любой открытый интервал, вне которого плотность равна нулю.

62| 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.*

$$f_\eta(y) = \left|(h^{-1}(y))'\right| f_\xi\big(h^{-1}(y)\big), \qquad y \in (C;D).$$

⊫ Пусть h возрастает. Для любого конечного интервала $(u;v] \subset (C;D)$ в силу монотонности и непрерывности (а значит и взаимной однозначности) обратной функции $\widetilde{h} = h^{-1}$

$$\mathbf{P}\{u < \eta \leqslant v\} = \mathbf{P}\big\{\,\widetilde{h}(u) < \xi \leqslant \widetilde{h}(v)\,\big\} = \int_{\widetilde{h}(u)}^{\widetilde{h}(v)} f_\xi(x)\,dx = \\ = \int_{\widetilde{h}(u)}^{\widetilde{h}(v)} f_\xi\big(\widetilde{h}(h(x))\big)\,dx.$$

Замена $t = h(x)$ по формуле замены переменных в интеграле Лебега (163 , стр. 180) переводит этот интеграл относительно меры Лебега λ в интеграл относительно индуцированной меры μ_h, которая на любом конечном интервале $(k;m]$ равна

$$\mu_h(k;m] = \lambda\big(h^{-1}(k); h^{-1}(m)\big] = \widetilde{h}(m) - \widetilde{h}(k).$$

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,

$$\mathbf{P}\{u < \eta \leqslant v\} = \int_u^v f_\xi\big(\widetilde{h}(t)\big)\,\mu_h(dt) = \int_u^v f_\xi\big(\widetilde{h}(t)\big)\,\widetilde{h}'(t)\,dt.$$

Такое представление ф.р. сл.величины η в виде интеграла эквивалентно абсолютной непрерывности распределения η с заявленной в теореме плотностью. Случай убывающей функции h аналогичен рассмотренному. ⫷

63| П р и м е р. Пусть $\xi \overset{d}{\backsim} \mathfrak{U}\mathrm{n}(-1,1)$, тогда плотность $f_\xi(x) = {}^1/_2$ при $x \in (-1;1)$. 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 задействованы при описании плотности, значение этой плотности предполагается равным нулю. ◎

Метод доказательства предыдущей теоремы может быть применён и к кусочно–монотонным функциям.

64| П р и м е р. Пусть $\xi \overset{d}{\backsim} \mathfrak{U}\mathrm{n}(-1,1)$ с ф.плотности $f(x) = {}^1/_2$, $x \in [-1;1]$. Функция $h(x) = x^2$ переводит отрезок

$[-1;1]$ в $[0;1]$. Вероятность

$$\mathbf{P}\{u \leqslant \xi^2 \leqslant v\} = \mathbf{P}\{-\sqrt{v} \leqslant \xi \leqslant -\sqrt{u}\} + \mathbf{P}\{\sqrt{u} \leqslant \xi \leqslant \sqrt{v}\} =$$
$$= \frac{1}{2}\int_{-\sqrt{v}}^{-\sqrt{u}} f(x)\,dx + \frac{1}{2}\int_{\sqrt{u}}^{\sqrt{v}} f(x)\,dx = \int_{\sqrt{u}}^{\sqrt{v}} dx.$$

для $\forall\, 0 < u \leqslant v < 1$. Замена $x = \sqrt{t}$, $t \in [u;v]$, в силу её непрерывной дифференцируемости на интервалах, не содержащих ноль, даёт

$$\mathbf{P}\{u \leqslant \xi^2 \leqslant v\} = \int_u^v \frac{1}{2\sqrt{t}}\,dt.$$

Итак, плотность $\eta = \xi^2$ равна $1/(2\sqrt{y})$, $0 < y < 1$. ⊚

Другой способ получения плотности связан с утверждением 26, стр. 38. Пусть h — возрастающая функция, плотность f_ξ всюду непрерывна. Представим ф.р. $\eta = h(\xi)$ в виде

$$F_\eta(y) = \mathbf{P}\{\eta \leqslant y\} = \mathbf{P}\{\,\xi \leqslant \widetilde{h}(y)\,\} = \int_{-\infty}^{\widetilde{h}(y)} f_\xi(x)\,dx.$$

По основной теореме анализа и правилу дифференцирования сложной функции

$$\frac{d}{dy}F_\eta(y) = f_\xi(\widetilde{h}(y))\widetilde{h}'(y).$$

Полученное выражение будет плотностью распределения η, если обратная функция $\widetilde{h}$ непрерывно дифференцируема. В случае кусочной непрерывности f и (или) $\widetilde{h}'$ абсолютную непрерывность F_η можно доказать, проверив равенство $\int_{\mathbb{R}} \left({}^{d}\!/_{dy} F_\eta(y)\right) dy = 1$ (см. 26, стр. 38).

65| Пример. Пусть $\xi \overset{d}{\sim} \mathcal{N}(0,1)$, найдём распределение $\chi_1^2 = \xi^2$. Искомая ф.р. равна (при $y > 0$)

$$F(y) = \mathbf{P}\{\xi^2 \leqslant y\} = \Phi(\sqrt{y}\,) - \Phi(-\sqrt{y}\,) = 2\Phi(\sqrt{y}\,) - 1.$$

Ясно, что $F(y) = 0$ при $y \leqslant 0$. Функция F, очевидно, всюду непрерывна и её производная (при $y > 0$) равна

$$F'(y) = \frac{1}{\sqrt{2\pi}} y^{1/2-1} \exp\left(-\tfrac{1}{2}y\right) = \frac{(1/2)^{1/2}}{\Gamma(1/2)} y^{1/2-1} \exp\left(-\tfrac{1}{2}y\right).$$

Равенство $\int_{\mathbb{R}} F'(y)\,dy = 1$ проверять не надо, т.к. F' совпадает с плотностью гамма-распределения $\mathcal{G}\mathtt{am}\left({}^1\!/_2\,, 2\right)$ с параметром формы ${}^1\!/_2$ и параметром масштаба 2 (см. стр. 249). ◎

Плотность функции вектора случайных величин. Аналогичные схемы используются и при многомерных преобразованиях.

66| Т е о р е м а. *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 h^{-1}*

$$J_{h^{-1}}(\vec{y}) = \left\| \left(\frac{\partial \widetilde{h}_j}{\partial y_l} \right)_{j,l} \right\| > 0 \qquad \text{для } \forall\, \vec{y} \in \mathscr{Y}^o.$$

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

$$f_{\vec{\eta}}\,(\vec{y}) \;=\; J_{h^{-1}}(\vec{y})\, f_{\vec{\xi}}\big(h^{-1}(\vec{y})\big), \qquad \vec{y} \in \mathscr{Y}^o.$$

67| П р и м е р ы. 1) Пусть $\xi_1, \xi_2 \overset{d}{\backsim} \mathcal{E}\mathrm{x}(1)$ — 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 $x_1 = y_1 y_2$, $x_2 = y_1 - y_1 y_2$, $0 < y_1 < \infty$, $0 < y_2 < 1$, равен

$$J_{H^{-1}} = \left\| \begin{matrix} y_2 & y_1 \\ 1-y_2 & -y_1 \end{matrix} \right\| = y_1 > 0.$$

Так как плотность $\vec{\xi} = (\xi_1, \xi_2)$ равна $f_{\vec{\xi}}(x_1, x_2) = e^{-(x_1+x_2)}$, то

$$f_{\vec{\eta}}(y_1, y_2) = J_{H^{-1}}\, e^{-(x_1+x_2)} = y_1 e^{-y_1} \times 1, \qquad y_1 > 0,\ 0 < y_2 < 1.$$

Другими словами, сл.в. η_1, η_2 независимы, $\eta_1 \overset{d}{\backsim} \mathcal{G}\mathtt{am}(2,1)$ с плотностью $f_{\eta_1}(y_1) = y_1 e^{-y_1}$, $\eta_2 \overset{d}{\backsim} \mathcal{U}\mathtt{n}(0;1)$ с плотностью $f_{\eta_2}(y_2) = 1$.

2) [*Генерирование нормальных сл.в.*] Пусть независимые равномерные сл.в. $\xi_1, \xi_2 \overset{d}{\backsim} \mathcal{U}\mathtt{n}(0,1)$. Покажем, что сл.в.

$$\eta_1 = \sqrt{-2\ln\xi_1}\cos(2\pi\xi_2), \quad \eta_2 = \sqrt{-2\ln\xi_1}\sin(2\pi\xi_2)$$

также независимы и имеют стандартное нормальное $\mathcal{N}(0,1)$ распределение. Будем считать, что $\xi_1, \xi_2 \in (0;1)$, тогда обратное преобразование определяется формулами

$$x_1 = \exp\big(-\tfrac{1}{2}(y_1^2 + y_2^2)\big), \quad x_2 = \frac{1}{2\pi}\mathrm{Atan}(y_2, y_1), \qquad y_1, y_2 \neq 0,$$

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 66 выполнены, т.к. якобиан обратного преобразования

$$J = \left| \begin{vmatrix} -y_1 e^{-(y_1^2+y_2^2)/2} & -y_2 e^{-(y_1^2+y_2^2)/2} \\ -\frac{y_2}{2\pi(y_1^2+y_2^2)} & \frac{y_1}{2\pi(y_1^2+y_2^2)} \end{vmatrix} \right| = \frac{1}{2\pi} e^{-(y_1^2+y_2^2)/2}$$

не обращается в нуль в области $y_1, y_2 \neq 0$. Плотность вектора (ξ_1, ξ_2) равна 1 при $0 < x_1, x_2 < 1$, поэтому плотность (η_1, η_2) равна $\dfrac{1}{2\pi} e^{-(y_1^2+y_2^2)/2}$, что и требовалось. ⊚

Особенно просто выглядит плотность линейного преобразования.

68| Лемма. (?!) *Пусть Q — невырожденная $(k \cdot k)$-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* $f_{\vec{\eta}}(\vec{y}) = \dfrac{1}{\|Q\|}\, f_{\vec{\xi}}(Q^{-1}(\vec{y} - \vec{a}))$.

Некоторые соображения к доказательству леммы, не связанные с теоремой 66, приведены в дополнении к этой главе (стр. 110).

69| $\mathfrak{Упр}$. 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.

70| Пример. Пусть $\vec{\xi} \overset{d}{\backsim} \mathcal{D}\mathtt{ir}_2(a_1, a_2, a_3)$ с плотностью (17), стр. 46. Вектор $(\xi_1 + \xi_2, \xi_2)$ получающийся 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 $(y_1, y_2) \in \mathscr{Y}$ равна

$$f(y_1, y_2) = \frac{\Gamma(a_1 + a_2 + a_3)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)} (y_1 - y_2)^{a_1 - 1} y_2^{a_2 - 1} (1 - y_1)^{a_3 - 1}.$$

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 $\xi_1 + \xi_2 \overset{d}{\backsim} \mathcal{B}\mathtt{et}(a_1 + a_2, a_3)$. Ясно, что справедливо и более общее:

$$(\xi_1, \ldots, \xi_k) \overset{d}{\backsim} \mathcal{D}\mathtt{ir}_k(a_1, \ldots, a_{k+1}) \quad \Rightarrow$$

$$\Big(\sum_{i=1}^{r_1} \xi_i, \ldots, \sum_{i=1}^{r_m} \xi_i\Big) \overset{d}{\backsim} \mathcal{D}\mathtt{ir}_m\Big(\sum_{i=1}^{r_1} a_i, \ldots, \sum_{i=1}^{r_m} a_i, a_{k+1}\Big)$$

для любых натуральных $r_1 + \ldots + r_m = k, \quad 1 \leqslant m \leqslant k.$ ⊚

§ 2. Распределение суммы случайных величин. Свёртка

Для того чтобы найти ф.р. суммы сл.величин, непосредственно воспользоваться приведёнными выше утверждениями нельзя, поскольку отображение $(x_1, \ldots, x_k) \mapsto \sum_1^k x_j$ не является взаимно-однозначным. В этом случае, как в примере 70, можно попытаться найти взаимно-однозначное отображение большей размерности, в которое искомое отображение будет входить как одна из компонент

Рассмотрим преобразование $y_1 = \sum_1^k x_j,\ y_i = x_i,\ i = \overline{2,k}$, с обратным $x_1 = y_1 - \sum_2^k y_j, x_i = y_i,\ i = \overline{2,k}$. Якобиан матрицы преобразования равен 1. Поэтому в силу леммы 68 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 29, стр. 43, о частной плотности имеем.

71| Т е о р е м а. *Если $f(x_1, \ldots, x_k)$ — ф.пл. сл.вектора $(\xi_1, \ldots, \xi_k)$, то плотность распределения суммы $S = \xi_1 + \ldots + \xi_k$*

$$f_S(t) = \int_{\mathbb{R}^{k-1}} f(t - y_2 - \ldots - y_k,\ y_2, \ldots, y_k)\, dy_2 \cdots dy_k.$$

72| $\mathfrak{Y}$пр. Докажите, что

$$f_S(t) = \int_{\mathbb{R}^{k-1}} f(y_1, \ldots, y_{k-1},\ t - y_1 - \ldots - y_{k-1})\, dy_1 \cdots dy_{k-1}.$$

Наиболее востребован случай с независимыми слагаемыми.

О п р е д е л е н и е. Ф.р. (ф.плотности) 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

$$F(\vec{u}) \;=\; \int\limits_{\mathbb{R}^m} \mathbf{I}_{\vec{u}}(\vec{y})\, dF(\vec{y}) \;=\; \int\limits_{\mathbb{R}^m} \mathbf{I}_{\vec{u}}(\vec{y})\, f(\vec{y})\, d\vec{y},$$

где $\mathbf{I}_{\vec{u}}(\vec{y})$ — индикаторная функция множества $\{\vec{y} : \ \vec{y} \leqslant \vec{u}\}$, и последнее равенство верно для абсолютно-непрерывных распределений с плотностью f.

73 Теорема. [О свёртке.] I) *Свёртка двух ф.р.* F_1, F_2

$$F_1 * F_2\,(\vec{u}) \;=\; \int\limits_{\mathbb{R}^k} F_1(\vec{u}-\vec{y})\, dF_2(\vec{y}) = \int\limits_{\mathbb{R}^k} F_2(\vec{u}-\vec{x})\, dF_1(\vec{x}). \quad (1)$$

II) *Если* f_1 *— плотность* F_1, *то*

$$f_1 * F_2\,(\vec{u}) \;=\; \int\limits_{\mathbb{R}^k} f_1(\vec{u}-\vec{y})\, dF_2(\vec{y}) \quad (2)$$

есть плотность свёртки $F_1 * F_2$.

III) *Если обе ф.р.* F_1, F_2 *имеют плотности* f_1 *и* f_2, *то плотность их свёртки равна*

$$f_1 * f_2\,(\vec{u}) \;=\; \int\limits_{\mathbb{R}^k} f_1(\vec{u}-\vec{y})\, f_2(\vec{y})\, d\vec{y} \;=\; \int\limits_{\mathbb{R}^k} f_2(\vec{u}-\vec{y})\, f_1(\vec{y})\, d\vec{y}. \quad (3)$$

⊫ I) Пусть $\vec{\xi} \overset{d}{\backsim} F_1(\vec{x})$, $\vec{\eta} \overset{d}{\backsim} F_2(\vec{y})$. По условию теоремы совместное распределение $(\vec{\xi}, \vec{\eta})$ равно $F_1(\vec{x})F_2(\vec{y})$. В силу теоремы Фубини

$$F_1 * F_2\,(\vec{u}) = \mathbf{P}\big\{\, \vec{\xi} + \vec{\eta} \leqslant \vec{u}\,\big\} = \int_{\mathbb{R}^{2k}} \mathbf{I}_{\vec{u}}(\vec{x}+\vec{y})\, dF_1(\vec{x})F_2(\vec{y}) =$$

$$= \int_{\mathbb{R}^k}\left(\int_{\mathbb{R}^k} \mathbf{I}_{\vec{u}}(\vec{x}+\vec{y})\, dF_1(\vec{x})\right) dF_2(\vec{y}) \;=\; \int_{\mathbb{R}^k} F_1(\vec{u}-\vec{y})\, dF_2(\vec{y}).$$

Последнее равенство следует из того, что при фиксированном $\vec{y}$ индикаторная функция $\mathbf{I}_{\vec{u}}(\vec{x}+\vec{y}) = \mathbf{I}_{(\vec{u}-\vec{y})}(\vec{x})$, $\vec{x} \in \mathbb{R}^k$.

II) Запишем ф.р. $F_1(\vec{u}-\vec{y})$ при фиксированном $\vec{y}$ через функцию плотности и произведём замену переменных $\vec{x} \dashrightarrow (\vec{v}-\vec{y})$, $\vec{v}\in\mathbb{R}^k$:

$$F_1(\vec{u}-\vec{y}) = \int_{\mathbb{R}^k} \mathbf{I}_{\vec{u}}(\vec{x}+\vec{y})\, f_1(\vec{x})\, d\vec{x} \;=\; \int_{\mathbb{R}^k} \mathbf{I}_{\vec{u}}(\vec{v})\, f_1(\vec{v}-\vec{y})\, d\vec{v},$$

где индикатор $\mathbf{I}_{\vec{u}}(\vec{v})$ не зависит от $\vec{y}$. Таким образом, по теореме Фубини

$$\begin{aligned}\mathbf{P}\{\,\vec{\xi}+\vec{\eta}\leqslant\vec{u}\,\} \;&=\; \int_{\mathbb{R}^k}\left(\int_{\mathbb{R}^k} \mathbf{I}_{\vec{u}}(\vec{v})\, f_1(\vec{u}-\vec{y})\, d\,\vec{v}\right) dF_2(\vec{y}) = \\ &=\; \int_{\mathbb{R}^k} \mathbf{I}_{\vec{u}}(\vec{v})\left(\int_{\mathbb{R}^k} f_1(\vec{v}-\vec{y})\, dF_2(\vec{y})\right) d\,\vec{v},\end{aligned}$$

что эквивалентно утверждению теоремы с заявленной плотностью. III) Следует из II). ⫷

74| $\mathfrak{Y}$пр. Докажите пункт III) теоремы в одномерном случае, воспользовавшись утверждением теоремы 71 .

⚠ Под свёрткой часто понимают только операцию над функциями типа (1), (3). Терминология, связанная со сл.в., представляется нам более наглядной. Так, пункт I) можно доказать, воспользовавшись свойствами условного матем.ожидания. Заметим, что если сл.в. ξ_1,ξ_2 независимы, то справедливо равенство 34 , стр. 76, и потому

$$\mathbf{E}\big[\,\mathbf{I}(\xi_1+\xi_2\leqslant u)\,|\,\xi_2=x_2\big] = \mathbf{E}\,[\,\mathbf{I}(\xi_1\leqslant u-x_2)] = F_1(u-x_2).$$

при почти всех x_2. Следовательно,

$$\begin{gathered}\mathbf{P}\{\xi_1+\xi_2\leqslant u\} = \mathbf{E}\big[\,\mathbf{E}\big[\,\mathbf{I}(\xi_1+\xi_2\leqslant u)\,|\,\xi_2\big]\,\big] = \\ = \int_{\mathbb{R}} \mathbf{E}\big[\,\mathbf{I}(\xi_1+\xi_2\leqslant u)\,|\,\xi_2=x_2\big]\, dF_2(x_2) = \int_{\mathbb{R}} F_1(u-x_2)\, dF_2(x_2).\end{gathered}$$

Если сл.в. ξ_1,ξ_2 неотрицательны, то вид свёртки их плотностей немного упрощается.

75| Лемма. (?!) *Пусть независимые сл.в.* ξ_1, ξ_2 ($\geqslant 0$) *имеют плотности* f_1, f_2 *соответственно. Тогда плотности распределений их суммы и разности равны соответственно*

$$f_{\xi_1+\xi_2}(u) = \int_0^u f_2(u-y) f_1(y)\, dy, \quad u > 0,$$

$$f_{\xi_1-\xi_2}(u) = \int\limits_{\max\{0,-u\}}^{\infty} f_1(u+y) f_2(y)\, dy, \quad u \in \mathbb{R}^1. \tag{4}$$

76| Пример. Пусть сл.в. $\xi \overset{d}{\backsim} \mathcal{G}\mathtt{am}(p,1)$, $\eta \overset{d}{\backsim} \mathcal{G}\mathtt{am}(q,1)$. Носители этих сл.в. совпадают с положительной полуосью $\mathbb{R}^1_+$, поэтому носитель суммы $\zeta = \xi + \eta$ также будет совпадать с $\mathbb{R}^1_+$. По предыдущей лемме, при $u > 0$ свёртка

$$f_\xi * f_\eta(u) = \int_0^u \frac{(u-y)^{p-1} y^{q-1}}{\Gamma(p)\Gamma(q)} e^{-(u-y)} e^{-y}\, dy =$$

$$= \frac{e^{-u} u^{p-1}}{\Gamma(p)\Gamma(q)} \int_0^u \left(1 - \frac{y}{u}\right)^{p-1} y^{q-1} dy.$$

Замена $y \dashrightarrow tu$, $0 \leqslant t \leqslant 1$, переводит последний интеграл 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, $f_\xi * f_\eta(u) \;=\; u^{p+q-1} e^{-u} / \Gamma(p+q)$, т.е. в неформальном виде: $\mathcal{G}\mathtt{am}(p,1) * \mathcal{G}\mathtt{am}(q,1) \overset{d}{\backsim} \mathcal{G}\mathtt{am}(p+q,1)$. ◎

77| $\mathfrak{Upr}$. Покажите, что

(а) $\mathcal{G}\mathtt{am}(p,\lambda) * \mathcal{G}\mathtt{am}(q,\lambda) \overset{d}{\backsim} \mathcal{G}\mathtt{am}(p+q,\lambda)$ при любых $\lambda > 0$,

(б) если $\xi_1, \xi_2 \overset{d}{\backsim} \mathcal{E}\mathtt{x}(1)$ — независимые показательные сл.в., то их разность имеет распределение Лапласа: $\xi_1 - \xi_2 \overset{d}{\backsim} \mathcal{L}\mathtt{ap}(0,1)$.

Свёртка дискретного распределения с непрерывным. 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

78| Лемма. (?!) *Если сл.в.* η *имеет дискретное распределение, сосредоточенное на носителе* $\mathcal{Y}_\eta = \langle y_j \rangle_1^N$ *с вероятностями* $p_j = \mathbf{P}\{\eta = y_j\}$, $y_j \in \mathcal{Y}_\eta$, $1 \leqslant j \leqslant N$, $N \leqslant \infty$, *а сл.в.* ξ *имеет абсолютно-непрерывный тип распределения с плотностью* f_ξ, *то плотность свёртки их распределений равна*

$$f_{\xi+\eta}(x) = \sum\nolimits_1^N f_\xi(x - y_j)p_j.$$

79| Пример. Пусть сл.в. $\xi \stackrel{d}{\backsim} \mathfrak{Un}(0,1)$, а сл.в. η принимает два 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 $\mathfrak{Un}(0,1) * \mathfrak{Bern}({}^1\!/_2) \stackrel{d}{\backsim} \mathfrak{Un}[0;2]$. ◎

80| $\mathfrak{Упр}$. Докажите, что если одна из ф.р. непрерывна, то свёртка также будет иметь всюду непрерывную ф.р.

Свёртка дискретных распределений также имеет простой вид.

81| Лемма. (?!) *Пусть* ξ, η *— независимые сл.в., сосредоточенные на носителях* $\mathcal{X}_\xi = \langle x_k \rangle_1^N$, $\mathcal{X}_\eta = \langle y_j \rangle_1^M$, $1 \leqslant N, M \leqslant \infty$. *Тогда* $\mathcal{X}_S = \langle u = x_k + y_j : \ k = \overline{1,N},\ j = \overline{1,M} \rangle$ *— носитель* $\xi + \eta$ *и для* $\forall\, u \in \mathcal{X}_S$

$$\mathbf{P}\{\xi + \eta = u\} = \sum\nolimits_1^M \mathbf{P}\{\xi = u - y_j\}\,\mathbf{P}\{\eta = y_j\}\ .$$

Подсказка. *Использовать* $\mathbf{P}\{\xi + \eta = u\} = \sum_j \mathbf{P}\{\xi + \eta = u, \eta = y_j\}$.

82| Пример. Пусть независимые сл.в. $\xi, \eta \stackrel{d}{\backsim} \mathfrak{Pois}(\lambda)$.

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 $y_j = j \in \mathbb{N}_0$, при которых слагаемые формулы 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

$$\sum_{j=0}^{\infty} \mathbf{P}\{\xi = z - j\}\ \mathbf{P}\{\eta = j\} = \sum_{j=0}^{z} \frac{\lambda^{z-j}}{(z-j)!} e^{-\lambda}\, \frac{\lambda^j}{j!} e^{-\lambda} =$$

$$= \lambda^z e^{-2\lambda} \sum_{j=0}^{z} \frac{1}{(z-j)!\,j!} = \frac{\lambda^z}{z!} e^{-2\lambda} \underbrace{\sum_{j=0}^{z} \frac{z!}{(z-j)!\,j!}}_{2^z} = \frac{(2\lambda)^z}{z!} e^{-2\lambda}$$

по формуле бинома Ньютона. Следовательно, $\xi + \eta \overset{d}{\backsim} \mathcal{P}\mathtt{ois}(2\lambda)$ или в неформальном виде $\mathcal{P}\mathtt{ois}(\lambda) * \mathcal{P}\mathtt{ois}(\lambda) \overset{d}{\backsim} \mathcal{P}\mathtt{ois}(2\lambda)$. ⊚

В следующей теореме приведены распределения сумм независимых одинаково распределённых сл.в. для основных вероятностных моделей (см. стр. 249). Доказательство утверждений мы оставляем читателю. Заметим, что распределение свёртки дискретных сл.в. может быть найдено непосредственно из определения соответствующей вероятностной модели.

Как видно из таблицы, многие классические вероятностные модели обладают свойством замкнутости относительно свёртки — свёртка принадлежит тому же типу $\mathbb{P}$, что и слагаемые: $(F_1, \ldots, F_n) \subset \mathbb{P} \;\;\Rightarrow\;\; \mathop{*}\limits_{j=1}^{n} F_j \in \mathbb{P}$.

83| Т е о р е м а. *Пусть* $S = \sum_1^n \xi_j$, *где* $\xi_1, \ldots, \xi_n$ *— независимые одинаково распределённые сл.в. Тогда имеет место*

Распределение слагаемых	*Распределение суммы*	*Параметры распределения* S
$\xi_j \overset{d}{\backsim} \mathcal{B}\mathtt{ern}(p)$	$S \overset{d}{\backsim} \mathcal{B}\mathtt{in}(n,p)$	$n,\ \ p$
$\xi_j \overset{d}{\backsim} \mathcal{B}\mathtt{in}(k_j,p)$	$S \overset{d}{\backsim} \mathcal{B}\mathtt{in}(K,p)$	$K=\sum_1^n k_j,\ \ p$
$\xi_j \overset{d}{\backsim} \mathcal{P}\mathtt{sc}(k_j,p)$	$S \overset{d}{\backsim} \mathcal{P}\mathtt{sc}(K,p)$	$K=\sum_1^n k_j,\ \ p$
$\xi_j \overset{d}{\backsim} \mathcal{P}\mathtt{ois}(\lambda_j)$	$S \overset{d}{\backsim} \mathcal{P}\mathtt{ois}(\Lambda)$	$\Lambda=\sum_1^n \lambda_j$
$\xi_j \overset{d}{\backsim} \mathcal{N}(\mu_j,\sigma_j^2)$	$S \overset{d}{\backsim} \mathcal{N}(\mu,\mathfrak{S}^2)$	$\mu=\sum_1^n \mu_j,\ \ \mathfrak{S}^2=\sum_1^n \sigma_j^2$
$\xi_j \overset{d}{\backsim} \mathcal{E}\mathtt{x}(\lambda)$	$S \overset{d}{\backsim} \mathcal{G}\mathtt{am}(n,\lambda)$	$\lambda,\ \ n$
$\xi_j \overset{d}{\backsim} \mathcal{G}\mathtt{am}(r_j,\lambda)$	$S \overset{d}{\backsim} \mathcal{G}\mathtt{am}(R,\lambda)$	$\lambda,\ \ R=\sum_1^n r_j$
$\xi_j \overset{d}{\backsim} \mathcal{C}\mathtt{auch}(\mu_j,\sigma_j)$	$S \overset{d}{\backsim} \mathcal{C}\mathtt{auch}(\mu,\mathfrak{S})$	$\mu=\sum_1^n \mu_j,\ \ \mathfrak{S}=\sum_1^n \sigma_j$

§ 3. Распределение немонотонных преобразований

Распределение Стьюдента. В математической статистике особую роль играет распределение Стьюдента. Не вдаваясь в подробности, скажем, что оно описывает поведение отношения $\zeta = \sqrt{2p}\,{}^{\xi}/_{\sqrt{\eta}}$ с независимыми сл.величинами $\xi \overset{d}{\backsim} \mathcal{N}(0,1)$ и $\eta \overset{d}{\backsim} \mathcal{G}\mathtt{am}(p,2)$; $p>0$ — фиксированный параметр. Найдём сначала совместное распределение вектора (ζ,η).

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

$$(z,v) \ = \ H(x,y) = \left(\sqrt{2p}\,\frac{x}{\sqrt{y}}\,,\ y\right), \qquad x\in\mathbb{R}^1, y>0,$$

$$(x,y) \ = \ H^{-1}(z,v) = \left(\frac{1}{\sqrt{2p}}\,z\sqrt{v}\,,\ v\right), \qquad z\in\mathbb{R}^1, v>0.$$

Якобиан $J_{H^{-1}} = {}^{\sqrt{v}}/_{\sqrt{2p}}$. Поэтому плотность вектора $(\zeta,V) = \left(\sqrt{2p}\,{}^{\xi}/_{\sqrt{\eta}},\ \eta\right)$ при $z\in\mathbb{R}^1$, $v>0$ равна

$$f(z,v) \ = \ J_{H^{-1}}\cdot f_\xi(x)f_\eta(y)\Big|_{(x,y)=H^{-1}(z,v)} = \frac{\sqrt{v}}{\sqrt{2p}}\ f_\xi\!\left(\frac{z\sqrt{v}}{\sqrt{2p}}\right) f_\eta(v).$$

Чтобы найти частную плотность ζ, подставим сюда плотности нормального $\mathcal{N}(0,1)$ и гамма $\mathcal{G}\mathtt{am}(p,2)$ законов и проинтегрируем по $v>0$:

$$f_\zeta(z) \;=\; \frac{1}{\sqrt{2p}}\int\limits_0^\infty \sqrt{v}\frac{1}{\sqrt{2\pi}}\exp\Big\{-\frac{1}{2}\,\frac{z^2}{2p}v\Big\}\frac{1}{2^p\Gamma(p)}v^{p-1}e^{-v/2}\,dv =$$

$$= \frac{1}{\sqrt{\pi}\sqrt{p}\,2^{p+1}\,\Gamma(p)}\int\limits_0^\infty v^{p+\frac{1}{2}-1}\,\exp\Big\{-\Big(\,\frac{1}{2}+\frac{z^2}{4p}\,\Big)\,v\Big\}\,dv.$$

Последний интеграл равен $\Gamma(p+1/2)(1/2+z^2/4p)^{-p-1/2}$. Итак,

$$f_\zeta(z) =$$

$$= \frac{\Gamma(p+{}^1\!/_2)}{\sqrt{\pi\,p}\,2^{p+1}\,\Gamma(p)}\Big(\frac{1}{2}+\frac{z^2}{4p}\Big)^{-p-\frac{1}{2}} = \frac{\Gamma(p+{}^1\!/_2)}{\sqrt{\pi\,2p}\,\Gamma(p)}\Big(1+\frac{z^2}{2p}\Big)^{-\frac{(2p+1)}{2}}.$$

У классического распределения Стьюдента $p={}^k\!/_2$:

$$f_\zeta(z) \;=\; \frac{\Gamma((k+1)/2)}{\sqrt{\pi k}\,\Gamma(k/2)}\cdot\Big(1+\frac{z^2}{k}\Big)^{-\frac{1}{2}(k+1)}, \qquad z\in\mathbb{R}^1, \tag{5}$$

где k — целое число, называемое числом степеней свободы распределения Стьюдента. Последнее связано с тем, что гамма-распределение $\mathcal{G}\mathtt{am}({}^k\!/_2,2)$ совпадает с хи-квадрат распределением $\mathcal{K}\mathtt{hi}(k)$ с k степенями свободы. Легко видеть (тем, кто знаком со «вторым замечательным пределом»), что при больших k эта плотность близка к нормальной $\mathcal{N}(0,1)$ плотности.

⚠ Вывод распределения Стьюдента может быть осуществлён без обращения к теореме 66 (напомним, что эта теорема осталась у нас без доказательства). Идея вывода здесь аналогична идее доказательства леммы 68 (стр. 110) — функция $x/\sqrt{y}$ линейна при каждом y по переменной x. Продемонстрируем схему применения этого свойства на следующем примере, где заодно покажем способ нивелирования возможной неоднозначности рассматриваемого преобразования, который заключается, конечно, в разбиении пространства значений сл.векторов на подходящие подмножества.

84| Т е о р е м а. *Пусть* $f(x,y)$ — *ф.пл. сл.вектора* (ξ,η). *Тогда плотность распределения отношения* $\zeta = {}^{\xi}/_{\eta}$ *равна*

$$f_\zeta(z) = \int_{\mathbb{R}} |y|\, f(zy,y)\, dy.$$

$\Rrightarrow$ По теореме Фубини ф.р. ζ

$$F_\zeta(z) = \mathbf{P}\Big\{\, \frac{\xi}{\eta} \leqslant z \,\Big\} = \int_{\mathbb{R}^2} \mathbf{I}\big({}^{x}/_{y} \leqslant\ z\big) f(x,y)\, dxdy =$$
$$= \int\limits_{y<0} \Big[\int\limits_{x/y\leqslant z} f(x,y)\, dx \Big]\, dy + \int\limits_{y>0} \Big[\int\limits_{x/y\leqslant z} f(x,y)\, dx \Big]\, dy,$$

где точку $y = 0$ можно исключить из области интегрирования, т.к. она имеет нулевую лебегову меру. В каждом внутреннем интеграле (в квадратных скобках) по теореме 164, стр. 181, мы вправе при фиксированном $y \neq 0$ осуществить подстановку $x = yu$, т.е. ${}^{x}/_{y} = u$, $u \leqslant z$. Таким образом,

$$F_\zeta(z) = \int\limits_{y<0} \Big[|y| \int\limits_{u\leqslant z} f(uy,y)\, du \Big]\, dy + \int\limits_{y>0} \Big[|y| \int\limits_{u\leqslant z} f(uy,y)\, du \Big]\, dy,$$

что после применения теоремы Фубини даёт

$$F_\zeta(z) = \int\limits_{u\leqslant z} \Big[\int\limits_{y\in\mathbb{R}} |y| f(uy,y)\, dy \Big]\, du.$$

Такая запись ф.р. эквивалентна утверждению теоремы. $\Lleftarrow$

85| $\mathfrak{Упр}$. Покажите, что отношение независимых стандартных нормальных сл.величин имеет стандартное распределение Коши:

$${}^{\xi}/_{\eta} \overset{d}{\backsim} \mathfrak{Cauch}(0,1), \quad \text{если} \quad \xi,\eta \overset{d}{\backsim} \mathcal{N}(0,1).$$

§ 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 ф.р. F. Эта выборка упорядочивается по возрастанию $x_{(1)} \leqslant \leqslant x_{(2)} \leqslant \ldots \leqslant x_{(n)}$, т.е. строится так называемый вариационный ряд. Сл.величину, описываюшую k-ое по порядку значение, будем обозначать $\xi_{(k)}$. Требуется найти распределение какой-то части вариационного ряда, например, первого члена (минимального значения исходной выборки) $\xi_{(1)}$ или (и) 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

$$F_{(n)}(x) := \mathbf{P}\{\xi_{(n)} \leqslant x\} = \prod_{1}^{n} \mathbf{P}\{\xi_k \leqslant x\} = F^n(x).$$

Аналогично, $F_{(1)}(x) := \mathbf{P}\{\xi_{(1)} \leqslant x\} = 1 - (1 - F(x))^n$. (?!)

Если ф.р. F имеет плотность f, то формальным дифференцированием отсюда можно получить ф.пл. Например, плотность максимума выборки

$$f_{(n)}(x) = nF^{n-1}(x)f(x).$$

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 приведём следующий факт.

86| Лемма. *Если F — некоторая непрерывная ф.р., то для любой функции* $h : [0; 1] \mapsto \mathbb{R}$

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

$\Rrightarrow$ В левой части здесь стоит интеграл Лебега–Стилтьеса относительно меры Лебега–Стилтьеса μ_F, порождённой ф.р. F. Применим к нему теорему о замене переменных в интеграле Лебега. Для того чтобы осуществить замену $t = F(x),\ 0 < t < 1$, необходимо найти меру μ_F^{-1} на борелевском интервале $(0; 1)$, порождённую отображением $F(x)$, т.е. найти меру множеств вида $\mu_F^{-1}(0; t] = \mu_F\{x:\ F(x) \leqslant t\}, \quad t \in (0; 1)$.

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 такая точка только одна, то неравенство $F(x) \leqslant t$ эквивалентно неравенству $x \leqslant$ $\leqslant x_t$, т.е. по определению меры μ_F

$$\mu_F^{-1}(0; t] = \mu_F\{x:\ x \leqslant x_t\} = F(x_t) = t.$$

Если решений уравнения $F(x) = t$ много, то в качестве x_t возьмём максимальное из них; в силу непрерывности F такая точка существует. Снова получим $\mu_F^{-1}(0; t] = t$. Таким образом, мера μ_F^{-1} равна мере Лебега. Доказательство леммы следует из теоремы о замене переменных в интеграле и того факта, что в интеграле относительно меры Лебега наличие или отсутствие конечного числа точек в области интегрирования не влияет на результат. $\Lleftarrow$

Как уже отмечалось, производная ф.р. является плотностью т. т. т. когда эта производная удовлетворяет свойствам ф.плотности. В данном случае

$$\int_{\mathbb{R}} f_{(n)}(x)\,dx = \int_{\mathbb{R}} nF^{n-1}(x)f(x)\,dx = n\int_{\mathbb{R}} F^{n-1}(x)\,dF(x) = \\ = n\int_0^1 t^{n-1}\,dt = 1,$$

т.е. $f_{(n)}(x)$ 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 $u_2 \leqslant u_1$ равна

$$F_{(1,n)}(u_1, u_2) = \mathbf{P}\{\xi_{(1)} \leqslant u_1, \xi_{(n)} \leqslant u_2\} = \mathbf{P}\{\xi_{(n)} \leqslant u_2\} = F^n(u_2).$$

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 $(-\infty; u_1]$, или иначе — вероятность события B_2 за вычетом события, что вся выборка лежит в интервале $(u_1; u_2]$:

$$F_{(1,n)}(u_1, u_2) = F^n(u_2) - (F(u_2) - F(u_1))^n.$$

Окончательно получаем ф.р.

$$F_{(1,n)}(u_1, u_2) = \begin{cases} F^n(u_2) - (F(u_2) - F(u_1))^n \ , & \text{при } \ u_1 < u_2, \\ F^n(u_2) & , \text{при } \ u_1 \geqslant u_2, \end{cases}$$

и ф.пл. (при $u_1 < u_2$)

$$f_{(1,n)}(u_1, u_2) = n(n-1)(F(u_2) - F(u_1))^{n-2} f(u_1) f(u_2)$$

87| $\mathfrak{Y}$пр. Проверьте свойства ф.р. и ф.пл.

Применим полученный результат для нахождения плотности 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 всюду при $y_2 > 0$. Поэтому плотность размаха выборки равна

$$f_\rho(y) = n(n-1) \int_{-\infty}^{\infty} (F(y + y_1) - F(y_1))^{n-2} f(y_1) f(y + y_1)\, dy_1.$$

при $y > 0$.

88| Пример. Пусть $\xi_k \stackrel{d}{\backsim} \mathcal{E}\mathrm{x}(1)$, 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 ф.пл. размаха

$$\begin{aligned} f_\rho(y) = n(n-1)\int_0^\infty (e^{-y_1} - e^{-y-y_1})^{n-2} e^{-y_1} e^{-y-y_1}\,dy_1 = \\ = n(n-1)(1-e^{-y})^{n-2}e^{-y}\int_0^\infty e^{-y_1(n-2)}e^{-2y_1}\,dy_1 = \\ = \ (n-1)(1-e^{-y})^{n-2}e^{-y}, \end{aligned}$$

а его ф.р. $F_\rho(y) = (1-e^{-y})^{n-1}$, $y > 0$. ⊚

Перейдём к распределению некоторых произвольных членов вариационного ряда, например, $(\xi_{(k)}, \xi_{(m)})$, $k < m$. Можно по аналогии с предыдущим выписать ф.р. этого вектора, однако здесь удобнее оперировать с ф.пл. При $u_1 < u_2$ и малом $\epsilon > 0$ событие $\{\xi_{(k)} \in (u_1; u_1+\epsilon], \xi_{(m)} \in (u_2; u_2+\epsilon]\}$ заведомо выполняется, если количество элементов выборки, попавших в тот или иной интервал, удовлетворяет следующей схеме:

$(-\infty; u_1]$	$(u_1; u_1+\epsilon]$	$(u_1+\epsilon; u_2]$	$(u_2; u_2+\epsilon]$	$(u_2+\epsilon; \infty)$
$k-1$	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

$$\begin{aligned} &\mathbf{P}\{\xi_{(k)} \in (u_1; u_1+\epsilon], \xi_{(m)} \in (u_2; u_2+\epsilon]\} \geqslant \\ &\qquad \geqslant \mathrm{C}_n^{k,m}\, F^{k-1}(u_1)(F(u_1+\epsilon) - F(u_1)) \times \\ &\times (F(u_2) - F(u_1+\epsilon))^{m-k-1}(F(u_2+\epsilon) - F(u_2))(1-F(u_2))^{n-m}. \end{aligned}$$

Следовательно, правая «производная» совместной ф.р.

$$\lim_{\epsilon \searrow 0} \frac{1}{\epsilon^2}\,\mathbf{P}\{\xi_{(k)} \in (u_1; u_1+\epsilon], \xi_{(m)} \in (u_2; u_2+\epsilon]\} \ \geqslant$$

$$\geqslant f_{(k,m)}(u_1,u_2) := \mathrm{C}_n^{k,m} f(u_1)f(u_2)\times$$
$$\times F^{k-1}(u_1)\left(F(u_2)-F(u_1)\right)^{m-k-1}(1-F(u_2))^{n-m}.$$

Применяя (два раза) лемму 86, получаем после замен $F(u_2) \dashrightarrow t_2 \dashrightarrow x_2+x_1,\ F(u_1)\dashrightarrow t_1 \dashrightarrow x_1$

$$\int_{\mathbb{R}^2} f_{(k,m)}(u_1,u_2)\,du_1du_2 =$$
$$= \mathrm{C}_n^{k,m}\int_0^1 dt_1 \int_{t_1}^1 \left[t_1^{k-1}(t_2-t_1)^{m-k-1}(1-t_2)^{n-m}\right]dt_2 =$$
$$= \mathrm{C}_n^{k,m}\int_0^1 dx_1 \int_0^{1-x_1} \left[x_1^{k-1}x_2^{m-k-1}(1-x_2-x_1)^{n-m}\right]dx_2 = 1,$$

т.к. подынтегральное выражение есть ф.пл. распределения Дирихле. В силу теоремы 26, стр. 38, отсюда следует, что найденное 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).

По аналогичным соображениям плотность всего вектора вариационного ряда равна

$$f_{(1,\ldots,n)}(u_1,\ldots,u_n) = n!\prod_1^n f(u_k), \qquad u_1<u_2<\ldots<u_n.$$

89| П р и м е р. Пусть $\xi_k \overset{d}{\backsim} \mathrm{Un}(0,T)$, т.е. имеет равномерное распределение в отрезке $[0;T]$ с плотностью $f(x) = {}^1/_T$, $0\leqslant x\leqslant T$. Функция плотности вектора $(\xi_{(1)},\ldots,\xi_{(n)})$ равна $n!/_{T^n}$ в области $0\leqslant u_1\leqslant\ldots\leqslant u_n\leqslant T$. Можно вычислить и ф.р.:

$$\mathbf{P}\{\xi_{(k)}\leqslant u_k,\ k=1,\ldots,n\} = \frac{n!}{T^n}\int_0^{u_1}dt_1\int_{t_1}^{u_2}dt_2\cdots\int_{t_{n-1}}^{u_n}dt_n,$$
$$0\leqslant u_1\leqslant\ldots\leqslant u_n\leqslant T.$$

Поскольку неравенство $\xi_{(k)}\leqslant u$ влечёт справедливость неравенства $\xi_{(j)}\leqslant u$ 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 $u_k \dashrightarrow u'_k = \min\{u_k, \ldots, u_n, T\}$.

⊚

A. Доказательство леммы о линейном преобразовании

⊫ Применим способ решения систем линейных уравнений, при котором шаг за шагом строки матрицы Q приводятся к виду, когда все элементы каждой строки обнуляются, кроме диагональных элементов, которые становятся равными 1, т.е. матрица Q приводится к единичной матрице (см. уточняющее замечание ниже). Пусть $\vec{e}_j$ — базисный вектор-столбец с единицей на j-ом месте. Для любой матрицы M произведение j-ой строки M^{-1} на M совпадает со строкой $\vec{e}_j^{\flat}$. Таким образом, j-ый шаг процесса приведения матрицы Q может быть описан как умножение матрицы Q слева на матрицу A_j, у которой все строки, кроме j-ой, совпадают со строками единичной матрицы, а j-ая строка совпадает с j-ой строкой Q_{j-1}^{-1}, где Q_{j-1} — преобразованная матрица Q после $(j-1)$-ого шага; $Q_0 = Q$. Поскольку по построению $A_k A_{k-1} \cdots A_1 Q = \dot{\mathbb{I}}$, то имеем представление для матрицы $Q = A_1^{-1} \cdots A_k^{-1}$; в частности, отсюда $|Q| = \left(\prod_1^k |A_j|\right)^{-1}$.

Пусть плотность сл.вектора $\vec{\xi}$ равна $f(\vec{x})$. Очевидно, матрица A_1^{-1} имеет схожую с A_1 структуру. 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}, $a_1^{-1} = |A_1|$. Первая строка A_1 равна $\vec{b}^{\flat} = (1, -a_2, \ldots, -a_k)/a_1$. Отсюда,

$$F_{\vec{\eta}}(\vec{y}) = \mathbf{P}\left\{\eta_1 \leqslant y_1,\, \eta_2 \leqslant y_2, \ldots, \eta_k \leqslant y_k\right\} =$$

$$= \mathbf{P}\left\{\, a_1\xi_1 + \ldots + a_k\xi_k \leqslant y_1,\, \xi_2 \leqslant y_2, \ldots, \xi_k \leqslant y_k \,\right\} =$$

$$= \int\limits_{\substack{x_i \leqslant y_i,\\ i=\overline{2,k}}} \ldots \int \Big[\int_{\mathbb{R}} \mathbf{I}\big(a_1x_1 + \ldots + a_kx_k \leqslant y_1\big) f(\vec{x})\, dx_1 \Big]\, dx_2 \cdots dx_k.$$

Во внутреннем интеграле (в квадратных скобках) при фиксированных значениях $(x_2, \ldots, x_k)$ можно осуществить линейную замену переменной 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 164, стр. 181, 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

$$f_{\vec{\eta}}(\vec{y}) = \frac{1}{|a_1|} f(b^\flat \vec{y}, y_2, \ldots, y_k) = |A_1|\, f\big(A_1\vec{y}\big).$$

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.	$A_k^{-1}\vec{\xi}$	равна $\lvert A_k\rvert\ f(A_k\vec{y})$,
ф.пл.	$A_{k-1}^{-1}A_k^{-1}\vec{\xi} = A_{k-1}^{-1}\vec{\eta}^{(1)}$	равна $\lvert A_{k-1}\rvert\,\lvert A_k\rvert\ f(A_kA_{k-1}\vec{y})$, $\ldots$,
ф.пл.	$A_1^{-1}\cdots A_k^{-1}\vec{\xi}$	равна $\lvert A_1\rvert\cdots\lvert A_k\rvert\ f(A_k\cdots A_1\vec{y})$.

Поскольку, очевидно, $A_k \cdots A_1 = Q^{-1}$, это доказывает лемму.

⚠ Описанный процесс преобразования матрицы Q осуществим только тогда, когда у этой матрицы отсутствуют строки, в которых диагональный элемент и все элементы справа от него равны нулю. Если такое встретится, можно сначала переставить соответствующим образом строки матрицы Q, что возможно, т.к. матрица Q невырожденная. Такая перестановка эквивалентна умножению матрицы Q слева на матрицу E, получающуюся из единичной матрицы соответствующей перестановкой строк. Для матрицы E и её обратной $E^{-1} = E^{\flat}$ утверждение леммы доказывается (?!) простыми рассуждениями. Так как $\|E\| = 1$, то вместе с предыдущим имеем

$$EQ\vec{\xi} \overset{d}{\sim} \frac{1}{\|Q\|} f(Q^{-1}E^{-1}\vec{y}) \Rightarrow$$
$$\Rightarrow Q\vec{\xi} = E^{-1}(EQ\vec{\xi}) \overset{d}{\sim} \frac{1}{\|Q\|} f(Q^{-1}E^{-1}E\vec{y}) = \frac{1}{\|Q\|} f(Q^{-1}\vec{y}),$$

что и требовалось. $\Lleftarrow$

B. Упражнения и указания к решению

Вид распределений сл.в. для следующих задачах приведён на стр. 249.

Упр. B.1 Найти функцию вероятностей:

(a) $[\![\xi]\!] + 1$, если $\xi \overset{d}{\sim} \mathcal{E}\mathtt{x}(1)$;

(b) ξ, если $\xi \overset{d}{\sim} \mathcal{P}\mathtt{ois}(\vartheta)$ со случайным $\vartheta \overset{d}{\sim} \mathcal{E}\mathtt{x}(1)$;

(c) $\xi_1 - \xi_2$, если $\xi_j \overset{d}{\sim} \mathcal{P}\mathtt{ois}(1)$, $j = 1, 2$, и независимы.

Упр. B.2 Найти плотность распределения:

(a) ${}^1/_{\xi}$, если $\xi \overset{d}{\sim} \mathcal{C}\mathtt{auch}(0, 1)$;

(b) $|\xi|$, если $\xi \overset{d}{\sim} \mathcal{L}\mathtt{ap}(0, 1)$;

(c) ξ, если $\ln \xi \overset{d}{\sim} \mathcal{N}(0, 1)$ (логнормальное распределение);

(d) ξ, если $\xi \overset{d}{\sim} \mathcal{E}\mathtt{x}(1/\vartheta)$ со случайным параметром $\vartheta \overset{d}{\sim} \mathcal{E}\mathtt{x}(1)$;

(e) (ρ, ϑ), где $\rho = \xi^2 + \eta^2$, $\vartheta = \xi/\eta$, если $(\xi, \eta) \overset{d}{\sim} \mathcal{U}\mathtt{n}(\Omega)$, где Ω — единичный круг с центром в начале координат;

(f) $\min(\xi_1, \ldots, \xi_k)$, если $\xi_j \overset{d}{\sim} \mathcal{E}\mathtt{x}(1)$, $j = \overline{1, k}$, и независимы;

(g) $\xi_1 - \xi_2$, если $\xi_j \overset{d}{\sim} \mathcal{E}\mathtt{x}(\lambda_j)$, $j = 1, 2$, и независимы;

(h) $\xi_1/(\xi_1 + \xi_2)$, если $\xi_j \overset{d}{\sim} \mathcal{E}\mathtt{x}(1)$, $j = 1, 2$, и независимы.

Упр. B.3 Найти $h(x)$ такое, что $h(\xi) \overset{d}{\sim} \mathcal{C}\mathtt{auch}(0, 1)$, если

$\xi \overset{d}{\backsim} \mathcal{E}\mathrm{x}(1)$.

$\mathfrak{Y}$пр. B.4 Пусть $\left\langle \xi_k \overset{d}{\backsim} F \right\rangle_1^n$ — независимые сл.в. с общей абсолютно-непрерывной ф.р. F. Покажите, что при $\forall\, k \leqslant n$

$$\text{(a)} \qquad \mathbf{P}\left\{\xi_k \leqslant x \,|\, \xi^{(n)} = t\right\} = \begin{cases} \dfrac{n-1}{n}\dfrac{F(x)}{F(t)} & \text{при } x < t, \\ 1 & \text{при } x \geqslant t; \end{cases}$$

$$\text{(б)} \qquad \mathbf{E}\left[\xi_k \,|\, \xi^{(n)} = t\right] = \frac{n-1}{n}\frac{1}{F(t)}\int_{-\infty}^{t} x\, dF(x) + \frac{t}{n}.$$

Указания к решению задач.

✦ (?!), стр. 88. Проверить свойства вероятности. Например, $\mathbb{P}_\eta\{B_1 \uplus B_2\} \overset{\mathrm{df}}{=} \mathbb{P}_\xi\{h^{-1}(B_1 \uplus B_2)\} = \mathbb{P}_\xi\{h^{-1}(B_1)\} + \mathbb{P}_\xi\{h^{-1}(B_2)\} = \mathbb{P}_\eta\{B_1\} + \mathbb{P}_\eta\{B_2\}$.

✦ 61, стр. 90. Событие $\{\eta \leqslant y\} \Leftrightarrow \neg\{\xi < (y-a)/b\}$. Учесть появление знака “$<$”.

✦ 68, стр. 95. Найти обратное преобразование и его якобиан, применить теорему 66, стр. 93.

✦ 69, стр. 95. Найти обратное преобразование.

✦ 74, стр. 98. Воспользоваться теоремой 71.

✦ 75, стр. 99. Подынтегральное выражение в формуле для свёртки отлично от нуля a) в первом случае при $y \geqslant 0, u - y \geqslant \geqslant 0$, b) во втором случае при $y \geqslant 0, u + y \geqslant 0$.

✦ 77, стр. 99. Применить (4), стр. 99. При отрицательном u произвести замену переменной интегрирования.

✦ 78, стр. 100. В формуле 2, стр. 97, записать матем.ожидание дискретной сл.в. в виде ряда.

✦ 80, стр. 100. Применить теорему Лебега к 1, стр. 97.

✦ 85, стр. 104. Подынтегральное выражение плотности в теореме 84, стр. 104, равно

$$|y| \exp\left(-\tfrac{1}{2}y^2(z^2+1)\right)dy = \exp\left(-\tfrac{1}{2}y^2(z^2+1)\right)d\,\tfrac{1}{2}y^2.$$

✦ (?!), стр. 105. $\{\xi_{(1)} > x\} \Leftrightarrow \xi_k > x,\ k = \overline{1,n}$.

✦ В.1. a) $\xi \stackrel{d}{\backsim} \mathcal{G}\mathtt{eo}(e^{-1})$;

b) $\mathbf{P}\{\xi = k\} = {}^{1}/{}_{2^{k+1}}$, применить 54, стр. 78;

c) $\mathbf{P}\{\xi = k\} = e^{-2} \sum_{j=\max(0,-k)}^{\infty} 1/j!(j+k)!$ (функция Бесселя).

✦ В.2. a) $\mathcal{C}\mathtt{auch}(0,1)$;

b) $\mathcal{E}\mathtt{x}(1)$, найти ф.р.;

d) найти частную плотность из 55, стр. 79;

e) независимы, $\rho \stackrel{d}{\backsim} \mathcal{U}\mathtt{n}(0,1)$, $\vartheta \stackrel{d}{\backsim} \mathcal{C}\mathtt{auch}(0,1)$, найти площадь части $x^2 + y^2 < r, x/y < \theta$ единичного круга;

f) $\mathcal{E}\mathtt{x}(1/k)$, см. указание к (?!), стр. 105, выше.

h) $\mathcal{U}\mathtt{n}(0,1)$; найти частную плотность $(\xi_1/(\xi_1+\xi_2), \xi_2)$.

✦ В.3. Проверить и применить следующие утверждения: если $\zeta \stackrel{d}{\backsim} F$ с непрерывной, строго монотонной ф.р. F и $\omega \sim \mathcal{U}\mathtt{n}(0,1)$, то (i) $F(\zeta) \stackrel{d}{\backsim} \omega$, (ii) $F^{-1}(\omega) \stackrel{d}{\backsim} \zeta$.

✦ В.4. По аналогии с 53, стр. 77, проверить (32), стр. 75.

III 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. 217).

§ 1. Функции множеств

Сначала очертим круг естественных требований, предъявляемых к 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 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. 226), в частности следующие утверждения. Напомним, что знак $\biguplus$ вместо привычного знака объединения применяется в ситуациях, когда предполагается или подразумевается, что множества, входящие в объединение, не пересекаются.

90 $\mathfrak{Y}\mathfrak{n}\mathfrak{p}$. Пусть $\langle Q_n \rangle_{n=1}^{\infty}$ — подмножества Ω. Докажите, что

а) $\mathbf{I}\Big(\bigcap\limits_{n=1}^{\infty} Q_n\Big) = \prod\limits_{n=1}^{\infty} \mathbf{I}(Q_n), \quad \mathbf{I}\Big(\biguplus\limits_{n=1}^{\infty} Q_n\Big) = \sum\limits_{n=1}^{\infty} \mathbf{I}(Q_n);$

б) $\bigcup\limits_{n=1}^{\infty} Q_n = \biguplus\limits_{k=1}^{\infty} B_k,$ где $B_1 = Q_1,\ B_k = Q_k \bigcap\limits_{n=1}^{k-1} Q_n^c,\ \ k \geqslant 2;$

в) $\mathbf{I}(\varliminf\limits_{n} Q_n) = \varliminf\limits_{n} \ \mathbf{I}(Q_n), \quad \mathbf{I}(\varlimsup\limits_{n} Q_n) = \varlimsup\limits_{n} \ \mathbf{I}(Q_n).$

г) для монотонной последовательности $\langle Q_n \rangle$

$$\varliminf_{n} Q_n = \varlimsup_{n} Q_n.$$

Обозначим через $\overline{\mathbb{R}}_+^1$ положительную часть числовой прямой $[0; +\infty]$ с добавленной точкой $\{+\infty\}$.

О п р е д е л е н и я. Неотрицательная функция $\mu : \mathcal{F} \mapsto \overline{\mathbb{R}}_+^1$, заданная на некотором 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 $S_j \in \mathcal{F},\ j = \overline{1, N}$, равно сумме значений μ на этих подмножествах:

$$B,\ \{S_j\}_1^N \in \mathcal{F}: \quad B = \biguplus\nolimits_1^N S_j \ \ \Rightarrow \ \ \mu(B) \ = \ \sum\nolimits_1^N \mu(S_j); \qquad (1)$$

(✓) *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:

$$B,\ S_1, S_2, \ldots \in \mathcal{F}:\ B = \lim_n {}_{\Downarrow(\Uparrow)} S_n \quad \Rightarrow \quad \mu(B) = \lim_n \mu(S_n).$$

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.

91| $\mathfrak{Y}\mathfrak{n}\mathfrak{p}$. Докажите, что

✦ для аддитивной или σ-аддитивной функции $\mu : \mathcal{F} \mapsto \overline{\mathbb{R}}^1_+$ значение $\mu(\varnothing) = 0$, если $\varnothing \in \mathcal{F}$ и $\exists\, B \in \mathcal{F}:\ \ \mu(B) < \infty$;

✦ σ-аддитивная функция аддитивна, если $\mu(\varnothing) = 0$.

Случай $\mu(B) = \infty$ для $\forall\, B \in \mathcal{F}$ по понятным причинам не рассматривается.

Определение. Неотрицательная функция множеств $\mu : \mathcal{F} \mapsto \overline{\mathbb{R}}^1_+$ называется *мерой на* $\mathcal{F}$, если она σ-аддитивная и $\mu(\varnothing) = 0$.

92| 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 $\mu(A) = \#(A \cap \mathcal{S})$ — количество элементов $\mathcal{S}$, вошедших в A. Легко показать (?!), что μ — 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. ⊚

⚠ Конструирование нетривиальной меры может начинаться только с некоторого узкого класса множеств. При этом, необходимо указывать 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 интервала на прямой и прямоугольника на плоскости:

$$\lambda\big(\lfloor a;b\rceil\big) = (b-a), \quad \lambda\big(\lfloor a_1;b_1\rceil \times \lfloor a_2;b_2\rceil\big) = (b_1-a_1)(b_2-a_2).$$

Здесь символы $\lfloor$ и $\rceil$ означают, что для этих мер неважно, 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. Как показывает следующая лемма, здесь оказалось важным, что класс интервалов образует полукольцо (см. стр. 228).

93| $\mathfrak{Упр}$. Докажите, что класс $\mathcal{H}$ подмножеств есть 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 $B = Q + \biguplus_1^k Q_j$.

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

$$B, \langle\, S_j \,\rangle_{j=1}^N \in \mathcal{F} : \quad B \subset \bigcup\nolimits_1^N S_j \quad \Rightarrow \quad \mu(B) \leqslant \sum\nolimits_1^N \mu(S_j).$$

94| 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.*

⊫ (i_1) Если $Q \subset W$, то ввиду 93 $W = Q + \biguplus_1^k q_j$, 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 93, $W = \biguplus_{j=0}^{k_i} q_{ij}$, 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

$$Q_1 = q_{10} \cap W = q_{10} \bigcap_{i=2}^{N} \biguplus_{j=0}^{k_i} q_{ij} = \biguplus_{j_2=0}^{k_2} \cdots \biguplus_{j_N=0}^{k_N} (q_{10} \cap q_{2j_2} \cap \ldots \cap q_{Nj_N}).$$

Поэтому 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 $\gamma(W) \leqslant \leqslant \sum_1^N \gamma(Q_i)$.

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. 90). Применяя два раза лемму 206, стр. 228, получаем, что для $\forall\, n \geqslant 1$ существуют элементы полукольца $Q_{ni}, G_{nj} \in \in \mathcal{H},\ i = \overline{1, L_n}, j = \overline{1, Z_n}$, 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 $W, \{Q_i\}_1^\infty$ — 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 $N < \infty$ конечная сумма $\biguplus_1^N Q_i \subset W$. Снова применяя лемму 206, стр. 228, представим W в виде конечной суммы:

$$W = \biguplus_{i=1}^{N} Q_i + \biguplus_{j=1}^{L} q_j, \qquad q_j \in \mathcal{H},\ L < \infty.$$

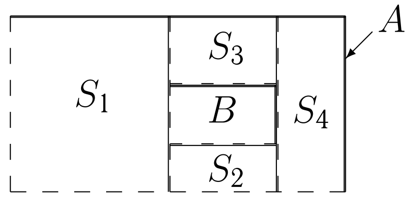
Следовательно, $\gamma(W) \geqslant \sum_1^N \gamma(Q_i)$. Полагая здесь $N \to \infty$, получаем требуемое неравенство $\gamma(W) \geqslant \sum_1^\infty \gamma(Q_i)$. ⫷

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 плоскоски.

95| **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 $\big\langle\, (a;b],\ A \leqslant a \leqslant b \leqslant B \,\big\rangle$ на интервале $\Omega = (A;B]$ образует уже полуалгебру.

2) На плоскости $\mathbb{R}^2$ полукольцо образует класс прямоугольников
$\mathcal{K} = \left\langle\, (a_1; b_1] \times (a_2; b_2], a_1 < b_1, a_2 < b_2 \,\right\rangle.$



На рисунке справа $A = B + \biguplus_1^4 S_i$.

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.

96| $\mathfrak{Y}\mathfrak{n}\mathfrak{p}$. Докажите, что если в пространстве Ω_i выделено полукольцо множеств $\mathcal{H}_i$, $\;i = \overline{1, k}$, то на прямом произведении $\Omega_1 \times \ldots \times \Omega_k$ 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 $\lambda\bigl(\biguplus_1^n (a_i; b_i]\bigr) \;=\; \sum_1^n \lambda\bigl((a_i; b_i]\bigr)$.

Система всевозможных конечных объединений непересекающихся элементов полукольца образует (минимальное) кольцо множеств, включающее все элементы полукольца (стр. 229). 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.

97| 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* $B = \biguplus_1^n S_i$, $n < \infty$:

$$\tilde{\mu}(B) = \sum\nolimits_1^n \mu(S_i). \qquad (3)$$

$\Rrightarrow$ 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:

$$B = \biguplus\nolimits_1^n S_i = \biguplus\nolimits_1^k D_j .$$

Тогда любое $S_i = S_i\, B = \biguplus_{j=1}^k D_j\, S_i$ и в силу аддитивности μ на полукольце $\mu(S_i) = \sum_{j=1}^k \mu(D_j\, S_i)$. 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

$$B = \biguplus\nolimits_{j=1}^n B_j, \quad S_k = \biguplus\nolimits_{i=1}^{N_k} D_{ki}, \qquad n, N_k < \infty, \quad B_j, D_{ki} \in \mathcal{H},$$

где все множества D_{ki} попарно не пересекаются.

Так как мера μ σ-аддитивна на полукольце и $B_j = B_j B$, то

$$\mu(B_j) = \mu\Bigg(\biguplus_{k=1}^\infty \biguplus_{i=1}^{N_k} B_j\, D_{ki}\Bigg) = \sum_{k=1}^\infty \sum_{i=1}^{N_k} \mu(B_j\, D_{ki}) = \sum_{k=1}^\infty \tilde{\mu}(B_j\, S_k).$$

Т.к. члены бесконечного ряда с положительными слагаемыми можно произвольно переставлять, то 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 его обозначения можно оставить тот же символ, что использовался для меры на полукольце: $\tilde{\mu} = \mu$.

Эквивалентность σ-аддитивности и непрерывности.

98| Т е о р е м а. *Пусть функция множеств $\mu : \mathcal{A} \mapsto \overline{\mathbb{R}}^1_+$ задана на кольце $\mathcal{A}$ подмножеств Ω.*

I) *Если μ непрерывна и аддитивна, то она σ-аддитивна.*

II) *Если μ σ-аддитивна, то для любой монотонной последовательности $\langle S_n \rangle_1^\infty \subset \mathcal{A}$ ($\mu(S_k) < \infty$ при некотором $k \geqslant 1$)*

$$\lim_n \mu(S_n) = \mu(\lim_n S_n).$$

$\Rrightarrow$ I) Пусть $B = \biguplus_1^\infty S_i$, тогда $B = \lim_{n\uparrow} \biguplus_{i=1}^n S_i$, причём конечные суммы принадлежат кольцу. Поэтому в силу непрерывности и аддитивности

$$\mu(B) = \lim_n \mu\Big(\biguplus_1^n S_i\Big) = \lim_n \sum_1^n \mu(S_i) = \sum_1^\infty \mu(S_i).$$

II) Пусть $B = \lim_{n}{}_{\uparrow} S_n = \bigcup_{n=1}^\infty S_n = \bigcup_{n=k}^\infty S_n$ и $\mu(S_n) < < \infty, \forall\, n$. Так как $S_1 \subset S_2 \subset \ldots$, то в силу 90, стр. 116,

$$\bigcup_{n=k}^{\infty} S_n = S_k + \biguplus_{n=k}^{\infty} (S_{n+1} \setminus S_n)$$

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

$$\begin{aligned}\mu(B) &= \mu(S_k) + \sum\nolimits_{n=k}^{\infty}(\mu(S_{n+1}) - \mu(S_n)) = \\ &= \lim_N\Big(\mu(S_k) + \sum\nolimits_{n=k}^{N}(\mu(S_{n+1}) - \mu(S_n))\Big) \;=\; \lim_N \mu(S_N).\end{aligned}$$

Пусть $B = \lim_n{\downarrow} S_n = \bigcap\limits_{n=1}^{\infty} S_n = \bigcap\limits_{n=k}^{\infty} S_n$ для $\forall\, k \geqslant 1$. По правилу двойственности d'Morgána, с учетом того, что $S_k \setminus S_{k+1} \subset \subset S_k \setminus S_{k+2}$,

$$S_k \setminus \bigcap_{n=1}^{\infty} S_n = \bigcup_{n=k+1}^{\infty} (S_k \setminus S_n) \;=\; \lim_n{\uparrow}\,(S_k \setminus S_n).$$

По доказанному свойству непрерывности меры $(\mu(S_k) < \infty)$

$$\mu(S_k) - \mu\Big(\bigcap\nolimits_{n=1}^{\infty} S_n\Big) = \mu(S_k) - \lim_n \mu(S_n) \Rightarrow \mu(B) = \lim_n \mu(S_n).$$

⫤

⚠ Для функций множеств на кольце условия σ-аддитивности и непрерывности эквивалентны. Кроме того, любое кольцо есть полукольцо, следовательно, для таких функций справедливы утверждения леммы 94, стр. 119. Поэтому определение меры иногда удобнее начинать с кольца подмножеств.

99| $\mathfrak{Y}\mathfrak{np}$. 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 $\mu(\lim_{\downarrow} S_n) = 0$.

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

$$\inf_{G^*\in\overline{\mathcal{G}}} \sum_{A\in G^*} \lambda(A) = \sup_{G_*\in\underline{\mathcal{G}}} \sum_{A\in G_*} \lambda(A),$$

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, $\underline{\mathcal{G}} = \Big\langle\, G_* = \big\{A_i\big\}_1^n \subset \mathcal{H} : \ \biguplus_1^n A_i \subset G, \ \ n < < \infty \,\Big\rangle$ задаёт приближения G изнутри. К сожалению, класс измеримых по Жордану множеств не слишком широк (см. задачу ниже). Более того, этот класс не замкнут относительно счётных объединений и пересечений, что неудобно: каждое утверждение асимптотического характера будет требовать обязательной проверки измеримости соответствующего множества.

100 $\mathfrak{Y}$пр. Докажите, что множество всех рациональных точек на отрезке $[0;1]$ не измеримо 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 $G \subset \Omega$:

$$\mu^*(G) = \inf \Big\langle \sum_{B_j\in G^\sigma} \mu(B_j) : \quad G^\sigma = \big\{B_j\big\}_{j=1}^\infty \subset \mathcal{A}, \ \ \bigcup_{j=1}^\infty B_j \supset G \Big\rangle.$$

При этом полагается по определению $\mu^*(\varnothing) = 0$ и $\mu^*(G) = \infty$, если не существует ни одного $\mathcal{A}$-покрытия множества G.

101 $\mathfrak{Y}$пр. Воспользовавшись утверждением б) задачи 90, стр. 116, покажите, что в определении внешней меры можно ограничиться покрытиями $\biguplus B_j \supset G$ с непересекающимися B_j. Более того, эти подмножества можно выбирать только из полукольца, порождающего кольцо $\mathcal{A}$.

102| 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 (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 лишь счётно-полуаддитивна.

Определение. Множество $G \subset \Omega$ *измеримо по мере* μ^*, если для $\forall\, B \subset \Omega$

$$\mu^*(B) = \mu^*(BG) + \mu^*(B\, G^c).$$

⚠ 1) Если $B = \Omega$, то для любого покрытия $\bigcup_j^\infty A_j \supset G^c$ подмножество $\left(\bigcup_j^\infty A_j\right)^c$ $(\subset G)$ 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 $\mu^*(B) \leqslant \leqslant \mu^*(B\, G) + \mu^*(B\, G^c)$.

2) Измеримость множества не означает, что его внешняя мера конечна.

103| $\mathfrak{Y}\mathfrak{n}\mathfrak{p}$. Докажите, что пустое множество $\varnothing$ и всё пространство Ω измеримы и, кроме того, измеримы все множества нулевой меры (внешней).

Здесь так же можно 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 ([2], стр. 150, 152), на эту роль может претендовать σ-алгебра, порождённая полукольцом (см. стр. 230 -233).

104| $\mathfrak{Y}$пр. Докажите, что если $\mathcal{A} = \mathfrak{C}(\mathcal{H})$ — кольцо, порождённое полукольцом $\mathcal{H}$, то σ-алгебра $\sigma(\mathcal{H}) = \sigma(\mathcal{A})$.

О п р е д е л е н и е. 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 такая возрастающая последовательность подмножеств $F_n \subset F_{n+1} \in \mathcal{F}$, что $F_n \nearrow \Omega$ и $\mu(F_n) < \infty$ при $\forall\, n \geqslant 1$.

105| Т е о р е м а. $\big[$*К. Каратеодори.*$\big]$ *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, то μ^* — единственная мера на σ-алгебре $\sigma(\mathcal{A})$, совпадающая с μ на кольце $\mathcal{A}$.*

⇛ Стр. 217-221. ⇚

106 Следствие. *Для любой σ-конечной меры μ, заданной на полукольце $\mathcal{H}$, существует единственное продолжение μ^* до минимальной σ-алгебры $\sigma(\mathcal{H})$, порождённой $\mathcal{H}$.*

Мера на σ-алгебре, порождённой полукольцом. Итак, можно выделить следующие этапы процесса построения меры на подмножествах пространства Ω.

(M_1) Задаётся полукольцо $\mathcal{H}$ подмножеств Ω.

(M_2) На элементах $\mathcal{H}$ определяется функция $\mu : \mathcal{H} \mapsto \overline{\mathbb{R}}^1_+$.

(M_3) Устанавливается σ-аддитивность (или аддитивность и σ-полуаддитивность) и σ-конечность μ на $\mathcal{H}$.

(M_4) Мера μ продолжается на кольцо $\mathfrak{C}$ всех конечных объединений непересекающихся элементов полукольца $\mathcal{H}$.

(M_5) По теореме Каратеодори делается вывод 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 98, 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 задана мера μ, то 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 $A \vartriangle B = (A \setminus B) + (B \setminus A) = AB^c + BA^c$ называется симметрической разностью и представляет собой как бы теоретико-множественную характеристику расхождения двух множеств.

107| 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 $\forall\, \epsilon > 0$ и $\forall\, G \in \sigma(\mathcal{A})$, $\mu(G) < \infty$, существует такое $B \in \mathcal{A}$, что $\mu(B \vartriangle G) < \epsilon$.*

$\Vdash\!\!\!>$ 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 $\forall\, \epsilon > 0$ найдётся подмножество $B \in \mathcal{A}$ такое, что $\mu(B \vartriangle (QG)) < \epsilon$. 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 $\forall\, \epsilon > 0$ найдутся подмножества $B_1, B_2 \in \mathcal{A}$ такие, что

$$\mu(B_i \vartriangle (QG_i)) = \mu(B_i Q^c \cup B_i G_i^c) + \mu(QG_i B_i^c) < \frac{1}{2}\epsilon, \quad i = 1, 2.$$

Тогда

$$\big(Q(G_1 G_2)\big) \vartriangle (B_1 B_2) = QG_1G_2(B_1^c \cup B_2^c) + B_1B_2(Q^c \cup G_1^c \cup G_2^c)$$
$$\subset \big(B_1Q^c \cup B_1G_1^c + QG_1B_1^c\big) \bigcup \big(B_2Q^c \cup B_2G_2^c + QG_2B_1^c\big).$$

Откуда $\mu(Q(G_1G_2) \vartriangle (B_1B_2)) < \epsilon$. Аналогично 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 $\forall\,\epsilon > 0$ найдётся k такое, что $\mu(QG \setminus (QG_k)) < {}^{\epsilon}/_{2}$. Так как $G_k \in \mathcal{G}$, то найдётся $B \in \mathcal{A}$ такое, что $\mu(B \vartriangle QG_k) < {}^{\epsilon}/_{2}$. Тогда ввиду включения $G \supset G_k$

$$QGB^c + B(QG)^c \subset QG_kB^c \cup B(Q^c \cup G_k^c) \cup QGG_k^c.$$

Следовательно, $\mu(B \vartriangle QG) \leqslant \mu(B \vartriangle QG_k) + \mu(QG \setminus (QG_k)) < \epsilon$.

Таким образом, класс $\mathcal{G}$ есть λ-система, замкнутая относительно конечных пересечений. Так как, очевидно, $\Omega \in \mathcal{G}$, то по лемме 213, стр. 231, класс $\mathcal{G}$ образует σ-алгебру. Кольцо $\mathcal{A} \subset \mathcal{G}$, поэтому $\mathcal{G} = \sigma(\mathcal{A})$, т.е. для любого подмножества $G \in \sigma(\mathcal{A})$ и $\forall\,\epsilon > 0$ найдётся подмножество из кольца $B \in \mathcal{A}$ такое, что $\mu(B \vartriangle (QG)) < \epsilon$.

Пусть $G \in \sigma(\mathcal{A})$ и $\mu(G) < \infty$. Т.к. мера μ σ–конечна на кольце, то найдётся множество $Q \in \mathcal{A}$ такое, что $\mu(G \setminus (QG)) < \epsilon$. По только что доказанному можно указать множество $B \in \mathcal{A}$, для которого $\mu(B \vartriangle (QG)) < \epsilon$, откуда $\mu(B \vartriangle G) < 2\epsilon$. ⫤

«Аппроксимирующее свойство» меры может быть расширено до всех подмножеств исходного пространства.

108| 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 $G \in \mathscr{F}$ такое, что $G \supset A$ и $\mu^*(A) = \mu(G)$.*

⊫ По определению внешней меры найдётся последовательность $\left\langle F_{in} \right\rangle_{i=1}^{\infty}$, $n \geqslant 1$, счётных $\mathscr{F}$-покрытий множества A таких, что $\mu^*(A) = \lim_n \sum_{i=1}^{\infty} \mu(F_{in})$. По лемме 90, стр. 116, каждое из этих покрытий можно выбрать так, чтобы все его элементы не пересекались. Поэтому для множеств $B_n = \biguplus_i F_{in} \in \mathscr{F}$ последовательность $\mu(B_n) \to \mu^*(A)$. Легко понять, что любое конечное пересечение $\bigcap_1^n B_i$ также образует покрытие A и, кроме того,

$$\min_{i\leqslant n}\mu(B_i)\geqslant\mu\Big(\bigcap\nolimits_1^n B_i\Big)\geqslant\mu^*(A).$$

Т.к. $\mu(B_i)\geqslant\mu^*(A)$ при $\forall\, i\geqslant 1$, то $\min_{i\leqslant n}\mu(B_i)\underset{n}{\to}\mu^*(A)$. Поэтому в силу свойства непрерывности меры $\mu(\cap_1^\infty B_i)=\mu^*(A)$. ⫤

Для множеств меры нуль, опираясь только на определение внешней меры, можно уточнить предыдущие утверждения.

109| Лемма. (?!) *Пусть* $\mu:\mathcal{A}\mapsto\overline{\mathbb{R}}_+^1$ — *мера на кольце подмножеств. Внешняя мера* $\mu^*(N)=0$ *т. т. т. когда для* $\forall\,\epsilon>0$ *найдётся семейство непересекающихся подмножеств* $\big\langle\, S_n\,\big\rangle_1^\infty\subset\mathcal{A}$, *что* $\biguplus_n S_n\supset N$ *и* $\sum_n\mu(S_n)<\epsilon$.

Пополнение по мере. Представляется весьма удобным, когда доказав равенство нулю какого-либо множества, мы можем перенести это свойство на более мелкие подмножества: $B\in\mathscr{F},\ \mu(B)=0,\ A\subset B\ \Rightarrow\ A\in\mathscr{F},\ \mu(A)=0$. Мера (или σ-алгебра) с таким свойством называется *полной*. Любую σ-алгебру можно пополнить по мере. Рассмотрим класс

$$\mathscr{F}_\mu\ =\ \Big\langle\, B\subset\Omega:\ \ \exists A,C\in\mathscr{F},\ A\subset B\subset C,\ \mu(C\setminus A)=0\,\Big\rangle. \quad (4)$$

Этот класс образует уже σ-алгебру (?!) относительно естественного продолжения μ :

$$\tilde{\mu}(B)\ =\ \mu(C),\quad \text{если}\quad A\subset B\subset C,\mu(C\setminus A)=0.$$

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 90, стр. 116, 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^*, то

$$\mu^*(\Omega_n) = \mu^*(B\Omega_n) + \mu^*(B^c\Omega_n).$$

По лемме 108 найдутся множества $C_n, G_n \in \mathscr{F}$ такие, что

$$\mu^*(B\Omega_n) = \mu(C_n), \quad B\Omega_n \subset C_n,$$
$$\mu^*(B^c\Omega_n) = \mu(G_n), \quad B^c\Omega_n \subset G_n.$$

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

$$\mu(\Omega_n) = \mu(C_n) + \mu(G_n), \text{ т.е. } \mu(\Omega_n \setminus G_n) = \mu(C_n).$$

Ясно, что $A_n = \Omega_n \setminus G_n \subset B\Omega_n$, поэтому 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 в евклидовом пространстве

Борелевская σ-алгебра в $\mathbb{R}^1$. Как уже отмечалось, минимальный естественный класс подмножеств $\mathbb{R}^1$, с которого может начинаться 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 $\mathscr{B}$.

110 $\mathfrak{Y}\mathfrak{n}\mathfrak{p}$. Докажите $\mathscr{B}$-измеримость

- одноточечных множеств и 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 217, стр. 233.

111| 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* $(-\infty;b]$, $b\in\mathbb{R}^1$.

⊫ $\left[\sigma(\mathscr{O}) = \sigma(\mathcal{K}_{(\,)})\right]$ Т.к. любое открытое множество $\mathbb{R}^1$ есть счётное объединение открытых интервалов, то ввиду 110 класс $\mathscr{O}\subset\mathscr{B}$. Обратно, класс интервалов $\mathcal{K}_{(\,)}$ порождает $\mathscr{B}$ и $\mathcal{K}_{(\,)}\subset\mathscr{O}\subset\mathscr{B}$. Поэтому $\mathscr{B}=\sigma(\mathcal{K}_{(\,)})\subset\sigma(\mathscr{O})\subset\mathscr{B}$. ⫤

⚠ Э. Борель (1898) ввёл определение меры любого открытого множества $\mu(B) = \sum_n (c_n - a_n)$, 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. 235) σ-алгебру $\mathscr{B}(\Omega) := \sigma(\mathscr{O})$ называют борелевской. Эта σ-алгебра будет совпадать с σ-алгеброй, порождённой базой топологии, только если эта база счётная, что верно для польских (например, евклидовых) пространств.

Из 106 непосредственно вытекает следующая важная

112| 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 219, стр. 233. Эта характеризация объясняет причину, по которой часто задание какой-либо меры начинают не с полукольца, а с порождённого кольца.

113| 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.

114| Лемма. (?!) *Пусть $\mathscr{F}$ — σ-алгебра в Ω и $Q \subset \Omega$. Тогда совокупность пересечений Q со всеми множествами $\mathscr{F}$:*

$$Q \Cap \mathscr{F} := \Big\langle Q \cap B : \quad B \in \mathscr{F} \Big\rangle,$$

образует σ-алгебру подмножеств Q.

Подсказка. *Роль основного пространства выполняет не Ω, а Q.*

Здесь важна связь индуцированной (на Q) σ-алгебры с семейством, порождающим исходную σ-алгебру.

115| Теорема. *Пусть $\sigma(Q \Cap \mathcal{H}; Q)$ — минимальная σ-алгебра на Q, порождённая $Q \Cap \mathcal{H}$. Тогда*

$$Q \Cap \sigma(\mathcal{H}) \;=\; \sigma(Q \Cap \mathcal{H}; Q).$$

$\Rrightarrow$ 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 $\mathcal{H}_Q = \big\langle\, B \in \sigma(\mathcal{H}) \;:\; Q\,B \in \sigma(Q \Cap \mathcal{H}; Q) \,\big\rangle$, 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 $Q\,B^c = Q \setminus (Q\,B) \in$ $\in \sigma(Q \Cap \mathcal{H}; Q)$. Следовательно, $\mathcal{H}_Q$ — σ-алгебра на Ω.

По построению $\mathcal{H} \subset \mathcal{H}_Q \subset \sigma(\mathcal{H})$, поэтому $\mathcal{H}_Q = \sigma(\mathcal{H})$, т.е. $\forall\, B \in \sigma(\mathcal{H}) \;\Rightarrow\; Q\,B \in \sigma(Q \Cap \mathcal{H}; Q)$. $\Lleftarrow$

⚠ Другими словами, если подмножество $B \subset Q$ $(\subset \Omega)$, то, установив его измеримость в более узком пространстве — $B \in \sigma(Q \Cap \mathcal{H}; Q)$, 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 $\big\langle\, [0;b], 0 \leqslant b \leqslant 1 \,\big\rangle$, или семейством $\big\langle\, (a;b], 0 \leqslant a \leqslant b \leqslant 1 \,\big\rangle$, или … (см. 111).

Расширенная прямая $\overline{\mathbb{R}}^1$. 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

$$\tilde{\mathcal{K}} = \left\langle\, [-\infty; b], \;\; (a; +\infty], \;\; [-\infty; +\infty], \;\; (c; d] \;\; a, b, c, d \in \mathbb{R}^1 \,\right\rangle$$

или класса интервалов $\widehat{\mathcal{K}} = \left\langle\, [-\infty; b], \;\; [-\infty; +\infty], \;\; b \in \mathbb{R}^1 \,\right\rangle$.

Заметим, что, например, $(0; +\infty] \neq (0; +\infty)$, поэтому, в частности, $\lim\limits_{b \to +\infty} (a; b] \neq (a; +\infty]$.

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 $\mathscr{B}(\mathbb{R}^1)$.

116 Лемма. (?!) *Множество $B \in \mathscr{B}(\overline{\mathbb{R}}^1)$ т. т. т. когда подмножество его конечных точек $B \setminus \{-\infty, +\infty\} \in \mathscr{B}(\mathbb{R}^1)$.*

Борелевская σ-алгебра в $\mathbb{R}^k$. Борелевскую σ-алгебру в евклидовом пространстве $\mathbb{R}^k$ 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 $\mathscr{O} = \left\langle\, U_r(\vec{c}), \;\; \vec{c} \in \mathbb{R}^k, 0 < r < \infty \,\right\rangle$, состоящей из всех открытых шаров $U_r(\vec{c}) = \{\vec{x} \in \mathbb{R}^k : \rho(\vec{x}, \vec{c}) < r\}$. Тогда σ-алгебра $\mathscr{B}(\mathbb{R}^k) := \sigma(\mathscr{O})$ — 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, $\Omega_\times = \Omega_1 \times \ldots \times \Omega_k$ — их прямое произведение. Рассмотрим 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 $B_i \in$

$\in \mathscr{F}_i$, $i = \overline{1,k}$, и хотя бы одно $B_i = \Omega_i$, принято называть цилиндрами. Сигма-алгебра, порождённая $\mathscr{F}_\times$:

$$\mathscr{F}_1 \otimes \ldots \otimes \mathscr{F}_k := \sigma(\mathscr{F}_1 \times \ldots \times \mathscr{F}_k),$$

называется *прямым произведением* σ-алгебр или *цилиндрической* σ-*алгеброй*. Если $(\Omega_i, \mathscr{F}_i)$ — копии пространства $(\Omega, \mathscr{F})$, то обозначают $\Omega_\times =: \Omega^k$, $\mathscr{F}_1 \otimes \ldots \otimes \mathscr{F}_k =: \mathscr{F}^{\otimes k}$ (или $\mathscr{F}^k$).

Определять меру, начиная с классов $\mathcal{O}$ и $\mathscr{F}_\times$, не всегда удобно. Рассмотрим класс подмножеств, полученных прямым произведением элементов полукольца $\mathcal{K}$ на прямой. Пусть $\mathcal{K}^k$ — полукольцо в $\mathbb{R}^k$, образованное параллелепипедами (прямоугольниками) вида $(\vec{a}; \vec{c}]\ :=(a_1; c_1] \times \ldots \times (a_k; c_k]$, $-\infty < a_i \leqslant c_i < \infty$, $i = \overline{1,k}$, и σ-алгебра $\sigma(\mathcal{K}^k)$ — минимальная σ-алгебра, порождённая этим полукольцом.

117 Т е о р е м а. *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 «углов» вида $\big\langle (-\infty; c_1] \times \ldots \times (-\infty; c_k] \big\rangle$.

⊫ 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 $B_1 \times B_2 \in \sigma(\mathcal{K} \times \mathcal{K})$. Отсюда будет следовать, что $\mathscr{B}^{\otimes 2} = \sigma(B_1 \times B_2 : B_1, B_2 \in \mathscr{B}) \subset \sigma(\mathcal{K} \times \mathcal{K})$.

Если $B_1 \in \mathcal{K}$, то совокупность $\mathcal{A}$ борелевских подмножеств

(‡) Здесь $\mathscr{B}^k$ только обозначение, но никак не k-кратное декартово произведение σ-алгебр.

$A \subset \mathbb{R}^1$, 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 217, стр. 233, и следующего упражнения. ⇚

118| $\mathfrak{Упр}$. Докажите $\sigma(\mathcal{K}^k)$-измеримость

- ✦ 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 $\forall\, \vec{a} < \vec{c}$ прямоугольник $(\vec{a}; \vec{c}] \in \sigma(\mathcal{K}^{\urcorner})$.

⚠ Различные характеризации борелевской σ-алгебры используются для описания различных объектов. Так, соотношение $\mathscr{B}^k = \sigma(\mathcal{K}^k)$ 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.

119| Лемма. *Пусть* $(\Omega_n, \mathscr{F}_n)$, $n = \overline{1, k}$, — *измеримые пространства. Тогда*

$$\bigotimes_{n=1}^{k} \mathscr{F}_n = (\bigotimes_{i=1}^{m} \mathscr{F}_i) \otimes (\bigotimes_{j=m+1}^{k} \mathscr{F}_j), \qquad 1 \leqslant m < k.$$

⊫ Так как $\mathscr{F}_\alpha \times \mathscr{F}_\beta \subset \mathscr{F}_\alpha \otimes \mathscr{F}_\beta = \sigma(\mathscr{F}_\alpha \times \mathscr{F}_\beta)$, то

$$\mathop{\times}_{n=1}^{k} \mathscr{F}_n = (\mathop{\times}_{i=1}^{m} \mathscr{F}_i) \times (\mathop{\times}_{j=m+1}^{k} \mathscr{F}_j) \subset (\bigotimes_{i=1}^{m} \mathscr{F}_i) \times (\bigotimes_{j=m+1}^{k} \mathscr{F}_j) \subset$$
$$\subset (\bigotimes_{i=1}^{m} \mathscr{F}_i) \otimes (\bigotimes_{j=m+1}^{k} \mathscr{F}_j).$$

Обратно, произведение $\bigotimes\limits_{i=1}^{m} \mathscr{F}_i \times \big(\mathop{\times}\limits_{j=m+1}^{k} \Omega_j\big)$, т.е. класс подмножеств вида $A \times \big(\mathop{\times}\limits_{j=m+1}^{k} \Omega_j\big)$, $A \in \bigotimes\limits_{i=1}^{m} \mathscr{F}_i$, есть σ-подалгебра σ-алгебры $\bigotimes\limits_{i=1}^{k} \mathscr{F}_i$. Поэтому

$$\bigotimes_{i=1}^{m} \mathscr{F}_i \times \bigotimes_{j=m+1}^{k} \mathscr{F}_j = \Big(\bigotimes_{i=1}^{m} \mathscr{F}_i \times (\mathop{\times}_{j=m+1}^{k} \Omega_j)\Big) \cap \Big((\mathop{\times}_{i=1}^{m} \Omega_i) \times \bigotimes_{j=m+1}^{k} \mathscr{F}_j\Big)$$
$$\subset \bigotimes_{i=1}^{k} \mathscr{F}_i,$$

что доказывает лемму. ⫤

Пространство последовательностей $\mathbb{R}^\infty$. Формализация понятия последовательности функций с заданными свойствами требует введения не только конечномерных пространств, но и бесконечных произведений измеримых пространств. Здесь при построении меры удобнее оперировать не 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 (правильнее было бы обозначить это пространство как $\mathbb{R}^\mathbb{N}$). Наиболее естественную топологию открытых множеств здесь определяет так называемая тихоновская топология (см. стр. 238), база окрестностей которой состоит из всех цилиндрических множеств вида $\prod_{i=1}^\infty B_i$, где конечное число множеств $B_{i_j} = (a_j; c_j)$, $a_j, c_j \in \mathbb{R}^1$, $j = \overline{1, n}$, $n < \infty$, остальные $B_i = \mathbb{R}^1$. Тогда борелевской естественно назвать минимальную σ-алгебру $\mathscr{B}(\mathbb{R}^\infty)$, порождённую открытыми множествами в этой топологии.

Остальные способы эксплуатируют схожую идею применительно к прямому произведению классов множеств, определяющих борелевские σ-алгебры на координатных прямых. Пусть $B \subset \mathbb{R}^n$ — подмножество n-мерного пространства. Цилиндром в $\mathbb{R}^\infty$ с основанием B будем называть множество

$$B \times \mathbb{R}_{n+1}^\infty = \{\boldsymbol{x} \in \mathbb{R}^\infty : (x_1, \ldots, x_n) \in B\}.$$

Заметим, что любой цилиндр с n-мерным основанием можно рассматривать как цилиндр с m-мерным основанием при $\forall\, m > n$. Действительно, если $B \subset \mathbb{R}^n$, то $B' = B \times \mathbb{R}^{m-n} \subset \mathbb{R}^m$ и $B \times \mathbb{R}_{n+1}^\infty = B' \times \mathbb{R}_{m+1}^\infty$.

Рассмотрим три класса цилиндрических множеств:

$$\mathcal{B} = \left\langle\, B \times \mathbb{R}_{n+1}^\infty \ : \ B \in \mathscr{B}^n, n \geqslant 1 \,\right\rangle$$

— алгебру (?!) цилиндров с борелевскими основаниями;

$$\mathcal{C} = \left\langle\, (B_1 \times \ldots \times B_k) \times \mathbb{R}_{n+1}^\infty \ : \ \{B_i\}_1^n \subset \mathscr{B}^1, n \geqslant 1 \,\right\rangle$$

— полукольцо (?!) цилиндров с основаниями в виде произведения одномерных борелевских подмножеств (борелевских параллелепипедов);

$$\mathcal{K} = \left\langle\, (\vec{a};\vec{c}\,]\times\mathbb{R}^{\infty}_{n+1} \;:\; \vec{a},\vec{c}\in\mathbb{R}^n, n\geqslant 1 \,\right\rangle$$

— класс цилиндров с параллелепипедами в основании, который будет составлять полукольцо, если к нему добавить цилиндры, в основании которых есть интервалы вида $(-\infty; c], (a; +\infty)$.

120 Т е о р е м а. *Борелевская* σ*-алгебра в* $\mathbb{R}^{\infty}$

$$\mathscr{B}^{\infty} := \mathscr{B}(\mathbb{R}^{\infty}) = \sigma(\mathcal{B}) = \sigma(\mathcal{C}) = \sigma(\mathcal{K}).$$

$\Rrightarrow$ 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 $\mathscr{B}^n = \sigma(\mathcal{K}^{(n)})$, то минимальная σ-алгебра, порождённая классом цилиндров с n-мерными параллелепипедами в основании, $\sigma(\mathcal{K}^{(n)}\times\mathbb{R}^{\infty}_{n+1}) = \mathscr{B}^n\times\mathbb{R}^{\infty}_{n+1}$. Следовательно, для любого n-мерного подмножества $B \in \mathscr{B}^n$ справедливо включение $B\times\mathbb{R}^{\infty}_{n+1} \in \sigma(\mathcal{K}^{(n)}\times\mathbb{R}^{\infty}_{n+1}) \subset \sigma(\mathcal{K})$. $\Lleftarrow$

121 $\mathfrak{Y}\mathfrak{np}$. Покажите, что $\mathscr{B}(\mathbb{R}^{\infty}) = \sigma(\mathcal{K})$.

Пространство числовых 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 индексов t. Чаще всего t интерпретируется как время, 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

$\mathbb{R}^k$ — здесь $\mathbb{T} = \{1, 2, \ldots, k\}$, и пространством последовательностей $\mathbb{R}^\infty$ — здесь $\mathbb{T} = \mathbb{N}$).

Для того чтобы определить борелевскую σ-алгебру, необходимо описать топологию открытых множеств. Наиболее удобна здесь тихоновская топология (стр. 238), в которой базис окрестностей образуют все цилиндрические множества вида $B_{t'} \times \prod_{t \neq t'} \mathbb{R}^1 = \{X \in \mathbb{R}^{\mathbb{T}} : X_{t'} \in B_{t'}\}$, $t' \in \mathbb{T}, B_{t'} \subset \mathbb{R}^1$, 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 $B_{t'}$. Непрерывность функционалов, принимающих значения в $\mathbb{R}^{\mathbb{T}}$, в тихоновской топологии эквивалентна покоординатной непрерывности, т.е. функционал $\xi : \Omega \mapsto \mathbb{R}^{\mathbb{T}}$ непрерывен т. т. т. когда для $\forall\, t_0 \in \mathbb{T}$ непрерывно отображение $\xi(\omega)_{t_0}$, $\omega \in \Omega$, в $\mathbb{R}^1$. 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.

Пусть $\mathcal{B}$ — некоторый класс борелевских подмножеств на числовой прямой $\mathbb{R}^1$. Рассмотрим класс цилиндрических множеств пространства $\mathbb{R}^{\mathbb{T}}$ с «основаниями» из $\mathcal{B}$:

$$\mathcal{Q}(\mathcal{B}) := \\ \Big\langle \bigcap_{k=1}^{n} \{X \in \mathbb{R}^{\mathbb{T}} : X_{t_k} \in B_{t_k}\} : \quad \{t_k\}_1^n \subset \mathbb{T}, \{B_{t_k}\}_1^n \subset \mathcal{B},\ n \geqslant 1 \Big\rangle.$$

⚠ Базу тихоновской топологии в $\mathbb{R}^{\mathbb{T}}$ образуют системы множеств $\mathcal{Q}(\mathcal{B})$, в которых классы $\mathcal{B}$ состоят из всех открытых подмножеств (или всех открытых интервалов) на прямой, т.е. как раз те классы множеств, которые порождают борелевскую σ-алгебру на $\mathbb{R}^1$. Было бы очень удобно, если бы, как и в конечномерных пространствах $\mathbb{R}^k$, σ-алгебра, порождённая базой топологии, совпадала с σ-алгеброй, порождённой всей топологией. Этот факт справедлив для так называемых польских пространств (полных, сепарабельных, метрических пространств), в которых топология может быть задана счётной базой.

В учебном пособии А. Н. Ширяева [12, Теорема 4, стр. 186] приведено доказательство совпадения σ-алгебр, порождённых $\mathcal{Q}(\mathcal{B})$ с естественными классами $\mathcal{B} = \mathscr{B}$ и $\mathcal{B} = \mathcal{K} = \big\langle (a;b] : a,b \in \mathbb{R}^1 \big\rangle$, с минимальной σ-алгеброй, порождённой классом цилиндров с конечномерными борелевскими основаниями:

$$\mathcal{Q}^* := \\ \Big\langle \{X \in \mathbb{R}^{\mathbb{T}} : (X_{t_1}, \ldots, X_{t_1}) \in B\} : \quad \{t_k\}_1^n \subset \mathbb{T}, B \in \mathscr{B}^n,\ n \geqslant 1 \Big\rangle.$$

Вполне естественно эту «цилиндрическую» σ-алгебру $\mathscr{B}(\mathbb{R}^{\mathbb{T}})$ назвать борелевской на $\mathbb{R}^{\mathbb{T}}$(если договориться всегда называть таковой σ-алгебру, порождённую базой топологии). Приведём без доказательства теорему с учётом используемых здесь обозначений.

122 Т е о р е м а. *Пусть* $\mathbb{T}$ — *несчётное множество. Тогда*

$$\mathscr{B}(\mathbb{R}^{\mathbb{T}}) := \sigma(\mathcal{Q}(\mathscr{B})) = \sigma(\mathcal{Q}(\mathcal{K})) = \sigma(\mathcal{Q}^*)$$

и любое подмножество $A \in \mathscr{B}(\mathbb{R}^{\mathbb{T}})$ *имеет следующую структу-*

ру: найдётся не более чем счётное множество точек $t_1, t_2, \ldots$ *из* $\mathbb{T}$ *и борелевское множество* $B \in \mathscr{B}(\mathbb{R}^{\mathbb{N}})$, *для которых* $A = \left\langle\, X : (X_{t_1}, X_{t_2}, \ldots) \in B \,\right\rangle$.

Таким образом, множества 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 $\sup_{T_0} X_t$, $X \in \mathbb{R}^{\mathbb{T}}$, или утверждениями о непрерывности или дифференцируемости X_t и т.п. К счастью, чаще всего можно предположить, что наблюдаемые реализации функции X_t принадлежат более узкому пространству функций, например, что эти реализации непрерывны всюду или не имеют разрывов второго рода.

Пространства $\mathcal{C}(\mathbb{T})$ **и** $\mathcal{D}(\mathbb{T})$. 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 [6]. 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, множества непрерывных функций вида $\{X \in \mathcal{C}(\mathbb{T}) : \ X_t \leqslant M\} \in \mathscr{B}(\mathcal{C}(\mathbb{T}))$ при каждом фиксированном $t \in \mathbb{T}$. 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 $t > 0$), если это пространство снабдить метрикой Скорохода. 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

$$\rho_S(X, Y) = \inf\{\epsilon > 0 : \exists \lambda \in \Lambda : \sup_t |X_t - Y_{\lambda(t)}| + \sup_t |t - \lambda(t)| < \epsilon\}.$$

Отметим, что сходимость $X^{(n)} \to X$ в этой метрике влечёт сходимость $X_t^{(n)} \to X_t$ во всех точках 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. Мера Лебега–Стилтьеса

Числовая прямая $\mathbb{R}^1$. Применим сначала описанный выше процесс к построению меры Лебега–Стилтьеса на числовой прямой. Пусть $F(x)$ — конечная, непрерывная справа, неубывающая функция действительного аргумента $x \in \mathbb{R}^1$.

($\mathtt{M}_1$) Полукольцо $\mathcal{K}$ состоит из интервалов вида $(a; b]$, $a \leqslant b$, $a, b \in \mathbb{R}^1$.

($\mathtt{M}_2$) Определим меру $\mu_F(a; b] = F(b) - F(a)$, неотрицательность которой следует из условий на функцию F.

($\mathtt{M}_3$) Поскольку прямую $\mathbb{R}^1$ можно представить в виде счётного объединения интервалов с целыми границами, то σ-конечность μ_F очевидна.

По лемме 94, стр. 119, для доказательства σ-аддитивности μ_F необходимо установить её аддитивность и σ-полуаддитивность.

(✓) Аддитивность покажем, используя геометрические свойства прямой. Рассмотрим конечное $\mathcal{K}$-разбиение $(A\,;B] = \biguplus_{i=1}^{N}(a_i; b_i]$. Перегруппируем интервалы так, чтобы $A = a_1 < b_1 = a_2 < \ldots < b_{N-1} = a_N < b_N = B$. Тогда

$$\mu_F(A;B] = F(B) - F(A) \ = \ F(b_N) - F(a_N) + F(b_{N-1}) - \ldots - F(a_2) + F(b_1) - F(a_1) \ = \ \sum_{i=1}^{N} \mu_F(a_i; b_i].$$

(✓) Для доказательства σ-полуаддитивности μ_F воспользуемся топологическими свойствами $\mathbb{R}^1$ и непрерывностью справа функции F. Рассмотрим счётное $\mathcal{K}$-разбиение интервала $(A\,;B] = \biguplus_{i=1}^{\infty}(a_i; b_i]$. Выберем (пока произвольно) точки $\widetilde{A} \in (A; < B)$ и $\tilde{b}_i > b_i$. Для этих точек замкнутый отрезок

$$[\widetilde{A}\,;B] \ \subset \ (A\,;B] \ = \ \biguplus_{i=1}^{\infty}(a_i; b_i] \ \subset \ \bigcup_{i=1}^{\infty}(a_i; \tilde{b}_i).$$

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, $(\widetilde{A}\,;B]$) :

$$(\widetilde{A}\,;B] \ \subset \ [\widetilde{A}\,;B] \ \subset \ \bigcup_{i=1}^{N}(a_i; \tilde{b}_i) \ \subset \ \bigcup_{i=1}^{N}(a_i; \tilde{b}_i].$$

По доказанному в 94, стр. 119, свойству полуаддитивности

$$\mu_F(\widetilde{A}\,;B] \ \leqslant \ \sum_{i=1}^{N} \mu_F(a_i; \tilde{b}_i].$$

Воспользуемся теперь непрерывностью справа функции F и выберем для фиксированного $\epsilon > 0$ точки $\widetilde{A}, \tilde{b}_i$ так, чтобы

$$0 \leqslant F(\widetilde{A}) - F(A) < \epsilon, \qquad 0 \leqslant F(\tilde{b}_i) - F(b_i) < \frac{\epsilon}{2^i}, \quad i \geqslant 1.$$

Тогда $\mu_F(A\,;B] - \mu_F(\widetilde{A}\,;B] = F(\widetilde{A}) - F(A) < \epsilon$, и при $\forall\, i \geqslant$ $\geqslant 1$ разность мер $\mu(a_i;\tilde{b}_i] - \mu_F(a_i;b_i] = F(\tilde{b}_i) - F(b_i) < \epsilon 2^{-i}$. Следовательно,

$$\mu_F(A;B] < \mu_F(\widetilde{A}\,;B] + \epsilon \leqslant \sum\nolimits_1^N \mu_F(a_i;\tilde{b}_i] + \epsilon <$$
$$< \sum\nolimits_1^\infty \mu_F(a_i;b_i] + \sum\nolimits_1^\infty \frac{\epsilon}{2^i} + \epsilon = \sum\nolimits_1^\infty \mu_F(a_i;b_i] + 2\epsilon.$$

Полагая здесь $\epsilon \to 0$, получаем искомое неравенство, а вместе с ним доказательство σ-аддитивности μ_F.

($\mathtt{M}_4$-$\mathtt{M}_5$) Борелевская σ-алгебра $\mathscr{B}$ порождается полукольцом интервалов, поэтому по теореме Каратеодори (следствию 106, стр. 129) следует, что на $\mathscr{B}$ существует единственная мера μ, для которой $\mu(a;b] = \mu_F(a;b] = F(b) - F(a)$.

О п р е д е л е н и я. 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. 133), 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. Более того, 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 определить меру Лебега–Стилтьеса с помощью непрерывной *слева* функции F на полукольце интервалов вида $[a;b)$ по формуле $\mu_F[a;b) = F(b) - F(a)$.

123 | 𝔘𝔭𝔯. Покажите, что в любом случае определения меры посредством функции F всюду непрерывной справа или слева справедливы равенства:

$$\mu_F(a;b] = F(b+) - F(a+), \qquad \mu_F[a;b) = F(b-) - F(a-),$$
$$\mu_F[a;b] = F(b+) - F(a-), \qquad \mu_F(a;b) = F(b-) - F(a+),$$
$$\mu_F(\{b\}) = F(b+) - F(b-).$$

Кстати, можно вовсе не требовать непрерывности функции F, если определить меру различных интервалов в соответствии с равенствами этой задачи.

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: $(0,0) < (1,0)$.

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.

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 92, стр. 117).

Мера Дира́ка δ_{x_0}, сосредоточенная в т. x_0, порождается индикаторной функцией $F(x) = \mathbf{I}_{[x_0;\infty)}(x)$.

Меры на индуцированном подпространстве. Очень часто вместо всего пространства $\mathbb{R}^1$ рассматривают его конечную часть, например, отрезок $[0;1]$. Поскольку этот отрезок измерим относительно борелевской σ-алгебры $\mathscr{B}(\mathbb{R}^1)$, то индуцированная борелевская σ-алгебра $\mathscr{B}([0;1]) \subset \mathscr{B}(\mathbb{R}^1)$. Поэтому любая мера, заданная на $\mathscr{B}(\mathbb{R}^1)$, естественным образом индуцирует меру на отрезке $[0;1]$. Следует сказать, что мера Лебега–Стилтьеса, порождённая функцией F, индуцированная на $[0;1]$, 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 борелевские множества из этого отрезка. В то же время, функция $F(x) = \llbracket x \rrbracket$ 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 является мерой Лебега–Стилтьеса в том узком понимании, как мера, порождённая непрерывной справа (или слева) генерирующей функцией.

Чтобы избежать подобных трудностей, вместо замкнутого отрезка $[0;1]$ рассматривают 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.

Евклидово пространство $\mathbb{R}^k$. Как и в одномерном случае, рассмотрим непрерывную справа функцию $F(\vec{x})$, $\vec{x} \in \mathbb{R}^k$: $F(\vec{x} + \vec{\epsilon}) \to F(\vec{x})$, $\vec{\epsilon} \geqslant \vec{0}$, $\vec{\epsilon} \to \vec{0}$, где неравенство $\vec{\epsilon} \geqslant \vec{0}$ понимается покомпонентно: $\epsilon_i \geqslant 0$, $i = \overline{1,k}$. 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

$$\begin{aligned}\Delta^{(j)}_{a,c}F(\vec{x}) &= F(x_1,\ldots,x_{j-1},c,x_{j+1},\ldots,x_k) - \\ &\quad - F(x_1,\ldots,x_{j-1},a,x_{j+1},\ldots,x_k),\\ \Delta_{\vec{a},\vec{c}}F &= \Delta^{(1)}_{a_1,c_1}\ldots\Delta^{(k)}_{a_k,c_k}F(\vec{x}).\end{aligned} \tag{5}$$

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 $F(\vec{x}) \equiv x_1\cdot\ldots\cdot x_k$, так что мера параллелепипеда равна его объёму: $\mu_F(\vec{a};\vec{c}] = \prod_1^k(c_i - a_i)$, называется *мерой Лебега* в $\mathbb{R}^k$.

⚠ Функция F не обязана возрастать в «обычном смысле». Например, для $F(x_1,x_2) = x_1x_2$ значение $F(-1,-1) > F(0,0)$, хотя $(-1,-1) < (0,0)$.

124| Лемма. (?!) *Если $F(\vec{x}) = \prod_1^k F_i(x_i)$, то мера параллелепипеда $\Delta_{\vec{a},\vec{c}}F = \prod_1^k(F_i(c_i) - F_i(a_i))$.*

Доказательство корректности такого определения почти полностью повторяет одномерный случай. Единственное, что требует чуть большего внимания — это доказательство аддитивности такой меры. Рассмотрим только случай $k = 2$. Пусть прямоугольник $\Pi = (a_1;b_1]\times(a_2;b_2] = \Pi_1 + \Pi_2$, т.е. 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 $\mu_F(\Pi) = \mu_F(\Pi_1) + \mu_F(\Pi_2)$. По индукции свойство аддитивности следует для любых разбиений в виде сетки

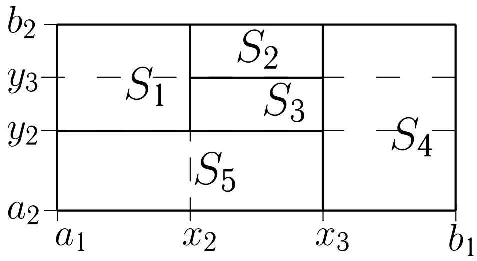
$$(a_1; b_1] \times (a_2; b_2] = \biguplus_{i=1}^{k-1} \biguplus_{j=1}^{m-1} (x_i; x_{i+1}] \times (y_j; y_{j+1}],$$

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

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.

125 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.

§ 4. 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 подмножеством $B \subset \Omega_2$.

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 пространства Ω_2 называется совокупность прообразов элементов $\mathcal{H}$:

$$\xi^{-1}(\mathcal{H}) = \left\langle\, \xi^{-1}(B) : B \in \mathcal{H} \,\right\rangle.$$

126| Лемма. (?!) *Для любых наборов* $\left\langle\, B_\alpha \,\right\rangle_{\alpha\in\mathfrak{A}} \subset \Omega_2$:

а) $\xi^{-1}(\Omega_2) = \Omega_1, \qquad \xi^{-1}(\varnothing) = \varnothing;$

б) $\xi^{-1}\Big(\bigcap\limits_{\alpha\in\mathfrak{A}} B_\alpha\Big) = \bigcap\limits_{\alpha\in\mathfrak{A}} \xi^{-1}(B_\alpha), \quad \xi^{-1}\Big(\bigcup\limits_{\alpha\in\mathfrak{A}} B_\alpha\Big) = \bigcup\limits_{\alpha\in\mathfrak{A}} \xi^{-1}(B_\alpha);$

в) $\xi^{-1}(B_1 \setminus B_2) = \xi^{-1}(B_1) \setminus \xi^{-1}(B_2), \qquad \xi^{-1}(B^c) = (\xi^{-1}(B))^c.$

⊫ Докажем только первое соотношение в):

$$\omega \in \xi^{-1}(B_1 \setminus B_2) \ \Leftrightarrow\ \xi(\omega) \in B_1 \setminus B_2 \ \Leftrightarrow\ \xi(\omega) \in B_1,\ \xi(\omega) \notin B_2$$
$$\Leftrightarrow\ \omega \in \xi^{-1}(B_1),\ \omega \notin \xi^{-1}(B_2) \ \Leftrightarrow\ \omega \in \xi^{-1}(B_1) \setminus \xi^{-1}(B_2).$$

Остальные соотношения доказываются аналогично. ⫤

127| 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*

$$\xi^{-1}(\mathscr{F}_2) \ = \ \left\langle\, \xi^{-1}(B) : \ B \in \mathscr{F}_2 \,\right\rangle,$$
$$\tilde{\xi}(\mathscr{F}_1) \ = \ \left\langle\, A \subset \Omega_2 : \ \xi^{-1}(A) \in \mathscr{F}_1 \,\right\rangle$$

образуют σ-алгебры на Ω_1 и Ω_2 соответственно.

$\Rrightarrow$ Покажем, к примеру, замкнутость $\tilde{\xi}(\mathscr{F}_1)$ относительно счётных объединений. Выберем произвольные подмножества $A_j \in \tilde{\xi}(\mathscr{F}_1)$, $j \geqslant 1$ (т.е. для них $\xi^{-1}(A_j) \in \mathscr{F}_1$). В силу 126

$$\xi^{-1}\Big(\bigcup_1^\infty A_j\Big) = \bigcup_1^\infty \xi^{-1}(A_j) \in \mathscr{F}_1. \qquad \Lleftarrow$$

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.

Определения. Сигма-алгебру $\sigma(\xi) := \xi^{-1}(\mathscr{F}_2)$ называют *сигма-алгеброй, порождённой (индуцированной) функцией ξ.*

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 $\mathscr{F}_1$: $\xi^{-1}(\mathscr{F}_2) \subset \mathscr{F}_1$. Ясно, что $\mathscr{F} = \sigma(\xi)$ 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 127 есть максимальная σ-алгебра в Ω_2

[‡] 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 $\omega_1 \in \Omega_1$ и принимает значения в Ω_2.

такая, что функция ξ $(\mathscr{F}_1|\mathscr{F})$-измерима.

Суперпозиция измеримых функций. Весьма важно, что суперпозиция измеримых функций также измерима.

128| Л е м м а. *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, тогда измерима суперпозиция* $(\eta \circ \xi)(\omega_1) = \eta(\xi(\omega_1)) : (\Omega_1, \mathscr{F}_1) \mapsto (\Omega_3, \mathscr{F}_3)$.

⊫ Доказательство почти очевидно:

$$(\eta \circ \xi)^{-1}(\mathscr{F}_3) \;=\; \xi^{-1}(\eta^{-1}(\mathscr{F}_3)) \;\subset\; \xi^{-1}(\mathscr{F}_2) \;\subset\; \mathscr{F}_1. \qquad ⫤$$

Обратно, измеримость действительной функции η относительно σ-алгебры, порождённой действительной (векторной) функцией ξ, можно интерпретировать как зависимость η от ξ.

129| Т е о р е м а. *Если* $\vec{\xi} : \Omega \mapsto \mathbb{R}^k, \eta : \Omega \mapsto \mathbb{R}$ — *функции, заданные на одном пространстве, причём* $\eta^{-1}(\mathscr{B}) \subset \vec{\xi}^{\,-1}(\mathscr{B}^k)$, *т.е.* η *измерима относительно* σ*-алгебры, порождённой функцией* $\vec{\xi}$, *то найдётся такая действительная измеримая функция* $\psi : (\mathbb{R}^k, \mathscr{B}^k) \mapsto (\mathbb{R}, \mathscr{B})$, *что* $\eta(\omega) = \psi(\vec{\xi}(\omega))$.

⊫ Рассмотрим только случай $k = 1$. Воспользуемся идеей представления измеримой функции как предела простых функций (см. теорему 135, стр. 161). Пусть $\eta = \mathbf{I}_A$ — индикаторная функция A, измеримость которой относительно σ-алгебры $\xi^{-1}(\mathscr{B})$ означает, что $A \in \xi^{-1}(\mathscr{B})$, т.е. $A = \xi^{-1}(B)$, для некоторого множества $B \in \mathscr{B}$. Таким образом, $\eta(\omega) = \mathbf{I}(\omega; \xi(\omega) \in B)$ и можно взять $\psi(x) = \mathbf{I}(x; B)$.

Если $\eta = \sum_1^n y_i \mathbf{I}(\omega; A_i)$ — простая функция с $\biguplus_1^n A_i = \Omega$, то функция $\psi(x) = \sum_1^n y_i \mathbf{I}(x; B_i)$, где $\xi^{-1}(B_i) = A_i$, причём B_i, $i = \overline{1, n}$, — непересекающиеся борелевские подмножества.

Пусть теперь η — произвольная $\xi^{-1}(\mathscr{B})$-измеримая функция. Тогда (см. теорему 137, стр. 162) существует последовательность простых функций $\eta_n(\omega) \underset{n}{\to} \eta(\omega)$ при $\forall\, \omega \in \Omega$. Пусть

$\psi_n(x)$ — последовательность простых борелевских функций, соответствующих последовательности η_n : $\eta_n(\omega) = \psi_n(\xi(\omega)) \to$ $\to \eta(\omega),\ \forall\, \omega \in \Omega$.

Множество $B = \{x : \exists \lim\limits_n \psi_n(x)\}$ измеримо по Борелю, поскольку по критерию Коши сходимости последовательностей

$$B = \bigcap_{l=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{j,m=N}^{\infty} \left\{x : \|\psi_j(x) - \psi_m(x)\| < \frac{1}{l}\right\} \in \mathscr{B}.$$

По построению $\xi(\omega) \in B,\ \forall\, \omega \in \Omega$. Определим $\psi(x) =$ $= \mathbf{I}(x; B) \lim\limits_n \psi_n(x)$. Тогда $\eta(\omega) = \psi(\xi(\omega)),\ \forall\, \omega \in \Omega$. ⫤

Измеримость функции и порождающий класс. Чтобы установить измеримость какой-либо функции, не обязательно проверять измеримость прообразов всех $\mathscr{F}_2$-измеримых множеств — достаточно проделать это для всех подмножеств из порождающего $\mathscr{F}_2$ семейства.

130| 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*

$$\xi^{-1}(\mathscr{F}_2) \subset \mathscr{F}_1 \quad \Leftrightarrow \quad \xi^{-1}(\mathcal{X}) \subset \mathscr{F}_1.$$

⊫ 1) Так как σ-алгебра $\xi^{-1}(\sigma(\mathcal{X}))$ содержит $\xi^{-1}(\mathcal{X})$, то она должна содержать и порождаемую σ-алгебру: $\xi^{-1}(\sigma(\mathcal{X})) \supset$ $\supset \sigma(\xi^{-1}(\mathcal{X}))$.

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. 127). 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. 127), то $\sigma(\mathcal{X}) \subset \tilde{\xi}(\mathscr{A}_1)$. Другими словами, прообраз любого множества из класса $\sigma(\mathcal{X})$ 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 $\xi^{-1}(\mathcal{X}) \subset \mathscr{F}_1$:

$$\xi^{-1}(\mathscr{F}_2) = \xi^{-1}(\sigma(\mathcal{X})) = \sigma(\xi^{-1}(\mathcal{X})) \subset \mathscr{F}_1. \qquad \Lleftarrow$$

Из этой леммы и теоремы 117, стр. 138, следует корректность классического определения борелевской функции.

Определение. Функция $\vec{\xi}(\omega) : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^k, \mathscr{B}^k)$ со значениями в евклидовом пространстве $\mathbb{R}^k$ называется *измеримой по Борелю* (или просто *измеримой*), если для $\forall\, \vec{x} \in \mathbb{R}^k$

$$\{\vec{\xi} \leqslant \vec{x}\} := \{\omega \in \Omega : \;\; \vec{\xi}(\omega) \leqslant \vec{x}\} \;\; \in \mathscr{F}.$$

Если исходное пространство $(\Omega, \mathscr{F}) = (\mathbb{R}^m, \mathscr{B}^m)$ также борелевское, то измеримая функция называется *борелевской*.

131 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* $\xi_j, \;\; j = \overline{1, k}$;

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* $\forall\, \vec{x} \in \mathbb{R}^k, \; r > 0$.

$\Rrightarrow$ Все утверждения следуют из того, что соответствующие классы подмножеств порождают борелевскую σ-алгебру в $\mathbb{R}^k$ (теорема 117, стр. 138). Например, если все компоненты векторной функции измеримы, то для $\forall\, B_1, \ldots, B_k \in \mathscr{B}$ прообраз $\vec{\xi}^{-1}(B_1 \times \ldots \times B_k) = \bigcap_j \xi_j^{-1}(B_j) \in \mathscr{F}$. Т.е. для класса $\mathcal{X} = \mathscr{B} \times \ldots \times \mathscr{B}$, порождающего борелевскую σ-алгебру в $\mathbb{R}^k$, имеет место: $\vec{\xi}^{-1}(\mathcal{X}) \subset \mathscr{F}$, что по предыдущей лемме гарантирует измеримость $\vec{\xi}$. Обратное утверждение в а) следует из соотношений вида: $\xi_1^{-1}(B) = \vec{\xi}^{-1}(B \times \mathbb{R}^1 \times \ldots \times \mathbb{R}^1) \in \mathscr{F}$. $\Lleftarrow$

132| 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 $\mathscr{B}(\Omega)$.*

$\Rrightarrow$ 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 нигде 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 $\mathscr{B}(\Omega)$. $\Lleftarrow$

133| Примеры. 1) Индикаторная функция $\mathbf{I}(\omega; A)$ подмножества $A \subset \Omega$ измерима т. т. т. когда измеримо A.

2) Если Ω разбито на счётное число непересекающихся измеримых подмножеств: $\Omega = \biguplus_1^\infty A_j$, то измерима функция

$$\xi(\omega) = \sum_1^\infty c_j \, \mathbf{I}(\omega; A_j), \qquad c_j \in \mathbb{R}^1, \ j = 1, 2, \ldots$$

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 $1/x : (\Omega, \mathscr{B}(\Omega)) \mapsto (\mathbb{R}^1, \mathscr{B})$, 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, измери-

ма 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 ${}^x/_y : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B})$, 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.

134| Т е о р е м а. *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.

⊫ Условие $\xi\ (= \lim_n \xi_n) \leqslant x$ имеет место, только когда «для $\forall\, k>0\ \exists N \geqslant 1$, что для $\forall\, n \geqslant N\ \ \xi_n \leqslant x + {}^1/_k$». Заменив здесь все логические операции на соответствующие теоретико-множественные операции, получаем, что множество $\{\xi \leqslant x\}$ измеримо, т.к. представимо посредством счётных операций объединения и пересечения измеримых множеств:

$$\{\omega : \xi(\omega) \leqslant x\} = \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \{\omega : \xi_n(\omega) \leqslant x + {}^1/_k\}.$$

Измеримость остальных функций следует из равенств:

$$\{\sup_n \xi_n \leqslant x\} = \bigcap\nolimits_{n=1}^{\infty} \{\xi_n \leqslant x\}, \qquad \{\inf_n \xi_n < x\} = \bigcup\nolimits_{n=1}^{\infty} \{\xi_n < x\},$$

$$\varliminf_n \xi_n = \lim_n \Big(\inf_{k\geqslant n} \xi_k\Big), \qquad \varlimsup_n \xi_n = \lim_n \Big(\sup_{k\geqslant n} \xi_k\Big)$$

с учётом утверждений теоремы 131 . ⇚

Простые функции.

Определение. Пусть $\Omega = \biguplus_1^K A_j$ — конечное $\mathscr{F}$-разбиение Ω. Действительная функция $h(\omega)$, $\omega \in \Omega$, постоянная на множествах из разбиения:

$$h(\omega) = \sum\nolimits_1^K c_j \, \mathbf{I}(\omega; A_j), \qquad c_j \in \mathbb{R}^1, \;\; A_j \in \mathscr{F}, \;\; j = \overline{1, K},$$

называется *простой функцией*.

135 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 $\xi(\omega)$:*

$$0 \leqslant h_{n-1}(\omega) \leqslant h_n(\omega) \underset{n\to\infty}{\longrightarrow} \xi(\omega), \qquad \omega \in \Omega.$$

⇛ Если ξ измерима, то для любых целых $k, n \geqslant 1$ измеримы множества

$$A_{nk} = \left\{\omega : \frac{(k-1)}{2^n} \leqslant \xi(\omega) < \frac{k}{2^n}\right\}, \qquad B_n = \{\omega : \xi(\omega) \geqslant n\}.$$

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

$$h_n(\omega) = \sum\nolimits_{k=1}^{n2^n} c_{nk}\,\mathbf{I}(\omega; A_{nk}) + n\,\mathbf{I}(\omega; B_n),$$

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, любое ω, начиная с некоторого $n > 1$, попадёт в одно из множеств A_{nk}, т.е. $0 \leqslant \xi(\omega) - h_n(\omega) < 1/2^n \underset{n}{\to} 0$.

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 $\xi^+(\omega) = \xi(\omega)\,\mathbf{I}(\omega; \xi(\omega) > 0)$ — «положительная часть» ξ, равная $\xi(\omega)$, если $\xi(\omega) > 0$, и равная 0, если $\xi(\omega) \leqslant 0$; функция $\xi^-(\omega) = (-\xi(\omega))^+ = -\xi(\omega)\,\mathbf{I}(\omega; \xi(\omega) < 0)$ — «отрицательная часть» ξ.

136| $\mathfrak{Y}\mathfrak{n}\mathfrak{p}$. 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 следует

137| Теорема. *Для любой измеримой по Борелю функции ξ найдётся 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 $\forall\,M > 0$.*

Свойства измеримых функций. Утверждение следующей леммы основано на факте измеримости суперпозиции измеримых функций (лемма 128, стр. 156).

138| 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 132 и 128. ⫤

139| 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 $\xi \neq 0$ ($Q = (-\infty; 0) \cup (0; \infty)$). ⊚

140| $\mathfrak{Upr}$. Покажите измеримость функции $\mathrm{tg}(\xi)$, если измеримая функция $\xi \neq \pi/_2 + \pi k$, $k = 0, \pm 1, \pm 2, \ldots$

141| Теорема. *Пусть функции* $\xi, \eta : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B})$ *измеримы. Тогда измеримы функции*

$$\xi + \eta, \quad \xi - \eta, \quad \xi \cdot \eta, \quad \xi/_\eta \ (\eta \neq 0), \quad \max\{\xi, \eta\}, \quad \min\{\xi, \eta\}.$$

⊫ 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 131). Следовательно, сумма $\xi + \eta : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B})$ измерима как суперпозиция измеримых функций. Аналогично для $\xi - \eta$, $\xi\eta$, $\max\{\xi, \eta\}$, $\min\{\xi, \eta\}$.

Измеримость отношения $^\xi/_\eta$ может быть доказана тем же способом. Ещё один способ доказательства использует теоремы 135 и 137 о представлении измеримых функций в виде предела простых. Пусть

$$h(\omega) = \sum\nolimits_{i=1}^{K} h_i \, \mathbf{I}(\omega; A_i), \qquad q(\omega) = \sum\nolimits_{j=1}^{N} q_j \, \mathbf{I}(\omega; B_j), \qquad (6)$$

$$\Omega = \biguplus\nolimits_{i=1}^{K} A_i = \biguplus\nolimits_{j=1}^{N} B_j, \qquad h_i, q_j \in \mathbb{R}^1, \ \ i = \overline{1, K}, \ j = \overline{1, N}.$$

Тогда $A_i = A_i \, \Omega = \biguplus_{j=1}^{N} A_i \, B_j$, $B_j = \biguplus_{i=1}^{K} B_j \, A_i$, $\Omega = \biguplus_{i=1}^{K} \biguplus_{j=1}^{N} A_i \, B_j$ и функции h, q можно записать в виде

$$h(\omega) = \sum_{i=1}^{K} \sum_{j=1}^{N} h_i \, \mathbf{I}(\omega; A_i \, B_j), \ \ q(\omega) = \sum_{i=1}^{K} \sum_{j=1}^{N} q_j \, \mathbf{I}(\omega; A_i \, B_j). \ \ (7)$$

Следовательно, функция

$$\frac{h}{q}\,(\omega) = \sum_{i=1}^{K} \sum_{j=1}^{N} \frac{h_i}{q_j} \, \mathbf{I}(\omega; A_i \, B_j)$$

есть простая функция, т.к. пересечения измеримых множеств (и объединения этих пересечений, если понадобятся) измеримы.

Для функций ξ и η по теореме 137 найдутся последовательности простых функций, что $\xi(\omega) = \lim_n h_n(\omega)$, $\eta(\omega) = \lim_n q_n(\omega)$, причём функции q_n можно выбрать так, чтобы $q_n(\omega) \neq 0$, например положив $q_n(\omega) = \pm{}^1/_{2^n}$ в области $A_{n1} = \{\omega : -{}^1/_{2^n} \leqslant \eta(\omega) \leqslant {}^1/_{2^n}\}$ (см. теорему 135). Тогда

$$\frac{\xi}{\eta}\,(\omega) = \frac{\lim_n h_n(\omega)}{\lim_n q_n(\omega)} = \lim_n \frac{h_n(\omega)}{q_n(\omega)} \quad -$$

измеримая функция, поскольку есть поточечный предел простых функций. ⋞

142| $\mathfrak{Ynp}$. Обобщите теорему 141 на счётное число слагаемых: докажите, что если ряд $\eta(\omega) = \sum_j \xi_j(\omega)$ из измеримых функций при $\forall\, \omega \in \Omega$ сходится, то функция η измерима.

§ 5. 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, 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, 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 $(\Omega, \mathscr{F})$.

Интеграл Лебега от простой неотрицательной функции.

Интегралом Лебега от простой неотрицательной функции

$$h(\omega) = \sum\nolimits_1^K c_j \, \mathbf{I}_{A_j}(\omega) \in \mathfrak{P}_+(\Omega, \mathscr{F}), \; c_j \geqslant 0, \; j = \overline{1, K}, \; K < \infty,$$

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.

143 Лемма. *Пусть* $g, h \in \mathfrak{P}_+(\Omega, \mathscr{F})$, *тогда*

($\mathrm{L_{S1}}$) $\int_\Omega h\,d\mu$ *не зависит от представления функции* h;

($\mathrm{L_{S2}}$) $\int_\Omega c\,\mathbf{I}_A\,d\mu = c\,\mu(A)$;

($\mathrm{L_{S3}}$) *если* $\mu(A) = 0$, *то* $\int_\Omega h\,\mathbf{I}_A\,d\mu = 0$;

($\mathrm{L_{S4}}$) *(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* $\int_\Omega h\,d\mu \leqslant \int_\Omega g\,d\mu$.

$\Rrightarrow$ ($\mathrm{L_{S1}}$) Пусть $h(\omega) = \sum_{i=1}^{K} c_i\,\mathbf{I}(\omega; A_i) = \sum_{j=1}^{N} q_j\,\mathbf{I}(\omega; B_j)$ 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 (6, 7). По определению

$$\int_\Omega (ah + bg)\, d\mu = \sum_{i=1}^{K} \sum_{j=1}^{N} (ac_i + bz_j)\mu(A_i B_j),$$

что после очевидных преобразований доказывает утверждение.

($\mathrm{L_{S5}}$) Снова используем (6, 7). Как и при доказательстве ($\mathrm{L_{S1}}$), легко устанавливается неравенство $c_i\, \mu(A_i B_j) \leqslant \leqslant z_j\, \mu(A_i B_j)$ при всех i, j, что доказывает теорему. $\Lleftarrow$

Интеграл Лебега от неотрицательной функции.

Пусть $\xi : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1_+, \mathscr{B})$ — неотрицательная, измеримая по Борелю функция. 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 ($\mathrm{L_{S5}}$) последовательность интегралов, рассмотренных в определении, возрастает и, следовательно, предел всегда существует или равен $+\infty$.

144| Лемма. *Интеграл $\int_\Omega \xi\, d\mu$ от неотрицательной измеримой функции не зависит от способа её приближения монотонной последовательностью простых функций:*

$$\xi(\omega) = \lim_n \uparrow x_n(\omega) = \lim_n \uparrow y_n(\omega), \omega \in \Omega; \quad \big\langle\, x_n,\ y_n \big\rangle_1^\infty \subset \mathfrak{P}_+(\Omega, \mathscr{F}),$$

$$\Downarrow$$

$$\lim_n \int\limits_\Omega x_n(\omega)\, d\mu = \lim_n \int\limits_\Omega y_n(\omega)\, d\mu.$$

$\Rrightarrow$ Здесь нам понадобится неравенство Маркова — лемма 145. Прежде всего отметим, что в силу этого неравенства мера

$\mu(\xi > \epsilon) < \infty$ при $\forall\, \epsilon > 0$, 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 $0 < \epsilon < y_*$. Тогда для множества $Q_n = \{\omega \in \Omega:\ x_n(\omega) + \epsilon > y_m(\omega)\}$, $n \geqslant 1$,

$$\begin{aligned} y_m = \ & y_m\, \mathbf{I}_{Q_n\{y_m>0\}} + y_m\, \mathbf{I}_{Q_n^c} \leqslant (x_n + \epsilon)\, \mathbf{I}_{Q_n\{y_m>0\}} + y^*\, \mathbf{I}_{Q_n^c} \leqslant \\ & \leqslant x_n + \epsilon\, \mathbf{I}_{Q_n\{y_m>0\}} + y^*\, \mathbf{I}_{Q_n^c}. \end{aligned}$$

Так как $x_n \nearrow \xi \geqslant y_m$, то $\lim_n{\uparrow}\, Q_n = \Omega$. Если $\omega \in Q_n^c$, то $\xi \geqslant y_m \geqslant x_n + \epsilon \geqslant \epsilon$, поэтому $\mu(Q_n^c) \leqslant \mu(\xi \geqslant \epsilon) < \infty$ и в силу свойства непрерывности меры $\lim_n{\downarrow}\, \mu(Q_n^c) = 0$.

Если $\omega \in Q_n \cap \{y_m > 0\} = Q_n \cap \{y_m \geqslant y_*\}$, то $x_n > y_* - \epsilon > > 0$. Следовательно, при каждом $n \geqslant 1$

$$\begin{aligned} \mu(Q_n\{y_m > 0\}) & \leqslant \lim_n{\uparrow}\, \mu(Q_n\{y_m > 0\}) \leqslant \\ & \leqslant \lim_n{\uparrow}\, \mu(x_n > y_* - \epsilon) = \mu(\xi > y_* - \epsilon) < \infty. \end{aligned}$$

Поэтому для $\forall\, \epsilon < y_*$

$$\begin{aligned} & \int_\Omega y_m\, d\mu \leqslant \int_\Omega x_n\, d\mu + \epsilon\, \mu(\xi > y_* - \epsilon) + y^* \mu(Q_n^c) \\ & \qquad \underset{n\to\infty}{\longrightarrow} \lim_n \int_\Omega x_n\, d\mu + \epsilon \mu(\xi > y_* - \epsilon) \underset{\epsilon\to 0}{\to} \lim_n \int_\Omega x_n\, d\mu. \end{aligned}$$

Таким образом, если $\lim_n \int x_n\, d\mu < \infty$, то $\lim_m \int y_m\, d\mu \leqslant \leqslant \lim_n \int x_n\, d\mu < \infty$. Произведя перестановку $x \leftrightarrow y$, получаем требуемое. $\Lleftarrow$

145 Лемма. [*Неравенство Маркова.*] *Пусть $\xi \geqslant 0$ и для последовательности $\langle\, x_n \,\rangle_1^\infty \subset \mathfrak{P}_+(\Omega, \mathscr{F})$ такой, что $\lim\limits_{n}{}_{\uparrow}\, x_n = \xi$, существует $\lim_n \int_\Omega x_n\, d\mu = \int_\Omega \xi\, d\mu$*[‡]. *Тогда для $\forall\, A > 0$*

$$\mu(\xi > A) \leqslant \frac{1}{A} \int\limits_\Omega \xi\, d\mu\,. \tag{8}$$

$\Rrightarrow$ По свойствам линейности и монотонности интеграла от простых функций и свойству непрерывности меры

$$\int_\Omega \xi\, d\mu = \lim_n \Big(\int_\Omega x_n\, \mathbf{I}_{\{x_n > A\}}\, d\mu + \int_\Omega x_n\, \mathbf{I}_{\{x_n \leqslant A\}}\, d\mu \Big) \geqslant$$
$$\geqslant A \lim_n{}_{\uparrow}\, \mu(x_n > A) = A\, \mu(\xi > A). \qquad \Lleftarrow$$

Другое определение интеграла от положительной функции, не требующее проверки корректности, содержит следующая теорема. Обозначим через $L(\xi) = \langle\, h \in \mathfrak{P}_+(\Omega, \mathscr{F}) : \ 0 \leqslant h \leqslant \leqslant \xi \,\rangle$ класс всех положительных простых функций, «подпирающих» неотрицательную функцию ξ снизу.

146 Теорема. *Пусть $\xi : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1_+, \mathscr{B})$ — измеримая функция. Тогда*

$$\int\limits_\Omega \xi\, d\mu = \sup_{h \in L(\xi)} \int\limits_\Omega h\, d\mu.$$

$\Rrightarrow$ 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 $\langle\, q_n \,\rangle_1^\infty \subset \subset L(\xi)$ — 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, очевидно, возрастает и принадлежит $L(\xi)$. Так как $h_n \leqslant z_n$, то

[‡] Здесь $\int_\Omega \xi\, d\mu$, понимаемое пока как обозначение для $\lim_n \int_\Omega x_n\, d\mu$, после леммы 144 становится полновесным интегралом.

$\int z_n\, d\mu \to J$. С другой стороны, $q_n \leqslant z_n \leqslant \xi$, следовательно, $z_n \nearrow \xi$ и по лемме 144 $\int \xi\, d\mu = \lim_n \int z_n\, d\mu = J$. ⫤

Приведём свойства интеграла от неотрицательных функций.

147 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 $\int_\Omega \eta\, d\mu < \infty$

$$\int_\Omega (a\xi - b\eta)\, d\mu = a\int_\Omega \xi\, d\mu - b\int_\Omega \eta\, d\mu.$$

⊫ ($\mathrm{L_{p1}}$) Пусть $W = \{\omega : \xi(\omega) = \eta(\omega)\}$, тогда для любой функции $h \in L(\xi)$ простая функция $q = h\,\mathbf{I}_W \leqslant \eta$, т.е. $q \in L(\eta)$, причём по свойству ($\mathrm{L_{S3}}$) для простых функций интегралы от h и q равны:

$$\int_\Omega h\, d\mu = \int_\Omega h\,\mathbf{I}_W\, d\mu + \int_\Omega h\,\mathbf{I}_{W^c}\, d\mu = \int_\Omega q\, d\mu.$$

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.

($\mathrm{L_{p2}}$) По неравенству Маркова (лемма 145) для $\forall\, k > 0$

$$\mu\Big(\xi > \frac{1}{k}\Big) \leqslant k \int_\Omega \xi\, d\mu.$$

Если интеграл равен нулю, то $\mu(\xi > {}^1/_k) = 0$ и $\mu(\xi > 0) = = \lim_k{}_{\uparrow}\, \mu(\xi > {}^1/_k) = 0$. Обратно, если $\mu(\xi > 0) = 0$, то по свойству ($\mathrm{L_{p1}}$) $\int \xi\, d\mu = 0$.

($\mathrm{L_{p3}}$) В силу свойства ($\mathrm{L_{p1}}$) достаточно рассмотреть случай $\xi(\omega) \leqslant \eta(\omega)$ при $\forall\, \omega \in \Omega$. В соответствии с ($\mathrm{L_{p4}}$)

$$\int_\Omega \eta\, d\mu = \int_\Omega \xi\, d\mu + \int_\Omega (\eta - \xi)\, d\mu \ \geqslant\ \int_\Omega \xi\, d\mu;$$

причём ввиду ($\mathrm{L_{p2}}$) равенство здесь будет достигаться, лишь когда $\mu(\xi < \eta) = 0$, если, конечно, $\int \eta\, d\mu \neq \infty$.

($\mathrm{L_{p4}}$) Если $h_n, q_n \in \mathfrak{P}_+(\Omega, \mathscr{F})$ и $\lim_n{}_{\uparrow}\, h_n = \xi$, $\lim_n{}_{\uparrow}\, q_n = \eta$, то $\lim_n{}_{\uparrow}\, (a h_n + b q_n) = (a\xi + b\eta)$ и по определению интеграла

$$\int_\Omega (a\xi + b\eta)\, d\mu \ =\ \lim_n{}_{\uparrow} \int_\Omega (a h_n + b z_n)\, d\mu = a \int_\Omega \xi\, d\mu + b \int_\Omega \eta\, d\mu,$$

в силу линейности интеграла от простых функций.

Второе утверждение этого пункта следует из первого:

$$a \int_\Omega \xi\, d\mu \ =\ \int_\Omega \big((a\xi - b\eta) + b\eta\big)\, d\mu = \int_\Omega (a\xi - b\eta)\, d\mu + b \int_\Omega \eta\, d\mu.$$

Заметим, что справедливость заявленного соотношения следует из приведённого равенства только в случае $\int_\Omega \eta\, d\mu < \infty$. ⫥

⚠ Однозначность определения интеграла Лебега посредством произвольно взятой возрастающей последовательности простых функций выгодно отличает его от интеграла Римана. Интеграл Лебега можно определить, воспользовавшись конкретным приближением, построенным в теореме 135, стр. 161:

$$\int\limits_{\Omega} \xi\, d\mu \;=\; \lim_{n} \left[\sum_{j=1}^{n2^n} \frac{j-1}{2^n} \mu \left(\frac{j-1}{2^n} \leqslant \xi < \frac{j}{2^n} \right) + n\mu\,(\xi \geqslant n) \right],$$

где предел существует или равен $+\infty$.

Интеграл Лебега от измеримой функции. Пусть $\xi : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1, \mathscr{B})$ — измеримая по Борелю функция, ξ^+, ξ^- — её положительная и отрицательная части. Интегралом от функции 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.

148| Теорема. [*Неравенство треугольника.*] *Функция ξ интегрируема т. т. т. когда интегрируема функция $|\xi|$, при этом*

$$\Big| \int\limits_{\Omega} \xi\, d\mu \Big| \;\leqslant\; \int\limits_{\Omega} |\xi|\, d\mu.$$

$\Rrightarrow$ 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:

$$\Big| \int_{\Omega} \xi^+\, d\mu - \int_{\Omega} \xi^-\, d\mu \Big| \leqslant \int_{\Omega} \xi^+\, d\mu + \int_{\Omega} \xi^-\, d\mu. \qquad \Lleftarrow$$

⚠ Если интеграл существует, но функция не интегрируема, то неравенство треугольника также верно: $+\infty \leqslant +\infty$.

149| $\mathfrak{Y}\mathfrak{n}\mathfrak{p}$. Докажите, что если $\int \eta\, d\mu < \infty$ и $|\xi| \leqslant \eta$, то ξ интегрируема и $\left|\int \xi\, d\mu\right| \leqslant \int \eta\, d\mu$.

Что можно сказать об интеграле $\int \xi\, d\mu$, если $\int |\xi|\, d\mu = \infty$?

Свойства интеграла Лебега.

150| 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* $\forall\, c \in \mathbb{R}^1$;

(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.*

⊫ Т.к. 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
$$\int (c\xi)^+ - \int (c\xi)^- = c\left(-\int \xi^- + \int \xi^+\right) = c\int \xi.$$

(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 слагаемых, получаем утверждение теоремы.

(L_4) Нам понадобятся следующие соотношения:

$$(*) \qquad (\xi+\eta)^+ \leqslant \xi^+ + \eta^+, \qquad (\xi+\eta)^- \leqslant \xi^- + \eta^-,$$

$$\binom{*}{*} \qquad (\xi+\eta)^- = \left[\xi^- + \eta^-\right] - \left[(\xi^+ + \eta^+) - (\xi+\eta)^+\right] \ (\geqslant 0).$$

Неравенства $(*)$ проверяются перебором. Равенство $\binom{*}{*}$ следует из равенств $\xi^+ - \xi^- + \eta^+ - \eta^- = \xi+\eta = (\xi+\eta)^+ - (\xi + {} + \eta)^-$. Заметим, что в $\binom{*}{*}$ оба выражения, стоящие в квадратных скобках, неотрицательны.

Из всех ситуаций, когда выражение $\int\xi + \int\eta$ имеет смысл, достаточно рассмотреть три.

✓ В случае $\int\xi^+ < \infty$, $\int\eta^+ < \infty$, $\int\xi^- = \int\eta^- = \infty$ из $(*)$ следует, что $\int(\xi+\eta)^+ < \infty$, а из $\binom{*}{*}$, ввиду линейности интеграла от положительных функций, что $\int(\xi+\eta)^- = \infty$. Поэтому

$$\int(\xi+\eta)^+ - \int(\xi+\eta)^- = -\infty = \int\xi + \int\eta.$$

✓ Ситуация $\int\xi^\pm < \infty$, $\int\eta^+ < \infty$, $\int\eta^- = \infty$ разбирается аналогично предыдущему.

✓ Если $\int\xi^\pm < \infty$, $\int\eta^\pm < \infty$, то, как и раньше, $\int(\xi+\eta)^\pm < {} < \infty$ и из $\binom{*}{*}$ имеем:

$$\int(\xi+\eta)^+ = \int\xi^+ + \int\eta^+ - \int\xi^- - \int\eta^- + \int(\xi+\eta)^-.$$

Откуда (L_4) следует простыми рассуждениями. ⫷

151| $\mathfrak{Ynp}$. Докажите свойство (L_1).

152| $\mathfrak{Ynp}$. Покажите, что для конечной функции h интеграл относительно меры Дира́ка δ_{ω_0}, сосредоточенной в т. ω_0,

равен
$$\int_\Omega h\,d\delta_{\omega_0} = h(\omega_0).$$

Интеграл от расширенной функции. 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 (8), если функция ξ интегрируема, то $\mu(\phi_\xi^c) = = \lim\limits_{N}{}_{\downarrow}\, \mu(\xi > N) \leqslant \lim\limits_{N} \frac{1}{N} \int_\Omega \xi\,d\mu = 0$.

153| 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* $\lim\limits_{n}{}_{\uparrow} \int_\Omega h_n\,d\mu = +\infty$.

Итак, можно положить по определению $\int_\Omega \xi\,d\mu = +\infty$, если $\mu(\xi = +\infty) > 0$, и $\int_\Omega \xi\,d\mu = \int_\Omega \xi\,\mathbf{I}_{\phi_\xi}\,d\mu$ в противном случае.

Интеграл по сумме мер. Иногда мера, по которой производится интегрирование, может быть представлена в виде конечной суммы мер.

154| 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* $\xi : (\Omega, \mathscr{F}) \mapsto (\mathbb{R}^1_+, \mathscr{B})$
$$\int_\Omega \xi\,d(a_1\mu_1 + a_2\mu_2) = a_1 \int_\Omega \xi\,d\mu_1 + a_2 \int\limits_\Omega \xi\,d\mu_2.$$

155| $\mathfrak{Y}$пр. Используя 152, 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 о предельном переходе под знаком интеграла. В редких случаях, как при доказательстве линейности, можно обойтись алгебраическими методами — чаще приходится прибегать к методам, использующим асимптотические свойства интеграла.

156| Теорема. [*О монотонной сходимости. Б.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*

$$\lim_n{}_\uparrow \int\limits_\Omega \xi_n \, d\mu = \int\limits_\Omega \eta \, d\mu.$$

$\Rrightarrow$ 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 $\int \xi_n \, d\mu < \infty$, $n \geqslant 1$, — в общем случае необходимо 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 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 $\zeta = \lim\limits_{k}{}_{\uparrow}\, z_k = \eta$ и

$$\int_\Omega \eta\, d\mu = \lim_{k}{}_{\uparrow} \int_\Omega z_k\, d\mu \leqslant \lim_{k}{}_{\uparrow} \int_\Omega \xi_k\, d\mu \leqslant \int_\Omega \eta\, d\mu. \qquad \Lleftarrow$$

⚠ Теорема о монотонной сходимости верна и для последовательности, предельные значения которой могут равняться $+\infty$.

157| Лемма. (?!) *Пусть* $\xi_n \geqslant 0$, $n \geqslant 1$, *тогда*

$$\int_\Omega \sum\nolimits_1^\infty \xi_n\, d\mu = \sum\nolimits_1^\infty \int_\Omega \xi_n\, d\mu.$$

Следующее утверждение даёт одно из наиболее действенных средств, наряду с теоремой о монотонной сходимости и теоремой Лебега, для анализа последовательностей интегралов.

158| Лемма. [*П. Фату.*] *Если* $\xi_n \geqslant 0$ — *последовательность измеримых функций* $(n \geqslant 1)$, *то*

$$\underline{\lim_{n}} \int_\Omega \xi_n\, d\mu \geqslant \int_\Omega \underline{\lim_{n}}\, \xi_n\, d\mu.$$

$\Rrightarrow$ Напомним, что $\underline{\lim\limits_{n}}\, \xi_n = \lim\limits_{n}{}_{\uparrow} (\inf_{k\geqslant n} \xi_k)$. 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

$$\int\limits_\Omega \underline{\lim_{n}}\, \xi_n\, d\mu = \int\limits_\Omega \lim_{n}{}_{\uparrow} \gamma_n\, d\mu = \lim_{n}{}_{\uparrow} \int\limits_\Omega \gamma_n\, d\mu \leqslant \underline{\lim_{n}} \int\limits_\Omega \xi_n\, d\mu.$$

$\Lleftarrow$

159| Следствие.

а) *Если* $\xi_n \geqslant \eta$, $n \geqslant 1$, *причём* $\int \eta\, d\mu > -\infty$, *то*

$$\underline{\lim_{n}} \int_\Omega \xi_n\, d\mu \geqslant \int_\Omega \underline{\lim_{n}}\, \xi_n\, d\mu.$$

б) *Если* $\xi_n \leqslant \eta$, $n \geqslant 1$, *причём* $\int \eta\, d\mu < +\infty$, *то*

$$\overline{\lim_n}\int_\Omega \xi_n\,d\mu \leqslant \int_\Omega \overline{\lim_n}\,\xi_n\,d\mu.$$

в) *Если* $|\xi_n| \leqslant \eta,\ n \geqslant 1,$ *причём* $\int \eta\,d\mu < +\infty,$ *то*

$$\int_\Omega \underline{\lim_n}\,\xi_n\,d\mu \leqslant \underline{\lim_n}\int_\Omega \xi_n\,d\mu \leqslant \overline{\lim_n}\int_\Omega \xi_n\,d\mu \leqslant \int_\Omega \overline{\lim_n}\,\xi_n\,d\mu.$$

$\Rrightarrow$ б) Если $\int \eta\,d\mu = -\infty$, то по свойству монотонности $\int \xi_n\,d\mu \leqslant \int \eta\,d\mu = -\infty$ и, аналогично, $\int \overline{\lim_n}\,\xi_n\,d\mu \leqslant \int \eta\,d\mu = -\infty$. 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 $\underline{\lim_n}\,\zeta_n = \eta - \overline{\lim_n}\,\xi_n$. Следовательно, по лемме Фату

$$\int_\Omega \eta\,d\mu - \underline{\lim_n}\int_\Omega \xi_n\,d\mu \geqslant \int_\Omega \eta\,d\mu - \int_\Omega \overline{\lim_n}\,\xi_n\,d\mu,$$

что доказывает б), т.к. $\left|\int \eta\,d\mu\right| < \infty$. Утверждение а) доказывается переходом к $-\xi_n$, в) следует из а), б). $\Lleftarrow$

При анализе возможности предельного перехода под знаком интеграла зачастую требуется расширенный вариант сходимости последовательности функций. Пусть $\mathfrak{A}$ — некоторое высказывание об элементах ω пространства $(\Omega, \mathscr{F}, \mu)$ с мерой, 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 $\mathfrak{A}\,[\mu]$.

160 Примеры.

1) $\xi_n \underset{n\to\infty}{\to} \xi_0$ [п.в.] $\Leftrightarrow \mu\big(\omega : \{\nexists \lim_n \xi_n(\omega)\} \cup \{\exists \lim_n \xi_n(\omega) \neq \neq \xi_0(\omega)\}\big) = 0.$

2) $\quad \xi = \eta\,[\mu] \quad \Leftrightarrow \quad \mu(\omega : \xi(\omega) \neq \eta(\omega)) = 0.$ ⊚

161| Т е о р е м а. [*О мажорируемой сходимости. Лебег.*] *Если измеримые функции* $\eta, \xi_0, \xi_1, \ldots$ *таковы, что* μ*-п.в.*

$$\text{a)}\ \lim_n \xi_n = \xi_0, \quad \text{b)}\ |\xi_n| \leqslant \eta,\ n \geqslant 1, \quad \textit{и} \quad \text{c)}\ 0 \leqslant \int_\Omega \eta\, d\mu < \infty,$$

то ξ_0 *интегрируема и*

$$\text{i)}\ \lim_n \int_\Omega \xi_n\, d\mu = \int_\Omega \xi_0\, d\mu, \qquad \text{ii)}\ \lim_n \int_\Omega |\xi_n - \xi_0|\, d\mu = 0.$$

⊫ 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 $\forall\, n \geqslant 0$

$$\int_\Omega \xi_n\, \mathbf{I}_A\, d\mu = \int_\Omega \xi_n\, d\mu, \qquad \int_\Omega \eta\, \mathbf{I}_A\, d\mu = \int_\Omega \eta\, d\mu$$

и, кроме того, $\lim_n \xi_n(\omega)\, \mathbf{I}_A(\omega) = \xi_0(\omega)\, \mathbf{I}_A(\omega)$ и $\xi_n(\omega)\, \mathbf{I}_A(\omega) \leqslant \leqslant \eta(\omega)\, \mathbf{I}(\omega)$ при всех $\omega \in \Omega,$ что доказывает теорему. ⫤

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

$$\mu_\xi(B) \ = \ \mu(\xi^{-1}(B)), \quad B \in \mathscr{Q}.$$

162| Л е м м а. (?!) *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.

163 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* $\int_{\mathfrak{X}} \eta(x)\mu_\xi(dx) = \int_\Omega \eta(\xi(\omega))\,\mu(d\omega)$.

⊫ Воспользуемся схемой построения интеграла Лебега.

✓ Пусть сначала $\eta = \mathbf{I}_B$, 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, $\eta(\xi(\omega)) = \mathbf{I}(\omega; \xi(\omega) \in B) = \mathbf{I}(\omega; \xi^{-1}(B))$. 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 в силу предыдущего $\int_{\mathfrak{X}} \eta^\pm\, d\mu_\xi = \int \eta^\pm \circ \xi\, d\mu$. Если, например, $\int_{\mathfrak{X}} \eta^+\, d\mu_\xi < \infty$, то $\int \eta^+ \circ \xi\, d\mu < \infty$. Таким образом, если существует один из интегралов, то существует другой 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 $\xi \cdot \mathbf{I}_A$ обозначается $\int_A \xi(\vec{x})\, dF(\vec{x})$ (по поводу возможных интерпретаций этой записи см. 192, стр. 204).

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 $\int_a^b \xi(x)\, dx$.

164 Лемма. *Для* $\forall\, b \neq 0, \vec{c} \in \mathbb{R}^k$

$$\int\limits_{\mathbb{R}^k} \eta(\vec{x})\, d\vec{x} = |b| \int\limits_{\mathbb{R}^k} \eta(b\,\vec{t} + \vec{c})\, d\vec{t},$$

если хотя бы один из интегралов существует.

$\Rrightarrow$ 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 $b < 0$ оставляем читателю.

$\Lleftarrow$

165 $\mathfrak{Y}\mathfrak{n}\mathfrak{p}$. Докажите, что мера Лебега инвариантна относительно сдвигов и однородна:

$$\lambda(bA+c) = |b|\lambda(A),\ A \in \mathscr{B}^k,\ b, c \in \mathbb{R}^1.$$

§ 6. Некоторые интегральные неравенства

Обсуждение популярных неравенств, потребность в которых возникает при изучении интегральных выражений, начнём с неравенства Йенсена для выпуклых функций (см. определения и свойства на стр. 241). 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

$$\mathbf{E}\vec{\xi} = \mathbf{E}(\xi_1, \ldots, \xi_k) := (\mathbf{E}\xi_1, \ldots, \mathbf{E}\xi_k).$$

166 [*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* $(\vec{a}; \vec{c})$ *пространства* $\mathbb{R}^k$. *Если измеримая по Борелю функция* $\vec{\xi} = (\xi_1, \ldots, \xi_k) : \Omega \mapsto (\vec{a}; \vec{c})$ *такова, что* $\mathbf{E}\big|\xi_j\big| < \infty$, $j = \overline{1, k}$, *то справедливо соотношение*

$$\mathbf{E}\big[\, h \circ \vec{\xi}\, \big] \geqslant h\big(\mathbf{E}\vec{\xi}\,\big), \tag{9}$$

со строгим неравенством, когда функция h *строго выпуклая и* $\mu(\vec{\xi} \neq \vec{x}_0) > 0$ *при* $\forall\, \vec{x}_0 \in (\vec{a}; \vec{c})$. *Для вогнутой функции* h *справедливо противоположное неравенство.*

$\Rrightarrow$ Выберем $\vec{x}_0 = \mathbf{E}\vec{\xi}$; очевидно, что $\vec{a} < \vec{x}_0 < \vec{c}$. Построим опорную («касательную») плоскость $(\circledast)$, стр. 241, 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 $\vec{x} = \vec{\xi}$) 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 $\Big(\int_\Omega |\xi|\,d\mu\Big)^a \leqslant \int_\Omega |\xi|^a\,d\mu$. Применяя это неравенство к $a = {}^r/_s$, $r > s > 0$, и заменяя функцию $\xi \dashrightarrow \xi^s$, получаем следующее утверждение.

167| [*Неравенство Ляпунова.*] *Пусть* $\xi : (\Omega, \mathscr{F}, \mu) \mapsto (\mathbb{R}^1, \mathscr{B})$ *— измеримая функция на пространстве с вероятностной мерой. Тогда для любых* $r > s > 0$

$$\Big(\int_\Omega |\xi|^s\,d\mu\Big)^{1/s} \leqslant \Big(\int_\Omega |\xi|^r\,d\mu\Big)^{1/r}$$

со строгим неравенством, если $\mu(\xi \neq x_0) > 0$ *при* $\forall\, x_0 \in \mathbb{R}^1$.

168| Упр. Приведите пример с не вероятностной мерой, когда неравенство Ляпунова нарушается.

Как следует из примера, рассмотренного в дополнении (стр. 244), функция $h(x_1, x_2) = x_1^a x_2^{1-a}$, $x_1 > 0, x_2 > 0$, вогнута при $0 < a < 1$, поэтому согласно неравенству Йенсена $\mathbf{E}\big[|\xi_1|^a|\xi_2|^{1-a}\big] \leqslant \big(\mathbf{E}|\xi_1|\big)^a\big(\mathbf{E}|\xi_2|\big)^{1-a}$. Положим $a = {}^1/_p, 1 - a = {}^1/_q, {}^1/_p + {}^1/_q = 1$ и $|\xi_1|^a \dashrightarrow |\xi_1|$, $|\xi_2|^{1-a} \dashrightarrow |\xi_2|$.

169| [*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* $p, q > 1$, ${}^1/_p + {}^1/_q = 1$,

$$\int_\Omega |\xi_1\,\xi_2|\,d\mu \leqslant \Big(\int_\Omega |\xi_1|^p\,d\mu\Big)^{1/p}\Big(\int_\Omega |\xi_2|^q\,d\mu\Big)^{1/q}.$$

⊫ Выше неравенство было доказано для вероятностной меры μ. Для произвольной меры μ неравенство выполняется тривиальным образом, если хотя бы один интеграл в правой части равен $+\infty$. Если оба эти интеграла конечны, достаточно рассмотреть вспомогательную вероятностную меру (см. 172 ниже) $\tilde{\mu}(A) = \int_A |\xi_1|^p \, d\mu / C$, где $C = \int_\Omega |\xi_1|^p \, d\mu$, и применить неравенство Гёльдера к функциям $\xi_1 = 1, \xi_2 = \xi_2 / |\xi_1|^{p/q}$:

$$\int_\Omega 1 \frac{|\xi_2|}{|\xi_1|^{p/q}} \frac{|\xi_1|^p}{C} d\mu \leqslant 1 \cdot \Big(\int_\Omega \frac{|\xi_2|^q}{|\xi_1|^p} \frac{|\xi_1|^p}{C} d\mu \Big)^{1/q}.$$

Осталось только заметить, что $p - {}^p\!/_q = 1$, $C^{1-1/q} = C^{1/p}$, и если $C = 0$, то неравенство Гёльдера, очевидно, справедливо по свойству интеграла Лебега от положительной функции. ⟸

⚠ Если один из интегралов правой части неравенства Гёльдера равен нулю, а второй $+\infty$, то это неравенство также остаётся верным (?!), если договориться считать $0 \cdot \infty = 0$.

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.).

170 [`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,* $\xi_1 = c\xi_2$ [μ-*п.в.*] *с некоторой константой* c.

⊫ Примем сокращённую форму записи $\int_\Omega \eta \, d\mu = \int \eta$.

II) 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 $c = -\left[\int \xi_1^2 / \int \xi_2^2\right]^{1/2}$. После возведения в квадрат правой части равенства $0 = \left(\sqrt{\int \xi_1^2} + c\sqrt{\int \xi_2^2}\right)^2$ и применения равенства I) (с учётом знака z), получим $\int (\xi_1 - c\,\xi_2)^2 = 0$. По свойству интеграла отсюда следует, что $\xi_1 = c\xi_2$ [μ-п.в.]. ⫷

171 [*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* $\forall\, p > 1$

$$\left(\int_\Omega |\xi_1 + \xi_2|^p \, d\mu\right)^{1/p} \leqslant \left(\int_\Omega |\xi_1|^p \, d\mu\right)^{1/p} + \left(\int_\Omega |\xi_2|^p \, d\mu\right)^{1/p}.$$

⫸ Понятно, что доказательства требует только случай, когда оба интеграла в правой части конечны. В этом случае интеграл в левой части также конечен, поскольку ввиду неравенства (1), стр. 243, 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. ${}^1/_q + {}^1/_p = 1$, тогда в силу неравенства треугольника и неравенства Гёльдера имеем (в сокращённой записи интеграла по Ω относительно меры μ):

$$\begin{aligned}\int |\xi_1 + \xi_2|^p = \int |\xi_1 + \xi_2||\xi_1 + \xi_2|^{p-1} &\leqslant \\ \leqslant \int |\xi_1||\xi_1 + \xi_2|^{p-1} + \int |\xi_2||\xi_1 + \xi_2|^{p-1} &\leqslant \\ \leqslant \left(\int |\xi_1|^p\right)^{1/p} \left(\int |\xi_1 + \xi_2|^{(p-1)q}\right)^{1/q} &+ \\ + \left(\int |\xi_2|^p\right)^{1/p} \left(\int |\xi_1 + \xi_2|^{(p-1)q}\right)^{1/q} &=\end{aligned}$$

$$= \Big(\int |\xi_1+\xi_2|^p\Big)^{1-1/p}\Big[\Big(\int |\xi_1|^p\Big)^{1/p} + \Big(\int |\xi_2|^p\Big)^{1/p}\Big],$$

что эквивалентно доказываемому (при конечной левой части).

⇚

⚠ В анализе пространство функций, μ-интегрируемых со степенью p, 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 $p > 1$ 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 $\mathcal{L}^p \subset \mathcal{L}^m, \ 1 < p < m.$

Из неравенства Коши–Буняковского следует, что L^2-норма может быть задана скалярным произведением $(\xi_1, \xi_2) := := \int_\Omega \xi_1 \xi_2 \, d\mu,$ каковое превращает $\mathcal{L}^2$ в гильбертово пространство (снова с оговоркой об отождествлении функций, совпадающих почти всюду).

Предыдущие неравенства могут быть сформулированы как неравенства для L^p-норм:

$$\|\xi\|_s \leqslant \|\xi\|_r \quad (s < r); \qquad \|\xi_1\xi_2\|_1 \leqslant \|\xi_1\|_p \, \|\xi_2\|_q \quad ({}^1/_p + {}^1/_q = 1).$$

§ 7. Теорема Радона–Никодима

Для дальнейшего нам понадобятся следующие

О п р е д е л е н и я. Пусть ψ, μ — меры на измеримом пространстве $(\Omega, \mathscr{F})$. Мера ψ называется

— *абсолютно-непрерывной* относительно $\mu : \psi \ll \mu$, 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 $N \in \mathscr{F}$, что $\psi(N^c) = \mu(N) = 0$.

Пусть $\xi(\omega) : (\Omega, \mathscr{F}, \mu) \mapsto (\overline{\mathbb{R}}^1_+, \mathscr{B})$ — измеримая неотрицательная функция на пространстве с мерой. Определим функцию множеств

$$M_\xi(B) \;=\; \int_B \xi \, d\mu := \int_\Omega \xi \, \mathbf{I}_B \, d\mu, \qquad B \in \mathscr{F}.$$

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 $\forall\, B \in \mathscr{F}$.

172 Т е о р е м а. *Для любой неотрицательной $\mathscr{F}$-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, тогда 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*

$$\int_\Omega \eta \, dM_\xi = \int_\Omega \eta\xi \, d\mu.$$

$\Rrightarrow$ a) Пусть измеримое множество $B = \biguplus_n A_n$, тогда $\mathbf{I}_B = \sum_n \mathbf{I}_{A_n}$. Из предложения 157 следует, что M_ξ — неотрицательная σ-аддитивная функция:

$$M_\xi(B) = \int_\Omega \sum_1^\infty \xi \, \mathbf{I}_{A_n} \, d\mu = \sum_1^\infty \int_\Omega \xi \, \mathbf{I}_{A_n} \, d\mu = \sum_1^\infty M_\xi(A_n).$$

b) Пусть $\mu(B) = 0$, тогда $\xi \, \mathbf{I}_B = 0 \, [\mu\text{-п.в.}]$ и по свойству интеграла Лебега $\int \xi\, \mathbf{I}_B\, d\mu = 0$.

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

$$M_\xi(\widehat{\Omega}_n \cup \{\xi = \infty\}) = M_\xi(\widehat{\Omega}_n) = \int\limits_{\widehat{\Omega}_n} \xi\, d\mu \leqslant n\mu(\widehat{\Omega}_n) \leqslant n\mu(\Omega_n) < \infty$$

и, очевидно, $\widehat{\Omega}_n \cup \{\xi = \infty\} \nearrow \Omega$, $n \to \infty$.

Обратно, если найдётся возрастающая последовательность подмножеств $\Omega_n \nearrow \Omega$, $n \to \infty$, с $M_\xi(\Omega_n) < \infty$, то и $M_\xi(\widetilde{\Omega}_n) < < \infty$, для $\widetilde{\Omega}_n = \Omega_n \cap \{\xi = \infty\}$, что в силу определения интеграла от расширенной функции возможно, только если $\mu(\widetilde{\Omega}_n) = 0$. Поэтому $\mu(\xi = \infty) = \lim\limits_{n}{}_{\uparrow}\, \mu(\widetilde{\Omega}_n) = 0$.

Для индикаторных функций $\eta = \mathbf{I}_B$ равенство d) выполняется по определению меры M_ξ; для произвольных измеримых функций — по конструкции интеграла Лебега. ⫤

173| $\mathfrak{Y}$пр. Пусть $\int |\xi|\, d\mu < \infty$. Докажите, что если для последовательности множеств $\langle A_n, n \geqslant 1 \rangle$ 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

174| Теорема. [*И. Радон, О. Никодим.*] *Если σ-конечная мера γ абсолютно–непрерывна относительно σ-конечной меры μ, то найдётся измеримая функция ξ $(\geqslant 0)$ 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 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.

175 Лемма. [О разложении мер. Лебег.] *Пусть γ, μ — σ-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*

$$\gamma(A) = \psi(A) + \int_A \xi\, d\mu, \qquad A \in \mathscr{F};$$

b) *существует единственное разложение меры γ в виде суммы двух мер, одна из которых μ-непрерывная, а другая μ-сингулярная.*

⊫ Стр. 221–224. ⫤

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 175 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. ⫤

⚠ Можно доказать вариант теоремы Радона–Никодима для произвольной меры γ $(\ll \mu)$ и σ-конечной меры μ. Мы приведём здесь доказательство, заимствованное из монографии Ж. Невё [1], с произвольной мерой γ и конечной мерой μ — вариант, востребованный при определении условного математического ожидания в теории вероятностей.

⊫ Определим класс множеств конечной γ-меры: $\mathfrak{S} = \langle S : \gamma(S) < \infty \rangle$. Выберем последовательность $S_n \in \mathfrak{S}$: $\lim_n \mu(S_n) = \sup\langle \mu(S), S \in \mathfrak{S} \rangle$. Т.к. 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 $\gamma(A) = \int_A \xi \, d\mu$, $A \in \mathscr{F}$. ⫤

176| С л е д с т в и е. *Если мера $\gamma \ll \mu$, то для любой измеримой функции h*

$$\int_\Omega h \, d\gamma = \int_\Omega h \, \frac{d\gamma}{d\mu} \, d\mu. \tag{10}$$

⊫ Для индикаторных и простых функций это равенство справедливо в силу теоремы Радона–Никодима. Отсюда по теореме о монотонной сходимости делаем вывод о справедливости следствия для измеримых неотрицательных функций и далее стандартным приёмом — для произвольных измеримых функций (если существует правая или левая часть равенства). ⫤

177| $\mathfrak{Ynp}$. Докажите, что если меры на 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.

178 Теорема. *Для любых мер* $\gamma, \mu \; : \; (\Omega, \mathscr{F}) \mapsto \overline{\mathbb{R}}^1_+$, $\gamma(\Omega) < \infty$, *следующие условия эквивалентны:*

а) $\gamma \ll \mu$, *т.е.* $\mu(N) = 0 \;\Rightarrow\; \gamma(N) = 0$;

б) *для* $\forall\, \epsilon > 0$ *найдётся такое* $\delta > 0$, *что*

$$A \in \mathscr{F},\; \mu(A) < \delta \quad\Rightarrow\quad \gamma(A) < \epsilon;$$

в) *если* σ*-алгебра* $\mathscr{F} = \sigma(\mathcal{K})$ *порождена полукольцом* $\mathcal{K}$, *то для* $\forall\, \epsilon > 0\; \exists\, \delta > 0$ *такое, что для любого счётного семейства элементов полукольца*

$$\forall \left\langle\, A_j \,\right\rangle_1^J \subset \mathcal{K},\; A_i A_j = \varnothing,\;\; i \neq j,\;\; J = \infty,\; \sum_{j=1}^{J} \mu(A_j) < \delta \;\;\Rightarrow$$

$$\Rightarrow \sum_{j=1}^{J} \gamma(A_j) < \epsilon.$$

Условие $J = \infty$ *можно заменить на условие* $\forall\, J < \infty$.

$\Rrightarrow$ а) $\Rightarrow$ б) Предположим от противного, что для некоторого $\epsilon > 0$ и любого $n \geqslant 1$ найдётся $A_n \in \mathscr{F}$, что $\mu(A_n) < {}^1\!/_{2^n}$ и $\gamma(A_n) > \epsilon$. Положим $A = \lim\limits_{n}{}_{\downarrow} \bigcup_{i=n}^{\infty} A_i$. Так как конечные меры непрерывны относительно убывающих последовательностей, то

$$\gamma(A) = \lim_n \gamma\Big(\bigcup\nolimits_{i=n}^{\infty} A_i\Big) \geqslant \varlimsup_n \gamma(A_n) > \epsilon.$$

С другой стороны, мера $\mu(A) \leqslant \sum_{i=n}^{\infty} \mu(A_i) < {}^2\!/_{2^n}$ при любом $n \geqslant 1$, т.е. $\mu(A) = 0$, что противоречит $\gamma \ll \mu$.

б)$\Rightarrow$а) Условие б) означает, что если $\mu(N) = 0$, то $\gamma(N) < \epsilon$ для $\forall\,\epsilon > 0$, т.е. $\gamma(N) = 0$.

б)$\Rightarrow$в) Следует из σ-аддитивности мер.

в)$\Rightarrow$а) Пусть $\mu(N) = 0$. Для $\epsilon > 0$ выберем $\delta > 0$ из условия в). В силу 109, стр. 132, 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 $\mu(S) = \sum_n \mu(S_n) < \delta$.

Т.к. подмножества из кольца представимы в виде конечных объединений непересекающихся элементов из полукольца, то можно считать все $S_n \in \mathcal{K}$. Из условия в) следует, что $\gamma(N) \leqslant \gamma(S) = \sum_n \gamma(S_n) < \epsilon$, т.е. $\gamma(N) = 0$. ⫤

Рассмотрим теперь некоторую меру Лебега–Стилтьеса на борелевской прямой $\mathbb{R}^1$, порождённую конечной непрерывной функцией F. Полукольцо $\mathcal{K}$, фигурирующее в теореме, состоит из интервалов вида $(a;b]$, для которых мера Лебега–Стилтьеса 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 функции 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:

$$F(b) - F(a) = \int_{(a;b]} F'(x)\,dx, \tag{11}$$

где, для определённости, полагается $F'(x) = 0$ в тех точках, в которых производная не существует.

179 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*

$$\mu_F(A) = \int\limits_A F'(x)\,dx, \qquad \forall\, A \in \mathscr{B}.$$

⊫ Для интервалов $A = (a;b\,]$ это равенство следует из (11). Справедливость его для любых борелевских множеств обеспечивает теорема Каратеодори. ⫤

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. функции F :

$$\bigl|F(x) - F(y)\bigr| \leqslant L\bigl|x-y\bigr|, \qquad \forall\, x, y \in Q \subset \mathbb{R}^1,$$

абсолютно-непрерывны на Q;

✦ непрерывно-дифференцируемые функции — Липшицевы (значит, абсолютно-непрерывные) на конечных интервалах;

✦ существуют равномерно-непрерывные на отрезке, ограниченные функции (например, канторова функция: 184, стр. 197), не являющиеся абсолютно-непрерывными на этом отрезке.

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

$$F'^{+}(x) = \varliminf_{h\to+0} \frac{1}{h}(F(x+h) - F(x)),$$

т.е. (в силу теоремы Лебега) для почти всех x эта «правая производная» $F'^{+}(x)$ совпадает с обычной производной функции F. Поэтому 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.

180 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, если*

$$\int_{\mathbb{R}^1} F'(x)\, dx \geqslant \mu_F(\mathbb{R}^1) = F(+\infty) - F(-\infty).$$

В обоих случаях плотность $d\mu_F/d\lambda = F'$ и для $\forall\, A \in \mathscr{B}$ и любой борелевской функции ξ

$$(\checkmark)\quad \mu_F(A) = \int_A F'(x)\, dx,$$

$$(\checkmark)\quad \int_{\mathbb{R}^1} \xi(x)\, dF(x) = \int_{\mathbb{R}^1} \xi(x) F'(x)\, dx.$$

⚠ *Существуют меры Лебега–Стилтьеса с $\mu_F(\mathbb{R}^1) = \infty$, для которых утверждение теоремы неверно.*

⊫ 1) Повторяет утверждение 179. Доказательство пункта 2) приведено в главе I, теорема 26, стр. 38, для многомерной функции распределения. ⫤

[‡] Как всегда, $F(\pm\infty) = \lim\limits_{x\to\pm\infty} F(x)$.

О разложении Лебега для мер Лебега–Стилтьеса Сингулярную часть в разложении Лебега для меры Лебега–Стилтьеса на $\mathbb{R}^1$ можно представить в виде суммы двух мер, одна из которых дискретная, а другая есть мера Лебега–Стилтьеса с непрерывной всюду генерирующей функцией.

Определение. Функция $F(x) : \mathbb{R}^1 \mapsto \mathbb{R}^1$ 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

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

181| Лемма. (?!) *Функция скачков не убывает, непрерывна при $x \notin \mathscr{X}$, непрерывна справа при всех $x = x_k \in \mathscr{X}$, причём скачок в каждой из этих точек равен $F(x_k) - F(x_k - 0) = p_k$.*

182| Лемма. (?!) *Монотонная на всей числовой прямой функция имеет не более чем счётное число точек разрыва.*

183| Теорема. [Разложение Жордана.] *Пусть F — непрерывная справа, неубывающая функция на $\mathbb{R}^1$. Тогда существует единственное представление $F = F_c + F_d$, где F_c — всюду непрерывная, а F_d — дискретная функции.*

⊫ Определим интервал

$$\Theta(x) = \begin{cases} (x; 0], & \text{если } \ x < 0, \\ (0; x], & \text{если } \ x \geqslant 0. \end{cases} \tag{12}$$

Пусть $\mathscr{X} = \{x_k,\ k = 1, 2, \ldots\}$ — не более чем счётное множество точек разрыва функции F, $p_k = F(x_k) - F(x_k - 0)$ — величина скачка (> 0) в каждой из точек $x_k \in \mathscr{X}$. Определим дискретную функцию

$$F_d(y) = \operatorname{sign}(y) \sum_{x_k \in \Theta(y)} p_k + F(0).$$

В силу 181 эта функция непрерывна справа. 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 181 $F_c(x+0) - F_c(x-0) = 0$, 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 скачков функции F на фиксированном конечном интервале $(a;\, b]$. 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

$$\begin{aligned} F(b) - F(a) = \big[F(b) - F(y_{nn})\big] + \sum_{k=2}^{n} \big[F(y_{nk}) - F(y_{n(k-1)})\big] + \\ + \big[F(y_{n1}) - F(a)\big] \geqslant \\ \geqslant \big[F(b) - F(y_{nn})\big] + \sum_{k=2}^{n} \big[F(y_{nk}) - F(y_{nk} - 0)\big] + \\ + \big[F(y_{n1}) - F(y_{n1} - 0)\big] \geqslant \end{aligned}$$

$$\geqslant \sum_{k=1}^{n} \big[F(y_{nk}) - F(y_{nk}-0)\big] = \sum_{x_k \in A_n} p_k \geqslant 0,$$

что доказывает монотонность F_c.

В силу равенства $F = F_c + F_d$ и непрерывности F_c все точки разрыва дискретной функции F_d совпадают с точками разрыва функции F с равными скачками. Поэтому в таком представлении дискретная функция определяется однозначно, следовательно, однозначно определяется и непрерывная функция F_c. ⫷

Определения. Точка x называется *точкой роста* неубывающей функции F, если $F(x-\epsilon) < F(x+\epsilon)$ для $\forall\, \epsilon > 0$.

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

184 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 присутствуют только нули и двойки. Для чисел x с $J = N_x < \infty$ определим функцию

$$\mathscr{K}(x) := \{0, \tfrac{x_1}{2}\tfrac{x_2}{2} \ldots \tfrac{x_{J-1}}{2}\, 1\}_2 \overset{\mathrm{df}}{=} \sum_{k=1}^{J-1} \frac{x_k}{2^{k+1}} + \frac{1}{2^J},$$

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

$$\mathscr{K}\big(\tfrac{3}{4}\big) = \frac{1}{2} + \frac{1}{8} + \ldots = \frac{2}{3}.$$

Функция $\mathscr{K}(x)$, $x \in [0;1]$, называется функцией Кантора (канторовой лестницей). Легко понять, что она непрерывна и не убывает. Далее будет показана её сингулярность. На рис. 1 приведён вариант графика канторовой лестницы.

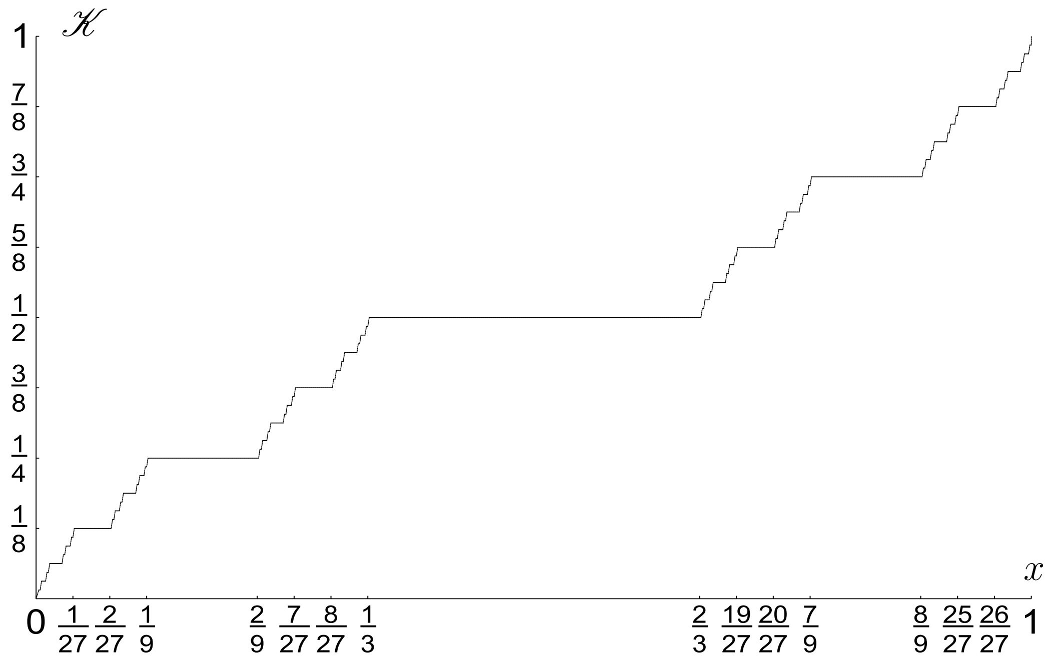


Рис. 1. График функции Кантора.

Пусть $x \in (^1/_3; ^2/_3)$, тогда $x = \{0,1x_2x_3\ldots\}_3$ и значение функции $\mathscr{K}(x) = \{0,1\}_2 = ^1/_2$ для всех таких x. Точки $x = ^1/_3$ и $x = ^2/_3$ могут быть представлены двумя способами, например, $^1/_3 = \{0,1\}_3 = \{0,022\ldots\}_3$. В обоих случаях функция $\mathscr{K}(x) = \{0,1\}_2 = \{0,011\ldots\}_2 = ^1/_2$. Таким образом, на отрезке $x \in [^1/_3; ^2/_3]$ функция $\mathscr{K}(x) = ^1/_2$, т.е. интервал $(^1/_3; ^2/_3)$ длины $^1/_3$ является интервалом постоянства функции $\mathscr{K}$.

Аналогично, $\mathscr{K}(x) = ^1/_4$ при $x \in (^1/_9; ^2/_9)$ и $\mathscr{K}(x) = ^3/_4$ при $x \in (^7/_9; ^8/_9)$. Суммарная длина последних двух интервалов равна $^2/_9$. На краях этих интервалов значения функции $\mathscr{K}$ совпадают со значениями во внутренних точках.

Продолжая подобным образом, получим семейство всё более измельчающихся интервалов постоянства функции $\mathscr{K}$, суммарная длина которых по формуле 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 $\mathscr{K}\left(\frac{1}{3} - \frac{1}{3^3}\right) = 0.375 < 0.5 = \mathscr{K}\left(\frac{1}{3} + \frac{1}{3^3}\right).$

Канторово множество — пример несчётного множества лебеговой меры нуль. ◎

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

а) *дополнение $Q_F^c = \cup_j (a_j; b_j)$ есть не более чем счётное объединение непересекающихся открытых интервалов, на каждом из которых функция F постоянна: $F(x) = F(y)$, $\forall\, x, y \in \in (a_j; b_j)$;*

б) *множество Q_F имеет полную меру относительно меры Лебега–Стилтьеса, порождённой F: $\mu_F(Q_F^c) = 0$.*

⊫ а) Из определения следует, что для $\forall\, x \notin Q_F$ найдётся $\epsilon > 0$ такое, что $F(x - \epsilon) = F(x) = F(x + \epsilon)$. Таким образом, с каждым $x \notin Q_F$ можно связать открытый интервал $(l_x; r_x)$, где

$$\begin{aligned} l_x &= \inf\{y : F(y) = F(x),\ y < x\}, \\ r_x &= \sup\{y : F(y) = F(x),\ y > x\}. \end{aligned}$$

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 достаточно заметить теперь, что любой набор непересекающихся открытых интервалов числовой прямой не может содержать более чем счётное число элементов.

б) По только что доказанному $Q_F^c = \cup_1^N (a_j; b_j)$, $N \leqslant \infty$, т.е. это множество измеримо по Борелю. Кроме того, для каждого из интервалов мера

$$\mu_F(a_j; b_j) = \lim_{\epsilon\to 0}{\uparrow}\, \mu_F(a_j + \epsilon; b_j - \epsilon) =$$
$$= \lim_{\epsilon\to 0}(F(b_j + \epsilon) - F(a_j - \epsilon)) = 0. \qquad \Lleftarrow$$

186| Следствие. *Пусть ψ — мера на $(\mathbb{R}^1, \mathscr{B})$. Мера Лебега–Стилтьеса μ_F, порождённая функцией F, имеющей ψ-нулевое множество точек роста Q_F, сингулярна относительно меры ψ.*

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

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

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

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

Заметим, что дискретные функции не являются сингулярными, хотя меры Лебега–Стилтьеса, ими порождённые, сингулярны относительно меры Лебега. Связь между понятиями сингулярности меры и функции поясняет следующая

188| Теорема. [*Разложение Лебега для функций.*] *Любую неубывающая непрерывную справа функция $F(x)$, $x \in \mathbb{R}^1$, можно единственным образом представить в виде*

$$F(x) = F_a(x) + F_d(x) + F_{cs}(x),$$

где F_a, F_d — соответственно абсолютно-непрерывная и дискретная функции, F_{cs} — непрерывная функция, порождающая сингулярную меру.

⊫ Воспользовавшись разложением Лебега для меры Лебега–Стилтьеса, представим меру $\mu_F = \Psi + \mu_a$. По лемме 179 функция $F_a(x) = \mu_a(\Theta(x))$, генерирующая абсолютно-непрерывную меру μ_a, с интервалом $\Theta(x)$, определённым в (12), есть абсолютно-непрерывная функция. Функцию $\widetilde{F}(x) = \Psi(\Theta(x))$, генерирующую сингулярную меру Ψ, в соответствии с разложением Жордана можно представить в виде $\widetilde{F} = F_{cs} + F_d$, с непрерывной функцией F_{cs}. ⫤

189| Пример. Существуют сингулярные меры, у которых генерирующие функции непрерывны и всюду возрастают. Рассмотрим функцию Кантора $\mathscr{K}(x)$, генерирующую меру Лебега–Стилтьеса $\mu_{\mathscr{K}}$ 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

$$G(x) \overset{(a)}{=} \int_0^1 F(x-y)\,\mu_{\mathscr{K}}(dy) \overset{(b)}{=} \int_0^1 \mathscr{K}(x-y)\,\mu_F(dy), \quad x \in [0;2].$$

Из представления (a) для свёртки следует строгая монотонность G, а из представления (b) — её непрерывность. Как известно, мера μ_G, генерируемая этой функцией, сосредоточена на множестве точек вида $\{x + r :\ x \in \mathcal{M}, r \in Q\} = \cup_{r \in Q}(\mathcal{M} + r)$. Поскольку мера Лебега λ инвариантна относительно сдвигов, то $\lambda(\mathcal{M} + r) = \lambda(\mathcal{M}) = 0$, и, следовательно, мера Лебега всего множества сосредоточения μ_G равна нулю. ◎

§ 8. Связь с интегралом Римана-Стилтьеса

Разложение Лебега–Жордана и теорема Радона–Никодима дают прямой путь вычисления интегралов Лебега относительно мер Лебега–Стилтьеса. Для вычисления интеграла $\int_\Omega \xi(x) F'(x)\, dx$ относительно абсолютно-непрерывной части этого разложения можно воспользоваться 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 1-ого рода в конечном или счётном числе точек. Поэтому при формулировке обратных 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.

190 Лемма. *Пусть $\xi : (\mathbb{R}^k, \mathscr{B}^k) \mapsto (\mathbb{R}, \mathscr{B})$ — 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) и 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 не интегрируема 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 «матем.ожидание» $\int_{\mathbb{R}} x f(x)\, dx$ условно конечно, 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. Точную формулировку утверждения оставляем читателю.

191| Пример. Пусть мера Лебега–Стилтьеса определяется функцией

$$F(x) = e^x\, \mathbf{I}(x < 0) + (4 - e^{-x})\, \mathbf{I}(x \geqslant 0)$$

с разрывом в $x = 0$. Эту функцию запишем в виде суммы двух функций, одна из которых (F_a) абсолютно-непрерывна, а другая (F_d) кусочно-постоянна с единственным скачком в $x = 0$:

$$F(x) = F_a(x) + F_d(x) :=$$
$$= \Big[e^x\, \mathbf{I}(x < 0) + (2 - e^{-x})\, \mathbf{I}(x \geqslant 0)\Big] + \Big[2\, \mathbf{I}(x \geqslant 0)\Big].$$

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, интеграл от функции ξ равен (см. 154, стр. 175)

$$\int_{\mathbb{R}} \xi(x)\, dF(x) = \int_{-\infty}^{\infty} \xi(x) e^{-|x|}\, dx + 2\xi(0),$$

если существует $\int_{-\infty}^{\infty} |\xi(x)|\, e^{-|x|}\, dx < \infty$. ⊙

192| ⚠ 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 даже 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 $\int_{\mathbb{R}} \mathbf{I}_{[a;b]}\xi(x)\,d\mu_F(x)$, понимаемой как интеграл Лебега подынтегральной функции 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 такого обозначения проявляется, когда мера Лебега–Стилтьеса, суженая с $\mathbb{R}^1$ до области интегрирования A не совпадает с мерой, построенной на функции F, суженой до A. Например, непрерывная справа функция $F(x) = \mathbf{I}(x;[0;\infty))$ порождает на $\Omega = \mathbb{R}^1$ меру Дирáка 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 $F \equiv 1$ и 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 $\int_{\mathbb{R}} \mathbf{I}_{[0;1]}\xi\, dF$, 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.

193 Пример. Пусть $F(x) = \llbracket x \rrbracket$ — функция, генерирующая меру, «считающую» целые точки (см. стр. 150), $\xi(x) = 1 + x$. 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 обуславливает положительность меры Лебега–Стилтьеса в этой точке. Ту же самую считающую меру порождает непрерывная слева функция $\widetilde{F}(x) = \lceil\!\lceil x \rceil\!\rceil$, округляющая числа вверх. Ситуация изменится, если рассматриваются интегралы по мере Лебега–Стилтьеса на соответствующем отрезке (обозначим сейчас такие интегралы со звёздочкой $\int^*$). Например,

$$\int^*_{[0;1]} \xi\, dF = 2, \quad \int^*_{[0;1]} \xi\, d\widetilde{F} = 1 \quad (?!). \qquad \circledcirc$$

Монотонная замена переменных в интеграле. Докажем классическую формулу замены переменных для интеграла Лебега на $\mathbb{R}$. Обозначим через $\lfloor a;b\rceil$ любой возможный интервал на числовой прямой — $[a;b]$, или $(a;b]$, или $(-\infty;b)$ и т.п.

194| Теорема. *Если функция* $\xi : \lfloor a;b\rceil \to \lfloor u;v\rceil$ *строго возрастает и обратная функция* $\tilde{\xi}\,(x) = \inf\{t \in \lfloor a;b\rceil : \xi(t) \geqslant \geqslant x\}$, $x \in \lfloor u;v\rceil$, *абсолютно-непрерывна, то для любой борелевской функции* $\eta : \lfloor u;v\rceil \to \mathbb{R}^1$

$$\int\limits_{\lfloor a;b\rceil} \eta(\xi(t))\, dt = \int\limits_{\lfloor u;v\rceil} \eta(x)\, |\tilde{\xi}\,'(x)|\, dx.^{(\ddagger)} \tag{13}$$

⊫ По теореме 163 , стр. 180, 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 $= \lfloor a;\tilde{\xi}\,(x))$ в зависимости от того, $\xi(\tilde{\xi}\,(x)) \leqslant x$ или $\xi(\tilde{\xi}\,(x)) > x$. В любом случае индуцированная мера интервала $\lambda_\xi(x;y] = \tilde{\xi}\,(y) - \tilde{\xi}\,(x)$. Итак, λ_ξ есть мера Лебега–Стилтьеса на $\lfloor u;v\rceil$, порождённая абсолютно-непрерывной функцией $\tilde{\xi}$, что вместе с утверждением 179 и равенством (10), стр. 190, доказывает теорему. ⫤

Доказательство многомерного варианта этой теоремы требует несколько больших усилий, поэтому мы приведём только формулировку результата (см., например, [5, стр. 78], [4, стр. 162], [9, 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 $\tilde{\xi}(x) = \inf\{t : \xi(t) \leqslant x\}$).

195| Теорема. *Пусть функция* $\vec{\xi} = (\xi_1, \dots, \xi_k) : \Omega \mapsto E$ *взаимно-однозначно отображает борелевскую область* $\Omega \subseteq \mathbb{R}^k$ *в борелевскую область* $E \subseteq \mathbb{R}^k$, $\vec{\xi}^{-1} = (\tilde{\xi}_1, \dots, \tilde{\xi}_k)$ *— обратная функция. Если компоненты функции* $\vec{\xi}^{-1}$ *непрерывно дифференцируемы и всюду на* E *якобиан*

$$J(x_1, \dots, x_k) = \left\| \left(\frac{\partial \tilde{\xi}_j(x_1, \dots, x_k)}{\partial x_n} \right)_{nj} \right\| > 0,$$

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.

§ 9. 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 всем пространстве представима в виде произведения мер на «координатных» пространствах. Соответствующее утверждение называется теоремой Фубини.

Определение. Пространство $(\Omega_1 \times \Omega_2,\ \mathscr{F}_1 \otimes \mathscr{F}_2,\ \mu_1 \times \mu_2)$, где $(\Omega_i, \mathscr{F}_i, \mu_i)$, $i = 1, 2,$ — пространства с σ-конечными мерами, называется *прямым произведением пространств*, если σ-алгебра $\mathscr{F}_1 \otimes \mathscr{F}_2 \ = \ \sigma(\mathscr{F}_1 \times \mathscr{F}_2)$ — минимальная σ-алгебра, содержащая все подмножества $\Omega_1 \times \Omega_2$ вида $B_1 \times B_2$, $B_i \in \in \mathscr{F}_i, i = 1, 2,$ и мера (также σ-конечная) на подмножествах такого вида равна $\mu_1 \times \mu_2(B_1 \times B_2) = \mu_1(B_1)\mu_2(B_2)$.

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.

196 Лемма. 1) *Для любого множества* $B \in \mathscr{F}_1 \otimes \mathscr{F}_2$ *и* $\forall\, \omega_1 \in \Omega_1$ *сечение* B *вдоль* ω_1 *измеримо относительно* $\mathscr{F}_2$:

$$B_{\omega_1} := \{\omega_2 :\ (\omega_1, \omega_2) \in B\} \quad \in \mathscr{F}_2.$$

2) *Пусть* $\xi : (\Omega_1 \times \Omega_2, \mathscr{F}_1 \otimes \mathscr{F}_2) \mapsto (\mathfrak{O}, \mathscr{A})$ *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.*

$\Rrightarrow$ 1) Покажем, что совокупность всех множеств, удовлетворяющих утверждению теоремы, образует σ-алгебру. Пусть

$$\mathcal{U} = \big\langle\, B :\ B \in \mathscr{F}_1 \otimes \mathscr{F}_2,\ \ B_{\omega_1} \in \mathscr{F}_2 \,\big\rangle.$$

✦ 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

$$(B^c)_{\omega_1} = \{\omega_2 :\ (\omega_1, \omega_2) \notin B\} = \Omega_2 \setminus \{\omega_2 :\ (\omega_1, \omega_2) \in B\} \in \mathscr{F}_2.$$

✦ 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

$$\{\omega_1 : \xi(\omega_1, \omega_2) \in A\} = \{(\omega_1, \omega_2) : \xi(\omega_1, \omega_2) \in A\}_{\omega_2} \quad \in \mathscr{F}_1,$$

ввиду измеримости ξ и по утверждению 1) леммы. $\Lleftarrow$

⚠ 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. Для мер Лебега–Стилтьеса этот факт вытекает из того, что функция $F_1(\vec{x})F_2(\vec{y})$ удовлетворяет всем свойствам, необходимым для построения меры Лебега–Стилтьеса в пространстве $\mathbb{R}^{k+m}$.

197| 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. 224 – 226. ⫤

198| 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) $\left| \int_{\Omega_2} \xi(\omega_1, \omega_2)\, \mu_2(d\omega_2) \right| < \infty$ *для* μ_1*-почти всех* ω_1 *(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*

$$\int\limits_{\Omega_1\times\Omega_2} \xi\, d\mu_1{\times}\mu_2 = \int\limits_{\Omega_1}\biggl(\int\limits_{\Omega_2} \xi(\omega_1,\omega_2)\,\mu_2(d\omega_2)\biggr)\,\mu_1(d\omega_1) = \\ = \int\limits_{\Omega_2}\biggl(\int\limits_{\Omega_1} \xi(\omega_1,\omega_2)\,\mu_1(d\omega_1)\biggr)\,\mu_2(d\omega_2). \tag{14}$$

Часто легче проанализировать конечность повторного интеграла, отсюда

199 Теорема. [*Л. Тонелли.*] *Если*

$$\int\limits_{\Omega_1}\biggl(\int\limits_{\Omega_2} \bigl|\xi(\omega_1,\omega_2)\bigr|\,\mu_2(d\omega_2)\biggr)\,\mu_1(d\omega_1) < \infty,$$

то $\Bigl|\int\limits_{\Omega_1\times\Omega_2} \xi\, d\mu_1\times\mu_2\Bigr| < \infty$ *и справедливы равенства (14).*

Для неотрицательных функций $\xi \geqslant 0$ не обязательно требовать конечности интегралов.

200 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* (14) *в том смысле, что если конечна (бесконечна) какая-либо из частей этих соотношений, то конечны (бесконечны) и две другие его части и они совпадают.*

⊫ Докажем все три теоремы одновременно. Заметим, что по предыдущему 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 $\omega_2 \in \Omega_2$.

Следуя схеме построения интеграла Лебега, рассмотрим $\xi = \mathbf{I}_B$, $B \in \mathscr{F}_1\otimes\mathscr{F}_2$, тогда при фиксированном ω_1 функция $\xi_{\omega_1}(\omega_2) \equiv \mathbf{I}_{B_{\omega_1}}(\omega_2)$, $\omega_2 \in \Omega_2$, и интеграл $\int_{\Omega_2} \xi(\omega_1,\omega_2)\,\mu_2(d\omega_2) = \mu_2(B_{\omega_1})$. Равенство (14) верно, поскольку совпадает с определением меры $\mu_1 \times \mu_2$ (см. доказательство предыдущей леммы, где была также показана измеримость $\mu_2(B_{\omega_1})$). Отсюда сразу следует (14) 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 $(h_{n-1})_{\omega_1} \leqslant \leqslant (h_n)_{\omega_1} \underset{n}{\to} \xi_{\omega_1}$, 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 (14) $(\mu = \mu_1 \times \mu_2)$:

$$\int_{\Omega_1 \times \Omega_2} \xi \, d\mu = \lim_n{}_\uparrow \int_{\Omega_1} \left(\int_{\Omega_2} (h_n)_{\omega_1} \, d\mu_2 \right) d\mu_1 = \\ = \int_{\Omega_1} \left(\int_{\Omega_2} \xi_{\omega_1}(\omega_2) \, d\mu_2 \right) d\mu_1.$$

Важно отметить, что равенство здесь симметрично. Другими словами, если интеграл справа конечен, то и интеграл слева также конечен, и наоборот. Т.е. можно считать доказанным вариант теоремы Фубини в редакции Тонелли. Кроме того, из этого равенства следует справедливость теоремы Фубини для положительных функций (даже если интегралы равны $+\infty$).

Произвольная функция есть разность двух неотрицательных $\xi = \xi^+ - \xi^-$, интегралы от которых по условию теоремы Фубини конечны, следовательно, к ним можно применить доказанные выше утверждения. По свойству линейности интеграла отсюда следует справедливость теоремы Фубини. ⫥

§ 10. Формула интегрирования по частям

Сначала докажем простейший вариант формулы интегрирования по частям. При выводе этой формулы используется простой алгебраический факт:

$$(\star) \qquad \sum_{k=0}^{M-1}(b_{k+1}-b_k)=\sum_{k=0}^{M-1}b_{k+1}-\sum_{k=0}^{M-1}b_k=b_M-b_0.$$

Отсюда легко получается так называемая формула суммирования по частям (преобразование Абеля):

$$\sum_{k=0}^{M-1}U_k(V_{k+1}-V_k)=\sum_{k=0}^{M-1}U_kV_{k+1}-\sum_{k=0}^{M-1}U_kV_k=$$
$$=\sum_{k=0}^{M-1}U_{k+1}V_{k+1}-\sum_{k=0}^{M-1}U_kV_k-\sum_{k=0}^{M-1}(U_{k+1}-U_k)V_{k+1}=$$
$$=-\sum_{k=0}^{M-1}(U_{k+1}-U_k)V_{k+1}+U_MV_M-U_0V_0.$$

Снова применяя $(\star)$, отсюда получаем равенство

$$\binom{*}{*} \qquad \sum_{k=0}^{M-1}(U_k-U_0)(V_{k+1}-V_k)=\sum_{k=0}^{M-1}(U_{k+1}-U_k)(V_M-V_{k+1}).$$

201| Т е о р е м а. *Пусть* G — *неубывающая и непрерывная, а* F — *неубывающая и непрерывная справа функции со значениями в интервале* $(A;B)$, *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 $n \to \infty$ последовательность простых неотрицательных функций

$$G_n(x) - G(0) = \sum_{k=0}^{2^n-1} \Big(G\Big(\frac{k}{2^n}\Big) - G(0)\Big)\, \mathbf{I}_{A_{nk}}(x) \nearrow G(x) - G(0),$$

$$F(1) - F_n(x) = \sum_{k=0}^{2^n-1} \Big(F(1) - F\Big(\frac{k+1}{2^n}\Big)\Big)\, \mathbf{I}_{A_{nk}}(x) \nearrow F(1) - F(x).$$

Действительно, если $x \in A_{nk}$ при некотором k, то $x \leqslant {}^{(k+1)}/{}_{2^n} \leqslant x + {}^{1}/{}_{2^n}$ 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 $\binom{*}{*}$:

$$\int\limits_{(0;1\,]} \big(G_n(x) - G(0)\big)\, dF(x) = \sum_{k=0}^{2^n-1} \Big(G\Big(\frac{k}{2^n}\Big) - G(0)\Big)\Big(F\Big(\frac{k+1}{2^n}\Big) - F\Big(\frac{k}{2^n}\Big)\Big) = \sum_{k=0}^{2^n-1} \Big(F(1) - F\Big(\frac{k+1}{2^n}\Big)\Big)\Big(G\Big(\frac{k+1}{2^n}\Big) - G\Big(\frac{k}{2^n}\Big)\Big) = \int\limits_{(0;1\,]} \big(F(1) - F_n(x)\big)\, dG(x).$$

В соответствии с определением интеграла Лебега отсюда следует равенство интегралов от функций $G(x)-G(0)$ и $F(1)-F(x)$. Эквивалентность этого утверждению теоремы очевидна. ⇚]

⚠ Функции F и G входят в формулу интегрирования по частям симметричным образом. Поэтому в условиях теоремы можно потребовать непрерывность любой из них, причём отказаться от этого условия, вообще говоря, нельзя. Рассмотрим в качестве примера две одинаковые функции $F(x) = G(x) = -\mathbf{I}(x; < 0) + \mathbf{I}(x; \geqslant 0)$. Для этих функций

$$\int\limits_{(-1;1]} G(x)\, dF(x) = \int\limits_{(-1;1]} F(x)\, dG(x) = 2,$$

в то же время $F(1)G(1) - F(-1)G(-1) = 0$. Здесь существенным оказалось то, что функции терпят разрыв в общей точке.

202 $\mathfrak{Упр}$. 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 от функции с ограниченной вариацией. Самый естественный способ обобщения формулы интегрирования по частям состоит в рассмотрении наряду с неубывающими функциями их линейных комбинаций. Действительное расширение произойдет, 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 $[a;b]: \ V_a^b[F] < \infty$, с

$$V_a^b[F] := \sup\Big\langle \sum_{k=1}^{n} \big|F(x_k) - F(x_{k-1})\big| \Big\rangle,$$

где супремум берётся по всем $a = x_0 < x_1 < \ldots < x_n = b$. Если F имеет ограниченную вариацию на $[a;b]$, то $F = F_1 - F_2$, 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, можно определить интеграл Лебега–Стилтьеса от 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 $+\infty - \infty$.

203| $\mathfrak{Y}$пр. Докажите, что формула интегрирования по частям справедлива 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.

204| 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*

$$\int\limits_{(a;b]} G(x)\,dF(x) + \int\limits_{(a;b]} F(x)G'(x)\,dx = G(b)F(b) - G(a)F(a).$$

Вспомогательный материал

A. Доказательства

✧ Доказательство теоремы Каратеодори (стр. 128).

⊫ Для доказательства сигма-полуаддитивности внешней меры возьмём $G \subset \bigcup\limits_{n=1}^{\infty} S_n$. 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

$$\forall\, \epsilon > 0\ \forall\, n \geqslant 1\ \exists \bigcup\nolimits_{j=1}^{\infty} B_{nj} \supset S_n :\ \mu^*(S_n) > \sum\nolimits_{j=1}^{\infty} \mu(B_{nj}) - \frac{\epsilon}{2^n},$$

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. представляет собой счётное $\mathcal{A}$-покрытие G, поэтому по определению внешней меры

$$\mu^*(G) \leqslant \sum\nolimits_n \sum\nolimits_j \mu(B_{nj}) < \sum\nolimits_n \mu^*(S_n) + \epsilon.$$

Полагая $\epsilon \to 0$, 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 $\forall\, B \subset \Omega$

$$\begin{aligned}\mu^*(B) &= \mu^*(B\,S) + \mu^*(B\,S^c) = \quad (\text{т.к. } G \in \Psi)\\ &= \mu^*(B\,S) + \mu^*((B\,S^c)\,G) + \mu^*((B\,S^c)\,G^c).\end{aligned}$$

Так как $S = S\,(S \cup G)$, $\;G\,S^c = (S \cup G)S^c$ 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 $\forall\, B \subset \Omega$

$$\mu^*\Big(B \biguplus\nolimits_1^n S_j\Big) \;=\; \sum\nolimits_1^n \mu^*(B\,S_j).$$

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

$$\begin{aligned}\mu^*\Big(B \biguplus_{i=1}^{n+1} S_i\Big) &= \mu^*\Big(\Big(B \biguplus_{i=1}^{n+1} S_i\Big) S_{n+1}\Big) + \mu^*\Big(\Big(B \biguplus_{i=1}^{n+1} S_i\Big) S_{n+1}^c\Big) =\\ &= \mu^*(B\,S_{n+1}) + \mu^*\Big(B \biguplus_{i=1}^{n} S_i\Big) \;=\; \sum_{i=1}^{n+1} \mu^*(B\,S_i).\end{aligned}$$

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. 102, стр. 127), поэтому

$$\begin{aligned}\mu^*(B) \;&=\; \mu^*\Big(B \biguplus\nolimits_{i=1}^n S_i\Big) + \mu^*\Big(B\,\Big(\biguplus\nolimits_{i=1}^n S_i\Big)^c\Big) \geqslant\\ &\geqslant\; \sum_{i=1}^{n} \mu^*(B\,S_i) + \mu^*\Big(B\,\Big(\biguplus\nolimits_{i=1}^\infty S_i\Big)^c\Big)\end{aligned}$$

при $\forall\, n \geqslant 1$. Полагая здесь $n \to \infty$ и воспользовавшись доказанной выше σ-полуаддитивностью μ^*, получаем

$$\mu^*(B) \geqslant \sum_{i=1}^{\infty} \mu^*(B\,S_i) + \mu^*\Big(B\,\Big(\biguplus_{i=1}^{\infty} S_i\Big)^c\Big) \geqslant \tag{15}$$

$$\geqslant \mu^*\Big(B \biguplus_{i=1}^{\infty} S_i\Big) + \mu^*\Big(B\,\Big(\biguplus_{i=1}^{\infty} S_i\Big)^c\Big) \geqslant \ \mu^*(B), \tag{16}$$

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 (16) (теперь уже равенства) следует, что $\biguplus_1^{\infty} S_i \in \Psi$. Во-вторых, полагая в (15) $B = \biguplus_1^{\infty} S_i$ $(\in \Psi)$, получаем σ-аддитивность μ^*.

Для произвольных множеств $S_n,\ n \geqslant 1,$ из класса Ψ

$$\bigcup_{n=1}^{\infty} S_n = \biguplus_{n=1}^{\infty}\Big(S_n \setminus \bigcup_{i=1}^{n-1} S_i\Big) \in \Psi,$$

так как 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 $\forall\, \epsilon > 0$ существует $\mathcal{A}$-покрытие $B \subset \bigcup_1^{\infty} S_n$, что

$$\mu^*(B) > \sum\nolimits_1^{\infty} \mu(S_n) - \epsilon.$$

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,

$$\mu^*(B) > \sum\nolimits_1^\infty \mu(S_n) - \epsilon = \sum\nolimits_1^\infty \mu(G\,S_n) + \sum\nolimits_1^\infty \mu(G^c\,S_n) - \epsilon \geqslant$$
$$\geqslant \mu^*(B\,G) + \mu^*(B\,G^c) - \epsilon.$$

Полагая $\epsilon \to 0$, 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 $\mu^*(G) \leqslant \mu(G)$. 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 $\gamma : \sigma(\mathcal{A}) \mapsto \mathbb{R}^1_+$ 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 $S_n \cap S_j = \varnothing, \;\; k \neq j$.

Зафиксируем $n \geqslant 1$ и определим класс

$$\mathscr{Q} = \big\langle\, G \in \sigma(\mathcal{A}) : \;\mu(G\,S_n) = \gamma(G\,S_n) \,\big\rangle.$$

По условию $\mathcal{A} \subset \mathscr{Q} \subset \sigma(\mathcal{A})$. 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 $G^c \in \mathscr{Q}$, так как значения мер в последнем равенстве конечны.

Докажем монотонность $\mathscr{Q}$. Пусть $B = \lim\limits_k {}_{\uparrow(\downarrow)} G_k, \quad G_k \in \in \mathscr{Q}$. 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 219, стр. 233, получаем, что $\mathscr{Q} = \mathfrak{M}(\mathcal{A}) = \sigma(\mathcal{A})$.

Таким образом, $\mu(B) = \sum_n^\infty \mu(B\, S_n) = \sum_n^\infty \gamma(B\, S_n) = \gamma(B)$ для $\forall\, B \in \sigma(\mathcal{A})$ с $\mu(B) < \infty$ или $\gamma(B) < \infty$. ⫤

✧ **Доказательство теоремы Лебега** (стр. 189).

⊫ Предположим, сначала, что меры γ, μ конечны.

a) Рассмотрим совокупность измеримых функций $H = \langle h \geqslant 0 : \int_A h\, d\mu \leqslant \gamma(A), \quad \forall\, A \in \mathscr{F} \rangle$, 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

$$\lim_{n\to\infty} \int_\Omega h_n\, d\mu \;=\; J \;:=\; \sup_{h\in H} \int_\Omega h\, d\mu \quad \leqslant \gamma(\Omega) < \infty.$$

Определим монотонную последовательность $g_n = \max\langle\, h_i, i = \overline{1, n}\,\rangle$. 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 156 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

$$D(A) = \left\langle B \in \mathscr{F}: \quad B \subset A, \quad \psi(B) < \frac{1}{n}\mu(B) \right\rangle.$$

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 $h_0 = \dfrac{1}{n}\mathbf{I}_A$ 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 $J < \infty$).

Определим $\forall\, k \geqslant 1$ множества $\tilde{B}_k = \bigcap\limits_{i=0}^{k-1} B_i^c$ $(B_0 = \varnothing)$ и

$$B_k \in D(\tilde{B}_k): \quad \mu(B_k) \geqslant \epsilon_k = \frac{1}{2}\,\mathsf{sup}\left\langle \mu(B): \quad B \in D(\tilde{B}_k) \right\rangle,$$

если $\mu(\tilde{B}_k) > 0$, иначе $B_k = \varnothing$, $\epsilon_k = 0$. Из монотонности семейства $D(A)$ следует, что $B_k \in D(\Omega)$, если $B_k \neq \varnothing$, и, кроме того, т.к. множества $B_k, k \geqslant 0$, не пересекаются, то и $Z = \biguplus_k B_k \in D(\Omega)$. Отсюда (ввиду конечности μ)

$$\infty > \mu(Z) = \sum\nolimits_1^\infty \mu(B_k) \geqslant \sum\nolimits_1^\infty \epsilon_k,$$

следовательно, $\epsilon_k \to 0$, $k \to \infty$.

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

$$2\epsilon_k = \mathsf{sup}\left\langle \mu(B): \quad B \in D(\tilde{B}_k) \right\rangle \geqslant$$
$$\geqslant \mathsf{sup}\left\langle \mu(B): \quad B \in D(Z^c) \right\rangle \geqslant \mu(Q) > 0.$$

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

$$\psi\Big(\bigcap_{n=1}^{\infty} Z_n\Big) \;\leqslant\; \lim_{n\to\infty}\psi(Z_n) = 0,$$

$$\mu\Big(\Big(\bigcap_{n=1}^{\infty} Z_n\Big)^c\Big) \;=\; \mu\Big(\bigcup_{n=1}^{\infty} Z_n^c\Big) \leqslant \sum_{n=1}^{\infty}\mu(Z_n^c) = 0,$$

что доказывает a) (единственность ξ легко следует из b)).

b) Пусть имеется 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 и 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 $\mathscr{F}$-разбиение $\Omega = \biguplus_1^{\infty} \Omega_k$, что $\mu(\Omega_k), \gamma(\Omega_k) < \infty$. Рассмотрим измеримое пространство $(\Omega_k, \mathscr{F}_k)$ с σ-алгеброй $\mathscr{F}_k = \Omega_k \Cap \mathscr{F} = \big\langle\, B\Omega_k : B \in \mathscr{F} \,\big\rangle$. 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 выше, найдутся мера $\psi_k \perp \mu_k$ и $\mathscr{F}_k$-измеримая функция $\xi_k \geqslant 0$, что для $\forall\, B \in \mathscr{F}$

$$\gamma_k(B\Omega_k) = \psi_k(B\Omega_k) + \int_{B\Omega_k} \xi_k\, d\mu_k$$

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 $\tilde{\xi} = \sum_k \tilde{\xi}_k\, \mathbf{I}_{\Omega_k}$ в силу теоремы о монотонном пределе под знаком интеграла

$$\int_B \tilde{\xi}\, d\mu = \sum\nolimits_1^\infty \int_{B\Omega_k} \xi_k\, d\mu_k.$$

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

$$\psi\Big(\biguplus\nolimits_1^\infty N_k\Big) = \sum\nolimits_1^\infty \psi_k(N_k) = 0,$$
$$\mu\Big(\Big(\biguplus\nolimits_1^\infty N_k\Big)^c\Big) = \sum\nolimits_1^\infty \mu_k(\Omega_k \setminus N_k) = 0.$$

Т.е. $\psi \perp \mu$ и пара $\tilde{\xi}, \psi$ удовлетворяет 175, a), стр. 189. Единственность разложения мер доказывается аналогично. $\Lleftarrow$

Доказательство леммы о произведении мер (стр. 210).

$\Rrightarrow$ 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 $(A+B)_{\omega_1} = A_{\omega_1} + B_{\omega_1}$ и, кроме того, $A_{\omega_1} \cap B_{\omega_1} = \varnothing$, 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 219, стр. 233, 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 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

$$\lim_{n}{}_{\uparrow} \mu_2((B_n)_{\omega_1}) = \mu_2(B_{\omega_1}) < \infty.$$

Т.к. 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 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:

$$\Omega_1 \times \Omega_2 = \Big(\biguplus_1^\infty Q_{1n} \times \Omega_2\Big) \bigcap \Big(\biguplus_1^\infty \Omega_1 \times Q_{2m}\Big) = \biguplus_{n,m=1}^\infty Q_{1n} \times Q_{2m}.$$

Для $B \in \mathscr{F}_1 \otimes \mathscr{F}_2$ положим $B^{(nm)} = B \cap (Q_{1n} \times Q_{2m})$, тогда, очевидно, $B_{\omega_1} = \biguplus_{nm} B_{\omega_1}^{(nm)}$ и в силу σ-аддитивности μ_2

$$\mu_2(B_{\omega_1}) = \sum\nolimits_{nm} \mu_2(B_{\omega_1}^{(nm)}).$$

Т.к. $\mu_2(B_{\omega_1}^{(nm)})$ измеримы (по ω_1) при каждых n, m, то и функция $\mu_2(B_{\omega_1})$ также измерима. Поэтому можно определить функцию множеств

$$\mu(B) := \sum\nolimits_{nm} \int_{\Omega_1} \mu_2(B_{\omega_1}^{(nm)})\, d\mu_1,$$

σ-аддитивность которой следует из σ-аддитивности μ_2, теоремы о монотонной сходимости для интеграла Лебега, а также из возможности перестановки слагаемых ряда с неотрицательными членами. Справедливость равенства $\mu(B_1 \times B_2) = \mu_1(B_1)\mu_2(B_2)$ проверяется простыми алгебраическими преобразованиями. ⫤

B. Основные понятия теории множеств

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;

✦ $\langle A_1, A_2, \ldots \rangle = \langle A_j \rangle_j = \langle A_j \rangle_1^n$ — совокупность выбранных элементов;

✦ $\mathscr{P}(\Omega) = \langle\, A \subseteq \Omega \,\rangle$ — класс всех подмножеств Ω, включая $\varnothing, \Omega$;

✦ $\biguplus$ (или $+$) — знак объединения непересекающихся множеств:

$$A_i\,A_j = \varnothing,\ i \neq j,\ \Leftrightarrow\ A_1 \cup A_2 = A_1 + A_2,\ \bigcup_{j=1}^{N} A_j = \biguplus_{j=1}^{N} A_j;$$

✦ $B \subset \bigcup_i S_i,\quad S_i \in \mathcal{H},$ — *покрытие* множества B элементами класса $\mathcal{H}$ или $\mathcal{H}$-*покрытие*;

✦ $B = \biguplus_i S_i,\quad S_i \in \mathcal{H},$ — *разбиение* множества B на непересекающиеся элементы класса $\mathcal{H}$ или $\mathcal{H}$-*разбиение*;

✦ индикаторная функция множества $A \subset \Omega$:

$$\mathbf{I}_A = \mathbf{I}(\omega; A) = \mathbf{I}(A) = \begin{cases} 1, & \text{если } \ \omega \in A, \\ 0, & \text{если } \ \omega \notin A. \end{cases}$$

✦ *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 множеств последовательности:

$$\overline{\lim_n}\, A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k;$$

✦ *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:

$$A, B \in \mathcal{H} \quad \Rightarrow \quad A\,B \in \mathcal{H};$$

C2) $A, B \in \mathcal{H} \quad \Rightarrow \quad \exists$ конечное $\mathcal{H}$-разбиение $A \setminus B$:

$$A \setminus B = \biguplus_{j=1}^{n} S_j, \quad \langle\, S_j \,\rangle_{j=1}^{n} \subset \mathcal{H}, \;\; n < \infty.$$

Полукольцо $\mathcal{H}$, содержащее основное пространство Ω, то есть полукольцо с «единицей», называется *полуалгеброй*.

205| Упр. а) Докажите, что системы множеств, описанные в примере 95, стр. 121, 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 $A \bigcup B \notin \mathcal{H}$ для подмножеств A, B из полукольца $\mathcal{H}$.

206| Лемма. *Пусть $Q_1, \ldots, Q_n \in \mathcal{H}$ — элементы полукольца. Тогда существует $\mathcal{H}$-разбиение $(B_1, \ldots, B_L \in \mathcal{H})$ разности*

$$Q_n \setminus \bigcup_{i=1}^{n-1} Q_i = \biguplus_{j=1}^{L} B_j.$$

⊫ 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 $S_k^{(3)} \setminus Q_1$)

$$Q_3 \setminus (Q_1 \cup Q_2) = \biguplus_{k=1}^{N_3} S_k^{(3)} \setminus Q_1 = \biguplus_{k=1}^{N_3} \biguplus_{i=1}^{M_k} S_{ki}^{(3)} = \biguplus_{j=1}^{L_3} B_j$$

с некоторыми наборами $\{S_{kj}^{(3)}, i = \overline{1, M_k}\}$, $\{S_k^{(3)}, k = \overline{1, N_3}\}$

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 $\forall\, n < \infty$

$$\texttt{R1)}\quad A, B \in \mathcal{D} \quad \Rightarrow\ B \setminus A \in \mathcal{D}$$

$$\texttt{R2)}\quad \langle A_i \rangle_1^n \subset \mathcal{D} \ \Rightarrow\ \bigcup_1^n A_i \in \mathcal{D}\ \text{ и }\ \bigcap_1^n A_i \in \mathcal{D}.$$

Кольцо, содержащее Ω, т.е. кольцо с «единицей», называется *алгеброй*.

207| $\mathfrak{Y}$пр. 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.

208| Лемма. (?!) *Класс множеств, состоящий из всевозможных конечных объединений непересекающихся элементов полукольца, образует кольцо. Это кольцо совпадает с классом всех конечных объединений (возможно, пересекающихся) элементов полукольца. Кроме того, это кольцо есть минимальное кольцо, содержащее все элементы полукольца.*

Определение. Кольцо $\mathfrak{C}(\mathcal{H})$ всех конечных объединений непересекающихся элементов полукольца $\mathcal{H}$ называется *кольцом, порождённым* $\mathcal{H}$.

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) $A \in \mathscr{F} \;\Rightarrow\; A^c = \Omega \setminus A \in \mathscr{F}$ — $\mathscr{F}$ замкнута относительно операции дополнения;

S3) $\langle A_i \rangle_1^\infty \subset \mathscr{F} \;\Rightarrow\; \bigcup\limits_{i=1}^{\infty} A_i \in \mathscr{F}$ — $\mathscr{F}$ замкнута относительно счётного объединения своих элементов.

209| $\mathfrak{Y}\mathfrak{p}\mathfrak{p}$. 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.

210| Примеры. Следующие системы подмножеств образуют σ-алгебры:

- $\mathscr{F} = \langle \varnothing, \Omega \rangle$, $\qquad$ - $\mathscr{F} = \langle \varnothing, \Omega, A, A^c \rangle$, где $A \subset \Omega$,
- $\mathscr{F} = \langle A \subseteq [0;1], \; A$ — измеримо по Лебегу $\rangle$,
- $\mathscr{F} = \mathscr{P}(\Omega)$ — совокупность всех подмножеств $A \subseteq \Omega$.

Последнюю σ-алгебру часто обозначают как 2^Ω, поскольку для Ω, состоящего из конечного числа k элементов, количество элементов этой σ-алгебры $\#(\mathscr{P}(\Omega))$ равно 2^k. ◎

211| $\mathfrak{Y}\mathfrak{n}\mathfrak{p}$. 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: $B_1, B_2 \in \mathcal{M}, B_1 \subset \subset B_2 \ \Rightarrow\ B_2 \setminus B_1 \in \mathcal{M}$.

212| 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). ◎

213| 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 $\Omega \in \mathcal{M}$.*

⊫ I) Так как $\mathcal{A}$ — алгебра, то для $\forall \left\langle S_i \right\rangle_1^\infty \subset \mathcal{A}$ имеем $B_n = \bigcup_1^n S_i \in \mathcal{A}$ 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 Морга̀на: $A \cup B = \Omega \setminus \big((\Omega \setminus A) \cap (\Omega \setminus B)\big)$. ⫤

214| Пример. Рассмотрим какое-либо счётное разбиение $\Sigma = \langle S_i \rangle_1^\infty$ основного пространства: $\Omega = \biguplus_1^\infty S_i$. Свяжем с этим разбиением класс множеств, полученных путём конечного или счётного объединения элементов Σ:

$$\mathfrak{S} = \Big\langle \biguplus_{i=1}^{\infty} S_i^{(j_i)} : j_i = \pm 1, i \geqslant 1 \Big\rangle,$$

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. ◎

215| Теорема. *Для любой совокупности $\mathcal{Q}$ подмножеств Ω здесь существуют единственные минимальные σ-алгебра $\sigma(\mathcal{Q})$, монотонный класс $\mathfrak{M}(\mathcal{Q})$ и λ-система $\Lambda(\mathcal{Q})$, содержащие все элементы $\mathcal{Q}$.*

⊫ 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 $\langle \mathscr{A}_t \rangle_{t \in \Upsilon}$ всех σ-алгебр, каждая из которых содержит все элементы $\mathcal{Q}$, не пуста. Тогда класс, состоящий из множеств, входящих во все σ-алгебры $\mathscr{A}_t$, $t \in \Upsilon$:

$$\sigma(\mathcal{Q}) = \bigcap\nolimits_{t \in \Upsilon} \mathscr{A}_t,$$

удовлетворяет утверждению теоремы. Аналогично для λ-системи монотонных классов. ⫷

216| $\mathfrak{Упр}$. Докажите, что пересечение любого набора σ-алгебр (монотонных классов, λ-систем) будет σ-алгеброй (монотонным классом, λ-системой).

При доказательстве совпадения σ-алгебр, определённых двумя различными способами, весьма востребованы утверждения следующего упражнения.

217| $\mathfrak{Упр}$. 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. 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).

218| $\mathfrak{Упр}$. Опишите минимальный монотонный класс $\mathfrak{M}(\mathcal{K})$, порождённый полукольцом интервалов на прямой вида $(a;b]$.

219| Теорема. *Если $\mathcal{H}$ — кольцо и минимальный монотонный класс $\mathfrak{M}(\mathcal{H}) \ni \Omega$, тогда $\mathfrak{M}(\mathcal{H}) = \sigma(\mathcal{H})$.*

⊫ 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 $\forall A_1, A_2 \in$ $\in \mathfrak{M}(\mathcal{H})$.

Покажем, что для любого $S \subset \Omega$ класс множеств

$$\mathfrak{T}(S) = \big\langle A \subset \Omega : \ S \cup A,\ S \setminus A,\ A \setminus S \in \mathfrak{M}(\mathcal{H}) \big\rangle$$

образует монотонный класс. Пусть последовательность $\langle C_n \rangle_1^\infty \subset \mathfrak{T}(S)$ возрастает и $C = \lim\limits_{n}{}_{\uparrow} C_n$. По определению $\mathfrak{T}(S)$

$$S \cup C_n, \;\; C_n \setminus S, \;\; S \setminus C_n \in \mathfrak{M}(\mathcal{H}),$$

причём последовательности $\left\langle\, S \cup C_n \,\right\rangle_1^\infty, \left\langle\, C_n \setminus S \,\right\rangle_1^\infty$ — возрастающие, а последовательность $\left\langle\, S \setminus C_n \,\right\rangle_1^\infty$ — 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 $\forall\, A_1, A_2 \in \mathfrak{M}(\mathcal{H})$, что и требовалось. ⫷

Кольцо в этом утверждении замкнуто относительно операций пересечения и объединения. В следующей теореме порождающий класс замкнут только относительно операции пересечения.

220| Теорема. *Пусть $\mathcal{H}$ — класс подмножеств Ω, замкнутый относительно конечных пересечений. Если минимальная λ-система $\Lambda(\mathcal{H}) \ni \Omega$, то $\sigma(\mathcal{H}) \subset \Lambda(\mathcal{H})$.*

⫸ 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 $\mathcal{H} \subset \mathrm{T}(F)$ (последнее 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.

Итак, 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 $B \cap F \in \Lambda(\mathcal{H})$. ⫤

C. Топологические пространства

Обсудим вкратце некоторые вопросы общей топологии, потребность в которых возникает при построении вероятностных пространств.

(✓) Задать *топологию* в 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).

Примеры. a) Естественная топология на числовой прямой задаётся открытыми множествами, которые с каждой своей точкой содержат и некоторый накрывающий эту точку открытый интервал.

б) В любом пространстве можно задать тривиальную (самую бедную) топологию $\mathcal{T} = \left\langle \varnothing, \Omega \right\rangle$ и дискретную (самую богатую) топологию $\mathcal{T} = \left\langle B : B \subset \Omega \right\rangle$, 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 $\rho(x, y)$ между точками $x, y \in \Omega$ базу топологии образуют все шары —

$$\mathfrak{B}_\mathcal{T} = \left\langle U_r(x) := \{\omega \in \Omega : \rho(x, \omega) < r\} : \ x \in \Omega, r > 0 \right\rangle.$$

Примеры. 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:

$$\bigcup\nolimits_\beta B_\beta \supset K, \quad \left\langle B_\beta \right\rangle_\beta \subset \mathcal{T} \Rightarrow$$

$$\bigcup_{i=1}^{n} B_{\beta_i} \supset K, \quad \left\langle B_{\beta_i} \right\rangle_{i=1}^{n} \subset \left\langle B_\beta \right\rangle_\beta, \ n < \infty.$$

✦ 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 $\{x_n\}_{n=1}^{\infty} \subset \Omega$ 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. *Прямым произведением* $\Omega = \prod_\alpha X_\alpha$ называется пространство всех функций $\omega(\alpha)$ $(= \omega_\alpha)$, заданных на $\aleph$ и при каждом α принимающих значения в 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 $\mathcal{T}_\alpha$.

✦ 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 $\left\langle\, \omega^{(n)}, n = 1, 2, \ldots \right\rangle \subset \Omega$ сходится к ω' в тихоновской топологии т. т. т. когда она сходится поточечно (покоординатно): $\omega^{(n)} \to \omega' \;\;\Leftrightarrow\;\; \omega^{(n)}_\alpha \to \omega'_\alpha$ 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-ый 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:

$$\forall\, \epsilon > 0\ \exists k\ \forall\, n, m > k : \rho(x_n, x_m) < \epsilon \Rightarrow \exists x \in \Omega : \ \rho(x_n, x) \to 0.$$

✦ Топологическое пространство $(\Omega, \mathcal{T})$ *сепарабельно*, если найдётся счётное всюду плотное подмножество $\mathfrak{S} \subset \Omega$, т.е. $A\mathfrak{S} \neq \varnothing$ для любого открытого подмножества $A \subset \Omega$.

✦ 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

$$1)\ (x,y)=(y,x), \qquad 2)\ (x,x)\geqslant 0,\ (x,x)=0 \ \Leftrightarrow\ x=0,$$

$$3)\ (x_1+x_2,y)=(x_1,y)+(x_2,y), \quad 4)\ (cx,y)=c(x,y),\ c\in\mathbb{R}.$$

✦ В евклидовом пространстве норма любого элемента определяется как $\|x\| = \sqrt{(x,x)}$.

✦ Если евклидово пространство полно относительно нормы, определяемой скалярным произведением, то такое пространство называется гильбертовым.

✦ Пусть $(\Omega, \mathscr{F}, \mu)$ — пространство с σ-конечной мерой. Совокупность 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 средне-квадратической ошибке $\mathbf{E}\left[(\xi-\eta)^2\right]$. Эту сл.в. можно назвать условным матем.ожиданием $\mathbf{E}\left[\,\xi \,|\, \mathscr{A}\,\right]$ относительно σ-алгебры $\mathscr{A}$.

D. 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: $\alpha\vec{x}_1 + (1-\alpha)\vec{x}_2 \in \mathcal{X}$, $\forall\,\alpha \in [0;1]$.

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 $\forall\,\alpha \in [0;1]$

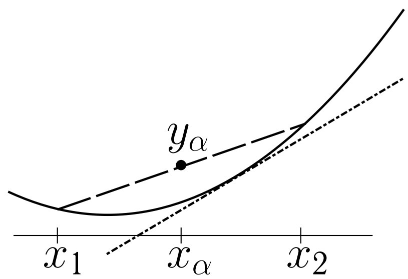
$$(*) \qquad h(\alpha\vec{x}_1 + (1-\alpha)\vec{x}_2) \leqslant \alpha h(\vec{x}_1) + (1-\alpha)h(\vec{x}_2);$$

— *строго выпуклой*, если неравенство $(*)$ строгое для любого $\alpha \in (0;1)$;

— *вогнутой*, если выпукла функция $-h$.

Другими словами, любая дуга графика выпуклой функции лежит ниже хорды, соединяющей края этой дуги. Рисунок справа 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.

(✓) 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):

$$(\circledast) \qquad \forall\,\vec{x}_0 \in \mathcal{X}\ \exists\vec{b}: \quad \vec{b}^\flat(\vec{x}-\vec{x}_0) + h(\vec{x}_0) \leqslant h(\vec{x}),\ \forall\,\vec{x} \in \mathcal{X}.$$

$\Rrightarrow$ Докажем этот факт для случая $k = 1$ и области $\mathcal{X} = (a; c)$. Пусть $x_0 \in \mathbb{R}^1$. 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,

$$(h(x') - h(x_0))\,\frac{x - x_0}{x - x'} \geqslant (h(x_0) - h(x))\,\frac{x_0 - x'}{x - x'},$$

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. $\Lleftarrow$

(✓) 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 первооткрывателя выпуклых функций:

$$h\Big(\sum_{i=1}^{n} \alpha_i \vec{x}_i\Big) \leqslant \sum_{i=1}^{n} \alpha_i h(\vec{x}_i),$$

$$\forall\, n \geqslant 1,\ \{\vec{x}_i\}_{i=1}^{n} \subset \mathcal{X},\ 0 \leqslant \alpha_1, \ldots, \alpha_n \leqslant 1,\ a_1 + \ldots + \alpha_n = 1.$$

Пример. Функция $h(\vec{x}) = \|\vec{x}\| = \sqrt{\vec{x}^\flat \vec{x}}$ на $\mathbb{R}^k$ выпукла, т.к. по неравенству треугольника

$$\|\alpha \vec{x}_1 + (1 - \alpha)\vec{x}_2\| \leqslant \alpha \|\vec{x}_1\| + (1 - \alpha)\|\vec{x}_2\|.$$

(✓) Ограниченная выпуклая функция непрерывна всюду.

(✓) 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 $\alpha_j = {}^1/_n$, $j = \overline{1,n}$, получаем неравенство $\ln(\sum_j x_j/n) \geqslant \frac{1}{n}\sum_j \ln(x_j)$ 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.

Пример. Функция $h(x) = x^a$, $x > 0$, строго выпукла при $a > 1$, т.к. $h''(x) = x^{a-2}/(a(a-1)) > 0$. Поэтому для $\forall\, x \neq \neq y$ (> 0)

$$(x+y)^a = 2^a\Big(\frac{1}{2}x + \frac{1}{2}y\Big)^a < 2^a\Big(\frac{1}{2}x^a + \frac{1}{2}y^a\Big) = 2^{a-1}\big(x^a + y^a\big). \tag{1}$$

При $a < 1$ функция $h(x) = x^a$, $x > 0$, строго вогнута, поэтому

$$(x+y)^a > \frac{x^a + y^a}{2^{1-a}}, \quad x, y > 0,\ x \neq y,\ 0 < a < 1.$$

Например, $\sqrt{2}\sqrt{x+y} > \sqrt{x}+\sqrt{y}$.

Рассмотрим функцию $h(u,v) = u^a v^{1-a}$, $u > 0, v > 0$, с $0 < a < 1$. Матрица вторых производных функции $-h$ 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.

Пример. Как отмечалось, функция $h(x_1,x_2) = \|(x_1,x_2)\| = \sqrt{x_1^2+x_2^2}$ 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

$$\sqrt{x_1^2+x_2^2} \geqslant \frac{ax_1+bx_2}{\sqrt{a^2+b^2}}\,.$$

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.

E. Указания к решению задач

✦ 90, стр. 116. б) $B_k B_m = \varnothing$. 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) $\mathbf{I}(\bigcap_k B_k) = \inf_k \mathbf{I}(B_k)$. Индикатор $\mathbf{I}(\underline{\lim}_n Q_n) = \mathbf{I}(\lim_{n}{}_{\uparrow} \bigcap_{k=n}^{\infty} Q_k) = \lim_{n}{}_{\uparrow} \mathbf{I}(\bigcap_{k=n}^{\infty} Q_k) = \lim_n \inf_{k\geqslant n} \mathbf{I}(Q_k) = \underline{\lim}_n\ \mathbf{I}(Q_n)$.

✦ 91, стр. 117. $B = B \biguplus \varnothing$, $\quad A \biguplus B = A \biguplus B \biguplus \varnothing \biguplus \varnothing \ldots$.

✦ 92, стр. 117. $\#((A \uplus B)\mathcal{S}) = \#(A\mathcal{S}) + \#(B\mathcal{S})$ — очевидно. Рассмотреть случаи $\mu(\uplus_1^\infty B_k) < \infty$ и $= \infty$. Во втором случае $\forall\, M\ \exists K:\ \mu(\uplus_1^K B_k) > M$.

✦ 93, стр. 118. Если $Q, B \in \mathcal{H}$, то $B \supset QB \in \mathcal{H} \;\Rightarrow\; B =$

$= QB + \biguplus_1^k Q_j = QB + (B \setminus Q)$.

✦ 95, стр. 121. Пересечение $(a;c] \cap (b;d]$ или пусто, или есть интервал того же вида. Если $(a;c] \subsetneq (b;d]$, то $(b;d] = (a;c] + \big[(a;b] \uplus (c;d]\big]$. Относительно прямоугольников см. 96, стр. 122.

✦ 96, стр. 122. В соответствии с 93, стр. 118, пусть $A_1 \times A_2 \supset B_1 \times B_2 \;\Rightarrow\; A_1 = B_1 + \uplus_j Q_j,\ A_2 = B_2 + \uplus_n R_n \;\Rightarrow\; A_1 \times A_2 = B_1 \times B_2 + \uplus_j Q_j \times A_2 + \uplus_n A_2 \times R_n + \uplus_{jn} Q_j \times R_n$.

✦ 99, стр. 125. $S_n = [n; \infty)$.

✦ 100, стр. 126. 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.

✦ 101, стр. 126. 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 $\mu(\biguplus_k \biguplus_j Z_j) = \sum_k \sum_j \mu(Z_j)$.

✦ 102, стр. 127. (i_1), (i_2) — определение. (i_3) — $\mu(\Omega) \leqslant \sum_i \mu(B_i)$ в силу σ-полуаддитивности.

✦ 103, стр. 127. В силу монотонности и полуаддитивности внешней меры: $\mu^*(B) \leqslant \mu^*(BG) + \mu^*(BG^c) = \mu^*(BG^c) \leqslant \mu^*(B)$.

✦ 104, стр. 128. $\mathcal{H} \subset \mathcal{A}$ и $\sigma(H) \supset \mathcal{A}$.

✦ 109, стр. 132. Воспользоваться 102, стр. 127, и тем, что инфимум равен нулю т. т. т. когда найдётся последовательность элементов, сходящаяся к нулю.

✦ (4), стр. 132. $\mu(C) = \mu(A) + \mu(C \setminus A) = \mu(A) \leqslant \mu(D)$ 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 $\mu(\bigcap C_k \setminus \bigcap A_k) \leqslant \sum_k \mu(C_k \setminus A_k)$.

✦ 110, стр. 133. $(a;b) = \bigcup_k (a; b+\frac{1}{k}]$. Любое открытое множество есть счётное объединение открытых интервалов; замкнутое множество есть дополнение какого-то открытого множества до $\mathbb{R}^1$.

✦ 111, стр. 134. "$\underset{4}{=}$" $[a;b) = \lim_k (a-\frac{1}{k}; b-\frac{1}{k}] \in \sigma(\mathcal{K})$, $(a;b] = \lim_k [a+\frac{1}{k}; b+\frac{1}{k}) \in \sigma(\mathcal{K}_{[)})$. "$\underset{5}{=}$" $\mathcal{K} \subset \mathcal{U} \subset \sigma(\mathcal{K})$. "$\underset{6}{=}$" $(a;b] = (-\infty; b] \setminus (-\infty; a]$.

✦ 114, стр. 135. $A_k \in Q \Cap \mathscr{F} \Rightarrow A_k = QB_k,\ B_k \in \mathscr{F} \Rightarrow \bigcup A_k = Q \bigcup B_k,\ \bigcup B_k \in \mathscr{F} \Rightarrow \bigcup A_k \in Q \Cap \mathscr{F}$.

✦ 116, стр. 137. Класс множеств $B \in \mathscr{B}(\overline{\mathbb{R}}^1)$, удовлетворяющих условию леммы, есть σ-алгебра, содержащая интервалы, определяющие σ-алгебру $\mathscr{B}(\overline{\mathbb{R}}^1)$.

✦ 118, стр. 139. Ср. с 110, стр. 133. Т.к. $(\vec{a}; \vec{c}] = \{\vec{x} : x_1 \leqslant \leqslant c_1, a_2 < x_1 \leqslant c_2, \ldots\} - \{\vec{x} : x_1 \leqslant a_1, a_2 < x_1 \leqslant c_2, \ldots\}$, то для принадлежности к σ-алгебре $\sigma(\mathcal{K}^{\urcorner})$ достаточно установить сей факт для каждого из этих подмножеств; далее по индукции.

✦ (?!), стр. 141. Любые конечные наборы множеств из $\mathcal{B}$ суть цилиндры с общим конечномерным пространством.

✦ 123, стр. 149. Пусть F непрерывна слева, тогда $\mu[a;b] = = \lim_k \mu[a; b+\frac{1}{k}) = \lim_k (F(b+\frac{1}{k}) - F(a)) = F(b+) - F(a) = = F(b+) - F(a-)$.

✦ 125, стр. 153. Функция $F(x_1, x_2) = \operatorname{sign}(x_1 x_2)\mu(\Pi)$, где Π — параллелепипед от $(\min(x_1, 0), \min(x_2, 0))$ до $(\max(x_1, 0), \max(x_2, 0))$, генерирует меру μ.

✦ 126, стр. 154. $\omega \in \xi^{-1}(\bigcup B_\alpha) \Leftrightarrow \xi(\omega) \in \bigcup B_\alpha \Leftrightarrow \big[\exists \alpha : \xi(\omega) \in B_\alpha\big] \Leftrightarrow \omega \in \bigcup_\alpha \xi^{-1}(B_\alpha)$.

✦ 136, стр. 162. 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 $\xi^- = -\xi \geqslant -\eta = \eta^-$.

✦ 142, стр. 164. Бесконечная сумма есть предел измеримых конечных сумм.

✦ 151, стр. 174. $\mu(\xi^\pm = \eta^\pm) = 0$;

✦ 152, стр. 174. $\delta(h \neq \tilde{h}) = 0$, где $\tilde{h} = h(\omega_0)\,\mathbf{I}(\omega = \omega_0)$.

✦ 153, стр. 175. По неравенству Маркова $\varliminf\limits_n \int x_n\,d\mu \geqslant$ $\geqslant a\mu(\xi > a)$, $a > 0$.

✦ 154, стр. 175. б) Для простых очевидно.

✦ 155, стр. 176. $\mu(\xi \neq \tilde{\xi}) = 0$, где $\tilde{\xi} = \sum_i \xi(\omega_i)\,\mathbf{I}(\omega = \omega_i)$.

✦ 157, стр. 177. В силу 153, стр. 175, можно считать $\mu(\eta = \infty) = 0$. Применить теорему Б. Леви к последовательности $\xi_n\,\mathbf{I}(\eta < \infty) \nearrow \eta\,\mathbf{I}(\eta < \infty)$.

✦ 162, стр. 179. $\mu_\xi(\uplus B_k) = \mu(\uplus \xi^{-1}(B_k)) = \sum \mu(\xi^{-1}(B_k)) =$ $= \sum \mu_\xi(B_k)$.

✦ 165, стр. 181. Меры $\lambda(bA + c)$ и $|b|\lambda(A)$ совпадают на параллелепипедах. Другой способ: $x \in bA + c \;\Leftrightarrow\; \dfrac{x - c}{b} \in A$; по предшествующей лемме $|b|\lambda(A) = |b|\int \mathbf{I}_A(t)dt =$ $= \int \mathbf{I}_A(\frac{x-c}{b})dx = \lambda(bA + c)$.

✦ 168, стр. 183. Использовать неравенство $\sqrt[2]{2} > \sqrt[3]{2}$.

✦ Замечание, стр. 184. $\int |\xi_1|^p = 0 \Rightarrow \xi_1 = 0 \Rightarrow \xi_1\xi_2 =$ $= 0$ [п.в.] .

✦ 173, стр. 188. Применить теорему Лебега об ограниченной сходимости к последовательности $\xi\,\mathbf{I}_{A_n}$.

✦ 177, стр. 190. Применить 176, стр. 190, к парам $\gamma \ll \rho$ и $\rho \ll \mu$.

✦ 181, стр. 195. Пусть c, $c \notin \mathscr{X}$, и $x_{i_j} \in (c - 1; c) \cap \mathscr{X}$. Если $a \nearrow c$, то $\exists j_m \to \infty$: $F(c) - F(a) = \sum_{j \geqslant j_m} p_{i_j} \to 0$ по свойству сходящихся рядов. То же самое для $\forall\, a$ и $c \searrow a$. Для $\forall\, c = x_j \in \mathscr{X}$ разность $F(c) - F(a) = p_j + \sum_{a < x_k < c} p_k$.

✦ 182, стр. 195. См. указание к 12, стр. 26, в первой главе.

✦ 205, стр. 228. 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.

✦ 207, стр. 229. а) $A\cap B=(A\cup B)\setminus(A\setminus B)(B\setminus A)$. б) Из определения полукольца $B\setminus A=\bigcup_i S_i\in\mathcal{D}$ (свойство $R2'$).

✦ 208, стр. 229. Если A, B_k из полукольца, то $B\setminus A = \big(\biguplus_k B_k\big)\setminus A=\biguplus_k \big(B_k\setminus A\big)=\biguplus_k\biguplus_j Q_{kj}$. Далее применить равенство $(B_1\uplus B_2)\setminus(A_1\uplus A_2)=(B_1\setminus A_1)\setminus A_2\uplus(B_2\setminus A_1)\setminus A_2$.

✦ 209, стр. 230. 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.

✦ 211, стр. 231. 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 $\omega_j\in A$, $j=\overline{1,k}$.

✦ 216, стр. 233. а) $\sigma(\mathcal{Q}_1),\sigma(\mathcal{Q}_2)$ — σ-алгебры, содержащие $\mathcal{Q}_1$, причём первая из них — минимальная. б, в) Применить а) и $\sigma(\sigma(\mathcal{Q}))=\sigma(\mathcal{Q})$.

✦ 218, стр. 233. $\lfloor a;b\rceil$, $\infty\leqslant a<b\leqslant\infty$, $\lfloor=[$ или $($, $\rceil =]$ или $)$.

✦ 202, стр. 215. Применить интегрирование по частям.

F. Основные вероятностные законы

Дискретные распределения.

✦ Биномиальное — $\mathcal{B}\mathtt{in}(n,p)$, $p \in (0;1), n \in \mathbb{N}$:

$$\mathbf{P}\{\xi = k\} = \mathrm{C}_n^k p^k (1-p)^{n-k}, \quad k = 0,1,\ldots,n, \quad \mathbf{E}\xi = np, \quad \mathbf{D}\xi = np(1-p).$$

✦ Геометрическое — $\mathcal{G}\mathtt{eo}(p)$, $p \in (0;1)$:

$$\mathbf{P}\{\xi = k\} = p(1-p)^{k-1}, \quad k = 1,2,\ldots,n, \quad \mathbf{E}\xi = \frac{1}{p}, \quad \mathbf{D}\xi = \frac{(1-p)}{p^2}.$$

✦ Пуассона — $\mathcal{P}\mathtt{ois}(\lambda)$, $\lambda > 0$:

$$\mathbf{P}\{\xi = k\} = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k \in \mathbb{N}, \quad \mathbf{E}\xi = \lambda, \quad \mathbf{D}\xi = \lambda.$$

✦ Гипергеометрическое — $\mathcal{H}\mathtt{g}(n, N_1, N_2)$, $1 \leqslant n \leqslant N_1 + N_2$:

$$\mathbf{P}\{\xi = k\} = \frac{\mathrm{C}_{N_1}^k \mathrm{C}_{N_2}^{n-k}}{\mathrm{C}_{N_1+N_2}^n}, \quad \max\{0, n - N_2\} \leqslant k \leqslant \min\{N_1, n\},$$

$$\mathbf{E}\xi = \frac{n\,N_1}{N_1 + N_2}, \quad \mathbf{D}\xi = \frac{n\,N_1 N_2 (N_1 + N_2 - n)}{(N_1 + N_2 - 1)(N_1 + N_2)^2}.$$

Абсолютно–непрерывные распределения.

✦ Равномерное — $\mathcal{U}\mathtt{n}(a,b)$, $a < b$, $a, b \in \mathbb{R}^1$:

$$f(x) = \frac{1}{b-a}, \quad x \in [a;b], \quad \mathbf{E}\xi = \frac{b+a}{2}, \quad \mathbf{D}\xi = \frac{(b-a)^2}{12}.$$

✦ Бета — $\mathcal{B}\mathtt{et}(\alpha_1, \alpha_2)$, $\alpha_1, \alpha_2 > 0$: $(x \in (0;1))$

$$f(x) = \frac{x^{\alpha_1 - 1}(1-x)^{\alpha_2 - 1}}{\mathrm{B}(\alpha_1, \alpha_2)}, \quad \mathbf{E}\xi = \frac{\alpha_1}{\alpha_1 + \alpha_2}, \quad \mathbf{D}\xi = \frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)^2 (\alpha_1 + \alpha_2 + 1)}.$$

✦ Показательное — $\mathcal{E}\mathtt{x}(\lambda)$, $\lambda > 0$:

$$f(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}}, \quad x \geqslant 0, \quad \mathbf{E}\xi = \lambda, \quad \mathbf{D}\xi = \lambda^2.$$

✦ Лапласа (двустороннее показательное) — $\mathcal{L}\mathtt{ap}(\lambda)$, $\lambda > 0$:

$$f(x) = \frac{1}{2\lambda} e^{-\frac{|x|}{\lambda}}, \quad x \in \mathbb{R}^1, \quad \mathbf{E}\xi = 0, \quad \mathbf{D}\xi = 2\lambda^2.$$

✦ Гамма — $\mathcal{G}\mathtt{am}(a, \lambda)$, $a, \lambda > 0$:

$$f(x) = \frac{1}{\Gamma(a)\lambda^a} x^{a-1} e^{-\frac{x}{\lambda}}, \quad x > 0, \quad \mathbf{E}\xi = a\lambda, \quad \mathbf{D}\xi = a\lambda^2.$$

✦ Нормальное (Гаусса) — $\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}^1, \sigma^2 > 0$:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\}, \quad x \in \mathbb{R}^1, \quad \mathbf{E}\xi = \mu, \quad \mathbf{D}\xi = \sigma^2.$$

✦ Коши — $\mathcal{C}\mathtt{auch}(\mu, \sigma)$, $\mu \in \mathbb{R}^1, \sigma > 0$:

$$f(x) = \frac{1}{\pi\sigma} \frac{1}{1 + \left(\frac{x-\mu}{\sigma}\right)^2}, \quad x \in \mathbb{R}^1, \qquad \mathbf{E}\xi^{\pm} = \infty.$$

Многомерные распределения.

✦ Многомерное гипергеометрическое — $\mathcal{H}\mathtt{g}_k(n; N_1, \dots, N_k)$ — стр. 17.

✦ Мульти(поли)номиальное — $\mathcal{M}\mathtt{ult}_k(n, q_1, \dots, q_k)$ — стр. 18.

✦ Дирихле — $\mathcal{D}\mathtt{ir}_k(\alpha_1, \dots, \alpha_{k+1})$ — стр. 46.

✦ Нормальное — $\mathcal{N}_2(\vec{\mu}, \mathbf{\Sigma})$ — стр. 46.

✦ Равномерное — $\mathcal{U}\mathtt{n}(\mathscr{X})$ — стр. 45.

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	$\mathfrak{W}$: дубльВ	$\mathfrak{X}$: Икс	$\mathfrak{Y}$: Игрик	$\mathfrak{Z}$: Зет

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Литература

[1] Dudley R.M. *Real analysis and probability.*/ R.M. Dudley. – Cambridge: University Press, 2004. – 555 p.

[2] Данфорд Н. *Линейные операторы. Общая теория.*/ Н. Данфорд, Дж. Шварц. – М.: ИЛ, 1962. – 896 стр.

[3] Дороговцев А.Я. *Элементы общей теории меры и интеграла*: уч.пособие/ А.Я. Дороговцев. – Киев: Вища школа, 1988. – 152 стр.

[4] Зорич В.А. *Математический анализ.* Часть II. – 4-е изд./ В.А. Зорич. – М.: МЦНМО, 2002. – 794 стр.

[5] Ильин В.А. *Основы математического анализа.* Ч. II: уч.пособие./ В.А. Ильин, Э.Г. Позняк. – М.: ФИЗМАТЛИТ, 2002. – 464 стр.

[6] Колмогоров А.Н. *Элементы теории функций и функционального анализа.* / А.Н. Колмогоров, С.В. Фомин. – М.: Наука, 1989. – 543 стр.

[7] Невё Ж. *Математические основы теории вероятностей.*/ Ж. Невё. – М.: Мир, 1969. – 310 стр.

[8] Феллер В. *Введение в теорию вероятностей и её приложения.* В 2-х томах / В. Феллер. – М.: Мир, 1967. – 1250 стр.

[9] Фихтенгольц П. *Курс дифференциального и интегрального исчисления.* В 3-х томах / П. Фихтенгольц. – М.: ФИЗМАТЛИТ, 2001.

[10] Chow Y. S. *Probability Theory.* – 3rd ed./ Y.S. Chow, H. Teiher. – New York: Springer-Verlag, 1997. – 488 p.

[11] Шилов Г.Е. *Интеграл, мера и производная. Общая теория.*/ Г.Е. Шилов, Б.Л. Гуревич. – М.: Наука, 1967. – 220 стр.

[12] Ширяев А.В. *Вероятность* — 1. 2./ А.В. Ширяев. – М.: Изд-во МЦНМО, 2004. – 928 стр.