МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА
Механико-математический факультет

Курс лекций по теории вероятностей

Лектор — Александр Вадимович Булинский

II курс, 4 семестр, поток математиков

Москва, 2004 г.

Оглавление

Предисловие

Данное издание является результатом совместного творчества DMVN Corporation и Дмитрия Горяшина. Пока работа полностью не завершена, трудно сказать, оказался ли этот эксперимент удачным или нет. Стандартизовать обозначения и исходные TEX-исходники непросто, но ряд усилий в этом направлении всё же было приложено. Случайные величины мы решили обозначать греческими буквами, а не латинскими, как это было на лекциях — все-таки, это общепринятый стандарт.

Порядок следования теорем, определений и т. д. в целом сохранён таким, каким он был на лекциях в 2004 году. Некоторые перестановки с целью более логичной группировки утверждений были осуществлены во второй и третьей главах настоящего издания. В частности, собраны воедино все теоремы, относящиеся к теории интеграла Лебега.

В данной версии исправлено несколько опечаток, за что спасибо Игорю Приходько.

Вопросы, комментарии, замечания и предложения направляйте на `dmvn@mccme.ru`, обновления электронной версии — на сайте `http://dmvn.mexmat.net`.

Набор и вёрстка: `DMVN Corporation, GorDmit`
Последняя компиляция: `15 февраля 2006 года`
Версия: `0.3`

1. Элементарная теория вероятностей

1.1. Вероятностные пространства

1.1.1. Предмет теории вероятностей

Теория вероятностей изучает математические модели случайных явлений. Один и тот же случайный эксперимент может быть описан с помощью разных пространств элементарных исходов. Одно и то же пространство может описывать разные случайные эксперименты.

1.1.2. Алгебры событий

Будем рассматривать некоторое (фиксированное) множество Ω и различные семейства его подмножеств. Его мы будем называть *множеством элементарных исходов*, а подмножества $A \subset \Omega$ — *событиями*.

Определение. *Алгеброй* подмножеств на Ω называется семейство $\mathscr{A} \subset 2^{\Omega}$ со следующими свойствами:
1° $\Omega \in \mathscr{A}$;
2° Если $A \in \mathscr{A}$, то и $\overline{A} := \Omega \setminus A \in \mathscr{A}$.
3° Если $A_1, \ldots, A_n \in \mathscr{A}$, то и $\bigcup\limits_{i=1}^{n} A_i \in \mathscr{A}$.

Из определения следует, что и пересечение конечного набора элементов алгебры принадлежит ей: по правилу Де Моргана $\bigcap\limits_{i=1}^{n} A_i = \overline{\bigcup\limits_{i=1}^{n} \overline{A}_i} \in \mathscr{A}$, так как $\overline{A}_i \in \mathscr{A}$.

Определение. Система $\mathscr{F}$ подмножеств Ω называется *σ-алгеброй*, если она является алгеброй, но к тому же выдерживает счётные объединения и пересечения множеств, т. е. если $\{A_i\}_{i=1}^{\infty} \subset \mathscr{F}$, то и $\bigcup\limits_{i=1}^{\infty} A_i \in \mathscr{F}$.

Элементы алгебр и σ-алгебр часто называют *измеримыми* множествами (по аналогии с классами множеств, измеримых по Жордану или по Лебегу).

Заметим, что для любой системы $\mathcal{M}$ подмножеств Ω существует минимальная σ-алгебра $\sigma\{\mathcal{M}\}$, порождённая этой системой. Хотя бы одна есть — 2^{Ω}, очевидно, является σ-алгеброй. Рассмотрим пересечение всех σ-алгебр, содержащих $\mathcal{M}$. Легко видеть, что оно тоже будет σ-алгеброй. Это и будет $\sigma\{\mathcal{M}\}$.

Определение. *Борелевской σ-алгеброй* топологического или метрического пространства называется σ-алгебра, порождённая всеми открытыми множествами.

Для обозначения объединения непересекающихся множеств часто будем использовать знак «$\sqcup$». Пересечение множеств A и B будем иногда обозначать через AB. Сигма-алгебры будем обозначать шрифтом `mathscr` ($\mathscr{A}, \mathscr{B}, \mathscr{C}, \ldots$), а пространства элементарных исходов — заглавными греческими буквами.

1.1.3. Вероятность. Аксиоматика Колмогорова

Определение. Пусть $\mathscr{F}$ — σ-алгебра на множестве Ω. Пара $(\Omega, \mathscr{F})$ называется *измеримым пространством.*

Определение. *Вероятностью* называется функция $\mathsf{P}\colon \mathscr{F} \to \mathbb{R}$ со свойствами:
1° Неотрицательность: $\mathsf{P}(A) \geqslant 0$ для любого события $A \in \mathscr{F}$;
2° Нормировка: $\mathsf{P}(\Omega) = 1$;
3° Счётная аддитивность: если $\{A_i\}_{i=1}^{\infty} \subset \mathscr{F}$, и A_i не пересекаются, то $\mathsf{P}\left(\bigcup\limits_{i=1}^{\infty} A_i\right) = \sum\limits_{i=1}^{\infty} \mathsf{P}(A_i)$.

Аксиомы в определении вероятности называют *аксиомами Колмогорова.*

Если на измеримом пространстве задана функция P, то тройка $(\Omega, \mathscr{F}, \mathsf{P})$ называется *вероятностным пространством.* Заметим, что вероятность есть частный случай σ-аддитивной меры на Ω.

Если пространство Ω не более чем счётно, то говорят, что вероятностное пространство *дискретно.* Пусть $\Omega = \{\omega_i\}_{i=1}^{\infty}$. Пусть каждой точке $\omega_i \in \Omega$ приписан вес $p_i \geqslant 0$, такой, что $\sum\limits_{i=1}^{\infty} p_i = 1$. Вероятность определим следующим образом. Пусть $A = \{w_{i_1} w_{i_2}, \ldots\}$. Положим

$$\mathsf{P}(A) := \sum_{i\colon\, \omega_i \in A} p_i. \tag{1}$$

Проверка того, что это действительно вероятность, предоставляется читателю.

Классическое определение вероятности таково: $|\Omega| = N$, все «веса» исходов одинаковы и равны $p = \frac{1}{N}$. Вероятность события $A \subset 2^{\Omega}$ определяется как $\mathsf{P}(A) := \frac{|A|}{N}$.

Задача 1.1. *На экзамене*[1] *по теории вероятностей студенту предлагают три билета. Один он знает хорошо, а два других — не знает. Он выбирает один билет, после чего ему открывают один из оставшихся билетов (которого он не знает). Что выгоднее сделать студенту: поменять свой выбор, или сохранить его?*

`TODO: Лемма о баллотировке`

1.2. Свойства вероятностных мер. Непрерывность

1.2.1. Простейшие свойства вероятности

1° Если $A \subset B$, то $\mathsf{P}(A) \leqslant \mathsf{P}(B)$.

$\square$ Действительно, $\mathsf{P}(B) = \mathsf{P}\big(A \sqcup (B \setminus A)\big) = \mathsf{P}(A) + \underbrace{\mathsf{P}(B \setminus A)}_{\geqslant 0} \geqslant \mathsf{P}(A)$. $\blacksquare$

2° Для любых $A, B \in \mathscr{F}$ имеет место равенство $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(AB)$.

$\square$ Имеем $A \cup B = A\overline{B} \sqcup AB \sqcup \overline{A}B$. Значит, $\mathsf{P}(A \cup B) = \mathsf{P}(A\overline{B}) + \mathsf{P}(AB) + \mathsf{P}(\overline{A}B)$. Заметим, что $\mathsf{P}(A\overline{B}) + \mathsf{P}(AB) = P(A)$, и $\mathsf{P}(AB) + \mathsf{P}(\overline{A}B) = \mathsf{P}(B)$. Подставим вероятности событий $A\overline{B}$ и $\overline{A}B$ и получим требуемое. $\blacksquare$

Теорема 1.1 (Формула включений-исключений). *Для любых $A_1, \ldots, A_n \in \mathscr{F}$ верна формула*

$$\mathsf{P}\left(A_1 \cup \ldots A_n\right) = \sum_{i=1}^{n} \mathsf{P}(A_i) - \sum_{i \leqslant j} \mathsf{P}(A_i A_j) + \sum_{i \leqslant j \leqslant k} \mathsf{P}(A_i A_j A_k) + \ldots + (-1)^{n-1} \mathsf{P}(A_1 \ldots A_n). \tag{2}$$

$\square$ Предыдущее доказательство очевидным образом по индукции обобщается на случай произвольного числа событий. $\blacksquare$

1.2.2. Непрерывность мер и её связь со счётной аддитивностью

Определение. Мера P называется *непрерывной (в нуле)*, если для любой последовательности вложенных событий $B_n \searrow \varnothing$ их вероятность стремится к нулю: $\mathsf{P}(B_n) \to 0$ при $n \to \infty$.

Теорема 1.2. *Счётная аддитивность функции P эквивалентна её конечной аддитивности и непрерывности в нуле.*

$\square$ Выведем непрерывность из счётной аддитивности. Пусть $\{A_i\}_{i=1}^{\infty} \subset \mathscr{F}$, и $A_{n+1} \subset A_n$. Рассмотрим события $B_n := A_n \setminus A_{n+1}$. Они не пересекаются, следовательно

$$\mathsf{P}(A_1) = \underbrace{\mathsf{P}(B_1) + \ldots + \mathsf{P}(B_n)}_{\downarrow \atop \mathsf{P}(A_1)} + \underbrace{\mathsf{P}(A_{n+1})}_{\downarrow \atop 0}, \tag{3}$$

так как вероятность события A_1 конечна, а все члены ряда неотрицательны.

Пусть теперь P аддитивна и непрерывна. Докажем её σ-аддитивность. Пусть $\{B_i\}_{i=1}^{\infty}$ — последовательность попарно непересекающихся событий. Пусть $A_n := \bigcup\limits_{i=n}^{\infty} B_i$ — «остаточные события». Они лежат в $\mathscr{F}$, так как это дополнение к конечному объединению. Очевидно, что $\bigcap\limits_n A_n = \varnothing$ и $A_{n+1} \subset A_n$. По непрерывности $\mathsf{P}(A_n) \to 0$. Значит, $\mathsf{P}\left(\bigcup\limits_{i=1}^{\infty} B_i\right) = \sum\limits_{i=1}^{\infty} \mathsf{P}(B_i)$. $\blacksquare$

Теорема 1.3 (Каратеодори[2]). *Вероятностная мера P на алгебре $\mathscr{A}$ однозначно продолжается на $\sigma\{\mathscr{A}\}$.*

1.3. Дискретные вероятностные пространства (примеры)

1.3.1. Схема Бернулли

В *схеме испытаний Бернулли* $\Omega = \{\omega = (k_1, \ldots, k_n)\}$, где $k_i \in \{0, 1\}$, т. е. проводится n испытаний, каждое из которых может быть либо успехом, либо неудачей. Вероятность успеха равна $p \in (0, 1)$, а неудачи — $q := 1 - p$. Вероятность задаётся корректно, так как строк с m единицами будет $\mathbf{C}_n^m$, а всего по формуле бинома Ньютона

$$\sum_{m=0}^{n} \mathbf{C}_n^m p^m (1-p)^{n-m} = \big(p + (1-p)\big)^n = 1. \tag{4}$$

[1]Нематематическая постановка задачи слегка изменена по сравнению с версией А. В. Булинского.
[2]Эта теорема доказывается в курсе действительного анализа.

1.3.2. Геометрическая вероятность

Модель *геометрического* распределения: монету бросают до первого выпадения герба. Вероятность этого события при каждом бросании равна p. Положим $\mathsf{P}(\{\omega_n\}) := (1-p)^{n-1}p$. Корректность проверяется аналогично.

1.3.3. Гипергеометрическая вероятность

Гипергеометрическое распределение моделируется так: в урне находится M белых и N чёрных шаров. Из неё наугад вынимают n шаров. Найдём вероятность того, что вынуто k белых шаров. Имеем

$$\mathsf{P}(\{\omega_k\}) = \frac{\mathbf{C}_m^k \mathbf{C}_N^{n-k}}{\mathbf{C}_{M+N}^n}. \tag{5}$$

Замечание. Будем считать, что $\mathbf{C}_n^k = 0$ при $k < 0$ или $k > n$.

1.3.4. Распределение Пуассона

В *схеме Пуассона* $\Omega = \{0\} \cup \mathbb{N}$, а $\mathscr{F} = 2^\Omega$. Положим

$$p_n := \mathsf{P}(\{\omega_n\}) = \frac{\lambda^n}{n!}e^{-\lambda}. \tag{6}$$

Задача 1.2. *Доказать, что если $np_n \to \lambda > 0$, то*

$$\mathbf{C}_n^k p_n^k (1-p_n)^{n-k} \to \frac{\lambda^k}{k!}e^{-\lambda},\ n \to \infty. \tag{7}$$

1.4. Условная вероятность и формула Байеса

1.4.1. Понятие условной вероятности

Определение. Система непересекающихся непустых событий $\{A_i\}$ называется *разбиением* Ω, если $\bigcup A_i = \Omega$.

Определение. Пусть $\mathsf{P}(B) > 0$. *Условной вероятностью события A при условии B* называется число

$$\mathsf{P}(A|B) := \frac{\mathsf{P}(AB)}{\mathsf{P}(B)}. \tag{8}$$

Теорема 1.4 (Формула полной вероятности). *Пусть $\{B_i\}$ — разбиение Ω. Тогда для $\forall\, A \in \mathscr{F}$*

$$\mathsf{P}(A) = \sum \mathsf{P}(A|B_n)\mathsf{P}(B_n). \tag{9}$$

□ Действительно, $\mathsf{P}(A) = \mathsf{P}(A\Omega) = \mathsf{P}(\bigcup AB_n) = \sum \mathsf{P}(AB_n) \overset{\text{def}}{=} \sum \mathsf{P}(A|B_n)\mathsf{P}(B_n)$. ■

Задача 1.3. *На экзамене по теории вероятностей N билетов, из них n хороших. Студенты подходят по очереди и вытягивают билеты. Доказать, что вероятность того, что попадётся хороший билет, не зависит от позиции в очереди.*

Решение. Пусть событие B_k — «первые m человек вынули k хороших билетов». Событие A — «достался хороший билет». Тогда

$$\mathsf{P}(A|B_k) = \frac{n-k}{N-m}, \quad \mathsf{P}(B_k) = \frac{\mathbf{C}_n^k \mathbf{C}_{N-n}^{m-k}}{\mathbf{C}_N^m}. \tag{10}$$

Отсюда

$$\mathsf{P}(A) = \sum_{k=0}^{n\wedge m} \frac{n-k}{N-m} \frac{\mathbf{C}_n^k \mathbf{C}_{N-n}^{m-k}}{\mathbf{C}_N^m} \frac{n}{N} = \frac{n}{N} \sum_{k=0}^{n\wedge m} \frac{\mathbf{C}_{n-1}^k \mathbf{C}_{(N-1)-(n-1)}^{m-k}}{\mathbf{C}_N^{m-1}} = \frac{n}{N}, \tag{11}$$

так как сумма в последней формуле есть в точности сумма вероятностей всех исходов в геометрическом распределении. Таким образом, искомая вероятность, вопреки бытующему мнению, зависит только степени подготовленности студентов, т. е. от концентрации хороших билетов. ■

1.4.2. Формула Байеса

Можно все вероятности рассматривать как условные: $\mathsf{P}(A) = \mathsf{P}(A\Omega)$. Если $\{B_i\}$ — разбиение Ω, то

$$\mathsf{P}(B_j|A) = \frac{\mathsf{P}(AB_j)}{\mathsf{P}(A)} = \frac{\mathsf{P}(A|B_j)\mathsf{P}(B_j)}{\sum\limits_{k=1}^{n}\mathsf{P}(A|B_k)\mathsf{P}(B_k)}. \tag{12}$$

К знаменателю дроби здесь была применена формула полной вероятности. Это и есть *формула Байеса*.

`TODO: пример`

1.5. Функции распределения и плотности

1.5.1. Распределения мер

Определение. Пусть P — вероятность на $\mathscr{B}(\mathbb{R})$. *Функцией распределения* меры P называется функция

$$F_{\mathsf{P}}(x) := \mathsf{P}\big((-\infty; x]\big). \tag{13}$$

Теорема 1.5. *Функция распределения F вероятностной меры P обладает следующими свойствами:*
1° *$F(x)$ неубывает;*
2° *$\lim\limits_{x\to-\infty} F(x) = 0$, $\lim\limits_{x\to\infty} F(x) = 1$;*
3° *$F(x)$ непрерывна справа на $\mathbb{R}$.*

□ **1°** Очевидно, что если $x \leqslant y$, то $\mathsf{P}\big((-\infty; x]\big) \leqslant \mathsf{P}\big((-\infty; y]\big)$.

2° Рассмотрим последовательность $x_n \to -\infty$. Тогда события $A_n := (-\infty; x_n] \searrow \varnothing$. По непрерывности меры $F(x_n) \to 0$. Переходя к дополнению к A_n, получим значение предела $F(x)$ на $+\infty$.

3° Фиксируем точку x_0 и рассмотрим последовательность $x_n \to +x_0$. Тогда $A_{n+1} \subset A_n$, и $A_n \searrow A = (-\infty; x_0]$. Снова по непрерывности меры $\mathsf{P}(A_n) \to \mathsf{P}(A)$. ■

Замечание. Теорема даёт исчерпывающее описание всех распределений мер на $\mathbb{R}$, т. е. если функция $F(x)$ обладает свойствами **1°**–**3°**, то существует единственная вероятностная мера P на $\mathscr{B}(\mathbb{R})$, такая, что её функция распределения есть F.

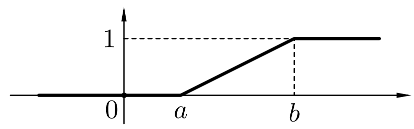
`TODO: пояснение`

1.5.2. Примеры распределений вероятностных мер

Пример 5.1. *Равномерное распределение* на отрезке $[a, b]$ имеет плотность

$$p(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b]; \\ 0, & x \notin [a, b]. \end{cases} \tag{14}$$

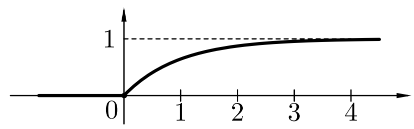
Функция распределения $F(x)$ имеет график



Пример 5.2. *Показательное распределение* с параметром λ имеет плотность

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0; \\ 0, & x < 0. \end{cases} \tag{15}$$

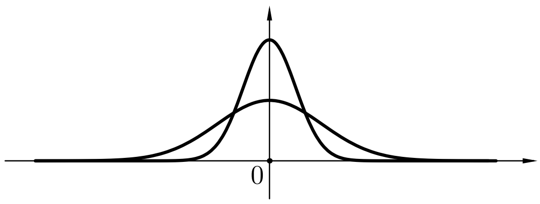
Функция распределения $F(x)$ имеет график



Пример 5.3. *Нормальное (гауссовское) распределение* с параметрами a и σ^2 имеет плотность

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right). \tag{16}$$

График этой плотности при разных значениях σ выглядит примерно так:



Чем меньше величина σ^2, тем явственнее выделяется «пик» в точке $x = a$. Обозначение: $\mathcal{N}(a, \sigma^2)$.

1.6. Независимость событий. Леммы Бореля – Кантелли

1.6.1. Попарная независимость и независимость в совокупности

Определение. Говорят, что событие A *не зависит* от события B, если $\mathsf{P}(A|B) = \mathsf{P}(A)$. По-другому можно переформулировать определение так: события A и B *независимы*, если $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B)$.

Замечание. Вторая форма определения удобнее, поскольку в ней не накладывается ограничений на вероятности событий. Невозможные события, т. е. такие, что их вероятность равна 0, являются независимыми с любыми событиями, ибо если $\mathsf{P}(A) = 0$, то $0 \leqslant \mathsf{P}(AB) \leqslant \mathsf{P}(A) = 0$, а значит, и $\mathsf{P}(AB) = 0$.

Определение. События $A_1, \ldots, A_n$ называются *попарно независимыми*, если $\mathsf{P}(A_iA_j) = \mathsf{P}(A_i)\mathsf{P}(A_j)$ для $\forall\, i \neq j$. События $A_1, \ldots, A_n$ называются *независимыми в совокупности*, если для любого набора индексов $i_1, \ldots, i_k$ имеем $\mathsf{P}(A_{i_1} \cdot \ldots \cdot A_{i_k}) = \mathsf{P}(A_{i_1}) \cdot \ldots \cdot \mathsf{P}(A_{i_k})$.

Замечание. Независимость в совокупности, очевидно, влечёт попарную независимость. Обратное, однако, неверно. Например, пусть $\Omega = \{1, 2, 3, 4\}$ и $\mathsf{P}(\{i\}) = \frac{1}{4}$. Рассмотрим события $A_1 = \{1, 2\}$, $A_2 = \{1, 3\}$, $A_3 = \{1, 4\}$. Имеем $A_iA_j = \{1\}$, значит, $\mathsf{P}(A_iA_j) = \frac{1}{4}$. С другой стороны, $\mathsf{P}(A_1A_2A_3) = \mathsf{P}(\{1\}) = \frac{1}{4} \neq \left(\frac{1}{2}\right)^3$.

Определение. Говорят, что $\{A_i\}_{i=1}^{\infty}$ — *последовательность независимых событий*, если для $\forall\, n \geqslant 2$ события $A_1, \ldots, A_n$ независимы в совокупности.

Определение. *Верхним пределом* системы событий $\{A_k\}$ называется множество

$$\overline{\lim}\, A_k = \limsup A_k := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k. \tag{17}$$

Непосредственно из определения следует, что верхний предел состоит их тех элементов Ω, принадлежащих бесконечному числу A_k.

Определение. *Нижним пределом* системы событий $\{A_k\}$ называется множество

$$\underline{\lim}\, A_k = \liminf A_k := \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k. \tag{18}$$

Непосредственно из определения следует, что нижний предел состоит их тех элементов Ω, принадлежащих всем A_k-м, начиная с некоторого.

1.6.2. Две леммы Бореля – Кантелли

Лемма 1.6 (Бореля – Кантелли №1). *Дана последовательность событий A_k. Пусть ряд $\sum\limits_{k=0}^{\infty} \mathsf{P}(A_k)$ сходится. Тогда вероятность события $A :=$ (произойдёт бесконечно много событий A_k) равна нулю.*

☐ Для $\forall\, n \in \mathbb{N}$ имеем $\mathsf{P}(A) \leqslant \mathsf{P}\left(\bigcup\limits_{k=n}^{\infty} A_k\right) \leqslant \sum\limits_{k=n}^{\infty} \mathsf{P}(A_k) \to 0$, так как это остаток сходящегося ряда. ■

Вторая лемма относится уже к независимым событиям.

Лемма 1.7 (Бореля – Кантелли №2). *Пусть $\{A_i\}_{i=1}^{\infty}$ — последовательность независимых событий. Пусть $\sum \mathsf{P}(A_k) = \infty$. Тогда $\mathsf{P}\left(\overline{\lim}\, A_k\right) = 1$.*

☐ Введём обозначение $B_n := \bigcup\limits_{n=k}^{\infty} A_k$. Заметим, что B_n убывают. Тогда по свойству непрерывности меры

$$\mathsf{P}(B_n) = \lim_{N \to \infty} \mathsf{P}\left(\bigcup_{k=n}^{N} A_k\right). \tag{19}$$

Рассмотрим дополнение:

$$\mathsf{P}\Big(\overline{\bigcup_{k=n}^{N} A_k}\Big) = \mathsf{P}\Big(\bigcap_{k=n}^{N} \overline{A}_k\Big) \overset{!}{=} \prod_{k=n}^{N} \mathsf{P}\left(\overline{A}_k\right) = \prod_{k=n}^{N} \big(1 - \mathsf{P}(A_k)\big) \overset{!!}{\leqslant} \prod_{k=n}^{N} \exp\big(-\mathsf{P}(A_k)\big) = \exp\Big(-\sum_{k=n}^{N} A_k\Big) \to 0,$$

так как ряд расходится. Переход «!» обоснован независимостью A_k, а «!!» — неравенством Бернулли. Переходя к пределу при $N \to \infty$, получаем, что $\mathsf{P}\left(\overline{\lim} A_k\right) = 1$. ■

Задача 1.4 (**). *Доказать вторую лемму Бореля – Кантелли для последовательности попарно независимых событий.*

Задача 1.5 (**). *Назовём событие A «атомом» вероятностного пространства, если оно обладает следующим свойством: если $B \subset A$, то либо $\mathsf{P}(B) = 0$, либо $\mathsf{P}(B) = \mathsf{P}(A)$. Доказать, что в любое вероятностное пространстве разбивается на не более чем счётное число атомов и «непрерывную» часть $\mathcal{C}$, обладающую следующим свойством: если $C \in \mathcal{C}$ и $\mathsf{P}(C) = p$, то для $\forall\, q \in [0, p]$ найдётся $C_q \in \mathcal{C}\colon \mathsf{P}(C_q) = q$.*

2. Случайные величины

2.1. Понятие случайного элемента

2.1.1. Измеримые отображения

Определение. Пусть $(\Omega, \mathscr{F})$ и $(\Theta, \mathscr{B})$ — измеримые пространства. Отображение $\xi\colon \Omega \to \Theta$ называется *$\mathscr{F}|\mathscr{B}$-измеримым*, если для $\forall\, B \in \mathscr{B}$ имеем $\xi^{-1}(B) \in \mathscr{F}$, т. е. прообраз измеримого множества измерим. В теории вероятностей измеримые отображения называются *случайными элементами*.

В дальнейшем, если из контекста ясно, какие алгебры заданы на наших пространствах, будем опускать префикс «$\mathscr{F}|\mathscr{B}$» и говорить просто «измеримая функция».

Определение случайной величины напоминает определение непрерывной функции в топологическом пространстве. Напомним, что *топологическим пространством* называется множество $\mathcal{X}$, на котором задана *топология*, т. е. отмечен класс подмножеств, называемых *открытыми*, такой, что объединение любого набора открытых множеств открыто, и пересечение конечного набора открытых множеств открыто. Открытыми считают также $\varnothing$ и все множество $\mathcal{X}$. *Непрерывным* отображением называется отображение, при котором прообраз открытого множества открыт. Иными словами, непрерывное отображение $f\colon \mathcal{X} \to \mathcal{Y}$ должно быль измеримо относительно топологий на $\mathcal{X}$ и $\mathcal{Y}$.

Если в определении случайной величины $\mathscr{B} = \mathscr{B}(\mathbb{R})$, то $\mathscr{F}|\mathscr{B}$-измеримые отображения называют *действительными случайными величинами.*

Определение. *Борелевским* называется отображение, при котором прообраз борелевского множества борелевский.

2.1.2. Свойства измеримых функций

Чтобы не загромождать формулировки следующих утверждений, оговоримся сразу о том, что мы будем рассматривать измеримые пространства $(\Omega, \mathscr{F})$ и $(\Theta, \mathscr{B})$ и случайную величину $\xi\colon \Omega \to \Theta$.

Теорема 2.1. *Пусть $\mathscr{B} = \sigma\left\{\mathcal{M}\right\}$. Чтобы ξ была измеримой функцией, необходимо и достаточно, чтобы для $\forall\, C \in \mathcal{M}$ было выполнено $\xi^{-1}(C) \in \mathscr{F}$. Иными словами, свойство измеримости можно проверять только на порождающих элементах σ-алгебры.*

□ Необходимость очевидна. Достаточность: рассмотрим класс множеств $\mathcal{D} := \left\{D \subset \Theta\colon \xi^{-1}(D) \in \mathscr{F}\right\}$, т. е. класс тех множеств, для которых свойство измеримости выполняется. Оно является σ-алгеброй, так как операция взятия прообраза сохраняет все теоретико-множественные операции, т. е. $\xi^{-1}\left(\bigcup B_n\right) = \bigcup \xi^{-1}(B_n)$ и т. д. По условию $\mathcal{M} \subset \mathcal{D}$, а значит, и $\sigma\left\{M\right\} \subset \mathcal{D}$. Но $\sigma\left\{\mathcal{M}\right\} = \mathscr{B}$, поэтому $\mathcal{D}$ совпадает с $\mathscr{B}$. ■

Далее $\xi\colon \Omega \to \mathbb{R}$ — действительная случайная величина, а $\mathscr{B} = \mathscr{B}(\mathbb{R})$.

Теорема 2.2. *Чтобы ξ была случайной величиной, необходимо и достаточно измеримости прообразов лучей $(-\infty; x]$, т. е. для $\forall\, x \in \mathbb{R}$ множество $\{\omega\colon \xi(\omega) \leqslant x\}$ должно содержаться в $\mathscr{F}$.*

□ Заметим, что $\mathscr{B}(\mathbb{R})$ порождается лучами $(-\infty, x]$, так как открытое множество на прямой есть объединение не более чем счётного множества интервалов (возможно, бесконечных), а интервал из лучей соорудить несложно. Остаётся применить предыдущую теорему. ■

Аналогичное утверждение справедливо для открытых лучей вида $(-\infty; x)$.

Теорема 2.3. *Пусть $\xi_n\colon \Omega \to \mathbb{R}$ — последовательность случайных величин. Тогда:*
1° $\sup \xi_n$, $\inf \xi_n$, $\overline{\lim}\, \xi_n$, $\underline{\lim}\, \xi_n$ *будут случайными величинами;*
2° *Если $\xi_n \to \xi$, то ξ — случайная величина.*

$\square$ **1°** В самом деле, имеем

$$\{\omega\colon \sup \xi_n(\omega) \leqslant x\} = \bigcap_{n=1}^{\infty} \underbrace{\{\xi_n(\omega) \leqslant x\}}_{\text{событие}}. \tag{1}$$

Счётное пересечение событий есть событие, поэтому $\sup \xi_n$ измерима. Аналогично $\inf \xi_n$ измерима. Докажем для верхних и нижних пределов. Имеем $\overline{\lim}\, \xi_n = \limsup \xi_n = \inf\limits_n \sup\limits_{k \geqslant n} \xi_k$. По доказанному функция $\eta_n := \sup\limits_{k \geqslant n} \xi_k$ будет случайной величиной. Точно также и $\inf \eta_n$ будет случайной величиной. Для нижнего предела имеем $\underline{\lim}\, \xi_n = \liminf \xi_n = \sup\limits_n \inf\limits_{k \geqslant n} \xi_k$, далее рассуждения аналогичны.

2° Существование предела равносильно совпадению верхнего и нижнего пределов. А по доказанному $\overline{\lim}$ и $\underline{\lim}$ измеримы, значит, и $\lim$ будет измерим. $\blacksquare$

Задача 2.1. *Рассмотрим последовательность случайных величин $\xi_n\colon \Omega \to \mathcal{S}$, где $\mathcal{S}$ — метрическое пространство. Доказать, что если $\xi_n \to \xi$, то $\xi \in \mathscr{F}|\mathscr{B}(\mathcal{S})$.*

Лемма 2.4. *Рассмотрим измеримые пространства $(\Omega, \mathscr{F})$, $(\Theta, \mathscr{B})$, $(\Lambda, \mathscr{A})$ и случайные величины $\xi \in \mathscr{F}|\mathscr{B}$ и $\eta \in \mathscr{B}|\mathscr{A}$. Тогда $\eta \circ \xi \in \mathscr{F}|\mathscr{A}$.*

$\square$ Очевидно: η измерима, поэтому $\eta^{-1}(A) \in \mathscr{B}$, а так как ξ измерима, то $\xi^{-1}\underbrace{\big(\eta^{-1}(A)\big)}_{\in \mathscr{B}} \in \mathscr{F}$. $\blacksquare$

Следствие 2.1. *Пусть $\xi\colon \Omega \to \mathbb{R}$ — случайная величина, а $\varphi\colon \mathbb{R} \to \mathbb{R}$ — борелевская функция. Тогда $\varphi \circ \xi$ — случайная величина.*

$\square$ Это частный случай предыдущего утверждения: $\varphi \in \mathscr{B}(\mathbb{R})|\mathscr{B}(\mathbb{R})$. $\blacksquare$

Следствие 2.2. *Пусть $\xi\colon \Omega \to \mathbb{R}$ — случайная величина, а φ — непрерывная функция. Тогда $\varphi \circ \xi$ — случайная величина.*

$\square$ Это опять-таки частный случай первого следствия: непрерывная функция, очевидно, борелевская (прообраз открытого множества открыт). $\blacksquare$

Утверждение 2.5. *Пусть ξ, η — действительные случайные величины. Тогда $\xi \pm \eta$, $\xi \cdot \eta$ и $\frac{\xi}{\eta}$ (если $\eta \neq 0$) также будут случайными величинами.*

$\square$ Доказывать можно тремя способами. Первый (короткий): сослаться на следствие предыдущей леммы, сказав, что арифметические операции на $\mathbb{R}$ непрерывны. Второй (подлиннее): открыть учебник М. И. Ульянова и П. Л. Дьяченко *«Мера и интеграл»* на странице 44 и найти там полное доказательство этого утверждения. Третий: немного подумать и доказать утверждение самостоятельно. $\blacksquare$

2.1.3. Примеры случайных величин

Пример 1.1. Пусть $\Omega = [0,1]$, $\mathscr{F} = \mathscr{B}(\Omega)$, а P — мера Лебега на Ω. Любая непрерывная функция $f\colon \Omega \to \mathbb{R}$ будет случайной величиной.

Пример 1.2. Рассмотрим схему Бернулли. Случайной величины будет, например, количество успехов при n испытаниях, т. е. $S_n(\omega) := \sum\limits_{i=1}^{n} k_i$, где $\omega = (k_1, \ldots, k_n)$, $k_n \in \{0, 1\}$.

Пример 1.3. Рассмотрим дискретное вероятностное пространство $(\Omega, \mathscr{F}, \mathsf{P})$ в котором $\mathscr{F} = 2^\Omega$. На таком пространстве любая функция $\xi\colon \Omega \to \mathbb{R}$ будет случайной величиной, так как прообраз любого множества содержится в 2^Ω.

2.2. Распределения случайных элементов. Пополнение вероятностного пространства

2.2.1. Вероятностные меры и распределения случайных величин

Определение. *Функцией распределения* (действительной) случайной величины ξ называется функция

$$F_\xi(x) := \mathsf{P}\big(\{\omega\colon \xi(\omega) \leqslant x\}\big). \tag{2}$$

В дальнейшем для сокращения записи будем в подобных формулах писать просто «$\mathsf{P}(\xi \leqslant x)$».

Пусть $(\Theta, \mathscr{B})$ — измеримое пространство. Покажем, что изучение вероятностных мер на $(\Theta, \mathscr{B})$ и случайных элементов со значениями в Θ — это одно и то же. В самом деле, пусть нам дана мера Q. Найдём такое вероятностное пространство, что распределение некоторого случайного элемента совпадает с Q. Положим

$$\Omega := \Theta, \quad \mathscr{F} := \mathscr{B}, \quad \xi := \mathrm{id}\colon \Theta \to \Theta, \quad \mathsf{P} := \mathsf{Q}. \tag{3}$$

Очевидно, что $F_\xi \equiv F_{\mathsf{Q}}$ (вспомните определение функции распределения меры!).

Определение. *Индикатором*[3] события $A \subset \Omega$ называется функция

$$\mathbb{I}_A(\omega) := \begin{cases} 1, & \omega \in A; \\ 0, & \omega \notin A. \end{cases} \tag{4}$$

Задача 2.2. *Рассмотрим измеримое пространство* $(\Omega, \mathscr{F})$*. Доказать, что* $\mathbb{I}_A \in \mathscr{F}|\mathscr{B}(\mathbb{R}) \Leftrightarrow A \in \mathscr{F}$.

2.2.2. Пополнение вероятностного пространства

В теории вероятностей всегда удобно считать вероятностное пространство $(\Omega, \mathscr{F}, \mathsf{P})$ *пополненным*. Поясним, что это означает. Рассмотрим класс $\mathcal{N}$ *нулевых* множеств, т. е. все события $A \in \mathscr{F}$ такие, что $\mathsf{P}(A) = 0$ и все их подмножества $B \subset A$ (множество B уже не обязательно событие!), т. е.

$$\mathcal{N} = \{B \colon B \subset A,\ A \in \mathscr{F},\ \mathsf{P}(A) = 0\}. \tag{5}$$

Пусть $\overline{\mathscr{F}} := \sigma\{\mathscr{F}, \mathcal{N}\}$.

Утверждение 2.6. $\overline{\mathscr{F}}$ *состоит из множеств вида* $A \cup C$*, где* $A \in \mathscr{F}$*, а* $C \in \mathcal{N}$*, и только из них.*

Доказательство этого утверждения предоставляется читателю в качестве упражнения (включения $\supset$ и $\subset$ легко устанавливаются).

Теперь, опираясь на это утверждение, определим на пространстве $(\Omega, \overline{\mathscr{F}})$ вероятностную меру $\overline{\mathsf{P}}$ следующим естественным образом:

$$\overline{\mathsf{P}}(A \cup C) := \mathsf{P}(A), \text{ где } A \in \mathscr{F},\ C \in \mathcal{N}. \tag{6}$$

Все аксиомы вероятности, очевидно, выполняются. Таким образом, получено вероятностное пространство $(\Omega, \overline{\mathscr{F}}, \overline{\mathsf{P}})$. Оно и называется *пополнением* пространства $(\Omega, \mathscr{F}, \mathsf{P})$.

Теперь поясним, собственно, зачем мы его построили. Сначала введём новый вид сходимости.

Определение. Говорят, что последовательность случайных величин ξ_n сходится к ξ *почти наверное (почти всюду)*, если $\xi_n \to \xi$ поточечно за исключением множества меры нуль, т. е. $\xi_n \to \xi$ на множестве $\Omega_0 \colon \mathsf{P}(\Omega_0) = 1$. Обозначение: $\xi_n \xrightarrow{\text{п.н.}} \xi$.

Утверждение 2.7. *Пусть действительные случайные величины* ξ_n *сходятся к* ξ *почти наверное. Тогда* ξ *будет случайной величиной на пополнении нашего пространства по мере* P.

Доказательство[4] этого утверждения мы также предоставляем читателю.

В дальнейшем, когда речь пойдёт о сходимости случайных величин и их математических ожиданий, всегда под словом «сходимость», как правило, будем иметь в виду сходимость почти наверное и считать пространство пополненным.

2.3. Независимость случайных величин

2.3.1. Независимые алгебры и случайные величины

Теперь дадим одно из самых важных в теории вероятностей определений.

Определение. Алгебры (или σ-алгебры) $\mathscr{A}_1, \ldots, \mathscr{A}_n \subset \mathscr{F}$ называются *независимыми*, если

$$\mathsf{P}(A_1 \cdot \ldots \cdot A_n) = \mathsf{P}(A_1) \cdot \ldots \cdot \mathsf{P}(A_n) \quad \forall\, A_i \in \mathscr{A}_i. \tag{7}$$

Определение. Пусть $(\Omega, \mathscr{F})$ и $(\Theta, \mathscr{B})$ — измеримые пространства, а $\xi \colon \Omega \to \Theta$ — случайная величина. Сигма-алгеброй, *порождённой* величиной ξ, называется множество

$$\sigma\{\xi\} := \left\{\xi^{-1}(B),\ B \in \mathscr{B}\right\}. \tag{8}$$

Определение. Величины $\xi_1, \ldots, \xi_n$ называются *независимыми*, если $\sigma\{\xi_1\}, \ldots, \sigma\{\xi_n\}$ независимы.

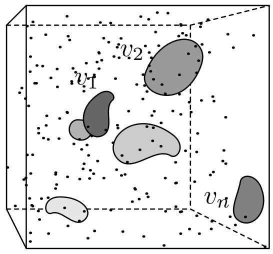
Иными словами, независимость означает, что

$$\mathsf{P}(\xi_1 \in B_1, \ldots, \xi_n \in B_n) = \bigcap_{i=1}^{n} \{\omega \colon \xi_i(\omega) \in B_i\} = \prod_{i=1}^{n} \mathsf{P}(\xi_i \in B_i). \tag{9}$$

Аналогично событиям определяется последовательность независимых случайных величин.

[3] Иногда используется термин *«характеристическая функция»*, однако в теории вероятностей он имеет совсем другой смысл.
[4] Его можно прочесть в книге А. В. Булинского и А. Н. Ширяева *«Теория случайных процессов»*.

2.3.2. Теорема о булочках с изюмом



Теперь докажем одну из теорем о сходимости последовательности случайных величин. Рассмотрим куб V с ребром M. Пусть в нём выделено n непересекающихся борелевских подмножеств v_i. Будем бросать точки в куб и считать, сколько их попадает в каждое подмножество. Вероятность попадания k-й точки в v_i равна $\mathsf{P}(\xi_k \in v_i) = \frac{|v_i|}{|V|}$ (распределение равномерное, а модуль обозначает меру Лебега). Пусть Y_i — число частиц, попавших в v_i. Это случайная величина, так как она равна сумме N индикаторов борелевских множеств (N — общее число частиц):

$$Y_i := \sum_{k=1}^{N} \mathbb{I}_{v_i}\big(\xi_k(\omega)\big). \tag{10}$$

Теорема 2.8. *Пусть длина ребра куба стремится к $+\infty$, а плотность частиц стремится к некоторому (конечному) числу: $\frac{N}{|V|} \to \lambda > 0$. Тогда*

$$\mathsf{P}\left(Y_1 = m_1, \ldots, Y_n = m_n\right) \to \frac{(\lambda|v_1|)^{m_1}}{m_1!}\exp\left(-\lambda|v_1|\right) \cdot \ldots \cdot \frac{(\lambda|v_n|)^{m_n}}{m_n!}\exp\left(-\lambda|v_n|\right) = \prod_{i=1}^{n} \mathsf{P}(z_i = m_i), \tag{11}$$

где величины z_i имеет пуассоновское распределение с параметром $\lambda|v_i|$.

□ Пусть m_0 — число частиц, не попавших ни в одно из v_i, а p_i — вероятность попадания в v_i.

$$m_0 := N - \sum_{i=1}^{n} m_i, \quad p_i := \frac{|v_i|}{|V|}, \quad p_0 := \frac{|V| - \sum\limits_{i=1}^{n}|v_i|}{|V|}. \tag{12}$$

Тогда искомая вероятность равна

$$\frac{N!}{m_0! \cdot m_1! \cdot \ldots \cdot m_n!} \cdot p_0^{m_0} \cdot p_1^{m_1} \cdot \ldots \cdot p_n^{m_n}. \tag{13}$$

Заметим, что

$$\frac{N!}{m_0!} = \frac{N!}{\left(N - \sum\limits_{i=1}^{n} m_i\right)!} \to N^{\sum\limits_{i=1}^{n} m_i}. \tag{14}$$

Тогда

$$N^{m_i} p_i^{m_i} = N^{m_i}\left(\frac{|v_i|}{|V|}\right)^{m_i} = \left(\frac{N|v_i|}{|V|}\right)^{m_i} \to (\lambda|v_i|)^{m_i}. \tag{15}$$

В следующей формуле все операции суммирования — по i от 1 до n. Имеем

$$p_0^{m_0} = \left(\frac{|V| - \sum|v_i|}{|V|}\right)^{N - \sum m_i} \to \left(1 - \frac{\sum|v_i|}{|V|}\right)^{N} = \left(1 - \frac{\sum|v_i|}{|V|}\right)^{\frac{N}{|V|}|V|} \to \exp\left(-\lambda\sum|v_i|\right), \tag{16}$$

так как показатель степени стремится к $\lambda|V|$. Теорема доказана. ■

Следствие 2.3. *Если $np_n \to \lambda > 0$ при $n \to \infty$, то*

$$\mathbf{C}_n^m p_n^m (1-p_n)^{n-m} \to \frac{\lambda^m}{m!}e^{-\lambda}. \tag{17}$$

2.4. Теорема о монотонных классах и её следствия

2.4.1. π-системы и λ-системы множеств

Определение. Система $\mathcal{M}$ подмножеств S называется π-*системой*, если $A, B \in \mathcal{M} \Rightarrow A \cap B \in \mathcal{M}$.

Замечание. Если $\mathcal{M}$ — π-система, то можно считать, что $S \in \mathcal{M}$.

Определение. Система $\mathcal{M}$ называется λ-*системой*, если:
1° $S \in \mathcal{M}$;
2° $A \subset B$ и $A, B \in \mathcal{M} \Rightarrow B \smallsetminus A \in \mathcal{M}$;
3° $A_n \nearrow A$ и $A_n \in \mathcal{M} \Rightarrow A \in \mathcal{M}$.

Очевидно, что пересечение π-систем будет π-системой. То же самое верно и для λ-систем.

Будем рассматривать $\pi\{\mathcal{K}\}$ и $\lambda\{\mathcal{K}\}$ — наименьшие π-системы и λ-системы, порождённые системой $\mathcal{K}$. Их существование доказывается аналогично соответствующему утверждению для σ-алгебр.

Лемма 2.9. *Множество $\mathcal{A}$ является σ-алгеброй тогда и только тогда, когда оно является π-λ-системой.*

$\square$ Честно проверяем аксиомы: замкнутость относительно пересечения уже есть, счётная аддитивность легко выводится из свойства **3°** определения λ-системы. Всё остальное совсем очевидно. $\blacksquare$

Теорема 2.10 (О монотонных классах). *Пусть π-система $\mathcal{M}$ содержится в λ-системе $\mathcal{D}$. Тогда σ-алгебра, натянутая на $\mathcal{M}$, совпадает с λ-системой, порождённой $\mathcal{M}$, т. е. $\sigma\{\mathcal{M}\} = \lambda\{\mathcal{M}\} \subset \mathcal{D}$.*

$\square$ Без потери общности можно считать, что $\mathcal{D} = \lambda\{\mathcal{M}\}$. Покажем, что $\mathcal{D}$ замкнуто относительно пересечения, откуда по предыдущей лемме и будет следовать, что $\mathcal{D}$ есть σ-алгебра, так как это будет π-λ-система.

Рассмотрим произвольные множества $A, B \in \mathcal{D}$ и докажем, что $A \cap B \in \mathcal{D}$. Пусть $B \in \mathcal{M}$. Рассмотрим класс $\mathcal{F}_B$ таких множеств A, для которых утверждение верно, т. е. $\mathcal{F}_B := \{A \in \mathcal{D} \colon A \cap B \in \mathcal{D}\}$. Легко видеть, что этот класс является λ-системой, а значит, он совпадает с $\mathcal{D}$. Аналогично, фиксировав множество A, рассмотрим класс $\mathcal{G}_A := \{B \in D \colon A \cap B \in \mathcal{D}\}$. Это также будет λ-система, поэтому $\mathcal{G}_A = \mathcal{D}$ для $\forall\, A \in \mathcal{D}$. Значит, $A \cap B \in \mathcal{D}$ для любых $A, B \in \mathcal{D}$. $\blacksquare$

2.4.2. Следствия

Теорема 2.11. *Рассмотрим вероятностное пространство $(\Omega, \mathscr{F}, \mathsf{P})$. Пусть $\mathscr{A} = \sigma\{\mathcal{K}\}$, где $\mathcal{K}$ — π-система подмножеств Ω, а Q — некоторая вероятностная мера. Пусть $\mathsf{P} = \mathsf{Q}$ на $\mathcal{K}$. Тогда $\mathsf{P} = \mathsf{Q}$ на $\mathscr{A}$.*

$\square$ Рассмотрим совокупность $\mathcal{M}$ множеств $A \in \mathscr{F}$ таких, что $\mathsf{P}(A) = \mathsf{Q}(A)$. Она является λ-системой, поэтому если $A_n \nearrow A$, где $A_n \in \mathcal{M}$, то и $A \in \mathcal{M}$. В силу непрерывности P имеем $\mathsf{P}(A_n) \to \mathsf{P}(A)$. Но $\mathsf{P}(A_n) = \mathsf{Q}(A_n)$ по определению класса $\mathcal{M}$. Значит, $\mathsf{P}(A) = \mathsf{Q}(A)$. $\blacksquare$

Следствие 2.4. *Пусть P, Q — меры на $\mathscr{B}(\mathbb{R})$, а $F(x)$ и $G(x)$ — их функции распределения. Пусть $F = G$. Тогда меры P и Q совпадают на $\mathscr{B}(\mathbb{R})$.*

$\square$ Действительно, очевидно, что класс $\mathcal{M} := \big\{(-\infty; x],\ x \in \mathbb{R}\big\}$ есть π-система. Кроме того, $\sigma\{\mathcal{M}\} = \mathscr{B}(\mathbb{R})$. Остается воспользоваться только что доказанной теоремой. $\blacksquare$

Задача 2.3. *Пусть $\mathscr{A} = \sigma\{\mathcal{K}\}$, где $\mathcal{K}$ — алгебра. Тогда для $\forall\, \varepsilon > 0$ и для $\forall\, A \in \mathscr{A}$ найдётся множество $A_\varepsilon \in \mathcal{K}$ такое, что $\mathsf{P}(A \triangle A_\varepsilon) < \varepsilon$.*

Решение. Нужно всего лишь вспомнить теорему о продолжении меры (Каратеодори). $\blacksquare$

Покажем теперь, что свойство независимости σ-алгебр можно проверять только на порождающих элементах, если они образуют π-системы.

Теорема 2.12. *Пусть $\mathscr{A}_1, \ldots, \mathscr{A}_n \subset \mathscr{F}$, где $\mathscr{A}_i = \sigma\{\mathcal{M}_i\}$, а $\mathcal{M}_i$ — π-системы. Пусть также для событий $A_i \in \mathcal{M}_i$ справедливо равенство*

$$\mathsf{P}(A_1 \cdot \ldots \cdot A_n) = \mathsf{P}(A_1) \cdot \ldots \cdot \mathsf{P}(A_n). \tag{18}$$

Тогда оно верно и для любых событий $A_i \in \mathscr{A}_i$.

$\square$ Зафиксируем $A_i \in \mathcal{M}_i$ для $i = \overline{2, n}$. Рассмотрим класс событий $\mathcal{K}_1$ таких, что (18) верно для любого $A_1 \in \mathcal{K}_1$. Легко видеть, что $\mathcal{K}_1$ будет λ-системой. Применяя теорему о монотонных классах, получаем, что формула верна для любого $A_1 \in \mathscr{A}_1$. Далее, аналогично зафиксируем все множества, кроме A_2, и т. д. Тем самым утверждение будет доказано для любых событий из $\mathscr{A}_1, \ldots, \mathscr{A}_n$. $\blacksquare$

Следствие 2.5. *Независимость действительных случайных величин $\xi_1, \ldots, \xi_n$ равносильна тому, что для $\forall\, x_1, \ldots, x_n \in \mathbb{R}$ совместная вероятность равна произведению вероятностей:*

$$\mathsf{P}(\xi_1 \leqslant x_1, \ldots, \xi_n \leqslant x_n) = \prod_{i=1}^{n} \mathsf{P}(\xi_i \leqslant x_i). \tag{19}$$

Лемма 2.13. *Пусть $\xi_1, \ldots, \xi_n$ — независимые случайные величины, а $f_1, \ldots, f_n \colon \mathbb{R} \to \mathbb{R}$ — борелевские функции. Тогда функции $f_1(\xi_1), \ldots, f_n(\xi_n)$ также независимы.*

$\square$ В самом деле,

$$\sigma\{f_i(\xi_i)\} = \left\{\big(f_i(\xi_i)\big)^{-1}(B)\right\} = \left\{\xi_i^{-1}\big(\underbrace{f_i^{-1}(B)}_{\in \mathscr{B}(\mathbb{R})}\big)\right\} \subseteq \sigma\{\xi_i\}. \tag{20}$$

Значит, определение независимости будет выполнено и для σ-алгебр $\sigma\{f_i(\xi_i)\}$. $\blacksquare$

Теперь обобщим утверждение леммы. Покажем, что борелевские функции от **непересекающихся** наборов независимых случайных величин независимы.

Теорема 2.14. *Рассмотрим независимые действительны случайные величины $\xi_1, \ldots, \xi_n$ и некоторые борелевские функции $f_1(x_1, \ldots, x_{k_1}), \ldots, f_s(x_{k_m}, \ldots, x_n)$. Тогда композиции $f(\xi_1, \ldots, \xi_{k_1}), \ldots, f(\xi_{k_m}, \ldots, \xi_n)$ также независимы.*

□ Легко видеть, что свойство независимости справедливо для «параллелепипедов», т. е. множеств вида $B = C_1 \times \ldots \times C_k$, где $C_i \in \mathscr{B}(\mathbb{R})$, для которых $\big\{(x_1, \ldots, x_k) \in B\big\} = \big\{x_1 \in C_1, \ldots, x_k \in C_k\big\}$. Параллелепипеды образуют π-систему. Остаётся применить первое следствие из теоремы о монотонных классах (теорему 2.11). ■

2.5. Построение случайных величин с заданным распределением

2.5.1. Последовательности равномерно распределённых величин

Рассмотрим независимые случайные величины ξ_k, принимающие только значения 0 и 1 с одинаковой вероятностью $p = \frac{1}{2}$. Рассмотрим также величину

$$\xi := \sum_{k=1}^{\infty} \frac{\xi_k(\omega)}{2^k}. \tag{21}$$

Лемма 2.15. *Величина ξ имеет равномерное распределение на* $[0, 1]$.

□ Заметим сначала, что ξ будет случайной величиной, так как это предел суммы сходящегося ряда. Пусть $x \in [0, 1]$ имеет двоичное разложение

$$x = \sum_{k=1}^{\infty} \frac{x_k}{2^k}. \tag{22}$$

Найдём функцию распределения ξ:

$$\begin{aligned}
\mathsf{P}(\xi < x) = \mathsf{P}\big(\{\xi_1 < x_1\} \cup \{\xi_1 = x_1, \xi_2 < x_2\} \cup \{\xi_1 = x_1, \xi_2 = x_2, \xi_3 < x_3\} \cup \ldots\big) &\overset{!}{=} \\
\overset{!}{=} \sum_{k=1}^{\infty} \mathsf{P}\left(\xi_1 = x_1, \ldots, \xi_{k-1} = x_{k-1}, \xi_k < x_k\right) &\overset{!!}{=} \\
\overset{!!}{=} \sum_{k=1}^{\infty} \big(\mathsf{P}(\xi_1 = x_1) \cdot \ldots \cdot \mathsf{P}\left(\xi_{k-1} = x_{k-1}\right) \cdot \mathsf{P}\left(\xi_k < x_k\right)\big) &= \sum_{k=1}^{\infty} \frac{x_k}{2^k} = x = \mathsf{P}\big([0, x]\big),
\end{aligned}$$

что и означает равномерную распределённость. Переход «!» обусловлен тем, что рассматриваемые события не пересекаются, а равенство «!!» — независимостью величин ξ_k. ■

Теперь построим не одну равномерно распределённую величину, а сразу много. Для этого запишем наши величины ξ_k в виде таблицы, заполняя её по диагоналям:

$$\begin{array}{lllll}
\xi_1 & \xi_2 & \xi_4 & \xi_7 & \ldots \\
\xi_3 & \xi_5 & \xi_8 & \ldots & \\
\xi_6 & \xi_9 & \ldots & & \\
\xi_{10} & \ldots & & &
\end{array} \tag{23}$$

Перенумеруем ξ_k индексами по строкам и столбцам, получим величины ξ_{ij}. Получатся последовательности $\{\xi_{ij}\}_{j=1}^{\infty}$ (при фиксированном i) независимых случайных величин, так как они являются подпоследовательностями в последовательности независимых величин $\{\xi_k\}$. Рассмотрим теперь величины

$$\eta_n := \sum_{k=1}^{\infty} \frac{\xi_{nk}(\omega)}{2^k}. \tag{24}$$

По только что доказанной лемме они имеют равномерное распределение на $[0, 1]$. Остаётся показать их независимость. Рассмотрим конечные суммы

$$\theta_{nm} := \sum_{k=1}^{m} \frac{\xi_{nk}(\omega)}{2^k} \to \eta_n,\ m \to \infty. \tag{25}$$

Они будут независимыми как (борелевские) функции от непересекающихся наборов случайных величин. Довести доказательство до конца, сославшись на непрерывность вероятности в нуле и равномерную распределённость величин η_n, мы предоставляем читателю.

2.5.2. Построение произвольной последовательности вероятностных мер

Определение. *Закон распределения* случайной величины $\xi\colon \Omega \to \mathbb{R}$ есть

$$\mathrm{Law}(\xi) = \mathsf{P}_\xi(B) := \mathsf{P}\big(\xi^{-1}(B)\big). \tag{26}$$

Теперь определим нечто вроде «обратной функции» для функций распределения.

Определение. Пусть $x \in (0,1)$. Положим

$$F^{-1}(y) := \inf\left\{x\colon F(x) \geqslant y\right\}, \tag{27}$$

то есть из всевозможных прообразов берём их нижнюю грань.

Лемма 2.16. *Для* $\forall\, z \in \mathbb{R}$ *имеет место равенство множеств*

$$\big\{x \in (0,1)\colon F^{-1}(x) \leqslant z\big\} = \left\{x \in (0,1)\colon x \leqslant F(z)\right\}. \tag{28}$$

□ Если $x \leqslant F(z)$, то по определению $F^{-1}(x) \leqslant z$. Таким образом, включение «$\supseteq$» установлено. Наоборот, если $F^{-1}(x) \leqslant z$, то $F\big(F^{-1}(x)\big) \leqslant F(z)$, так как F неубывает. Покажем, что $x \leqslant F\big(F^{-1}(x)\big)$ для $\forall\, x \in (0,1)$. Пусть $F^{-1}(x) = a$. Рассмотрим последовательность $y_n \searrow a$. Тогда $F(y_n) \geqslant a$ в силу неубывания F. Функция F непрерывна справа, поэтому $F(y_n) \to F(a)$. Переход к пределу даёт $F\big(F^{-1}(a)\big) \geqslant a$. ■

Следствие 2.6. *Пусть ν имеет равномерное распределение на $[0,1]$, а F — функция распределения некоторой меры Q на $\mathscr{B}(\mathbb{R})$. Тогда $\mathrm{Law}\big(F^{-1}(\nu)\big) = \mathsf{Q}$.*

□ Действительно, $\mathsf{P}\big(F^{-1}(\nu) \leqslant z\big) \overset{!}{=} \mathsf{P}\big(\omega\colon \nu(\omega) \leqslant F(z)\big) = F(z)$, так как распределение равномерно. Равенство «!» обусловлено леммой. ■

Таким образом, имея равномерно распределённую величину ξ, можно построить случайную величину с произвольной функцией распределения.

А вот теперь настало время прояснить, зачем всё это нам потребовалось.

Теорема 2.17. *Пусть задана последовательность $\{\mathsf{Q}_n\}$ вероятностных мер на $\mathscr{B}(\mathbb{R})$, т. е. последовательность $\{F_n\}$ их функций распределения. Тогда на некотором вероятностном пространстве существуют независимые случайные величины ξ_n с функциями распределения F_n, т. е. $\mathrm{Law}(\xi_n) = \mathsf{Q}_n$.*

□ Построим последовательность независимых случайных величин $\{\nu_n\}$, равномерно распределённых на $[0,1]$. Положим $\xi_n := F_n^{-1}(\nu_n)$. Величины ξ_n будут независимыми как борелевские функции от независимых величин. По следствию из леммы их законы распределения суть в точности Q_n. ■

3. Математическое ожидание

3.1. Интеграл Лебега по вероятностной мере

3.1.1. Определение интеграла Лебега

Как обычно, мы рассматриваем вероятностное пространство $(\Omega, \mathscr{F}, \mathsf{P})$.

Определение. Случайная величина $\xi\colon \Omega \to \mathbb{R}$ называется *простой*, если она представима в виде

$$\xi(\omega) = \sum_{i=1}^{n} a_i \mathbb{I}_{A_i}(\omega), \tag{1}$$

где $A_1, \ldots, A_n$ — разбиение Ω.

Математическое ожидание определяется в три этапа: сначала для простых случайных величин, потом для произвольных неотрицательных, и только потом для любых. Приступим…

Определение. *Интегралом Лебега* по вероятностной мере P *(математическим ожиданием)* простой случайной величины ξ называется число

$$\mathsf{E}\xi := \sum_{i=1}^{n} a_k \mathsf{P}(A_k). \tag{2}$$

Задача 3.1. *Показать, что значение $\mathsf{E}\xi$ не зависит от разбиения, т. е.*

$$\xi = \sum_{i=1}^{n} a_i \mathbb{I}_{A_i} = \sum_{j=1}^{m} b_j \mathbb{I}_{B_j} \;\Rightarrow\; \mathsf{E}\xi = \sum_{i=1}^{n} a_i \mathsf{P}(A_i) = \sum_{j=1}^{m} b_j \mathsf{P}(B_j). \tag{3}$$

Решение. Достаточно перейти к более мелкому разбиению, т. е. к попарным пересечениям всех A_i и B_j. ■

Очевидно, что для простых случайных величин верны свойства:
1° $\mathsf{E}(c\xi) = c\mathsf{E}\xi$;
2° $\mathsf{E}(\xi \pm \eta) = \mathsf{E}\xi \pm \mathsf{E}\eta$;
3° $\xi \leqslant \eta \ \Rightarrow\ \mathsf{E}\xi \leqslant \mathsf{E}\eta$ (в частности, $\xi \geqslant 0 \ \Rightarrow\ \mathsf{E}\xi \geqslant 0$).
Теперь определим математическое ожидание произвольной неотрицательной величины.

Определение. Пусть $\xi \geqslant 0$. Положим

$$\mathsf{E}\xi := \sup\left\{\mathsf{E}\eta\colon \eta \leqslant \xi\right\}, \tag{4}$$

где η — простые случайные величины.

Замечание. Для неотрицательных случайных величин математическое ожидание определено всегда, однако оно может принимать бесконечные значения.

Наконец, определим $\mathsf{E}\xi$ для произвольной величины ξ. Для этого введём следующие обозначения:

$$\xi^+ := \max\left\{\xi, 0\right\}, \quad \xi^- := -\min\left\{\xi, 0\right\}. \tag{5}$$

Определение. Пусть ξ — действительная случайная величина. Положим

$$\mathsf{E}\xi := \mathsf{E}\left(\xi^+\right) - \mathsf{E}\left(\xi^-\right). \tag{6}$$

При этом говорят, что математическое ожидание существует, если хотя бы одно из значений $\mathsf{E}\left(\xi^+\right)$ и $\mathsf{E}\left(\xi^-\right)$ конечно, причём $C + \infty := +\infty$, а $C - \infty := -\infty$. В противном случае $\mathsf{E}\xi$ не определено.

В дальнейшем будем говорить, что математическое ожидание существует, если интегралы от положительной и отрицательной части конечны.

Интеграл Лебега обозначается так:

$$\mathsf{E}\xi = \int\limits_{\Omega} \xi \, d\mathsf{P} = \int\limits_{\Omega} \xi(\omega)\ \mathsf{P}(d\omega). \tag{7}$$

Если ξ интегрируема по Лебегу, пишут $\xi \in \mathcal{L}^1$.

Обозначение «E» происходит от французского *«espérance mathématique»*. Иногда используется буква «M», происходящая от английского *«mean value»* (среднее значение), что лучше отражает смысл данного понятия.

Замечание. Интеграл Лебега относится к числу *абсолютных* интегралов, т. е. $f \in \mathcal{L}^1 \ \Leftrightarrow\ |f| \in \mathcal{L}^1$.

3.1.2. Свойства математического ожидания

Далее будем рассматривать случайную величину $\xi\colon \Omega \to \mathbb{R}$.

Следующее утверждение, несмотря на свою очевидность, чрезвычайно важно и будет многократно применяться ниже.

Лемма 3.1 (об аппроксимации). *Пусть $\xi \geqslant 0$. Тогда существует последовательность простых случайных величин $\{\xi_n\}$ такая, что $\xi_n \nearrow \xi$.*

□ Построим «ступенчатые» функции, уменьшая высоту ступенек со скоростью 2^{-n} и наращивая высоту «лестницы» со скоростью n. Там, где наша функция оказывается между ступеньками, берем нижнюю границу, а выше уровня n «срезаем» функцию. Более формально, положим

$$\xi_n := \begin{cases} k2^{-n}, & k2^n \leqslant \xi(\omega) < (k+1)2^{-n}; \\ n, & \xi(\omega) > n, \end{cases} \ \text{где } k = \overline{0,\, 2^n - 1}. \tag{8}$$

Это — искомая последовательность. Сходимость к ξ и неубывание последовательности очевидны. ■

Лемма 3.2. *Пусть $0 \leqslant \xi_n \nearrow \xi$, где ξ_n — простые случайные величины. Тогда $\mathsf{E}\xi_n \to \mathsf{E}\xi$.*

□ Положим $a := \lim \mathsf{E}\xi_n$ и покажем, что $a = \mathsf{E}\xi$. По определению математического ожидания $\mathsf{E}\xi_n \leqslant \mathsf{E}\xi$, а значит, и $a \leqslant \mathsf{E}\xi$. Докажем обратное неравенство, т. е. что $a \geqslant \mathsf{E}\eta$ для любой простой величины $0 \leqslant \eta \leqslant \xi$. Пусть η принимает значения $a_1, \ldots, a_k$ на множествах $A_1, \ldots, A_k$ соответственно. Фиксируем $\varepsilon \in (0, 1)$. Введём случайные величины

$$\eta_n := (1 - \varepsilon)\eta \cdot \mathbb{I}_{\{(1-\varepsilon)\eta \leqslant \xi_n\}}. \tag{9}$$

Рассмотрим множества $A_{in} := A_i \cap \{(1-\varepsilon)\eta \leqslant \xi_n\}$. Очевидно, что $A_{in} \nearrow A_i$ при $n \to \infty$. Значит, $\mathsf{P}(A_{in}) \to \mathsf{P}(A_i)$ при $n \to \infty$. Тогда

$$\mathsf{E}\eta_n = (1-\varepsilon)\sum_{i=1}^{k} a_i \mathsf{P}(A_{in}) \to (1-\varepsilon)\sum_{i=1}^{k} a_i \mathsf{P}(A_i) = (1-\varepsilon)\mathsf{E}\eta. \tag{10}$$

Значит, $(1-\varepsilon)\mathsf{E}\eta \leqslant a$, а так как ε произвольно, то $\mathsf{E}\eta \leqslant a$. ■

Теорема 3.3. *Множество интегрируемых функций $\mathcal{L}^1$ есть векторное пространство, а E — линейный оператор на $\mathcal{L}^1$. Если $\xi \geqslant 0$, то и $\mathsf{E}\xi \geqslant 0$. Если $0 \leqslant \eta \leqslant \xi$ и $\xi \in \mathcal{L}^1$, то и $\eta \in \mathcal{L}^1$. Если $\xi \overset{\text{п.н.}}{=} \eta$, то $\mathsf{E}\xi = \mathsf{E}\eta$.*

□ Докажем первое утверждение теоремы. Если ξ и η — простые случайные величины, то $\alpha\xi+\beta\eta$ — простая случайная величина, и $\mathsf{E}(\alpha\xi+\beta\eta) = \alpha\mathsf{E}\xi+\beta\mathsf{E}\eta$ для любых $\alpha, \beta \in \mathbb{R}$. Если $\xi, \eta \geqslant 0$, то построим последовательности $0 \leqslant \xi_n \nearrow \xi$ и $0 \leqslant \eta_n \nearrow \eta$. Тогда $\mathsf{E}(\xi_n + \eta_n) = \mathsf{E}\xi_n + \mathsf{E}\eta_n$, а по предыдущей лемме $\mathsf{E}(\xi_n + \eta_n) \nearrow \mathsf{E}(\xi + \eta)$ и $\mathsf{E}\xi_n + \mathsf{E}\eta_n \nearrow \mathsf{E}\xi + \mathsf{E}\eta$. Далее, если $C \geqslant 0$, то $C\xi_n \nearrow C\xi$, поэтому $\mathsf{E}(C\xi) = C\mathsf{E}\xi$.

Покажем теперь, что если $\xi, \eta \geqslant 0$, то

$$\mathsf{E}(\xi - \eta) = \mathsf{E}\xi - \mathsf{E}\eta. \tag{11}$$

По определению, $\mathsf{E}(\xi-\eta) = \mathsf{E}(\xi-\eta)^+ - \mathsf{E}(\xi-\eta)^-$. Так как $(\xi-\eta)^+ + \eta = \xi + (\xi-\eta)^-$, и каждое слагаемое справа и слева неотрицательно, то $\mathsf{E}(\xi-\eta)^+ + \mathsf{E}\eta = \mathsf{E}\xi + \mathsf{E}(\xi-\eta)^-$, а это равносильно (11). Представляя каждое из слагаемых в виде разности двух неотрицательных, получаем

$$\mathsf{E}(\xi+\eta) = \mathsf{E}\big((\xi^+ + \eta^+) - (\xi^- + \eta^-)\big) = (\mathsf{E}\xi^+ + \mathsf{E}\eta^+) - (\mathsf{E}\xi^- + \mathsf{E}\eta^-) = \mathsf{E}\xi + \mathsf{E}\eta. \tag{12}$$

Таким образом, первое утверждение доказано для любых $\xi, \eta \in \mathcal{L}^1$.

Второе и третье утверждение теоремы очевидны. Докажем четвёртое утверждение. Как видно из доказательства первой части, можно рассматривать лишь случай $\xi, \eta \geqslant 0$. Пусть событие $A := \{\xi = \eta\}$. Построим $0 \leqslant \xi_n \nearrow \xi$ и $0 \leqslant \eta_n \nearrow \eta$. Тогда $\mathsf{E}\eta_n \nearrow \mathsf{E}\eta$. Имеем

$$\mathsf{E}\eta_n = \mathsf{E}(\eta_n\mathbb{I}_A) + \mathsf{E}(\eta_n\mathbb{I}_{\overline{A}}). \tag{13}$$

Возьмём η_n из леммы об аппроксимации. Так как $\mathsf{P}(\overline{A}) = 0$, то $\mathsf{E}\eta_n = \mathsf{E}(\eta_n\mathbb{I}_A) = \mathsf{E}(\xi_n\mathbb{I}_A) \to \mathsf{E}(\xi\mathbb{I}_A) = \mathsf{E}\xi$. ■

Как известно, интеграл от произведения функций в общем случае не равен произведению интегралов. А вот если случайные величины независимы, то оказывается, что это всегда верно. Настало время это доказать.

Теорема 3.4. *Пусть $\xi, \eta \in \mathcal{L}^1$ — независимые случайные величины. Тогда $(\xi\cdot\eta) \in \mathcal{L}^1$ и $\mathsf{E}(\xi\cdot\eta) = \mathsf{E}\xi\cdot\mathsf{E}\eta$.*

□ Очевидно, что ξ^+, η^+ и ξ^-, η^- — также независимые величины (это борелевские функции от независимых величин). Покажем, что $\mathsf{E}(\xi^+\cdot\eta^+) = \mathsf{E}(\xi^+)\cdot\mathsf{E}(\eta^+)$. Снова построим величины $0 \leqslant \xi_n \nearrow \xi$ и $0 \leqslant \eta_n \nearrow \eta$. Величины ξ_n и η_n будут независимыми. Имеем

$$\xi_n = \sum_{i=1}^{k} a_i^{(n)}\mathbb{I}_{A_i^{(n)}}, \quad \eta_n = \sum_{j=1}^{m} b_j^{(n)}\mathbb{I}_{B_j^{(n)}}. \tag{14}$$

Пользуясь тем, что $\mathsf{P}\big(A_i^{(n)}\cdot B_j^{(n)}\big) = \mathsf{P}\big(A_i^{(n)}\big)\cdot\mathsf{P}\big(B_j^{(n)}\big)$, перемножим ряды, и получим, что $\mathsf{E}(\xi_n\eta_n) = \mathsf{E}\xi_n\mathsf{E}\eta_n$. Остаётся перейти к пределу при $n \to \infty$. ■

3.1.3. Теоремы о предельном переходе под знаком интеграла Лебега

Следующая теорема является обобщением уже полученных результатов.

Теорема 3.5 (о монотонной сходимости). *Пусть $0 \leqslant \xi_n \nearrow \xi$. Тогда $\mathsf{E}\xi_n \nearrow \mathsf{E}\xi$.*

□ Построим простые величины $\eta_{nk} \nearrow \xi_n$ при $k \to \infty$.

$$\begin{array}{lr}
\eta_{11} \leqslant \eta_{12} \leqslant \ldots \leqslant \eta_{1k} \leqslant \ldots \nearrow & \xi_1 \\
 & \rotatebox{90}{$\geqslant$} \\
\vdots & \vdots \\
 & \rotatebox{90}{$\geqslant$} \\
\eta_{k1} \leqslant \eta_{k2} \leqslant \ldots \leqslant \eta_{kk} \leqslant \ldots \nearrow & \xi_k
\end{array}$$

Положим

$$\zeta_k := \max_{1\leqslant i,j\leqslant k} \eta_{ij} = \max_{1\leqslant i\leqslant k} \eta_{nk}. \tag{15}$$

Это будет возрастающая последовательность простых случайных величин. Пусть $\zeta_k \nearrow \zeta$. Имеем

$$0 \leqslant \eta_{nk} \leqslant \zeta_k \leqslant \xi_k \leqslant \xi. \tag{16}$$

Устремим $k \to \infty$, получим $\xi_n \leqslant \zeta \leqslant \xi$. Теперь устремим $n \to \infty$, получим $\xi \leqslant \zeta \leqslant \xi$, откуда $\zeta = \xi$.

В силу неравенства (16) $\mathsf{E}\eta_{nk} \leqslant \mathsf{E}\zeta_k \leqslant \mathsf{E}\xi_k$. В пределе при $k \to \infty$ получаем $\mathsf{E}\xi_n \leqslant \mathsf{E}\zeta \leqslant \lim \mathsf{E}\xi_k$. Переходя к пределу при $n \to \infty$, получаем $\lim \mathsf{E}\xi_n \leqslant \mathsf{E}\zeta \leqslant \lim \mathsf{E}\xi_k$. Но $\zeta = \xi$, поэтому $\lim \mathsf{E}\xi_n = \mathsf{E}\xi$. ■

Следствие 3.1. *Пусть $\xi_n \geqslant 0$. Тогда $\mathsf{E} \sum\limits_{n=1}^{\infty} \xi_n = \sum\limits_{n=1}^{\infty} \mathsf{E}\xi_n$.*

□ Нужно рассмотреть (неубывающую) последовательность частичных сумм и применить теорему. ■

Следствие 3.2. *Пусть $\sum\limits_{n=1}^{\infty} \mathsf{E}|\xi_n|$ сходится. Тогда $\sum\limits_{n=1}^{\infty} \xi_n$ сходится почти наверное и $\mathsf{E} \sum\limits_{n=1}^{\infty} \xi_n = \sum\limits_{n=1}^{\infty} \mathsf{E}\xi_n$.*

□ В самом деле, если бы функциональный ряд расходится на множестве положительной меры, то расходится и ряд из модулей. Но тогда расходится и ряд из интегралов от модулей. Противоречие. ■

Задача 3.2. *Доказать более общую теорему: пусть $\eta \leqslant \xi_n \nearrow \xi$, где $\mathsf{E}\eta > -\infty$. Тогда $\mathsf{E}\xi_n \nearrow \mathsf{E}\xi$.*

Лемма 3.6 (Фату). *Пусть $\eta \leqslant \xi_n$, и $\mathsf{E}\eta > -\infty$. Тогда $\mathsf{E} \liminf \xi_n \leqslant \liminf \mathsf{E}\xi_n$.*

□ Имеем

$$\xi := \liminf \xi_n = \lim_n \underbrace{\inf_{k \geqslant n} \xi_k}_{\eta_n}. \tag{17}$$

Тогда $\eta_n \nearrow \xi$, и $\eta \leqslant \eta_n \leqslant \xi_n$. Из этого неравенства и теоремы о монотонной сходимости (точнее, из задачи 3.2) следует, что $\mathsf{E}\xi = \lim \mathsf{E}\eta_n \leqslant \liminf \mathsf{E}\xi_n$. ■

Теорема 3.7 (Лебега о мажорируемой сходимости). *Пусть ξ, η — случайные величины, и $\xi_n \to \xi$, и $|\xi| \leqslant \eta$. Пусть $\eta \in \mathcal{L}^1$. Тогда $\mathsf{E}\xi_n \to \mathsf{E}\xi$.*

□ Имеем $-\xi_n \geqslant -\eta$ и $\mathsf{E}(-\eta) > -\infty$. По лемме Фату

$$\mathsf{E}\big(\liminf(-\xi_n)\big) \leqslant \liminf \mathsf{E}(-\xi_n) \ \Rightarrow\ \mathsf{E}(-\limsup \xi_n) \leqslant -\limsup \mathsf{E}\xi_n. \tag{18}$$

Выносим минус за знак E, и получаем $\mathsf{E}(\limsup \xi_n) \geqslant \limsup \mathsf{E}\xi_n$. Таким образом,

$$\mathsf{E} \limsup \xi_n \geqslant \limsup \mathsf{E}\xi_n \geqslant \liminf \mathsf{E}\xi_n \geqslant \mathsf{E} \liminf \xi_n. \tag{19}$$

Отсюда всё следует, так как $\xi = \lim \xi_n = \liminf \xi_n = \limsup \xi_n$, а потому $\mathsf{E}\xi \geqslant \limsup \mathsf{E}\xi_n \geqslant \limsup \mathsf{E}\xi_n \geqslant \mathsf{E}\xi$. Следовательно, $\lim \mathsf{E}\xi_n = \mathsf{E}\xi$. ■

Теорема 3.8. *Пусть $(\Omega, \mathscr{F}, \mathsf{P})$ и $(\Theta, \mathscr{B}, \mathsf{P}_\xi)$ — вероятностные пространства, где $\xi\colon \Omega \to \Theta$ — случайная величина, а $\mathsf{P}_\xi = \mathrm{Law}(\xi)$. Пусть также $\varphi\colon \Theta \to \mathbb{R}$ — борелевская функция. Тогда*

$$\mathsf{E}\varphi\big(\xi(\omega)\big) = \int\limits_{\Omega} \varphi\big(\xi(\omega)\big)\, \mathsf{P}(d\omega) = \int\limits_{\Theta} \varphi(\theta)\, \mathsf{P}_\xi(d\theta). \tag{20}$$

□ Докажем утверждение сначала для индикаторов. Имеем

$$\mathsf{E}\mathbb{I}_B\big(\xi(\omega)\big) = \mathsf{P}\,(\xi \in B) \stackrel{\text{def}}{=} \mathsf{P}_\xi(B) = \int\limits_{\Theta} \mathbb{I}_B(\theta)\, \mathsf{P}_\xi(d\theta). \tag{21}$$

Для простой функции $\varphi = \sum\limits_{k=1}^{n} a_k \mathbb{I}_{A_k}$ утверждение также справедливо по линейности интеграла.

Теперь пусть $\varphi \geqslant 0$. Построим простые функции $0 \leqslant \varphi_n \nearrow \varphi$. По теореме о монотонной сходимости

$$\mathsf{E}\varphi_n(\xi) = \int\limits_{\Theta} \varphi_n(\theta)\, \mathsf{P}_\xi(d\theta) \to \int\limits_{\Theta} \varphi(\theta)\, \mathsf{P}_\xi(d\theta) = \mathsf{E}\varphi(\xi),\ n \to \infty. \tag{22}$$

В общем случае рассмотрим φ^+ и φ^- и повторим предыдущее рассуждение. ■

Замечание. Как уже было замечено выше, все теоремы о сходимости случайных величин можно переформулировать в терминах сходимости почти наверное, пополнив вероятностное пространство, поскольку значение интеграла не зависит от значений функции на множестве меры нуль.

Далее будем рассматривать вероятностное пространство с *σ-конечной* мерой, т. е. такое, которое можно разбить на счётное число подмножеств, каждое из которых имеет конечную меру.

Определение. Пусть ξ — случайная величина. Говорят, что она обладает *плотностью*, если её функция распределения представляется интегралом от неотрицательной функции $p(t)$:

$$F_\xi(x) = \int\limits_{-\infty}^{x} p(t)\, dt. \tag{23}$$

Очевидно, что если функция f интегрируема по Риману, то она интегрируема по Лебегу и интегралы совпадают. Если же существует несобственный интеграл Римана у $|f|$, то существует и интеграл Лебега и они также совпадают.

Пусть $p(t)$ — плотность величины ξ. Введём функцию

$$\mathsf{Q}(B) := \int\limits_B p(t)\,dt = \int\limits_{\mathbb{R}} p(t)\mathbb{I}_B(t)\,dt. \tag{24}$$

Эта функция является мерой на $\mathscr{B}(\mathbb{R})$, так как

$$\mathsf{Q}\left(\bigcup B_i\right) = \int\limits_{\mathbb{R}} p(t)\sum \mathbb{I}_{B_i}(t)\,dt = \sum \int\limits_{\mathbb{R}} p(t)\mathbb{I}_{B_i}(t)\,dt = \sum \mathsf{Q}(B_i). \tag{25}$$

Имеем $\mathsf{Q}\big((-\infty, x]\big) = \int\limits_{-\infty}^{x} p(t)\,dt = F_\xi(x)$. Значит, $\mathsf{Q} = \mathsf{P}_\xi$.

Теорема 3.9. *Пусть случайная величина ξ имеет плотность $p(t)$, а φ — борелевская функция. Тогда*

$$\mathsf{E}\big(\varphi(\xi)\big) = \int\limits_{\mathbb{R}} \varphi(x)\,\mathsf{P}_\xi(dx) = \int\limits_{\mathbb{R}} h(t)p(t)\,dt. \tag{26}$$

3.2. Дисперсия и ковариация. Пространство L^p

3.2.1. Определения и свойства

Определим понятие «момента инерции» случайной величины — величину среднеквадратичного отклонения от своего среднего значения.

Определение. Пусть существует $\mathsf{E}\xi^2$. *Дисперсией* величины ξ называется число

$$\mathsf{D}\xi := \mathsf{E}(\xi - \mathsf{E}\xi)^2. \tag{27}$$

Рассмотрим множество $\mathcal{L}^p$ случайных величин ξ на вероятностном пространстве $(\Omega, \mathscr{F}, \mathsf{P})$, у которых существует $\mathsf{E}|\xi|^p$, где $p \in (1, \infty)$ — некоторое (фиксированное) число. Введём на нём отношение эквивалентности: $\xi \sim \eta \;\Leftrightarrow\; \xi \overset{\text{п.н.}}{=} \eta$. Положим $L^p := {\mathcal{L}^p}/{\sim}$ — пространство классов эквивалентных функций, интегрируемых вместе со своей p-й степенью.

Пространство L^2 является полным нормированным *(гильбертовым)* пространством. Скалярное произведение вводится так:

$$\langle X, Y\rangle := \int\limits_{\Omega} XY\,d\mathsf{P}. \tag{28}$$

Очевидно, что это симметричная билинейная положительно определённая форма. Как будет показано ниже, $\langle X, X\rangle = 0 \;\Leftrightarrow\; X \overset{\text{п.н.}}{=} 0$.

Таким образом, $\mathsf{D}\xi$ имеет смысл, если $\xi \in L^2$.

Раскрывая скобки в определении дисперсии, по свойству линейности получаем

$$\mathsf{D}\xi = \mathsf{E}\left(\xi - \mathsf{E}\xi\right)^2 = \mathsf{E}\left[\xi^2 - 2\xi\mathsf{E}\xi + (\mathsf{E}\xi)^2\right] = \mathsf{E}(\xi^2) - \left(\mathsf{E}\xi\right)^2. \tag{29}$$

Теперь определим «скалярное произведение» случайных величин.

Определение. *Ковариацией* величин $\xi, \eta \in L^2$ называется число

$$\operatorname{cov}(\xi, \eta) := \mathsf{E}\big[(\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta)\big]. \tag{30}$$

Определение корректно в силу неравенства $|XY| \leqslant \frac{1}{2}\left(|X|^2 + |Y|^2\right)$, где $X := \xi - \mathsf{E}\xi$, а $Y := \eta - \mathsf{E}\eta$.

Если в определении раскрыть скобки, получится ещё одна хорошая формула для ковариации:

$$\operatorname{cov}(\xi, \eta) = \mathsf{E}(\xi \cdot \eta) - \mathsf{E}\xi \cdot \mathsf{E}\eta. \tag{31}$$

Замечание. Если ξ, η — независимые случайные величины, то $\operatorname{cov}(\xi, \eta) = 0$, так как $\mathsf{E}\xi\mathsf{E}\eta = \mathsf{E}\xi\eta$. Обратное, однако, неверно: существуют зависимые случайные величины, с нулевой ковариацией. В качестве примера годятся, скажем, $\sin\nu$ и $\cos\nu$, где величина ν равномерно распределена на $[0, \pi]$.

Теперь сформулируем некоторые свойства дисперсии и ковариации. Очевидно, что ковариация если симметричная билинейная функция на L^2, именно потому она имеет сходство со скалярным произведением. Кроме того, $\mathsf{D}\xi = \operatorname{cov}(\xi, \xi)$, и $\mathsf{D}(C\xi) = C^2\mathsf{D}\xi$.

Из свойства билинейности и симметричности ковариации следует, что

$$\mathsf{D}\sum_{k=1}^{n}\xi_k = \sum_{i=1}^{n}\sum_{j=1}^{n}\operatorname{cov}(\xi_i, \xi_j) = \sum_{k=1}^{n}\mathsf{D}\xi_k + \sum_{i\neq j}\operatorname{cov}(\xi_i\xi_j) = \sum_{k=1}^{n}\mathsf{D}\xi_k + 2\sum_{i<j}\operatorname{cov}(\xi_i, \xi_j). \tag{32}$$

Если же ξ_k попарно независимы, то $\mathsf{D}\sum\xi_k = \sum\mathsf{D}\xi_k$.

Пример 2.1. Пусть $\mathsf{P}(\xi = 1) = p$, $\mathsf{P}(\xi = 0) = 1 - p$. Тогда $\mathsf{E}\xi = p$, а $\mathsf{D}\xi = \mathsf{E}\left(\xi^2\right) - \left(\mathsf{E}\xi^2\right) = p(1-p)$.

3.2.2. Неравенство Чебышева

Лемма 3.10 (неравенство Чебышева[5] для L^p). *Для* $\forall\,\alpha \geqslant 0$ *выполняется неравенство*

$$\alpha^p \cdot \mathsf{P}\left\{x\colon |f(x)| \geqslant \alpha\right\} \leqslant \int\limits_{\Omega}|f(x)|^p\, d\mathsf{P}. \tag{33}$$

□ Пусть $L := \{\omega\colon\ |f(x)| < \alpha\}$, а $G := \Omega \setminus L$. Тогда

$$\int\limits_{\Omega}|f|^p\, d\mathsf{P} = \int\limits_{L}|f|^p\, d\mathsf{P} + \int\limits_{G}|f|^p\, d\mathsf{P} \geqslant \int\limits_{G}|f|^p\, d\mathsf{P}. \tag{34}$$

Последний интеграл оценим снизу числом $\alpha^p\mathsf{P}(G)$, ибо $|f|^p \geqslant \alpha^p$ на G. Значит, $\int\limits_{\Omega}|f|^p\, d\mathsf{P} \geqslant \alpha^p \cdot \mathsf{P}\left\{|f| \geqslant \alpha\right\}$. ■

На языке теории вероятностей это неравенство при $p = 1$ переформулируется так:

$$\mathsf{P}\left(|\xi| \geqslant \varepsilon\right) \leqslant \frac{\mathsf{E}|\xi|}{\varepsilon}. \tag{35}$$

Следствие 3.3. $\mathsf{D}\xi = 0 \ \Leftrightarrow\ \mathsf{E}|\xi| = 0 \ \Leftrightarrow\ \xi \overset{\text{п.н.}}{=} 0$.

□ Справа налево утверждение очевидно. Наоборот: положим $\alpha = \frac{1}{n}$ для $\forall n \in \mathbb{N}$. Тогда по условию и доказанному неравенству $\mathsf{P}\left(|\xi| \geqslant \frac{1}{n}\right) \leqslant n \cdot \mathsf{E}|\xi| = 0$. Так как n произвольно, то $\mathsf{P}\left(|\xi| \neq 0\right) = 0$. ■

Ещё одно следствие леммы будет много раз применяться в следующей главе.

Следствие 3.4 (неравенство Чебышева).

$$\mathsf{P}\big(|\xi - \mathsf{E}\xi| \geqslant \varepsilon\big) \leqslant \frac{\mathsf{D}\xi}{\varepsilon^2}. \tag{36}$$

□ Рассмотрим случайную величину $\eta := \xi - \mathsf{E}\xi$ и запишем для неё неравенство Чебышева для L^2. Это и будет требуемое неравенство, так как $\mathsf{E}\eta^2 = \mathsf{D}\xi$. ■

4. Сходимость случайных величин. Закон больших чисел

4.1. Закон больших чисел

4.1.1. Простейший вариант ЗБЧ

Определение. Говорят, что $\xi_n \to \xi$ *по вероятности*, и пишут «$\xi_n \xrightarrow{\mathsf{P}} \xi$», если для $\forall\,\varepsilon > 0$ вероятность отклонения ξ_n от ξ на величину ε стремится к нулю:

$$\mathsf{P}\big\{|\xi_n - \xi| \geqslant \varepsilon\big\} \to 0,\ n \to \infty. \tag{1}$$

Докажем теперь один из фундаментальных результатов теории вероятностей.

Теорема 4.1 (Закон больших чисел). *Пусть* $\{\xi_n\}$ *— последовательность попарно независимых случайных величин. Пусть также их дисперсии ограничены в совокупности:* $\mathsf{D}\xi_n \leqslant C < \infty$ *для* $\forall\,n$. *Тогда*

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow{\mathsf{P}} \frac{1}{n}\sum_{k=1}^{n}\mathsf{E}\xi_k. \tag{2}$$

[5] Эта лемма в явном виде не доказывалась на лекциях, но она сильно сокращает дальнейшие выкладки.

□ Введём обозначение: $\xi := \frac{1}{n}\sum\limits_{k=1}^{n}\xi_k$. Нужно показать, что $\mathsf{P}\big(|\xi - \mathsf{E}\xi| \geqslant \varepsilon\big) \to 0$. По неравенству Чебышева

$$\mathsf{P}\big(|\xi - \mathsf{E}\xi| \geqslant \varepsilon\big) \leqslant \frac{\mathsf{D}\xi}{\varepsilon^2} = \frac{1}{n^2\varepsilon^2}\mathsf{D}\sum_{k=1}^{n}\xi_k \overset{!}{=} \frac{1}{n^2\varepsilon^2}\sum_{k=1}^{n}\mathsf{D}\xi_k \leqslant \frac{C\cdot n}{n^2\varepsilon^2} = \frac{C}{n\varepsilon^2} \to 0,\ n\to\infty. \tag{3}$$

Равенство «!» обусловлено попарной независимостью. ■

Иными словами, при большом количестве случайных испытаний отклонение от среднего значения неслучайно. Это явление, замеченное уже давно, как мы сейчас убедились, имеет строгое математическое обоснование.

Следствие 4.1. *В схеме испытаний Бернулли средняя частота успехов стремится к p.*

4.1.2. Теорема Вейерштрасса

Покажем, как теперь можно сравнительно просто доказать теорему о плотности многочленов в $\mathbf{C}[0,1]$.

Теорема 4.2. *Пусть $f \in \mathbf{C}[0,1]$. Тогда найдётся последовательность многочленов $\{P_n\}$ таких, что $P_n \rightrightarrows f$.*

□ В силу непрерывности на компакте f ограничена по модулю числом M и равномерно непрерывна на нём. Определим *многочлены Бернштейна*

$$B_n(f,p) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\mathbf{C}_n^k p^k(1-p)^{n-k}. \tag{4}$$

Для удобства считаем, что $0^0 = 1$ и $B_n(f,0) = f(0)$, а $B_n(f,1) = f(1)$. Введём независимые величины

$$\xi_k := \begin{cases} 1 & \text{с вероятностью } p; \\ 0 & \text{с вероятностью } (1-p). \end{cases} \tag{5}$$

Пусть $\sigma_n := \xi_1 + \ldots + \xi_n$, тогда $\mathsf{P}(\sigma_n = k) = \mathbf{C}_n^k p^k(1-p)^{n-k}$. Теперь заметим, что

$$\mathsf{E}f\left(\frac{\sigma_n}{n}\right) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\cdot\mathsf{P}(\sigma_n = k) = B_n(f,p). \tag{6}$$

Приступим у оценке разности. Так как $\mathsf{E}(\mathrm{const}) = \mathrm{const}$, и $|\mathsf{E}\xi| \leqslant \mathsf{E}|\xi|$, то имеем

$$|f(p) - B_n(f,p)| = \left|\mathsf{E}f(p) - \mathsf{E}f\left(\frac{\sigma_n}{n}\right)\right| \leqslant \mathsf{E}\left|f(p) - f\left(\frac{\sigma_n}{n}\right)\right| = (*). \tag{7}$$

Пусть событие $A := \big\{\big|\frac{\sigma_n}{n} - p\big| < \delta\big\}$, где $\delta > 0$. Тогда, разбивая разность на две части и оценивая вторую по неравенству Чебышева, получаем

$$(*) = \mathsf{E}\left|f(p) - f\left(\frac{\sigma_n}{n}\right)\right|\cdot\mathbb{I}_A + \mathsf{E}\left|f(p) - f\left(\frac{\sigma_n}{n}\right)\right|\cdot\mathbb{I}_{\overline{A}} \leqslant \omega_f(\delta)\cdot\underbrace{\mathsf{P}\left(\left|\frac{\sigma_n}{n} - p\right| < \delta\right)}_{\leqslant 1} + 2M\cdot\mathsf{P}\left(\left|\frac{\sigma_n}{n} - p\right| \geqslant \delta\right) \leqslant$$

$$\leqslant \omega_f(\delta) + 2M\cdot\frac{\mathsf{D}\left(\frac{\sigma_n}{n}\right)}{\delta^2} = \omega_f(\delta) + 2M\cdot\frac{1}{n^2\delta^2}\sum_{k=0}^{n}\mathsf{D}\xi_k = \omega_f(\delta) + 2M\cdot\frac{np(1-p)}{n^2\delta^2} \leqslant \omega_f(\delta) + \frac{M}{2n\delta^2} \to 0,\ n\to\infty,$$

так как $\max\limits_{[0,1]}\big\{p(1-p)\big\} = \frac{1}{4}$, а модуль непрерывности f стремится к нулю при достаточно малых δ. ■

4.2. Различные виды сходимости и их взаимосвязь

4.2.1. Теорема Пуассона

4.2.2. Сходимость по вариации

4.2.3. Сходимость по распределению и слабая сходимость

4.2.4. Сходимость почти всюду и сходимость по вероятности

4.3. Усиленные законы больших чисел и их следствия

4.3.1. УЗБЧ для некоррелированных величин

Теорема 4.3. *Пусть $X_1,\ \ldots,\ X_n,\ \ldots$ — случайные величины такие, что $\mathsf{D}X_k \leqslant c < \infty$ $\forall k$ и $\operatorname{cov} X_iX_j = 0$ $\forall i \neq j$ (т.е. случайные величины $X_1,\ \ldots,\ X_n,\ \ldots$ некоррелированы). Тогда*

$$\frac{1}{n}\sum_{k=1}^{n}X_k - \frac{1}{n}\sum_{k=1}^{n}\mathsf{E}X_k \xrightarrow{\text{п.н.}} 0, \quad n\to\infty. \tag{8}$$

$\square$ Без ограничения общности можно считать, что $\mathsf{E}X_k = 0\ \forall k$. Действительно, иначе рассмотрим случайные величины $\widetilde{X}_k = X_k - \mathsf{E}X_k$. Для них $\mathsf{D}\widetilde{X}_k = \mathsf{D}X_k \leqslant c$, $\operatorname{cov}\widetilde{X}_i\widetilde{X}_j = \operatorname{cov}X_iX_j = 0$, $\mathsf{E}\widetilde{X}_k = 0\ \forall k$ и $\frac{1}{n}\sum\limits_{k=1}^{n} X_k - \frac{1}{n}\sum\limits_{k=1}^{n}\mathsf{E}X_k = \frac{1}{n}\sum\limits_{k=1}^{n}\widetilde{X}_k$.

Для каждого $n \geqslant 2$ найдем $m(n) \in \mathbb{N}$: $m(n)^2 < n \leqslant (m(n)+1)^2$. Очевидно, $m(n) \nearrow \infty$. Обозначим $S_n = X_1 + \ldots + X_n$. Тогда $\dfrac{|S_n|}{n} \leqslant \dfrac{|S_{m(n)^2}| + Z_{m(n)}}{n}$, где $Z_{m(n)} := \max\limits_{1\leqslant k\leqslant 2m(n)+1}\left|X_{m(n)^2+1} + \ldots + X_{m(n)^2+k}\right|$ (напоминаем, что $(m(n)+1)^2 = m(n)^2 + 2m(n) + 1$). Далее, имеем $\dfrac{|S_n|}{n} \leqslant \dfrac{|S_{m(n)^2}| + Z_{m(n)}}{n} \leqslant \dfrac{S_{m(n)^2}}{m(n)^2} + \dfrac{Z_{m(n)}}{m(n)^2}$, поэтому достаточно доказать, что $\dfrac{S_{m(n)^2}}{m(n)^2} \xrightarrow{\text{п.н.}} 0$ и $\dfrac{Z_{m(n)}}{m(n)^2} \xrightarrow{\text{п.н.}} 0$ при $n \to \infty$.

Для $\forall\varepsilon > 0$ рассмотрим $\mathsf{P}\left(\omega : \frac{|S_{m^2}(\omega)|}{m^2} > \varepsilon\right)$. По неравенству Чебышёва и используя условия $\mathsf{D}X_k \leqslant c < \infty$ $\forall k$ и $\operatorname{cov}X_iX_j = 0\ \forall i \neq j$, получаем

$$\mathsf{P}\left(\omega : \frac{|S_{m^2}(\omega)|}{m^2} > \varepsilon\right) \leqslant \frac{\mathsf{D}S_{m^2}}{\varepsilon^2 m^4} = \frac{\sum\limits_{k=1}^{m^2}\mathsf{D}X_k}{\varepsilon^2 m^4} \leqslant \frac{cm^2}{\varepsilon^2 m^4} = \frac{c}{\varepsilon^2 m^2}.$$

Следовательно, $\sum\limits_{m=1}^{\infty}\mathsf{P}\left(\frac{|S_{m^2}(\omega)|}{m^2} > \varepsilon\right) \leqslant \frac{c}{\varepsilon^2}\sum\limits_{m=1}^{\infty}\frac{1}{m^2} < \infty$. По лемме Бореля–Кантелли отсюда следует, что $\forall\omega \in \Omega_0$, $\mathsf{P}(\Omega_0) = 1$ $\exists N(\omega) : \dfrac{|S_{m^2}(\omega)|}{m^2} \leqslant \varepsilon\ \forall m \geqslant N$. А это и означает, что $\dfrac{S_{m^2}}{m^2} \xrightarrow{\text{п.н.}} 0$, $m \to \infty$.

Теперь рассмотрим $\frac{Z_m}{m^2}$. Имеем:

$$\mathsf{P}\left(\frac{|Z_m|}{m^2} > \varepsilon\right) = \mathsf{P}\left(\bigcup_{k=1}^{2m+1}\left\{|X_{m^2+1} + \ldots + X_{m^2+k}| > \varepsilon m^2\right\}\right) \leqslant \sum_{k=1}^{2m+1}\mathsf{P}\Big(|X_{m^2+1} + \ldots + X_{m^2+k}| > \varepsilon m^2\Big),$$

и применяя неравенство Чебышёва и условия теоремы, отсюда получаем:

$$\mathsf{P}\left(\frac{|Z_m|}{m^2} > \varepsilon\right) \leqslant \sum_{k=1}^{2m+1}\frac{\mathsf{D}(X_{m^2+1} + \ldots + X_{m^2+k})}{\varepsilon^2 m^4} \leqslant \frac{c(2m+1)}{\varepsilon^2 m^4} \sim \frac{2c}{\varepsilon^2 m^3}, \quad m \to \infty,$$

поэтому $\sum\limits_{m=1}^{\infty}\mathsf{P}\left(\frac{Z_m}{m^2} > \varepsilon\right) < \infty$. Снова по лемме Бореля–Кантелли получаем $\dfrac{Z_m}{m^2} \xrightarrow{\text{п.н.}} 0$, $m \to \infty$. А так как $m(n) \nearrow \infty$ при $n \to \infty$, то и $\dfrac{Z_{m(n)}}{m(n)^2} \xrightarrow{\text{п.н.}} 0$, $n \to \infty$. Теорема доказана. $\blacksquare$

4.3.2. Теорема Этемади

Теорема 4.4 (Этемади). *Пусть $X_1, X_2, \ldots$ — попарно независимые одинаково распределенные случайные величины, $\mathsf{E}|X_1| < \infty$. Тогда*

$$\frac{S_n}{n} \xrightarrow{\text{п.н.}} \mathsf{E}X_1, \quad n \to \infty, \tag{9}$$

где $S_n = X_1 + \ldots + X_n$.

Замечание. Если случайные величины $X_n \geqslant 0$, одинаково распределены, некоррелированы (т.е. $\operatorname{cov}X_iX_j = 0\ \forall i \neq j$) и существует $\mathsf{E}|X_1| < \infty$, то соотношение (9) следует из теоремы 4.3.

$\square$ Покажем сначала, что достаточно рассмотреть случай неотрицательных случайных величин X_n. Действительно, если они любого знака, то представляем их в виде $X_n = X_n^+ - X_n^- = x^+(X_n) - x^-(X_n)$, где $x^+ = \max(0, x)$, $x^- = x^+ - x$. Если $X_i \perp\!\!\!\perp X_j$, то $x^+(X_i) \perp\!\!\!\perp x^+(X_j)$ и $x^-(X_i) \perp\!\!\!\perp x^-(X_j)$. Поэтому если теорема будет доказана для неотрицательных величин, то применяя ее для $X_1^+, X_2^+, \ldots$ и $X_1^-, X_2^-, \ldots$, получим $\frac{S_n}{n} = \frac{\sum_1^n X_i^+}{n} - \frac{\sum_1^n X_i^-}{n} \xrightarrow{\text{п.н.}} \mathsf{E}X_1^+ - \mathsf{E}X_1^- = \mathsf{E}X_1$, $n \to \infty$.

Итак, далее считаем, что $X_n \geqslant 0$, $n = 1, 2, \ldots$. Введем «урезанные» случайные величины $Y_i = X_i\,\mathbb{I}\{X_i \leqslant i\}$, $i = 1, 2, \ldots$. Пусть $\widetilde{S}_n = \sum\limits_{i=1}^{n} Y_i$. Определим последовательность $k_n = [\alpha^n]$, $n = 1, 2, \ldots$, где $\alpha > 1$.

Лемма 4.5. $\dfrac{\widetilde{S}_{k_n} - \mathsf{E}\widetilde{S}_{k_n}}{k_n} \xrightarrow{\text{п.н.}} 0$, $n \to \infty$.

$\square$ Для доказательства применим лемму Бореля–Кантелли. А именно, докажем, что $\forall \varepsilon > 0$ сходится ряд $\sum\limits_{n=1}^{\infty} \mathsf{P}\left(\left|\frac{\widetilde{S}_{k_n} - \mathsf{E}\widetilde{S}_{k_n}}{k_n}\right| > \varepsilon\right) < \infty$. Применяя неравенство Чебышёва, получаем $\sum\limits_{n=1}^{\infty} \mathsf{P}\left(\left|\frac{\widetilde{S}_{k_n} - \mathsf{E}\widetilde{S}_{k_n}}{k_n}\right| > \varepsilon\right) \leqslant \sum\limits_{n=1}^{\infty} \frac{\mathsf{D}\widetilde{S}_{k_n}}{\varepsilon^2 k_n^2}$. В силу некоррелированности имеем $\mathsf{D}\widetilde{S}_{k_n} = \sum\limits_{i=1}^{k_n} \mathsf{D}Y_i$, поэтому

$$\sum_{n=1}^{\infty} \mathsf{P}\left(\left|\frac{\widetilde{S}_{k_n} - \mathsf{E}\widetilde{S}_{k_n}}{k_n}\right| > \varepsilon\right) \leqslant \sum_{n=1}^{\infty} \frac{\mathsf{D}\widetilde{S}_{k_n}}{\varepsilon^2 k_n^2} = \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{k_n^2} \sum_{i=1}^{k_n} \mathsf{D}Y_i \leqslant \frac{1}{\varepsilon^2} \sum_{i=1}^{\infty} \mathsf{E}Y_i^2 \sum_{n:\, k_n \geqslant i} \frac{1}{k_n^2}.$$

В последнем неравенстве мы изменили порядок суммирования и воспользовались оценкой $\mathsf{D}Y_n \leqslant \mathsf{E}Y_n^2$. Далее, оцениваем последнюю сумму в правой части: $\sum\limits_{n:[\alpha^n] \geqslant i} \frac{1}{[\alpha^n]^2} \leqslant \sum\limits_{n:[\alpha^n] \geqslant i} \frac{1}{c_\alpha \alpha^{2n}} = \frac{1}{c_\alpha} \sum\limits_{n:[\alpha^n] \geqslant i} \frac{1}{\alpha^{2n}}$, т.к. $[\alpha^n]^2 \geqslant c_\alpha \alpha^{2n}$ для некоторой константы c_α. Так как $\sum\limits_{n \geqslant m} \frac{1}{q^{2n}} = \frac{\frac{1}{q^{2m}}}{1 - \frac{1}{q^2}} = \frac{1}{q^{2(m-1)}(1-q^2)}$, то $\sum\limits_{n:\, [\alpha^n] \geqslant i} \frac{1}{\alpha^{2n}} = \sum\limits_{n \geqslant [\log_\alpha(i+1)]} \frac{1}{\alpha^{2n}} \leqslant \frac{1}{\alpha^{2\log_\alpha(i+1)}\left(\frac{1-\alpha^2}{\alpha^2}\right)} = \widetilde{c}_\alpha \frac{1}{(i+1)^2}$.

Итак, мы доказали, что $\sum\limits_{n=1}^{\infty} \mathsf{P}\left(\left|\frac{\widetilde{S}_{k_n} - \mathsf{E}\widetilde{S}_{k_n}}{k_n}\right| > \varepsilon\right) \leqslant \frac{C(\alpha)}{\varepsilon^2} \sum\limits_{i=1}^{\infty} \frac{\mathsf{E}Y_i^2}{(i+1)^2}$. Далее, по формуле математического ожидания $\mathsf{E}f(X) = \int\limits_{-\infty}^{\infty} f(x) \mathsf{P}_X(dx)$ и учитывая, что $\mathsf{P}_{X_i} = \mathsf{P}_{X_1}$ $\forall i$ (так как по условию $X_1, X_2, \ldots$ одинаково распределены) имеем $\sum\limits_{i=1}^{\infty} \frac{\mathsf{E}Y_i^2}{(i+1)^2} = \sum\limits_{i=1}^{\infty} \frac{1}{(i+1)^2} \int\limits_0^i x^2 \mu(dx)$, где $\mu = \mathsf{P}_{X_1}$ (в нашем случае функция f имеет вид $f(x) = x^2 \mathbb{I}_{\{0 \leqslant x \leqslant i\}}$). Применяя аддитивность интеграла и меняя порядок суммирования, получаем:

$$\sum_{i=1}^{\infty} \frac{\mathsf{E}Y_i^2}{(i+1)^2} = \sum_{i=1}^{\infty} \frac{1}{(i+1)^2} \int\limits_0^i x^2 \mu(dx) = \sum_{i=1}^{\infty} \frac{1}{(i+1)^2} \sum_{k=0}^{i-1} \int\limits_k^{k+1} x^2 \mu(dx) = \sum_{k=0}^{\infty} \int\limits_k^{k+1} x^2 \mu(dx) \sum_{i \geqslant k+1} \frac{1}{(i+1)^2}.$$

Далее, воспользуемся известной из математического анализа оценкой $\sum\limits_{n=m}^{\infty} f(n) \leqslant \int\limits_{m-1}^{\infty} f(x)\, dx$; получим: $\sum\limits_{i \geqslant k+1} \frac{1}{(i+1)^2} = \sum\limits_{j \geqslant k+2} \frac{1}{j^2} \leqslant \int\limits_{k+1}^{\infty} \frac{dx}{x^2} = \frac{1}{k+1}$. Следовательно, $\sum\limits_{k=0}^{\infty} \int\limits_k^{k+1} x^2 \mu(dx) \sum\limits_{i \geqslant k+1} \frac{1}{(i+1)^2} \leqslant \sum\limits_{k=0}^{\infty} \frac{1}{k+1} \int\limits_k^{k+1} x^2\, \mu(dx) \leqslant \sum\limits_{k=0}^{\infty} \int\limits_k^{k+1} x\, \mu(dx) = \mathsf{E}X_1 < \infty$ (по условию). Подставляя эту оценку в исходную, получаем $\sum\limits_{n=1}^{\infty} \mathsf{P}\left(\left|\frac{\widetilde{S}_{k_n} - \mathsf{E}\widetilde{S}_{k_n}}{k_n}\right| > \varepsilon\right) \leqslant \frac{C(\alpha)\mathsf{E}X_1}{\varepsilon^2} < \infty$ $\forall \varepsilon > 0$. Значит, по лемме Бореля–Кантелли $\frac{\widetilde{S}_{k_n} - \mathsf{E}\widetilde{S}_{k_n}}{k_n} \xrightarrow{\text{п.н.}} 0$, $n \to \infty$, что и требовалось. $\blacksquare$

Вернемся к доказательству теоремы. Мы имеем $\frac{\mathsf{E}\widetilde{S}_{k_n}}{k_n} = \frac{\sum_{i=1}^{k_n} \mathsf{E}Y_i}{k_n}$. Покажем, что $\mathsf{E}Y_i = \mathsf{E}X_i \mathbb{I}\{0 \leqslant X_i \leqslant i\} \to \mathsf{E}X_1$, $i \to \infty$. Действительно, $\mathsf{E}X_i \mathbb{I}\{0 \leqslant X_i \leqslant i\} = \int\limits_0^i x\, \mu(dx) = \int\limits_0^{\infty} x \mathbb{I}\{x \leqslant i\}\, \mu(dx) \to \int\limits_0^{\infty} x\, \mu(dx) = \mathsf{E}X_1$, $i \to \infty$ по теореме о монотонной сходимости. Следовательно, $\frac{\mathsf{E}\widetilde{S}_{k_n}}{k_n} \to \mathsf{E}X_1$, $n \to \infty$. А так как по лемме 4.5 $\frac{\widetilde{S}_{k_n} - \mathsf{E}\widetilde{S}_{k_n}}{k_n} \xrightarrow{\text{п.н.}} 0$, $n \to \infty$, то отсюда получаем $\frac{\widetilde{S}_{k_n}}{k_n} \xrightarrow{\text{п.н.}} \mathsf{E}X_1$, $n \to \infty$. Докажем, что отсюда вытекает $\frac{S_{k_n}}{k_n} \xrightarrow{\text{п.н.}} \mathsf{E}X_1$, $n \to \infty$. Действительно, так как $\sum\limits_i \mathsf{P}(X_i \neq Y_i) = \sum\limits_i \mathsf{P}(X_i > i) = \sum\limits_i \mathsf{P}(X_1 \geqslant i) \overset{\text{лемма}}{<} \mathsf{E}X_1 + 1 < \infty$, то по лемме Бореля–Кантелли $\mathsf{P}(X_i \neq Y_i) = 0$. Следовательно, для почти всех $\omega \in \Omega$ и $\forall n > N(\omega)$ $X_i(\omega) = Y_i(\omega) \implies \frac{1}{k_n} \sum\limits_{i=1}^{k_n} X_i \xrightarrow{\text{п.н.}} \mathsf{E}X_1$, $n \to \infty$.

Наконец, докажем, что $\frac{S_n}{n} \xrightarrow{\text{п.н.}} \mathsf{E}X_1$. Фиксируем $\alpha > 1$. Для каждого $n \geqslant N(\alpha)$ найдем такое $m(n)$, что $[\alpha^{m(n)}] \leqslant n < [\alpha^{m(n)+1}]$. Так как все $X_i \geqslant 0$, то $\varliminf\limits_n \frac{S_n}{n} \geqslant \varliminf\limits_n \frac{S_{[\alpha^{m(n)}]}}{n} = \varliminf\limits_n \frac{S_{[\alpha^{m(n)}]}}{[\alpha^{m(n)}]} \frac{[\alpha^{m(n)}]}{n} \geqslant \mathsf{E}X_1 \frac{1}{\alpha}$, так как

$\dfrac{[\alpha^{m(n)}]}{n} \geqslant \dfrac{[\alpha^{m(n)}]}{[\alpha^{m(n)+1}]} \longrightarrow \dfrac{1}{\alpha}$, $n \longrightarrow \infty$. Аналогично, $\overline{\lim\limits_n} \dfrac{S_n}{n} \leqslant \overline{\lim\limits_n} \dfrac{S_{[\alpha^{m(n)+1}]}}{n} = \overline{\lim\limits_n} \dfrac{S_{[\alpha^{m(n)+1}]}}{[\alpha^{m(n)+1}]} \dfrac{[\alpha^{m(n)+1}]}{n} \leqslant \mathsf{E}X_1 \cdot \alpha$ (воспользовались тем, что $\dfrac{[\alpha^{m(n)+1}]}{n} \leqslant \dfrac{[\alpha^{m(n)+1}]}{[\alpha^{m(n)}]} \longrightarrow \alpha$, $n \longrightarrow \infty$). Итак, $\dfrac{1}{\alpha}\mathsf{E}X_1 \leqslant \underline{\lim\limits_n} \dfrac{S_n}{n} \leqslant \overline{\lim\limits_n} \dfrac{S_n}{n} \leqslant \alpha \mathsf{E}X_1$, где $\alpha > 1$ — любое. Устремим $\alpha \Downarrow 1$, тогда получим $\lim\limits_n \frac{S_n}{n} = \mathsf{E}X_1$ п.н., что и требовалось доказать. ■

Следствием доказанной теоремы является важная классическая теорема Колмогорова:

Следствие 4.2 (Колмогоров). *Пусть X_1, X_2, … — независимые в совокупности одинаково распределенные случайные величины такие, что $\mathsf{E}|X_1| < \infty$. Тогда $\dfrac{S_n}{n} \xrightarrow{\text{п.н.}} \mathsf{E}X_1$, $n \longrightarrow \infty$, где $S_n = X_1 + \ldots + X_n$.*

4.3.3. Классическая теорема Колмогорова. Закон 0 и 1

Теорема 4.6 (Закон 0 или 1 Колмогорова). *Пусть ξ_1, ξ_2, … — независимые случайные величины и пусть $\mathscr{F}_n = \sigma\{\xi_n, \xi_{n+1}, \ldots\}$. Обозначим $\mathscr{F}_\infty = \bigcap\limits_n \mathscr{F}_n$. Тогда если $A \in \mathscr{F}_\infty$, то $\mathsf{P}(A) = 0$ или $\mathsf{P}(A) = 1$.*

□ Пусть $A \in \mathscr{F}_\infty$. Тогда $A \in \mathscr{F}_n$ $\forall n$. Обозначим $\mathscr{A}_{nk} = \sigma\{\xi_n, \ldots, \xi_{n+k}\}$, тогда $\mathscr{F}_n = \sigma\Big\{\bigcup\limits_{k \geqslant 1} \mathscr{A}_{nk}\Big\}$. Но $\mathscr{A}_n = \bigcup\limits_{k \geqslant 1} \mathscr{A}_{nk}$ — это алгебра, следовательно, $\forall \varepsilon > 0$ $\exists A_\varepsilon \in \mathscr{A}_{nk} = \sigma\{\xi_n, \ldots, \xi_{n+k}\}:$ $\mathsf{P}(A \triangle A_\varepsilon) < \varepsilon$ (см. задачу такую-то). Но $A \in \mathscr{F}_{n+k+1} = \sigma\{\xi_{n+k+1}, \ldots\}$, поэтому A и A_ε — независимые события, т.е. $\mathsf{P}(A \cap A_\varepsilon) = \mathsf{P}(A)\mathsf{P}(A_\varepsilon)$. Отсюда получаем, что $\mathsf{P}(A) = \mathsf{P}(A)^2$, откуда либо $\mathsf{P}(A) = 0$, либо $\mathsf{P}(A) = 1$, что и требовалось доказать. ■

5. Центральная предельная теорема

5.1. Теорема Александрова

5.2. Характеристические функции

5.2.1. Определение, основные свойства и примеры

Пусть Q — вероятностная мера на $\mathscr{B}(\mathbb{R})$.

Определение. Характеристической функцией вероятностной меры Q называется функция

$$\varphi_{\mathsf{Q}}(t) := \int\limits_{\mathbb{R}} e^{itx} \mathsf{Q}(dx) = \int\limits_{\mathbb{R}} \cos(tx) \mathsf{Q}(dx) + i \int\limits_{\mathbb{R}} \sin(tx) \mathsf{Q}(dx).$$

Если $X\colon \Omega \longrightarrow \mathbb{R}$ — случайная величина, то на $\mathscr{B}(\mathbb{R})$ возникает вероятностная мера P_X, определяемая по формуле $\mathsf{P}_X(B) = \mathsf{P}(X^{-1}(B))$. Тогда характеристической функцией случайной величины X называется характеристическая функция вероятностной меры P_X:

$$\varphi_{\mathsf{P}_X}(t) = \int\limits_{\mathbb{R}} e^{itx} \mathsf{P}_X(dx) = \int\limits_{\Omega} e^{itX}\, d\mathsf{P} = \mathsf{E}e^{itX} = \varphi_X(t).$$

Таким образом, характеристическая функция случайной величины X записывается в виде $\varphi_X(t) = \mathsf{E}e^{itX}$.

Имеется взаимно-однозначное соответствие между вероятностными мерами Q на Ω, их функциями распределения F и характеристическими функциями φ: $\mathsf{Q} \leftrightarrow F \leftrightarrow \varphi$.

Приведем несколько примеров характеристических функций.

Пример 2.1. Пусть $X(\omega) = C = \mathrm{const}$ $\forall\, \omega \in \Omega$ — постоянная случайная величина. Тогда ее характеристическая функция: $\varphi_X(t) = \mathsf{E}e^{itX} = \mathsf{E}e^{itC} = e^{itC}$.

Пример 2.2. Пусть $X \sim \pi_\lambda$ (пуассоновское распределение с параметром $\lambda > 0$), $X = \left\{ \begin{array}{cccc} 0, & \ldots, & k, & \ldots \\ e^{-\lambda}, & \ldots, & \frac{\lambda^k e^{-\lambda}}{k!} & \ldots \end{array} \right.$

Найдем характеристическую функцию этого распределения:

$$\varphi_X(t) = \mathsf{E}e^{itX} = \sum_{k=0}^{\infty} e^{itk} \mathsf{P}(X = k) = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it} - 1)}.$$

Итак, характеристическая функция пуассоновской случайной величины $X \sim \pi_\lambda$ с параметром $\lambda > 0$ равна $\varphi_X(t) = e^{\lambda(e^{it} - 1)}$.

Пример 2.3. Пусть случайная величина $X \sim \mathcal{N}(0,1)$ имеет стандартное нормальное распределение. Тогда, по определению, его плотность равна $p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $x \in \mathbb{R}$. Характеристическая функция равна $\varphi(t) = \varphi_X(t) = \int\limits_{-\infty}^{\infty} e^{itx}\, dF(x) = \int\limits_{-\infty}^{\infty} e^{itx} p(x)\, dx = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{itx - \frac{x^2}{2}}\, dx$. Легко проверить, что эта функция от t удовлетворяет дифференциальному уравнению $\varphi'(t) = -t\varphi(t)$ (продифференцируйте интеграл по параметру t и убедитесь в этом!), решая которое, получаем решение $\varphi(t) = e^{-\frac{t^2}{2}}$ с начальным условием $\varphi(0) = 1$, которое должно быть выполнено для любой характеристической функции.

Итак, характеристическая функция случайной величины $X \sim \mathcal{N}(0,1)$, имеющей стандартное нормальное распределение, равна $\varphi_X(t) = e^{-\frac{t^2}{2}}$.

Теорема 5.1 (Свойства характеристических функций). *Все характеристические функции обладают следующими свойствами:*

1) $\varphi(0) = 1$*;*

2) $|\varphi(t)| \leqslant 1$*;*

3) φ *равномерно непрерывна на* $\mathbb{R}$*;*

4) $\varphi(-t) = \overline{\varphi(t)}$*;*

5) X *имеет симметричное распределение (т.е.* $\mathrm{Law}(X) = \mathrm{Law}(-X)$*) тогда и только тогда, когда* φ_X *— действительная функция;*

6) Если $Y = aX + b$, $a, b \in \mathbb{R}$, *то* $\varphi_Y(t) = e^{itb}\varphi_X(at)$.

$\square$ Свойство 1) очевидно: $\varphi_X(0) = \mathsf{E}e^0 = 1$. Свойство 2): $|\varphi_X(t)| = \left|\int\limits_{-\infty}^{\infty} e^{itx}\, dF(x)\right| \leqslant \int\limits_{-\infty}^{\infty} |e^{itx}|\, dF(x) = \int\limits_{-\infty}^{\infty} dF(x) = 1$, так как $|e^{itx}| = 1$.

Докажем свойство 3) — равномерную непрерывность. Имеем:

$$\varphi(t+h) - \varphi(t) = \int\limits_{-\infty}^{\infty} e^{itx}(e^{ihx} - 1)\, dF(x), \quad |\varphi(t+h) - \varphi(t)| \leqslant \int\limits_{-\infty}^{\infty} |e^{ihx} - 1|\, dF(x) \quad \forall\, x \in \mathbb{R}.$$

Так как $|e^{ihx} - 1| \longrightarrow 0$, $h \longrightarrow 0$ $\forall\, x \in \mathbb{R}$, то по теореме о мажорируемой сходимости получаем требуемое.

Докажем свойство 4). Вспомним, что комплексно сопряженное число к $z = u + iv$ $(u, v \in \mathbb{R})$ — это число $\bar{z} = u - iv$. Имеем: $\varphi(-t) = \int\limits_{-\infty}^{\infty} e^{-itx}\, dF(x) = \overline{\int\limits_{-\infty}^{\infty} e^{itx}\, dF(x)} = \overline{\varphi(t)}$.

Свойство 5) сразу следует из свойства 4): действительно, пусть φ — действительная функция. Тогда $\varphi_{-X}(t) = \mathsf{E}e^{it(-X)} = \mathsf{E}e^{i(-t)X} = \varphi_X(-t) = \overline{\varphi_X(t)} = \varphi_X(t)$, следовательно $\varphi_{-X}(t) = \varphi_X(t)$ $\forall\, t \in \mathbb{R}$. По теореме единственности (формула обращения; см. следующий пункт) $\mathrm{Law}(-X) = \mathrm{Law}(X)$. Производя те же выкладки в обратном порядке, получаем обратное утверждение (достаточность).

Осталось доказать свойство 6). Имеем: $\varphi_{aX+b}(t) = \mathsf{E}e^{it(aX+b)} = \mathsf{E}e^{i(at)X}e^{itb} = e^{itb}\varphi_X(at)$, что и требовалось доказать. $\blacksquare$

Доказанная теорема дает необходимые условия, которым должны удовлетворять характеристические функции произвольных случайных величин. В частности, если некоторая (данная) функция φ не удовлетворяет хотя бы одному из этих условий, то теорема позволяет сделать вывод о том, что эта функция не является характеристической ни для какой случайной величины X. Если же все свойства для функции φ выполнены, то этот вопрос остается открытым.

Следующая теорема дает необходимое и достаточное условие того, что функция φ является характеристической для некоторой случайной величины X. Чтобы сформулировать ее, дадим следующее

Определение. Функция $\varphi(t)$ называется неотрицательно определенной, если $\forall\, n \in \mathbb{N}$, $\forall\, t_1, \ldots, t_n \in \mathbb{R}$ и $\forall\, z_1, \ldots, z_n \in \mathbb{C}$

$$\sum_{k=1}^{n} \sum_{l=1}^{n} \varphi(t_k - t_l) z_k \bar{z}_l \geqslant 0.$$

Теорема 5.2 (Бохнер–Хинчин). *Функция* $\varphi(t)$, $t \in \mathbb{R}$ *является характеристической функцией для некоторой случайной величины* X *тогда и только тогда, когда выполнены следующие три условия:*

1) $\varphi(0) = 1$*;*

2) φ *непрерывна в нуле;*

3) φ *является неотрицательно определенной функцией.*

$\square$ Мы докажем только необходимость. Пусть φ — характеристическая функция. Тогда 1) и 2), очевидно, выполнены. Докажем 3), т.е. проверим, что φ — неотрицательно определенная функция. Имеем: $\sum\limits_{k=1}^{n}\sum\limits_{l=1}^{n}\varphi(t_k - t_l)z_k\bar{z}_l = \sum\limits_{k=1}^{n}\sum\limits_{l=1}^{n}\left(\mathsf{E}e^{i(t_k-t_l)X}\right)z_k\bar{z}_l = \sum\limits_{k=1}^{n}\sum\limits_{l=1}^{n}\mathsf{E}(z_k e^{it_kX})(\bar{z}_l e^{it_lX}) = \mathsf{E}\left(\sum\limits_{k=1}^{n} z_k e^{it_kX}\right)\left(\sum\limits_{l=1}^{n}\bar{z}_l e^{it_lX}\right) = \mathsf{E}\omega\bar{\omega} = \mathsf{E}|\omega|^2 \geqslant 0$ (применили линейность мат. ожидания), что и требовалось. $\blacksquare$

5.2.2. Формула обращения

Теорема 5.3 (Формула обращения). *1) Для любых точек $a, b \in \mathscr{C}(F)$ имеет место формула*

$$F(b) - F(a) = \lim_{c\longrightarrow\infty}\frac{1}{2\pi}\int\limits_{-c}^{c}\frac{e^{-ita} - e^{-itb}}{it}\varphi(t)\,dt.$$

2) Если $\varphi \in \mathscr{L}^1(\mathbb{R})$, т.е. $\int\limits_{-\infty}^{\infty}|\varphi(t)|\,dt < \infty$, то F абсолютно непрерывна и имеет плотность

$$p(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}e^{-itx}\varphi(t)\,dt, \quad x \in \mathbb{R}; \quad p \in \mathscr{L}^1(\mathbb{R}), \quad p \geqslant 0.$$

$\square$ По формуле для характеристической функции $\varphi(t)$ имеем:

$$\frac{1}{2\pi}\int\limits_{-c}^{c}\frac{e^{-ita} - e^{-itb}}{it}\varphi(t)\,dt = \frac{1}{2\pi}\int\limits_{-c}^{c}\left(\frac{e^{-ita} - e^{-itb}}{it}\int\limits_{-\infty}^{\infty}e^{itx}\,dF(x)\right)dt.$$

Изменим в последнем интеграле порядок интегрирования. Для этого применим теорему Фубини. Чтобы ее условия были выполнены, необходимо, чтобы $\int\limits_{-c}^{c}\int\limits_{\mathbb{R}}\left|\frac{e^{-ita} - e^{-itb}}{it}e^{itx}\right|dF(x)\,dt < \infty$. Так как $|e^{i\alpha}| = 1\ \ \forall\,\alpha \in \mathbb{R}$ и $\left|\frac{e^{-ita}-e^{-itb}}{it}\right| = \left|\int\limits_a^b e^{-itx}\,dx\right| \leqslant \int\limits_a^b|e^{-itx}|\,dx = \int\limits_a^b dx = b - a$, то имеет место оценка $\int\limits_{-c}^{c}\int\limits_{\mathbb{R}}\left|\frac{e^{-ita} - e^{-itb}}{it}e^{itx}\right|dF(x)\,dt < 2c(b-a) < \infty$. Итак, можно поменять порядок интегрирования; получим:

$$\frac{1}{2\pi}\int\limits_{-c}^{c}\frac{e^{-ita} - e^{-itb}}{it}\varphi(t)\,dt = \frac{1}{2\pi}\int\limits_{-c}^{c}\left(\frac{e^{-ita} - e^{-itb}}{it}\int\limits_{-\infty}^{\infty}e^{itx}\,dF(x)\right)dt = \int\limits_{-\infty}^{\infty}dF(x)\left(\frac{1}{2\pi}\int\limits_{-c}^{c}\frac{e^{-it(x-a)} - e^{-it(x-b)}}{it}\,dt\right) = (*)$$

Так как $e^{it\alpha} = \cos\alpha + i\sin\alpha$ и $\int\limits_{-c}^{c}\frac{\cos t(x-a)}{t}\,dt = 0$ (подынтегральная функция нечетна), то отсюда получаем:

$$(*) = \int\limits_{-\infty}^{\infty}dF(x)\left\{\frac{1}{2\pi}\int\limits_{-c}^{c}\frac{\sin t(x-a) - \sin t(x-b)}{t}\,dt\right\} = \int\limits_{-\infty}^{\infty}\frac{1}{\pi}\left\{\int\limits_{0}^{c(x-a)}\frac{\sin z}{z}\,dz - \int\limits_{0}^{c(x-b)}\frac{\sin z}{z}\,dz\right\}dF(x) = \int\limits_{-\infty}^{\infty}\psi_c(x)\,dF(x).$$

Функция $f(z) = \frac{\sin z}{z}$ не интегрируема по Лебегу на $(0,\infty)$, но интегрируема по Риману на этом множестве в несобственном смысле (интеграл Дирихле): $\int\limits_0^{\infty}\frac{\sin z}{z}\,dz = \frac{\pi}{2}$. Поэтому здесь мы понимаем этот интеграл как несобственный интеграл Римана, т.е. как предел $\lim\limits_{u\longrightarrow\infty}\int\limits_0^u\frac{\sin z}{z}\,dz$. Так как функция $\int\limits_0^u\frac{\sin z}{z}\,dz$ непрерывна по u, то $\exists\,A : \left|\int\limits_0^u\frac{\sin z}{z}\,dz\right| \leqslant A = \mathrm{const}\ \ \forall\,u \geqslant 0$. Значит, $|\psi_c(x)| \leqslant 2A\ \ \forall\,x$ и

$$\psi_c(x) \longrightarrow \psi(x) = \begin{cases} 0 & \text{если } x < a \text{ или } x > b,\\ \frac{1}{2} & \text{если } x = a \text{ или } x = b,\\ 1 & \text{если } a < x < b.\end{cases} \qquad c \longrightarrow \infty.$$

По теореме о мажорируемой сходимости получаем $\int\limits_{-\infty}^{\infty} \psi_c(x)\,dF(x) \longrightarrow \int\limits_{-\infty}^{\infty} \psi(x)\,dF(x) = \int\limits_{-\infty}^{\infty} \psi(x)\,d\mathsf{Q}(x) = \frac{\mathsf{Q}(\{a\})}{2} + \frac{\mathsf{Q}(\{b\})}{2} + F(b-) - F(a) = \frac{F(a)-F(a-)}{2} + \frac{F(b)-F(b-)}{2} + F(b-) - F(a) = \frac{F(b)+F(b-)}{2} - \frac{F(a)+F(a-)}{2}$. Если $a, b \in \mathscr{C}(F)$, то $\frac{F(a)+F(a-)}{2} = F(a)$, $\frac{F(b)+F(b-)}{2} = F(b)$, откуда и следует утверждение 1) теоремы.

Докажем утверждение 2). А именно, докажем, что функция $p(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-itx}\varphi(t)\,dt$, $x \in \mathbb{R}$ является плотностью нашего распределения, $F(x) = \int\limits_{-\infty}^{x} p(u)\,du$. Функция $p(x)$ является непрерывной на $\mathbb{R}$ по теореме о мажорируемой сходимости, $p(x+h) - p(x) = \int\limits_{\mathbb{R}} (e^{-i(x+h)t} - e^{-ixt})\varphi(t)\,dt$, $|e^{-itx}| = 1$.

Применим теорему Фубини: $\forall\, a < b$ имеем

$$\int\limits_a^b p(x)\,dx = \int\limits_a^b \left(\frac{1}{2\pi}\int\limits_{\mathbb{R}} e^{-itx}\varphi(t)\,dt\right)dx = \frac{1}{2\pi}\int\limits_{\mathbb{R}}\left(\int\limits_a^b e^{-itx}\varphi(t)\,dx\right)dt = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} \varphi(t)\frac{e^{-ita}-e^{-itb}}{it}\,dt = (*)$$

По теореме о мажорируемой сходимости и применяя уже доказанный пункт 1) теоремы, получаем:

$$(*) = \frac{1}{2\pi}\lim_{c\longrightarrow\infty}\int\limits_{-c}^{c} \varphi(t)\frac{e^{-ita}-e^{-itb}}{it}\,dt = F(b) - F(a), \quad a, b \in \mathscr{C}(F).$$

Итак, мы получили, что $p(x)$ — непрерывная функция, такая что $\int\limits_a^b p(x)\,dx = F(b) - F(a)$, $\forall\, a, b \in \mathscr{C}(F)$. Следовательно, $p(x) \geqslant 0$ $\forall\, x \in \mathbb{R}$. Устремим $a \longrightarrow -\infty$, тогда получим: $\int\limits_{-\infty}^{b} p(x)\,dx = F(b)$ $\forall\, b \in \mathscr{C}(F)$. Этот интеграл можно понимать как интеграл Лебега (так как $p(x) \geqslant 0$). Итак, $F(x) = \int\limits_{-\infty}^{x} p(u)\,du$ $\forall\, x \in \mathbb{R}$, что и требовалось. ■

5.2.3. Теорема Леви

Теорема 5.4 (Леви) (Теорема непрерывности). *1) Пусть последовательность функций распределения $F_n(x)$ сходится к $F(x)$ $\forall\, x \in \mathscr{C}(F)$: $F_n(x) \xrightarrow{\omega} F(x)$. Тогда $\varphi_n(t) \longrightarrow \varphi(t)$ $\forall\, t \in \mathbb{R}$, где $F_n \leftrightarrow \varphi_n$, $F \leftrightarrow \varphi$.*

2) Пусть характеристические функции $\varphi_n(t) \longrightarrow \varphi(t)$, $n \longrightarrow \infty$, причем φ непрерывна в 0. Тогда φ является характеристической функцией, и $F_n \xrightarrow{\omega} F$, где $F_n \leftrightarrow \varphi_n$, $F \leftrightarrow \varphi$.

Пример 2.4. Докажем, что $\displaystyle\sum_{\lambda+a\sqrt{\lambda}\leqslant k\leqslant\lambda+b\sqrt{\lambda}} \frac{\lambda^k e^{-\lambda}}{k!} \longrightarrow \frac{1}{\sqrt{2\pi}}\int\limits_a^b e^{-\frac{x^2}{2}}\,dx, \quad \lambda \longrightarrow \infty.$

Рассмотрим пуассоновскую случайную величину $Z_\lambda \sim \pi_\lambda$. Тогда

$$\sum_{\lambda+a\sqrt{\lambda}\leqslant k\leqslant\lambda+b\sqrt{\lambda}} \frac{\lambda^k e^{-\lambda}}{k!} = \sum_{\lambda+a\sqrt{\lambda}\leqslant k\leqslant\lambda+b\sqrt{\lambda}} \mathsf{P}(Z_\lambda = k) = \mathsf{P}\Big(Z_\lambda \in [\lambda + a\sqrt{\lambda}, \lambda + b\sqrt{\lambda}]\Big) = \mathsf{P}\left(a \leqslant \frac{Z_\lambda - \lambda}{\sqrt{\lambda}} \leqslant b\right).$$

Вспоминая свойство 6) характеристических функций и выражение для характеристической функции пуассоновской случайной величины $\Big(\varphi_{Z_\lambda}(s) = e^{\lambda(e^{is}-1)}\Big)$, получаем:

$$\varphi_{\frac{Z_\lambda-\lambda}{\sqrt{\lambda}}}(t) = \varphi_{\frac{1}{\sqrt{\lambda}}Z_\lambda - \sqrt{\lambda}}(t) = e^{-it\sqrt{\lambda}}\varphi_{Z_\lambda}\left(\frac{t}{\sqrt{\lambda}}\right) = e^{-it\sqrt{\lambda}}e^{\lambda\left(e^{i\frac{t}{\sqrt{\lambda}}}-1\right)} = e^{\lambda\left(e^{i\frac{t}{\sqrt{\lambda}}}-1-\frac{it}{\sqrt{\lambda}}\right)} \longrightarrow e^{-\frac{t^2}{2}} \quad \forall\, t, \ \lambda \longrightarrow \infty,$$

так как $e^w = 1 + w + \frac{w^2}{2} + o(w^2)$, $w \longrightarrow 0$. Итак, $\varphi_{\frac{Z_\lambda-\lambda}{\sqrt{\lambda}}}(t) \longrightarrow \varphi_X(t) = e^{-\frac{t^2}{2}}$, $\lambda \longrightarrow \infty$, $\forall\, t \in \mathbb{R}$, поэтому по теореме непрерывности $\frac{Z_\lambda-\lambda}{\sqrt{\lambda}} \xrightarrow{\mathcal{D}} X \sim \mathcal{N}(0,1)$, значит $\forall\, a < b$ будет $\mathsf{P}\left(a \leqslant \frac{Z_\lambda-\lambda}{\sqrt{\lambda}} \leqslant b\right) \longrightarrow \mathsf{P}\left(a \leqslant X \leqslant b\right) = \frac{1}{\sqrt{2\pi}}\int\limits_a^b e^{-\frac{x^2}{2}}\,dx$, $\lambda \longrightarrow \infty$, так как функция распределения случайной величины $X \sim \mathcal{N}(0,1)$ непрерывна на $\mathbb{R}$.

Доказательство теоремы непрерывности.

□ 1) Заметим сначала, что если $F_n \longrightarrow F$, то $\int\limits_{\mathbb{R}} e^{itx}\,dF_n(x) \longrightarrow \int\limits_{\mathbb{R}} e^{itx}\,dF(x)$, поскольку e^{itx} — непрерывная ограниченная функция.

Лемма 5.5. *Пусть имеется соответствие функции распределения, вероятностной меры и характеристической функции: $F \leftrightarrow \mathsf{Q} \leftrightarrow \varphi$. Тогда $\forall\, a > 0$*

$$\mathsf{Q}\left(\mathbb{R}\backslash\left(-\frac{1}{a}, \frac{1}{a}\right)\right) = \int\limits_{\mathbb{R}\backslash\left(-\frac{1}{a}, \frac{1}{a}\right)} dF(x) \leqslant \frac{C}{a}\int\limits_0^a (1 - \operatorname{Re}\varphi(t))\, dt, \text{ где } C = \text{const}.$$

□ $\int\limits_0^a (1 - \operatorname{Re}\varphi(t))\, dt = \frac{1}{a}\int\limits_0^a \left(\int\limits_{-\infty}^{\infty} (1 - \cos tx)\, dF(x)\right) dt = (*)$, так как $\operatorname{Re}\varphi(t) = \int\limits_{-\infty}^{\infty} \cos tx\, dF(x)$. По теореме Фубини $(*) = \int\limits_{-\infty}^{\infty} \left(1 - \frac{\sin ax}{ax}\right) dF(x) \geqslant \int\limits_{|ax| \geqslant 1} \left(1 - \frac{\sin ax}{ax}\right) dF(x) \geqslant \inf\limits_{u \geqslant 1} \left(1 - \frac{\sin u}{u}\right) \int\limits_{|x| \geqslant \frac{1}{a}} dF(x) = (1 - \sin 1) \int\limits_{|x| \geqslant \frac{1}{a}} dF(x)$. Полагая $1/C = 1 - \sin 1$, получаем нужное неравенство. Лемма доказана. ■

Лемма 5.6. *Пусть $\varphi_n(t) \longrightarrow \varphi(t)$, где $\varphi(t)$ непрерывна в нуле. Тогда семейство распределений Q_n ($\mathsf{Q}_n \leftrightarrow \varphi_n$) плотно, т.е. $\forall\, \varepsilon > 0\ \exists\, K_\varepsilon : \sup\limits_n \mathsf{Q}_n(\mathbb{R}\backslash K_\varepsilon) < \varepsilon$.*

□ По предыдущей лемме $\forall\, a > 0$ имеем $\mathsf{Q}_n\left(\mathbb{R}\backslash\left(-\frac{1}{a}, \frac{1}{a}\right)\right) \leqslant$ ■

■

5.3. Центральная предельная теорема в условиях Линдеберга

5.3.1. Ещё два (а может, три) свойства характеристических функций

5.4. Свёртки распределений

5.5. Случайные векторы

5.5.1. Основные определения

5.5.2. Многомерная центральная предельная теорема