# Теория вероятностей

(Ещё не)алгебраист

17 июня 2025 г.

## Предисловие

Эти записки созданы с целью аккуратно формализовать и заполнить пробелы в лекциях Е. Б. Яровой. Многие примеры, приведённые здесь взяты из лекций Елены Борисовны или были предложены А. Е. Кондратенко на семинарах. В разделе 0 будут содержаться основные принятые в курсе обозначения, а также сведения и определения из разных разделов математики, которыми автор будет пользоваться. Поскольку автор считает полезным взгляд на всякий раздел математики с точки зрения теории категорий и её приложений, этот язык также будет упоминаться (тем не менее, не замещая собой прочие подходы).

## Содержание


**0 Предварительные сведения** **4**

- 0.1 Обозначения . . . 4
- 0.2 Предварительные сведения из действительного анализа . . . 4
  - 0.2.1 Системы множеств и структуры на них . . . 4
  - 0.2.2 Минимальное кольцо и минимальная алгебра . . . 6
  - 0.2.3 Мера на полукольце и её продолжение на минимальное кольцо 7
  - 0.2.4 Лебеговское продолжение меры . . . 8
  - 0.2.5 Единственность продолжения меры на минимальную $\sigma$-алгебру 9
  - 0.2.6 Непрерывность и полнота меры . . . 11
  - 0.2.7 Мера Лебега-Стилтьеса . . . 12
  - 0.2.8 Измеримые функции . . . 13
  - 0.2.9 Интеграл Лебега . . . 16
  - 0.2.10 Прямой образ меры (pushforward measure) . . . 20
- 0.3 Теория категорий и взгляд на измеримые пространства с её точки зрения . . . 21
  - 0.3.1 Категория измеримых пространств . . . 22
  - 0.3.2 Прямой образ $\sigma$-алгебры . . . 22
  - 0.3.3 Обратный образ $\sigma$-алгебры . . . 23

0.3.4 Связь между минимальной $\sigma$-алгеброй, прямым и обратным образами $\sigma$-алгебры . . . 24
0.3.5 Произведение в категории измеримых пространств . . . 25
0.3.6 Функтор борелевской $\sigma$-алгебры . . . 26
0.4 Предварительные сведения из анализа Фурье . . . 27
0.5 Предварительные сведения из линейной алгебры . . . 27
0.5.1 Билинейные функции и квадратичные формы . . . 27
0.5.2 Полуторалинейные функции . . . 29
0.6 Элементарная комбинаторика . . . 30
0.6.1 Элементарные факты для подсчёта числа комбинаций . . . 30
0.6.2 Классические комбинаторные величины . . . 30
0.6.3 Свойства комбинаторных величин . . . 32

**1 Вероятностное пространство, случайные события** **33**

**2 Условные вероятности, формула Байеса, независимость событий** **37**
2.1 Условная вероятность . . . 37
2.2 Формула полной вероятности и формула Байеса . . . 38
2.3 Независимость событий . . . 42
2.4 Произведение вероятностных пространств . . . 44

**3 Случайные величины, их распределения, функции распределения и плотности** **45**
3.1 Случайные величины . . . 45
3.2 Функции распределения . . . 46
3.3 Разновидности распределений . . . 50
3.4 Преобразования случайных величин . . . 51

**4 Классические примеры распределений** **52**
4.1 Дискретные распределения. . . . 52
4.1.1 Распределение константы . . . 52
4.1.2 Распределение Бернулли . . . 52
4.1.3 Дискретное равномерное распределение . . . 53
4.1.4 Биномиальное распределение . . . 53
4.1.5 Распределение Пуассона . . . 54
4.1.6 Геометрическое распределение . . . 54
4.1.7 Гипергеометрическое распределение . . . 55
4.1.8 Отрицательное биномиальное распределение . . . 55
4.2 Абсолютно непрерывные случайные величины . . . 56
4.2.1 Равномерное распределение . . . 56
4.2.2 Экспоненциальное (показательное) распределение . . . 56
4.2.3 Нормальное распределение (распределение Гаусса) . . . 58
4.2.4 Распределение Коши . . . 58

4.3 Сингулярные распределения . . . . . . . . . . . . . . . . . . . . . . 59

**5 Совместные распределения случайных величин** **60**

**6 Независимость случайных величин** **62**
6.1 Независимость конечного набора случайных величин . . . . . . . . . 62
6.2 Последовательность независимых случайных величин . . . . . . . . 66

**7 Численные характеристики случайных величин** **66**
7.1 Математическое ожидание . . . . . . . . . . . . . . . . . . . . . . . 67
7.2 Моменты и абсолютные моменты случайной величины . . . . . . . . 71
7.3 Дисперсия, ковариация и корреляция . . . . . . . . . . . . . . . . . 74
7.4 Связь с независимостью случайных величин . . . . . . . . . . . . . 77
7.5 Квантили и медиана . . . . . . . . . . . . . . . . . . . . . . . . . . 78
7.6 Мода . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 78
7.7 Асимметрия и эксцесс . . . . . . . . . . . . . . . . . . . . . . . . . 79
7.8 Правило трёх $\sigma$-м . . . . . . . . . . . . . . . . . . . . . . . . . 80
7.9 Вычисление матожиданий и дисперсий для распределений, связанных с нормальным распределением . . . . . . . . . . . . . . . . . . 81
7.9.1 $\chi^2$-распределение . . . . . . . . . . . . . . . . . . . . . . . 81
7.9.2 Распределение Стьютенда . . . . . . . . . . . . . . . . . . . . 81
7.9.3 $F$-распределение . . . . . . . . . . . . . . . . . . . . . . . . 82
7.9.4 Логнормальное распределение . . . . . . . . . . . . . . . . . . 82
7.10 Вычисление матожиданий через совместные распределения . . . . . 82

**8 Свёртки случайных величин** **82**

**9 Сходимости случайных величин** **83**
9.1 Сходимость почти наверное . . . . . . . . . . . . . . . . . . . . . . 83
9.2 Сходимость по вероятности . . . . . . . . . . . . . . . . . . . . . . 83
9.3 Пространство $\mathcal{L}_p$ и сходимость в нём . . . . . . . . . . . . . . . 83
9.4 Сходимость по распределению . . . . . . . . . . . . . . . . . . . . . 84
9.5 Связь сходимостей . . . . . . . . . . . . . . . . . . . . . . . . . . . 84

**10 Производящие функции** **85**

**11 Характеристические функции** **85**

**12 Предельные теоремы** **88**
12.1 Закон нуля или единицы . . . . . . . . . . . . . . . . . . . . . . . 88
12.2 Неравенства . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 91
12.3 Закон больших чисел . . . . . . . . . . . . . . . . . . . . . . . . . 91
12.4 Центральная предельная теорема . . . . . . . . . . . . . . . . . . . 95
12.5 Теорема Муавра-Лапласа . . . . . . . . . . . . . . . . . . . . . . . 98

12.6 Закон повторного логарифма . . . . . . . . . . . . . . . . . . . . . . . . . . 100
12.7 Закон арксинуса . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 100

**13 Указатель терминов** **101**

**14 Указатель теорем** **101**

# 0 Предварительные сведения

## 0.1 Обозначения

- $\Omega$ — пространство элементарных исходов;
- $\omega$ — элементарный исход;
- $\mathfrak{F}$ — $\sigma$-алгебра событий;
- $\mathrm{P}$ — вероятностная мера;
- $\xi, \eta, \zeta$ — случайные величины;
- $\mathrm{E}\,\xi$ — математическое ожидание случайной величины $\xi$;
- $\mathrm{D}\,\xi$ — дисперсия случайной величины $\xi$;
- $\mathrm{Cov}(\xi, \eta)$ — ковариация случайных величин $\xi$ и $\eta$;
- $\rho(\xi, \eta)$ — корреляция случайных величин $\xi$ и $\eta$;

## 0.2 Предварительные сведения из действительного анализа

### 0.2.1 Системы множеств и структуры на них

Система множеств (следует понимать как синоним термина «семейство множеств») $S$ называется полукольцом, если она удовлетворяет следующим аксиомам:

(1) $\varnothing \in S$;

(2) $\forall A, B \in S : A \cap B \in S$;

(3) $\forall A, B \in S, A \subset B\ \exists n \in \mathbb{N}\ \exists C_1, \ldots, C_n \in S : A = B \sqcup \bigsqcup\limits_{k=1}^{n} C_k$.

Множество $\Omega \in U$ называется единицей системы множеств $U$, если всякий элемент $A \in U$ является подмножеством $\Omega$.

Система множеств $R$ называется кольцом, если она удовлетворяет следующим аксиомам:

(1) $\forall\ A, B \in R : A \cap B \in R$;

(2) $\forall\ A, B \in R : A \vartriangle B \in R$.

Следующее утверждение проверяется непосредственно, исходя из теоретико-множественных тождеств, но его доказательство приведено, например, в книге [2].

**Предложение 0.1.** *Пусть $R$ — кольцо. Тогда $R$ является полукольцом. Кроме того, для любых элементов $A, B \in R$ в $R$ также содержатся их объединение $A \cup B$ и разность $A \setminus B$.*

Кольцо называется $\sigma$-кольцом, если для любого счётного набора его элементов $\{A_k\}_{k \in R} \subset R$ их объединение содержится в $R$ ($\bigcup\limits_{k \in \mathbb{N}} A_k \in R$) и $\delta$-кольцом, если для любого счётного набора его элементов $\{A_k\}_{k \in R} \subset R$ их пересечение содержится в $R$.

Кольцо с единицей $\Omega$ называется алгеброй (подмножеств множества $\Omega$).

В книгах по теории вероятностей понятие алгебры часто вводится с использование другого равносильного набора аксиом, что выражает следующее

**Предложение 0.2** (Определение алгебры в традиции теории вероятностей)**.** *Система множеств $R$ является алгеброй подмножеств множества $\Omega$ тогда и только тогда, когда $R$ удовлетворяет следующим аксиомам*

(1) $\Omega \in R$;

(2) $\forall\ A, B \in R : A \cup B, A \cap B \in R$;

(3) $\forall\ A \in R :\ \Omega \setminus A := \overline{A} \in R$.

Мы снова опускаем доказательство, сводящееся к тождествам теории множеств.

Если алгебра является $\sigma$-кольцом или $\delta$-кольцом, то её называют $\sigma$-алгеброй или $\delta$-алгеброй, соответственно.

**Предложение 0.3.** *Имеет место следующее:*

(1) *Всякое $\sigma$-кольцо является $\delta$-кольцом, обратное вообще говоря не верно.*

(2) *Всякая $\sigma$-алгебра является $\delta$-алгеброй и наоборот.*

**Лемма 0.4.** *Пусть $R$ — ($\sigma$-)кольцо и $A \in R$. Тогда множество*

$$R \cap A := \{\, B \cap A |\, B \in R\}$$

*является ($\sigma$-)алгеброй подмножеств $A$. Также $R \cap A \subset R$.*

*Доказательство.* По построению $\Omega \cap A = A$ содержится в $R \cap A$ и всякий элемент $R \cap A$ есть подмножество $A$. Так как кольцо замкнуто относительно пересечений, то $R \cap A \subset R$.

Пусть теперь $C_1 = B_1 \cap A, C_2 = B_2 \cap A \in R \cap A$ — два множества. Тогда $C_1 \cap C_2 = (B_1 \cap B_2) \cap A \in R \cap A$, так как $B_1 \cap B_2 \in R$. Далее, $C_1 \cup C_2 = (B_1 \cup B_2) \cap A \in R \cap A$, так как $B_1 \cup B_2 \in R$. Окончательно, $A \setminus C_1 = (A \setminus B_1) = (\Omega \setminus B_1) \cap A \in R \cap A$, поскольку $\Omega \setminus B_1 \in R$.

Предположим, что $R$ являлось $\sigma$-алгеброй. Пусть $\{C_k\}$ — счётное семейство элементов $R \cap A$ и $C_k = B_k \cap A$. Тогда

$$\bigcup_{i=1}^{\infty} C_k = \bigcup_{i=1}^{\infty} (B_k \cap A) = \left( \bigcup_{i=1}^{\infty} B_k \right) \cap A \in R \cap A,$$

принадлежность справедлива в силу того, что $\bigcup\limits_{i=1}^{\infty} B_k \in R$. □

Можно показать, что кольцо множеств является кольцом в алгебраическом смысле этого слова с операциями сложения $\triangle$ и умножения $\cap$, а алгебра множеств является булевой алгеброй (в частности, $\mathbb{F}_2$-алгеброй).

### 0.2.2 Минимальное кольцо и минимальная алгебра

Следующее утверждение сводится к проверке аксиом кольца или алгебры, но его доказательство также можно прочитать в книге [2].

**Предложение 0.5.** *Пусть $\{R_\alpha\}_{\alpha \in \mathcal{A}}$ — семейство $(\sigma$-,$\delta$-$)$колец множеств. Тогда система $R = \bigcap\limits_{\alpha \in \mathcal{A}} R_\alpha$ является $(\sigma$-,$\delta$-$)$кольцом. Кроме того, если все кольца $R_\alpha$ являются $(\sigma$-$)$алгебрами подмножеств множества $\Omega$ (то есть у них есть общая единица), то $R$ также является $(\sigma$-$)$алгеброй подмножеств множества $\Omega$.*

**Теорема 0.6** (О минимальной $(\sigma$-$)$алгебре)**.** *Пусть $U$ — система множеств. Тогда существует как минимум одно $(\sigma$-$)$кольцо, содержащее $U$. Пересечение всех таких $(\sigma$-$)$колец $R(U)$ $(R_\sigma(U))$ само является $(\sigma$-$)$кольцом. Всякое $(\sigma$-$)$кольцо, содержащее $U$, содержит и $R(U)$. Если $\Omega \in U$ — единица $U$, то $R(U)$ $(R_\sigma(U))$ является $(\sigma$-$)$алгеброй подмножеств множества $\Omega$.*

*Доказательство.* В качестве $(\sigma$-$)$кольца, содержащего $U$ можно взять булеан $2^\Sigma$, где множество $\Sigma$ определено как объединение $\bigcup\limits_{A \in U} A$.

Пересечение всех таких $(\sigma$-$)$колец существует, поскольку имеется хотя бы одно кольцо и по предложению 0.5 это пересечение само является $(\sigma$-$)$кольцом.

Пусть $(\sigma$-$)$кольцо $R'$ содержит множество $U$. Тогда по построению $R(U)$ $(R_\sigma(U))$ содержится в пересечении $2^\Sigma \cap R'$, откуда $R(U) \subset R'$ $(R_\sigma(U) \subset R')$.

Если $\Omega$ — единица $U$, то по построению $\Sigma = \Omega$. Для всякого $(\sigma$-$)$кольца $R'$, содержащего $U$ имеем $\Omega \in R$ и по лемме 0.4 система множеств $R' \cap \Omega \subset R'$ является

($\sigma$-)алгеброй подмножеств $\Omega$. Так как $\Omega$ являлось единицей $U$, то $U$ содержится в $R' \cap \Omega$. Следовательно, достаточно рассматривать пересечение только ($\sigma$-)алгебр подмножеств множества $\Omega$, содержащих $U$. По предложению 0.5 их пересечение является ($\sigma$-)алгеброй подмножеств $\Omega$. $\square$

Опираясь на предложение 0.6 дадим определение. Для системы множеств $U$ пересечение всех колец, содержащих $U$ называется минимальным кольцом, порождённым $U$ и обозначается $R(U)$.

**Предложение 0.7.** *Пусть $S$ — полукольцо и $R(S)$ — минимальное кольцо, порождённое $S$. Тогда $R(S)$ допускает следующие описания*

- $R(S) = \{\, A_1 \cup \ldots \cup A_n |\, n \in \mathbb{N}, \{A_i\}_{i=1}^n \subset S \}\,;$
- $R(S) = \{\, A_1 \sqcup \ldots \sqcup A_n |\, n \in \mathbb{N}, \{A_i\}_{i=1}^n \subset S \}\,.$

*Доказательство.* НУЖНО: дописать доказательство или сослаться $\square$

### 0.2.3 Мера на полукольце и её продолжение на минимальное кольцо

Пусть $S$ — некоторое полукольцо. Будем называть неотрицательную функцию $m\colon S \to \mathbb{R}$ мерой на полукольце $S$, если $m$ удовлетворяет аксиоме аддитивности

$$\forall A, B \in S, A \cap B = \varnothing, A \cup B \in S : m(A \sqcup B) = m(A) + m(B).$$

Если дополнительно для любой последовательности попарно непересекающихся подмножеств $\{A_k\}_{k \in \mathbb{N}}$, объединение которых есть элемент из $S$ (отметим, что это автоматически выполнено, если $S$ является $\sigma$-кольцом) имеет место равенство

$$m\left(\bigsqcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} m(A_k),$$

то мера $m$ называется $\sigma$-аддитивной (аксиома $\sigma$-аддитивности). Подразумевается, что ряд в правой части сходится для любой последовательности попарно не пересекающихся $\{A_k\}$, объединение которых лежит в $S$. Можно показать, что из этой аксиомы следует, что $m(\varnothing) = 0$ и поэтому из неё следует аксиома аддитивности.

Пользуясь предложением 0.7 введём функцию $\nu\colon R(S) \to \mathbb{R}$ по правилу $\nu(A_1 \sqcup \ldots \sqcup A_n) = \sum\limits_{i=1}^{n} m(A_i)$, где $A_i \in S$. Следующее предложение позволяет назвать $\nu$ продолжением меры $m$ с полукольца $S$ на его минимальное кольцо.

**Предложение 0.8.** *Справедливо следующее*

(1) *функция $\nu$ определена корректно, то есть значение $\nu$ не зависит от выбора представления $A_1 \sqcup \ldots \sqcup A_n$;*

(2) *функция $\nu$ является мерой на кольце $R(S)$;*

(3) *ограничение функции* $\nu$ *на полукольцо* $S$ *совпадает с* $m$;

(4) *если мера* $m$ *была* $\sigma$*-аддитивной, то функция* $\nu$ *также является* $\sigma$*-аддитивной.*

*Доказательство.* См. доказательство в [2, Глава 1, Теорема 2.2]. □

**Лемма 0.9.** *Пусть* $S$ *— полукольцо и* $m\colon S \to \mathbb{R}$ *— мера на* $S$*. Тогда* $m$ *удовлетворяет следующим свойствам*

(1) *если для* $A, B \in S$ *выполнено* $A \subset B$*, то* $m(A) \leqslant m(B)$*;*

(2) *если* $A, A_1, \ldots, A_n \in S$ *и* $A \subset \bigcup\limits_{i=1}^{n} A_i$*, то*

$$m(A) \leqslant \sum_{i=1}^{n} m(A_i);$$

(3) *если* $A_1, \ldots, A_n \subset S$ *— попарно не пересекающиеся множества и* $\bigsqcup\limits_{i=1}^{n} A_i \subset A \in S$*, то*

$$\sum_{i=1}^{n} m(A_i) \leqslant m(A);$$

(4) *если* $\{A_i\}_{i=1}^{+\infty} \subset S$ *— попарно не пересекающиеся множества и* $\bigsqcup\limits_{i=1}^{+\infty} A_i \subset A \in S$*, то*

$$\sum_{i=1}^{+\infty} m(A_i) \leqslant m(A).$$

*Доказательство.* НУЖНО: дописать доказательство или сослаться □

Последнее свойство, доказанное в лемме, называют полуаддитивностью.

### 0.2.4 Лебеговское продолжение меры

Далее будем рассматривать полукольцо $S$ с единицей $\Omega$ и $\sigma$-аддитивной мерой $m$. Пусть $\nu\colon R(S) \to \mathbb{R}$ — продолжение этой меры на минимальное кольцо.

Введём функцию внешней меры $\mu^*\colon 2^\Omega \to \mathbb{R}$, заданную по правилу

$$\mu^*(A) := \inf_{A \subset \bigsqcup\limits_{i=1}^{+\infty} B_i,\, B_i \in S} \sum_{i=1}^{+\infty} m(B_i).$$

**Предложение 0.10.** *Для всякого* $A \subset \Omega$ *в определении внешней меры можно заменить дизъюнктные объединения на произвольные:*

$$\mu^*(A) = \inf_{A \subset \bigcup\limits_{i=1}^{+\infty} B_i,\, B_i \in S} \sum_{i=1}^{+\infty} m(B_i).$$

*Доказательство.* НУЖНО: дописать доказательство или сослаться □

Множество $A \subset \Omega$ называется измеримым, если для любого $\varepsilon > 0$ найдётся множество $B \in R(S)$ такое, что $\mu^*(A \vartriangle B) < \varepsilon$. Если $A$ измеримо, то его мерой называется значение $\mu(A) := \mu^*(A)$. Обозначим через $\mathcal{M}$ системы всех измеримых подмножеств единицы $\Omega$.

**Лемма 0.11.** *Пусть $\{A_i\}_{i=1}^{+\infty} \subset \mathcal{M}$ — последовательность множеств, $A \in M$ и $A \subset \bigcup\limits_{i=1}^{n} A_i$. Тогда*

$$\mu^*(A) \leqslant \sum_{i=1}^{n} \mu^*(A_i);$$

*Доказательство.* НУЖНО: дописать доказательство или сослаться □

**Теорема 0.12.** *Система измеримых множеств $\mathcal{M}$ является алгеброй.*

*Доказательство.* НУЖНО: дописать доказательство или сослаться □

**Теорема 0.13.** *Функция $\mu$ на алгебре множеств $\mathcal{M}$ является мерой.*

*Доказательство.* НУЖНО: дописать доказательство или сослаться □

**Теорема 0.14.** *Алгебра измеримых множеств $\mathcal{M}$ является $\sigma$-алгеброй.*

*Доказательство.* НУЖНО: дописать доказательство или сослаться □

**Теорема 0.15.** *Мера $\mu$ на $\sigma$-алгебре измеримых множеств $\mathcal{M}$ является $\sigma$-аддитивной.*

*Доказательство.* НУЖНО: дописать доказательство или сослаться □

Ограничение внешней меры $\mu^*$ на $\sigma$-алгебру измеримых подмножеств $\mathcal{M}$ мы будем называть лебеговским продолжением меры $m$.

### 0.2.5 Единственность продолжения меры на минимальную $\sigma$-алгебру

Прежде, чем приступить к доказательству теорема Каратеодори мы докажем лемму, характеризующую минимальную $\sigma$-алгебру, порождённую данной алгеброй.

**Лемма 0.16.** *Пусть $\mathcal{A}$ — некоторая алгебра. Тогда существует наименьшая по включению система множеств $Mon(\mathcal{A})$, удовлетворяющая свойствам*

(1) $\mathcal{A} \subset Mon(\mathcal{A})$;

(2) *если имеется последовательность множеств $\{A_i\} \subset Mon(\mathcal{A})$ и либо $A_1 \subset A_2 \subset \ldots$, либо $A_1 \supset A_2 \supset \ldots$, то либо $\bigcup\limits_{i=1}^{+\infty} A_i \in Mon(\mathcal{A})$, либо $\bigcap\limits_{i=1}^{+\infty} A_i \in Mon(\mathcal{A})$, соответственно (мы будем называть это свойство монотонностью).*

*Более того, система множеств $Mon(\mathcal{A})$ совпадает с минимальной $\sigma$-алгеброй $R_\sigma(\mathcal{A})$, порождённой $\mathcal{A}$.*

*Доказательство.* Из определения $\sigma$-алгебры следует, что $R_\sigma(\mathcal{A})$ удовлетворяет свойствам из условия. Кроме того, пересечение произвольного семейства систем множеств, удовлетворяющих данным условиям снова будет удовлетворять им. Поэтому взяв пересечение всех таких систем множеств, мы получим наименьшую по включению систему $Mon(\mathcal{A})$. Так как всегда можно ограничиться системами множеств с той же единицей, что была в $\mathcal{A}$, то $Mon(\mathcal{A})$ обладает единицей.

Из сказанного выше следует, что $Mon(\mathcal{A})$ содержится в $\sigma$-алгебре $R_\sigma(\mathcal{A})$. Мы докажем, что система $Mon(\mathcal{A})$ сама является $\sigma$-алгеброй. Отсюда по минимальности (теорему 0.6) будет следовать обратное включение.

Для множества $B \in Mon(\mathcal{A})$ обозначим через

$$L(B) = \{\, A \in Mon(\mathcal{A}) |\, A \setminus B, B \setminus A, A \cup B \in Mon(\mathcal{A})\}$$

систему множеств из $Mon(\mathcal{A})$, разности и объединение которых с $B$ снова лежат в $Mon(\mathcal{A})$. По построению $A \in L(B)$ равносильно тому, что $B \in L(A)$.

Рассмотрим случай, когда система множеств $L(B)$ непуста и в ней содержится вложенную последовательность $\{A_i\} \subset L(B)$ то есть либо $A_1 \subset A_2 \subset \ldots$, либо $A_1 \supset A_2 \supset \ldots$. Тогда из того, что $Mon(\mathcal{A})$ удовлетворяет свойству монотонности, то $\bigcup\limits_{i=1}^{+\infty} A_i \in L(B)$, либо $\bigcap\limits_{i=1}^{+\infty} A_i \in L(B)$, соответственно. Поэтому система множеств $L(B)$ тоже удовлетворяет свойству монотонности.

Пусть теперь $A \in \mathcal{A}$. Тогда $\mathcal{A} \subset L(A)$, то есть $L(A)$ удовлетворяет обоим свойствам из условия. Отсюда по минимальности $Mon(\mathcal{A}) \subset L(A)$. Тогда для любого $B \in Mon(\mathcal{A})$ имеем $\mathcal{A} \subset L(B)$. Поэтому, аналогично, $Mon(\mathcal{A}) \subset L(B)$, то есть $Mon(\mathcal{A}) = L(B)$ для любого $B \in Mon(\mathcal{A})$. Следовательно, система множеств $Mon(\mathcal{A})$ является алгеброй.

Из монотонности следует, что $Mon(\mathcal{A})$ является $\sigma$-алгеброй. □

Мы приведём доказательство теоремы Каратеодори для случая конечной меры, он отметим, что она остаётся верной и для $\sigma$-конечной меры.

**Теорема 0.17** (Каратеодори)**.** *Пусть $\mathcal{A}$ — алгебра подмножеств множества $\Omega$ и $\nu\colon \mathcal{A} \to \mathbb{R}$ — конечная мера на $\mathcal{A}$. Тогда существует единственная $\sigma$-аддитивная мера $\mu$ на минимальной $\sigma$-алгебре $R_\sigma(\mathcal{A})$, ограничение которой на $\mathcal{A}$ совпадает с $\nu$.*

*Доказательство.* Для доказательства существования рассмотрим лебеговское продолжение меры $\nu$ на $\sigma$-алгебру измеримых подмножеств $\mathcal{M}$. По минимальности (теореме 0.6) имеем включение $R_\sigma(\mathcal{A}) \subset \mathcal{M}$. Тогда ограничение лебеговского продолжения на $R_\sigma(\mathcal{A})$ будет являться $\sigma$-аддитивной мерой.

Докажем единственность. Предположим, что $\mu_1$ и $\mu_2$ — два продолжения меры $\nu$. Положим $\mathcal{A}_= = \{\, B \in R_\sigma(\mathcal{A}) |\, \mu_1(B) = \mu_2(B)\} \subset R_\sigma(\mathcal{A})$ — система множеств, на которых меры $\mu_1$ и $\mu_2$ совпадают.

Промерим, что $\mathcal{A}_=$ удовлетворяет условиям леммы 0.16. По предположению о совпадении ограничений $\mu_1$ и $\mu_2$ на алгебре $\mathcal{A}$ с мерой $m$ имеем включение $\mathcal{A} \subset \mathcal{A}_=$. Остаётся проверить только второе условие.

Пусть имеется вложенная последовательность $\{A_i\} \subset \mathcal{A}_=$, то есть либо $A_1 \subset A_2 \subset \ldots$, либо $A_1 \supset A_2 \supset \ldots$. По непрерывности мер в обоих случаях имеем равенства

$$\mu_1\left(\bigcup_{i=1}^{+\infty} A_i\right) = \lim_{i\to+\infty} \mu_1(A_i) = \lim_{i\to+\infty} \mu_2(A_i) = \mu_2\left(\bigcup_{i=1}^{+\infty} A_i\right),$$

$$\mu_1\left(\bigcap_{i=1}^{+\infty} A_i\right) = \lim_{i\to+\infty} \mu_1(A_i) = \lim_{i\to+\infty} \mu_2(A_i) = \mu_2\left(\bigcap_{i=1}^{+\infty} A_i\right).$$

Тогда $\mathcal{A}_=$ удовлетворяет условию монотонности и по лемме 0.16 содержит (поскольку мы не проверили минимальность лемма даёт только такой результате) $\sigma$-алгебру $R_\sigma(\mathcal{A})$.

Из всего сказанного следует равенство $\mathcal{A}_= = R_\sigma(\mathcal{A})$. □

### 0.2.6 Непрерывность и полнота меры

Мера $\mu$ на кольце $R$ называется непрерывной, если для любой последовательности вложенных подмножеств $\{A_i\}_{i=1}^{+\infty} \subset R$, $A_1 \supset A_2 \supset A_3 \supset \ldots$, пересечения которых $A = \bigcap\limits_{i=1}^{\infty} A_i$ лежит в $R$, имеет место равенство

$$\lim_{i\to+\infty} \mu(A_i) = \mu(A) = \mu\left(\bigcap_{i=1}^{\infty} A_i\right).$$

**Предложение 0.18.** *Конечная мера $\mu$ на кольце $R$ непрерывна тогда и только тогда, когда она $\sigma$-аддитивна.*

*Доказательство.* НУЖНО: дописать доказательство □

**Следствие 0.19.** *Пусть задана мера $\mu$ на кольце $R$ и последовательность вложенных подмножеств $\{A_i\}_{i=1}^{+\infty} \subset R$, $A_1 \supset A_2 \supset A_3 \supset \ldots$, пересечение которых $A = \bigcap\limits_{i=1}^{\infty} A_i$ лежит в $R$ и, кроме того, мера $\mu(A_1) < +\infty$. Тогда имеет место равенство*

$$\lim_{i\to+\infty} \mu(A_i) = \mu(A) = \mu\left(\bigcap_{i=1}^{\infty} A_i\right).$$

*Доказательство.* НУЖНО: дописать доказательство □

**Предложение 0.20.** *Пусть задана конечная $\sigma$-аддитивная мера $\mu$ на кольце $R$. Тогда для любой последовательности вложенных подмножеств $\{A_i\}_{i=1}^{+\infty} \subset R$, $A_1 \subset A_2 \subset A_3 \subset \ldots$, объединение которых $A = \bigcup\limits_{i=1}^{\infty} A_i$ лежит в $R$, имеет место равенство*

$$\lim_{i\to+\infty} \mu(A_i) = \mu(A) = \mu\left(\bigcup_{i=1}^{\infty} A_i\right).$$

*Доказательство.* НУЖНО: дописать доказательство □

Мера $\mu$ на кольце $R$ называется полной, если для любого множества $A \in R$ из равенства нулю меры $\mu(A) = 0$ вытекает, что всякое подмножество $B \subset A$ лежит в $R$. Из свойств меры в этом случае $\mu(B) = 0$.

### 0.2.7 Мера Лебега-Стилтьеса

Пусть $f\colon \mathbb{R} \to \mathbb{R}$ — неубывающая ограниченная функция. Можно рассматривать и случай, когда функция не ограничена и для неё развить теорию, используя $\sigma$-конечную меры, но мы не будем сталкиваться с такими случаями в дальнейшем и поэтому остановимся на рассмотрении ограниченной функции $f$. Рассмотрим полукольцо $S$ полуинтервалов вида $(a, b]$ или $(a; +\infty), (-\infty; b]$ и введём на нём функцию $m_f$ по правилу $m_f((a, b]) = f(b) - f(a)$, $m_f((a; +\infty)) = \lim\limits_{x\to+\infty} f(x) - f(a)$, $m_f((-\infty; b]) = f(b) - \lim\limits_{x\to-\infty} f(x)$ и $\mu_f(\mathbb{R}) = \lim\limits_{x\to+\infty} f(x) - \lim\limits_{x\to-\infty} f(x)$.

**Предложение 0.21.** *Функция $m_f$ является мерой на полукольце $S$. Функция $f$ непрерывна справа тогда и только тогда, когда мера $m_f$ является $\sigma$-аддитивной.*

*Доказательство.* Докажем, что $m_f$ удовлетворяет аксиоме аддитивности. Пусть $(a, b] = \bigsqcup\limits_{i=1}^{n} (a_i, b_i]$. Можно считать, что $a = a_1 < b_1 = a_2 < b_2 < \ldots < b_n = b$. Тогда

$$m_F((a, b]) = f(b) - f(a) = \sum_{i=1}^{n} (f(b_i) - f(a_i)) = \sum_{i=1}^{n} m_F((a_i, b_i]).$$

Доказательство для бесконечных интервалов аналогично, нужно лишь заменить числа $a$ или $b$ на соответствующие пределы.

Предположим, что $f$ непрерывна справа. Пусть $(a, b] = \bigsqcup\limits_{i=1}^{+\infty} (a_i, b_i]$. По полуаддитивности

$$\sum_{i=1}^{+\infty} m_f((a_i, b_i]) \leqslant m_f((a, b]).$$

Фиксируем $\varepsilon > 0$. По непрерывности справа найдём такое $a' \in (a, b]$, что $F(a') - F(a) < \frac{\varepsilon}{2}$ и такие $b'_i > b_i$, чтобы $F(b'_i) - F(b_i) < \frac{\varepsilon}{2^{i+1}}$. Тогда имеем

$$[a', b] \subset (a, b] = \bigsqcup_{i=1}^{+\infty} (a_i, b_i] \subset \bigsqcup_{i=1}^{+\infty} (a_i, b'_i).$$

Поскольку отрезок $[a', b]$ компактен, то из его покрытия $\{(a_i, b'_i)\}$ можно выбрать конечное подпокрытие $\{(a_{i_k}, b'_{i_k})\}_{k=1}^{n}$. Тогда

$$m_f((a,b]) \leqslant m_f((a',b]) + \tfrac{\varepsilon}{2} \leqslant \sum_{k=1}^{n} m_f((a_{i_k}, b'_{i_k}]) + \tfrac{\varepsilon}{2} \leqslant \sum_{i=1}^{+\infty} m_f((a_i, b'_i]) + \varepsilon.$$

Отсюда следует равенство $\sum\limits_{i=1}^{+\infty} m_f((a_i, b_i]) = m_f((a,b])$.

Случай с бесконечными концами сводится к случаю конечного полуинтервала путём выбора точки $c$ такой, что $F(c) - F(-\infty) < \varepsilon$ или $F(+\infty) - F(c) < \varepsilon$ и рассмотрения пересечения полуинтервалов разбиения с $(-\infty; c]$ или $(c; +\infty)$. □

Мерой Лебега-Стилтьеса (порождённой функцией $f$) мы назовём лебеговское продолжение меры $m_f$ и обозначим её через $\mu_f$.

### 0.2.8 Измеримые функции

Пусть $\mathfrak{F}$ — $\sigma$-алгебра подмножеств $\Omega$. Назовём отображение $f\colon \Omega \to \mathbb{R}(\mathbb{R} \cup \{\pm\infty\})$ измеримой функцией, если для любого борелевского множества $B \subset \mathbb{R}(B \subset \mathbb{R} \cup \{\pm\infty\})$ его прообраз является элементом $\mathfrak{F}$ (то есть $f$ является измеримым отображением $f\colon (\Omega, \mathfrak{F}) \to (\mathbb{R}, \mathcal{B})$). Здесь подразумевается, что пространство $\mathbb{R} \cup \{\pm\infty\}$ с рассматриваемой топологией гомеоморфно отрезку. Если мы подразумеваем, что функция принимает только значения из $\mathbb{R}$ и это не явно из контекста, то будем называть её конечной. Можно показать (см. лемму 0.43), что данное требование равносильно тому, что прообраз любого интервала или тому, что прообраз любого бесконечного полуинтервала $(-\infty, b]([-\infty, b])$ измерим. Будем говорить, что функция $f$, определённая на подмножестве $A \subset \Omega$, измерима на $A$, если также для любого борелевского множества $B$ его прообраз $f^{-1}(B) \subset A$ лежит в $\mathfrak{F}$ (это равносильно тому, что ограничение $f|_A$ является измеримым отображением $f|_A : (A, \mathfrak{F} \cap A) \to (\mathbb{R}, \mathcal{B})$).

Далее будем считать, что на $\sigma$-алгебре $\mathfrak{F}$ задана $\sigma$-аддитивная мера $\mu$ и элементы $\mathfrak{F}$ мы будем называет измеримыми множествами.

**Предложение 0.22.** *Пусть функция $f$ измерима и $g$ — непрерывная на* $\operatorname{Im} f (\operatorname{Im} f \cap \mathbb{R})$ *функция. Тогда композиция $g \circ f$ измерима.*

*Доказательство.* Пусть $U$ — открытое множество. Тогда $g^{-1}(U)$ открыто и, следовательно, $f^{-1}(g^{-1}(U))$ измеримо. □

**Предложение 0.23.** *Пусть $f, g$ — измеримые функции. Тогда множество $A_{f \leqslant g} = \{x \in \Omega |\, f(x) \leqslant g(x)\}$ измеримо. Функции $a + f, af, |f|, f^2, f + g$ и $fg$, где $a$ — константа, измеримы. Если функция $g$ не принимает значения 0, то функции $\frac{1}{g}$ и $\frac{f}{g}$ измеримы.*

*Доказательство.* Как было замечено в определении измеримой функции, для проверки измеримости функции $h$ достаточно доказать, что для любой константы $t$ множество $A_{h\leqslant t}$ измеримо.

Первое утверждение следует из представления

$$A_{f\leqslant g}=\bigcup_{q\in\mathbb{Q}}\{x\in\Omega|\, f(x)<q<g(x)\}=\bigcup_{q\in\mathbb{Q}}\left(\{x\in\Omega|\, f(x)<q\}\cap\{x\in\Omega|\, q<g(x)\}\right),$$

где все множества справа измеримы, так как измеримы функции $f$ и $g$.

Измеримость функций $a+faf, |f|, f^2, \frac{1}{g}$ следует из предложения 0.22.

По доказанному выше функция $a-g$ измерима, Тогда также по доказанному выше множество $A_{f\leqslant a-g}$ измеримо и, поэтому измерима функция $f+g$.

Измеримость функции $fg$ следует из представления $fg=\frac{1}{4}(f+g)^2-\frac{1}{4}(f-g)^2$ и доказанного выше. Из этого следует измеримость функции $\frac{f}{g}$. □

**Предложение 0.24.** *Пусть $\{f_n\}$ — последовательность простые функций на $\Omega$. Тогда $\sup\limits_{n\in\mathbb{N}} f_n, \inf\limits_{n\in\mathbb{N}} f_n, \limsup\limits_{n\to+\infty} f_n, \liminf\limits_{n\to+\infty} f_n$ являются измеримыми. Множество определения функции $\lim\limits_{n\to+\infty} f_n$ измеримо и она сама измерима на нём.*

*Доказательство.* Имеем $\{x\in\Omega|\sup f_n(x)<a\}=\bigcap_{n\in\mathbb{N}}\{x\in\Omega|\, f_n(x)<a\}$ — измеримо. Для инфимума выполнено так как $\inf f_n=-\sup(-f_n)$. Для верхнего и нижнего пределов выполнено так как $\limsup\limits_{n\to+\infty} f_n=\inf\limits_{n\to+\infty}\sup\limits_{k\geqslant n} f_k$ и $\liminf\limits_{n\to+\infty} f_n=\sup\limits_{n\to+\infty}\inf\limits_{k\geqslant n} f_k$

Множество определения функции $\lim\limits_{n\to+\infty} f_n$ в точности равно множеству, на котором совпадают верхний и нижний предел, поэтому оно измеримо. Предел на этом множестве совпадает с верхним пределом, поэтому на нём он является измеримой функцией. □

Простой функцией мы будем называть конечную измеримую функцию $f$ такую, что множество значений $f$ конечно и мера множества, на котором $f$ принимает ненулевые значения конечна.

Обобщённой простой функцией мы будем называть произвольную измеримую функцию, принимающую не более чем счётное число значений. Всякая простая функция является обобщённой простой функцией.

Индикатором множества $A$ мы будем называть функцию $\mathbb{1}_A$, заданную по правилу

$$\mathbb{1}_A(x)=\begin{cases}1, & x\in A;\\ 0, & x\notin A.\end{cases}.$$

**Предложение 0.25.** *Верно следующее.*

(1) *Индикатор $\mathbb{1}_A$ измеримого множества $A$ является простой функцией.*

(2) *Всякая (обобщённая) простая функция единственным образом может быть представления в виде конечной линейной комбинации (ряда) с попарно различными коэффициентами индикаторов попрано непересекающихся измеримых множеств, объединение которых равно $\Omega$.*

(3) *Всякая (обобщённая) простая функция единственным образом может быть представления в виде линейной комбинации (ряда) с попарно различными ненулевыми коэффициентами индикаторов попрано непересекающихся измеримых множеств.*

(4) *Всякая простая функция является обобщённой простой и её представления из пунктов выше как для простой и как для обобщённой простой функции совпадают.*

(5) *Ряд из индикаторов с коэффициентами (не обязательно различными) попарно непересекающихся измеримых множеств является обобщённой простой функций.*

*Доказательство.* Индикатор множества принимает всего два значения — 0 и 1. Так как $A$ измеримо, то прообраз любого борелевского множества относительно индикатора есть одно из четырёх измеримых множеств: $\varnothing, A, \Omega \setminus A, \Omega$. Следовательно, $\mathbb{1}_A$ — простая функция.

Пусть $\operatorname{Im} f = \{c_i\}_{i=1}^{n(+\infty)}$ — образ функции $f$. Множества $A_i = f^{-1}(\{c_i\})$ измеримы, так как $f$ измерима, попарно не пересекаются и покрывают всё $\Omega$. Тогда

$$f = \sum_{i=1}^{n(+\infty)} c_i \mathbb{1}_{A_i}.$$

Пусть имеется другое представление $f$ в конечной суммы (ряда), удовлетворяющее условию пункта (2). Пусть множество $B$, индикатор которого входит в это представление не совпадает ни с одним из множеств $A_i$. Если $B$ строго содержится в некотором $A_i$, то найдётся её одно множество $C$ из второго разбиения, пересекающееся с $A_i$. Тогда функция $f$ принимает на $B$ и на $C$ равные значения, что противоречит условию. Иначе $B$ пересекается как минимум с двумя множествами $A_i$ и функция $f$ принимает на $B$ не менее двух различных значений, что снова противоречит условию. Следовательно, данное представление единственно.

Чтобы получить представление для пункта (3) удалим из построенной суммы (ряда) слагаемое, соответствующее $c_i = 0$. Доказательство единственности аналогично.

Построенное выше представление для простой функции содержит только конечное число ненулевых слагаемых, поэтому выполнен пункт (4).

Функция, построенная в пункте (5) принимает счётное число различных значений, причём каждое конкретное значение — на измеримом множестве. Тогда прообраз всякого борелевского множества является объединением (не более, чем счётным) этих множеств (и, возможно, дополнения до их объединения) и поэтому измерим. □

**Предложение 0.26.** *Сумму двух (обобщённых) простых функций является (обобщённой) простой функцией. Функция, пропорциональная (обобщённой) простой функции, сама является таковой. Модуль (обобщённой) простой функции, является таковой функцией.*

*Доказательство.* Сумма двух измеримых функций принимающих конечное (не более, чем счётное) число значений снова является измеримой функцией по предложению 0.23 и снова принимает конечное (не более, чем счётное) число значений.

Функция, пропорциональная измеримой, измерима и модуль измеримой функции измерим по предложению 0.23. Модуль функции, принимающей не более, чем счётное число значений и эта функция умноженная на константу также принимают не более, чем счётное число значений. □

**Предложение 0.27.** *Для всякой измеримой функции $f$ существует неубывающая последовательность обобщённых простых функций $\{f_n\}$, непревосходящих $f$, равномерно сходящаяся к $f$. Для всякой неотрицательной измеримой функции $f$ существует неубывающая последовательность неотрицательных простых функций $f_n$, непревосходящих $f$, поточечно сходящаяся к $f$ (обозначение: $f_n \nearrow f$).*

*Доказательство.* Поскольку $f$ измерима, то для всяких натурального $n$ и целого $k$ множество $A_{n,k} = f^{-1}([\frac{k}{2^n}, \frac{k+1}{2^n}))$, а также $A_{-\infty} = f^{-1}(\{-\infty\}), A_{+\infty} = f^{-1}(\{+\infty\})$ измеримы и для фиксированного $n$ и различных $k$ эти множества образуют разбиение $\Omega$. Положим

$$f_n(x) = \sum_{k\in\mathbb{Z}} \tfrac{k}{2^n} \mathbb{1}_{A_{n,k}} + (+\infty)\mathbb{1}_{A_{+\infty}} + (-\infty)\mathbb{1}_{A_{-\infty}}.$$

Тогда $f_n$ удовлетворяет определению обобщённой простой функции. Из конструкции последовательность функций $f_n$ поточечно неубывает и не превосходит $f$. Кроме того, для всякого $n$ и $x \in \Omega$ выполнены неравенства $0 \leqslant f(x) - f_n(x) \leqslant \frac{1}{2^n}$. Поэтому последовательность $\{f_n\}$ сходится к $f$ равномерно.

Для построения последовательности простых функций поправим конструкцию выше. Если мера всего пространства бесконечна, по будем поочереди доопределять её на новом куске исчерпания. Там, где $f$ равна $+\infty$ положим $f_n$ равной $n$, там, где $f$ принимает значения из промежутка $[0, 2^n]$ — как $f_n$ выше и наконец, на оставшемся множестве — нулём. □

### 0.2.9 Интеграл Лебега

Пусть $f = \sum\limits_{i=1}^{n} c_i \mathbb{1}_{A_i}, c_i \neq 0$ — простая функция в представлении из предложения 0.25. Интегралом простой функции $f$ на пространстве $\Omega$ по мере $\mu$ мы назовём сумму $\sum\limits_{i=1}^{n} c_i \cdot \mu(A_i)$ и обозначим через $\int\limits_{\Omega} f \mathrm{d}\mu$.

Обобщённую простую функцию $f = \sum\limits_{i=1}^{+\infty} c_i \mathbb{1}_{A_i}$ (в представлении из предложения 0.25) назовём интегрируемой по Лебегу на пространстве $\Omega$ по мере $\mu$, если для $c_i \neq 0$ меры $\mu(A_i)$ конечны, для $c_i = \pm\infty$ мера $\mu(A_i) = 0$ и ряд $\sum\limits_{i=1}^{+\infty} c_i \cdot \mu(A_i)$ сходится абсолютно (предполагается, что $0 \cdot \mu(A) = 0$, даже если мера $\mu(A)$ бесконечна). В этом случае обозначим через $\int\limits_{\Omega} f \mathrm{d}\mu$ сумму ряда и назовём её интегралом обобщённой простой функции $f$ на пространстве $\Omega$ по мере $\mu$.

Из определения простой функции следует, что она является интегрируемой на пространстве $\Omega$ по мере $\mu$ обобщённой простой функцией.

**Лемма 0.28.** *Пусть $f, g$ — обобщённые простые функции. Тогда*

(1) *если $f, g$ — интегрируемые на $\Omega$ по мере $\mu$, то функция $f+g$ интегрируема и имеет место формула*

$$\int\limits_{\Omega} f \mathrm{d}\mu + \int\limits_{\Omega} g \mathrm{d}\mu = \int\limits_{\Omega} (f+g) \mathrm{d}\mu;$$

(2) *интегрируемость функции $f$ равносильна интегрируемости функции $|f|$ и в случае интегрируемости справедливо неравенство*

$$\left| \int\limits_{\Omega} f \mathrm{d}\mu \right| \leqslant \int\limits_{\Omega} |f| \mathrm{d}\mu;$$

(3) *если $|f| \leqslant g$ и $g$ интегрируема, то $|f|$ интегрируема и выполнено неравенство*

$$\int\limits_{\Omega} |f| \mathrm{d}\mu \leqslant \int\limits_{\Omega} g \mathrm{d}\mu.$$

*Доказательство.* Первые два свойства следуют из свойств сходимости абсолютно сходящихся рядов. Третье — из признака абсолютной сходимости. □

**Лемма 0.29.** *Пусть мера $\mu$ конечна на $\Omega$. Тогда выполнены следующие утверждения.*

(1) *Всякая ограниченная обобщённая простая функция интегрируема на $\Omega$ по $\mu$.*

(2) *Интеграл $\int\limits_{\Omega} \mathbb{1}_A \mathrm{d}\mu$ от индикатора $\mathbb{1}_A$ равен мере $A$.*

*Доказательство.* Пусть константа $C$ ограничивает функцию $f$. Так как мера $\Omega$ конечна, то функция $C \cdot \mathbb{1}_{\Omega}$ интегрируема, откуда по лемме 0.28 следует, что $f$ интегрируема. Второе утверждение выполнено из построения интеграла обобщённой простой функции □

**Теорема 0.30.** *Пусть мера $\mu$ конечна на $\Omega$. Пусть $f$ — измеримая функция и существует последовательность интегрируемых обобщённых простых функций $\{f_n\}$, равномерно сходящаяся к $f$. Тогда*

(1) *существует предел $\int\limits_{\Omega} f_n \mathrm{d}\mu =: I$;*

(2) *для любой другой последовательности интегрируемых обобщённых простых функций $\{g_n\}$, равномерно сходящихся к $f$ предел $\int\limits_{\Omega} g_n \mathrm{d}\mu$ также равен $I$;*

(3) *если $f$ — обобщённая простая функция, то $f$ интегрируема и $\int\limits_{\Omega} f \mathrm{d}\mu = I$.*

*Доказательство.* Фиксируем $\varepsilon > 0$. Так как функции $f_n$ равномерно сходятся к $f$, то начиная с некоторого $N$ для $n, m > N$ и для любого $x \in \Omega$ выполнено неравенство $|f_n(x) - f_m(x)| < \varepsilon$. Тогда по леммам 0.29 и 0.29 имеем

$$\left| \int\limits_{\Omega} f_n \mathrm{d}\mu - \int\limits_{\Omega} f_m \mathrm{d}\mu \right| \leqslant \int\limits_{\Omega} |f_n - f_m| \mathrm{d}\mu \leqslant \varepsilon\mu(X).$$

По критерию Коши существует предел $\int\limits_{\Omega} f_n \mathrm{d}\mu =: I$.

Заметим, что если $f = 0$, то для $n > N(\varepsilon)$ выполнено равенство

$$\left| \int\limits_{\Omega} f_n \mathrm{d}\mu \right| < \varepsilon\mu(X).$$

Поэтому в данном случае $I = 0$. Поскольку, если последовательности $f_n$ и $g_n$ сходятся равномерно к $f$, то последовательность $f_n - g_n$ сходится равномерно к нулевой функции, то предел $I$ будет одинаковым для обеих последовательностей.

Пусть $f$ сама является обобщённой простой функцией. Найдётся индекс $n$ такой, что для любого $x \in \Omega$ выполнено неравенство $|f(x) - f_n(x)| < 1$. По леммам 0.29 и 0.28 функция $f - f_n$ интегрируема. Так как $f_n$ интегрируема, то снова по лемме 0.28 функция $f = (f - f_n) + f_n$ интегрируема. Теперь возьмём в качестве последовательности $f_n$ постоянную последовательность из функции $f$. Тогда предел этой последовательности равен одновременно $\int\limits_{\Omega} f \mathrm{d}\mu$ и $I$. □

Измеримая функция $f$ на $\Omega$ с $\mu(\Omega) < +\infty$ называется интегрируемой, если существует последовательность интегрируемых обобщённых простых функций $\{f_n\}$, равномерно сходящаяся к $f$. Интегралом (Лебега) $f$ на $\Omega$ по мере $\mu$ называется предел $\int\limits_{\Omega} f_n \mathrm{d}\mu$. В дальнейшем мы будем обозначать его через $\int\limits_{\Omega} f \mathrm{d}\mu$.

**Следствие 0.31.** *Путь мера $\mu$ конечна на $\Omega$. Пусть измеримая функция $f$ интегрируема на $\Omega$ по $\mu$ и существует последовательность обобщённых простых функций $\{f_n\}$, равномерно сходящаяся к $f$. Тогда начиная с некоторого номера $N$ все функции в этой последовательности интегрируемы.*

*Доказательство.* Рассмотрим некоторую последовательность интегрируемых обобщённых простых функций $\{f_n\}$, равномерно сходящуюся к $f$. Тогда последовательность простых функций $\{f_n - g_n\}$ равомерно сходится к тождественно нулевой функции. Следовательно, начиная с некоторого индекса $N$ для всех $n > N$ и $x \in \Omega$ выполнено неравенство $|f_n(x) - g_n(x)| < 1$. Тогда по леммам 0.25 и 0.28 функция $g_n - f_n$ интегрируема. Тогда интегрируема и функция $g_n = (g_n - f_n) + f_n$. □

Для пространства $\Omega$ с $\sigma$-конечной мерой $\mu$ назовём исчерпанием всякую возрастающую вложенную последовательность измеримых множеств $\{\Omega_i\}$ такую, что $\bigcup\limits_{i=1}^{+\infty} \Omega_i = \Omega$. Измеримая функция $f$ на пространстве $\Omega$ с $\sigma$-конечной мерой $\mu$ называется интегрируемой, если для всякого исчерпания $\{\Omega_i\}$ функция $f$ интегрируема на всех $\Omega_i$ и существует конечный предел, не зависящий от исчерпания

$$\lim_{i \to +\infty} \int\limits_{\Omega_i} f \mathrm{d}\mu =: \int\limits_{\Omega} f \mathrm{d}\mu.$$

**Лемма 0.32.** *Пусть неотрицательная измеримая функция $f$ интегрируема по некоторому одному исчерпанию на пространстве $\Omega$. Тогда $f$ интегрируема по любому исчерпанию и пределы интегралов не зависят от исчерпания. Другими словами $f$ интегрируема на $\Omega$.*

**Лемма 0.33.** *Пусть $f, g$ — измеримые неотрицательные функции, $f \leqslant g$ и $g$ интегрируема. Тогда $f$ интегрируема и $\int\limits_{\Omega} f \mathrm{d}\mu \leqslant \int\limits_{\Omega} g \mathrm{d}\mu$.*

*Доказательство.* Лемма верна для обобщённых простых функций на множестве конечной меры, поэтому она верна и для измеримых неотрицательных функций на множестве конечной меры. Чтобы доказать лемму для множества произвольной меры, заметим, что для любого исчерпания последовательность интегралов от $f$ или от $g$ по его членам неубывает, поэтому для всякого исчерпания из сходимости предела для $g$ следует абсолютная сходимость предела для $f$. Отсюда же следует, что предел для $f$ не зависит от выбора исчерпания (нужно разбить исчерпание и воспользоваться счётной аддитивностью). □

**Теорема 0.34.** *Неотрицательная измеримая функция $f$ интегрируема тогда и только тогда, когда существует конечный супремум*

$$\sup_{g \text{ — простая}, 0 \leqslant g \leqslant f} \int\limits_{\Omega} g \mathrm{d}\mu.$$

*Если супремум существует, то он равен $\int\limits_{\Omega} f \mathrm{d}\mu$.*

*Доказательство.* Если $f$ интегрируема, то по лемме выше для любой простой $g$, такой, что $0 \leqslant g \leqslant f$ выполнено неравенство $\int\limits_{\Omega} g \mathrm{d}\mu \leqslant \int\limits_{\Omega} f \mathrm{d}\mu$, поэтому супремум существует и не превосходит интеграла.

Построим возрастающую последовательность обобщённых простых функций, равномерно сходящуюся к $f$, ка это было сделано в предложении 0.27. По лемме 0.31 можем считать, что все функции в последовательности интегрируемы. Для каждой такой функции можно подобрать возрастающую последовательность простых функций, не превосходящих данную обобщённую простую, которая сходится к ней и поэтому последовательность их интегралов сходится к интегралу данной обобщённой простой функции. Так как интегралы от обобщённых простых функций сходятся к её интегралу, то суперемум в условии равен интегралу.

Предположим теперь, что существует супремум. Снова по предложению 0.27 построим последовательность обобщённых простых функций, сходящуюся равномерно к $f$ и последовательности простых функций, сходящуюся к данной обобщённой простой функции. Из формул для интеграла простой и обобщённой простой функций следует, что интегралы простых функций сходятся к интегралам обобщённых простых функций. Поскольку супремум интегралов конечен, то обобщённые простые функции в последовательности интегрируемы на всём пространстве и их интегралы не превосходят данного супремума. Тогда $f$ интегрируема на всяком множестве конечной меры. Фиксируем произвольное исчерпание. Тогда можно построить обобщённую простую функцию, которая интегрируема по его членам и интегралы которой отличаются от интегралов $f$ не более, чем на $\varepsilon$ для каждого члена. Тогда $f$ интегрируема по данному исчерпанию и её интеграл не превосходит супремума. По лемме $f$ интегриуема. $\square$

### 0.2.10 Прямой образ меры (pushforward measure)

Пусть $(X, \mathfrak{A}_X)$ и $(Y, \mathfrak{A}_Y)$ — измеримые пространства, $\mu$ — мера на $\sigma$-алгебре $\mathfrak{A}_X$ и $f\colon (X, \mathfrak{A}_X) \to (Y, \mathfrak{A}_Y)$ — измеримое отображение. Тогда прямым образом меры $\mu$ при отображении $f$ называется функция $f_*\mu$, заданная по правилу $f_*\mu(A) = \mu(f^{-1}(A))$ для любого $A \in \mathfrak{A}_Y$.

**Предложение 0.35.** *Прямой образ ($\sigma$-аддитивной) меры $f_*\mu$ является ($\sigma$-аддитивной) мерой на алгебре $\mathfrak{A}_Y$.*

*Доказательство.* Пусть $\{A_k\}$ — последовательность попрано не пересекающихся множеств из $\mathfrak{A}_Y$. Тогда множества из последовательности $\{f^{-1}(A_k)\}$ также попарно не пересекаются. Из определения прямого образа меры $f_*\mu$ имеем цепочку равенств

$$f_*\mu\left(\bigsqcup_{k=1}^{+\infty} A_k\right) = \mu\left(f^{-1}\left(\bigsqcup_{k=1}^{+\infty} A_k\right)\right) = \mu\left(\bigsqcup_{k=1}^{+\infty} f^{-1}(A_k)\right) = \sum_{k=1}^{+\infty} \mu(f^{-1})(A_k) = \sum_{k=1}^{+\infty} f_*\mu(A_k).$$

Доказательство аддитивности проводится аналогично. $\square$

**Теорема 0.36.** *Пусть мера $\mu$ конечна на $X$. Рассмотрим измеримую функцию $g\colon (Y, \mathfrak{A}_Y) \to (\mathbb{R}, \mathcal{B})$. Тогда $g$ интегрируема на $Y$ по прямому образу меры $f_*\mu$*

*тогда и только тогда, когда композиция $g \circ f$ интегрируема на $X$ по мере $\mu$. В случае интегрируемости равны интегралы*

$$\int\limits_X (g \circ f) \mathrm{d}\mu = \int\limits_Y g \mathrm{d} f_* \mu.$$

*Доказательство.* Из определения обобщённых простых функций $g$ и $g \circ f$ одновременно являются или не являются простыми обобщёнными функциями.

Для обобщённой простой функции истинность утверждения следует из того, что интегралы $\int\limits_X (g \circ f) \mathrm{d}\mu$ и $\int\limits_Y g \mathrm{d} f_* \mu$ равны сумме одного и того же ряда (поскольку $f_*\mu(g^{-1}(\{a\})) = \mu(f^{-1}(g^{-1}(\{a\}))) = \mu((g \circ f)^{-1}(\{a\}))$ для $a \in \mathbb{R}$).

Если существует последовательность интегрируемых обобщённых простых функций $\{g_n\}$ на $Y$, равномерно сходящаяся к $g$, то последовательность $\{g_n \circ f\}$ будет являться последовательностью интегрируемых обобщённых простых функций на $X$, равномерно сходящейся к $g \circ f$. Следовательно, из интегрируемости $g$ следует интегрируемость $g \circ f$.

Обратно, пусть $g \circ f$ интегрируема. По предложению 0.27 существует последовательность обобщённых простых функций $\{g_n\}$ на $Y$, равномерно сходящаяся к $g$. Тогда последовательность обобщённых простых функций $\{g_n \circ f\}$ равномерно сходится к $g \circ f$. По следствию 0.31 начиная с некоторого номера $N$ функции $g_n \circ f$ являются интегрируемыми. По доказанному выше, это означает, что для всех $n > N$ функции $g_n$ интегрируемы, поэтому $g$ интегрируема.

Пусть теперь $g$ и $g \circ f$ интегрируемы. Фиксируем последовательность интегрируемых обобщённых простых функций $\{g_n\}$ на $Y$, равномерно сходящуюся к $g$. Имеем цепочку равенств

$$\int\limits_X (g \circ f) \mathrm{d}\mu = \lim_{n \to +\infty} \int\limits_X (g_n \circ f) \mathrm{d}\mu = \lim_{n \to +\infty} \int\limits_Y g_n \mathrm{d} f_* \mu = \lim_{n \to +\infty} \int\limits_Y g \mathrm{d} f_* \mu.$$

□

## 0.3 Теория категорий и взгляд на измеримые пространства с её точки зрения

Теория категорий в её лучшем проявлении выражает собой формализацию понятия «математическая конструкция» через понятия объектов, морфизмов, функторов, естественных преобразований, а также формализует интуицию в виде универсальных свойств, сопряжённости функторов, эквивалентности категорий и так далее.

В этом подразделе будет предполагаться, что вы знакомы с определением категории (например, основанном на теории множеств) и знакомы с понятиями объекта, морфизма между объектами и функтором из одной категории в другую. Остальные определения по возможности будут приведены здесь.

### 0.3.1 Категория измеримых пространств

Пусть $\mathfrak{A}_X$ — $\sigma$-алгебра подмножеств множества $X$. Пара $(X, \mathfrak{A}_X)$ называется измеримым пространством.

Пусть $(X, \mathfrak{A}_X)$ и $(Y, \mathfrak{A}_Y)$ — пара измеримых пространств. Морфизмом измеримых пространств (измеримым отображением) $f\colon (X, \mathfrak{A}_X) \to (Y, \mathfrak{A}_Y)$ называется отображение множеств $f\colon X \to Y$ такое, что для любого $U \in \mathfrak{A}_Y$ выполнено $f^{-1}(U) \in \mathfrak{A}_X$, где $f^{-1}$ — это полный прообраз.

В категории измеримых пространств $\mathfrak{Meas}$ объектами являются измеримые пространства, а морфизмами — морфизмы измеримых пространств.

### 0.3.2 Прямой образ $\sigma$-алгебры

Пусть $(X, \mathfrak{A}_X)$ — измеримое пространство и $f\colon X \to Y$ — отображение множеств. Существует способ естественным образом построить $\sigma$-алгебру подмножеств $Y$.

Положим $f_*(\mathfrak{A}_X) = \{\, B \subset Y \mid f^{-1}(B) \in \mathfrak{A}_X \}$ — все подмножества $Y$, полный прообраз которых лежит в $\mathfrak{A}_X$. Построенная система множеств называется прямым образом $\sigma$-алгебры $\mathfrak{A}_X$ при отображении $f$.

**Предложение 0.37.** *Конструкция $f_*$ обладает следующими свойствами.*

(1) *Система множеств $f_*(\mathfrak{A}_X)$ является $\sigma$-алгеброй подмножеств множества $Y$.*

(2) *Отображение $f$ является морфизмом измеримых пространств $f\colon (X, \mathfrak{A}_X) \to (Y, f_*(\mathfrak{A}_X))$.*

(3) *Если $\mathfrak{A}_Y$ — $\sigma$-алгебра подмножеств $Y$, то отображение $f$ является морфизмом измеримых пространств $f\colon (X, \mathfrak{A}_X) \to (Y, \mathfrak{A}_Y)$ тогда и только тогда, когда $\mathfrak{A}_Y \subset f_*(\mathfrak{A}_X)$.*

(4) *Пусть имеются отображения множеств $f\colon X \to Y$, $g\colon X \to Z$ и $h\colon Y \to Z$ такие, что $h \circ f = g$, то $h$ является морфизмом измеримых пространств $h\colon (Y, f_*(\mathfrak{A}_X)) \to (Z, g_*(\mathfrak{A}_X))$.*

$$\begin{tikzcd} & Y \arrow[dd, "h"] \\ X \arrow[ur, "f"] \arrow[dr, "g"'] & \\ & Z, \end{tikzcd} \qquad \begin{tikzcd} & (Y, f_*(\mathfrak{A}_X)) \arrow[dd, "h"] \\ (X, \mathfrak{A}_X) \arrow[ur, "f"] \arrow[dr, "g"'] & \\ & (Z, g_*(\mathfrak{A}_X)). \end{tikzcd}$$

*Доказательство.* Первый пункт проверяется непосредственно и следует из теоретико-множественных тождеств. Второй пункт — из построения $f_*(\mathfrak{A}_X)$, третий — из определения морфизма измеримых пространств.

Докажем четвёртый пункт. Пусть $U \in g_*(\mathfrak{A}_X)$. Тогда $g^{-1}(U) = f^{-1}(h^{-1}(U)) \in \mathfrak{A}_X$. По определению прямого образа $f_*(\mathfrak{A}_X)$ множество $h^{-1}(U)$ содержится в нём. Следовательно, $h$ является морфизмом измеримых пространств. □

### 0.3.3 Обратный образ $\sigma$-алгебры

Пусть $(Y, \mathfrak{A}_Y)$ — измеримое пространство и $f\colon X \to Y$ — отображение множеств. Существует способ естественным образом построить $\sigma$-алгебру подмножеств $X$.

Положим $f^*(\mathfrak{A}_Y) = \{ f^{-1}(B) | \, B \in \mathfrak{A}_Y \}$ — полные прообразы всех элементов $\sigma$-алгебры $\mathfrak{A}_Y$. Построенная система множеств называется обратным образом $\sigma$-алгебры $\mathfrak{A}_X$ при отображении $f$.

**Предложение 0.38.** *Конструкция $f^*$ обладает следующими свойствами.*

(1) *Система множеств $f^*(\mathfrak{A}_Y)$ является $\sigma$-алгеброй подмножеств множества $X$.*

(2) *Отображение $f$ является морфизмом измеримых пространств $f\colon (X, f^*(\mathfrak{A}_Y)) \to (Y, \mathfrak{A}_Y)$.*

(3) *Если $\mathfrak{A}_X$ — $\sigma$-алгебра подмножеств $X$, то отображение $f$ является морфизмом измеримых пространств $f\colon (X, \mathfrak{A}_X) \to (Y, \mathfrak{A}_Y)$ тогда и только тогда, когда $f^*(\mathfrak{A}_Y) \subset \mathfrak{A}_X$.*

(4) *Пусть имеются отображения множеств $f\colon X \to Y, g\colon Z \to Y$ и $h\colon X \to Z$ такие, что $g \circ h = f$, то $h$ является морфизмом измеримых пространств $h\colon (X, f^*(\mathfrak{A}_Y)) \to (Z, g^*(\mathfrak{A}_Y))$.*

$$\begin{tikzcd} X \arrow[d, "h"] \arrow[dr, "f"] & \\ Z \arrow[r, "g"'] & Y, \end{tikzcd} \qquad \begin{tikzcd} (X, f^*(\mathfrak{A}_Y)) \arrow[d, "h"] \arrow[dr, "f"] & \\ (Z, g^*(\mathfrak{A}_Y)) \arrow[r, "g"'] & (Y, \mathfrak{A}_Y). \end{tikzcd}$$

*Доказательство.* Первый пункт проверяется непосредственно и следует из теоретико-множественных тождеств. Второй пункт — из построения $f^*(\mathfrak{A}_X)$, третий — из определения морфизма измеримых пространств.

Докажем четвёртый пункт. Пусть $U \in g^*(\mathfrak{A}_Y)$. Тогда для некоторого $V \in \mathfrak{A}_Y$ выполнено $g^{-1}(V) = U$. Далее, $h^{-1}(U) = h^{-1}(g^{-1}(V)) = f^{-1}(V)$. По определению обратного образа $f^*(\mathfrak{A}_X)$ множество $h^{-1}(U)$ содержится в нём. Следовательно, $h$ является морфизмом измеримых пространств. □

### 0.3.4 Связь между минимальной $\sigma$-алгеброй, прямым и обратным образами $\sigma$-алгебры

**Лемма 0.39.** *Пусть $X, Y$ — множества, $f\colon X \to Y$ — отображение множеств. Если $\mathfrak{A}_X$ — $\sigma$-алгебра подмножеств $X$, то $f^*(f_*(\mathfrak{A}_X)) \subset \mathfrak{A}_X$. Если $\mathfrak{A}_Y$ — $\sigma$-алгебра подмножество $Y$, то $\mathfrak{A}_Y \subset f_*(f^*(\mathfrak{A}_Y))$.*

*Доказательство.* По предложению 0.37 отображение $f$ является морфизмом измеримых пространств $f\colon (X, \mathfrak{A}_X) \to (Y, f_*(\mathfrak{A}_X))$. Тогда по предложению 0.38 имеем включение $f^*(f_*(\mathfrak{A}_X)) \subset \mathfrak{A}_X$.

Теперь По предложению 0.38 отображение $f$ является морфизмом измеримых пространств $f\colon (X, f^*(\mathfrak{A}_Y)) \to (Y, \mathfrak{A}_Y)$. Тогда по предложению 0.37 имеем включение $\mathfrak{A}_Y \subset f_*(f^*(\mathfrak{A}_Y))$. $\square$

Рассмотрим категорию $\sigma\text{-}\mathfrak{Alg}_X$ всех $\sigma$-алгебр с единицей $X$, в которой объектами выступают $\sigma$-алгебры подмножеств $X$, а единственный существующий морфизм из $\sigma$-алгебры $\mathfrak{A}$ идёт в $\sigma$-алгебру $\mathfrak{A}'$, если $\mathfrak{A} \subset \mathfrak{A}'$. Построенные нами для отображения множеств $f\colon X \to Y$ конструкции $f_*$ и $f^*$ являются функторами между категориями $\sigma\text{-}\mathfrak{Alg}_X$ и $\sigma\text{-}\mathfrak{Alg}_Y$ и, более того, что эти функторы сопряжены.

**Теорема 0.40.** *Пусть $f\colon X \to Y$ — отображение множеств. Тогда $f_*\colon \sigma\text{-}\mathfrak{Alg}_X \to \sigma\text{-}\mathfrak{Alg}_Y$ и $f^*\colon \sigma\text{-}\mathfrak{Alg}_Y \to \sigma\text{-}\mathfrak{Alg}_X$ — функторы. Кроме того, функтор $f^*$ является левым сопряжённым к функтору $f_*$ (и обратно, функтор $f_*$ является правым сопряжённым к $f^*$). Другими словами, для любого объекта $\mathfrak{A}_X \in \sigma\text{-}\mathfrak{Alg}_X$ и любого объекта $\mathfrak{A}_Y \in \sigma\text{-}\mathfrak{Alg}_Y$ и имеется естественный изоморфизм множеств $\mathrm{Hom}(f^*(\mathfrak{A}_Y), \mathfrak{A}_X)) \cong \mathrm{Hom}(\mathfrak{A}_Y, f_*(\mathfrak{A}_X))$.*

*Доказательство.* Из построения $f_*$ и $f^*$ сохраняют отношения включения и равенства (и, следовательно, композицию включений), поэтому они функториальны.

Теперь по предложению 0.38 имеет место включение $f^*(\mathfrak{A}_Y) \subset \mathfrak{A}_X$ тогда и только тогда, когда $f\colon (X, \mathfrak{A}_X) \to (Y, \mathfrak{A}_Y)$ является измеримым отображением. По предложению 0.37 последнее равносильно включению $\mathfrak{A}_X \subset f_*(\mathfrak{A}_Y)$. Множества $\mathrm{Hom}(f^*(\mathfrak{A}_Y), \mathfrak{A}_X))$ и $\mathrm{Hom}(\mathfrak{A}_Y, f_*(\mathfrak{A}_X))$ одновременно либо пусты, либо состоят из одного элемента. Тогда пусть данный изоморфизм будет пустым в первом случае и сопоставляет единственный элемент единственному элементу во втором случае.

Если имеются включения $\mathfrak{A}_{X1} \subset \mathfrak{A}_{X2}$ и $\mathfrak{A}_{Y1} \subset \mathfrak{A}_{Y2}$, то возникает коммутативная диаграмма, индуцированная этими включениями, выражающая естественность изоморфизма

$$\begin{tikzcd}
 & \mathrm{Hom}(f^*(\mathfrak{A}_{Y2}), \mathfrak{A}_{X2})) \arrow[rr] \arrow[dd, dotted] & & \mathrm{Hom}(\mathfrak{A}_{Y2}, f_*(\mathfrak{A}_{X2})) \arrow[dd] \\
\mathrm{Hom}(f^*(\mathfrak{A}_{Y2}), \mathfrak{A}_{X1})) \arrow[ur] \arrow[rr] \arrow[dd] & & \mathrm{Hom}(\mathfrak{A}_{Y2}, f_*(\mathfrak{A}_{X1})) \arrow[ur] \arrow[dd] & \\
 & \mathrm{Hom}(f^*(\mathfrak{A}_{Y1}), \mathfrak{A}_{X2})) \arrow[rr, dotted] & & \mathrm{Hom}(\mathfrak{A}_{Y1}, f_*(\mathfrak{A}_{X2})) \\
\mathrm{Hom}(f^*(\mathfrak{A}_{Y1}), \mathfrak{A}_{X1})) \arrow[ur, dotted] \arrow[rr] & & \mathrm{Hom}(\mathfrak{A}_{Y1}, f_*(\mathfrak{A}_{X1})) \arrow[ur] &
\end{tikzcd}$$

□

**Лемма 0.41.** *(О коммутировании конструкций минимальной $\sigma$-алгебры и обратного образа) Пусть $X, Y$ — множества, $T$ — система подмножеств $Y$, $Y \in T$. Пусть $f\colon X \to Y$ — отображение множеств. Обозначим через $f^{-1}(T)$ систему множеств $\{f^{-1}(U) | U \in T\}$. Тогда $R_\sigma(f^{-1}(T)) = f^*(R_\sigma(T))$.*

*Доказательство.* Так как $T \subset R_\sigma(T)$, то по построению $f^*$ имеем включение $f^{-1}(T) \subset f^*(R_\sigma(T))$. По предложению 0.38 система множеств $f^*(R_\sigma(T))$ является $\sigma$-алгеброй. Тогда по минимальности (теореме 0.6) имеем включение $R_\sigma(f^{-1}(T)) \subset f^*(R_\sigma(T))$.

Далее, система множеств $f_*(R_\sigma(f^{-1}(T)))$ является $\sigma$-алгеброй по предложению 0.37. По построению $f_*$ имеем включение $T \subset f_*(R_\sigma(f^{-1}(T)))$. Снова по минимальности (теореме 0.6) имеем включение $R_\sigma(T) \subset f_*(R_\sigma(f^{-1}(T)))$.

По лемме 0.39 имеем включение $f^*(f_*(R_\sigma(f^{-1}(T))) \subset R_\sigma(f^{-1}(T))$. Собирая вместе все включения и пользуясь тем, что $f^*(A) \subset f^*(B)$ для $A \subset B$ получаем

$$f^*(R_\sigma(T)) \subset f^*(f_*(R_\sigma(f^{-1}(T))) \subset R_\sigma(f^{-1}(T)) \subset f^*(R_\sigma(T)),$$

откуда следует требуемое. □

### 0.3.5 Произведение в категории измеримых пространств

Пусть $(X, \mathfrak{A}_X)$ и $(Y, \mathfrak{A}_Y)$ — измеримые пространства. Построим по ним новые измеримые пространства следующим образом.

Произведением измеримых пространств $(X, \mathfrak{A}_X)$ и $(Y, \mathfrak{A}_Y)$ мы назовём измеримое пространство $(X \times Y, \mathfrak{A}_X \otimes \mathfrak{A}_Y)$, где $\mathfrak{A}_X \otimes \mathfrak{A}_Y$ есть минимальная $\sigma$-алгебра, порождённая полукольцом всех возможных декартовых произведений $A_X \times A_Y$, где $A_X \in \mathfrak{A}_X$ и $A_Y \in \mathfrak{A}_Y$.

**Теорема 0.42.** *Пусть $(X, \mathfrak{A}_X)$ и $(Y, \mathfrak{A}_Y)$ — измеримые пространства, $(X \times Y, \mathfrak{A}_X \otimes \mathfrak{A}_Y)$ — их произведение. Тогда*

(1) *отображения проекций $\mathrm{pr}_X\colon X \times Y \to X$ и $\mathrm{pr}_Y\colon X \times Y \to Y$ являются морфизмами измеримых пространств.*

(2) *пространство* $(X \times Y, \mathfrak{A}_X \otimes \mathfrak{A}_Y)$ *удовлетворяет универсальному свойству произведения, то есть для всякого измеримого пространства* $(Z, \mathfrak{A}_Z)$ *и всяких двух морфизмов* $f\colon (Z, \mathfrak{A}_Z) \to (X, \mathfrak{A}_X)$, $g\colon (Z, \mathfrak{A}_Z) \to (Y, \mathfrak{A}_Y)$ *существует единственный морфизм* $f \times g\colon (Z, \mathfrak{A}_Z) \to (X \times Y, \mathfrak{A}_X \otimes \mathfrak{A}_Y)$*, удовлетворяющий условию* $\mathrm{pr}_X \circ (f \times g) = f$ *и* $\mathrm{pr}_Y \circ (f \times g) = g$.

$$\begin{tikzcd}
& (Z, \mathfrak{A}_Z) \arrow[dl, "f"'] \arrow[dr, "g"] \arrow[dd, dashed, "f\times g"] & \\
(X, \mathfrak{A}_X) & & (Y, \mathfrak{A}_Y) \\
& (X \times Y, \mathfrak{A}_X \otimes \mathfrak{A}_Y) \arrow[ul, "\mathrm{pr}_X"] \arrow[ur, "\mathrm{pr}_Y"'] &
\end{tikzcd}$$

*Доказательство.* Пусть $A \in \mathfrak{A}_X$. Тогда $\mathrm{pr}_X^{-1}(A) = A \times Y \in \mathfrak{A}_X \otimes \mathfrak{A}_Y$, поэтому отображение проекции на $X$ является морфизмом. Аналогично доказывается для второй проекции.

Пусть теперь $f, g$ — морфизмы из условия. Положим $(f \times g)(z) = (f(z), g(z))$ — единственный возможные вариант определить это отображение так, чтобы выполнялись условия на проекции. Проверим, что $f \times g$ является морфизмом. Пусть $A = A_X \times A_Y$, где $A_X \in \mathfrak{A}_X$ и $A_Y \in \mathfrak{A}_Y$. Тогда $(f \times g)^{-1}(A) = \{z \in Z \mid f(z) \in A_X, g(z) \in A_Y\} = f^{-1}(A_X) \cap g^{-1}(A_Y) \in \mathfrak{A}_Z$. Положим $S = \{A = A_X \times A_Y \mid A_X \in \mathfrak{A}_X, A_Y \in \mathfrak{A}_Y\}$. По определению $\mathfrak{A}_X \otimes \mathfrak{A}_Y = R_\sigma(S)$. По лемме 0.41 имеем $R_\sigma((f \times g)^{-1}(S)) = (f \times g)^*(R_\sigma(S)) = (f \times g)^*(\mathfrak{A}_X \otimes \mathfrak{A}_Y)$. Так как $(f \times g)^{-1}(S) \subset \mathfrak{A}_Z$, то по минимальности (по теореме 0.6) имеем включение $R_\sigma((f \times g)^{-1}(S)) \subset \mathfrak{A}_Z$. Тогда $(f \times g)^*(\mathfrak{A}_X \otimes \mathfrak{A}_Y) \subset \mathfrak{A}_Z$ и по свойствам обратного образа (см. предложение 0.38) отображение $f \times g$ является морфизмом. □

### 0.3.6 Функтор борелевской $\sigma$-алгебры

Пусть $(X, \tau)$ — топологическое пространство. Минимальная $\sigma$-алгебра, порождённая системой открытых множеств $\tau$ называется борелевской $\sigma$-алгеброй, а её элементы называются борелевскими множествами. Мы будем обозначать её через $\mathcal{B}(\tau)$, а соответствующее измеримое пространство через $\mathrm{Bor}((X, \tau)) = (X, \mathcal{B}(\tau))$. Мы докажем, что конструкция $\mathrm{Bor}\colon \mathfrak{Top} \to \mathfrak{Meas}$, сопоставляющая топологическому пространству $(X, \tau)$ измеримое пространство $(X, \mathcal{B}(\tau))$, а непрерывному отображению $f$ его же как отображение множеств, функториальна.

**Лемма 0.43.** *Пусть* $f\colon X \to Y$ *— отображение множеств,* $\theta$ *— топология* $Y$, $\mathfrak{A}_X$ *—* $\sigma$*-алгебра подмножеств* $X$*. Пусть также* $S$ *— база топологии* $\theta$ *такая, что всякое открытое подмножество представляется в виде не более, чем счётного объединения элементов базы, и* $T$ *— предбаза топологии* $\theta$ *такая, что всякое открытое множество представляется в виде не более, чем счётного объединения конечных пересечений элементов* $T$*. Тогда следующие утверждения равносильны*

(1) *прообраз всякого элемента борелевской $\sigma$-алгебры $\mathcal{B}(\theta)$ лежит в $\mathfrak{A}_X$;*

(2) *прообраз всякого открытого множества лежит в $\mathfrak{A}_X$;*

(3) *прообраз всякого элемента базы $S$ лежит в $\mathfrak{A}_X$;*

(4) *прообраз всякого элемента предбазы $T$ лежит в $\mathfrak{A}_X$.*

*Доказательство.* Все пункты являются частным случаем пункта (1), а пункты (3) и (4) — пункта (2). Так как база топологии (с данным дополнительным условие) является частным случаем предбазы топологии (с дополнительным условием), то достаточно вывести из пункта (4) пункт (1).

Рассмотрим систему множеств $T_X = f^{-1}(T) \cup \{X\} := \{f^{-1}(U) | U \in T\} \cup \{X\}$. По условию $T_X \subset \mathfrak{A}_X$. Пусть $R_\sigma(T_X)$ — минимальная $\sigma$-алгебра, содержащая $T_X$. Так как $\mathfrak{A}_X$ является $\sigma$-алгеброй, то $R_\sigma(T_X) \subset \mathfrak{A}_X$. Из построения $f_*$ имеем $f_*(R_\sigma(T_X)) \subset f_*(\mathfrak{A}_X)$. Так же из построения $f_*$ имеем включение $T \subset f_*(R_\sigma(T_X))$. Из условия наложенного на $T$ следует, что минимальная $\sigma$-алгебра, порождённая $T$ совпадает с $\mathcal{B}(\theta)$. Тогда по минимальности (теореме 0.6) имеем включение $\mathcal{B}(\theta) \subset f_*(R_\sigma(T_X))$. Следовательно, $\mathcal{B}(\theta) \subset f_*(\mathfrak{A}_X)$ и по предложению 0.37 $f$ является морфизмом измеримых пространств $f\colon (X, \mathfrak{A}_X) \to (Y, \mathcal{B}(\theta))$, что и утверждается в пункте (1). □

**Теорема 0.44.** *Пусть $(X, \tau), (Y, \theta)$ — топологические пространства, $f\colon (X, \tau) \to (Y, \theta)$ — непрерывное отображение. Тогда отображение $f$ является морфизмом измеримых пространств $f\colon (X, \mathcal{B}(\tau)) \to (Y, \mathcal{B}(\theta))$.*

*Доказательство.* Следует из эквивалентности пунктов (1) и (2) леммы 0.43 для случая, когда $\mathfrak{A}_X = \mathcal{B}(\tau)$. □

## 0.4 Предварительные сведения из анализа Фурье

## 0.5 Предварительные сведения из линейной алгебры

### 0.5.1 Билинейные функции и квадратичные формы

Пусть $\Bbbk$ — некоторое поле (в нашем случае будут рассматриваться только поля вещественных чисел $\mathbb{R}$) и $V$ — векторное пространство над $\Bbbk$.

Отображение $B\colon V \times V \to \Bbbk$ называется билинейной функцией, если выполнены следующие аксиомы

(1) $\forall v, u, w \in V\ B(u + v, w) = B(u, w) + B(v, w)$;

(2) $\forall v, u \in V, \lambda \in \Bbbk\ B(\lambda u, v) = \lambda B(u, v)$;

(3) $\forall v, u, w \in V\ B(u, v + w) = B(u, w) + B(u, v)$;

(4) $\forall v, u \in V, \lambda \in \Bbbk\ B(u, \lambda v) = \lambda B(u, v)$.

Билинейная функция называется симметрической, если дополнительно для любых $u, v \in V$ выполнено $B(u,v) = B(v,u)$.

**Пример.** Пусть $V = \Bbbk$ и $B(a,b) = a \cdot b$, где $\cdot$ — умножение в поле $\Bbbk$. Тогда $B$ — симметрическая билинейная функция. ★

**Пример.** Пусть в векторном пространстве $V$ фиксирован базис $e_1, \ldots, e_n$. Тогда если $B(x,y) = \sum\limits_{i=1}^{n} x_i y_i$, где $x = \sum\limits_{i=1}^{n} x_i e_i$ и $y = \sum\limits_{i=1}^{n} y_i e_i$, то $B$ — также билинейная симметрическая форма. ★

Квадратичной формой называется отображение $Q\colon V \to \Bbbk$ такое, что для некоторой билинейной формы и любой вектора $v \in V$ имеет место равенство $Q(v) = B(v,v)$. Если $B$ — билинейная функция, то квадратичная форма $Q$, заданная формулой $Q(v) = B(v,v)$ называется квадратичной формой соответствующей билинейной функции $B$. Пусть $\Bbbk = \mathbb{R}$, $Q$ — квадратичная форма и для любого ненулевого вектора $v \in V$ выполнено неравенство $Q(v) > 0$. Тогда форма $Q$ называется положительно определённой. Если для любого $v \in V$ выполнено неравенство $Q(v) \geqslant 0$, то форма $Q$ называется неотрицательно определённой.

Симметрическую билинейную форму с положительно определённой соответствующей квадратичной формой называют скалярным произведением. Вместо $B(u,v)$ часто пишут $(u,v)$ или $\langle u, v\rangle$.

Примеры. Квадратичные формы, соответствующие билинейным функциям из примеров выше являются положительно определёнными.

**Теорема 0.45** (Коши, Буняковский, Шварц)**.** *Пусть $V$ — векторное пространство над полем $\mathbb{R}$ и $B$ — скалярное произведение на $V$. Тогда дл любых двух векторов $u, v \in V$ выполнено равенство*

$$B(u,v)^2 \leqslant B(u,u)B(v,v),$$

*причём равенство достигается тогда и только тогда, когда $u$ и $v$ коллинеарны.*

*Доказательство.* Рассмотрим вектор $u + tv$, где $t \in \mathbb{R}$ и значение квадратичной формы на нём. По билинейности, симметричности и положительной определённости имеем

$$B(u+tv, u+tv) = B(u,u) + tB(u,v) + tB(v,u) + t^2 B(v,v) = B(u,u) + 2tB(u,v) + t^2 B(v,v) \geqslant 0,$$

причём последнее равенство достигается тогда и только тогда, когда $u + tv = 0$.

Многочлен второй степени принимает только неотрицательные (положительные) значения тогда и только тогда, когда его дискриминант меньше или равен 0 (меньше 0). Итого

$$D = 4B(u,v)^2 - 4B(u,u)B(v,v) \leqslant 0 \Leftrightarrow B(u,v)^2 \leqslant B(u,u)B(v,v)$$

и $D = 0 \Leftrightarrow B(u,v)^2 = B(u,u)B(v,v)$. Последнее равносильно тому, что многочлен имеет корень $t$ и $u + tv = 0$, то есть $u$ и $v$ пропорциональны. □

Заметьте, что доказательство этого неравенства в случае поля комплексных чисел требует добавления дополнительной «поправки» $\lambda$.

### 0.5.2 Полуторалинейные функции

В этом подразделе будем рассматривать только векторные пространства над полем комплексных чисел.

Отображение $S\colon V \times V \to \Bbbk$ называется полуторалинейной функцией (по второму аргументу), если выполнены следующие аксиомы

(1) $\forall v, u, w \in V \ S(u+v, w) = S(u, w) + S(v, w)$;

(2) $\forall v, u \in V, \lambda \in \Bbbk \ S(\lambda u, v) = \lambda S(u, v)$;

(3) $\forall v, u, w \in V \ S(u, v+w) = S(u, w) + S(u, v)$;

(4) $\forall v, u \in V, \lambda \in \Bbbk \ S(u, \lambda v) = \overline{\lambda} S(u, v)$, где надчёркивание означает комплексное сопряжение.

Полуторалинейная функция называется эрмитовой, если для любых векторов $u$ и $v$ дополнительно выполнено равенство $S(u, v) = \overline{S(v, u)}$.

Эрмитова функция называется скалярным произведением, если для любого ненулевого вектора $v$ выполнено неравенство $S(v, v) > 0$.

**Теорема 0.46** (Коши, Буняковский, Шварц)**.** *Пусть $V$ — векторное пространство над полем $\mathbb{C}$ и $S$ — скалярное произведение на $V$. Тогда для любых двух векторов $u, v \in V$ выполнено равенство*

$$S(u, v)\overline{S(u, v)} \leqslant S(u, u) S(v, v),$$

*причём равенство достигается тогда и только тогда, когда $u$ и $v$ коллинеарны.*

*Доказательство.* Если $S(u, v) = 0$, то неравенство выполнено. При таком условии $u$ и $v$ пропорциональны тогда и только тогда, когда один из этих векторов равен 0. Последнее в свою очередь равносильно тому, что правая часть неравенства обращается в нуль. Далее будем считать, что $S(u, v) \neq 0$.

Рассмотрим вектор $u + t\lambda v$, где $t \in \mathbb{R}$ и $\lambda = S(u, v)$. Поскольку $S$ — скалярное произведение и из условий наложенных на $\lambda$, то

$$S(u + t\lambda v, u + t\lambda v) = S(u, u) + t\overline{\lambda} S(u, v) + t\lambda S(v, u) + t^2 \lambda\overline{\lambda} S(v, v) =$$

$$= S(u, u) + 2tS(u, v)S(v, u) + t^2 S(u, v)S(v, u)S(v, v) \leqslant 0$$

причём последнее равенство достигается тогда и только тогда, когда $u + t\lambda v = 0$.

Многочлен второй степени принимает только неотрицательные (положительные) значения тогда и только тогда, когда его дискриминант меньше или равен 0 (меньше 0). Итого

$$D = 4S(u, v)^2 S(v, u)^2 - 4S(u, u)S(v, v)S(u, v)S(v, u) \leqslant 0 \Leftrightarrow S(u, v)S(v, u) \leqslant S(u, u)S(v, v)$$

и $D = 0 \Leftrightarrow S(u, v)^2 = S(u, u)S(v, v)$. Последнее равносильно тому, что многочлен имеет корень $t_0$ и $u + t_0 S(u, v) v = 0$, то есть $u$ и $v$ пропорциональны. $\square$

## 0.6 Элементарная комбинаторика

### 0.6.1 Элементарные факты для подсчёта числа комбинаций

Центральной задачей комбинаторной математики можно считать задачу расположения объектов в соответствии со специальными правилами и нахождения числа способов, которыми это может быть сделано. Для нахождения числа способов полезны следующие результаты.

**Лемма 0.47.** *Мощность объединения двух конечных множеств $A$ и $B$ равна*

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

**Доказательство** проводится изображением диаграммы Эйлера-Венна и подсчётом элементов в её частях. □

**Лемма 0.48** (формула включений–исключений)**.**

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \ldots$$

$$\ldots + (-1)^{n-1}|A_1 \cap A_2 \cap \ldots \cap A_n|.$$

**Доказательство** проводится по индукции с помощью предыдущей леммы, которая используется как база и шаг индукции одновременно. □

**Лемма 0.49** (принцип умножения)**.** *Число пар, которые можно составить, когда первый элемент выбирается из множества $A$, а второй независимо от первого берется из множества $B$ равно $|A| \cdot |B|$.*

**Доказательство.** Следует из формулы $|A \times B| = |A| \cdot |B|$. □

**Лемма 0.50.** *Число кортежей $(a_1, \ldots, a_n)$ в которых независимо выбираются первый элемент $a_1$ из множества $A_1$, второй $a_2$ — из множества $A_2$, …, $a_n$ — из множества, равно $|A_1| \cdot |A_2| \cdot \ldots \cdot |A_n|$.*

**Доказательство.** По индукции из предыдущей леммы. □

### 0.6.2 Классические комбинаторные величины

Перестановка — это расположение $n$ различных объектов по различным местам. Число перестановок множества в $n$ элементов обозначаем $P_n$. Например, число способов разложить девять подписанных писем в девять подписанных конвертов — это число перестановок $P_9$.

Перестановка с повторениями — это расположение $n$ объектов, некоторые из которых могут быть тождественны друг другу, по различным местам. Число перестановок множества в $n$ элементов, где $n_1$ элементов 1-го типа, …, $n_k$ элементов

$k$-того типа (причём $n_1 + \ldots + n_k = n$) обозначаем $P(n_1, \ldots, n_k)$. Например, число способов составить из букв слова «математика» всевозможные (абстрактные) слова — это число перестановок с повторениями $P(2, 3, 2, 1, 1, 1)$.

Упорядоченная выборка (кортеж) в $m$ элементов из множества в $n$ элементов называется размещением, неупорядоченная выборка (подмножество) из $n$ элементов по $m$ называется сочетанием. Выборка может быть без повторений, если элементы повторяться не могут, и может быть с повторениями, если элементы в выборке повторяются. Число размещений без повторений обозначается $A_n^m$, с повторениями — $\widetilde{A}_n^m$. Число сочетаний без повторений обозначается $C_n^m$ или $\binom{n}{m}$, с повторениями — $\widetilde{C}_n^m$ или $\left(\binom{n}{m}\right)$.

Примеры:

(1) Число способов расставить шесть книг из имеющихся восьми на книжкой полке — это число размещений без повторения $A_8^6$.

(2) Число способов выбрать команду в пять человек из девяти данных игроков — это число сочетаний без повторения $C_9^5$.

(3) Число способов купить семь ручек из продающихся ручек пяти фирм — это число сочетаний с повторениями $\widetilde{C}_5^7$.

**Теорема 0.51.** *Числа перестановок, размещений и сочетаний находятся для неотрицательных целых $n, n_i, m$ по следующим формулам (для комбинаций без повторений дополнительно требуется $m \leqslant n$):*

| *Комбинация* | *без повторений* | *с повторениями* |
|---|---|---|
| *перестановка* | $P_n = n!$ | $P(n_1, \ldots, n_k) = \frac{(n_1+\ldots+n_k)!}{n_1! \cdot \ldots \cdot n_k!}$ |
| *размещение* | $A_n^m = n^{[m]} = \frac{n!}{(n-m)!}$ | $\widetilde{A}_n^m = n^m$ |
| *сочетание* | $C_n^m = \frac{n!}{m!(n-m)!}$ | $\widetilde{C}_n^m = C_{n+m-1}^m = \frac{(n+m-1)!}{m!(n-1)!}$ |

**Доказательство.**

Для $A_n^m$ применяем принцип умножения: $n \cdot (n-1) \cdot \ldots \cdot (n-m+1) = \frac{n!}{(n-m)!}$.

Для $\widetilde{A}_n^m$ также работает принцип умножения: $n \cdot n \cdot \ldots \cdot n = n^m$.

Для $P_n$ имеем $P_n = A_n^n = n!$.

Для $C_n^m$ замечаем, что выбрать $m$ и упорядочить их можно $A_n^m$ способами, причём каждый набор упорядочен $m!$ раз, поэтому $C_n^m = \frac{A_n^m}{m!} = \frac{n!}{m!\,(n-m)!}$.

Для $P(n_1, \ldots, n_k)$ аналогично предыдущему замечаем, что верно равенство $P_{n_1+\ldots+n_k} = P(n_1, \ldots, n_k) \cdot n_1! \cdot \ldots \cdot n_k!$, откуда $P(n_1, \ldots, n_k) = \frac{P_{n_1+\ldots+n_k}}{n_1! \cdot \ldots \cdot n_k!} = \frac{(n_1+\ldots+n_k)!}{n_1! \cdot \ldots \cdot n_k!}$.

Для $\widetilde{C}_n^m$ рассуждаем следующим образом. Необходимо выбрать из набора из $n$ различных типов предметов ровно $m$ предметов, причём каждый тип можно выбрать несколько раз (т.е. с повторениями), и порядок выбора не важен. Представим каждую из выбранных $m$ «штук» как звёздочку $(*)$, а разделители между типами — как палочки $(|)$. Всего типов $n$, значит между этими типами нужно разместить

$n-1$ разделитель. Расположение звёзд и палочек вдоль одной оси однозначно кодирует, сколько звёзд (предметов) взято каждого типа. Всего символов («звёзд» и «палочек») $= m + (n-1)$. Из них нужно выбрать $m$ позиций для звёзд (или, эквивалентно, $n-1$ позиций для палочек). Таким образом число всех конфигураций равно числу способов выбрать $m$ позиций из $(m+n-1)$, т.е. $\widetilde{C}_n^m = C_{n+m-1}^{\,m}$.□

### 0.6.3 Свойства комбинаторных величин

**Теорема 0.52** (о свойствах числа сочетаний)**.**

$$C_n^m = P(m, n-m) \quad (m \leqslant n);$$

$$C_n^m = C_n^{n-m} \quad (m \leqslant n);$$
$$C_n^m = \tfrac{n}{m} \cdot C_{n-1}^{m-1} \quad (m \leqslant n);$$
$$C_n^m + C_n^{m+1} = C_{n+1}^{m+1} \quad (m < n).$$

**Доказательство** проводится выписыванием с помощью факториалов левых и правых частей равенств и их сравнением. □

**Теорема 0.53** (о биноме Ньютона)**.**

$$(x+y)^n = \sum_{k=0}^{n} C_n^k x^{n-k} y^k =$$

$$= C_n^0 x^n + C_n^1 x^{n-1} y + \ldots + C_n^k x^{n-k} y^k + \ldots + C_n^n y^n.$$

**Доказательство.** Чтобы получить член $x^{n-k}y^k$ при раскрытии $n$ скобок $(x+y)$ нужно выбрать $k$ скобок, из которых возьмём $y$ (из остальных автоматически выбирается $x$). Следовательно, при раскрытии скобок будет ровно $C_n^k$ слагаемых вида $x^{n-k}y^k$. □

В связи с этой теоремой числа сочетаний $C_n^m = \binom{n}{m}$ называют биномиальными коэффициентами.

Бином Ньютона может быть обобщён до мультинома Ньютона — возведения в степень суммы произвольного числа слагаемых:

**Теорема 0.54** (Мультином Ньютона)**.**

$$(x_1 + x_2 + \cdots + x_k)^n = \sum_{\substack{n_j \geqslant 0 \\ n_1+n_2+\cdots+n_k=n}} P(n_1, \cdots, n_k) x_1^{n_1} \ldots x_k^{n_k}.$$

В связи с формулой из теоремы число $P(n_1, \cdots, n_k)$ также называется мультиномиальным (полиномиальным) коэффициентом и обозначается $\binom{n}{n_1, n_2, \ldots, n_k}$.

При $k = 2$ мультиномиальный коэффициент совпадает с биномиальным $\binom{n_1+n_2}{n_1, n_2} = \binom{n_1+n_2}{n_1} = \binom{n_1+n_2}{n_2}$.

**Урновая схема.** Пусть имеется $n$ урн, пронумерованных от 1 до $n$. Разместим в урны произвольным образом $m$ шаров, где $m \leqslant n$. Число способов разместить шары зависит от двух условий: (1) различимы ли шары или нет; (2) имеет ли место принцип исключения, который не позволяет положить второй шар в урну, уже содержащую один шар. Это число способов в зависимости от комбинации условий находится как одно из известных комбинаторных чисел:

| Комбинация | с принципом исключений | без принципа исключений |
|---|---|---|
| шары различимы | $A_n^m = n^{[m]} = \frac{n!}{(n-m)!}$ | $\widetilde{A}_n^m = n^m$ |
| шары неразличимы | $C_n^m = \frac{n!}{m!(n-m)!}$ | $\widetilde{C}_n^m = C_{n+m-1}^m = \frac{(n+m-1)!}{m!(n-1)!}$ |

**Число решений** $(x_1, \ldots, x_n)$ уравнения

$$x_1 + \ldots + x_n = k$$

в неотрицательных целых $x_i$ равно числу $\widetilde{C}_n^k$ сочетаний с повторениями из $n$ элементов по $k$.

**Задача о беспорядках.** Требуется найти число перестановок $(p_1, p_2, \ldots, p_n)$ множества $\{1, 2, \ldots, n\}$ таких что $p_i \neq i$ для всех $i$. Такие перестановки называются беспорядками.

**Теорема 0.55** (о числе беспорядков)**.** *Число беспорядков на множестве из $n$ элементов равно* $n!\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots + (-1)^n \frac{1}{n!}\right)$.

**Доказательство** следует из формулы включений-исключений. □

# 1 Вероятностное пространство, случайные события

Пусть $\Omega$ — некоторое множество, $\mathfrak{F}$ — $\sigma$-алгебра с единицей $\Omega$ и $\mathrm{P}$ — $\sigma$-аддитивная мера на $\mathfrak{F}$, удовлетворяющая свойству $\mathrm{P}(\Omega) = 1$. Тогда тройка $(\Omega, \mathfrak{F}, \mathrm{P})$ называется вероятностным пространством. Множество $\Omega$ называется пространством элементарных событий (исходов), элементы $\sigma$-алгебры $\mathfrak{F}$ называются событиями, а мера $\mathrm{P}$ вероятностной мерой.

Иногда мы будем называть вероятностное пространство «экспериментом» или «испытанием» или говорить, что «эксперименту» или «испытанию» соответствует вероятностное пространство, выражая таким образом физический смысл этого понятия: проходит эксперимент (испытание), у которого есть различные элементарные исходы. Эти исходы могут в результате этого эксперимента в разных комбинациях возникнуть с разной вероятностью.

Вероятностное пространство называется дискретным, если множество $\Omega$ не более, чем счётно.

Для кратности, если множество $\{\omega\}$ является событием, вместо $\mathrm{P}(\omega)$ будем писать $\mathrm{P}(\omega)$.

**Пример.** (Классическое определение вероятностей). В урне $M$ шаров, которые занумерованы числами от 1 до $M$. Из них выбираются номера $n$ шаров ($n < M$) по следующей схеме: из урны случайным образом вынимается один шар, фиксируется его номер, после чего шар возвращается обратно. В этом случае:

$$\Omega = \{\omega : \omega = (a_1, \ldots, a_n), \quad a_i \in \{1, 2, \ldots, M\}\},$$

$$N = N(\Omega) = M^n.$$

Рассмотрим событие $A = \{\omega : \omega = (a_1, \ldots, a_n),\ a_i \neq a_j\}$ — все номера различны. Тогда число благоприятных исходов равно

$$N(A) = M(M-1)(M-2)\cdots(M-n+1).$$

Вероятность события $A$ равна

$$P(A) = \frac{N(A)}{N} = \frac{M(M-1)(M-2)\cdots(M-n+1)}{M^n}.$$

Вероятность каждого отдельного элементарного исхода:

$$p(\omega) = \frac{1}{M^n}.$$

Событие $A$ можно интерпретировать как последовательное извлечение шаров из урны **без возвращения**, при этом каждый следующий номер определяется случайным образом из числа оставшихся. ★

**Пример.** Пусть $\Omega = [0, 1]$ и $\mathfrak{F} = M$ — $\sigma$-алгебра измеримых относительно меры Лебега подмножеств отрезка $[0, 1]$, $\mathrm{P} = \mu$ — классическая мера Лебега. Тогда $\mathrm{P}((\frac{1}{2}, \frac{3}{4})) = \frac{1}{4}$, $\mathrm{P}(\frac{2}{9}) = 0$ и $\mathrm{P}((0, \frac{1}{2}) \cup (\frac{2}{3}, 1)) = \frac{5}{6}$. ★

**Пример.** Пусть теперь $\Omega = [0, 2]$ и $\mathfrak{F} = M$ — $\sigma$-алгебра измеримых относительно меры Лебега подмножеств отрезка $[0, 2]$, $\mathrm{P} = \frac{1}{2}\mu$ — мера, пропорциональная мере Лебега (мы выбрали имеено такую меру, чтобы удовлетворить условию $\mathrm{P}(\Omega) = 1$). Тогда $\mathrm{P}((0, 1)) = \frac{1}{2}$, $\mathrm{P}(\frac{4}{5}) = 0$ и $\mathrm{P}((0, \frac{1}{2}) \cup (1, \frac{4}{3})) = \frac{5}{12}$. ★

**Пример.** Пусть $\Omega = \{1, 2, 3, 4, 5, 6\}$ — числа, возникающие при броске игральной кости. Будем считать, что все элементарные исходы равновероятны, то есть $\mathrm{P}(1) = \mathrm{P}(2) = \mathrm{P}(3) = \mathrm{P}(4) = \mathrm{P}(5) = \mathrm{P}(6) = \frac{1}{6}$. Тогда вероятность события $A = \{2, 4, 6\}$ — «выпало чётное число» равна $\mathrm{P}(A) = \mathrm{P}(2) + \mathrm{P}(4) + \mathrm{P}(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$. ★

Рассмотренный пример мотивирует нас ввести параллельные определения для дискретного пространства. Дискретным вероятностным пространством мы будем называть пару $(\Omega, \mathrm{P})$, где $\Omega = \{\omega_k\}_{k \in \mathbb{N}}$ — не более чем счётное множество (также называемое пространством элементарных исходов), а $\mathrm{P}\colon \Omega \to \mathbb{R}$ — неотрицательная функция, удовлетворяющая свойству $\sum\limits_{k \in \mathbb{N}} \mathrm{P}(\omega_k) = 1$. Говорят, что в этом случае на $\Omega$ заданы вероятности элементарных событий и что функция $\mathrm{P}$ задаёт на $\Omega$ распределение вероятностей. Событиями называются подмножества $\Omega$. Вероятностью события $A \subset \Omega$ называется величина

$$\mathrm{P}(A) = \sum_{\omega \in A} P(\omega),$$

которую мы также будем обозначать буквой $\mathrm{P}$. Последнее данное определение корректно, поскольку ряд в правой части сходится абсолютно.

**Предложение 1.1.** *Пусть* $(\Omega, \mathrm{P})$ *— дискретное вероятностное пространство в смысле последнего определения. Пусть* $\mathrm{P}\colon 2^{\Omega} \to \mathbb{R}$ *— функция, сопоставляющая событию его вероятность. Тогда тройка* $(\Omega, 2^{\Omega}, \mathrm{P})$ *является вероятностным пространством в смысле исходного определения.*

*Доказательство.* Множество $2^{\Omega}$ является $\sigma$-алгеброй, поэтому достаточно проверить, что функция $\mathrm{P}$ удовлетворяет аксиомам вероятностной меры.

Из определения $\mathrm{P}$ имеем

$$\mathrm{P}(\Omega) = \sum_{i=1}^{+\infty} \mathrm{P}(\omega_i) = 1.$$

Пусть $A, B \subset \Omega$ и $A \cap B = \varnothing$. Положим $A = \{\omega_i\}_{i \in I_A}$, $B = \{\omega_i\}_{i \in I_B}$ и $A \sqcup B = \{\omega_i\}_{i \in I_{A \sqcup B}}$. Поскольку $A$ и $B$ не пересекаются, то $I_A \sqcup I_B = I_{A \sqcup B}$. Тогда, так как ряды в формуле ниже сходятся абсолютно, имеем

$$\mathrm{P}(A \sqcup B) = \sum_{i \in I_{A \sqcup B}} \omega_i = \sum_{i \in I_A} \omega_i + \sum_{i \in I_B} \omega_i = \mathrm{P}(A) + \mathrm{P}(B).$$

Пусть теперь $\{A_k\}_{k \in \mathbb{N}}$ — счётное семейство непересекающихся подмножеств множества $\Omega$. Положим $A_k = \{\omega_i\}_{i \in I_k}$, $A = \bigsqcup\limits_{k \in I} A_k$. Снова, поскольку $A_k$ попарно не пересекаются, то $I = \bigsqcup\limits_{k \in \mathbb{N}} I_k$. Поскольку все ряды ниже сходятся абсолютно, то выполнены равенства

$$\mathrm{P}(A) = \sum_{i \in I} \mathrm{P}(\omega_i) = \sum_{k \in \mathbb{N}} \sum_{i \in I_k} \mathrm{P}(\omega_i) = \sum_{k \in \mathbb{N}} \mathrm{P}(A_k).$$

$\square$

Пусть $A, B \in \mathfrak{F}$ — события. Введём основные операции над событиями и приведём их классические наименования и обозначения в теории вероятностей.

Событие $\Omega \setminus A$ называется дополнением к событию $A$ и обозначается $\overline{A}$ («событие $A$ не произошло»).

Событие $A \cup B$ называется суммой событий $A$ и $B$ и обозначается $A + B$ («произошло событие $A$ или $B$»). В курсе лекций это обозначение использовалось для случаев, когда $A \cap B = \varnothing$.

Событие $A \cap B$ называется произведением событий $A$ и $B$ и обозначается $AB$ («произошло и событие $A$ и событие $B$»).

События $\Omega$ и $\varnothing$ называются достоверным и невозможным, соответственно.

Если $AB = \varnothing$, то события $A$ и $B$ называются несовместными. («события $A$ и $B$ не происходят одновременно»).

**Предложение 1.2** (Начальные свойства вероятностной меры)**.** *Пусть* $A, B, A_k \in \mathfrak{F}$ *— события. Тогда имеет место следующее:*

(1) $\mathrm{P}(\overline{A}) = 1 - \mathrm{P}(A)$;

(2) *если* $A \subset B$, *то* $\mathrm{P}(B \setminus A) = \mathrm{P}(B) - \mathrm{P}(A)$;

(3) *если* $A \subset B$, *то* $\mathrm{P}(A) \leqslant \mathrm{P}(B)$;

(4) $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(AB)$;

(5) $\mathrm{P}(A \cup B) \leqslant \mathrm{P}(A) + \mathrm{P}(B)$;

(6) $\mathrm{P}(\bigcup\limits_{k=1}^{n} A_k) = \sum\limits_{k=1}^{n} \left( \sum\limits_{i_1 < i_2 < \ldots < i_k} (-1)^{k-1} \mathrm{P}(A_{i_1} \ldots A_{i_k}) \right)$;

(7) $\mathrm{P}\left(\bigcup\limits_{k=1}^{+\infty} A_k\right) \leqslant \sum\limits_{k=1}^{+\infty} \mathrm{P}(A_k)$ *(это свойство называется субаддитивностью).*

*Доказательство.* Равенство (1) следует из цепочки

$$1 = \mathrm{P}(\Omega) = \mathrm{P}(A \sqcup \overline{A}) = \mathrm{P}(A) + \mathrm{P}(\overline{A}).$$

Равенство (2) — из цепочки

$$\mathrm{P}(B) = \mathrm{P}(A \cup (B \setminus A)) = \mathrm{P}(A) + \mathrm{P}(B \setminus A).$$

Неравенство (3) следует из этого равенства и неотрицательности вероятности.

Равенство (4) — из цепочки

$$\mathrm{P}(A \cup B) = \mathrm{P}((A \setminus B) \sqcup (A \cap B) \sqcup (B \setminus A)) =$$

$$= \mathrm{P}(A \setminus B) + \mathrm{P}(A \cap B) + \mathrm{P}(B \setminus A) + \mathrm{P}(A \cap B) - \mathrm{P}(A \cap B) =$$
$$= \mathrm{P}((A \setminus B) \sqcup (A \cap B)) + \mathrm{P}((B \setminus A) \sqcup (A \cap B)) - \mathrm{P}(A \cap B) =$$
$$= \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B).$$

Неравенство (5) немедленно следует из равенства (4).

Докажем (6) по индукции.

База $n = 2$ была доказана в пункте 3.

Докажем шаг. Положим $B = \bigcup\limits_{k=1}^{n-1} A_k$. По базе индукции

$$\mathrm{P}(B \cup A_n) = \mathrm{P}(B) + \mathrm{P}(A_n) - \mathrm{P}(BA_n).$$

Далее, положим $B_k = A_k A_n$. Тогда $BA_n = \bigcup\limits_{k=1}^{n-1} B_k$. По индукционному предположению вероятность $\mathrm{P}(B \cup A_n)$ равна

$$\sum_{k=1}^{n-1}\left(\sum_{i_1<i_2<\ldots<i_k}(-1)^{k-1}\,\mathrm{P}(A_{i_1}\ldots A_{i_k})\right)+\mathrm{P}(A_n)-\sum_{k=1}^{n-1}\left(\sum_{i_1<i_2<\ldots<i_k}(-1)^{k-1}\,\mathrm{P}(A_{i_1}\ldots A_{i_k}A_n)\right)=$$

$$=\sum_{k=1}^{n}\left(\sum_{i_1<i_2<\ldots<i_k}(-1)^{k-1}\,\mathrm{P}(A_{i_1}\ldots A_{i_k})\right).$$

Докажем неравенство (7). Положим $B_k = A_k \setminus \bigcup\limits_{i=1}^{k-1} A_i$. Тогда $\bigcup\limits_{k=1}^{+\infty} B_k = \bigcup\limits_{k=1}^{+\infty} A_k$, причём $B_k$ попарно не пересекаются и $\mathrm{P}(B_k) \leqslant \mathrm{P}(A_k)$ по (2). Тогда по $\sigma$-аддитивности имеем

$$\bigcup_{k=1}^{+\infty} A_k = \bigcup_{k=1}^{+\infty} B_k = \sum_{k=1}^{+\infty}\mathrm{P}(B_k) \leqslant \sum_{k=1}^{+\infty}\mathrm{P}(A_k).$$

□

# 2 Условные вероятности, формула Байеса, независимость событий

## 2.1 Условная вероятность

В задачах бывает полезно рассмотреть вероятность того, что произойдёт некоторое событие $B$ при условии, что произойдёт событие $A$. Пусть $\mathrm{P}(A) > 0$. Тогда вероятность $\mathrm{P}(B \mid A) = \frac{\mathrm{P}(AB)}{\mathrm{P}(A)}$ называется условной вероятностью события $B$ при условии того, что событие $A$ произойдёт с вероятностью $\mathrm{P}(A) > 0$. Вероятность $\mathrm{P}(B)$ также иногда называется априорной вероятностью события $B$.

**Предложение 2.1.** *Пусть* $(\Omega, \mathfrak{F}, \mathrm{P})$ *— вероятностное пространство. Пусть* $A \in \mathfrak{F}$ *— событие, удовлетворяющее условию* $\mathrm{P}(A) > 0$. *Тогда тройка*

$$(\Omega, \mathfrak{F}, \mathrm{P}|_A),$$

*где* $\mathrm{P}|_A(B) := \mathrm{P}(B \mid A) = \frac{\mathrm{P}(AB)}{\mathrm{P}(A)}$, *является вероятностным пространством.*

*Доказательство.* Достаточно проверить аксиомы вероятностной меры (аксиомы $\sigma$-аддитивной меры и равенство $\mathrm{P}|_A(\Omega) = 1$).

Так как обе величины $\mathrm{P}(AB)$ и $\mathrm{P}(A)$ неотрицательны (а последняя и вовсе положительна), то $\mathrm{P}(B \mid A) \leqslant 0$.

Справедливость упомянутого равенства выводится из определения условной вероятности:

$$\mathrm{P}|_A(\Omega) = \tfrac{\mathrm{P}(A\Omega)}{\mathrm{P}(A)} = \tfrac{\mathrm{P}(A)}{\mathrm{P}(A)} = 1.$$

Пусть $\{B_k\}_{k\in\mathbb{N}}$ — счётная последовательность попарно не пересекающихся элементов алгебры $\mathfrak{F}$. Тогда элементы последовательности $\{B_k \cap A\}_{k\in\mathbb{N}}$ также попарно не пересекаются. Тогда

$$\mathrm{P}|_A\left(\bigsqcup_{k=1}^{+\infty} B_k\right) = \tfrac{1}{\mathrm{P}(A)}\,\mathrm{P}\left(A \cap \bigsqcup_{k=1}^{+\infty} B_k\right) = \tfrac{1}{\mathrm{P}(A)}\,\mathrm{P}\left(\bigsqcup_{k=1}^{+\infty} AB_k\right) = \sum_{k=1}^{+\infty} \tfrac{\mathrm{P}(AB_k)}{\mathrm{P}(A)} = \sum_{k=1}^{+\infty} \mathrm{P}|_A(B_k).$$

□

**Следствие 2.2.** *Пусть $A \in \mathfrak{F}$ — событие, вероятность которого больше* $0$, *$B_1, B_2 \in \mathfrak{F}$. Тогда справедливы следующие свойства*

(1) *если $B_1 \supset A$, то* $\mathrm{P}(B_1 \mid A) = 1$;

(2) $\mathrm{P}(B_1 \cup B_2 \mid A) = \mathrm{P}(B_1 \mid A) + \mathrm{P}(B_2 \mid A) - \mathrm{P}(B_1 B_2 \mid A)$;

(3) *если $B_1$ и $B_2$ несовместны, то* $\mathrm{P}(B_1 + B_2 \mid A) = \mathrm{P}(B_1 \mid A) + \mathrm{P}(B_2 \mid A)$.

**Пример.** В урне $n$ белых и $N - n$ чёрных шаров. Рассматривается выбор без возвращения. Вычислим вероятность $P(A)$ события $A$ — извлечь при первой попытке белый шар.

$$P(A) = \frac{n}{N}.$$

Теперь найдём вероятность $P(B \mid A)$ события $B$ — извлечь при второй попытке белый шар, если первый извлечённый шар — белый.

$$P(B \mid A) = \frac{n-1}{N-1}.$$

★

## 2.2 Формула полной вероятности и формула Байеса

Теперь мы покажем, как связаны условные вероятности с вероятностями произведений событий, как можно вычислять вероятность события, зная его условные вероятности для несовместных событий (формула полной вероятности) и как можно вычислить условную вероятность «с переставленными причиной и следствием» (формула Байеса).

**Лемма 2.3.** *Пусть $A, B \in \mathfrak{F}$ — события и $\mathrm{P}(A), \mathrm{P}(B) > 0$. Тогда имеют место равенства*

$$\mathrm{P}(AB) = \mathrm{P}(A \mid B)\,\mathrm{P}(B) = \mathrm{P}(B \mid A)\,\mathrm{P}(A),$$

$$\mathrm{P}(A \mid B) = \tfrac{\mathrm{P}(B|A)\,\mathrm{P}(A)}{\mathrm{P}(B)}.$$

*Доказательство.* Первое равенство немедленно следует из определения условной вероятности, второе – немедленно из первого и предположения, что $\mathrm{P}(B) > 0$. $\square$

Второе равенство, доказанное в лемме иногда (особенно в школьных программах), называют формулой Байеса. Ниже, пользуясь этим простым свойством, мы докажем более общую формулу и в дальнейшем будем называть формулой Байеса её.

**Теорема 2.4** (Формула произведения вероятностей)**.** *Пусть $A_1, \ldots, A_n \in \mathfrak{F}$ — события. Если вероятности событий $A_2A_3\ldots A_n, \ldots, A_{n-1}A_n, A_n$ не равны нулю, то имеет место формула*

$$\mathrm{P}(A_1A_2\ldots A_n) = \mathrm{P}(A_1 \mid A_2\ldots A_n)\,\mathrm{P}(A_2 \mid A_3\ldots A_n)\ldots\mathrm{P}(A_{n-1} \mid A_n)\,\mathrm{P}(A_n).$$

*Если вероятности событий $A_1A_2\ldots A_n, A_1A_2\ldots A_{n-1}, \ldots, A_1$ не равны нулю, то имеет место формула*

$$\mathrm{P}(A_1A_2\ldots A_n) = \mathrm{P}(A_n \mid A_1\ldots A_{n-1})\,\mathrm{P}(A_{n-1} \mid A_1\ldots A_{n-2})\ldots\mathrm{P}(A_2 \mid A_1)\,\mathrm{P}(A_1).$$

*Доказательство.* Докажем первое утверждение индукцией по $n$, второе получается из первого перестановкой индексов в обратном порядке.

База: $n = 2$ есть определение условной вероятности.

Докажем шаг индукции. Пусть для $n-1$ утверждение выполнено. Положим $B = A_2A_3\ldots A_n$. Тогда по базе индукции (здесь мы пользуемся тем, что $\mathrm{P}(B) > 0$) и затем по индукционному предположению (а здесь всеми остальными условиями) имеем

$$\mathrm{P}(A_1B) = \mathrm{P}(A_1 \mid B)\,\mathrm{P}(B) = \mathrm{P}(A_1 \mid A_2A_3\ldots A_n)\,\mathrm{P}(A_2 \mid A_3\ldots A_n)\ldots\mathrm{P}(A_{n-1} \mid A_n)\,\mathrm{P}(A_n).$$

$\square$

Набор событий $A_1, \ldots, A_n \in \mathfrak{F}$ называется разбиением пространства $\Omega$ (или просто «разбиение $\Omega$»), если $\mathrm{P}(A_i) > 0$ для каждого $i$, $A_i$ попарно несовместны ($A_iA_j = \varnothing$ при $i \neq j$) и $A_1 + A_2 + \ldots + A_n = \Omega$.

**Пример.** Нужно перечислить все возможные разбиения пространства исходов $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Действуя перебором, получим:

(1) $\{\omega_1, \omega_2, \omega_3\} = \Omega$;

(2) $\{\omega_1\} \cup \{\omega_2\} \cup \{\omega_3\} = \Omega$;

(3) $\{\omega_1\} \cup \{\omega_2, \omega_3\} = \Omega$;

(4) $\{\omega_2\} \cup \{\omega_1, \omega_3\} = \Omega$;

(5) $\{\omega_3\} \cup \{\omega_1, \omega_2\} = \Omega$.

Тогда общее число разбиений $d(3)$ для пространства $\Omega$, состоящего из трёх элементов, будет равно 5. ★

**Теорема 2.5** (Формула полной вероятности)**.** *Пусть* $A_1, \ldots, A_n \in \mathfrak{F}$ *— разбиение* $\Omega$*. Тогда для всякого события* $B$ *имеет место равенство формула полной вероятности*

$$\mathrm{P}(B) = \sum_{i=k}^{n} \mathrm{P}(B \mid A_k)\,\mathrm{P}(A_k).$$

*Доказательство.* Так как события $A_k$ попарно несовместны, то события $A_k B$ также попарно несовместны. Имеем цепочку равенств

$$\mathrm{P}(B) = \mathrm{P}(B\Omega) = \mathrm{P}(B(A_1 + \ldots + A_n)) = \mathrm{P}(BA_1 + \ldots + BA_n) \overset{\text{несовместность}}{=}$$

$$\overset{\text{несовместность}}{=} \sum_{k=1}^{n} \mathrm{P}(BA_k) = \sum_{k=1}^{n} \mathrm{P}(B \mid A_k)\,\mathrm{P}(A_k).$$

□

Формула полной вероятности остаётся справедливой, если отказаться от требования $A_1 + A_2 + \ldots + A_n = \Omega$ и заменить его на условие $B \subset A_1 + \ldots + A_n$ (сохраняя требования попарной несовместности событий $A_i$ и $\mathrm{P}(A_i) > 0$).

**Пример.** Пусть снова в урне $n$ белых и $N - n$ чёрных шаров. Рассматривается выбор без возвращения.

Вероятность $P(A)$ события $A$ — извлечь белый шар на первом шаге равна

$$P(A) = \frac{n}{N}.$$

Вероятность $P(B \mid A)$ события $B$ — извлечь белый шар на втором шаге, если на первом шаге тоже был извлечён белый шар равна

$$P(B \mid A) = \frac{n-1}{N-1}.$$

Теперь определим вероятность $P(B)$ события $B$ — извлечь на втором шаге белый шар, если цвет первого извлечённого шара неизвестен. Рапишем вероятность $P(\bar{A})$ события $\bar{A}$ — извлечь чёрный шар на первом шаге:

$$P(\bar{A}) = \frac{N-n}{N}.$$

Затем найдём вероятность $P(B \mid \bar{A})$ — извлечь белый шар на втором шаге, если на первом шаге был извлечён чёрный шар:

$$P(B \mid \bar{A}) = \frac{n}{N-1}.$$

Тогда по формуле полной вероятности имеем:

$$P(B) = P(B \mid A) \cdot P(A) + P(B \mid \bar{A}) \cdot P(\bar{A}) = \frac{n-1}{N-1} \cdot \frac{n}{N} + \frac{n}{N-1} \cdot \frac{N-n}{N}.$$

После упрощения поучаем

$$P(B) = \frac{n(n-1)}{(N-1)N} + \frac{n(N-n)}{(N-1)N} = \frac{n}{N}.$$

Таким образом, то обстоятельство, что цвет первого извлечённого шара неизвестен, не изменило вероятность того, что извлечённый на втором шаге шар оказался белым. ★

**Теорема 2.6** (Формула Байеса)**.** *Пусть события $A_1, \ldots, A_n \in \mathfrak{F}$ образуют разбиение $\Omega$, пусть $B \in \mathfrak{F}$ — ещё одно событие и $\mathrm{P}(B) > 0$. Тогда справедлива формула Байеса*

$$\mathrm{P}(A_i \mid B) = \frac{\mathrm{P}(B|A_i)\,\mathrm{P}(A_i)}{\sum\limits_{k=1}^{n} \mathrm{P}(B|A_k)\,\mathrm{P}(A_k)}.$$

*Доказательство.* По лемме 2.3 («простейшая формула Байеса») имеем равенство

$$\mathrm{P}(A_i \mid B) = \frac{\mathrm{P}(B|A_i)\,\mathrm{P}(B)}{\mathrm{P}(B)}.$$

По формуле полной вероятности имеем

$$\mathrm{P}(B) = \sum_{i=k}^{n} \mathrm{P}(B \mid A_k)\,\mathrm{P}(A_k),$$

откуда следует искомая формула. □

**Пример.** Приведём стандартный пример на применение простейшего вида формулы Байеса. Пусть в популяции заболевание встречается с вероятностью $\mathrm{P}(\text{Б}) = 0.1$. ПЦР-тест на выявление заболевания устроен так, что:

- При наличии заболевания он даёт положительный результат с вероятностью $\mathrm{P}(+ \mid \text{Б}) = 0.9$,
- При отсутствии заболевания он даёт ложноположительный результат с вероятностью $\mathrm{P}(+ \mid \text{З}) = 0.2$.

Найдём вероятность того, что человек действительно болен, если результат теста оказался положительным. По формуле Байеса:

$$\mathrm{P}(\text{Б} \mid +) = \frac{\mathrm{P}(+ \mid \text{Б}) \cdot \mathrm{P}(\text{Б})}{\mathrm{P}(+ \mid \text{Б}) \cdot \mathrm{P}(\text{Б}) + \mathrm{P}(+ \mid \text{З}) \cdot \mathrm{P}(\text{З})}.$$

Подставим известные значения:

$$\mathrm{P}(\text{Б} \mid +) = \frac{0.99 \cdot 0.01}{0.99 \cdot 0.01 + 0.05 \cdot 0.99} = \frac{0.0099}{0.0099 + 0.0495} = \frac{0.0099}{0.0594} \approx 0.1667.$$

Формально мы рассматривали дискретное вероятностное пространство $\Omega = \{+\text{Б}, -\text{З}, +\text{З}, -\text{Б}\}$ с четырьмя элементарными исходами, выражающими все возможные комбинации результата тестирования и реального состояния тестируемого человека. Здесь событие Б «болен» являлось объединением $\{+\text{Б}, -\text{Б}\}$, событие «пцр-тест дал положительный результат» — объединением $\{+\text{Б}, +\text{З}\}$ и так далее. ★

## 2.3 Независимость событий

Интуиция говорит нам, что события $A$ и $B$ «независимы», когда от того с какой вероятностью произойдёт событие $A$ не зависит вероятность того, что произойдёт событие $B$ и наоборот. Математически это выражается формулами $\mathrm{P}(B \mid A) = \mathrm{P}(B)$ и $\mathrm{P}(A \mid B) = \mathrm{P}(A)$. Чтобы не ограничиваться случаями, когда вероятности событий больше 0, мы определим независимость следствием формул выше. События $A$ и $B$ называются независимыми, если справедливо равенство $\mathrm{P}(AB) = \mathrm{P}(A)\,\mathrm{P}(B)$.

**Предложение 2.7** (Свойства независимости)**.** *Имеют место следующие утверждения:*

(1) *если* $\mathrm{P}(B) > 0$, *то независимость* $A$ *и* $B$ *равносильна равенству* $\mathrm{P}(A \mid B) = \mathrm{P}(A)$*;*

(2) *если* $A$ *и* $B$ *независимы, то* $\overline{A}$ *и* $B$ *независимы;*

(3) *если события* $B_1$ *и* $B_2$ *несовместны,* $A$ *и* $B_1$ *независимы, а также* $A$ *и* $B_2$ *независимы, то* $A$ *и* $B_1 + B_2$ *независимы.*

*Доказательство.* Проверим (1). Если $\mathrm{P}(B) > 0$, то по независимости имеем

$$\mathrm{P}(A \mid B) = \tfrac{\mathrm{P}(AB)}{\mathrm{P}(B)} = \tfrac{\mathrm{P}(A)\,\mathrm{P}(B)}{\mathrm{P}(B)} = \mathrm{P}(A).$$

Обратно, если $\mathrm{P}(A \mid B) = \mathrm{P}(A)$, то $\frac{\mathrm{P}(AB)}{\mathrm{P}(B)} = \mathrm{P}(A)$, откуда $\mathrm{P}(AB) = \mathrm{P}(A)\,\mathrm{P}(B)$.

Для доказательства (2) выпишем цепочку равенств

$$\mathrm{P}(\overline{A}B) = \mathrm{P}((\Omega \setminus A)B) = \mathrm{P}(B \setminus AB) = \mathrm{P}(B) - \mathrm{P}(AB) \overset{\text{независимость}}{=}$$

$$\overset{\text{независимость}}{=} \mathrm{P}(B) - \mathrm{P}(A)\,\mathrm{P}(B) = \mathrm{P}(B)(1 - \mathrm{P}(A)) = \mathrm{P}(B)\,\mathrm{P}(\overline{A}).$$

Наконец, для доказательства (3) заметим, что события $B_1A$ и $B_2A$ несовместны. Тогда

$$\mathrm{P}((B_1 + B_2)A) = \mathrm{P}(B_1A + B_2A) = \mathrm{P}(B_1A) + \mathrm{P}(B_2A) = \mathrm{P}(B_1)\,\mathrm{P}(A) + \mathrm{P}(B_2)\,\mathrm{P}(A) =$$
$$= (\mathrm{P}(B_1) + \mathrm{P}(B_2))\,\mathrm{P}(A) = \mathrm{P}(B_1 + B_2)\,\mathrm{P}(A).$$

□

Теперь определим независимость для набора событий. Пусть $B_1, \ldots, B_n \in \mathfrak{F}$ — события. Будем говорить, что они попарно независимы, если для всяких двух индексов $i \neq j$ выполнено равенство $\mathrm{P}(B_iB_j) = \mathrm{P}(B_i)\,\mathrm{P}(B_j)$ (то есть $B_i$ и $B_j$ независимы). Будем называть эти события независимыми, если для всякого набора индексов $i_1 < \ldots < i_k$ (здесь $2 \leqslant k \leqslant n$) имеет место равенство

$$\mathrm{P}\left(\bigcap_{s=1}^{k} B_{i_s}\right) = \prod_{s=1}^{k} \mathrm{P}(B_{i_s}).$$

**Предложение 2.8.** *Если события $B_1, \ldots, B_n$ независимы, то они попарно независимы.*

**Пример.** Вообще говоря из попарной независимости не следует независимость, что демонстрируется следующим примером. Рассмотрим тетраэдр, три грани которого покрашены в красный, зелёный и синий цвета, соответственно, а последняя разбита на три треугольника, покрашенных в те же цвета. Пусть вероятности выпадения граней равны $\frac{1}{4}$. покажем, что события «выпала грань с цветом $A$», где $A$ — цвет попарно независимы, но не являются таковыми в совокупности. Формально ситуация выглядит записывается так: $\Omega = \{\omega_R, \omega_G, \omega_B, \omega_{RGB}\}$ — элементарное событие — выпала грань с данной раскраской. По условию $\mathrm{P}(\omega_R) = \mathrm{P}(\omega_G) = \mathrm{P}(\omega_B) = \mathrm{P}(\omega_{RGB}) = \frac{1}{4}$. Обозначим через $R = \{\omega_R, \omega_{RGB}\}$ ($G = \{\omega_G, \omega_{RGB}\}, B = \{\omega_B, \omega_{RGB}\}$) события «выпала грань с красным (зелёным, синим) цветом», соответственно. Тогда $\mathrm{P}(R) = \mathrm{P}(G) = \mathrm{P}(B) = \frac{1}{2}$, $\mathrm{P}(RG) = \mathrm{P}(GB) = \mathrm{P}(BR) = \frac{1}{4} = \mathrm{P}(R)\,\mathrm{P}(G) = \mathrm{P}(G)\,\mathrm{P}(B) = \mathrm{P}(B)\,\mathrm{P}(R)$, но $\mathrm{P}(RGB) = \mathrm{P}(\omega_{RGB}) = \frac{1}{4} \neq \mathrm{P}(R)\,\mathrm{P}(G)\,\mathrm{P}(B) = \frac{1}{8}$. ★

**Пример.** Пользуясь примером выше, можно показать, что условие несовместности в пункте (3) предложения 2.7 нельзя опустить. Положим $A = R$ и $B_1 = G$ и $B_2 = B$. Тогда $\mathrm{P}((B_1 \cup B_2)A) = \mathrm{P}(\omega_{RGB}) = \frac{1}{4}$, но $\mathrm{P}(A) = \frac{1}{2}$, $\mathrm{P}(B_1 + B_2) = \frac{3}{4}$ и $\frac{1}{2} \cdot \frac{3}{4} \neq \frac{1}{4}$. Таким образом, события $B_1 + B_2$ и $A$ не являются независимыми. ★

**Пример.** Покажем, что из условия независимости нельзя удалить ни одно из равенств. Более того, мы докажем, что для всякого натурального $n$ и семейства наборов индексов $S_J = \{(i_{1,j}, \ldots, i_{k_j,j})\}_{j \in J}$ можно построить пример вероятностного пространства $(\Omega, \mathfrak{F}, \mathrm{P})$ и событий $A_1, \ldots, A_n \in \mathfrak{F}$ для которых множество наборов, на которых выполнены равенства

$$\mathrm{P}\left(\bigcap_{s=1}^{k} A_{i_s}\right) = \prod_{s=1}^{k} \mathrm{P}(A_{i_s})$$

в точности совпадает с $J$.

Построим пример для дискретного вероятностного пространства. Положим $\Omega = \{\,(\varepsilon_1, \ldots, \varepsilon_n) |\, \varepsilon_i \in \{0, 1\}\}$ — множество всех кортежей из нулей и единиц длины $n$, $\mathrm{P}((\varepsilon_1, \ldots, \varepsilon_n)) = p_{(\varepsilon_1,\ldots,\varepsilon_n)}$ — будущее распределение вероятностей. Также положим $A_k = \{\,(\varepsilon_1, \ldots, \varepsilon_n) |\, \varepsilon_i \in \{0, 1\}, \varepsilon_k = 1\}$. Рассмотрим отображение $\varphi\colon \mathbb{R}^{2^n} \to \mathbb{R}^{2^n}$, заданное в некоторых фиксированных базисах этих пространств по правилу

$$\varphi\colon \begin{pmatrix} \cdots \\ p_{(\varepsilon_1,\ldots,\varepsilon_n)} \\ \cdots \end{pmatrix} \mapsto \begin{pmatrix} \cdots \\ \mathrm{P}(A_{i_1} \ldots A_{i_k}) \\ \cdots \end{pmatrix},$$

где для $k = 0$ предполагается, что в матрице стоит $\mathrm{P}(\Omega)$. Поскольку $\mathrm{P}(A_{i_1} \ldots A_{i_k}) = \mathrm{P}(\{\,(\varepsilon_1, \ldots, \varepsilon_n) |\, \varepsilon_i \in \{0, 1\}, \varepsilon_{i_s} = 1, 1 \leqslant s \leqslant k\}) = \sum\limits_{\varepsilon_{i_s}=1, 1\leqslant s\leqslant k} p_{(\varepsilon_1,\ldots,\varepsilon_n)}$, то $\varphi$ — линейное отображение. Можно показать, что $\varphi$ сюръективно (проверьте с помощью элементарных преобразований, что его матрица имеет ранг $2^n$) и, следовательно, биективно. Достаточно подобрать значения вероятностей все возможных произведений $A_i$ так, чтобы вероятности $p_{(\varepsilon_1,\ldots,\varepsilon_n)}$ были неотрицательны, в сумме давали $1$ ($\mathrm{P}(\Omega) = 1$) и при этом выполнялись в точности все желаемые равенства на вероятности произведений событий $A_i$. Положим $\mathrm{P}(A_i) = \frac{1}{2^{2n}}$, $\mathrm{P}(\Omega) = 1$. Если $(i_1, \ldots, i_k) \in S_J$, то положим $\mathrm{P}(A_{i_1} \ldots A_{i_k}) = \frac{1}{2^{2kn}}$. Иначе положим $\mathrm{P}(A_{i_1} \ldots A_{i_k}) = \frac{1}{2^{2kn+1}}$. Проверим, что имеют место неравенство $\frac{1}{2^{2kn+2}} \leqslant p_{(\varepsilon_1,\ldots,\varepsilon_n)} \leqslant \frac{1}{2^{2kn}}$, для кортежей с $k > 0$ числом единиц. Для кортежа $(1, \ldots, 1)$ неравенство выполнено по построению. Докажем неравенства для оставшихся кортежей с данным условием индукцией по числу нулей в кортеже. Пусть в текущем кортеже $(\varepsilon_1, \ldots, \varepsilon_n)$ присутствует $n - k \geqslant 1$ нулей. Прибавим к $p_{(\varepsilon_1,\ldots,\varepsilon_n)}$ все остальные значения вероятностей элементарных исходов — кортежей, в которых некоторые нули из данного кортежа заменены на единицы. Тогда $p_{(\varepsilon_1,\ldots,\varepsilon_n)} \leqslant \frac{1}{2^{2kn}}$, так как по предположению индукции все остальные слагаемые положительны, а сумма не превосходит $\frac{1}{2^{2kn}}$. С другой стороны, $p_{(\varepsilon_1,\ldots,\varepsilon_n)} \geqslant \frac{1}{2^{2kn+1}} - \frac{2^k-1}{2^{2(k+1)n}} = \frac{1}{2^{2kn+1}} - \frac{1}{2^{2kn+2+(2n-k-2)}}$. Так как $n - k - 1 \leqslant 0$ и $n - 1 \leqslant 0$, то последнее слагаемое по модулю не превосходит $\frac{1}{2^{2kn+2}}$. Тогда $p_{(\varepsilon_1,\ldots,\varepsilon_n)}$ не меньше $\frac{1}{2kn+2}$. Остаётся убедиться в том, что $p_{(0,\ldots,0)} = 1 - \sum\limits_{(\varepsilon_1,\ldots,\varepsilon_n)\neq(0,\ldots,0)} p_{(\varepsilon_1,\ldots,\varepsilon_n)} \geqslant 1 - \frac{2^n-1}{2^{2n}} > 0$. ★

**Предложение 2.9.** *Пусть событие $A$ независимо с самим собой тогда и только тогда, когда* $\mathrm{P}(A) = 0$ *или* $\mathrm{P}(A) = 1$.

*Доказательство.* По определению независимости $\mathrm{P}(AA) = \mathrm{P}(A)\,\mathrm{P}(A)$, откуда $\mathrm{P}(A)^2 - \mathrm{P}(A) = 0$ и либо $\mathrm{P}(A) = 0$, либо $\mathrm{P}(A) = 1$. Обратно, если одно из этих равенств выполнено, то $\mathrm{P}(A)\,\mathrm{P}(A) = \mathrm{P}(A) = \mathrm{P}(AA)$. □

## 2.4 Произведение вероятностных пространств

Если считать, что каждый исход получен в результате отдельного испытания, то мы обнаружим, что любое событие, относящееся к фиксированному испытанию будет независимым от любого события, относящегося к другим испытаниям. В таких случаях говорят о последовательности независимых испытаний.

Формализуем данную ситуацию. Рассмотрим два вероятностных пространства $(\Omega_1, \mathfrak{F}_1, \mathrm{P}_1)$ и $(\Omega_2, \mathfrak{F}_2, \mathrm{P}_2)$. Их произведением мы будем называть вероятностное пространство $(\Omega, \mathfrak{F}, \mathrm{P})$, где $\Omega = \Omega_1 \times \Omega_2$ — декартово произведение, $\mathfrak{F} = \mathfrak{F}_1 \otimes \mathfrak{F}_2$ — $\sigma$-алгебра, порождённая всеми возможными произведениями $A_1 \times A_2$, где $A_1 \in \mathfrak{F}_1, A_2 \in \mathfrak{F}_2$ — события в исходных вероятностных пространствах (заметьте, что сами по себе таким произведения образуют полукольцо, но вообще говоря не образуют даже кольца), а мера $\mathrm{P}$ определяется как продолжение меры $m$ с полукольца произведений, заданной по правилу $m(A_1 \times A_2) := \mathrm{P}_1(A_1) \cdot \mathrm{P}_2(A_2)$.

Пусть теперь $(\Omega, \mathfrak{F}, \tilde{\mathrm{P}})$ — вероятностное пространство с теми же $\Omega$ и $\mathfrak{F}$, то возможно другой вероятностной мерой. Это можно представлять себе так: происходят два испытания и результаты одного могу «повлиять» на результаты другого. Неформальное «повлиять» выражается понятием независимости испытаний. Мы будем говорить, что испытания, которым соответствуют вероятностные пространства $(\Omega_1, \mathfrak{F}_1, \mathrm{P}_1)$ и $(\Omega_2, \mathfrak{F}_2, \mathrm{P}_2)$, а составному эксперименту, состоящему из двух этих испытаний, независимы, если для любых двух событий $A_1 \in \mathfrak{F}_1$ и $A_2 \in \mathfrak{F}_2$ выполнено равенство

$$\tilde{\mathrm{P}}(A_1 \times A_2) = \mathrm{P}_1(A_1)\,\mathrm{P}_2(A_2) = \tilde{\mathrm{P}}(A_1 \times \Omega_2)\tilde{\mathrm{P}}(\Omega_2 \times A_2).$$

# 3 Случайные величины, их распределения, функции распределения и плотности

## 3.1 Случайные величины

Пусть $(\Omega, \mathfrak{F}, \mathrm{P})$ — вероятностной пространство. Функция $\xi \colon \Omega \to \mathbb{R}$ называется случайной величиной, если прообраз любого борелевского множества $B \in \mathcal{B}$ при отображении $f$ лежит в $\mathfrak{F}$. Соотнося это из определениями из действительного анализа, мы можем сказать, что $\xi$ является измеримой функцией. Также, вспомним, что мы получим эквивалентное определение, если вместо всех борелевских множеств будем рассматривать все интервалы, все отрезки, все интервалы одного из двух видов $(-\infty, b)$ и $(a, +\infty)$ или все полуинтервалы одного из двух видов $(-\infty, b]$ и $[a, +\infty)$.

Случайные величины удовлетворяют всем тем стандартным свойствам, которым удовлетворяют измеримые функции:

**Предложение 3.1.** *Пусть $\xi$ — случайная величины и $g$ — непрерывная на $\operatorname{Im} \xi$ функция. Тогда композиция $g(\xi)$ является случайной величиной.*

*Доказательство.* Переформулировка предложения 0.22. □

**Предложение 3.2.** *Пусть $\xi, \eta$ — случайные величины. Тогда множество $A_{\xi \leqslant \eta} = \{\omega \in \Omega |\, \xi(\omega) \leqslant \eta(\omega)\}$ измеримо. Функции $a + \xi, a\xi, |\xi|, \xi^2, \xi + \eta$ и $\xi\eta$, где $a$ — константа, являются случайными величинами. Если случайная величина $\eta$ не принимает значения 0, то функции $\frac{1}{\eta}$ и $\frac{\xi}{\eta}$ также являются случайными величинами.*

*Доказательство.* Переформулировка предложения 0.23. □

Случайная величина называется дискретной, если она принимает не более, чем счётное число различных значений. Случайные величины на дискретном вероятностном пространстве всегда дискретны. Кроме того, всякая функция на дискретном вероятностном пространстве (в смысле второго определения) является случайной величиной.

**Пример.** Пусть $\Omega = \{\text{ГГ}, \text{ГР}, \text{РГ}, \text{РР}\}$ — все возможные результаты выпадения на двух монетках с гербом (Г) на одной стороне и решкой (Р) на другой. Будем считать, что вероятности всех элементарных событий равны $\frac{1}{4}$. Пусть $\xi\colon \Omega \to \{0, 1\}$ сопоставляет элементам ГГ, ГР и РГ единицу, а РР — ноль. Тогда $\xi$ — дискретная случайная величина. Мы можем интерпретировать 0 как «неудачу» в эксперименте «бросить две монетки и получить хотя бы один герб» и 1 — «успех». В нашем случае $\xi$ принимает значение 1 с вероятностью $\mathrm{P}(\xi^{-1}(\{1\})) = \mathrm{P}(\{\text{ГГ}, \text{ГР}, \text{РГ}\}) = \frac{3}{4}$ и значение 0 с вероятностью $\mathrm{P}(\xi^{-1}(\{0\})) = \mathrm{P}(\{\text{РР}\}) = \frac{1}{4}$. ★

Вместо $\mathrm{P}(\xi^{-1}(\{a\}))$ мы будем пользоваться записью $\mathrm{P}(\xi = a)$, отражающей смысл этого выражения — «вероятность случайной величины $\xi$ принять значение $a$». Также вместо перегруженных скобками выражений $\mathrm{P}(\xi^{-1}((-\infty; b))), \mathrm{P}(\xi^{-1}((a; +\infty)))$, $\mathrm{P}(\xi^{-1}((a; b)))$ и им подобных мы будем писать $\mathrm{P}(\xi \leqslant b), \mathrm{P}(\xi \geqslant a)$ и $\mathrm{P}(a \leqslant \xi \leqslant b)$.

**Пример.** В примере выше $\mathrm{P}(\xi = 1) = \frac{3}{4}$ и $\mathrm{P}(\xi = 0) = \frac{1}{4}$. ★

**Пример.** Дискретная случайная велиичина может быть задана и не на дискретном вероятностном пространстве. Пусть $\Omega = [-1; 1]$ — отрезок, $\mathfrak{F} = \mathcal{B}$ — сигма алгебра его борелевских подмножеств, а вероятностная мера — это умноженная на $\frac{1}{2}$ классическая мера Лебега $\mu$. Тогда случайная величина $\xi(\omega) = sgn(\omega)$, заданная функцией знак, является дискретной случайной величиной. ★

**Пример.** Вернёмся одному из первых рассмотренных примеров. Пусть $\Omega = [0, 2]$ и $\mathfrak{F} = M$ — $\sigma$-алгебра измеримых относительно меры Лебега подмножеств отрезка $[0, 2]$, $\mathrm{P} = \frac{1}{2}\mu$ — мера, пропорциональная мере Лебега. Пусть $\xi(\omega) = \omega^2$. Эта функция непрерывна, поэтому является случайной величиной (измеримой функцией). Найдём $\mathrm{P}(\frac{1}{4} \leqslant \xi < \frac{16}{9})$. Так как неравенства $\frac{1}{4} \leqslant \xi(\omega) < \frac{16}{9}$ выполнены тогда и только тогда, когда $\frac{1}{2} \leqslant \omega < \frac{4}{3}$. Тогда $\mathrm{P}(\frac{1}{4} \leqslant \xi < \frac{16}{9}) = \frac{1}{2}(\frac{4}{3} - \frac{1}{2}) = \frac{5}{12}$. ★

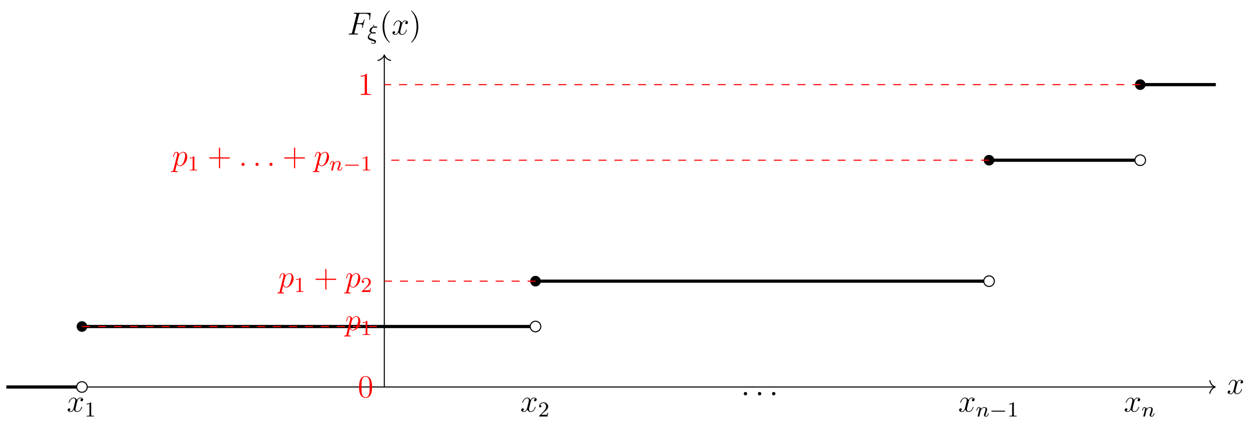
## 3.2 Функции распределения

Введём некоторые объекты, характеризующие случайную величину $\xi$ через её вероятность принять некоторые значения. Функцией распределения случайной величины $\xi$ будем называть функцию $F_\xi\colon \mathbb{R} \to \mathbb{R}$, заданную по правилу $F_\xi(x) = \mathrm{P}(\xi \leqslant x)$. Заметим, что построенная функция $F_\xi$ корректно определена на всей числовой прямой, так как $\xi$ — случайная величина и прообраз бесконечного полуинтервала относительно $\xi$ является событием (измерим).

**Пример.** Функция распределения дискретной случайной величины $\xi$ имеет «ступенчатую форму». Пусть $\xi$ принимает конечное число значений $x_1 < x_2 < \ldots < x_n$ и $\mathrm{P}(\xi = x_i) = p_i$. Тогда функция распределения $F_\xi$ может быть задана как

$$F_\xi(x) = \begin{cases} 0, & x < x_1; \\ p_1, & x_1 \leqslant x < x_2; \\ p_1 + p_2, & x_2 \leqslant x < x_3; \\ \ldots & \\ p_1 + p_2 + \ldots + p_n = 1, & x \geqslant x_n. \end{cases}$$

График $F_\xi$ будет выглядеть следующим образом.




★

Выделим некоторые свойства этой функции.

**Предложение 3.3.** *Пусть $\xi$ — случайная величина и $F_\xi$ — её функция распределения. Тогда выполнены следующие свойства*

(1) *функция $F_\xi$ неубывает;*

(2) *функция $F_\xi$ непрерывна справа;*

(3) $F_\xi(+\infty) := \lim\limits_{x \to +\infty} F_\xi(x) = 1$;

(4) $F_\xi(-\infty) := \lim\limits_{x \to -\infty} F_\xi(x) = 0$.

*Доказательство.* Для доказательства неубывания (1) воспользуемся свойствами вероятностной меры. Пусть $x < y$. Тогда

$$F_\xi(y) = \mathrm{P}(\xi \leqslant y) = \mathrm{P}(\xi \leqslant x) + \mathrm{P}(x < \xi \leqslant y) \geqslant \mathrm{P}(\xi \leqslant x) = F_\xi(x).$$

Проверим непрерывность справа (2). Пусть последовательность $\{x_n\}$ сходится к $x_0$ справа (можно считать, что последовательность убывает). Тогда $\bigcap\limits_{n=1}^{+\infty}(-\infty; x_n] = (-\infty; x_0]$ и по непрерывности вероятностной меры

$$\lim_{x\to x_0+0} F_\xi(x) = \lim_{n\to+\infty} F_\xi(x_n) = \lim_{n\to+\infty} \mathrm{P}(\xi \leqslant x_n) = \mathrm{P}(\xi \leqslant x_0) = F_\xi(x_0).$$

Свойства (3) и (4) также следуют из непрерывности

$$\lim_{x\to+\infty} F_\xi(x) = \lim_{n\to+\infty} F_\xi(x_n) = \lim_{n\to+\infty} \mathrm{P}(\xi \leqslant x_n) = \mathrm{P}(\xi \in \mathbb{R}) = 1;$$

$$\lim_{x\to-\infty} F_\xi(x) = \lim_{n\to+\infty} F_\xi(x_n) = \lim_{n\to+\infty} \mathrm{P}(\xi \leqslant x_n) = \mathrm{P}(\xi \in \varnothing) = 0.$$

Здесь в первом случае $\bigcup\limits_{n=1}^{+\infty}(-\infty; x_n] = (-\infty; +\infty) = \mathbb{R}$, а во втором $\bigcap\limits_{n=1}^{+\infty}(-\infty; x_n] = (-\infty; -\infty] = \varnothing$. □

Будем обозначать через $F_\xi(x-)$ или $F_\xi(x-0)$ левосторонний предел функции $F_\xi$ в точке $x$. Тогда $\mathrm{P}(\xi = x) = F_\xi(x) - F_\xi(x-)$.

Случайная величина $\xi$ называется непрерывной, если её функция распределения $F_\xi$ непрерывна на $\mathbb{R}$.

**Предложение 3.4.** *Случайная величина $\xi$ непрерывна тогда и только тогда, когда для любого числа $a \in \mathbb{R}$ вероятность того, что $\xi$ примет значение $a$ равна нулю:* $\mathrm{P}(\xi = a) = 0$.

*Доказательство.* Функция $F_\xi$ монотонна и поэтому у неё существуют левые пределы во всех точках. Так как $F_\xi$ ещё и непрерывна справа, то непрерывность в точке $a$ для неё равносильна равенству $F_\xi(a) = F_\xi(a-)$, что в свою очередь равносильно равенству $\mathrm{P}(\xi = a) = F_\xi(a) - F_\xi(a-) = 0$. □

**Теорема 3.5.** *Пусть функция $F$ удовлетворяет свойствам из предложения 3.3. Тогда существует единственная вероятностная мера $\mathrm{P}_F$ на борелевской $\sigma$-алгебре $\mathcal{B}$ подмножеств $\mathbb{R}$ такая, что функция $F$ является функцией распределения случайной величины* $\mathrm{id}\colon (\mathbb{R}, \mathcal{B}, \mathrm{P}_F) \to (\mathbb{R}, \mathcal{B})$.

*Доказательство.* Существование следует из конструкции меры Лебега-Стилтьеса $\mu_F$ для функции $F$. Для неё функцией распределения тождественного отображения автоматически становится функция $F$ (так как буквально теми же формулами определяется эта мера по полуинтервалах).

Пусть теперь P — некоторая вероятностная мера и функция $F$ оказывается функцией распределения для тождественного отображения $\mathrm{id}\colon x \mapsto x$ (далее мы будем также обозначать его через $x$). Тогда имеем

$$\mathrm{P}((a, b]) = \mathrm{P}(a < x \leqslant b) = \mathrm{P}(x \leqslant b) - \mathrm{P}(x \leqslant a) = F(b) - F(a);$$

$$\mathrm{P}((-\infty, b]) = \mathrm{P}(x \leqslant b) - 0 = F(b) - F(-\infty);$$

$$\mathrm{P}((a, +\infty)) = \mathrm{P}(x > a) = 1 - \mathrm{P}(x \leqslant a) = 1 - F(a) = F(+\infty) - F(a).$$

Мы показали, что меры P и $\mu_F$ совпадают на полуинтервалах, следовательно, они совпадают на минимальной алгебре $\mathcal{A}$, порождённой ими. Так как борелевская $\sigma$-алгебра $\mathcal{B}$ является минимальной $\sigma$-алгеброй, содержащей $\mathcal{A}$, то по теореме Каратеодори 0.17 меры P и $\mu_F$ совпадают на ней. $\square$

Для вероятностной меры P на измеримом пространстве $(\mathbb{R}, \mathcal{B})$ функцию распределения случайной величины тождественного отображения id мы будем иногда называть функцией распределения вероятностной меры P.

Следующая теорема не формулируется в основном курсе, но даёт ещё одно описание для меры $\mathrm{P}_F$.

Прямой образ меры $\xi_* \mathrm{P}$ (напомним, что он определяется правилом $\xi_* \mathrm{P}(B) = \mathrm{P}(\xi^{-1}(B)) = \mathrm{P}(\xi \in B)$) на борелевской $\sigma$-алгебре $\mathcal{B}$ подмножеств $\mathbb{R}$ называется распределением (вероятностей) случайной величины $\xi$.

**Теорема 3.6.** *Пусть $\xi$ — случайная величина и $F_\xi$ её функция распределения. Пусть $\mu_\xi$ — мера Лебега-Стилтьеса на подмножествах $\mathbb{R}$, порождённая функцией $F_\xi$. На борелевской $\sigma$-алгебре $\mathcal{B}$ мера $\mu_\xi$ совпадает с распределением $\xi_* \mathrm{P}$ вероятностной меры* P.

*Доказательство.* Достаточно проверить, что эти меры совпадают на минимальном кольце, порождённом всеми полуинтвералами вида $(a, b]$, $(-\infty; b]$ и $(a; +\infty)$. Тогда равенство на борелевской $\sigma$-алгебре будет следовать из теоремы Каратеодори 0.17.

Поскольку мера на минимальном кольце, порождённом данным полукольцом, однозначной определяется по мере на этом полукольце, то достаточно проверить равенства на самих полуинтервалах.

Все равенства немедленно следуют из определений функции распределения, меры Лебега-Стилтьеса и прямого образа меры, а также свойств меры и функции распределения:

$$\mu_f((a, b]) = F_\xi(b) - F_\xi(a) = \mathrm{P}(\xi \leqslant b) - \mathrm{P}(\xi \leqslant a) = \mathrm{P}(a < \xi \leqslant b) = \mathrm{P}(\xi^{-1}((a, b])) = \xi_* \mathrm{P}((a, b]);$$

$$\mu_f((a, +\infty)) = F_\xi(+\infty) - F_\xi(a) = 1 - \mathrm{P}(\xi \leqslant a) = \mathrm{P}(\xi > a) = \mathrm{P}(\xi^{-1}((a, +\infty))) = \xi_* \mathrm{P}((a, +\infty));$$

$$\mu_f((-\infty; b]) = F_\xi(b) - F_\xi(-\infty) = \mathrm{P}(\xi \leqslant b) - 0 = \mathrm{P}(\xi \leqslant b) = \mathrm{P}(\xi^{-1}((-\infty; b])) = \xi_* \mathrm{P}((-\infty; b]).$$

$\square$

В условиях последних двух теорем $\mu_F = \xi_* \mathrm{P}_F$. Если случайная величина $\xi$ имеет распределение $\Gamma$ мы будем писать $\xi \sim \Gamma$.

## 3.3 Разновидности распределений

Распределение дискретной случайной величины $\xi$ называется дискретным. Если случайная величина $\xi$ может принимать только значения $x_1, x_2, \dots$ с вероятностями $p_i = \mathrm{P}(\xi = x_i)$ (при этом мы будем считать, что вероятности $\mathrm{P}(\xi = x_i)$ положительны, чтобы не рассматривать вырожденные случаи), то можно удобно изобразить это с помощью таблицы

| $\xi$ | $x_1$ | $x_2$ | $\dots$ |
|---|---|---|---|
| | $p_1$ | $p_2$ | $\dots$ |

или $\xi \sim \begin{matrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{matrix}$

Поскольку $\sum\limits_{i=1}^{+\infty} p_i = \sum\limits_{i=1}^{+\infty} \mathrm{P}(\xi = x_i) = \mathrm{P}(\Omega) = 1$, то дискретное распределение $\{p_i\}$ можно задать на дискретном вероятностном пространстве.

Распределение случайной величины $\xi$ и она сама называются абсолютно непрерывными, если существует неотрицательная функция $p_\xi \colon \mathbb{R} \to \mathbb{R}$ такая, что для любого борелевского множества $B \in \mathcal{B}$ выполнено равенство

$$\mathrm{P}(\xi \in B) = \xi_* \mathrm{P}(B) = \int\limits_B p_\xi \mathrm{d}\mu,$$

где $\mu$ — классическая мера Лебега. С точки зрения теории меры и интеграла Лебега это определение эквивалентно тому, что мера $\xi_* \mathrm{P}$ абсолютно непрерывна относительно классической меры Лебега $\mu$ на прямой (это значит, что случайная величина $\xi$ попадает в множество меры нуль с вероятностью нуль). Теорема Радона-Никодима даёт формулу для $\xi_* \mathrm{P}$, принятую нами за определение. Функция $p_\xi$ называется функцией плотности
абсолютно непрерывной случайной величины $\xi$.

В силу доказанных выше теорем мы можем дать эквивалентное определение в терминах функции распределения. Будем говорить, что распределение случайной величины $\xi$ и она сама называются абсолютно непрерывными, если функция распределения $F_\xi$ может быть выражена как интеграл

$$F_\xi(x) = \int\limits_{(-\infty;x]} p_\xi \mathrm{d}\mu.$$

В действительном анализе такая функция $F_\xi$ называется абсолютно непрерывной, почти всюду существует производная $F'_\xi$, которая совпадает с $p(x)$ в точках непрерывности последней, а также имеет место аналог формулы Ньютона-Лейбница: $F_\xi(b) - F_\xi(a) = \int\limits_a^b p \mathrm{d}\mu$. Можно доказать, что функция абсолютно непрерывна тогда и только тогда, когда она непрерывна, имеет ограниченную вариацию и переводим множества меры нуль по Лебегу в множества меры нуль.

Во многих рассматриваемых случаях (в частности, во всех примерах следующего раздела) функция $p_\xi$ оказывается интегрируемой по Риману на любом отрезке и абсолютно интегрируема на $\mathbb{R}$, поэтому можно записать равенства

$$F_\xi(x) = \int\limits_{-\infty}^{x} p_\xi(t)\mathrm{d}t, \int\limits_{-\infty}^{+\infty} p_\xi(t)\mathrm{d}t = 1.$$

## 3.4 Преобразования случайных величин

Пусть $g$ — непрерывная строго возрастающая функция, $\xi$ — случайная величина. Мы попытаемся выяснить, как связаны функции распределения (и плотности для абсолютно непрерывных случайных величин) случайных величин $\xi$ и $g(\xi)$.

**Предложение 3.7.** *Имеет место равенство* $F_{g(\xi)}(x) = F_\xi(g^{-1}(x))$.

*Доказательство.* Действительно, $g(\xi) \leqslant x$ равносильно тому, что $\xi \leqslant g^{-1}(x)$ в силу непрерывности и строго монотонности $g$. □

Если заменить $g$ на строго убывающую функцию, то выполнено равенство $F_{g(\xi)(x)} = 1 - F_\xi(g^{-1}(x)-)$.

**Предложение 3.8.** *Пусть $\xi$ — абсолютно непрерывная случайная величина с плотностью $p_\xi$ и $g$ монотонно возрастает и дифференцируема на $\mathbb{R}$. Тогда распределение случайной величины $g(\xi)$ абсолютно непрерывно и* $p_{g(\xi)}(x) = p_\xi(g^{-1}(x))\frac{\mathrm{d}g^{-1}}{\mathrm{d}x}(x)$.

*Доказательство.* По предыдущему предложению 3.7 имеем $F_{g(\xi)}(x) = F_\xi(g^{-1}(x))$. Функция в правой части дифференцируема почти всюду, поэтому композиция $g(\xi)$ абсолютно непрерывна. Продифференцировав, по формуле производной сложной функции получим $p_{g(\xi)} = p_\xi(g^{-1}(x)) \cdot \frac{\mathrm{d}g^{-1}}{\mathrm{d}x}(x)$. □

Если заменить возрастание функции $g$ на неубывание, то мы получим формулу $p_{g(\xi)}(x) = p_\xi(g^{-1}(x))\left|\frac{\mathrm{d}g^{-1}}{\mathrm{d}x}(x)\right|$, которую можно использовать в общем случае для монотонных дифференцируемых функций $g$.

**Пример.** Пусть случайная величина $\eta = a \cdot \xi + b$, и известна плотность $f_\xi(x)$. Тогда

$$\xi = \frac{\eta - b}{a}, \quad \text{то есть} \quad g^{-1}(y) = \frac{y-b}{a},$$

и можно найти плотность $\eta$ по формуле перехода к новой случайной величине:

$$f_\eta(y) = f_\xi\left(\frac{y-b}{a}\right) \cdot \left|\frac{1}{a}\right|.$$

★

**Пример.** Пусть $\eta = g(\xi)$ и $g$ не является взаимно однозначной функцией, а также известна плотность $p_\xi(x)$. К примеру, возьмём $\eta = \xi^2$. Тогда распишем функцию распределения $F_\eta(y)$ по определению:

$$F_\eta(y) = \mathbb{P}\{\eta \le y\} = \mathbb{P}\{\xi^2 \le y\} = \mathbb{P}\{\xi^2 - y \le 0\}$$

$$= \mathbb{P}\{(\xi - \sqrt{y})(\xi + \sqrt{y}) \le 0\} = \mathbb{P}\{-\sqrt{y} \le \xi \le \sqrt{y}\} = F_\xi(\sqrt{y}) - F_\xi(-\sqrt{y}).$$

Тогда при дифференцировании по $y$ получим плотность $\eta$:

$$f_\eta(y) = \frac{d}{dy} F_\eta(y) = \frac{d}{dy}\left[F_\xi(\sqrt{y}) - F_\xi(-\sqrt{y})\right]$$

$$= \frac{1}{2\sqrt{y}} f_\xi(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_\xi(-\sqrt{y}) = \frac{1}{2\sqrt{y}}\left(f_\xi(\sqrt{y}) + f_\xi(-\sqrt{y})\right).$$

★

# 4 Классические примеры распределений

В этом разделе мы рассмотрим классические примеры распределений, повсеместно встречающихся в теории вероятностей и в её приложениях.

Важно помнить, что мы описываем распределение случайной величины, а не непосредственно её саму (случайные величины с одинаковыми распределениями могут быть заданы на разных вероятностных пространствах). Вместо данной случайной величины всегда можно рассмотреть случайную величину id на вероятностном пространстве $(\mathbb{R}, \mathcal{B}, P)$ для некоторого $P$ (или на пространстве $(U, \mathcal{B} \cap U, P)$ для некоторого подмножества $U \subset \mathbb{R}$), то есть распределение некоторой вероятностной меры на $\mathbb{R}$ или подмножестве $\mathbb{R}$. Случайную величину, обладающую распределением Γ также называют Γ-ой (бернуллиевской, геометрической, биномиальной и т.д.) случайной величиной.

НУЖНО: вписать описания для всех классических распределений

## 4.1 Дискретные распределения.

Приведём примеры дискретных распределений. Каждый пример мы будем снабжать описывающей его таблицей.

### 4.1.1 Распределение константы

Пусть случайная величина $\xi$ принимает значение $C$ с вероятность 1: $\mathrm{P}(\xi = C) = 1$. Тогда распределение такой случайной величины называется распределением константы или «вырожденным распределением». Используя табличку можем записать

$$\xi \sim \begin{array}{c} C \\ 1 \end{array}.$$

### 4.1.2 Распределение Бернулли

Пусть случайная величина $\xi$ принимает значения 1 и 0 с вероятностями, соответственно $p$ и $q$ ($p + q = 1$, $p, q \geqslant 0$). Её распределение называется распределением

Бернулли и обозначается $\mathrm{Be}(p)$. И снова в виде таблицы

$$\xi \sim \begin{matrix} 0 & 1 \\ q & p \end{matrix} \sim \mathrm{Be}(p).$$

**Пример.** Рассмотрим дискретное вероятностное пространство $\Omega = \{1, 2, 3, 4, 5, 6\}$ результатов броска игральной кости. Будем считать, что наша игральная кость «иделальная», то есть вероятности всех элементарных событий равны $\frac{1}{6}$. Пусть $\xi\colon \Omega \to \mathbb{R}$ — случайная величина, принимающая значение 1 на числах множестве $\{1, 3\}$ и 0 иначе и выражающая смысл «выпала степень тройки». Тогда $\xi$ обладает биномиальным распределением и

$$\xi \sim \begin{matrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{matrix}.$$

★

### 4.1.3 Дискретное равномерное распределение

Рассмотрим случайную величину $\xi$ принимающую значения из набора $x_1, x_2, \ldots, x_n$, каждое с вероятностью $\frac{1}{n}$. Распределение случайной величины $\xi$ называется дискретным распределением и обозначается $\mathrm{R}\{1, \ldots, n\}$. В виде таблицы:

$$\xi \sim \begin{matrix} x_1 & \ldots & x_n \\ \frac{1}{n} & \ldots & \frac{1}{n} \end{matrix} \sim \mathrm{R}\{1, \ldots, n\}.$$

**Пример.** В качестве примера случайной величины с равномерным распределением можно взять случайную величину $\xi$ на дискретном пространстве $\Omega = \{1, 2, 3, 4, 5, 6\}$ (вероятности всех элементарных событий снова равны $\frac{1}{6}$), сопоставляющую числу из $\Omega$ его же само. ★

### 4.1.4 Биномиальное распределение

Представим, что происходит $n$ последовательных независимых испытаний, в каждом из которых с вероятностью $0 < p < 1$ происходит «успех», а с вероятностью $q = 1 - p$ «неудача». Формально, имеется пространство элементарных исходов $\Omega = \{0, 1\}^n$ с заданным на нём распределением вероятностей, полученное как произведение вероятностных пространств $(\{0, 1\}, 2^{\{0,1\}}, \mathrm{P})$, где $\mathrm{P}(1) = p, \mathrm{P}(0) = q$. Из построения получаем вероятность элементарного исхода $(\varepsilon_1, \ldots, \varepsilon_n)$ равной $p^k q^{n-k}$, где $k$ — это «число успехов», то есть число единиц в кортеже $(\varepsilon_1, \ldots, \varepsilon_n)$.

Рассмотрим случайную величину $\xi$ принимающую на $\Omega$ значения, равные «числу успехов» (сумме единиц в кортеже). Тогда $\xi$ принимает значения от 0 до $n$, причём значение $k$ принимается ей с вероятностью $C_n^k p^k q^{n-k}$. Распределение такой случайной величины $\xi$ называется биномиальным распределением и обозначается $\mathrm{B}(n, p)$. В виде таблицы:

$$\xi \sim \begin{matrix} 0 & 1 & \ldots & k & \ldots & n \\ q^n & npq^{n-1} & \ldots & C_n^k p^k q^{n-k} & \ldots & p^n \end{matrix}.$$

### 4.1.5 Распределение Пуассона

Теперь мы введём счётный аналог биномиального распределение. Воспользуемся разложением экспоненты в ряд

$$e^{\lambda} = \sum_{k=1}^{+\infty} \tfrac{\lambda^k}{k!}.$$

Пусть случайная величина $\xi$ принимает целые неотрицательные значения и принимает значение $k \in \mathbb{Z}_{\geqslant 0}$ с вероятностью $\frac{\lambda^k}{k!}e^{-\lambda}$, $\lambda > 0$ (это условие нужно для того, чтобы все вероятности были положительными). Распределение такой случайной величины называется распределение Пуассона и обозначается $\Pi(\lambda)$. Выписывая это в виде таблицы получаем

$$\xi \sim \begin{matrix} 0 & 1 & 2 & \dots & k & \dots \\ e^{-\lambda} & \lambda e^{-\lambda} & \frac{\lambda^2}{2}e^{-\lambda} & \dots & \frac{\lambda^k}{k!}e^{-\lambda} & \dots \end{matrix} \sim \Pi(\lambda).$$

### 4.1.6 Геометрическое распределение

Рассмотрим эксперимент с двумя исходами «успех» и «неудача», который повторяется до достижения «успеха». Будем считать, что в каждом отдельном испытании «успех» достигается с вероятностью $0 < p < 1$, а «неудача» с вероятностью $q = 1 - p$. Рассмотрим вероятностное пространство $\Omega = \mathbb{N}$ с распределением $\mathrm{P}(n) = pq^{n-1}$ — вероятность того эксперимент закончится на $n$-м этапе (произошли $n-1$ «неудача» и затем «успех»). Пусть случайная величина $\xi$ сопоставляет числу из $\Omega$ его же само. Таким образом, $\xi$ несёт в себя смысл «количество испытаний, проведённых до первого успеха». Распределение такой случайной величины называется геометрическим распределением и обозначается $\mathrm{G}(p)$. Изобразим таблицу:

$$\xi \sim \begin{matrix} 1 & 2 & 3 & \dots & k & \dots \\ p & pq & pq^2 & \dots & pq^{k-1} & \dots \end{matrix} \sim \mathrm{G}(p).$$

**Пример.** Пусть $\Omega = \{0,1\}^{\mathbb{N}}$ — множество всех последовательностей из нулей и единиц. Пусть $\mathfrak{F}$ — $\sigma$-алгебра, полученная лебеговским продолжением с полукольца последовательностей, имеющих фиксированное общее начало (в нём содержатся, например, последовательности, начинающиеся с 1, с 1101 и т.д., также будем включать в полукольцо всё $\Omega$), по мере $m$ заданной по правилу $m(S) = p^k q^m$ для множества $S$ последовательностей с общим началом $\varepsilon_1\varepsilon_2\dots\varepsilon_{k+m}$. Здесь $k$ равно числу единиц в начальном префиксе, а $m$ — числу нулей. Мы опустим проверку того, что это последовательности с общим фиксированным началом образуют полукольцо, а заданная на них функция $m$ является $\sigma$-аддитивной мерой. Теперь пусть случайная величина $\xi$ равна номеру позиции на которой встретилась первая единица последовательности и 1 на последовательности из одних нулей. Тогда $\xi$ обладает биномиальным распределением. ★

### 4.1.7 Гипергеометрическое распределение

Теперь рассмотрим эксперимент с шарами. Пусть в мешке лежат $N$ шаров (уникальных), каждый из которых покрашен в чёрный или в белый цвет. Путь число белых шаров в мешке равно $M$. Вытаскивается $1 \leqslant n \leqslant N$ шаров (считается, что вероятности вытащить все возможные наборы шаров одинаковы). Обозначим через $\xi$ случайную величину, равную числу вытащенных белых шаров и выясним, чему равна вероятность вытащить $m$ белых шаров среди всех $n$ шаров. Формально пространство элементарных исходов $\Omega$ состоит из всевозможных подмножеств мощности $n$ множества всех шаров, а $\xi$ принимает значение равное количеству белых шаров в этом подмножестве. Вероятность будет равна 0, если $m > n$ или $m > M$ или $n - m > N - M$. Разберём остальные случаи. Имеем $|\Omega| = C_N^n$. Число способов выбрать $m$ белых шаров из общего числа $M$ и $n - m$ чёрных шаров из общего числа $N - M$ равно $C_M^m$ и $C_{N-M}^{n-m}$, соответственно. Поскольку вероятность выбрать данные $n$ шаров равна $\frac{1}{C_N^n}$, то $\mathrm{P}(\xi = m) = \frac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n}$ при всех наложенных ограничениях. Распределение такой случайной величины называется гипергеометрическим распределением

### 4.1.8 Отрицательное биномиальное распределение

Кроме подсчёта числа «успехов», как это делается биномиальной случайной величиной, и подсчёта числа попыток, после которых достигается «успех», как в случае с геометрической случайной величиной, можно считать число попыток, после которых достигается $r$ успехов. Снова будем считать, что в каждом отдельном испытании «успех» достигается с вероятностью $p$, а неудача происходит с вероятностью $q$. В качестве вероятностного пространства можно снова взять пространство $\Omega = \mathbb{N}$ с распределением $\mathrm{P}(k) = C_{k-1}^{r-1} p^r q^{k-r}$ для $k \geqslant r$ — вероятность того эксперимент закончится на $k$-м этапе (то есть в этот момент будет достигнут $r$-й «успех», а среди предыдущих $k - 1$ будут как-то разбросаны ещё $r - 1$ «успех»). Пусть случайная величина $\xi$ сопоставляет числу из $\Omega$ его же само. Величина $\xi$ несёт в себя смысл «количество испытаний, проведённых до $r$-го успеха». Распределение такой случайной величины называется геометрическим распределением и обозначается $\mathrm{NB}(r, p)$. Изобразим таблицу:

$$\xi \sim \begin{array}{cccccc} r & r+1 & r+2 & \dots & k & \dots \\ p^r & C_r^{r-1} p^r q & C_{r+1}^{r-1} p^r q^2 & \dots & C_{k-1}^{r-1} p^r q^{k-r} & \dots \end{array} \sim \mathrm{NB}(r, p).$$

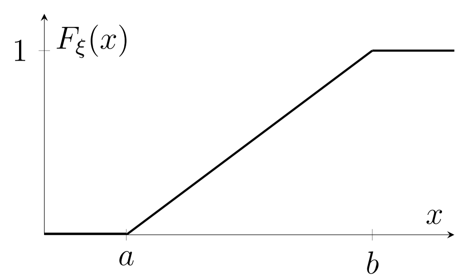
**Пример.** Пусть как и в примере с геометрической случайной величиной $\Omega = \{0, 1\}^{\mathbb{N}}$ — множество всех последовательностей из нулей и единиц. Пусть $\mathfrak{F}$ — $\sigma$-алгебра, полученная лебеговским продолжением с полукольца последовательностей, имеющих фиксированное общее начало (в нём содержатся, например, последовательности, начинающиеся с 1, с 1101 и т.д., также будем включать в полукольцо всё $\Omega$), по мере $m$ заданной по правилу $m(S) = p^k q^m$ для множества $S$ последовательностей с общим началом $\varepsilon_1 \varepsilon_2 \dots \varepsilon_{k+m}$. Здесь $k$ равно числу единиц в начальном префиксе, а $m$ — числу нулей. Мы опустим проверку того, что это последовательности с общим фиксированным началом образуют полукольцо, а заданная на них функция $m$ является $\sigma$-аддитивной мерой. Теперь пусть случайная величина $\xi$ равна номеру позиции на которой встретилась $r$-я единица последовательности и $r$, если в последовательности содержится меньше $r$ единиц. Тогда $\xi$ обаладает отрицательным биномиальным распределением. ★

## 4.2 Абсолютно непрерывные случайные величины

### 4.2.1 Равномерное распределение

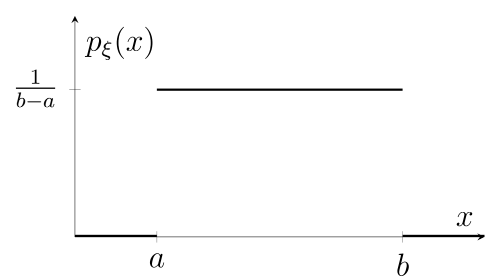
Абсолютно непрерывным аналогом дискретного равномерного распределения выступает равномерное распределение. Пусть $\Omega = [a, b]$ — отрезок, $\mathfrak{F} = \mathcal{B} \cap [a, b]$ — борелевская $\sigma$-алгебра и вероятностная мера задана как $\mathrm{P} = \frac{1}{b-a}\mu$ — нормированная мера Лебега. Рассмотрим распределение случайной величины $\xi$, которая сопоставляет точке на отрезке её координату. Можно думать об этом, как об эксперименте в котором на отрезок «бросается» точка и изучаются её вероятности попасть в разные подмножества отрезка. Распределение случайной величины $\xi$ называется равномерным распределением и обозначается $\mathrm{R}[a, b]$. Функция распределения $F_\xi$ это случайной величины может быть задана кусочно.

$$F_\xi(x) = \begin{cases} 0, & x < a, \\ \dfrac{x-a}{b-a}, & x \in [a, b], \\ 1, & x > b. \end{cases}$$




Данна случайная величина обладает функцией плотности.

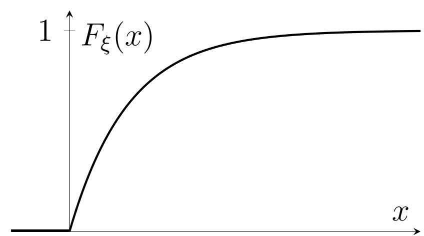
$$p_\xi(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$



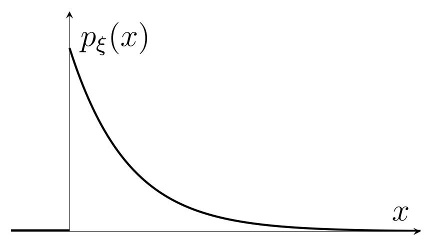

### 4.2.2 Экспоненциальное (показательное) распределение

Как равномерное распределение выступает аналогом дискретного равномерного распределения, так и экспоненциальное распределение является абсолютно непрерывным аналого геометрического распределения. Пусть $\Omega = [0, +\infty)$ — бесконечный полуинтервал, $\mathfrak{F} = \mathcal{B} \cap [0, +\infty)$ — борелевская $\sigma$-алгебра и вероятностная мера задана на бесконечных полуинтервалах как $\mathrm{P}((u;+\infty)) = e^{-\alpha u}$, где $\alpha > 0$ — некоторое число. Рассмотрим распределение случайной величины $\xi$, которая сопоставляет точке на отрезке её координату. В реальной жизни данная конструкция соответствует, например, ситуации, когда измеряется время до первой поломки электроприбора или до первого прихода автобуса, если такие события происходят часто и независимо, а частота их наступления («интенсивность» $\alpha$) постоянна. Распределение случайной величины $\xi$ называется экспоненциальным (или показательным) распределением и обозначается $\mathrm{E}(\alpha)$. Экспоненциальная случайная величина обладает функцией плотности, которую, как и функцию распределения можно задать кусочно. Формулы для функции распределения и плотности приведены ниже.

$$F_\xi(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\alpha x}, & x \geqslant 0. \end{cases}$$




$$p_\xi(x) = \begin{cases} \alpha e^{-\alpha x}, & x \geqslant 0, \\ 0, & x < 0. \end{cases}$$




**Пример.** Будем рассматривать переменную $x$ на прямой как переменную означающую «прошедшее время». Для абсолютно непрерывной случайной величины $\xi$ введём функцию надёжности $\overline{F}_\xi(x) := 1 - F_\xi(x) = \mathrm{P}(\xi > x)$ выражающую вероятность того, что нечто произойдёт (приедет автобус) или не произойдёт (поломка электроприбора) в течении времени $x$. Введём также функцию интенсивности отказов полагаемую равной $\lambda(x) = \lim\limits_{\Delta x \to 0} \frac{\overline{F}_\xi(x+\Delta x) - \overline{F}_\xi(x)}{\overline{F}_\xi(x)\Delta x}$. Интенсивность отказов выражает «долю отказавших в единицу времени приборов». Почти всюду функции $\lambda(x)$ и $-\frac{\overline{F}'_\xi(x)}{\overline{F}_\xi(x)} = \frac{p_\xi(x)}{\overline{F}_\xi(x)}$ совпадают, поэтому будем также писать $\lambda(x) = \frac{p_\xi(x)}{\overline{F}_\xi(x)}$.

Пусть теперь интенсивность отказов постоянна на полуинтервале $[0;+\infty)$ и равна $\alpha > 0$. Из формулы выше получаем дифференциальное уравнение $\overline{F}'_\xi = -\alpha \overline{F}_\xi$, откуда $\overline{F}_\xi(x) = Ce^{-\alpha x}$ на $[0;+\infty)$. Если считать, что $\overline{F}_\xi(0) = 1$ («изначально всё работает»), то $C = 1$ и $F_\xi(x) = 1 - e^{-\alpha x}$.

В общем случае, решив такую же задачу Коши с начальным условием $\overline{F}_\xi(0) = 1$, получаем формулу для функции распределения:

$$F_\xi(x) = 1 - e^{-\int\limits_0^x \lambda(x)\mathrm{d}x}.$$

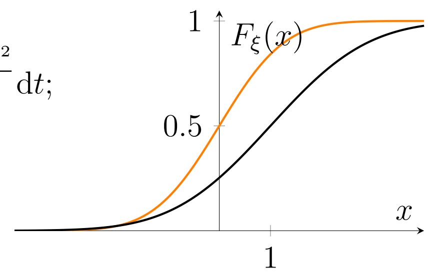
★

### 4.2.3 Нормальное распределение (распределение Гаусса)

Распределение, которое мы рассмотрим сейчас возникает как предел распределения суммы большого количества малых, независимых случайных величин (см. Центральную предельную теорему ниже). Нормальное распределение $\mathcal{N}(a, \sigma^2)$ задаётся двумя параметрами $a$ и $\sigma > 0$, которые, как мы выясним позже равны её математическому ожиданию и квадратному корню из дисперсии, соответственно. Если функция распределения случаной величины $\xi$ имеет вид, указанный ниже, то её распределение называется нормальным распределением. Если к тому же параметры имеют особый вид: $a = 0$ и $\sigma^2 = 1$, то распределение называется стандартным нормальным распределением.

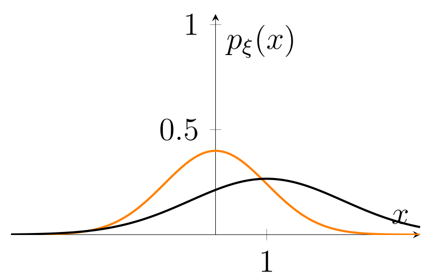
$$\Phi_{a,\sigma^2}(x) := F_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(x-a)^2}{2\sigma^2}}\mathrm{d}t;$$

$$\Phi(x) := F_\xi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{x^2}{2}}\mathrm{d}t.$$




$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}};$$

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$



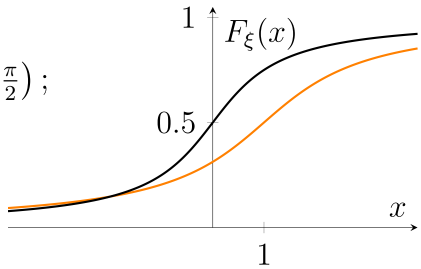

Функция распределения $\Phi_{a,\sigma^2}$ выражается через $\Phi$ по формуле $\Phi_{a,\sigma^2}(x) = \Phi(\frac{x-a}{\sigma})$.

### 4.2.4 Распределение Коши

Ещё одно абсолютно непрерывное распределение, которое мы рассмотрим, похоже на нормальное распределение. Как мы увидим позже, между ними есть существенные различия — распределение Коши является классическим примером распределения, которое не имеет матожидания. Если случайная величина имеет описываемую формулой ниже функцию распределения, то её распределение называется распределением Коши.

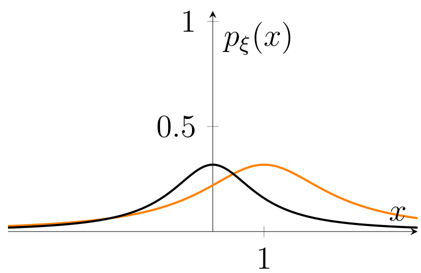
$$K_{a,\sigma^2}(x) := F_\xi(x) = \frac{1}{\pi}\left(\arctan\left(\tfrac{x-a}{\sigma}\right) + \tfrac{\pi}{2}\right);$$

$$K_{0,1}(x) := F_\xi(x) = \frac{1}{\pi}\left(\arctan(x) + \tfrac{\pi}{2}\right).$$




$$p_\xi(x) = \frac{1}{\pi}\frac{1}{1+\left(\frac{x-a}{\sigma}\right)^2};;$$

$$p_\xi(x) = \frac{1}{\pi}\frac{1}{1+x^2}.$$




Аналогично нормальному распределению имеет место формула $K_{a,\sigma^2}(x) = K_{0,1}(\frac{x-a}{\sigma})$.

## 4.3 Сингулярные распределения

Распределение случайной величины $\xi$ называется сингулярным распределение, если её функция распределения $F_\xi$ непрерывна и множество её точек роста

$$\{x \in \mathbb{R} |\, \forall \varepsilon > 0 \ F_\xi(x-\varepsilon) < F_\xi(x+\varepsilon)\}$$

имеет меру нуль по Лебегу.

**Пример.** Стандартным примером сингулярного распределения служит распределение, соответствующее канторовой лестнице $c\colon [0,1] \to [0,1]$. Рассмотрим канторово множество

$$K = \bigcap_{i=1}^{+\infty}\bigcup_{j=0}^{2^i-1}\left[\tfrac{k_j}{3^i}, \tfrac{k_j+1}{3^i}\right] = \bigcap_{i=1}^{+\infty}\bigcup_{j=0}^{3^{i-1}-1}\left[\tfrac{3j}{3^i}, \tfrac{1+3j}{3^i}\right]\cup\left[\tfrac{2+3j}{3^i}, \tfrac{3+3j}{3^i}\right] = [0,1]\setminus\bigcup_{i=1}^{+\infty}\bigcup_{j=0}^{3^{i-1}-1}\left(\tfrac{1+3j}{3^i}, \tfrac{2+3j}{3^i}\right),$$

где $k_j$ — $j$-е число от 0 до $3^i$, в троичной записи которого присутствуют только нули и двойки. Всякая точка канторова множеств в троичной системе счисления имеет вид бесконечной троичной дроби с нулями и двойками. Тогда канторова лестница задаётся следующими формулами

$$c(x) = \begin{cases} \frac{j}{2^i}, & x \in \left(\frac{1+3j}{3^i}, \frac{2+3j}{3^i}\right); \\ \sum\limits_{i=1}^{+\infty} \frac{a_j/2}{2^i}, & x \in K, x = \sum\limits_{i=1}^{+\infty}\frac{a_j}{3^j}. \end{cases}$$

Можно показать, что заданная такой формулой функция непрерывна. Множество точек роста $c$ совпадает канторовым множеством, и поэтому имеет меру нуль по Лебегу.

В нашем примере оказалось так, что канторова лестница $c$ имеет производную почти всюду, как и абсолютно непрерывные функции. Однако, в отличии от абсолютно непрерывных функций эта производна равна 0 во всех точках своего определения и поэтому канторову лестницу нельзя вычислить как интеграл от своей производной. ★

# 5 Совместные распределения случайных величин

Пусть $\xi_1, \ldots, \xi_n$ — случайные величины, заданные на вероятностном пространстве $(\Omega, \mathfrak{F}, \mathrm{P})$. Тогда функция $\xi\colon \Omega \to \mathbb{R}^n$, заданная по правилу $\xi(\omega) = (\xi_1(\omega), \ldots, \xi_n(\omega))$ называется случайным вектором или многомерной случайной величиной.

**Предложение 5.1.** *Справедливы следующие утверждение о случайных величинах и многомерных случайных величинах.*

(1) *Многомерная случайная величина является измеримым отображением $\xi\colon (\Omega, \mathfrak{F}) \to (\mathbb{R}^n, \mathcal{B}_n)$, где $\mathcal{B}_n$ — борелесвкая $\sigma$-алгебра подмножеств $\mathbb{R}$.*

(2) *Если $\xi\colon (\Omega, \mathfrak{F}) \to (\mathbb{R}^n, \mathcal{B}_n)$ — измеримое отображение, то каждая композиция $\xi_i = \mathrm{pr}_i \circ \xi$ $(\xi(\omega) = (\xi_1(\omega), \ldots, \xi_n(\omega)))$ является случайной величиной и тем самым $\xi$ является многомерной случайной величиной.*

*Доказательство.* Оба утверждения являются следствиями теоремы 0.42. □

Аналогично одномерным случайным величинам для случайного вектора определяются его распределение как прямой образ вероятностной меры $\xi_* \mathrm{P}$. Снова напомним, что $\xi_* \mathrm{P}(B) := \mathrm{P}(\xi^{-1}(B)) = \mathrm{P}(\xi \in B)$. Аналогично же определяется функция распределения случайного вектора (функция совместного распределения)

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = \mathrm{P}(\xi_1 \leqslant x_1, \ldots, \xi_n \leqslant x_n).$$

Функция совместного распределения удовлетворяет свойствам, аналогичным свойствам одномерной функции распределения.

**Предложение 5.2.** *Пусть $\xi_1, \ldots, \xi_n$ — случайные величины и $F_{\xi_1,\ldots,\xi_n}$ — её функция распределения. Тогда выполнены следующие свойства*

(1) *функция $F_{\xi_1,\ldots,\xi_n}$ неубывает по каждому аргументу;*

(2) *функция $F_{\xi_1,\ldots,\xi_n}$ непрерывна справа по каждому аргументу;*

(3) $\lim\limits_{x_n \to +\infty} F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = F_{\xi_1,\ldots,\xi_{n-1}}(x_1, \ldots, x_{n-1})$;

(4) $\lim\limits_{x_n\to-\infty} F_{\xi_1,\dots,\xi_n}(x_1,\dots,x_n)=0;$

(5) $\lim\limits_{x_1\to+\infty\dots x_n\to+\infty} F_{\xi_1,\dots,\xi_n}(x_1,\dots,x_n)=1;$

По функции $F$, удовлетворяющей свойствам из предложения 5.2 модно построить многомерную меру Лебега-Стилтьеса $\mu_F$. Остаются верны следующие результаты.

**Теорема 5.3.** *Пусть функция $F$ удовлетворяет свойствам из предложения 5.2. Тогда существует единственная вероятностная мера $\mathrm{P}_F$ на борелевской $\sigma$-алгебре $\mathcal{B}_n$ подмножеств $\mathbb{R}^n$ такая, что функция $F$ является функцией совместного распределения многомерной случайной величины* $\mathrm{id}\colon(\mathbb{R}^n,\mathcal{B}_n,\mathrm{P}_F)\to(\mathbb{R}^n,\mathcal{B}_n)$.

**Теорема 5.4.** *Пусть $\xi$ — случайный вектор и $F_\xi$ его функция распределения. Пусть $\mu_\xi$ — мера Лебега-Стилтьеса на подмножествах $\mathbb{R}^n$, порождённая функцией $F_\xi$. На борелевской $\sigma$-алгебре $\mathcal{B}_n$ мера $\mu_\xi$ совпадает с распределением $\xi_*\mathrm{P}$ вероятностной меры $\mathrm{P}$.*

Многомерная случайная величина называется дискретной, если она принимает не более, чем счётное число значений. Распределение такой случайной величины называется дискретным.

Распределение случайного вектора $\xi=(\xi_1,\dots,\xi_n)$ и он сам называются абсолютно непрерывными, если существует функция $p\colon\mathbb{R}^n\to\mathbb{R}$ такая, что для любого борелевского множества $B\in\mathcal{B}_n$ выполнено равенство

$$\mathrm{P}(\xi\in B)=\xi_*\mathrm{P}(B)=\int\limits_B f\mathrm{d}\mu^{(n)},$$

где $\mu^{(n)}$ — классическая мера Лебега на подмножествах $\mathbb{R}^n$. С точки зрения теории меры и интеграла Лебега это определение эквивалентно тому, что мера $\xi_*\mathrm{P}$ абсолютно непрерывна относительно классической меры Лебега $\mu^{(n)}$ в многомерном пространстве. (это значит, что случайная величина $\xi$ попадает в множество меры нуль с вероятностью нуль). Теорема Радона-Никодима даёт формулу для $\xi_*\mathrm{P}$, принятую нами за определение. Функция $p_\xi=p_{\xi_1,\dots,\xi_n}$ называется функцией совместной плотности абсолютно непрерывного случайного вектора $\xi=(\xi_1,\dots,\xi_n)$.

В силу доказанных выше теорем мы можем дать эквивалентное определение в терминах функции совместного распределения. Будем говорить, что совместное распределение случайных величин $\xi_1,\dots,\xi_n$ называется абсолютно непрерывными, если функция совместного распределения $F_{\xi_1,\dots,\xi_n}$ может быть выражена как интеграл

$$F_{\xi_1,\dots,\xi_n}(x_1,\dots,x_n)=\int\limits_{\prod\limits_{i=1}^n(-\infty;x_i]} p_{\xi_1,\dots,\xi_n}\mathrm{d}\mu^{(n)}.$$

Почти всюду существует частная производная $\frac{\partial^n F_{\xi_1,\dots,\xi_n}}{\partial x_1 \dots \partial x_n}(\tilde{x_1},\dots,\tilde{x_n})$, которая почти всюду совпадает с $p_{\xi_1,\dots,\xi_n}(\tilde{x_1},\dots,\tilde{x_n})$.

Во многих рассматриваемых случаях (в частности, во всех примерах выше) функция $p_\xi$ оказывается интегрируемой по Риману на любом отрезке и абсолютно интегрируема на $\mathbb{R}$, поэтому можно записать равенства

$$F_{\xi_1,\dots,\xi_n}(x_1,\dots,x_n) = \int\limits_{-\infty}^{x_n} \dots \int\limits_{-\infty}^{x_1} p_{\xi_1,\dots,\xi_n}(t_1,\dots,t_n)\mathrm{d}t_1\dots\mathrm{d}t_n,$$

$$\int\limits_{-\infty}^{+\infty} \dots \int\limits_{-\infty}^{+\infty} p_{\xi_1,\dots,\xi_n}(t_1,\dots,t_n)\mathrm{d}t_1\dots\mathrm{d}t_n = 1.$$

Также верна формула

$$p_{\xi_1,\dots,\xi_{n-1}}(t_1,\dots,t_{n-1}) = \int\limits_{-\infty}^{+\infty} p_{\xi_1,\dots,\xi_n}(t_1,\dots,t_n)\mathrm{d}t_n.$$

**Пример.** Пусть $a = (a_1,\dots,a_n)^T$ — вектор, $B = (\sigma_{ij})$ — положительно определённая симметрическая матрица. Будем говорить, что вектор $\xi = (\xi_1,\dots,\xi_n)$ имеет многомерное нормальное распределение, если это распределение имеет совместную плотность

$$p_\xi(x) = \frac{1}{(2\pi)^{n/2}\sqrt{\det B}} e^{-\frac{(x-\alpha)^T B^{-1}(x-\alpha)}{2}}.$$

Используя замену координат и приведение матрицы $B$ к единичному виду, можно показать, что

$$\int\limits_{-\infty}^{+\infty} \dots \int\limits_{-\infty}^{+\infty} p_{\xi_1,\dots,\xi_n}(t_1,\dots,t_n)\mathrm{d}t_1\dots\mathrm{d}t_n = 1.$$

★

# 6 Независимость случайных величин

## 6.1 Независимость конечного набора случайных величин

Введём понятие независимости случайных величин, с которым мы будем в дальнейшем сталкиваться повсеместно. Интуиция говорит там, что случайные величины независимы, если вероятности их принять какое-то значения не связаны или другими словами независимы события вида $\xi_k \in B_k$, где $B_k$ — борелевские множества. Случайные величины $\xi_1,\dots,\xi_n$ называются независимыми, если для любого набора борелевских множеств $B_1,\dots,B_n$ выполнено равенство

$$\mathrm{P}(\xi_1 \in B_1,\dots,\xi_n \in B_n) = \mathrm{P}(\xi_1 \in B_1)\dots\mathrm{P}(\xi_n \in B_n).$$

По всякой случайной величине можно построить $\sigma$-алгебру событий $\mathfrak{F}_\xi$, состоящую из всевозможных прообразов борелевских множеств при $\xi$: $\mathfrak{F}_\xi = \{\xi^{-1}(B) | \, B \in \mathcal{B}\}$. Построенную $\sigma$-алгебры мы будем называть $\sigma$-алгеброй, порождённой случайной величиной $\xi$. Отметим, что данная $\sigma$ алгебра есть не что иное как, обратный образ борелевской $\sigma$-алгебры при отображении $\xi$. Проверка того, что это $\mathfrak{F}_\xi$ действительно является $\sigma$-алгеброй осуществляется непосредственной проверкой всех аксиом. Кроме обозначения $\mathfrak{F}_\xi$, будем также писать $\sigma(\xi)$.

**Предложение 6.1.** *Пусть $\xi, \eta$ — две случайные величины и $\mathfrak{F}_\xi \subset \mathfrak{F}_\eta$ (в таких случаях ещё говорят, что $\xi$ измерима относительно $\mathfrak{F}_\eta$, поскольку включение означает, что прообразы борелевских множеств при отображении $\xi$ лежат в $\mathfrak{F}_\eta$). Тогда существует такая функция борелевская функция $g$, что $\xi = g(\eta)$.*

*Доказательство.* Положим $A_{k,n} = \xi^{-1}\left(\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right)$. Так как $A_{k,n} \in \mathfrak{F}_\xi \subset \mathfrak{F}_\eta$, то найдутся борелевские множества $B_{k,n}$ такие, что $\eta^{-1}(B_{k,n}) = A_{k,n}$. Положим $g_n(x) = \frac{k}{2^n}$ на $B_{k,n} \setminus \bigcup\limits_{k' \neq k, n' \neq n} B_{k',n'}$ и нулём там, где хотя бы два из множеств $B_{k,n}$ пересекаются (все такие точки лежат вне образа случайной величины $\eta$, так как множества $A_{k,n}$ попарно не пересекаются). Тогда $g_n$ является борелевской и $g_n(\eta(\omega)) = \frac{[2^n \xi(\omega)]}{2^n}$. Кроме того, функции $g_n$ образуют поточечно возрастающую последовательность, поточечно ограниченную $\xi$ или нулём и такую, что композиции $g_n \circ \eta$ стремятся к $\xi$. Следовательно, во всех точках существует предел $g(x) = \lim\limits_{n \to +\infty} g_n(x)$, являющийся борелевской функцией и выполнено равенство $g(\eta) = \xi$. □

Две ($\sigma$-)алгебры $\mathfrak{F}_1, \ldots, \mathfrak{F}_n \subset \mathfrak{F}$ будем называть независимыми, если события из любого набора $A_1 \in \mathfrak{F}_1, \ldots, A_n \in \mathfrak{F}_n$ независимы. Поскольку всегда в качестве некоторых $A_i$ можно взять $\Omega$, то независимость $\sigma$-алгебр $\mathfrak{F}_1, \ldots, \mathfrak{F}_n \subset \mathfrak{F}$ равносильна истинности для всех наборов $A_1 \in \mathfrak{F}_1, \ldots, A_n \in \mathfrak{F}_n$ равенства

$$\mathrm{P}(A_1 \ldots A_n) = \mathrm{P}(A_1) \ldots \mathrm{P}(A_n).$$

Ввиду данного определения возникает описание независимости случайных величин в терминах порождённых ими $\sigma$-алгебр.

**Предложение 6.2.** *Случайные величины независимы тогда и только тогда, когда порождённые ими $\sigma$-алгебры независимы.*

*Доказательство.* Взяв в качестве $A_i = \xi^{-1}(B_i)$, где $B_i$ борелевское, дословно перепишем друг в друга равенства из обоих определений. □

**Следствие 6.3.** *Пусть $g, h$ — борелевские функции, а $\xi, \eta$ — независимые случайные величины. Тогда случайные величины $g(\xi)$ и $h(\eta)$ независимы.*

*Доказательство.* Имеем включения $\sigma$-алгебр $\mathfrak{F}_{g(\xi)} \subset \mathfrak{F}_\xi$ и $\mathfrak{F}_{h(\eta)} \subset \mathfrak{F}_\eta$. По предложению 6.2 большие $\sigma$-алгебры $\mathfrak{F}_\xi$ и $\mathfrak{F}_\eta$ независимы. Тогда из определения независимости алгебр следует, что содержащиеся в них $\sigma$-алгебры $\mathfrak{F}_{g(\xi)}$ и $\mathfrak{F}_{h(\eta)}$ тоже независимы. Вновь по предложению 6.2 случайные величины $g(\xi)$ и $h(\eta)$ независимы. □

**Теорема 6.4** (Об аппроксимации)**.** *Пусть* $(\Omega, \mathfrak{F}, \mathrm{P})$ *— вероятностное пространство и* $\mathfrak{A}$ *—* $\sigma$*-алгебра, порождённая алгеброй* $\mathcal{A} \subset \mathfrak{F}$*. Тогда для всякого элемента* $U \in \mathfrak{A}$ *существует последовательность элементов* $\{U_i\} \subset \mathcal{A}$ *такая, что* $\lim\limits_{i\to+\infty} \mathrm{P}(U_i \vartriangle U) = 0$*, и, следовательно,* $\lim\limits_{i\to+\infty} \mathrm{P}(U_i) = \mathrm{P}(U)$*.*

*Доказательство.* Построим по алгебра $\mathcal{A}$ и ограничении на неё меры P их лебеговское продолжение, как это делалось в предварительных сведениях. Тогда, новая $\sigma$-алгебра измеримых множеств будет содержать в себе алгебру $\mathfrak{A}$. По теореме Каратеодори 0.17 меры P и продолжение по Лебегу этой же меры с $\mathcal{A}$ совпадают на $\mathfrak{A}$. Из измеримости множества $U$ и конструкции легебогвского продолжения следует существование множеств $U_i \in \mathcal{A}$, удовлетворяющих условию. Второе выражение с пределом следует из неравенства $|\,\mathrm{P}(U) - \mathrm{P}(U_i)| \leqslant \mathrm{P}(U \vartriangle U_i)$. □

**Теорема 6.5.** *Пусть* $\mathfrak{F}_1, \ldots, \mathfrak{F}_n \subset \mathfrak{F}$ *—* $\sigma$ *алгебры и* $\sigma$*-алгебра* $\sigma_i$ *порождается алгеброй* $\mathcal{A}_i$*. Тогда независимость данных* $\sigma$*-алгебр равносильна независимости всех возможных наборов* $A_i \in \mathcal{A}_i$ *(независимости алгебр* $A_i$*).*

*Доказательство.* Независимость наборов из условия следует из определения независимости $\sigma$-алгебр.

Пусть теперь все алгебры $\mathcal{A}_i$ из условия независимы. Пусть $U_i \in \mathfrak{F}_i$ — набор событий. По лемме об аппроксимации 6.4 существуют последовательности $U_{ij} \in \mathcal{A}_i$ такие, что $\lim\limits_{j\to+\infty} \mathrm{P}(U_{ij}) = \mathrm{P}(U_i)$. Так как

$$\mathrm{P}((U_1 \ldots U_n) \vartriangle (U_{1j} \ldots U_{nj})) \leqslant \sum_{i=1}^{n} \mathrm{P}(U_i \vartriangle U_{ij}),$$

то существует и предел $\lim\limits_{j\to+\infty} \mathrm{P}(U_{1j} \ldots U_{nj}) = \mathrm{P}(U_1 \ldots U_n)$. Тогда

$$\mathrm{P}(U_1) \ldots \mathrm{P}(U_n) = \lim_{j\to+\infty} \mathrm{P}(U_{1j}) \ldots \mathrm{P}(U_{nj}) = \lim_{j\to+\infty} \mathrm{P}(U_{1j} \ldots U_{nj}) = \mathrm{P}(U_1 \ldots U_n).$$

□

Следующая теорема демонстрирует связь между независимостью, функциями распределения и совместным распределением.

**Теорема 6.6.** *Случайные величины* $\xi_1, \ldots, \xi_n$ *независимы тогда и только тогда, когда для любого набора* $(x_1, \ldots, x_n)$ *имеет место равенство*

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n).$$

*Доказательство.* Если случайные величины независимы, то взяв в качестве $B_i = (-\infty; x_i]$ получим требуемое равенство.

Предположим, что выполнено равенство из условия. В силу предложения 6.2 достаточно доказать, что независимы $\sigma$-алгебры $\mathfrak{F}_{\xi_1}, \ldots, \mathfrak{F}_{\xi_n}$. Обозначим через

$B_i(x)$ множество $\xi_i^{-1}((-\infty;x])$ и через $C_i(a,b]$ множество $\xi_i^{-1}((a,b])$. По условию, для любых $x_1,\ldots,x_n$ события $B_1(x_1),\ldots,B_n(x_n)$ независимы. Так как $C_1(a,b] = B_1(b)\setminus B_1(a)$ и выполнены равенства

$$\mathrm{P}(C_i(a,b]B_2(x_2)\ldots B_n(x_n)) = \mathrm{P}(B_1(b)B_2(x_2)\ldots B_n(x_n)) - \mathrm{P}(B_1(a)B_2(x_2)\ldots B_n(x_n)) =$$

$$= \mathrm{P}(B_1(b))\,\mathrm{P}(B_2(x_2))\ldots\mathrm{P}(B_n(x_n)) - \mathrm{P}(B_1(a))\,\mathrm{P}(B_2(x_2))\ldots\mathrm{P}(B_n(x_n)) =$$
$$= \mathrm{P}(C_i(a,b])\,\mathrm{P}(B_2(x_2))\ldots\mathrm{P}(B_n(x_n)),$$

по индукции, а также аналогично для интервалов $(a;+\infty)$ можем вывести, что для любых $a_1,\ldots,a_n,b_1,\ldots,b_n$ независимы события $A_1,\ldots,A_n$, где $A_i\in S_i$ — принадлежат полукольцу полуинвервалов. Из построения минимального кольца 0.7 и свойств независимости событий 2.7 вытекает, что алгебры, порождённые прообразами всех полуинтервалов независимы. Так как минимальная сигма-алгебра, порождённая полуинтвервалами совпадает с борелевской $\sigma$-алгеброй, то по лемме 0.41 их прообразы порождают $\sigma$-алгебры $\mathfrak{F}_{\xi_i}$. По теореме 6.5 эти $\sigma$-алгебры независимы. □

**Следствие 6.7.** *Пусть абсолютно непрерывные случайные величины $\xi_1,\ldots,\xi_n$ имеют плотности $p_{\xi_1},\ldots,p_{\xi_n}$. Тогда их независимость равносильна тому, что их совместное распределение абсолютно непрерывно с плотностью*

$$p_{\xi_1,\ldots,\xi_n}(x_1,\ldots,x_n) = p_{\xi_1}(x_1)\ldots p_{\xi_n}(x_n).$$

*Доказательство.* В доказательстве мы будем писать интегралы в нотации интеграла Римана, подразумевая, что они вычисляются любым допустимым способом (например, по Лебегу).

По теореме 6.6 из независимости имеем равенство

$$F_{\xi_1,\ldots,\xi_n}(x_1,\ldots,x_n) = F_{\xi_1}(x_1)\cdot\ldots\cdot F_{\xi_n}(x_n).$$

Тогда

$$F_{\xi_1,\ldots,\xi_n}(x_1,\ldots,x_n) = \int\limits_{-\infty}^{x_1} p_{\xi_1}(t_1)\mathrm{d}t_1\ldots\int\limits_{-\infty}^{x_n} p_{\xi_n}(t_n)\mathrm{d}t_n =$$

$$= \int\limits_{-\infty}^{x_1}\ldots\int\limits_{-\infty}^{x_n} p_{\xi_1}(t_1)\ldots p_{\xi_n}(t_n)\mathrm{d}t_n\ldots\mathrm{d}t_1.$$

Обратно, если верна формула выше, то проведя выкладки в обратном порядке и воспользовавшись теоремой 6.6 в другу сторону, получим независимость. □

## 6.2 Последовательность независимых случайных величин

Последовательность случайных величин $\{\xi_i\}$ будем называть последовательностью независимых случайных величин, если для всякого $n \in \mathbb{N}$ случайные величины $\xi_1, \ldots, \xi_n$ независимы.

Назовём $\sigma$-алгеброй, порождённой случайными величинами $\xi_k, \ldots, \xi_n$ $(n \leqslant +\infty)$, $\sigma$-алгебру, порождённую событиями вида $\bigcap\limits_{k \leqslant m \leqslant n} A_m$, где $A_m \in \sigma(\xi_m)$. Будем обозначать эту $\sigma$-алгебру через $\sigma(\xi_k, \ldots, \xi_n)$ для конечных $n$ и через $\sigma(\xi_k, \xi_{k+1}, \ldots)$ для $n = +\infty$.

**Предложение 6.8.** *Справедливо равенство*

$$\sigma(\xi_k, \xi_{k+1}, \ldots) = R_\sigma(\bigcup_{n=k}^{+\infty} \sigma(\xi_k, \ldots, \xi_n)).$$

*Доказательство.* Включения $\sigma(\xi_k, \ldots, \xi_n) \subset \sigma(\xi_k, \ldots, \xi_{n+1}), \sigma(\xi_k, \xi_{k+1}, \ldots)$ следуют из построения. Таким образом, объединение в правой части является алгеброй, и имеется включение

$$\sigma(\xi_k, \xi_{k+1}, \ldots) \supset R_\sigma(\bigcup_{n=k}^{+\infty} \sigma(\xi_k, \ldots, \xi_n)).$$

С другой стороны, так как пересечения $\bigcap\limits_{k \leqslant m} A_m$ все содержатся в правой части, то выполнено равенство. $\square$

Хвостовой $\sigma$-алгеброй для данной последовательности независимых случайных величины будем называть пересечение $\mathfrak{X} = \bigcap\limits_{k=1}^{+\infty} \sigma(\xi_k, \xi_{k+1}, \ldots)$. События $A \in \mathcal{X}$ называются хвостовыми. Позже мы докажем, что вероятностная мера всякого хвостового события равна либо 0, либо 1 (см. закон нуля или единицы Колмогорова 12.3).

**Предложение 6.9.** *Для любого $k \geqslant 1$ $\sigma$-алгебры $\sigma(\xi_{n+k})$ и $\sigma(\xi_1, \ldots, \xi_n)$ независимы.*

*Доказательство.* Из независимости $\sigma$-алгебр $\sigma(\xi_1), \ldots, \sigma(\xi_n), \sigma(\xi_{n+k})$ следует независимость $\sigma$-кольца, порождённого всеми возможными произведениями $\bigcap\limits_{k \leqslant m \leqslant n} A_m$, где $A_m \in \sigma(\xi_m)$, $1 \leqslant m \leqslant n$ и $\sigma$-алгебры $\sigma(\xi_{n+k})$. Тогда по 6.5 утверждение верно и для минимальны $\sigma$-алгебр $\sigma(\xi_1, \ldots, \xi_n)$ и $\sigma(\xi_{n+k})$. $\square$

# 7 Численные характеристики случайных величин

В этом разделе мы введём различные величины, характеризующие случайные величины, такие как математические ожидание, выражающее «усреднённое значение» случайной величины, дисперсию, показывающую, насколько случайная величина отклоняется от своего «среднего» и другие.

## 7.1 Математическое ожидание

Пусть $\xi$ — случайная величина. Если существует интеграл Лебега $\int\limits_{\Omega} \xi \, \mathrm{d}\,\mathrm{P}$, то говорят, что $\xi$ обладает математическим ожиданием, величина интеграла называется математическим ожиданием случайной величины $\xi$ и обозначается $\mathrm{E}\,\xi$. Интегрируемость случайной величины $\xi$ также будем записывать формулой $\xi \in \mathcal{L}^1(\Omega)$ или $\xi \in \mathcal{L}^1$.

Для дискретной случайной величины $\xi$, принимающей значения $x_1, x_2, \ldots$ с вероятностями $p_1, p_2, \ldots$, было бы естественно определить её математические ожидание как сумму ряда $\sum\limits_{i=1}^{+\infty} x_i p_i$, а его существование как его абсолютную сходимость. И следующее предложение позволяет нам сделать это

**Предложение 7.1.** *Пусть $\xi$ — дискретная случайная величина, принимающая значения $x_1, x_2, \ldots$ с вероятностями $p_1, p_2, \ldots$, соответственно. Тогда математическое ожидание существует тогда и только тогда, когда абсолютно сходится ряд $\sum\limits_{i=1}^{+\infty} x_1 p_1$, и в случае сходимости равно его сумме.*

*Доказательство.* Дискретная случайная величина является обобщённой простой функцией на $\Omega$, поэтому утверждение следует из предложения 0.28. □

**Пример.** Есть $N$ лотерейных билетов, причём на $m_1, \ldots, m_n$ из них приходятся соответственно выигрыши $a_1, \ldots, a_n$. Разыгрывается денежная сумма

$$S = a_1 m_1 + a_2 m_2 + \ldots + a_n m_n.$$

Тогда средний выигрыш, приходящийся на один билет, равен

$$\frac{S}{N} = \frac{a_1 m_1 + \ldots + a_n m_n}{N} = a_1 \frac{m_1}{N} + \ldots + a_n \frac{m_n}{N} = a_1 p_1 + \ldots + a_n p_n,$$

где $p_i = \frac{m_i}{N}$.

Введём в данном примере вероятностное пространство $\Omega$, состоящее из $N$ точек, и распределение вероятностей, сопоставляющие точкам из непересекающихся подмножеств по $m_i$ точек вероятности $p_i$. Тогда, если случайная величина $\xi$ сопоставляет точкам из $i$-го подмножества число («выигрыш») $a_i$, то $\mathrm{E}\,\xi$ существует и равно $\frac{S}{N}$. ★

**Пример.** Пусть на прямой в точках $x_1, \ldots, x_n$ сосредоточены массы $p_1, \ldots, p_n$. Тогда центр тяжести этой системы равен

$$\bar{x} = \frac{x_1 p_1 + x_2 p_2 + \ldots + x_n p_n}{p_1 + p_2 + \ldots + p_n}.$$

Формально можно рассмотреть меру введённую на числовой прямой и сосредоточенную в точках с координатами $x_1, \ldots, x_n$, а также случайную величину, сопоставляющую точке её координату. Тогда матожидание оказывается равным в точности $\bar{x}$. ★

**Пример.** Если $\xi = c$ — константа, то $\mathbb{E}\xi = c$. Если $\xi = \mathbb{1}_A$, то $\mathbb{E}\xi = \mathrm{P}(A)$.

Таким образом, для бернуллиевской случайной величины $\xi \sim \mathrm{Be}(p)$ с параметром $p$ выполнено равенство:

$$\mathbb{E}\xi = 0 \cdot q + 1 \cdot p = p.$$

★

**Пример.** Пусть случайная величина $\Pi$ распределена по закону Пуассона с параметром $\lambda > 0$:

$$\mathrm{E}\,\Pi = \sum_{k=0}^{\infty} k \cdot \mathrm{P}(\Pi = k) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!}.$$

Так как $k = 0$ даёт вклад 0, начинаем сумму с $k = 1$:

$$\mathrm{E}\,\Pi = \sum_{k=1}^{\infty} \frac{k\lambda^k e^{-\lambda}}{k!}.$$

Представим $k = j + 1$, тогда $j = k - 1$, и:

$$\mathrm{E}\,\Pi = \sum_{j=0}^{\infty} \frac{(j+1)\lambda^{j+1} e^{-\lambda}}{(j+1)!} = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

★

**Предложение 7.2.** *Случайная величина $\xi$ обладает математическим ожиданием тогда и только тогда, когда $|\xi|$ обладает математическим ожиданием.*

*Доказательство.* Пусть $\{\xi_n\}$ — последовательность интегрируемых обобщённых простых функций, равномерно сходящаяся к $\xi$. Тогда функции $|\xi_n|$ интегрируемы и равномерно сходятся к $|\xi|$ из чего следует, что $|\xi|$ интегрируема.

Обратно, пусть $\{\xi_n\}$ — произвольная последовательность, равномерно сходящаяся к $\xi$. Снова $\{|\xi_n|\}$ равномерно сходится к $|\xi|$. Поскольку $|\xi|$ интегрируема, то начиная с некоторого члена функции $|\xi_n|$ интегрируемы по 0.31. Отюсюда следует интегрируемость $\xi$. □

**Предложение 7.3.** *Имеют место следующие свойства.*

(1) $\mathrm{E}\,c = c$ *для любой константы* $c$;

(2) $\mathrm{E}\,\mathbb{1}_A = \mathrm{P}(A)$ *для любого события* $A$;

(3) *если* $\eta, \xi$ *— неотрицательные случайные величины, для всякого* $\omega \in \Omega$ *имеем место* $\eta(\omega) < \xi(\omega)$ $(=: \eta \leqslant \xi)$ *и* $\xi \in \mathcal{L}^1$, *то* $\eta \in \mathcal{L}^1$;

(4) *если* $\xi, \eta \in \mathcal{L}^1$, *то* $a\xi, \xi + \eta \in \mathcal{L}^1$ *для любой константы* $a \in \mathbb{R}$;

(5) *если* $\xi, \eta \in \mathcal{L}^1$, *то* $\mathrm{E}(a\xi) = a\,\mathrm{E}\,\xi$ *и* $\mathrm{E}(\xi + \eta) = \mathrm{E}\,\xi + \mathrm{E}\,\eta$ *для любой константы* $a \in \mathbb{R}$;

(6) *если* $\xi, \eta \in \mathcal{L}^1$ *и* $\xi \leqslant \eta$, *то* $\mathrm{E}\,\xi \leqslant \mathrm{E}\,\eta$.

*Доказательство.* Свойства (1) и (2) следуют из предложения о дискретных случайных величинах.

Свойство (3) следует из признака сравнения абсолютной сходимости рядов (мы применяем его к последовательности равномерно сходящихся к данной обобщённых простых функций).

Свойства (4) и (5) также следуют из абсолютной сходимости суммы двух абсолютно сходящихся рядов и абсолютно сходящегося ряда, умноженного на коэффициент.

Свойство (6) следует из того, что аналогичное неравенство возникает для абсолютно сходящихся рядов — интегралов обобщённых простых функций. $\square$

Теперь покажем, как с помощью функций распределения и плотностей можно посчитать математическое ожидание.

**Лемма 7.4.** *Пусть* $\lambda$ *— конечная абсолютно непрерывная относительно меры* $\mu$ *мера на измеримом пространстве* $(\Omega, \mathfrak{F})$ *и* $\lambda(A) = \int\limits_A p\mathrm{d}\mu$ *для некоторой интегрируемое по* $\mu$ *функции* $p$. *Пусть* $\xi\colon (\Omega, \mathfrak{F}) \to (\mathbb{R}, \mathcal{B})$ *— измеримая функция. Тогда* $\xi$ *интегрируема относительно меры* $\lambda$ *тогда и только тогда, когда* $\xi \cdot p$ *интегрируема относительно меры* $\mu$ *и в случае интегрируемости равны интегралы*

$$\int\limits_\Omega \xi \mathrm{d}\lambda = \int\limits_\Omega \xi \cdot p\mathrm{d}\mu.$$

*Доказательство.* Докажем утверждение для случая конечной меры $\mu$. Если мера $\mu$ является $\sigma$-конечной, то сначала воспользуемся теоремой для множеств $\Omega_k$ из исчерпания $\Omega$, а после воспользуемся сходимость интегралов по исчерпанию для одной из сторон, чтобы доказать сходимость другой.

Пусть сначала $\xi = \sum\limits_{i=1}^{+\infty} a_i \mathbb{1}_{A_i}$ — обобщённая простая функция. Имеем цепочку равенств

$$\int\limits_\Omega \xi \mathrm{d}\lambda = \sum_{i=1}^{+\infty} a_i \lambda(A_i) = \sum_{i=1}^{+\infty} a_i \int\limits_{A_i} p\mathrm{d}\mu = \sum_{i=1}^{+\infty} \int\limits_{A_i} \xi \cdot p\mathrm{d}\mu = \int\limits_\Omega \xi \cdot p\mathrm{d}\mu,$$

где последнее равенство и одновременная сходимость выполнены в силу $\sigma$-аддитивности интеграла Лебега. $\square$

**Теорема 7.5** (О способах вычислении матожидания). *Пусть $\xi$ — случайная величина. Тогда $\xi$ обладает математическим ожиданием тогда и только тогда, когда сходится интеграл Лебега-Стилтьеса*

$$\int\limits_{-\infty}^{+\infty} x \mathrm{d} F_\xi.$$

*Кроме того, когда интеграл сходится, он равен математическому ожиданию $\mathrm{E}\,\xi$. Если $\xi$ абсолютно непрерывна с плотностью $p_\xi$, то существование математического ожидания $\mathrm{E}\,\xi$ равносильно сходимости интеграла Лебега*

$$\int\limits_{-\infty}^{+\infty} x \cdot p_\xi \mathrm{d}\mu,$$

*а в случае сходимости равно этому интегралу.*

*Доказательство.* Поскольку $x \circ \xi = \xi$, то по теореме о прямом образе меры 0.36 имеем одновременную интегрируемость и равенство

$$\int\limits_{\Omega} \xi \mathrm{d}\,\mathrm{P} = \int\limits_{\mathbb{R}} x \mathrm{d} \xi_* \,\mathrm{P}\,.$$

Поскольку мы рассматриваем для вычисления второго интеграла борелевскую $\sigma$-алгебру и на ней по теореме 3.6 совпадают мера Лебега-Стилтьеса, порождённая $F_\xi$, и прямой образ меры $\xi_*\,\mathrm{P}$, то справедливо первое утверждение теоремы.

Эквивалентность второму способу следует из леммы 7.4. □

Из этой теоремы следует, что матожидания одинаково распределённых случайных величин равны.

**Теорема 7.6.** *Пусть $g$ — борелевская функция и $\xi$ — случайная величина. Тогда матожидание $g(\xi)$ существует тогда и только тогда, когда сходится интеграл*

$$\int\limits_{-\infty}^{+\infty} g(x) \mathrm{d} F_\xi,$$

*и равно этому интегралу в случае сходимости. Если $\xi$ абсолютно непрерывна с плотностью $p_\xi$, то существование математического ожидания $\mathrm{E}\,\xi$ равносильно сходимости интеграла Лебега*

$$\int\limits_{-\infty}^{+\infty} g \cdot p_\xi \mathrm{d}\mu,$$

*а в случае сходимости равно этому интегралу.*

*Доказательство.* Полностью повторяет доказательство теоремы 7.5, с точностью до замены $x$ на $g(x)$. □

Поскольку функция распределения монотонна и ограничена, то она является функцией ограниченной вариации и для непрерывных $g$ первый интеграл равен соответствующему (существующему) интегралу Римана-Стилтьеса. Если функция плотности оказывается интегрируемой по Риману на всех промежутках, а $g$ непрерывна, то второй интеграл в условии теоремы можно заменить на интеграл Римана.

Для многомерного распределения верна аналогичная теорема

**Теорема 7.7** (О способах вычислении матожидания)**.** *Пусть $\xi$ — случайный вектор и $g\colon \mathbb{R}^n \to \mathbb{R}$ — борелевская функция. Тогда $g(\xi)$ обладает математическим ожиданием тогда и только тогда, когда сходится многомерный интеграл Лебега-Стилтьеса*

$$\int\limits_{\mathbb{R}^n} g \mathrm{d}F_\xi,$$

*где $F_\xi$ — функция совместного распределения. Когда интеграл сходится, он равен математическому ожиданию $\mathrm{E}\, g(\xi)$. Если $\xi$ абсолютно непрерывен с совместной плотностью $p_\xi$, то существование математического ожидания $\mathrm{E}\, g(\xi)$ равносильно сходимости интеграла Лебега*

$$\int\limits_{\mathbb{R}^n} g \cdot p_\xi \mathrm{d}\mu,$$

*а в случае сходимости равно этому интегралу.*

## 7.2 Моменты и абсолютные моменты случайной величины

Величины $\mathrm{E}\,\xi^n$ обозначаются $\alpha_n$ и называются моментами случайной величины $\xi$. Величины $\mathrm{E}\,|\xi|^n$ обозначаются $\beta_n$ и называются абсолютными моментами случайной величины $\xi$. Доказанное выше предложение говорит нам, что $k$-й момент существует тогда и только тогда, когда существует $k$-й абсолютный момент и $\alpha_n \leqslant \beta_n$.

**Лемма 7.8.** *Пусть у случайной величины $\xi$ существует $k$-й момент $\mathrm{E}\,\xi^k$. Тогда у $\xi$ существуют все моменты с меньшими номерами.*

*Доказательство.* По сказанному выше, у $\xi$ существует $\mathrm{E}\,|\xi|^k$. Достаточно доказать, что для всякого $j < k$ существует момент $\mathrm{E}\,|\xi|^j$. Воспользуемся теоремой 7.6 и рассмотрим интеграл:

$$\int\limits_{\mathbb{R}} |x|^j \mathrm{d}F_\xi = \int\limits_{[-1,1]} |x|^j \mathrm{d}F_\xi + \int\limits_{\mathbb{R}\setminus[-1,1]} |x|^j \mathrm{d}F_\xi.$$

Первый интеграл всегда сходится, так как $|x|^n$ непрерывна, а $F_\xi$ монотонна. Второй же интеграл существует, так как неотрицательная функция $|x|^j$ не превосходит $|x|^k$ при $j < k$ и $|x| > 1$. $\square$

Матожидание $\mu_k = \mathrm{E}(\xi - \mathrm{E}\,\xi)^k$ называется $k$-м центральным моментом случайной величины $\xi$.

**Предложение 7.9.** *Случайная величина $\xi$ обладает $k$-м центральным моментом тогда и только тогда, когда она обладает $k$-м моментом.*

*Доказательство.* Если существует $k$-й центральный момент, то из неравенства $\frac{1}{2}|x|^k < |x - \mathrm{E}\,\xi|^k$ для достаточно больших $x$ следует существование $k$-го момента.

Если существует $k$-й момент, то существуют моменты с меньшими номерами и поэтому существует $\mathrm{E}(\xi - \mathrm{E}\,\xi)^k$. $\square$

**Пример.** Поскольку для бернуллиевской случайной величины $\xi$ имеет место равенство $\xi^k = \xi$, то все (абсолютные) моменты бернуллиеской случайной величины существуют и равны $p$. Вычислим центральные моменты:

$$\mathrm{E}(\xi - p)^k = \sum_{i=0}^{k} C_k^i\,\mathrm{E}\,\xi^i(-p)^{k-i} = p\sum_{i=0}^{k-1} C_k^i(-p)^{k-i} + (-p)^k$$

$$= p(1-p)^k - p(-p)^k + (-p)^k = pq^k + (-1)^k p^k q.$$

★

**Пример.** Вычислим центральные моменты биномиальной случайной величины $\xi \sim \mathrm{B}(n,p)$. Как уже отмечалось ранее $\mathrm{E}\,\xi = np$. Вместо $\xi$ рассмотрим сумму из $n$ независимых бернуллиевских случайных величин $\theta_k \sim \mathrm{Be}(p)$. Тогда

$$\mathrm{E}(\xi - np)^k = \mathrm{E}((\theta_1 - p) + \ldots + (\theta_n - p))^k = \sum_{k_1+\ldots+k_n=k} P(k_1,\ldots,k_n)\,\mathrm{E}(\theta_1-p)^{k_1}\ldots\mathrm{E}(\theta_n-p)^{k_n} =$$

$$= p^k q^k \sum_{k_1+\ldots+k_n=k} P(k_1,\ldots,k_n)(q^{k_1-1} + (-1)^{k_1}p^{k_1-1})\ldots(q^{k_n-1} + (-1)^{k_n}p^{k_n-1}).$$

★

**Пример.** Вычислим моменты для равномерно распределённой случайной величины $\xi \sim \mathrm{R}[a,b]$. Имеем

$$\mathrm{E}\,\xi^k = \int_{-\infty}^{+\infty} x^k p_\xi \mathrm{d}\mu = \int_a^b \tfrac{x^k}{b-a}\mathrm{d}x = \tfrac{1}{b-a}\cdot\tfrac{b^{k+1}-a^{k+1}}{k+1} = \tfrac{1}{k+1}\cdot\tfrac{b^{k+1}-a^{k+1}}{b-a}.$$

★

**Пример.** Вычислим моменты случайной величины $\xi \sim \mathrm{E}(\alpha)$ с экспоненциальным распределением. Имеем

$$\mathrm{E}\,\xi^k = \int\limits_{-\infty}^{+\infty} x^k p_\xi \mathrm{d}\mu = \int\limits_{0}^{+\infty} \alpha x^k e^{-\alpha x} \mathrm{d}x = -\int\limits_{0}^{+\infty} x^k \mathrm{d}e^{-\alpha x} =$$

$$= -x^k e^{-\alpha x}\Big|_0^{+\infty} + \int\limits_{0}^{+\infty} e^{-\alpha x} \mathrm{d}x^k = \int\limits_{0}^{+\infty} k x^{k-1} e^{-\alpha x} \mathrm{d}x = \tfrac{k}{\alpha}\,\mathrm{E}\,\xi^{k-1}.$$

Согласно это формуле

$$\mathrm{E}\,\xi^k = \tfrac{k!}{\alpha^k}\,\mathrm{E}\,\xi^0 = \tfrac{k!}{\alpha^k}.$$

Из формулы для плотности следует, что моменты и абсолютные моменты равны $\beta_n = \alpha_n$. ★

**Пример.** Вычислим моменты и абсолютные моменты случайной величины $\xi \sim \mathcal{N}(0,1)$ стандартного нормального распределения. Для нечётна $n$ функция $\frac{1}{\sqrt{2\pi}} x^n e^{-\frac{x^2}{2}}$ интегрируема на $\mathbb{R}$ по Риману и нечётна, поэтому $\alpha_{2k+1} = 0$. С другой стороны, для чётных $n$ она чётная и неотрицательна, поэтому $\beta_{2k} = \alpha_{2k}$. Так что достаточно посчитать абсолютные моменты. Для $k = 0$ и $k = 1$ из свойств матожидания и теоремы из матанализа имеем

$$\mathrm{E}\,|\xi|^0 = 1,$$

$$\mathrm{E}\,|\xi|^1 = \tfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} |x| e^{-\frac{x^2}{2}} \mathrm{d}x = \tfrac{2}{\sqrt{2\pi}} \int\limits_{0}^{+\infty} x e^{-\frac{x^2}{2}} \mathrm{d}x = -\tfrac{2}{\sqrt{2\pi}} \int\limits_{0}^{+\infty} \mathrm{d}e^{-\frac{x^2}{2}} = -\tfrac{2}{\sqrt{2\pi}}.$$

Далее для $k \leqslant 2$,

$$\mathrm{E}\,|\xi|^k = \int\limits_{-\infty}^{+\infty} |x|^k p_\xi \mathrm{d}\mu = \tfrac{2}{\sqrt{2\pi}} \int\limits_{0}^{+\infty} x^k e^{-\frac{x^2}{2}} \mathrm{d}x = \tfrac{2}{\sqrt{2\pi}} \int\limits_{0}^{+\infty} x^{k-1} e^{-\frac{x^2}{2}} \mathrm{d}\tfrac{x^2}{2} = -\tfrac{2}{\sqrt{2\pi}} \int\limits_{0}^{+\infty} x^{k-1} \mathrm{d}e^{-\frac{x^2}{2}} =$$

$$= -\tfrac{2}{\sqrt{2\pi}} x^{k-1} e^{-\frac{x^2}{2}}\Big|_{-\infty}^{+\infty} + \tfrac{2}{\sqrt{2\pi}} \int\limits_{0}^{+\infty} e^{-\frac{x^2}{2}} \mathrm{d}x^{k-1} = \tfrac{2}{\sqrt{2\pi}} \int\limits_{0}^{+\infty} (k-1) x^{k-2} e^{-\frac{x^2}{2}} \mathrm{d}x = (k-1)\,\mathrm{E}\,|\xi|^{k-2}.$$

Таким образом, $\beta_{2k} = \alpha_{2k} = (2k-1)!!$, $\alpha_{2k+1} = 0$ и $\beta_{2k+1} = (2k)!!\sqrt{\frac{2}{\pi}}$. ★

**Пример.** Для нормально распределённой случайной величины $\xi \sim \mathcal{N}(a, \sigma^2)$ аналогично получаем

$$\mathrm{E}\,|\tfrac{\xi-a}{\sigma}|^{2k} = (2k-1)!!, \mathrm{E}\,|\tfrac{\xi-a}{\sigma}|^{2k+1} = (2k+1)!! \cdot \sqrt{\tfrac{2}{\pi}}.$$

В частности, $\mathrm{E}\,|\xi - a| = 0$, откуда $\mathrm{E}\,\xi = a$. Поскольку $\mathrm{E}\,|\xi - a|^2 = \sigma^2$, то $\mathrm{E}\,\xi^2 - 2a^2 + a^2 = \sigma^2$. Тогда $\mathrm{E}\,\xi^2 = \sigma^2 + a^2$. ★

## 7.3 Дисперсия, ковариация и корреляция

Кроме «среднего значения» нам также хотелось бы знать «отклонение» от него, для чего мы и вводим понятие дисперсии случайной величины $\xi$ определяемой по формуле $\mathrm{D}\,\xi = \mathrm{E}(\xi - \mathrm{E}\,\xi)^2$. Исходя из формулы, дисперсия всегда неотрицательна, квадратный корень $\sqrt{\mathrm{D}\,\xi}$ называется среднеквадратичным отклонением случайной величины $\xi$

**Предложение 7.10.** *Пусть у случайной величины $\xi$ существует математическое ожидание $\mathrm{E}\,\xi$. Тогда дисперсия тогда и только тогда, когда существует матожидание $\mathrm{E}\,\xi^2$ (далее будет сказано, что эта величина называется вторым моментом случайной величины $\xi$) и в случае существования верна формула*

$$\mathrm{D}\,\xi = \mathrm{E}\,\xi^2 - (\mathrm{E}\,\xi)^2.$$

*Доказательство.* Имеем равенство $(\xi - \mathrm{E}\,\xi)^2 = \xi^2 - 2\xi \cdot \mathrm{E}\,\xi + (\mathrm{E}\,\xi)^2$. Поскольку $\mathrm{E}\,\xi$ есть константа и $\xi$ обладает математическим ожиданием, то сумма обладает им тогда и только тогда, когда существует $\mathrm{E}\,\xi^2$. По линейности имеем

$$\mathrm{D}\,\xi = \mathrm{E}(\xi^2 - 2\xi \cdot \mathrm{E}\,\xi + (\mathrm{E}\,\xi)^2) = \mathrm{E}\,\xi^2 - 2(\mathrm{E}\,\xi)(\mathrm{E}\,\xi) + (\mathrm{E}\,\xi)^2 = \mathrm{E}\,\xi^2 - (\mathrm{E}\,\xi)^2.$$

□

**Пример.** Пусть $\theta$ — бернуллиевская случайная величина с вероятностью «успеха» $p$ $(0 < p < 1)$ и вероятностью «неудачи» $1 - p = q$. Тогда случайная величина $\theta^2$ имеет то же распределение, что и $\theta$, поскольку $\theta \in \{0, 1\}$:

$$\theta^2 \sim \begin{array}{cc} 0 & 1 \\ 1-p & p. \end{array}$$

Найдём математическое ожидание квадрата (второй момент):

$$\mathrm{E}\,\theta^2 = 1 \cdot p + 0 \cdot (1 - p) = p.$$

Воспользуемся формулой дисперсии:

$$\mathrm{D}\,\theta = \mathrm{E}\,\theta^2 - (\mathrm{E}\,\theta)^2 = p - p^2 = p(1 - p) = pq.$$

★

**Пример.** Пусть случайная величина $\Pi$ распределена по закону Пуассона с параметром $\lambda > 0$. Ранее вы вычислили, что $\mathrm{E}\,\Pi = \lambda$, теперь найдём дисперсию $\Pi$. Случайная величина $\Pi^2$ имеет распределение, записанное в таблице ниже

$$\Pi^2 \sim \begin{array}{cccccc} 0 & 1 & 2^2 & \ldots & k^2 & \ldots \\ e^{-\lambda} & \lambda e^{-\lambda} & \frac{\lambda^2}{2} e^{-\lambda} & \ldots & \frac{\lambda^k}{k!} e^{-\lambda} & \ldots \end{array}.$$

Тогда

$$\mathrm{E}\,\Pi^2 = e^{-\lambda}\sum_{k=1}^{+\infty}\frac{k^2\lambda^k}{k!} = \lambda e^{-\lambda}\sum_{k=1}^{+\infty}\frac{k\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{k=2}^{+\infty}\frac{\lambda^{k-1}}{(k-2)!} + \lambda e^{-\lambda}\sum_{k=1}^{+\infty}\frac{\lambda^{k-1}}{(k-1)!} =$$
$$= \lambda^2 e^{-\lambda}e^{\lambda} + \lambda e^{-\lambda}e^{\lambda} = \lambda^2 + \lambda.$$

Отсюда

$$\mathrm{D}\,\Pi = \mathrm{E}\,\Pi^2 - (\mathrm{E}\,\Pi)^2 = (\lambda^2+\lambda) - \lambda^2 = \lambda.$$

★

**Пример.** Из прошлого подраздела для равномерно распределённой случайной величины $\xi$ имеем $\mathrm{E}\,\xi = \frac{a+b}{2}$ и $\mathrm{E}\,\xi^2 = \frac{a^2+ab+b^2}{3}$. Тогда $\mathrm{D}\,\xi = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{4(a^2+ab+b^2)-3(a+b)^2}{12} = \frac{a^2-2ab+b^2}{12} = \frac{(a-b)^2}{12}$. ★

**Пример.** Пользуясь вычислениями выше для экспоненциально распределённой случайной величины $\xi$ для первых двух моментов: $\mathrm{E}\,\xi = \frac{1}{\lambda}$ и $\mathrm{E}\,\xi^2 = \frac{2}{\lambda^2}$ получаем $\mathrm{D}\,\xi = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$. ★

**Пример.** Для нормально распределённой случайной величины $\xi \sim \mathcal{N}(a,\sigma^2)$ из доказанных выше равенств $\mathrm{E}\,\xi = a$ и $\mathrm{E}\,\xi^2 = \sigma^2 + a^2$ получаем $\mathrm{D}\,\xi = \sigma^2 + a^2 - a^2 = \sigma^2$. ★

**Предложение 7.11.** *Дисперсия удовлетворяет следующим свойствам:*

(1) $\mathrm{D}\,c = 0$ *для любой константы* $c$*;*

(2) *если существует* $\mathrm{D}\,\xi = 0$*, то для некоторой константы* $c$ *выполнено* $\mathrm{P}(\xi = c) = 1$ *(в частности* $c = \mathrm{E}\,\xi$*);*

(3) *если существует дисперсия* $\mathrm{D}\,\xi$*, то для любой константы* $c$ *существует* $\mathrm{D}(c\xi) = c^2\,\mathrm{D}\,\xi$*.*

(4) *если у случайных величин* $\xi$ *и* $\eta$ *существуют матожидания и дисперсии, то существует дисперсия у их суммы и имеет место формула*

$$\mathrm{D}(\xi+\eta) = \mathrm{D}(\xi) + \mathrm{D}(\eta) + 2\,\mathrm{E}(\xi - \mathrm{E}\,\xi)(\eta - \mathrm{E}\,\eta).$$

*Доказательство.* Все свойства следуют из определения и формулы доказанной выше. Мы докажем только второе и третье.

Если $\mathrm{D}\,\xi = 0$, то $\xi - \mathrm{E}\,\xi$ равно нулю почти наверное (:=почти всюду). Положим $c = \mathrm{E}\,\xi$. Тогда $\mathrm{P}(\xi = c) = 1$.

Поскольку $(\xi + \eta - \mathrm{E}\,\xi - \mathrm{E}\,\eta)^2 \leqslant 2\xi^2 + 2\eta^2 + 2(\mathrm{E}\,\xi + \mathrm{E}\,\eta)^2$, то функция в правой части неравенства интегрируема и, следовательно, существует дисперсия суммы $\mathrm{D}(\xi+\eta)$. Далее,

$$\mathrm{D}(\xi+\eta) = \mathrm{E}(\xi+\eta-\mathrm{E}(\xi+\eta))^2 = \mathrm{E}((\xi-\mathrm{E}\,\xi)+(\eta-\mathrm{E}\,\eta))^2 =$$
$$= \mathrm{E}(\xi-\mathrm{E}\,\xi)^2 + \mathrm{E}(\eta-\mathrm{E}\,\eta)^2 + 2\,\mathrm{E}(\xi-\mathrm{E}\,\xi)(\eta-\mathrm{E}\,\eta) = \mathrm{D}(\xi)+\mathrm{D}(\eta)+2\,\mathrm{E}(\xi-\mathrm{E}\,\xi)(\eta-\mathrm{E}\,\eta).$$

Поскольку все слагаемые, кроме последнего существуют, то и оно тоже существует. □

Последнее доказанное свойство показывает, что дисперсия является квадратичной формой на пространстве случайных величин (и даже более того, положительно определённой формой, но на факторпространстве случайных величин по величинам почти всюду равным константе). Соответствущую ей (поляризованную) билинейную функцию $\mathrm{E}(\xi - \mathrm{E}\,\xi)(\eta - \mathrm{E}\,\eta)$ называют ковариацией случайных величин $\xi$ и $\eta$ и обозначают $\mathrm{Cov}(\xi, \eta)$.

**Предложение 7.12.** *Пусть $\xi, \eta$ — случайные величины, обладающие вторыми моментами. Тогда существуют матожидание произведения $\mathrm{E}\,\xi\eta$ и ковариация $\mathrm{Cov}(\xi, \eta)$. В этом случе ковариация может быть вычислена по формуле*

$$\mathrm{Cov}(\xi, \eta) = \mathrm{E}\,\xi\eta - \mathrm{E}\,\xi\,\mathrm{E}\,\eta.$$

*Доказательство.* Поскольку $(\xi+\eta)^2, (\xi-\eta)^2 \leqslant 2(\xi^2+\eta^2)$, интегрируемы функции $(\xi+\eta)^2$ и $(\xi-\eta)^2$. Отсюда следует интегрируемость произведения

$$\xi\eta = \tfrac{1}{4}(\xi+\eta)^2 - \tfrac{1}{4}(\xi-\eta)^2.$$

Существование ковариации следует из формулы, доказанной ранее. Проверим формулу для ковариации

$$\mathrm{Cov}(\xi, \eta) = \mathrm{E}(\xi - \mathrm{E}\,\xi)(\eta - \mathrm{E}\,\eta) = \mathrm{E}(\xi\eta - \eta\,\mathrm{E}\,\xi - \xi\,\mathrm{E}\,\eta + \mathrm{E}\,\xi\,\mathrm{E}\,\eta) =$$

$$= \mathrm{E}\,\xi\eta - \mathrm{E}\,\eta\,\mathrm{E}\,\xi - \mathrm{E}\,\xi\,\mathrm{E}\,\eta + \mathrm{E}\,\xi\,\mathrm{E}\,\eta = \mathrm{E}\,\xi\eta - \mathrm{E}\,\xi\,\mathrm{E}\,\eta.$$

□

**Предложение 7.13.** *Пусть $L$ — векторное пространство случайных величин обладающих матожиданием и дисперсией рассмотренных с точностью до случайных величин почти наверное равных константе. Тогда* $\mathrm{Cov}$ *на $L$ является положительно определённой квадратичной формой (скалярным произведением), а* $\mathrm{D}$ *— соответствующей ей квадратичной формой.*

*Доказательство.* Если случайная величина $\xi$ почти наверное равна константе, то $\xi - \mathrm{E}\,\xi$ почти наверное равно нулю и $\mathrm{Cov}(\xi, \eta) = 0$ для любой $\eta$, так что $\mathrm{Cov}$ корректно определена на данном векторном пространстве. Билинейность $\mathrm{Cov}$ следует из линейности интеграла и билинейности произведения. Положительная определённость $\mathrm{Cov}$ доказана в предложении 7.11. □

**Предложение 7.14.** *Пусть $\xi_1, \ldots, \xi_n$ — случайные величины. Тогда*

$$\mathrm{D}\left(\sum_{i=1}^{n} \xi_i\right) = \sum \mathrm{D}\,\xi_i + 2\sum_{i<j} \mathrm{Cov}(\xi_i, \xi_j).$$

В евклидовом пространстве через скалярное произведение можно выразить угол между двумя векторами (точнее, его косинус) по формуле $\cos\alpha = \frac{(v,u)}{\sqrt{(v,v)(u,u)}}$.

Мы введём аналог косинуса этого угла, называемый корреляцией случайных величин $\xi$ и $\eta$, вычисляемый по той же формуле

$$\rho(\xi, \eta) = \frac{\mathrm{Cov}(\xi,\eta)}{\sqrt{\mathrm{D}\,\xi\,\mathrm{D}\,\eta}}.$$

Неравенство Коши-Буняковского-Шварца 0.45 говорит нам, что $-1 \leqslant \rho(\xi, \eta) \leqslant 1$ и, что значения $\pm 1$ достигаются когда, $\xi$ и $\eta$ (с точностью до прибавления константы) пропорциональны, то есть связаны соотношением вида $\eta = a\xi + b$.

## 7.4 Связь с независимостью случайных величин

Оказывается, что независимость случайных величин влечёт хорошие свойства матожидания их произведения.

**Предложение 7.15.** *Пусть $\xi, \eta$ — независимые случайные величины, обладающие матожиданием. Тогда выполнено следующее:*

(1) *существует* $\mathrm{E}\,\xi\eta = \mathrm{E}\,\xi\,\mathrm{E}\,\eta$;

(2) *если $\xi$ и $\eta$ обладают дисперсией, то* $\mathrm{D}(\xi + \eta) = \mathrm{D}(\xi) + \mathrm{D}(\eta)$ *и* $\mathrm{Cov}(\xi, \eta) = 0, \rho(\xi, \eta) = 0$.

*Доказательство.* Для простых функций утверждение проверяется непосредственно с использованием пересечения разбиений. Докажем для неотрицательных случайных величин, для произвольных будет следовать из представления в виде суммы положительной и отрицательной частей.

Фиксируем две последовательности простых функций $\xi_n \nearrow \xi$ и $\eta_n \nearrow \eta$. Тогда $\xi_n\eta_n \nearrow \xi\eta$ и по теореме Беппо-Леви интегрируема функция $\xi\eta$ и её интеграл равен

$$\int\limits_\Omega \xi\eta \mathrm{d}\,\mathrm{P} = \lim_{n\to+\infty} \int\limits_\Omega \xi_n\eta_n \mathrm{d}\,\mathrm{P} = \lim_{n\to+\infty} \int\limits_\Omega \xi_n \mathrm{d}\,\mathrm{P} \int\limits_\Omega \eta_n \mathrm{d}\,\mathrm{P} = \int\limits_\Omega \xi \mathrm{d}\,\mathrm{P} \int\limits_\Omega \eta \mathrm{d}\,\mathrm{P}.$$

Утверждения для дисперсии, ковариации и корреляции следуют из доказанных выше формул. □

**Теорема 7.16.** *Пусть $\xi_1, \ldots, \xi_n$ — независимые случайные величины, обладающие матожиданием. Тогда выполнено следующее:*

(1) *существует* $\mathrm{E}(\xi_1 \ldots \xi_n) = \mathrm{E}\,\xi_1 \ldots \mathrm{E}\,\xi_n$;

(2) *если $\xi_k$ обладают дисперсией, то* $\mathrm{D}(\xi_1 + \ldots + \xi_k) = \mathrm{D}(\xi_1) + \ldots + \mathrm{D}(\xi_n)$.

*Доказательство.* Для простых случайных величин доказывается по индукции, далее — аналогично рассуждению из предложения 7.15. □

Случайные величины $\xi$ и $\eta$ корреляция которых равна нулю (что также равносильно равенству ковариации нулю и равенству матожидания произведений произведению матожиданий) называют некоррелированными. Из некоррелированности не следует независимость, что мы проиллюстрируем следующим примером.

**Пример.** Пусть случайные величины $\xi$ и $\zeta$ независимы и их матожидания равны 0. Положим $\eta = \xi\zeta$. Тогда $\mathrm{E}\,\xi\,\mathrm{E}\,\eta = 0 = \mathrm{E}\,\xi^2\,\mathrm{E}\,\zeta = \mathrm{E}\,\xi^2\zeta = \mathrm{E}\,\xi\eta$. Если распределение $\xi$ и $\zeta$ соответствуют таблицам

$$\xi \sim \begin{array}{cc} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{array}, \zeta \sim \begin{array}{ccc} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{array},$$

то распределение $\eta$ записывается в виде

$$\eta \sim \begin{array}{ccc} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{array}.$$

Тогда $\mathrm{P}(\xi = -1, \eta = 0) = \mathrm{P}(\xi = -1, \zeta = 0) = \mathrm{P}(\xi = -1)\,\mathrm{P}(\zeta = 0) = \frac{1}{4}$. С другой стороны, $\mathrm{P}(\xi = -1)\,\mathrm{P}(\eta = 0) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$, так что $\xi$ и $\eta$ не являются независимыми. ★

## 7.5 Квантили и медиана

Рассмотрим ситуацию при которой функция распределения $F_\xi$ случайной величины $\xi$ строго возрастает. Тогда для случайной величины $\xi$ квантилем уровня $0 < \alpha < 1$ называется такое число $x_\alpha$, что выполнено равенство $F_\xi(x_\alpha) = \alpha$ ($\mathrm{P}(\xi \leqslant x_\alpha) = \alpha$).

**Предложение 7.17.** *Пусть $\xi$ — случайная величина с монотонной функцией распределения и $0 < \alpha_1 < \alpha_2 < 1$. Тогда $x_{\alpha_1} < x_{\alpha_2}$ и $\mathrm{P}(x_{\alpha_1} < x \leqslant x_{\alpha_2}) = \alpha_2 - \alpha_1$.*

**Предложение 7.18.** *Пусть $\xi$ — случайная величина с монотонной функцией распределения и $0 < \alpha < 1$. Тогда*

$$\mathrm{P}\left(x_{\frac{1-\alpha}{2}} < x \leqslant x_{\frac{1+\alpha}{2}}\right) = \alpha.$$

Медианой случайной величины $\xi$ называется квантиль уровня 0.5.

Верхней квартилью называется квантиль уровня 0.75 и нижней квартилью — квантиль уровня 0.25. Децилиями называются квантили уровней кратны 0.1.

## 7.6 Мода

Пусть $\xi$ — абсолютно непрерывная случайная величина с плотностью $p_\xi$. Тогда модой случайной величины $\xi$ называются каждая точка локального максимума функции плотности $p_\xi$. Если плотность имеет единственный максимум, то распределение называется унимадольным.

**Пример.** Нормальное распределение является унимодальным. ★

## 7.7 Асимметрия и эксцесс

Пусть $\xi$ — случайная величина, обладающая третьим моментом и $\sigma^2 = \mathrm{D}\,\xi$. Тогда коэффициентом асимметрии случайной величины $\xi$ будем называть число

$$K_{as} = \mathrm{E}\left(\frac{\xi - \mathrm{E}\,\xi}{\sqrt{\mathrm{D}\,\xi}}\right)^3 = \frac{\mu_3}{\sigma^3}.$$

Если $\xi$ обладает четвёртым моментом, то коэффициентом эксцесса случайной величины $\xi$ называется число

$$K_{ex} = \mathrm{E}\left(\frac{\xi - \mathrm{E}\,\xi}{\sqrt{\mathrm{D}\,\xi}}\right)^4 = \frac{\mu_3}{\sigma^3}.$$

**Пример.** Пользуясь формулами для моментов стандартной нормальной случайной величины $\xi$ получаем $K_{ex} = (4-1)!! = 3$. ★

Ввиду рассмотренного примера иногда рассматривают число $K_{ex}^* = K_{ex} - 3$.

**Пример.** Посчитаем ассиметрию и эксцесс биномиального распределения $\xi \sim \mathrm{B}(n, p)$. Для начала заметим, что если $\eta_1, \ldots, \eta_n \sim \mathrm{Be}(p)$ — независимые бернуллиевские случайные величины, то $\eta_1 + \ldots + \eta_n \sim \mathrm{B}(n, p)$. Тогда, пользуясь доказанными ранее свойствами можем сказать, что

$$\mathrm{E}\,\xi = \mathrm{E}\,\eta_1 + \ldots + \mathrm{E}\,\eta_n = np$$

и

$$\mathrm{D}\,\xi = \mathrm{D}\,\eta_1 + \ldots + \mathrm{D}\,\eta_n = npq.$$

Далее, некоторыми выкладками можно показать, что

$$K_{as} = \frac{q-p}{\sqrt{npq}}, K_{ex} = \frac{1-6pq}{npq} + 3.$$

★

**Пример.** Можно показать, что распределение Пуассона имеет следующие коэффициенты асимметрии и эксцесса

$$K_{as} = \frac{1}{\sqrt{\lambda}}, K_{ex} = \frac{1}{\lambda} + 3.$$

★

Распределение случайной величины $\xi$ называется симметричным, если для любого $x \in \mathbb{R}$ выполнено равенство $F_\xi(x) + F_\xi(-x) = 1$.

**Предложение 7.19.** *Пусть $\xi$ — абсолютно непрерывная случайная величина и её плотность чётна, то есть удовлетворяет соотношению $p_\xi(x) = p_\xi(-x)$. Тогда*

(1) *$\xi$ симметрична;*

(2) *если функция распределения $\xi$ строго монотонна, то для любого $0 < \alpha < 1$ имеет место равенство квантилей $x_\alpha = -x_{1-\alpha}$.*

*Доказательство.* Имеем

$$F_\xi(x)+F_\xi(-x)=\int\limits_{-\infty}^{x}p_\xi(u)\mathrm{d}u+\int\limits_{-\infty}^{x}p_\xi(u)\mathrm{d}u=\int\limits_{-\infty}^{x}p_\xi(u)\mathrm{d}u+\int\limits_{-\infty}^{-x}p_\xi(-t)\mathrm{d}(-t)=$$

$$=\int\limits_{-\infty}^{x}p_\xi(u)\mathrm{d}u+\int\limits_{x}^{+\infty}p_\xi(t)\mathrm{d}t=\int\limits_{-\infty}^{+\infty}p_\xi(u)\mathrm{d}u=1.$$

Далее, для квантилей:

$$F_(x_{1-\alpha})=1-\alpha=1-F_\xi(x_\alpha)=F_\xi(-x_\alpha),$$

откуда $x_{1-\alpha}=-x_\alpha$. □

## 7.8 Правило трёх $\sigma$-м

Обозначим через $u_\alpha$ квантиль уровня $\alpha$ у стандартно нормально распределённой случайной величины $\xi\sim\mathcal{N}(0,1)$.

**Предложение 7.20.** *Пусть $\varepsilon\in(0,1)$. Тогда для случайной величины $\eta\sim\mathcal{N}(a,\sigma^2)$ имеет место равенство*

$$\mathrm{P}(|\xi-a|>\sigma u_{1-\frac{\varepsilon}{2}})=\varepsilon.$$

*Доказательство.* Положим $\alpha_1=\frac{\varepsilon}{2}$ и $\alpha_2=1-\frac{\varepsilon}{2}$. В силу симметричности стандартного нормального распределения $u_{\alpha_1}=-u_{\alpha_2}$. Тогда

$$\mathrm{P}(|\xi-a|>\sigma u_{\alpha_2})=1-\mathrm{P}(|\xi-a|\leqslant\sigma u_{\alpha_2})=1-\mathrm{P}(-u_{\alpha_2}\leqslant\tfrac{\xi-a}{\sigma}\leqslant u_{\alpha_2})=$$

$$=1-\mathrm{P}(u_{\alpha_1}\leqslant\tfrac{\xi-a}{\sigma}\leqslant u_{\alpha_2})=1-F_\xi(u_{\alpha_2})+F_\xi(u_{\alpha_1})=1-(1-\tfrac{\varepsilon}{2})+\tfrac{\varepsilon}{2}=\varepsilon.$$

□

**Следствие 7.21** (Правило трёх сигм)**.** *Пусть $\eta\sim\mathcal{N}(a,\sigma^2)$ — нормально распределённая случайная величина. Тогда справедливо приближённое равенство*

$$\mathrm{P}(|\eta-a|>3\sigma)\approx0.0027.$$

*Доказательство.* Пусть $u_{1-\frac{\varepsilon}{2}}=3$. Тогда $\Phi(3)=1-\frac{\varepsilon}{2}$ и $\varepsilon=2-2\Phi(3)$. Поскольку $\Phi(3)\approx0.99865$, то

$$\varepsilon\approx2-2\cdot0.99865=2-1.9973=0.0027.$$

□

## 7.9 Вычисление матожиданий и дисперсий для распределений, связанных с нормальным распределением

### 7.9.1 $\chi^2$-распределение

Пусть $\xi_1, \ldots, \xi_n$ — независимые случайные величины и $\xi_k \sim \mathcal{N}(0,1)$. Распределение $\chi_n^2 := \sum\limits_{k=1}^{n} \xi_k^2$ называется $\chi^2$-распределением с $n$ степенями свободы.

Оно абсолютно непрерывно и имеет плотность, сосредоточенную на положительной полуоси:

$$p_{\chi_n^2}(x) = \frac{\frac{1}{2^n}}{\Gamma(\frac{n}{2})} \cdot x^{\frac{n}{2}-1} e^{-\frac{x}{2}},$$

где $\Gamma(z) = \int\limits_0^{+\infty} t^{z-1} e^{-t} \mathrm{d}t$ — гамма-функция.

Вычислим её численные характеристики.

Ранее мы доказали, что $\mathrm{E}\,\xi^2 = 1$ и $\mathrm{E}\,\xi^4 = 3$ для $\xi \sim \mathcal{N}(0,1)$. Поэтому по независимости имеем $\mathrm{E}\,\chi_n^2 = n\,\mathrm{E}\,\xi_1^2 = n$ и $\mathrm{D}\,\chi_n^2 = n\,\mathrm{D}(\xi_1^2) = n(\mathrm{E}\,\xi_1^4 - (\mathrm{E}\,\xi_1^2)^2) = n(3-1) = 2n$.

Далее,

$$K_{as} = \frac{\mathrm{E}(\chi_n^2 - \mathrm{E}\,\chi_n^2)^3}{\sigma^3} = \frac{\mathrm{E}(\sum\limits_{k=1}^{n}(\xi_k^2 - \mathrm{E}\,\xi_k^2))^3}{2n\sqrt{2n}} =$$

$$= \frac{n\sum\limits_{k=1}^{n}\mathrm{E}(\xi_k^2-1)^3 + 3\sum\limits_{k,j=1}^{n}\mathrm{E}(\xi_k^2-1)^2\,\mathrm{E}(\xi_j^2-1) + 6\sum\limits_{1\leqslant i<j<k\leqslant n}\mathrm{E}(\xi_k^2-1)\,\mathrm{E}(\xi_j^2-1)\,\mathrm{E}(\xi_i^2-1)}{2n\sqrt{2n}} =$$

$$= \frac{n\sum\limits_{k=1}^{n}(15-3\cdot 3+3-1)+0+0}{2n\sqrt{2n}} = \frac{8n}{2n\sqrt{2n}} = \sqrt{\frac{8}{n}}.$$

Для коэффициента эксцесса имеем

$$K_{ex} = \frac{\mathrm{E}(\chi_n^2 - \mathrm{E}\,\chi_n^2)^4}{\sigma^4} = \frac{\mathrm{E}(\sum\limits_{k=1}^{n}(\xi_k^2 - \mathrm{E}\,\xi_k^2))^4}{4n^2} =$$

$$= \frac{n\sum\limits_{k=1}^{n}\mathrm{E}(\xi_k^2-1)^4 + 6\sum\limits_{1\leqslant j<k\leqslant n}\mathrm{E}(\xi_j^2-1)^2\,\mathrm{E}(\xi_k^2-1)^2}{4n^2} =$$

$$= \frac{n\sum\limits_{k=1}^{n}(105-4\cdot 15+6\cdot 3-4+1)+3n(n-1)(3-2\cdot 1+1)^2}{4n^2} = \frac{3n^2+12n}{n^2} = 3 + \frac{12}{n}.$$

### 7.9.2 Распределение Стьютенда

Распределением Стьюдента с $n$ степенями свободы называется называется распределение случайной величины $t_n = \frac{\xi_0}{\sqrt{\frac{\chi_n^2}{n}}}$, где $\xi_0 \sim \mathcal{N}(0,1)$.

Плотность распределения Стьюдента имеет вид

$$p_{t_n}(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi n}}(1+\frac{x^2}{n})^{-\frac{n+1}{2}}.$$

Оно обладает следующими параметрами. Существует $\mathrm{E}\, t_n = 0$ при $n > 1$. Существует $\mathrm{D}\, t_n = \frac{n}{n-2}$ при $n > 2$. Существует $K_{as} = 0$ при $n > 3$. Существует $K_{ex} = \frac{6}{n-4}$ при $n > 4$.

### 7.9.3 $F$-распределение

Распределение случайной величины $F(\nu_1, \nu_2) = \frac{\chi^2_{\nu_1}/\nu_1}{\chi^2_{\nu_2}/\nu_2}$ называется $F$-распределением (распределением Фишера).

Для него при $\nu_2 > 2$ существует матожидание $\mathrm{E}\, F(\nu_1, \nu_2) = \frac{\nu_2}{\nu_2 - 2}$ и для $\nu_2 > 4$ существует дисперсия $\mathrm{D}\, F(\nu_1, \nu_2) = \frac{2\nu_2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$.

### 7.9.4 Логнормальное распределение

Пусть $\xi$ — случайная величина. Если существует такая константа $C$, что существует случайная величина $\ln(\xi - C)$ и она распределена нормальна, то говорят, что $\xi$ имеет логарифмическое нормальное распределение (логнормальное распределение). Имеем формулу для плотности $p_\xi(x) = \frac{1}{(x-C)\sqrt{2\pi}\sigma} e^{-\frac{(\ln(x-C)-a)^2}{2\sigma^2}}$. Для заданных параметров имеем $\mathrm{E}\, \xi = C + e^{a+\frac{\sigma^2}{2}}$, $\mathrm{D}\, \xi = (e^{\sigma^2} - 1) \cdot e^{2a+\sigma^2}$.

**Пример.** Пусть случайная величина $\xi$ распределена по логнормальному закону с параметрами $(0, a, \sigma^2)$. Тогда её плотность имеет вид:

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}\,\sigma x} \exp\left(-\frac{(\ln^2 x - a)^2}{2\sigma^2}\right).$$

Найдём плотность случайной величины $\eta = \ln \xi$. Так как $\xi = e^\eta$, то $g^{-1}(y) = e^y$ — непрерывная монотонно возрастающая функция. Применим предложение 3.8:

$$f_\eta(y) = f_\xi(e^y) \cdot \left|\frac{d}{dy} e^y\right| = f_\xi(e^y) \cdot e^y.$$

Подставим выражение для $f_\xi$:

$$f_\eta(y) = \frac{1}{\sqrt{2\pi}\,\sigma e^y} \exp\left(-\frac{(y-a)^2}{2\sigma^2}\right) \cdot e^y = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(y-a)^2}{2\sigma^2}\right).$$

Таким образом, $\eta = \ln \xi \sim \mathcal{N}(a, \sigma^2)$. ★

## 7.10 Вычисление матожиданий через совместные распределения

# 8 Свёртки случайных величин

d

# 9 Сходимости случайных величин

## 9.1 Сходимость почти наверное

Пусть $\{\xi_n\}$ — последовательность случайных величин. Будем говорить, что она сходится почти наверное к случайной величине $\xi$, если равна 1 вероятностная мера множества $\omega \in \Omega$ таких, что $\lim\limits_{n\to+\infty} \xi_n(\omega) = \xi(\omega)$. Мы будем записывать это как $\mathrm{P}(\xi_n \to \xi) = 1$ или $\mathrm{P}(\xi_n \nrightarrow \xi) = 0$ или $\xi_n \overset{\text{п.н.}}{\to} \xi$.

## 9.2 Сходимость по вероятности

Пусть $\{\xi_n\}$ — последовательность случайных величин. Будем говорить, что она сходится по вероятности к случайной величине $\xi$, если для любого $\gamma > 0$ равен нулю предел вероятностей $\lim\limits_{n\to+\infty} \mathrm{P}(|\xi_n - \xi| \geqslant \gamma) = 0$ или, что равносильно равен единице предел вероятностей $\lim\limits_{n\to+\infty} \mathrm{P}(|\xi_n - \xi| < \gamma) = 1$. Мы будем обозначать данный факт записью $\xi_n \overset{\mathrm{P}}{\to} \xi$.

## 9.3 Пространство $\mathcal{L}_p$ и сходимость в нём

Пусть $1 \leqslant p \leqslant +\infty$. Рассмотрим множество $\tilde{\mathcal{L}}^p$ случайных величин $\xi$, для которых существует математическое ожидание $\mathrm{E}\,|\xi|^p$ (если $p$ конечно).

**Предложение 9.1.** *Множество $\tilde{\mathcal{L}}^p$ является векторным пространством.*

*Доказательство.* Если $\xi \in \tilde{\mathcal{L}}^p$, то $|\xi|^p$ интегрируема. Тогда для любой константы $a \in \mathbb{R}$ интегрируема случайная величина $|a\xi|^p$, поэтому $a\xi \in \tilde{\mathcal{L}}^p$.

Пусть $\xi, \eta \in \tilde{\mathcal{L}}^p$. Тогда $|\xi + \eta|^p \leqslant |2\max(\xi, \eta)|^p \leqslant 2^p(|\xi|^p + |\eta|^p)$. Поэтому из интегрируемости $|\xi|^p$ и $|\eta|^p$ следует интегрируемость $|\xi + \eta|^p$. $\square$

Введём отношение эквивалентности на пространстве случайных величин. Будем говорить, что случайные величины эквивалентны, если они равны почти всюду. Так как данное отношение сохраняет принадлежность к пространству $\tilde{\mathcal{L}}^p$ и эквивалентные нулю случайные величины образую в нём подпространство $J$, то мы можем рассмотреть факторпространство $\mathcal{L}^p = \tilde{\mathcal{L}}^p / J$.

**Теорема 9.2.** *На пространстве $\mathcal{L}^p$ корректно определена функция $||\xi||_p := \sqrt[p]{\mathrm{E}\,|\xi|^p}$. Пара $(\mathcal{L}^p, ||\cdot||_p)$ является нормированным пространством.*

**Теорема 9.3.** *Пространство $(\mathcal{L}^p, ||\cdot||)$ полно. В частности, $(\mathcal{L}^p, ||\cdot||)$ — банахово пространство (то есть полное нормированное векторное пространство).*

Пусть даны последовательность и случайная величина $\{\xi_n\} \subset \mathcal{L}^p, \xi \in \mathcal{L}^p$. Будем говорить, что данная последовательность сходится к $\xi$ по норме $||\cdot||_p$, если $\lim\limits_{n\to+\infty} ||\xi_n - \xi||_p = 0$. Будем обозначать этот факт записью $\xi_n \overset{\mathcal{L}^p}{\to} \xi$.

## 9.4 Сходимость по распределению

Пусть $\{xi_n\}$ — последовательность случайных величин (возможно, заданных на разных вероятностных пространствах) и $\xi$ — ещё одна случайная величина. Тогда будем говорить, что последовательность $\xi_n$ сходится по распределению к $\xi$, если для любой непрерывной ограниченной функции $g$ на $\mathbb{R}$ имеется сходимость предела

$$\lim_{n\to+\infty} \mathrm{E}\, g(\xi_n) = \mathrm{E}\, g(\xi).$$

Будем обозначать это с помощью записи $\xi_n \xrightarrow{D} \xi$.

Следующая теорема объясняет данное нами название.

**Теорема 9.4.** *Следующие утверждения о случайных величинах $\{\xi_n\}, \xi$ и их функциях распределения равносильны:*

(1) $\xi_n \xrightarrow{D} \xi$;

(2) *для всякой точки $x \in \mathbb{R}$ непрерывности функции $F_\xi$ выполнено $\lim\limits_{n\to+\infty} F_{\xi_n}(x) = F_\xi(x)$.*

*Доказательство.* Построим непрерывные ограниченные функции $g_n$ и $h_n$, приближающие сверху и снизу индикатор множества $(-\infty; x]$, где $x$ — точка непрерывности функции $F_\xi$. Тогда

$$\mathrm{E}\, g_n(\xi_m) \leqslant \mathrm{E}\, \mathbb{1}_{(-\infty;x]}(\xi_m) \leqslant \mathrm{E}\, h_n(\xi_m).$$

Переходя к пределу по $m$ получаем

$$\mathrm{E}\, g_n(\xi) \leqslant \lim_{m\to+\infty} \mathrm{E}\, \mathbb{1}_{(-\infty;x]}(\xi_m) \leqslant \mathrm{E}\, h_n(\xi).$$

Так как $x$ — точка непрерывности то, при переходе к пределу по $n$ левая и права части неравенства совпадут. Поскольку

$$\mathrm{E}\, g_n(\xi) \leqslant \mathrm{E}\, \mathbb{1}_{(-\infty;x]}(\xi) \leqslant \mathrm{E}\, h_n(\xi),$$

то выполнено равенство

$$\lim_{m\to+\infty} F_{\xi_m}(x) = \mathrm{E}\, \mathbb{1}_{(-\infty;x]}(\xi_m) = \mathrm{E}\, \mathbb{1}_{(-\infty;x]}(\xi) = F_\xi(x).$$

НУЖНО: дописать явный вид $g_n$ и $h_n$. НУЖНО: дописать обратный переход □

## 9.5 Связь сходимостей

НУЖНО: записать вывод одних сходимостей из других

# 10 Производящие функции

НУЖНО: записать определение

# 11 Характеристические функции

Пусть $\xi$ — случайная величина. Характеристической функцией случайной величины $\xi$ называется функция $f_\xi(t) = \mathrm{E}\, e^{it\xi}$.

Если случайная величина $\xi$ абсолютно непрерывна, то по доказанному ранее (по теореме 7.5) имеем

$$f_\xi(t) = \int\limits_{-\infty}^{+\infty} e^{itx} p_\xi(x) \mathrm{d}x.$$

Таким образом, характеристическая функция является преобразованием Фурье функции плотности.

**Пример.** Пусть имеется соотношение между случайными величинами $\eta = a\xi + b$. Тогда

$$f_\eta(t) = \mathrm{E}\, e^{it\eta} = \mathrm{E}\, e^{it(a\xi+b)} = e^{itb} \cdot \mathrm{E}\, e^{i(at)\xi} = e^{itb} \cdot f_\xi(at).$$

★

**Пример.** Пусть $\xi \sim \Pi(\lambda)$ — случайная величина с пауссоновским распределением. Тогда

$$f_\xi(t) = \mathrm{E}\, e^{it\eta} = \sum_{k=0}^{+\infty} e^{itk} \cdot \tfrac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \cdot e^{e^{it}\lambda} = e^{\lambda(e^{it}-1)}.$$

★

**Пример.** Пусть $\xi \sim \mathrm{Be}(p)$ — бернуллиевская случайная величина. Тогда $f_\xi(t) = \mathrm{E}\, e^{it\xi} = e^{it}p + q$. ★

**Предложение 11.1.** *Пусть $\xi_1, \ldots, \xi_n$ — независимые случайные величины. Тогда*

$$f_{\xi_1+\ldots+\xi_n}(t) = f\xi_1(t) \cdot \ldots \cdot f_{\xi_n}(t).$$

*Если эти случайные величины одинаково распределены, то для любого $1 \leqslant k \leqslant n$ выполнено*

$$f_{\xi_1+\ldots+\xi_n}(t) = (f_{\xi_k}(t))^n$$

*Доказательство.* Поскольку $\xi_1, \ldots, \xi_n$ независимы, то независимы и композиции $e^{it\xi}$ (мы не рассматривали ранее комплексно значные случайные величины, но можно считать, что они являются случайными векторами, независимость определяется как попарная независимость компонент обеих величин) по следствию 6.3.

Тогда по теореме 7.16 имеем

$$f_{\xi_1+\ldots+\xi_n}(t) = \mathrm{E}\, e^{it(\xi_1+\ldots+\xi_n)} = \mathrm{E}(e^{it\xi_1} \cdot \ldots e^{it\xi_n}) = \mathrm{E}\, e^{it\xi_1} \cdot \ldots \cdot \mathrm{E}\, e^{it\xi_n} = f_{\xi_1}(t) \cdot \ldots \cdot f_{\xi_n}(t).$$

Из одинаковой распределённости следует равенство характеристических функции (как интегралов от композиции с данной случайной величиной), поэтому выполнено второе равенство. $\square$

**Пример.** Пусть $\xi \sim \mathcal{N}(0,1)$. Найдём её характеристическую функцию.

$$f_\xi(t) = \tfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{itx} e^{-\frac{x^2}{2}}\,\mathrm{d}x = \tfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} \cos(tx) e^{-\frac{x^2}{2}}\,\mathrm{d}x.$$

Интеграл справа сходится абсолютно и равномерно по $t$, кроме того, сходится абсолютно и равномерно по $t$ интеграл от производной подынтегральной функции по $t$. Тогда имеет место равенство

$$f'_\xi(t) = \tfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} -x\sin(tx) e^{-\frac{x^2}{2}}\,\mathrm{d}x = \tfrac{-1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} \sin(tx) e^{-\frac{x^2}{2}}\,\mathrm{d}\tfrac{x^2}{2} = \tfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} \sin(tx)\mathrm{d}e^{-\frac{x^2}{2}}$$

$$= \tfrac{}{\sqrt{2\pi}} \sin(tx) e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty} - \tfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,\mathrm{d}\sin(tx) = -\tfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} t\cos(tx) e^{-\frac{x^2}{2}}\,\mathrm{d}x = -tf_\xi(t).$$

Решением дифференциального уравнения являются функции вида $Ce^{-\frac{x^2}{2}}$. Поскольку $f_\xi(0) = 1$, то $C = 1$ и $f_\xi(t) = e^{-\frac{x^2}{2}}$. ★

**Пример.** Пользуясь вычислениями выше, получим формулу характеристической функции для случайной величины $\eta \sim \mathcal{N}(a, \sigma^2)$. Положим $\frac{\eta - a}{\sigma} \sim \mathcal{N}(0,1)$. Тогда $f_\xi(t) = e^{-\frac{t^2}{2}}$. Далее, $\eta = \sigma\xi + a$, поэтому

$$f_\eta(t) = e^{ita} f_\xi(\sigma t) = e^{ita - \frac{t\sigma^2}{2}}.$$

★

**Предложение 11.2.** *Пусть $\xi$ — случайная величина и $f_\xi$ — её характеристическая функция. Тогда $f_\xi$ обладает следующими свойствами*

(1) $|f_\xi(t)| \leqslant f_\xi(0) = 1$;

(2) *$f_\xi$ равномерно непрерывна по $t$ на $\mathbb{R}$;*

(3) *если $\xi$ обладает $n$-м моментом $\mathrm{E}\,\xi^n$, то для всех $k \leqslant n$ характеристическая функция $f_\xi$ обладает $k$-й производной и при этом выполнено равенство*

$$f_\xi^{(k)}(t) = \int\limits_{-\infty}^{+\infty} (ix)^k e^{itx} \mathrm{d}F_\xi;$$

(4) *если существует производная $f_\xi^{(2k)}$, то $\xi$ обладает $2k$-м моментом.*

*Доказательство.* Для доказательства (1) имеем

$$f_\xi(0) = \mathrm{E}\, e^{it\cdot 0} = \mathrm{E}\, 1 = 1,$$

$$|f_\xi(t)| = |\,\mathrm{E}\, e^{it\xi}| \leqslant \mathrm{E}\, |e^{it\xi}| = \mathrm{E}\, 1 = 1.$$

Фиксируем $\varepsilon > 0$. Имеем

$$|f_\xi(t) - f_\xi(t+h)| = |\,\mathrm{E}(e^{it\xi} - e^{i(t+h)\xi})| = |\,\mathrm{E}\, e^{it\xi}(1 - e^{ih\xi})| \leqslant \mathrm{E}\,|1 - e^{ih\xi}| =$$

$$= \int\limits_{-\infty}^{+\infty} |1 - e^{ihx}| \mathrm{d}F_\xi = \int\limits_{\mathbb{R}\setminus[-A,A]} |1 - e^{ihx}| \mathrm{d}F_\xi + \int\limits_{-A}^{+A} |1 - e^{ihx}| \mathrm{d}F_\xi,$$

где мера Лебега-Стилтьеса множества $\mathbb{R} \setminus [-A, A]$ не превосходит $\frac{\varepsilon}{4}$. Тогда первый интеграл не превосходит $2 \cdot \frac{\varepsilon}{4} = \frac{\varepsilon}{2}$. Существует $h_0$ такое, что $h_0 A < \frac{\pi}{4}$ и $|1 - e^{ih_0 A}| < \frac{\varepsilon}{4A}$. Тогда для $|h| < h_0$ имеем $\int\limits_{-A}^{+A} |1 - e^{ihx}| \mathrm{d}F_\xi \leqslant 2A \cdot \frac{\varepsilon}{4A} = \frac{\varepsilon}{2}$. Отсюда следует оценка $|f_\xi(t) - f_\xi(t+h)| < \varepsilon$.

Докажем, что если $\xi$ обладает $n$-м моментом, то $f_\xi$ обладает, производными до $n$-го порядка. Для $n = 1$ имеем сходимость интеграла $\mathrm{E}\, \xi e^{it\xi}$, поскольку подынтегральная функция по модулю не превосходит $\xi$. Тогда этот интеграл сходится равномерно относительно $t$, поэтому можно применить теорему о дифференцировании под знаком интеграла. Кроме того, при $t = 0$ получим формулу, требуемую в условии. Продолжая применять рассуждение для моментов с большими номерами получим требуемое. □

**Теорема 11.3** (Теорема единственности)**.** *Пусть $F, G$ — две функции распределения, которым соответствует одна и также характеристическая функция. Тогда $F = G$.*

**Теорема 11.4** (Формула обращения)**.** *Пусть $F = F(x)$ — функция распределения и $f$ — соответствующая её характеристическая функция. Тогда*

(1) *для любых двух точек $a < b$ и непрерывной $F$ выполнено*

$$F(b) - F(a) = \lim_{c \to +\infty} \tfrac{1}{2\pi} \int\limits_{-c}^{c} \tfrac{e^{-iat} - e^{-ibt}}{it} f(t) \mathrm{d}t.$$

(2) *если интеграл $\int\limits_{-\infty}^{+\infty} |f(t)| \mathrm{d}t$ сходится, то функция распределения обладает плотностью $p$, причём выполнено равенство*

$$p(x) = \tfrac{1}{2\pi} \int\limits_{-\infty}^{+\infty} \tfrac{e^{-iat} - e^{-ibt}}{it} f(t) \mathrm{d}t.$$

**Теорема 11.5.** *Для дискретной случайной величины $\xi$ имеет место равенство.*

$$\mathrm{P}(\xi = k) = \tfrac{1}{2\pi}\int\limits_{-\pi}^{\pi} \tfrac{e^{-itk}-e^{-ibt}}{it} f(t)\mathrm{d}t.$$

**Теорема 11.6** (Бохнер, Хинчин)**.** *Пусть $f\colon \mathbb{R} \to \mathbb{C}$ — непрерывная функция и $f(0) = 1$. Тогда следующие утверждения о функции $f$ эквивалентны*

(1) *функция $f = f_\xi$ является характеристической функцией некоторой случайной величины $\xi$;*

(2) *функция $f$ неотрицательно определена,то есть для любого натурального $n$ и любых наборов $t_1, \ldots, t_n \in \mathbb{R}$, $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ вещественных и комплексных чисел неотрицательна сумма*

$$\sum_{i,j=1}^{n} f(t_i - t_j)\lambda_i\overline{\lambda_j} \geqslant 0.$$

**Теорема 11.7** (Теорема непрерывности)**.** *Пусть $F_n$ — последовательность функций распределения и $f_n$ — последовательность соответствующих им характеристических функций. Тогда если $F_n \xrightarrow{W} F$, то $f_n \to f$ поточечно. Если имеется поточечная сходимость $f_t \to f$ и $f$ непрерывна в нуле, то $f$ является характеристической функцией некоторого распределения $F$ и $F_n \xrightarrow{W} F$.*

# 12 Предельные теоремы

## 12.1 Закон нуля или единицы

Пусть $\{A_k\}$ — последовательность событий. Верхним пределом этой последовательности будем называть событие

$$\overline{\lim_n} A_n := \bigcap_{n=1}^{+\infty} \bigcup_{k\geqslant n} A_k.$$

Можно проверить, что $\overline{\lim_n} A_n$ в точности состоит из элементарных событий $\omega$, которые входят в бесконечное число событий $A_n$.

Нижним пределом этой последовательности будем называть событие

$$\underline{\lim_n} A_n := \bigcup_{n=1}^{+\infty} \bigcap_{k\geqslant n} A_k.$$

Можно проверить, что $\underline{\lim_n} A_n$ в точности состоит из элементарных событий $\omega$, которые входят во все события $A_n$ начиная с некоторого номера $n$.

**Лемма 12.1** (Борель, Кантелли)**.** *Пусть* $\{A_k\}$ *— последовательность событий и сходится* $\sum\limits_{k=1}^{+\infty} \mathrm{P}(A_k) < +\infty$*. Тогда* $\mathrm{P}(\overline{\lim\limits_n} A_n) = 0$.

*Доказательство.* Имеем оценку $\mathrm{P}(\bigcup\limits_{k\geqslant n} A_k) \leqslant \sum\limits_{k=n}^{+\infty} \mathrm{P}(A_k) =: s_n$. Поскольку ряд $\sum\limits_{k=1}^{+\infty} \mathrm{P}(A_k)$ сходится, то $\lim\limits_{n\to+\infty} s_n = 0$. Тогда для любого $\varepsilon > 0$ найдётся $n$ такое, что

$$\mathrm{P}(\overline{\lim_n} A_n) \leqslant \mathrm{P}(\bigcup_{k\geqslant n} A_k) \leqslant s_n < \varepsilon.$$

Отсюда следует, что $\mathrm{P}(\overline{\lim\limits_n} A_n) = 0$. □

**Лемма 12.2** (Борель, Кантелли)**.** *Пусть* $\{A_k\}$ *— последовательность независимых событий и расходится ряд* $\sum\limits_{k=1}^{+\infty} \mathrm{P}(A_k) = +\infty$*. Тогда* $\mathrm{P}(\overline{\lim\limits_n} A_n) = 1$.

*Доказательство.* Положим $B_k = \Omega \setminus A_k$. Мы докажем, что

$$\mathrm{P}(\Omega \setminus \overline{\lim_n} A_n) = \mathrm{P}(\underline{\lim_n} B_n) = 0.$$

Поскольку события $A_k$ независимы, то $B_k$ тоже независимы.

Имеем

$$\mathrm{P}(\bigcap_{k=n}^{m} B_k) = \prod_{k=n}^{m} \mathrm{P}(B_k) = \prod_{k=n}^{m}(1 - \mathrm{P}(A_k)) \leqslant e^{-\sum\limits_{k=n}^{m} \mathrm{P}(A_k)}.$$

Переходя к пределу при $m \to +\infty$ получим

$$\mathrm{P}(\bigcap_{k=n}^{m} B_k) \leqslant 0,$$

откуда $\mathrm{P}(\bigcap\limits_{k=n}^{m} B_k) = 0$ и $\mathrm{P}(\underline{\lim\limits_n} B_n) = 0$. □

**Теорема 12.3** (Закон нуля или единицы Колмогорова)**.** *Пусть* $\{\xi_k\}$ *— последовательность независимых случайных величин и* $\mathfrak{X} = \bigcap\limits_{k=1}^{+\infty} \sigma(\xi_k, \xi_{k+1}, \ldots)$ *— хвостовая $\sigma$-алгебра. Тогда для любого хвостового события* $A \in \mathfrak{X}$ *либо* $\mathrm{P}(A) = 0$*, либо* $\mathrm{P}(A) = 1$.

*Доказательство.* Докажем, что всякое хвостовое событие независимо с самим собой. Тогда по критерию 2.9 его вероятностная мера будет равна либо 0, либо 1.

Пусть $A \in \mathfrak{X}$ — хвостовое событие. Тогда $A \in\subset \sigma(\xi_k, \xi_{k+1}, \ldots)$ и по предложению 6.9 событие $A$ независимо со всяким событием из $\sigma$-алгебры $\sigma(\xi_1, \ldots, \xi_{k-1})$. Тогда

$A$ независимо со всяким событием из алгебры $\bigcup\limits_{n=k}^{+\infty} \sigma(\xi_1, \ldots, \xi_n)$ и по теореме 6.5 событие $A$ независимо со всяким событием из $\sigma$-алгебры (см. предложение 6.8)

$$\sigma(\xi_1, \xi_{k+1}, \ldots) = R_\sigma(\bigcup_{n=k}^{+\infty} \sigma(\xi_1, \ldots, \xi_n)).$$

Так как $A \in \sigma(\xi_1, \xi_{k+1}, \ldots)$, то $A$ независимо с самим собой. □

**Следствие 12.4.** *Пусть $\{\xi_n\}$ — последовательность независимых случайных величин. Тогда предел $\lim\limits_{n\to+\infty} \xi_n$ существует с вероятностью 1, либо не существует с вероятностью 1. Ряд $\sum\limits_{n=1}^{+\infty} \xi_n$ либо сходится с вероятностью 1, либо расходится с вероятностью 1.*

*Доказательство.* Рассмотрим событие $C_{m,k_1,k_2}$, выражающее истинность неравенства $|\xi_{k_1} - \xi_{k_2}| < \frac{1}{m}$. Тогда $C_{m,k_1,k_2} \in \sigma(\xi_{k_1}, \ldots, \xi_{k_2})$. Далее,

$$C_{m,k_1} := \bigcap_{k_2=k_1}^{+\infty} C_{m,k_1,k_2} \in \sigma(\xi_{k_1}, \xi_{k_1+1}, \ldots),$$

$$C_{m,\geqslant n} := \bigcap_{k_1=n}^{+\infty} C_{m,k_1} \in \sigma(\xi_n, \xi_{n+1}, \ldots),$$

$$C_{n_0} := \bigcap_{m=1}^{+\infty} \bigcup_{n=n_0}^{+\infty} C_{m,n} \in \sigma(\xi_{n_0}, \xi_{n_0+1}, \ldots).$$

Поскольку сходимость предела равносильна сходимости начиная с любого момента, а сходимость с момента $n$ попаданию в событие $C_n$, то для всех $n$ события $C_n$ совпадают и поэтому $C = C_n \in \mathfrak{X}$. По закону нуля или единицы 12.3 имеем $\mathrm{P}(C) = 0$ или $\mathrm{P}(C) = 1$.

Доказательство для сходимости ряда аналогично. □

**Пример.** Пусть $\{X_n\}$ — последовательность независимых случайных величин, принимающих только значения $\pm 1$. Ряд $\sum\limits_{n=1}^{+\infty} \frac{X_n}{n}$ либо сходится с вероятностью 1, либо расходится с вероятностью 1. ★

**Теорема 12.5.** *Пусть $\{\xi_n\}$ — последовательность независимых случайных величин. Тогда сходимость ряда $\sum\limits_{n=1}^{+\infty} \xi_n$ с вероятностью 1 влечёт сходимости рядов $\sum\limits_{n=1}^{+\infty} \mathrm{E}\,\xi_n$ и $\sum\limits_{n=1}^{+\infty} \mathrm{D}\,\xi_n$. Если случайные величины равномерно ограничены некоторой константой $C$, то есть для любого $n \geqslant 1$ выполнено $\mathrm{P}(|\xi_n| \leqslant C) = 1$, то верно и обратное.*

## 12.2 Неравенства

**Теорема 12.6** (Неравенства Чебышёва и Маркова)**.** *Пусть $\xi$ — неотрицательная случайная величина и $\varepsilon > 0$. Тогда при дополнительных условиях имеют место следующие неравенства.*

(1) *Если существует матожидание $\mathrm{E}\,\xi$, то справедливо первое неравенство Чебышёва*
$$\mathrm{P}(\xi \geqslant \varepsilon) \leqslant \tfrac{\mathrm{E}\,\xi}{\varepsilon};$$

(2) *Если существует матожидание $\mathrm{E}\,\xi^p$ для некоторого $p > 0$, то выполнено неравенство Маркова*
$$\mathrm{P}(\xi \geqslant \varepsilon) \leqslant \tfrac{\mathrm{E}\,\xi^p}{\varepsilon^p};$$

(3) *Если существует дисперсия $\mathrm{D}\,\xi$, то имеет место второе нервенство Чебышёва*
$$\mathrm{P}(|\xi - \mathrm{E}\,\xi| \leqslant \varepsilon) \leqslant \tfrac{\mathrm{D}\,\xi}{\varepsilon^2}.$$

*Доказательство.* Для доказательства первого неравенства Чебышёва рассмотрим случайную величину $\varepsilon \cdot \mathbb{1}_{\{\omega\in\Omega|\xi(\omega)\geqslant\varepsilon\}} \leqslant \xi$. Взяв матожидания от обеих частей и разделив на $\varepsilon$ получим неравенство Чебышёва.

Неравенство Маркова немедленно следует из первого неравенства Чебышёва, поскольку $\mathrm{P}(\xi^p \geqslant \varepsilon^p) = \mathrm{P}(\xi \geqslant \varepsilon)$.

Для доказательства второго неравенства Чебышёва рассмотрим случайную величину $|\xi - \mathrm{E}\,\xi|$. Тогда, применив к ней неравенство Маркова для $p = 2$ получим второе неравенство Чебышёва. □

## 12.3 Закон больших чисел

**Теорема 12.7** (Закон больших чисел в форме Чебышёва)**.** *Пусть $\{\xi_k\}$ — последовательность попарно независимых (некоррелированных) случайных величин и для некоторой константы $C > 0$ выполнены равенства $\mathrm{D}\,\xi_k \leqslant C$. Положим $S_n := \sum\limits_{i=1}^{n} \xi_n$. Тогда для любого $\varepsilon > 0$ существует предел*
$$\lim_{n\to+\infty} \mathrm{P}(|\tfrac{S_n}{n} - \mathrm{E}\,\tfrac{S_n}{n}| \geqslant \varepsilon) = 0$$
*или, в других обозначениях*
$$\tfrac{S_n}{n} - \mathrm{E}\,\tfrac{S_n}{n} \xrightarrow{\mathrm{P}} 0.$$

*Доказательство.* Из второго неравенства Чебышёва (3) имеем

$$\mathrm{P}(|\tfrac{S_n}{n} - \mathrm{E}\,\tfrac{S_n}{n}| \geqslant \varepsilon) \leqslant \frac{\mathrm{D}\,\frac{S_n}{n}}{\varepsilon^2}.$$

Далее, по свойствам дисперсии и независимости (некоррелированности)

$$|\,\mathrm{D}\,\tfrac{S_n}{n}| = \mathrm{D}\,\tfrac{S_n}{n} = \tfrac{\mathrm{D}\,S_n}{n^2} = \sum_{k=1}^{n} \tfrac{\mathrm{D}\,\xi_k}{n^2} \leqslant \tfrac{nC}{n^2} = \tfrac{C}{n}.$$

Тогда

$$\mathrm{P}(|\tfrac{S_n}{n} - \mathrm{E}\,\tfrac{S_n}{n}| \geqslant \varepsilon) \leqslant \tfrac{C}{n\varepsilon^2}.$$

Поскольку при $n$ стремящемся к бесконечности правая часть неравенства стремится к нулю, а левая часть неравенства неотрицательна, то она тоже стремится к нулю. □

**Следствие 12.8.** *Пусть $\{\xi_k\}$ — последовательность попарно независимых (некоррелированных) одинаково распределённых случайных величин и $\mathrm{E}\,\xi_k = m$, $\mathrm{D}\,\xi_k = \sigma^2 < +\infty$. Положим $S_n := \sum\limits_{i=1}^{n} \xi_n$. Тогда для любого $\varepsilon > 0$ существует предел*

$$\lim_{n\to+\infty} \mathrm{P}(|\tfrac{S_n}{n} - m| \geqslant \varepsilon) = 0$$

*или, в других обозначениях*

$$\tfrac{S_n}{n} \xrightarrow{\mathrm{P}} m.$$

*Более того,*

$$\mathrm{P}(|\tfrac{S_n}{n} - m| \geqslant \varepsilon\tfrac{\sigma}{\sqrt{n}}) \leqslant \tfrac{1}{\varepsilon^2}.$$

*Доказательство.* Первая часть выполнена, так как $\mathrm{E}\,\frac{S_n}{n} = m$. Вторая является следствием второго неравенства Чебышёва для $\varepsilon := \varepsilon\frac{\sigma}{\sqrt{n}}$. □

**Теорема 12.9** (Теорема Пуассона)**.** *Пусть $\nu_n \sim \mathrm{B}(n, p_n)$ — случайная величина числа «успехов» в $n$ независимых испытаниях Бернулли с вероятностью «успеха» $p_n$ (другими словами, $\nu_n$ — биномиальная случайная величина). Пусть при $n \to +\infty$ имеет место $np_n \to \lambda$ для некоторого фиксированного $\lambda$. Тогда*

$$\lim_{n\to+\infty} \mathrm{P}(\nu_n = k) = \tfrac{\lambda^k}{k!}e^{-\lambda}.$$

*Доказательство.* Поскольку $\lim\limits_{n\to+\infty} np_n = \lambda$, то при $n \to +\infty$ выполнено $p_n = \frac{\lambda}{n} + o(\frac{1}{n})$

Имеем

$$\mathrm{P}(\nu_n = k) = C_n^k p_n^k q_n^{n-k} = \tfrac{n!}{k!(n-k)!}\left(\tfrac{\lambda}{n} + o(\tfrac{1}{n})\right)^k \left(1 - \tfrac{\lambda}{n} + o(\tfrac{1}{n})\right)^{n-k} =$$

$$= \tfrac{1}{k!}\tfrac{n!}{(n-k)!n^k}\left(\lambda + o(1)\right)^k \left(1 - \tfrac{\lambda}{n} + o(\tfrac{1}{n})\right)^n \left(1 - \tfrac{\lambda}{n} + o(\tfrac{1}{n})\right)^{-k}.$$

После перехода к пределу $n \to +\infty$ множители $\frac{n!}{(n-k)!n^k}$ и $\left(1 - \frac{\lambda}{n} + o(\frac{1}{n})\right)^{-k}$ обратятся в единицы, а оставшиеся устремятся в точности к $\frac{\lambda^k}{k!}e^{-\lambda}$. □

НУЖНО: поправить лемму и теорему ниже

**Лемма 12.10.** *Пусть* $\nu_n \sim \mathrm{B}(n, p_n)$ *— случайная величина числа «успехов» в* $n$ *независимых испытаниях Бернулли с вероятностью «успеха»* $p_n$ *(другими словами,* $\nu_n$ *— биномиальная случайная величина). Тогда для любого подмножества* $B \subset \mathbb{Z}_{\geqslant 0}$ *выполнено неравенство*

$$\left|\mathrm{P}(\nu_n \in B) - \sum_{k \in B} \tfrac{\lambda^k}{k!} e^{-\lambda}\right| \leqslant \min(p, np^2).$$

**Теорема 12.11** (О скорости сходимости в теореме Пуассона)**.** *Пусть* $\nu_n \sim \mathrm{B}(n, p_n)$ *— случайная величина числа «успехов» в* $n$ *независимых испытаниях Бернулли с вероятностью «успеха»* $p_n$ *(другими словами,* $\nu_n$ *— биномиальная случайная величина). Пусть при* $n \to +\infty$ *имеет место* $np_n \to \lambda$ *для некоторого фиксированного* $\lambda$*. Тогда для любого подмножества* $B \subset \mathbb{Z}_{\geqslant 0}$ *выполнено неравенство*

$$\left|\mathrm{P}(\nu_n \in B) - \sum_{k \in B} \tfrac{\lambda^k}{k!} e^{-\lambda}\right| \leqslant \min(p, np^2).$$

**Теорема 12.12** (Усиленный закон больших чисел для схемы Бернулли)**.** *Пусть* $\theta_1, \ldots, \theta_n \sim \mathrm{Be}(p)$ *— независимые бернуллиевские случайные величины с вероятностью «успеха»* $p$*. Положим* $\nu_n = \theta_1 + \ldots + \theta_n \sim \mathrm{B}(n, p)$*. Тогда случайная величина* $\frac{\nu_n}{n}$ *сходится к* $p$ *почти наверное. Эквивалентно,*

$$\lim_{n \to +\infty} \mathrm{P}(\sup_{k \geqslant n} \left|\tfrac{\nu_k}{k} - p\right| > \varepsilon) = 0$$

*для любого* $\varepsilon > 0$*.*

*Доказательство.* Имеем оценку

$$\mathrm{P}(\sup_{k \geqslant n} \left|\tfrac{\nu_k}{k} - p\right| > \varepsilon) = \mathrm{P}(\bigcup_{k \geqslant n} (\left|\tfrac{\nu_k}{k} - p\right| > \varepsilon)) \leqslant \sum_{k \geqslant n} \mathrm{P}(\left|\tfrac{\nu_k}{k} - p\right| > \varepsilon).$$

и воспользуемся неравенством Маркова с параметром 4:

$$\mathrm{P}(\left(\tfrac{\nu_k}{k} - p\right) > \varepsilon) \leqslant \frac{\mathrm{E}\left(\frac{\nu_k}{k} - p\right)^4}{\varepsilon^4}.$$

Пользуясь формулой для центральных моментов биномиальных случайных величин (см. соответствующий раздел), перебирая разбиения $(4 = 4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1)$ получим

$$\mathrm{E}\left(\tfrac{\nu_k}{k} - p\right)^4 = \tfrac{1}{k^4}(k \cdot (pq^4 + qp^4) + 0 + C_4^2 C_k^2 (pq^2 + p^2 q) + 0 + 0) = \tfrac{pq(p^3 + q^4)}{k^3} + \tfrac{3pq(k-1)}{k^3}.$$

Тогда ряд $\sum\limits_{k=1}^{+\infty} \mathrm{P}(\left|\frac{\nu_k}{k} - p\right| > \varepsilon)$ сходится. Отсюда следует, что его хвостовые суммы $\sum\limits_{k \geqslant n} \mathrm{P}(\left|\frac{\nu_k}{k} - p\right| > \varepsilon)$ стремятся к нулю. □

В законе больших чисел для одинаково распределённых независимых случайных величин оказывается не обязательно требовать даже существования у них дисперсии. В такой формулировке для доказательства теоремы уже не удаётся применить второе неравенство Чебышёва, но её удаётся доказать как следствие закона нуля или единицы 12.3 или с использованием инструментария характеристических функций.

**Теорема 12.13** (Закон больших чисел в форме Хинчина)**.** *Пусть* $\{\xi_k\}$ *— последовательность попарно независимых одинаково распределённых случайных величин и* $\mathrm{E}\,\xi_k = m$. *Положим* $S_n := \sum\limits_{i=1}^{n} \xi_n$. *Тогда для любого* $\varepsilon > 0$ *существует предел*

$$\lim_{n\to+\infty} \mathrm{P}(|\tfrac{S_n}{n} - m| \geqslant \varepsilon) = 0$$

*или, в других обозначениях*

$$\tfrac{S_n}{n} \xrightarrow{\mathrm{P}} m.$$

*Доказательство.* Положим $f_{\xi_i}(t) = \mathrm{E}\, e^{it\xi_i}$ — характеристическая функция. Также положим $f_{\frac{S_n}{n}}(t) = \mathrm{E}\, e^{\frac{itS_n}{n}}$ — характеристическая функция случайной величины $\frac{S_n}{n}$.

В силу независимости $\xi_k$ имеем

$$f_{\frac{S_n}{n}}(t) = \mathrm{E}\, e^{\frac{itS_n}{n}} = \mathrm{E}\, e^{it\frac{\xi_1+\ldots+\xi_n}{n}} = \mathrm{E}\, e^{\frac{it\xi_1}{n}+\ldots+\frac{it\xi_n}{n}} = \mathrm{E}\, e^{\frac{it\xi_1}{n}} \cdot \ldots \cdot \mathrm{E}\, e^{\frac{it\xi_n}{n}}.$$

В силу одинаковой распределённости

$$\mathrm{E}\, e^{\frac{it\xi_1}{n}} \cdot \ldots \cdot \mathrm{E}\, e^{\frac{it\xi_n}{n}} = (\mathrm{E}\, e^{\frac{it\xi_1}{n}})^n = f_{\xi_1}^n(\tfrac{t}{n}).$$

Тогда по НУЖНО: вставить свойство хар функций имеем

$$f_{\xi_1}(\tfrac{t}{n}) = 1 + i\tfrac{t}{n}\,\mathrm{E}\,\xi_1 + o(\tfrac{t}{n}) == 1 + i\tfrac{t}{n}m + o(\tfrac{t}{n}), \tfrac{t}{n} \to 0.$$

Подставляя в эту формулу $t = t_0$ при $n \to +\infty$ получим

$$f_{\xi_1}(\tfrac{t_0}{n}) = 1 + i\tfrac{t_0}{n}m + o(\tfrac{1}{n}).$$

Тогда

$$f_{\xi_1}^n(\tfrac{t_0}{n}) = \left(1 + i\tfrac{t_0}{n}m + o(\tfrac{1}{n})\right)^n.$$

Отсюда

$$\lim_{n\to+\infty} f_{\frac{S_n}{n}}(t_0) = \lim_{n\to+\infty} f_{\xi_1}^n(\tfrac{t_0}{n}) = \lim_{n\to+\infty} \left(1 + i\tfrac{t_0}{n}m + o(\tfrac{1}{n})\right)^n = e^{it_0m} = f_m(t).$$

Из теоремы непрерывности следует, что $\frac{S_n}{n} \xrightarrow{D} m$. Так как $m$ — константа, то по предложению НУЖНО: дописать отсюда $\frac{S_n}{n} \xrightarrow{\mathrm{P}} m$. □

## 12.4 Центральная предельная теорема

**Теорема 12.14** (Центральная предельная теорема)**.** *Пусть $\{\xi_k\}$ — последовательность независимых одинаково распределённых случайных величин, обладающих конечными вторыми моментами* $\mathrm{E}\,\xi_k^2$ *(и, следовательно, конечными матожиданиями и дисперсиями). Положим $S_n = \sum\limits_{k=1}^{n} \xi_k$. Тогда для любого $x \in \mathbb{R}$ существует предел*

$$\lim_{n\to+\infty} \mathrm{P}(\tfrac{S_n - \mathrm{E}\,S_n}{\sqrt{\mathrm{D}\,S_n}} \leqslant x) = \Phi(x),$$

*где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{x^2}{2}}$ — функция распределения стандартного распределения. Другими словами, имеется сходимость по распределению*

$$\tfrac{S_n - \mathrm{E}\,S_n}{\sqrt{\mathrm{D}\,S_n}} \xrightarrow{D} \eta,$$

*где $\eta \sim \mathcal{N}(0,1)$.*

*Доказательство.* Положим $m = \mathrm{E}\,\xi_k$ и $\sigma^2 = \mathrm{D}\,\xi_k$. Имеем

$$f_{\xi_k - m}(t) = \mathrm{E}\,e^{it(\xi_k - m)}.$$

По независимости:

$$f_{\frac{S_n - \mathrm{E}\,S_n}{\sqrt{\mathrm{D}\,S_n}}}(t) = \mathrm{E}\exp\left(it\frac{S_n - \sum\limits_{k=1}^{n}\mathrm{E}\,\xi_k}{\sqrt{\sum\limits_{k=1}^{n}\mathrm{D}\,\xi_k}}\right) = \mathrm{E}\,e^{it\frac{\sum\limits_{k=1}^{n}(\xi_k - m)}{\sqrt{n}\sigma}} = \mathrm{E}\prod_{k=1}^{n} e^{it\frac{\xi_k - m}{\sqrt{n}\sigma}} = \prod_{k=1}^{n} \mathrm{E}\,e^{it\frac{\xi_k - m}{\sqrt{n}\sigma}}.$$

По одинаковой распределённости:

$$\prod_{k=1}^{n} \mathrm{E}\,e^{it\frac{\xi_k - m}{\sqrt{n}\sigma}} = \prod_{k=1}^{n} f_{\xi_k - m}(\tfrac{t}{\sqrt{n}\sigma}). = \left(f_{\xi_1 - m}(\tfrac{t}{\sqrt{n}\sigma})\right)^n.$$

При $\frac{t}{\sqrt{n}} \to 0$ из свойств характеристических функий имеем разложение по формуле Тейлора

$$f_{\xi_k - m}(\tfrac{t}{\sqrt{n}\sigma}) = 1 + \tfrac{it}{\sqrt{n}\sigma}\,\mathrm{E}(\xi_k - m) + \tfrac{(it)^2}{2n\sigma^2}\,\mathrm{E}(\xi_k - m)^2 + o(\tfrac{t^3}{n\sqrt{n}\sigma^3}) = 1 - \tfrac{t^2}{2n} + o(\tfrac{t^3}{n\sqrt{n}}).$$

При фиксированном $t = t_0$ и $n \to +\infty$ получаем

$$f_{\xi_k - m}(\tfrac{t_0}{\sqrt{n}\sigma}) = 1 - \tfrac{t_0^2}{2n} + o(\tfrac{1}{n\sqrt{n}}).$$

Отсюда

$$\lim_{n\to+\infty}\left(1 - \tfrac{t_0^2}{2n} + o(\tfrac{1}{n\sqrt{n}})\right)^n = e^{-\frac{t_0^2}{2}} = f_\eta(t_0),$$

где $\eta \sim \mathcal{N}(0,1)$.

Из теоремы непрерывности следует утверждение теоремы. □

Центральная предельная теорема допускает обобщение на случай, когда случайные величины не одинаково распределены. В этом случае требует выполнение дополнительных условий, как демонстрирует следующая

**Теорема 12.15** (Обобщение центральной предельной теоремы с условием Линдеберга)**.** *Пусть $\{\xi_k\}$ — последовательность независимых случайных величин, обладающих вторым моментом* $\mathrm{E}\,\xi^2$ *(и, следовательно, математическим ожиданием и дисперсией). Положим* $\sigma_k^2 = \mathrm{D}\,\xi_k > 0$ *и* $S_n = \sum\limits_{k=1}^{n} \xi_k$, $\mathrm{D}\,S_n := B_n^2$. *Пусть* $F_{\xi_k}$ *— функция распределения случайно величины* $\xi_k$. *Тогда, если для любого* $\varepsilon > 0$ *выполнено условие Линдеберга*

$$\lim_{n\to+\infty} \tfrac{1}{B_n^2} \sum_{k=1}^{n} \int\limits_{\{x\in\mathbb{R} \mid |x-\mathrm{E}\,\xi_k| \geqslant \varepsilon B_n\}} (x-\mathrm{E}\,\xi_k)^2 \mathrm{d}F_{\xi_k} = 0,$$

*то для любого* $x \in \mathbb{R}$ *имеет место сходимость*

$$\lim_{n\to+\infty} \mathrm{P}\left(\tfrac{S_n-\mathrm{E}\,S_n}{\sqrt{\mathrm{D}\,S_n}}\right) = \Phi(x),$$

*где* $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{x^2}{2}}$ *— функция распределения стандартного распределения. Другими словами, имеется сходимость по распределению*

$$\tfrac{S_n-\mathrm{E}\,S_n}{\sqrt{\mathrm{D}\,S_n}} \xrightarrow{D} \eta,$$

*где* $\eta \sim \mathcal{N}(0,1)$.

**Теорема 12.16** (Обобщение центральной предельной теоремы с условием Ляпунова)**.** *Пусть $\{\xi_k\}$ — последовательность независимых случайных величин, обладающих вторым моментом* $\mathrm{E}\,\xi^2$ *(и, следовательно, математическим ожиданием и дисперсией). Положим* $\sigma_k^2 = \mathrm{D}\,\xi_k > 0$ *и* $S_n = \sum\limits_{k=1}^{n} \xi_k$, $\mathrm{D}\,S_n := B_n^2$. *Пусть* $F_{\xi_k}$ *— функция распределения случайно величины* $\xi_k$. *Тогда, если для любого* $\varepsilon > 0$ *и некоторого* $\delta > 0$ *выполнено условие Ляпунова*

$$\lim_{n\to+\infty} \tfrac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} \mathrm{E}(\xi_k - \mathrm{E}\,\xi_k)^{2+\delta} = 0,$$

*то для любого* $x \in \mathbb{R}$ *имеет место сходимость*

$$\lim_{n\to+\infty} \mathrm{P}\left(\tfrac{S_n-\mathrm{E}\,S_n}{\sqrt{\mathrm{D}\,S_n}}\right) = \Phi(x),$$

*где* $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{x^2}{2}}$ *— функция распределения стандартного распределения. Другими словами, имеется сходимость по распределению*

$$\tfrac{S_n-\mathrm{E}\,S_n}{\sqrt{\mathrm{D}\,S_n}} \xrightarrow{D} \eta,$$

*где* $\eta \sim \mathcal{N}(0,1)$.

*Доказательство.* Докажем, что условие Ляпунова влечёт условие Линдеберга (то есть является более сильным).

Фиксируем $\varepsilon > 0$. Имеем

$$\mathrm{E}(\xi_k - \mathrm{E}\,\xi_k)^{2+\delta} = \int\limits_{-\infty}^{+\infty} (x - \mathrm{E}\,\xi_k)^{2+\delta} \mathrm{d}F_{\xi_k} \geqslant \int\limits_{\{x\in\mathbb{R} \mid |x - \mathrm{E}\,\xi_k| \geqslant \varepsilon B_n\}} (x - \mathrm{E}\,\xi_k)^{2+\delta} \mathrm{d}F_{\xi_k} \leqslant$$

$$\leqslant \varepsilon^\delta B_n^\delta \cdot \int\limits_{\{x\in\mathbb{R} \mid |x - \mathrm{E}\,\xi_k| \geqslant \varepsilon B_n\}} (x - \mathrm{E}\,\xi_k)^2 \mathrm{d}F_{\xi_k}.$$

Отсюда

$$\tfrac{1}{B_n^2} \sum_{k=1}^{n} \int\limits_{\{x\in\mathbb{R} \mid |x - \mathrm{E}\,\xi_k| \geqslant \varepsilon B_n\}} (x - \mathrm{E}\,\xi_k)^2 \mathrm{d}F_\xi \leqslant$$

$$\varepsilon^{-\delta} \cdot \tfrac{\varepsilon^\delta B_n^\delta}{B_n^{2+\delta}} \sum_{k=1}^{n} \int\limits_{\{x\in\mathbb{R} \mid |x - \mathrm{E}\,\xi_k| \geqslant \varepsilon B_n\}} (x - \mathrm{E}\,\xi_k)^2 \mathrm{d}F_\xi \leqslant$$

$$\leqslant \varepsilon^{-\delta} \cdot \tfrac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} \mathrm{E}(\xi_k - \mathrm{E}\,\xi_k)^{2+\delta}.$$

Переходя к пределу $n \to +\infty$ получим нуль для самого последнего выражения по условию Ляпунова. В силу неотрицательности первого выражения в последней цепочке оно также стремится к нулю. Таким образом, выполнено условие Линдеберга. □

**Теорема 12.17** (Оценка Берри-Эссена)**.** *Пусть $\{\xi_k\}$ — последовательность независимых одинаково распределённых случайных величин, обладающих моментами $\mathrm{E}\,\xi^k < +\infty$ (и, следовательно, обладающие математическим ожиданием и дисперсией). Положим $m = \mathrm{E}\,\xi_k$, $\sigma^2 = \mathrm{D}\,\xi_k$ и $S_n = \sum\limits_{k=1}^{n} \xi_k$. Если случайные величины последовательности также обладают третьими моментами $\mathrm{E}\,|\xi|^3 < +\infty$, то для любого натурального $n$ справедлива оценка*

$$\sup_{x\in\mathbb{R}} \left| \mathrm{P}\left( \tfrac{S_n - \mathrm{E}\,S_n}{\sqrt{\mathrm{D}\,S_n}} \leqslant x \right) - \Phi(x) \right| \leqslant \tfrac{C}{\sqrt{n}} \cdot \tfrac{\mathrm{E}\,|\xi_k - \mathrm{E}\,\xi_k|^3}{\sigma^3},$$

*где $C$ — некоторая константа, удовлетворяющая неравенствам*

$$0.3989 \approx \tfrac{1}{\sqrt{2\pi}} \leqslant C < 0.4784.$$

**Пример.** Рассмотрим схему Бернулли, где

$$\nu_n = \sum_{i=1}^{n} \theta_i, \quad \mathrm{E}\,\nu_n = np, \quad \mathrm{D}\,\nu_n = npq, \quad \mathrm{E}\,\theta_1 = p, \quad \mathrm{D}\,\theta_1 = pq.$$

Тогда оценка Берри–Эссена будет иметь вид:

$$\sup_{x\in\mathbb{R}}\left|\mathrm{P}\left(\frac{\nu_n-np}{\sqrt{npq}}\leqslant x\right)-\Phi(x)\right|\leqslant\frac{C\cdot\mathrm{E}\,|\theta_1-p|^3}{(pq)^{3/2}\sqrt{n}}.$$

Случайная величина $|\theta_1-p|^3$ распределена следующим образом:

$$|\theta_1-p|^3\sim\begin{array}{cc}|0-p|^3 & |1-p|^3\\ 1-p & p\end{array}.$$

Отсюда вычислим математическое ожидание:

$$\mathrm{E}\,|\theta_1-p|^3=p^3(1-p)+(1-p)^3p=p^3q+q^3p=pq(p^2+q^2).$$

Подставим в правую часть неравенства:

$$\frac{C\cdot\mathrm{E}\,|\theta_1-p|^3}{(pq)^{3/2}\sqrt{n}}=\frac{C\cdot pq(p^2+q^2)}{(pq)^{3/2}\sqrt{n}}=\frac{C(p^2+q^2)}{\sqrt{npq}}.$$

В результате, оценка Берри–Эссена для схемы Бернулли будет иметь вид:

$$\sup_{x\in\mathbb{R}}\left|\mathrm{P}\left(\frac{\nu_n-np}{\sqrt{npq}}\leqslant x\right)-\Phi(x)\right|\leqslant\frac{C(p^2+q^2)}{\sqrt{npq}}.$$

При $p=q=\frac{1}{2}$ правая часть неравенства будет равна $\frac{\sqrt{C}}{\sqrt{n}}$. Если $p\gg q$ или $q\gg p$, то оценка будет работать хуже. ★

## 12.5 Теорема Муавра-Лапласа

**Теорема 12.18** (Формула Стирлинга)**.** *При $n\to+\infty$ имеется эквивалентность*

$$n!\sim\sqrt{2\pi n}\left(\tfrac{n}{e}\right)^n\cdot e^{\frac{\Theta}{12n}},$$

*где $0<\Theta<1$. Также $n\to+\infty$ имеется асимптотическое представление*

$$n!=\sqrt{2\pi n}\left(\tfrac{n}{e}\right)^n(1+o(1)).$$

**Теорема 12.19** (Локальная теорема Муавра-Лапласа)**.** *Пусть $\nu_n\sim\mathrm{B}(n,p)$ — случайная величина числа «успехов» в $n$ независимых испытаниях Бернулли с вероятностью «успеха» $p$ (другими словами, $\nu_n$ — биномиальная случайная величина). Тогда если равномерно по всех целым $k$ таких, что при $n\to+\infty$ имеется асимптотическая оценка*

$$|k-np|=o\left((npq)^{\frac{2}{3}}\right),$$

*то равномерно при $n\to+\infty$ имеет место асимптотическая оценка*

$$\mathrm{P}(\nu_n=k)\sim\tfrac{1}{\sqrt{2\pi npq}}e^{-\frac{(k-np)^2}{2npq}}.$$

*Доказательство.* Воспользуемся разложением из формулы Стирлинга:

$$n! = \sqrt{2\pi n}\left(\tfrac{n}{e}\right)^n(1+o(1)).$$

Тогда имеем

$$\mathrm{P}(\nu_n = k) = \tfrac{n!}{k!(n-k)!}p^k q^{n-k} = \tfrac{n^n}{k^k(n-k)^{n-k}} \cdot \sqrt{\tfrac{n}{2\pi k(n-k)}} \cdot e^{k+(n-k)-n} p^k q^{n-k} =$$

$$= \tfrac{1}{\sqrt{2\pi n \frac{k}{n}\cdot\frac{n-k}{n}}} \cdot \left(\tfrac{k}{n}\right)^{-k}\left(\tfrac{n-k}{n}\right)^{-(n-k)} p^k q^{n-k}.$$

Положим $k = np + \delta_n$, где $\delta_n = o\left((npq)^{\frac{2}{3}}\right)$ по условию. Отсюда $\frac{k}{n} = p + \frac{\delta_n}{n}$ и $\frac{n-k}{n} = q - \frac{\delta_n}{n}$. Тогда, продолжая цепочку равенств выше получаем

$$\tfrac{1}{\sqrt{2\pi n(p+\frac{\delta_n}{n})(q-\frac{\delta_n}{n})}} \cdot \left(1+\tfrac{\delta_n}{np}\right)^{-np-\delta_n}\left(1-\tfrac{\delta_n}{nq}\right)^{-nq+\delta_n}.$$

Для первого множителя имеем $\frac{1}{\sqrt{2\pi n(p+\frac{\delta_n}{n})(q-\frac{\delta_n}{n})}} \sim \frac{1}{\sqrt{2\pi n}}$ из асимптотической оценки на $\delta_n$. Оставшиеся множители перепишем в виде

$$\left(1+\tfrac{\delta_n}{np}\right)^{-np-\delta_n}\left(1-\tfrac{\delta_n}{nq}\right)^{-nq+\delta_n} = e^{(-np-\delta_n)\ln(1+\frac{\delta_n}{np})+(-nq+\delta_n)\ln(1-\frac{\delta_n}{nq})}.$$

Используем разложение в ряд Маклорена логарифма $\ln(1+x) = x - \frac{x^2}{2} + O(x^3)$ для оценки степени:

$$(-np-\delta_n)\ln(1+\tfrac{\delta_n}{np}) + (-nq+\delta_n)\ln(1-\tfrac{\delta_n}{nq}) =$$

$$= (-np-\delta_n)(\tfrac{\delta_n}{np} - \tfrac{\delta_n^2}{2(np)^2} + O(\tfrac{\delta_n^3}{n^3 3})) + (-nq+\delta_n)(-\tfrac{\delta_n}{nq} - \tfrac{\delta_n^2}{2(nq)^2} + O(\tfrac{\delta_n^3}{n^3})) =$$

$$= \tfrac{\delta_n^2}{n}(-\tfrac{1}{p} + \tfrac{1}{2p} - \tfrac{1}{q} + \tfrac{1}{2q}) + \tfrac{\delta^3}{2(np)^2} + O(\tfrac{\delta^3}{n^2}) + \tfrac{\delta^3}{2(nq)^2} + O(\tfrac{\delta^3}{n^2}) = -\tfrac{\delta^2(p+q)}{2npq} + O(\tfrac{\delta^3}{n^2}) =$$

$$= -\tfrac{\delta^2}{2npq} + o(1),$$

где в последнем переходе мы воспользовались тем, что $\delta^3 = o(n^2p^2q^2)$.

Из всего доказанного выше вытекает требуемая оценка. □

**Теорема 12.20** (Интегральная теорема Муавра-Лапласа)**.** *Пусть $\nu_n \sim \mathrm{B}(n,p)$ — случайная величина числа «успехов» в $n$ независимых испытаниях Бернулли с вероятностью «успеха» $p$ (другими словами, $\nu_n$ — биномиальная случайная величина). Тогда для всяких (возможно бесконечных) $-\infty \leqslant c < d \leqslant +\infty$ существует предел*

$$\lim_{n\to+\infty} \mathrm{P}(c < \tfrac{\nu_n - np}{\sqrt{npq}} \leqslant d) = \Phi(d) - \Phi(c),$$

*где $\Phi$ — функция распределения стандартного нормального распределения.*

*Доказательство.* В данном случае имеем $\mathrm{E}\,\nu_n = np$ и $\mathrm{D}\,\nu_n = npq$, поэтому для $c = -\infty$ формула выше совпадает с соответствующей формулой в центральной предельной теореме. Для произвольных $c$ воспользуемся равенством

$$\mathrm{P}(c < \tfrac{\nu_n - np}{\sqrt{npq}} \leqslant d) = \mathrm{P}(\tfrac{\nu_n - np}{\sqrt{npq}} \leqslant d) - \mathrm{P}(\tfrac{\nu_n - np}{\sqrt{npq}} \leqslant c).$$

$\square$

## 12.6 Закон повторного логарифма

**Теорема 12.21** (Закон повторного логарифма)**.** *Пусть $\{\xi_k\}$ — последовательность независимых случайных величины с распределением*

$$\xi_k \sim \begin{matrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{matrix}.$$

*Положим $S_n = \xi_1 + \ldots + \xi_n$. Тогда имеет место сходимость почти наверное*

$$\mathrm{P}(\varlimsup_{n\to+\infty} \tfrac{S_n}{\sqrt{2n \ln\ln n}} = 1) = 1,$$

$$\mathrm{P}(\varliminf_{n\to+\infty} \tfrac{S_n}{\sqrt{2n \ln\ln n}} = -1) = 1.$$

## 12.7 Закон арксинуса

**Теорема 12.22** (Закон арксинуса)**.** *Пусть $\{\xi_k\}$ — последовательность независимых случайных величины с распределением*

$$\xi_k \sim \begin{matrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{matrix}.$$

*Пусть $\nu_n$ — случайная величина, выражающая количество целых точек $k \in [1, n]$ таких, что $\xi_1 + \ldots + \xi_k > 0$. Тогда*

$$\lim_{n\to+\infty} \mathrm{P}(\tfrac{\nu_n}{n} \leqslant x) = \tfrac{1}{\pi} \int\limits_0^x \tfrac{\mathrm{d}y}{\sqrt{y(1-y)}} = \tfrac{2}{\pi} \arcsin\sqrt{x}.$$

**Теорема 12.23** (Усиленный закон арксинуса)**.** *Пусть $\{\xi_k\}$ — последовательность независимых случайных величины с распределением*

$$\xi_k \sim \begin{matrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{matrix}.$$

*Пусть $\nu_n$ — случайная величина, выражающая количество целых точек $k \in [1, n]$ таких, что $\xi_1 + \ldots + \xi_k > 0$. Тогда для $1 \leqslant k \leqslant n$ выполнено*

$$\mathrm{P}(\nu_{2n} = 2k \mid S_{2n} = 0) = \tfrac{1}{n+1}.$$

# 13 Указатель терминов

d

# 14 Указатель теорем

d

# Список литературы

[1] Колмогоров А. Н., Фомин С. В. *Элементы теории функций и функционального анализа*, Физматлит, 2004, 572с.

[2] Дьяченко М. И., Ульянов П. Л. *Мера и интеграл*, Факториал, 1998, 160с.

[3] Боровков А. А. *Теория вероятностей*, Физматлит, 1986, 432с.