# Зимний коллоквиум курса «Теория вероятностей»

ФКН НИУ ВШЭ, 2-й курс ОП ПМИ, 2-й модуль, 2020 учебный год

Сергей Пилипенко Атаев Азнарур Аня «10 за коллок» Смирнова

Дата последнего обновления: 04/04/2021 14:42.

*Спасибо Васильеву Демиду за исходники своих ответов на вопросы и Косову Е. Д. за исходники лекций.*

## Содержание

# Билет 1

*Теорема Пуассона. Задача про булочки с изюмом.*

## Теорема Пуассона

Рассмотрим серию испытаний по схеме Бернулли, причем пусть $N$-я серия состоит из $N$ испытаний и вероятность успеха в этой серии равна $p_N$. Потребуем, чтобы произведение $N \cdot p_N = \lambda$ не зависело от $N$. Нас интересует вероятность $P(S_N = k)$ наступления ровно $k$ успехов в $N$-ой серии.

**Теорема.** *Пусть $N \cdot p_N = \lambda$ — не зависит от $N$. Тогда*

$$P(S_N = k) = C_N^k p_N^k (1 - p_N)^{N-k} \to \frac{\lambda^k}{k!} e^{-\lambda}, \quad N \to +\infty.$$

*Доказательство.* Пусть $N \cdot p_N = \lambda$ — не зависит от $N$. Тогда

$$P(S_N = k) = C_N^k p_N^k (1 - p_N)^{N-k}.$$

Распишем вероятность $P(S_N = k)$ в следующем виде:

$$P(S_N = k) = \frac{N!}{k!(N-k)!} \cdot \frac{\lambda^k}{N^k} \cdot \left(1 - \frac{\lambda}{N}\right)^{N-k} = \frac{\lambda^k}{k!} \cdot \frac{N \cdot (N-1) \cdot \ldots \cdot (N-k+1)}{N^k} \cdot \left(1 - \frac{\lambda}{N}\right)^{-k} \cdot \left(1 - \frac{\lambda}{N}\right)^N.$$

Заметим следующие вещи:

- $\frac{N \cdot (N-1) \cdot \ldots \cdot (N-k+1)}{N^k} = \frac{N^k + o(N^k)}{N^k} \underset{N \to +\infty}{\to} 1;$
- $\left(1 - \frac{\lambda}{N}\right)^{-k} \underset{N \to +\infty}{\to} 1;$
- $\left(1 - \frac{\lambda}{N}\right)^{N} \underset{N \to +\infty}{\to} e^{-\lambda}.$

Учитывая, что $\lambda$ и $k$ не меняются, устремляем $N \to \infty$ и получаем

$$P(S_N = k) \underset{N \to +\infty}{\to} e^{-\lambda} \frac{\lambda^k}{k!}.$$

$\square$

## Задача про булочки с изюмом

**Формулировка** Какое в среднем количество изюма должны содержать булочки, для того чтобы вероятность иметь хотя бы одну изюминку в булочке была не меньше 0.99?

**Решение** Предположим, что уже изготовлено тесто на некоторое количество булочек. В это тесто добавлено $N$ изюминок так, что отношение числа изюминок к количеству булочек равно $\lambda$. Значит количество булочек $b = \frac{N}{\lambda}$.

Выделим в тесте кусок, из которого будет изготовлена данная булочка. Вероятность попадания одной изюминки в эту булочку равна $\frac{1}{b} = \frac{\lambda}{N}$, а вероятность того, что хотя бы одна изюминка попала в булку, равна $1 - P$(булочка без изюма) и равна

$$1 - \left(1 - \frac{\lambda}{N}\right)^N.$$

Поскольку мы рассматриваем серийное производство булочек, то можно предполагать, что $N \to +\infty$, т. е. растет объем теста и количество изюма, но не меняется плотность $\lambda$. Как и выше, получаем $\left(1 - \frac{\lambda}{N}\right)^N \to e^{-\lambda}$. Для решения задачи надо найти $\lambda$ такое, что $e^{-\lambda} < 0.01$. Подходит $\lambda = 5$, т. е. плотность изюма должна быть не менее пяти изюминок на булочку.

# Билет 2

*Модель Эрдеша-Реньи случайного графа. Теорема о надежности сети.*

## Модель Эрдеша-Реньи случайного графа

**Определение** (модель Эрдеша-Реньи)**.** Пусть $V_n$ — конечное множество $\{1, 2, \ldots, n\}$, элементы которого мы называем *вершинами*. Будем проводить между двумя различными вершинами ребро (только одно) с вероятностью $p$ независимо от остальных пар вершин. Получающийся граф будем называть *случайным графом в модели Эрдеша-Реньи*.

Множество элементарных исходов $\Omega$ состоит из $C_n^2$ ребер. Событием называется любое подмножество ребер в клике на $n$ вершинах $E \subseteq \Omega$. Вероятность $E$ задается формулой

$$P(E) = p^{|E|}(1-p)^{C_n^2 - |E|}.$$

## Теорема о надежности сети

**Теорема** (о надежности сети в общем случае)**.** *Если $p = \frac{c \ln n}{n}$, то при $c > 1$ вероятность того, что граф связен, стремится к 1 (граф почти всегда связен), а при $c < 1$ вероятность того, что граф связен, стремится к 0 (граф почти всегда не связен).*

**Теорема** (о надежности сети в частном случае)**.** *Если $p = \frac{c \ln n}{n}$ и $c > 2$, то граф почти всегда связен.*

*Доказательство.* Пусть случайная величина $X_n$ — число компонент связности в случайном графе $G$, если граф не является связным, и $X_n = 0$ в случае связности $G$. Нам надо доказать, что $P(X_n > 1) \to 0$ при $n \to \infty$. По неравенству Чебышева

$$P(X_n > 1) \leqslant \mathbb{E}X_n.$$

Следовательно, достаточно доказать стремление к нулю $\mathbb{E}X_n$. Пусть $K_1, \ldots, K_{C_n^k}$ — все $k$-элементные подмножества $V_n$. Через $X_{n,k,i}$ обозначим случайную величину, которая равна единице в случае, когда $K_i$ является компонентой связности, и равна нулю в случае, когда это не так. Ясно, что

$$\mathbb{E}X_n = \sum_{k=1}^{n-1} \sum_{i=1}^{C_n^k} \mathbb{E}X_{n,k,i}.$$

Заметим, что $\mathbb{E}X_{n,k,i} = P(X_{n,k,i} = 1)$, а эта вероятность оценивается[1] через вероятность того, что вершины из множества $K_i$ не соединены ребрами с вершинами из $V_n \setminus K_i$. Пусть $q = 1 - p$, тогда имеет место оценка[2]

$$\mathbb{E}X_n \leqslant \sum_{k=1}^{n-1} C_n^k q^{k(n-k)}.$$

Эта сумма симметрична и, удваивая ее, можно считать, что суммирование идет по $k \leqslant \frac{n}{2}$. При таких $k$ имеет место неравенство $k(n-k) \geqslant k(n - \frac{n}{2}) = \frac{kn}{2}$. Добавим и вычтем $1 + q^{n^2/2}$, чтобы можно было свернуть по формуле бинома Ньютона

$$\sum_{k=1}^{n-1} C_n^k q^{k(n-k)} \leqslant 2 \sum_{k=1}^{n-1} C_n^k (q^{n/2})^k = 2(1 + q^{n/2})^n - 2 - 2q^{n^2/2}.$$

По условию, $q = 1 - p = \frac{2a \ln n}{n}$, где $a > 1$. Имеем

$$q^{n/2} = e^{2^{-1} n \ln\left(1 - \frac{2a \ln n}{n}\right)} = e^{-a \ln n + \beta_n} = \frac{1}{n^a} e^{\beta_n}, \quad \beta_n \to 0.$$

Следовательно,

$$(1 + q^{n/2})^n = \left(1 + \frac{1}{n^a} e^{\beta_n}\right)^n \to 1,$$

и

$$2(1 + q^{n/2})^n - 2 - 2q^{n^2/2} \to 0.$$

Таким образом, $\mathbb{E}X_n \to 0$ и теорема доказана. □

[1] Мы говорим, что любая компонента связности на $k$ вершинах никак не соединена с оставшимися $n-k$ вершинами, однако не все графы, для которых это верно, являются компонентами связности.

[2] Выбрали вершину (всего $k$ штук) и удалили ребра из нее в оставшиеся $n-k$ вершин.

# Билет 3

*Вероятностное пространство в общем случае: алгебра и $\sigma$-алгебра подмножеств. Примеры $\sigma$-алгебр, $\sigma$-алгебра, порожденная системой подмножеств, борелевская $\sigma$-алгебра.*

## Вероятностное пространство в общем случае: алгебра и $\sigma$-алгебра подмножеств

**Определение.** Класс $\mathcal{A}_0$ подмножеств пространства $\Omega$ называется *алгеброй*, если

1. $\Omega, \varnothing \in \mathcal{A}_0$;
2. $A \in \mathcal{A}_0 \implies \Omega \setminus A \in \mathcal{A}_0$;
3. $A, B \in \mathcal{A}_0 \implies A \cap B, A \cup B \in \mathcal{A}_0$.

**Определение.** Класс $\mathcal{A}$ подмножеств пространства $\Omega$ называется *$\sigma$-алгеброй*, если

1. $\Omega, \varnothing \in \mathcal{A}$;
2. $A \in \mathcal{A} \implies \Omega \setminus A \in \mathcal{A}$;
3. $A_n \in \mathcal{A},\ \forall n \in \mathbb{N} \implies \bigcap\limits_{n=1}^{\infty} A_n, \bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{A}$.

Отметим, что в силу формул

$$\Omega \setminus \bigcup_{\alpha} A_\alpha = \bigcap_{\alpha} (\Omega \setminus A_\alpha)$$

и

$$\Omega \setminus \bigcap_{\alpha} A_\alpha = \bigcup_{\alpha} (\Omega \setminus A_\alpha),$$

в пункте 3 каждого определения достаточно проверять включение либо только для объединений, либо только для пересечений.

## Примеры $\sigma$-алгебр, $\sigma$-алгебра, порожденная системой подмножеств, борелевская $\sigma$-алгебра

**Примеры $\sigma$-алгебр** Множество всех подмножеств $2^\Omega$, $\{\varnothing, \Omega\}$, $\{\varnothing, B, \Omega \setminus B, \Omega\}$ являются $\sigma$ алгебрами.

**Примеры алгебр** Множество всех конечных объединений попарно непересекающихся промежутков $(a, b]$ на $\mathbb{R}$ является алгеброй, но не является $\sigma$ алгеброй, поскольку она не содержит одноточечные множества — пересечения счетного числа полуинтервалов.

**Определение.** Говорят, что $\sigma$ алгебра *порождена набором множеств $S$*, если эта $\sigma$ алгебра является наименьшей по включению среди всех $\sigma$-алгебр, которые содержат данный набор множеств $S$. Такую $\sigma$-алгебру обозначают $\sigma(S)$.

**Определение.** $\sigma$-алгебра называется *борелевской $\sigma$-алгеброй* $\mathcal{B}(\mathbb{R})$ подмножеств прямой $\mathbb{R}$, если она порождена всеми промежутками (отрезками, интервалами, лучами).

Несложно показать, что в определении не обязательно в качестве порождающего множества брать все промежутки. Например, можно ограничиться только отрезками или только интервалами или только лучами $(-\infty, c]$. Например, проверим, что $\mathcal{B}(\mathbb{R})$ порождена всеми лучами вида $(-\infty, c]$. Действительно, $(-\infty, c] \in \mathcal{B}(\mathbb{R})$, как счетное объединение промежутков вида $(-n, c]$, поэтому $\sigma(\{(-\infty, c]\}) \subset \mathcal{B}(\mathbb{R})$. С другой стороны $(a, b] = (-\infty, b] \setminus (-\infty, a]$, отрезки получаются счетным пересечением промежутков вида $(a - \frac{1}{n}, b]$, интервалы получаются счетным объединением промежутков вида $(a, b - \frac{1}{n}]$, а полуинтервалы вида $[a, b)$ получаются объединением уже полученных отрезков вида $[a, b - \frac{1}{n}]$. Тем самым, все промежутки принадлежат $\sigma(\{(-\infty, c]\})$, а значит имеет место и включение $\mathcal{B}(\mathbb{R}) \subset \sigma(\{(-\infty, c]\})$.

# Билет 4

*Аддитивные и счетно аддитивные функции множества на алгебрах и $\sigma$-алгебрах. Вероятностная мера и определение вероятностного пространства. Эквивалентность счетной аддитивности и непрерывности в нуле для неотрицательной аддитивной функции множества на алгебре. Свойства непрерывности вероятностной меры.*

## Аддитивные и счетно аддитивные функции множества на алгебрах и $\sigma$-алгебрах

**Определение 0.1.** Пусть $\mathcal{A}_0$ — алгебра множеств. Функция $P\colon \mathcal{A}_0 \to [0,1]$ называется *аддитивной*, если для произвольных $A, B \in \mathcal{A}_0$, $A \cap B = \varnothing$ выполнено

$$P(A \cup B) = P(A) + P(B).$$

**Определение 0.2.** Функция $P\colon \mathcal{A}_0 \to [0,1]$ называется *счетно аддитивной*, если для всякого не более чем счетного набора попарно непересекающихся событий $A_n \in \mathcal{A}_0$, для которых $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{A}_0$ выполняется

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

## Вероятностная мера и определение вероятностного пространства

**Определение 0.3.** Пусть $\mathcal{A} - \sigma$ алгебра. Функция $P\colon \mathcal{A} \to [0,1]$ называется *вероятностной мерой* на $\mathcal{A}$, если $P(\Omega) = 1$ и $P$ — счетно аддитивна на $\mathcal{A}$.

**Определение 0.4.** Пусть $\mathcal{A}$ — $\sigma$-алгебра подмножеств $\Omega$, тогда тройку $(\Omega, \mathcal{A}, P)$ называют *вероятностным пространством.*

## Эквивалентность счетной аддитивности и непрерывности в нуле для неотрицательной аддитивной функции множества на алгебре

**Предложение.** Пусть $P\colon \mathcal{A}_0 \to [0,1]$ — аддитивная функция множества на алгебре $\mathcal{A}_0$. Функция $P$ счетно аддитивна на $\mathcal{A}_0$ тогда и только тогда, когда для произвольного набора $A_n \in \mathcal{A}_0$, $A_{n+1} \subset A_n$, $\bigcap\limits_{n=1}^{\infty} A_n = \varnothing$ выполнено

$$\lim_{n\to\infty} P(A_n) = 0.$$

*Доказательство.*

$\Longrightarrow$ Пусть $P$ счетно аддитивна на $\mathcal{A}_0$. Рассмотрим множества $C_n = A_n \setminus A_{n+1}$. Тогда

$$A_1 = \bigcup_{n=1}^{\infty} C_n, \ldots, A_{N+1} = \bigcup_{k=N+1}^{\infty} C_k,$$

и

$$P(A_1) = \sum_{n=1}^{N} P(C_n) + P(A_{N+1}).$$

Если $P$ счетно аддитивна, то $\sum\limits_{n=1}^{N} P(C_n) \to P(A_1)$, а $P(A_{N+1}) \to 0$.

$\Longleftarrow$ Пусть $C_n \in \mathcal{A}_0$ — набор попарно непересекающихся множеств, причем известно, что $\bigcup\limits_{n=1}^{\infty} C_n = A_1 \in \mathcal{A}_0$. Пусть

$$A_{N+1} = \bigcup_{k=N+1}^{\infty} C_k,$$

тогда $A_{N+1} \subset A_N$, причем $\bigcap\limits_{N=1}^{\infty} A_N = \varnothing$. Если $P(A_{N+1}) \to 0$, то $P(A_1) = \sum\limits_{n=1}^{N} P(C_n) + P(A_{N+1})$ и переходя к пределу, получаем

$$P(A_1) = \sum_{n=1}^{\infty} P(C_n). \qquad \square$$

## Свойства непрерывности вероятностной меры

**Следствие** (непрерывность вероятностной меры)**.** Пусть $(\Omega, \mathcal{A}, P)$ — вероятностное пространство. Тогда

1. Если $A_n \in \mathcal{A}$, $A_{n+1} \subset A_n$ и $A = \bigcap\limits_{n=1}^{\infty} A_n$, то $\lim\limits_{n\to\infty} P(A_n) = P(A)$;

   *Доказательство.* Рассмотрим $A'_n = A_n \setminus A$. Очевидно, $\bigcap\limits_{i=1}^{\infty} A'_i = \varnothing \implies \lim\limits_{n\to\infty} P(A'_n) = 0$. В то же время $P(A_n) = P(A'_n \sqcup A) = P(A'_n) + P(A)$. Значит, $\lim\limits_{n\to\infty} P(A_n) = P(A)$ □

2. Если $A_n \in \mathcal{A}$, $A_n \subset A_{n+1}$ и $A = \bigcup\limits_{n=1}^{\infty} A_n$, то $\lim\limits_{n\to\infty} P(A_n) = P(A)$.

   *Доказательство.* Рассмотрим $A'_n = \Omega \setminus A_n$. Тогда $\bigcap\limits_{i=1}^{\infty} A'_n = \Omega \setminus A$. По п. 1:

$$1 - P(A) = P(\Omega \setminus A) = P\left(\bigcap_{i=1}^{\infty} A'_n\right) = \lim_{n\to\infty} P(A'_n) = \lim_{n\to\infty} P(\Omega \setminus A_n) = 1 - \lim_{n\to\infty} P(A_n)$$

□

В частности,

$$P\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \lim_{N\to\infty} P\Big(\bigcup_{n=1}^{N} A_n\Big) \leqslant \sum_{n=1}^{\infty} P(A_n).$$

# Билет 5

*Случайные величины на общих вероятностных пространствах: определение и основные свойства (прообраз борелевских лежит в $\sigma$-алгебре, сумма и произведение случайных величин — случайная, предел случайных — случайная)*

## Случайные величины на общих вероятностных пространствах

**Определение 0.5.** Пусть задано вероятностное пространство $(\Omega, \mathcal{A}, P)$. Функция $X\colon \Omega \to \mathbb{R}$ называется *случайной величиной*, если для всякого числа $t \in \mathbb{R}$ выполнено

$$X^{-1}((-\infty, t]) = \{\omega \in \Omega \mid X(\omega) \leqslant t\} \in \mathcal{A}.$$

**Предложение.** Если $X$ случайная величина, то $\{\omega \mid X(\omega) \in B\} = X^{-1}(B) \in \mathcal{A}$ для всякого $B \in \mathcal{B}(\mathbb{R})$.

*Доказательство.* Напомним следующие соотношения для прообраза функции:

$$X^{-1}\Big(\bigcap_{n=1}^{\infty} A_n\Big) = \bigcap_{n=1}^{\infty} X^{-1}(A_n), \quad X^{-1}\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \bigcup_{n=1}^{\infty} X^{-1}(A_n), \quad X^{-1}(\mathbb{R} \setminus B) = \Omega \setminus X^{-1}(B).$$

Рассмотрим систему множеств

$$\mathcal{C} := \{B \subset \mathbb{R} \mid X^{-1}(B) \in \mathcal{A}\}.$$

Эта система образует $\sigma$-алгебру. Действительно, $X^{-1}(\varnothing) = \varnothing \in \mathcal{A}$ и $X^{-1}(\mathbb{R}) = \Omega \in \mathcal{A}$. Если $B \in \mathcal{C}$, то $X^{-1}(\mathbb{R} \setminus B) = \Omega \setminus X^{-1}(B) \in \mathcal{A}$. Наконец, если $B_n \in \mathcal{C}$, то

$$X^{-1}\Big(\bigcap_{n=1}^{\infty} B_n\Big) = \bigcap_{n=1}^{\infty} X^{-1}(B_n) \in \mathcal{A}.$$

По условию $\sigma$-алгебра $\mathcal{C}$ содержит все лучи вида $(-\infty, t]$. Мы знаем, что $\mathcal{B}(\mathbb{R})$ — наименьшая по включению $\sigma$ алгебра, содержащая все лучи такого вида, поэтому $\mathcal{B}(\mathbb{R}) \subset \mathcal{C}$, что и требовалось. □

**Замечание.** Т. к. $\{X^2 \leqslant t\} = \{-\sqrt{t} \leqslant x \leqslant \sqrt{t}\}$ (при $t \geq 0$) и отрезок $[-\sqrt{t}, \sqrt{t}]$ — борелевское множество, получаем, что $X^2$ — также случайная величина. Можно проверить, что для случайной величины $X$ и для любой «разумной» функции $f\colon \mathbb{R} \to \mathbb{R}$ (например, если $f$ непрерывная), $f(X)$ также будет случайной величиной.

**Предложение.** Пусть $X, Y$ — случайные величины. Тогда случайными величинами будут $\alpha X + \beta Y$, $X \cdot Y$.

*Доказательство.* Ясно, что $\alpha X$ и $\beta Y$ — случайные величины. Проверим, что $X + Y$ — случайная величина:

$$\{X + Y > t\} = \{X > t - Y\} = \bigcup_{r_n \in \mathbb{Q}} (\{\omega \mid X(\omega) > r_n\} \cap \{\omega \mid r_n > t - Y(\omega)\}) \in \mathcal{A}.$$

В последнем переходе мы воспользовались тем, что $\mathbb{Q}$ всюду плотно в $\mathbb{R}$, поэтому между любыми двумя вещественными числами есть рациональное число. Поэтому и $\{X + Y \leqslant t\} \in \mathcal{A}$, а значит $X + Y$ — случайная величина. Для произведения заметим, что $X \cdot Y = \frac{1}{2}\big((X+Y)^2 - X^2 - Y^2\big)$, и утверждение следует из уже доказанных. □

**Предложение.** Пусть $X_n$ — случайные величины и для всякого $\omega$ существует предел $\lim_{n\to\infty} X_n(\omega) = X(\omega)$. Тогда $X$ является случайной величиной.

*Доказательство.* Рассмотрим множество $\{\omega\colon\ X(\omega) \leqslant t\}$. Заметим, что $X(\omega) \leqslant t$ тогда и только тогда, когда для каждого натурального числа $k$ найдётся такой номер $N$, что для всех $n > N$ верно неравенство $X_n(\omega) \leqslant t + \frac{1}{k}$. На языке теории множеств эту формулу фразу можно записать так

$$\{\omega\colon\ X(\omega) \leqslant t\} = \bigcap_{k} \bigcup_{N} \bigcap_{n>N} \{\omega\colon\ X(\omega) \leqslant t + \frac{1}{k}\}$$

Остаётся заметить, что $\{\omega\colon\ X(\omega) \leqslant t + \frac{1}{k}\} \in \mathcal{A}$ □

Таким образом, со случайными величинами можно выполнять арифметические операции и переходить к пределу.

# Билет 6

*Распределение случайной величины и функция распределения. Три основных свойства функции распределения. Формулировка теоремы о продолжении счетно аддитивной функции множества с алгебры на $\sigma$-алгебру, порожденную исходной алгеброй. Идея построения меры Лебега равномерного распределения на отрезке с борелевской $\sigma$-алгебой. Формулировка теоремы об однозначности задания распределения функцией распределения и о существовании распределения с заданой функцией распределения (т.е. с функцией, удовлетворяющей трем свойствам). Идея доказательства.*

## Распределение случайной величины и функция распределения

**Определение 0.6.** *Распределением* случайной величины $X$ называется вероятностная мера $\mu_X$ на $\mathcal{B}(\mathbb{R})$, определяемая равенством

$$\mu_X(B) = P(\{\omega \mid X(\omega) \in B\}) = P(X^{-1}(B)).$$

Обратим внимание, что, как и в дискретном случае, распределение случайной величины это мера на значениях случайной величины, т. е. мера $\mu_X$ показывает с какой вероятностью принимаются те или иные значения $X$.

**Определение 0.7.** Функция

$$F_X(t) = \mu_X((-\infty, t]) = P(\{\omega \mid X(\omega) \leqslant t\}).$$

называется *функцией распределения* случайной величины $X$.

Из определения $F_X$ следует, что $P(a < X \leqslant b) = \mu_X((a, b]) = F(b) - F(a)$.

## Три основных свойства функции распределения

**Предложение.** Функция $F_X$ удовлетворяет следующим свойствам:

1. $F_X\colon \mathbb{R} \to [0, 1]$ не убывает;
2. $F_X$ непрерывна справа;
3. $\lim\limits_{t\to-\infty} F_X(t) = 0$ и $\lim\limits_{t\to+\infty} F_X(t) = 1$.

*Доказательство.* Т.к. $\{\omega \mid X(\omega) \leqslant t\} \subset \{\omega \mid X(\omega) \leqslant s\}$ при $t \leqslant s$, то получаем свойство 1.
Обоснуем пункт 2. Пусть $t_n \to t$, $t_n \geqslant t$. Заметим, что

$$\{X \leqslant t\} = \bigcap_{k=1}^{\infty} \{X \leqslant t + \tfrac{1}{k}\}.$$

В силу непрерывности вероятностной меры $P$ получаем

$$\lim_{k\to\infty} F_X(t + \tfrac{1}{k}) = \lim_{k\to\infty} P(X \leqslant t + \tfrac{1}{k}) = P(X \leqslant t) = F_X(t).$$

Значит для каждого $\varepsilon > 0$ найдется $k$, для которого

$$F_X(t) \leqslant F_X(t + \tfrac{1}{k}) \leqslant F_X(t) + \varepsilon.$$

Т.к. $t_n \to t$, $t_n \geqslant t$, то найдется номер $n_0$, начиная с которого $t \leqslant t_n < t + \frac{1}{k}$. В силу монотонности $F_x(t) \leqslant F_X(t_n) \leqslant F_X(t + \frac{1}{k}) \leqslant F_X(t) + \varepsilon$ при $n \geqslant n_0$. Это и означает, что $\lim\limits_{n\to\infty} F_X(t_n) = F(t)$.

Свойство 3 обосновывается аналогично. □

## Формулировка теоремы о продолжении счетно аддитивной функции множества с алгебры на $\sigma$-алгебру, порожденную исходной алгеброй

**Теорема 0.1** (б/д)**.** *Пусть $\mathcal{A}_0$ есть некоторая алгебра подмножеств пространства $\Omega$ и пусть $P_0\colon \mathcal{A}_0 \to [0, 1]$ счетно аддитивная функция множества на алгебре $\mathcal{A}_0$. Тогда существует единственная вероятностная мера $P$ на $\sigma(\mathcal{A}_0)$, продолжающая функцию $P_0$, т.е. $P(A) = P_0(A)$ для произвольного множества $A \in \mathcal{A}_0$.*

## Идея построения меры Лебега равномерного распределения на отрезке с борелевской $\sigma$-алгебой

Мера Лебега — обычная длина, т. е. $\lambda([a, b]) = b - a$.

**Схема построения меры Лебега** Рассмотрим алгебру $\mathcal{A}_0$ конечных объединений попарно непересекающихся промежутков вида $(a,b] \subset [0,1]$ и возможно одноточечного множества $\{0\}$. Для множества $A = \bigsqcup_{j=1}^{m}(a_j,b_j]$ с попарно непересекающимися $(a_j,b_j]$ зададим меру Лебега равенством

$$\lambda(A) := \sum_{j=1}^{m}(b_j - a_j).$$

Нетрудно проверить, что это корректно определенная аддитивная функция множества на $\mathcal{A}_0$. Если теперь проверить, что она оказывается счетно аддитивной на этой алгебре (что верно), то по теореме о продолжении меры существует единственная вероятностная мера на $\mathcal{B}([0,1])$, совпадающая с $\lambda$ на $\mathcal{A}_0$.

## Формулировка теоремы об однозначности задания распределения функцией распределения и о существовании распределения с заданой функцией распределения

**Теорема 0.2.** *Распределение $\mu_X$ однозначно определяется функцией распределения $F_X$. Кроме того, если задана функция $F$, удовлетворяющая свойствам 1, 2, 3, то существует вероятностное пространство $(\Omega, \mathcal{A}, P)$ и случайная величина $X$ с функцией распределения $F$.*

Эта теорема позволяет говорить о распределении случайной величины без уточнения, на каком вероятностном пространстве задана случайная величина и как именно она задана.

**Идея доказательства** Наметим основные идеи доказательства. Первая часть является прямым следствием теоремы о продолжении меры. Пусть $A = \bigcup_{j=1}^{n}(a_j,b_j]$, причем $(a_j,b_j] \cap (a_k,b_k] = \varnothing$ при $j \neq k$. Тогда

$$\mu_X(A) = \sum_j F_X(b_j) - F_X(a_j).$$

Кроме того, множества $A$ указанного вида образуют алгебру $\mathcal{A}_0$ подмножеств $\mathbb{R}$. Поэтому, если есть две случайные величины с одной и той же функцией распределения, то по теореме о продолжении меры (часть о единственности продолжения) их распределения также совпадают на всех множествах из $\sigma(\mathcal{A}_0) = \mathbb{B}(\mathbb{R})$.

Доказательство второй части аналогично рассуждению о построении меры Лебега. Будем строить вероятностную меру $P$ на $\Omega = \mathbb{R}$ с $\mathcal{A} = \mathcal{B}(\mathbb{R})$. Рассмотрим алгебру $\mathcal{A}_0$ множеств вида $A = \bigcup_{j=1}^{n}(a_j,b_j]$, где $(a_j,b_j] \cap (a_k,b_k] = \varnothing$ при $j \neq k$. Для такого множества $A$ положим $P(A) := \sum_{j=1}^{n} F(b_j) - F(a_j)$. Нетрудно видеть, что корректно определена (т.е. для разных представлений $A$ равенство дает одно и тоже число) аддитивная функция множества на алгебре $\mathcal{A}_0$. Если теперь суметь проверить счетную аддитивность $P$ на $\mathcal{A}_0$, то $P$ продолжается до счетно аддитивной меры на $\sigma(\mathcal{A}_0) = \mathcal{B}(\mathbb{R})$. Если теперь рассмотреть случайную величину $X(\omega) = \omega$, то $F_X(t) = F(t)$ при $t \in \mathbb{R}$ в силу того, что $P((-\infty,t])$, являясь продолжением, стовпадает с $F(t) - F(-\infty) = F(t)$.

# Билет 7

*Дискретные и абсолютно непрерывные распределения случайных величин. Функция распределения дискретной случайной величины. Определение абсолютно непрерывного распределения случайной величины и определение плотности распределения. Основные свойства плотности и связь с функцией распределения. Примеры абсолютно непрерывных случайных величин.*

*Еще будут пояснения для таких как Аня Смирнова чтобы осознать.*

## Функция распределения дискретной случайной величины.

**Определение 0.8.** Случайная величина $\xi$ называется *дискретной*, если множество ее значений конечно или счетно. Если $x_1, \ldots, x_N, \ldots$ — различные значения $\xi$, то множества $A_i = \xi^{-1}\{x_i\}$ попарно не пересекаются. Пусть $p_i = P(A_i)$. Тогда распределение $\mu_\xi$ имеет вид

$$\mu_\xi = p_1\delta_{x_i} + \ldots + p_N\delta_{x_N} + \ldots$$

и полностью определяется значениями $x_i$ и $p_i$. В этой формуле $\delta_{x_i}(A) := 1$, если $x_i \in A$ и $\delta_{x_i}(A) = 0$, если $x_i \notin A$ для каждого $A \in \mathcal{B}(\mathbb{R})$.

*Вот это (мю кси) $\mu_\xi$ — вероятностная мера, а $\mathcal{B}(\mathbb{R})$ — борелевская сигма-алгебра.*

## Определение абсолютно непрерывного распределения случайной величины и определение плотности распределения

**Определение 0.9.** Говорят, что случайная величина $X$ имеет *абсолютно непрерывное* распределение (или является абсолютно непрерывной), если существует такая неотрицательная (и интегрируемая) функция $\rho_X$, что

$$F_X(t) = \int_{-\infty}^{t} \rho_X(x)\mathrm{d}x,$$

Функция $\rho_X$ называется *плотностью* случайной величины $X$.

*$F_X(t)$ — это функция распределения случ. величины, и она абсолютно непрерывна, если ее можно задать какой-то функцией $\rho_X$ и бахнуть интеграл, а так обычно определение другое.*

**Факты** Отметим, что в данном случае

$$\mu_X((a,b]) = F_X(b) - F_X(a) = \int_a^b \rho_X(x)\mathrm{d}x,$$

кроме того

$$P(X = a) = \lim_{n\to\infty}[F_X(a) - F_X(a - 1/n)] = 0 \text{ (непрерывность интеграла с переменным пределом).}$$

На самом деле можно доказать, что

$$\mu_X(A) = \int_A \rho_X(x)\mathrm{d}x$$

для всякого множества $A$, для которого имеет смысл интеграл в правой части, т. е. функция $I_A\rho_X$ интегрируема по Риману, где $I_A(x) = 1$ при $x \in A$ и $I_A(x) = 0$ при $x \notin A$.

## Основные свойства плотности и связь с функцией распределения.

**Предложение.** Отметим несколько свойств плотности распределения:

1. $\rho_X \geqslant 0$;
2. $\displaystyle\int\limits_{-\infty}^{+\infty} \rho_X(x)\mathrm{d}x = 1$;
3. $F'_X(x) = \rho_X(x)$ для любой точки непрерывности функции $\rho_X$.

Последнее свойство следует из теоремы о дифференцируемости интеграла с переменным верхнем пределом.

*Конечно же мы ее не помним: Пусть функция интегрируема на $[a,b]$ и непрерывна в точке $x_0 \in [a,b]$. Тогда функция $F$ дифференцируема в точке $x_0$ и $F'(x_0) = f(x_0)$.*

## Примеры абсолютно непрерывных случайных величин.

### Равномерное распределение

Случайная величина имеет *равномерное распределение на отрезке* $[a, b]$, если ее распределение задано плотностью

$$\rho(x) = \begin{cases} \frac{1}{b-a} & x \in (a, b], \\ 0 & x \notin (a, b]. \end{cases}$$

Такая случайная величина описывает случайное бросание точки в отрезок $[a, b]$. Вероятность того, что точка попадёт в отрезок $[c, d] \subset [a, b]$ равна $\frac{d-c}{b-a}$.

### Нормальное распределение

Случайная величина имеет *нормальное распределение* с параметрами $a$ и $\sigma^2$, если ее распределение задано плотностью

$$\rho(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

В случае $a = 0$ и $\sigma = 1$ эта плотность появлялась в теореме Муавра-Лапласа. *Помним этот гроб билет первого коллока.*

### Экспоненциальное (показательное) распределение

Случайная величина имеет *экспоненциальное распределение* (которое еще иногда называется показательным) с параметром $\lambda > 0$, если ее распределение задано плотностью

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

Функция распределения такой случайной величины имеет вид $F(t) = 1 - e^{-\lambda x}$.

# Билет 8

*Совместное распределение случайных величин, корректность определения. Функция совместного распределения и четыре ее основных свойства. Формулировка теоремы об однозначности задания совместного распределения функцией совместного распределения и о существовании распределения с заданой функцией совместного распределения (т.е. с функцией, удовлетворяющей четырем свойствам). Неоднозначность задания совместного распределения распределениями компонент (пример).*

## Совместное распределение случайных величин, корректность определения

**Определение 0.10.** Пусть $X$ и $Y$ — случайные величины. *Совместным распределением* случайных величин $X, Y$ называется вероятностная мера $\mu_{X,Y}$ на $\mathcal{B}(\mathbb{R}^2)$, определяемая следующим образом:

$$\mu_{X,Y}(B) = P(\{\omega \mid (X(\omega), Y(\omega)) \in B\}).$$

**Предложение.** Определение выше корректно в том смысле, что для $B \in \mathcal{B}(\mathbb{R}^2)$

$$\{\omega \mid (X(\omega), Y(\omega)) \in B\} \in \mathcal{A}.$$

*Доказательство.* Рассмотрим отображение $g\colon \Omega \to \mathbb{R}^2$, $g(\omega) = (X(\omega), Y(\omega))$. Аналогично тому, как мы уже делали, проверяется, что система множеств

$$\mathcal{C} := \{B \subset \mathbb{R}^2 \mid g^{-1}(B) \in \mathcal{A}\}$$

является $\sigma$-алгеброй. Заметим, что параллелепипеды $[a, b] \times [c, d] \in \mathcal{C}$, т.к.

$$g^{-1}([a, b] \times [c, d]) = \{\omega \mid X(\omega) \in [a, b], Y(\omega) \in [c, d]\} = \{\omega \mid X(\omega) \in [a, b]\} \cap \{\omega \mid Y(\omega) \in [c, d]\}.$$

Тем самым, $\mathcal{C}$ — некотрая $\sigma$-алгебра, содержащая все параллелепипеды, а $\mathcal{B}(\mathbb{R}^2)$ — это наименьшая по включению $\sigma$-алгебра, содержащая все параллелепипеды. □

## Функция совместного распределения и четыре ее основных свойства.

**Определение 0.11.** Функцию

$$F_{X,Y}(x, y) = P(\{\omega \mid X(\omega) \leqslant x,\ Y(\omega) \leqslant y\}) = \mu_{X,Y}((-\infty, x] \times (-\infty, y]).$$

называют *функцией совместного распределения* случайных величин $X$ и $Y$ или функцией распределения случайного вектора $(X, Y)$.

**Предложение.** Функция $F$ совместного распределения пары случайных величин удовлетворяет следующим свойствам:

1. $F\colon \mathbb{R}^2 \to [0, 1]$ и $F(b, d) - F(a, d) - F(b, c) + F(a, c) \geqslant 0$ для всякого прямоугольника $(a, b] \times (c, d]$;
2. $F$ непрерывна справа по совокупности переменных;
3. $\lim\limits_{(x,y)\to(u,v)} F(x, y) = 0$ если хотя бы одна из переменных $u$ или $v$ равна $-\infty$;
4. $\lim\limits_{(x,y)\to(+\infty,+\infty)} F(x, y) = 1$.

*Доказательство.* Доказательство повторяет рассуждения одномерного случая. Например, докажем (2). Заметим, что

$$\bigcap_{k\in\mathbb{N}} \{\omega \mid X(\omega) \leqslant x + \tfrac{1}{k}, Y(\omega) \leqslant y + \tfrac{1}{k}\} = \{\omega \mid X(\omega) \leqslant x, Y(\omega) \leqslant y\}.$$

Поэтому $P(X \leqslant x + \frac{1}{k}, Y \leqslant y + \frac{1}{k}) \to P(X \leqslant x, Y \leqslant y)$ и для каждого $\varepsilon$ найдется такое $k$, что

$$P(X \leqslant x + \tfrac{1}{k}, Y \leqslant y + \tfrac{1}{k}) < P(X \leqslant x, Y \leqslant y) + \varepsilon.$$

Если теперь $x_n \to x$, $x_n \geqslant x$, $y_n \to y$, $y_n \geqslant y$, то для произвольного $k$ найдется номер $n_0$, начиная с которого выполняется

$$x \leqslant x_n < x + \tfrac{1}{k},\ y \leqslant y_n < y + \tfrac{1}{k}.$$

Поэтому при $n > n_0$

$$P(X \leqslant x, Y \leqslant y) \leqslant P(X \leqslant x_n, Y \leqslant y_n) \leqslant P(X \leqslant x + \tfrac{1}{k}, Y \leqslant y + \tfrac{1}{k}) < P(X \leqslant x, Y \leqslant y) + \varepsilon.$$

Утверждения (3) и (4) обосновываются аналогично. □

## Формулировка теоремы об однозначности задания совместного распределения функцией совместного распределения и о существовании распределения с заданой функцией совместного распределения

**Теорема 0.3.** *Совместное распределение пары случайных величин $\mu_{X,Y}$ однозначно задается функцией совместного распределения $F_{X,Y}$. Кроме того, для всякой функции $F$, удовлетворяющей свойствам $(1),(2),(3),(4)$, существует вероятностное пространство $(\Omega, \mathcal{A}, P)$ и пара случайных величин $X, Y$ с функцией совместного распределения $F$.*

## Неоднозначность задания совместного распределения распределениями компонент

**Пример** Пусть в квадрате $[0,1]\times[0,1]$ случайно выбирается точка $(x,y)$. Случайные величины $X(x,y)=x$ и $Y(x,y)=y$ имеют равномерное распределение на $[0,1]$ и их совместное распределение является равномерным на $[0,1]\times[0,1]$, т. е. вероятность попадания в множество $B$ равна площади этого множества. Будем теперь выбирать точку $(x,y)$ случайным образом на диагонали квадрата $[0,1]\times[0,1]$, а случайные величины останутся прежними. Для всякого отрезка $[a,b]\subset[0,1]$ вероятность того, что $(x,y)\in[a,b]\times\mathbb{R}$ равна вероятности попасть в отрезок длины $(b-a)\sqrt{2}$ при бросании точки на отрезок длины $\sqrt{2}$, т. е. равна $b-a$. Таким образом, $X$ и $Y$ опять имеют равномерное распределение на $[0,1]$, но совместное распределение у них совсем другое.

# Билет 9

*Случайные векторы с абсолютно непрерывным распределением и плотность совместного распределения. Связь с функцией совместного распределения. Вычисление плотности компонент по совместной плотности. Плотность случайного вектора, являющегося функцией от другого случайного вектора. Равномерное распределение на многомерных областях.*

## Случайные векторы с абсолютно непрерывным распределением и плотность совместного распределения

**Определение 0.12.** Если существует такая интегрируемая и неотрицательная функция $\rho_{X,Y}(x,y)$, что

$$F_{X,Y}(x,y) = \iint\limits_{(-\infty,x]\times(-\infty,y]} \rho_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y,$$

то говорят, что совместное распределение случайных величин $X, Y$ *абсолютно непрерывно.* Функцию $\rho_{X,Y}$ называют *плотностью* совместного распределения случайных величин $X, Y$ (случайного вектора).

## Связь с функцией совместного распределения

Ясно, что

$$P(a < X \leqslant b,\ c < Y \leqslant d) = \mu_{X,Y}((a,b]\times(c,d]) = \iint\limits_{(a,b]\times(c,d]} \rho_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y.$$

Можно доказать, что

$$P((X,Y)\in B) = \mu_{X,Y}(B) = \iint\limits_{B} \rho_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y.$$

для всякого множества $A$, для которого имеет смысл интеграл Римана в правой части. В каждой точке непрерывности плотности $\rho_{X,Y}$ выполнено равенство

$$\frac{\partial^2}{\partial x\partial y}F_{X,Y}(x,y) = \rho_{X,Y}(x,y).$$

## Вычисление плотности компонент по совместной плотности

Если известна плотность $\rho_{X,Y}$ совместного распределения $X$ и $Y$, то можно найти плотности распределения каждой из случайных величин. Например, для случайной величины $X$:

$$F_X(t) = P(X \leqslant t,\ Y\in\mathbb{R}) = \int_{-\infty}^{t}\Big(\int_{-\infty}^{+\infty}\rho_{X,Y}(x,y)\mathrm{d}y\Big)\mathrm{d}x.$$

и, следовательно,

$$\rho_X(x) = \int_{-\infty}^{+\infty}\rho_{X,Y}(x,y)\mathrm{d}y.$$

Если распределение каждой из случайных величин задается плотностью, то совместное распределение может не иметь плотность.

## Плотность случайного вектора, являющегося функцией от другого случайного вектора

**Теорема 0.4.** *Пусть распределение $X, Y$ задано плотностью $\rho_{X,Y}$. Рассмотрим две случайные величины $\xi = f(X,Y), \eta = g(X,Y)$ и предположим, что отображение $T\colon (x,y)\mapsto (f(x,y), g(x,y))$ удовлетворяет условиям теоремы о замене переменных в кратном интеграле Римана (например непрерывно дифференцируемо с невырожденным якобианом). Тогда*

$$\rho_{\xi,\eta}(u,v) = \rho_{X,Y}\big(T^{-1}(u,v)\big)\cdot\big|J\big(T^{-1}(u,v)\big)\big|^{-1},$$

*где $J$ — якобиан отображения $T$.*

*Доказательство.* Заметим, что

$$P((\xi,\eta)\in A) = P((X,Y)\in T^{-1}(A)) = \iint\limits_{T^{-1}(A)}\rho_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y.$$

Сделаем замену в интеграле $u = f(x,y), v = g(x,y)$, т.е. $(x,y) = T^{-1}(u,v)$. Тогда последний интеграл равен

$$\iint\limits_{A}\rho_{X,Y}\big(T^{-1}(u,v)\big)\cdot\big|J\big(T^{-1}(u,v)\big)\big|^{-1}\mathrm{d}u\mathrm{d}v,$$

что завершает доказательство. □

## Равномерное распределение на многомерных областях

Говорят, что вектор $(\xi, \eta)$ *равномерно распределен* на множестве $B$, имеющем положительную площадь, если его распределение задано плотностью

$$\rho(x, y) = \begin{cases} \frac{1}{|B|}, & (x, y) \in B, \\ 0, & (x, y) \notin B. \end{cases}$$

Аналогичным образом определяется равномерно распределенный вектор с любым конечным числом координат.

# Билет 10

*Независимые случайные величины: характеризация в терминах функций распределения, в терминах совместного распределения, в терминах плотностей. Независимость функций от независимых случайных величин. Формула свертки для плотности суммы независимых случайных величин, заданных плотностями.*

## Независимые случайные величины: характеризация в терминах функций распределения, в терминах совместного распределения, в терминах плотностей

**Напоминание** Пусть задано вероятностное пространство $(\Omega, \mathcal{A}, P)$. Функция $X : \Omega \to \mathbb{R}$ называется случайной величиной, если для всякого числа $t \in \mathbb{R}$ выполнено

$$X^{-1}((-\infty, t]) = \{\omega \in \Omega : X(\omega) \leq t\} \in A$$

**Определение.** Случайные величины $X$ и $Y$ называются *независимыми*, если:

$$F_{X,Y}(x, y) = F_X(x)F_Y(y).$$

**Предложение.** Случайные величины $X$ и $Y$ независимы тогда и только тогда, когда для произвольных $U, V \in \mathcal{B}(\mathbb{R})$ выполнено:

$$P(\{\omega\colon X(\omega) \in U,\, Y(\omega) \in V\}) = P(\{\omega\colon X(\omega) \in U\}) \cdot P(\{\omega\colon Y(\omega) \in V\}.$$

*Доказательство.* Если $V = (-\infty, y]$, то две меры $U \to \frac{\mu_{X,Y}(U\times V)}{\mu_Y(V)}$ и $U \to \mu_X(U)$ совпадают на всех лучах $(-\infty, x]$, т.е. имеют одинаковые функции распределения, а значит совпадают на всех борелевских множествах $U$. Теперь для произвольного борелевского множества $U$ меры $V \to \frac{\mu_{X,Y}(U\times V)}{\mu_X(U)}$ и $V \to \mu_Y(V)$ совпадают на всех лучах $(-\infty, y]$, а значит и на всех борелевских множествах $V$. □

**Предложение.** Пусть распределения $X$ и $Y$ заданы плотностями. Тогда независимость $X$ и $Y$ равносильна тому, что совместное распределение задано плотностью и эта плотность имеет вид:

$$\rho_{X,Y}(x, y) = \rho_X(x)\rho_Y(y).$$

*Доказательство.* Если $X$ и $Y$ независимы, то

$$F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y) = \int_{-\infty}^{x} \rho_X(t)\,dt \cdot \int_{-\infty}^{y} \rho_Y(s)\,ds = \iint\limits_{(-\infty,x]\times(-\infty,y]} \rho_X(t)\rho_Y(s)\,dtds.$$

Обратно,

$$F_{X,Y}(x, y) = \iint\limits_{(-\infty,x]\times(-\infty,y]} \rho_X(t)\rho_Y(s)\,dtds = \int_{-\infty}^{x} \rho_X(t)\,dt \cdot \int_{-\infty}^{y} \rho_Y(s)\,ds = F_X(x) \cdot F_Y(y).$$

□

## Независимость функций от независимых случайных величин

**Определение.** Функция $f\colon \mathbb{R} \to \mathbb{R}$ называется борелевской, если $f^{-1}((-\infty, t]) \in \mathcal{B}(\mathbb{R})$ для каждого $t \in \mathbb{R}$.

Например, такими функциями будут все монотонные функции или все непрерывные.

**Следствие.** Пусть $X$ и $Y$ независимы, а $f, g$ — борелевские функции. Тогда $f(X)$ и $g(Y)$ также независимы.

## Формула свертки для плотности суммы независимых случайных величин, заданных плотностями

**Теорема** (Формула свертки)**.** *Предположим, что $X$ и $Y$ независимы и их распределения заданы плотностями $\rho_X$ и $\rho_Y$. Тогда распределение суммы $Z = X + Y$ задано плотностью*

$$\rho_Z(z) = \int_{-\infty}^{+\infty} \rho_X(t)\rho_Y(z - t)\mathrm{d}t.$$

*Доказательство.* По определению $F_Z(t) = P(\{\omega \mid X(\omega) + Y(\omega) \leqslant t\})$. С другой стороны, эта вероятность выражается через интеграл:

$$F_Z(t) = P(\{\omega \mid X(\omega) + Y(\omega) \leqslant t\}) = \iint\limits_{x+y\leqslant t} \rho_X(x)\rho_Y(y)\mathrm{d}x\mathrm{d}y.$$

Переходя к новым переменным $u = x + y$, $v = x$, и, применяя теорему Фубини[3], преобразуем этот интеграл:

$$F_Z(t) = \int_{-\infty}^{t} \Big( \int_{-\infty}^{+\infty} \rho_X(v) \rho_Y(u - v) \mathrm{d}v \Big) \mathrm{d}u.$$

Следовательно, распределение $Z$ имеет плотность требуемого вида. □

[3]Также известная, как сведение двойного интеграла к повторному.

# Билет 11

*Математическое ожидание в общем случае: построение математического ожидания для ограниченных, неотрицательных и общих случайных величин. Корректность определений и основные свойства (линейность, монотонность, равенство нулю неотрицательной случайной величины с нулевым ожиданием, ожидание модуля случайной величины).*

## Математическое ожидание в общем случае

**Определение.** Случайные величины с конечным числом значений будем называть *простыми.*

**Определение.** Пусть $X$ — простая случайная величина на $(\Omega, \mathcal{A}, P)$, принимающая конечное число значений $\{x_1, \ldots, x_N\}$. Тогда по определению полагаем, что $\mathbb{E}X := \sum\limits_{j=1}^{N} x_j P(X = x_j)$.

**Определение.** Пусть $X$ — ограниченная случайная величина, тогда ее *математическим ожиданием* называют предел $\lim\limits_{n\to\infty} \mathbb{E}X_n$ математических ожиданий произвольной последовательности простых случайных величин $X_n$, равномерно сходящейся к $X$.

**Определение.** Пусть $X \geqslant 0$ — неотрицательная случайная величина. Скажем, что у нее есть конечное математическое ожидание, если конечен следующий супремум

$$\mathbb{E}X := \sup\{\mathbb{E}U \mid 0 \leqslant U \leqslant X;\ U\text{— ограниченная}\}.$$

**Определение.** Пусть $X$ — случайная величина и пусть $X^+ := \max\{X, 0\} \geqslant 0$, $X^- := \max\{-X, 0\} \geqslant 0$ (в частности, $X = X^+ - X^-$). Скажем, что $X$ обладает математическим ожиданием, если $X^+$ и $X^-$ имеют конечные математические ожидания. В этом случае определим математическое ожидание $X$ равенством:

$$\mathbb{E}X := \mathbb{E}X^+ - \mathbb{E}X^-.$$

## Корректность определений и основные свойства

**Лемма.** *Пусть $X$ — ограниченная случайная величина. Тогда найдется последовательность простых случайных величин $X_n$, равномерно сходящаяся к $X$.*

*Доказательство.* Пусть $|X(\omega)| < R$ для каждого $\omega \in \Omega$. Рассмотрим случайную величину[4]

$$X_n(\omega) := \sum_{j=1}^{n} (-R + \tfrac{2R}{n}(k-1)) I_{\{\omega | -R + \frac{2R}{n}(k-1) \leqslant X(\omega) < -R + \frac{2R}{n}k\}}.$$

Возьмем теперь произвольный элемент $\omega_0 \in \Omega$. Тогда, т.к. $|X(\omega_0)| < R$, то найдется[5] число $k_0 \in \{1, \ldots, n\}$, для которого $-R + \frac{2R}{n}(k_0 - 1) \leqslant X(\omega_0) < -R + \frac{2R}{n}k_0$. Таким образом

$$|X(\omega_0) - X_n(\omega_0)| \leqslant \frac{2R}{n} = \varepsilon.$$

□

**Предложение.** Определение матожидания неотрицательной случайной величины корректно в том смысле, что для произвольной ограниченной случайной величины $X$ и для произвольной последовательности простых случайных величин $X_n$, равномерно сходящейся к $X$, существует предел $\lim\limits_{n\to\infty} \mathbb{E}X_n$. Кроме того, для произвольной другой последовательности простых случайных величин $Y_n$, равномерно сходящейся к $X$, выполнено

$$\lim_{n\to\infty} \mathbb{E}Y_n = \lim_{n\to\infty} \mathbb{E}X_n.$$

*Доказательство.* Заметим что,

$$|\mathbb{E}X_n - \mathbb{E}X_k| \leqslant \mathbb{E}|X_n - X_k| \leqslant \sup_{\omega\in\Omega} |X_n(\omega) - X_k(\omega)| \leqslant \sup_{\omega\in\Omega} |X_n(\omega) - X(\omega)| + \sup_{\omega\in\Omega} |X(\omega) - X_k(\omega)|.$$

Тем самым, последовательность $\{\mathbb{E}X_n\}$ фундаментальна, а значит сходится. Если $Y_n$ другая последовательность простых случайных величин, равномерно сходящаяся к $X$, то последовательность $Z_m$, для которой $Z_{2k-1} := X_k$, $Z_{2k} := Y_k$, также образует последовательность простых случайных величин, равномерно сходящуюся к $X$. Тогда последовательность чисел $\mathbb{E}Z_m$ сходится, а значит $\lim\limits_{n\to\infty} \mathbb{E}Y_n = \lim\limits_{n\to\infty} \mathbb{E}X_n$, как пределы двух подпоследовательностей сходящейся последовательности чисел. □

[4] Мы поделили $[-R, R - \frac{2R}{n}]$ на $n$ частей и представили $X$ в виде взвешенной суммы индикаторов попадания в подотрезки.
[5] Тут мы пользуемся «взвешеностью» с.в., чтобы получить ограничение на область ее значений.

**Предложение.** Определение матожидания произвольной сулчайной величины корректно в слудющем смысле. Предположим, что $U \geqslant 0, V \geqslant 0$ — случайные величины с конечными математическими ожиданиеми, причём $X = U - V$. Тогда $\mathbb{E}X = \mathbb{E}U - \mathbb{E}V$.

*Доказательство.* Действительно, в этом случае $X^+ - X^- = U - V$, т.е. $X^+ + V = U + X^-$, откуда $\mathbb{E}(X^+ + V) = \mathbb{E}(U + X^-)$. В силу того, что все функции $U, V, X^+, X^-$ неотрицательны, получаем, что $\mathbb{E}X^+ + \mathbb{E}V = \mathbb{E}U + \mathbb{E}X^-$, т.е. $\mathbb{E}X^+ - \mathbb{E}X^- = \mathbb{E}U - \mathbb{E}V$.

Из определения в частности следует, что для случайной величины $X$, обладающей математическим ожиданием, $|X| = X^+ + X^-$ также будет иметь конечное математическое ожидание. Наоборот, если $|X|$ обладает конечным математическим ожиданием, то $X^+ \leqslant |X|, X^- \leqslant |X|$, поэтому $X^+$ и $X^-$ имеют конечные математические ожидания, а значит и у $X$ определено математическое ожидание. □

**Предложение.** Для ограниченных случайных величин $X, Y$ выполнены свойства:

1. $\mathbb{E}(\alpha X + \beta Y) = \alpha\mathbb{E}X + \beta\mathbb{E}Y$;

2. Если $X \geqslant 0$ п.н., то $\mathbb{E}X \geqslant 0$, в частности, если $X \geqslant Y$ п.н., то $\mathbb{E}X \geqslant \mathbb{E}Y$.

*Доказательство.*

1. Если $X_n \rightrightarrows X$, $Y_n \rightrightarrows Y$, $X_n, Y_n$ — простые, то $\alpha X_n + \beta Y_n \rightrightarrows \alpha X + \beta Y$. Отсюда
$$\mathbb{E}(\alpha X + \beta Y) = \lim_{n\to\infty} \mathbb{E}(\alpha X_n + \beta Y_n) = \lim_{n\to\infty} (\alpha\mathbb{E}X_n + \beta\mathbb{E}Y_n) = \alpha\mathbb{E}X + \beta\mathbb{E}Y.$$

2. Пусть сначала $X_n(\omega) \geqslant 0$ для каждой $\omega \in \Omega$. Пусть $X_n$ — последовательность простых случайных величн, равномерно сходящаяся к $\sqrt{X}$. Тогда $X_n^2 \rightrightarrows X$, откуда $\mathbb{E}X = \lim\limits_{n\to\infty} \mathbb{E}X_n^2 \geqslant 0$. Теперь, для произвольного $X \geqslant 0$ п.н., выполнено $\mathbb{E}X = \mathbb{E}[XI_{\{x\geqslant 0\}}] + \mathbb{E}[XI_{\{X<0\}}]$. Покажем, что $\mathbb{E}[XI_{\{X<0\}}] = 0$. По доказанному $\mathbb{E}[-XI_{\{X<0\}}] \geqslant 0$ и $\mathbb{E}[(M + X)I_{\{X<0\}}] \geqslant 0$, где $M = \sup\limits_{\omega\in\Omega} |X(\omega)|$. Таким образом, $0 \leqslant \mathbb{E}[-XI_{\{X<0\}}] \leqslant MP(X < 0) = 0$. В случае, когда $X \geqslant Y$ п.н., получаем, что $X - Y \geqslant 0$ п.н. и $\mathbb{E}X - \mathbb{E}Y = \mathbb{E}(X - Y) \geqslant 0$.

□

**Предложение.** Для неотрицательных случайных величин $X, Y$ выполнены свойства:

1. $\mathbb{E}(\alpha X + \beta Y) = \alpha\mathbb{E}X + \beta\mathbb{E}Y$, если $\alpha, \beta \geqslant 0$;

2. Если $X \geqslant Y \geqslant 0$, то $\mathbb{E}X \geqslant \mathbb{E}Y$, в частности, если $X$ имеет конечное математическое ожидание, то и $Y$ также имеет конечное математическое ожидание;

3. Если $X = 0$ п.н., то $\mathbb{E}X = 0$.

*Доказательство.*

1. Достаточно доказать утверждение при $\alpha = \beta = 1$. Если $0 \leqslant U \leqslant X$, $0 \leqslant V \leqslant Y$, $U, V$ — ограниченные, то $U + V \leqslant X + Y$, откуда $\mathbb{E}X + \mathbb{E}Y \leqslant \mathbb{E}(X + Y)$. Наоборот, пусть $0 \leqslant Z \leqslant X + Y$, $Z$ — ограниченная случайная величина. Пусть $U := \min(X, Z)$, $V := Z - U$. Тогда $0 \leqslant U \leqslant X$, $U$ — ограниченная, $V = (Z - X)I_{\{X<Z\}} \leqslant X + Y - X = Y$, $V$ — ограниченная. Таким образом, $\mathbb{E}Z = \mathbb{E}(U + V) = \mathbb{E}U + \mathbb{E}V \leqslant \mathbb{E}X + \mathbb{E}Y$. А значит аналогичная оценка верна и для $\mathbb{E}(X + Y)$.

2. Следует из определния;

3. Произвольная ограниченная случайная величина $U$, $0 \leq U \leq X$, также обращается в нуль п.н. Поэтому $\mathbb{E}U = \mathbb{E}[UI_{\{U\neq 0\}}] \leqslant [\sup U]P(U \neq 0) = 0$.

□

**Предложение.** Для случайных величин $X, Y$, обладающих математическим ожиданием, выполнены свойства:

1. $\mathbb{E}(\alpha X + \beta Y) = \alpha\mathbb{E}X + \beta\mathbb{E}Y$;

2. Если $X \geqslant 0$ почти наверное, то $\mathbb{E}X \geq 0$, в частности, если $X \geqslant Y$, то $\mathbb{E}X \geqslant \mathbb{E}Y$;

3. Если $X \geqslant 0$ почти наверное и $\mathbb{E}X = 0$, то $X = 0$ почти наверное;

4. $|\mathbb{E}X| \leqslant \mathbb{E}|X|$.

*Доказательство.*

1. Заметим, что $\mathbb{E}[-X] = -\mathbb{E}X$, т.к. для произвольного представления $X = U - V$, $U, V \geqslant 0$, выполнено $-X = V - U$, откуда $\mathbb{E}[-X] = \mathbb{E}V - \mathbb{E}U = -(\mathbb{E}U - \mathbb{E}V) = -\mathbb{E}X$. Тогда достаточно доказать линейность только в случае $\alpha, \beta \geqslant 0$. В этом случае $\alpha X + \beta Y = \alpha X^+ + \beta Y^+ - (\alpha X^- + \beta Y^-)$, причем $\alpha X^+ + \beta Y^+ \geqslant 0$ и $\alpha X^- + \beta Y^- \geqslant 0$. Поэтому

$$\begin{aligned}\mathbb{E}(\alpha X + \beta Y) = \mathbb{E}(\alpha X^+ + \beta Y^+) - \mathbb{E}(\alpha X^- + \beta Y^-) = \alpha \mathbb{E}X^+ + \beta \mathbb{E}Y^+ - \alpha \mathbb{E}X^- - \beta \mathbb{E}Y^- \\ = \alpha(\mathbb{E}X^+ - \mathbb{E}X^-) + \beta(\mathbb{E}Y^+ - \mathbb{E}\beta Y^-) = \alpha \mathbb{E}X + \beta \mathbb{E}Y.\end{aligned}$$

2. В этом случае $X^- = 0$ п.в., а значит и $\mathbb{E}X^- = 0$. Таким образом, $\mathbb{E}X = \mathbb{E}X^+ \geqslant 0$.

3. Заметим, что $k^{-1}P(X \geqslant k^{-1}) = \mathbb{E}[k^{-1}I_{\{X \geqslant k^{-1}\}}] \leqslant \mathbb{E}X = 0$, откуда получаем, что $P(X > 0) = P(\bigcup\limits_{k=1}^{\infty}\{X \geqslant k^{-1}\}) \leqslant \sum\limits_{k=1}^{\infty} P(X \geqslant k^{-1}) = 0$.

4. Заметим, что $-|X| \leqslant X \leqslant |X|$, откуда по свойствам (2) и (1) получаем, неравенства $-\mathbb{E}|X| \leqslant \mathbb{E}X \leqslant \mathbb{E}|X|$.

$\square$

# Билет 12

*Математическое ожидание функции от случайной величины с абсолютно непрерывным распределением. Математическое ожидание произведения независимых случайных величин. Дисперсия, ковариация, коэффициент корреляции и их свойства. Неравенство Чебышева.*

## Математическое ожидание функции от случайной величины с абсолютно непрерывным распределением

**Лемма 0.1.** *Пусть случайная величина $X \geqslant 0$ и пусть $A_n \subset A_{n+1}$, $A_n \in \mathcal{A}$, причем $\bigcup_{n=1}^{\infty} A_n = \Omega$. Тогда $X$ имеет конечное математическое ожидание тогда и только тогда, когда*

$$\sup_n \mathbb{E}[XI_{A_n}] := M < \infty.$$

*При этом $\mathbb{E}X = \lim_{n\to\infty} \mathbb{E}[XI_{A_n}] = M$.*

*Доказательство.* Пусть $U$ — произвольная ограниченная случайная величина, причем $0 \leqslant U \leqslant X$. Тогда $P(\Omega \backslash A_n) \to 0$ (из-за свойства непрерывности вероятностной меры) и $\mathbb{E}[UI_{\Omega\backslash A_n}] \leqslant [\max U]P(\Omega \setminus A_n) \to 0$. Отсюда

$$\mathbb{E}U = \lim_{n\to\infty} \mathbb{E}[UI_{A_n}] \leqslant M,$$

т. к. имеет место оценка $\mathbb{E}[UI_{A_n}] \leqslant \mathbb{E}[XI_{A_n}] \leqslant M$. Значит $\mathbb{E}X \leqslant M$. С другой стороны $X \geqslant XI_{A_n}$, откуда $\mathbb{E}X \geqslant \mathbb{E}[XI_{A_n}]$. $\square$

**Предложение.** Пусть $X$ — случайная величина, распределение которой имеет плотность $\rho_X$. Пусть задана непрерывная функция $f$. Тогда математическое ожидание $\mathbb{E}f(X)$ существует тогда и только тогда, когда сходится несобственный интеграл

$$\int_{-\infty}^{+\infty} |f(x)|\rho_X(x)\mathrm{d}x.$$

Более того, в случае сходимости

$$\mathbb{E}f(X) = \int_{-\infty}^{+\infty} f(x)\rho_X(x)\mathrm{d}x.$$

*Доказательство.* Пусть $R > 0$. Для непрерывной функции $f$ на отрезке $[-R, R]$ найдется последовательность ступенчатых функций $g_n$, равномерно сходящаяся к $f$ на $[-R, R]$. Функции $g_n$ имеют вид $g_n = \sum_{j=1}^{N_n} c_j I_{[a_j, b_j]}$, где $\{[a_j, b_j]\}$ — разбиение отрезка $[-R, R]$. Заметим, что

$$\mathbb{E}g_n(X) = \sum_{j=1}^{N_n} c_j \mathbb{E}I_{\{a_j \leqslant X \leqslant b_j\}} = \sum_{j=1}^{N_n} c_j P(X \in [a_j, b_j]) = \sum_{j=1}^{N_n} c_j \int_{a_j}^{b_j} \rho_X(x)\mathrm{d}x = \int_{-R}^{R} g_n(x)\rho_X(x)\mathrm{d}x.$$

Заметим, что

$$\left|\mathbb{E}[f(X)I_{\{-R\leqslant X\leqslant R\}}] - \mathbb{E}g_n(X)\right| \leqslant \mathbb{E}[|f(X) - g_n(X)|I_{\{-R\leqslant X\leqslant R\}}] \leqslant \sup_{x\in[-R,R]} |f(x) - g_n(x)| \to 0.$$

Аналогично $\int_{-R}^{R} g_n(x)\rho_X(x)\mathrm{d}x \to \int_{-R}^{R} f(x)\rho_X(x)\mathrm{d}x$, откуда получаем равенство

$$\mathbb{E}[f(X)I_{\{-R\leqslant X\leqslant R\}}] = \int_{-R}^{R} f(x)\rho_X(x)\mathrm{d}x.$$

Достаточно доказать исходное утверждение для неотрицательных функций $f$, для которых оно теперь следует из леммы 0.1. $\square$

## Математическое ожидание произведения независимых случайных величин

**Предложение.** Пусть $X, Y$ — независимые случайные величины, имеющие математическое ожидание. Тогда $X \cdot Y$ также обладает математическим ожиданием и

$$\mathbb{E}[X \cdot Y] = [\mathbb{E}X] \cdot [\mathbb{E}Y].$$

*Доказательство.* Заметим, что

$$X \cdot Y = (X^+ - X^-)(Y^+ - Y^-) = X^+Y^+ + X^-Y^- - X^+Y^- - X^-Y^+.$$

Таким образом, достаточно доказать утверждение только для неотрицательных $X, Y$. Если $X, Y$ ограничены, $|X| < R$, $|Y| < R$, то рассмотрим

$$X_n(\omega) := \sum_{j=1}^{n} (-R + \tfrac{2R}{n}(k-1)) I_{\{\omega | -R + \frac{2R}{n}(k-1) \leqslant X(\omega) < -R + \frac{2R}{n}k\}},$$

$$Y_n(\omega) := \sum_{j=1}^{n} (-R + \tfrac{2R}{n}(k-1)) I_{\{\omega | -R + \frac{2R}{n}(k-1) \leqslant Y(\omega) < -R + \frac{2R}{n}k\}}.$$

Т.к. $X_n$ имеет вид $f_n(X)$, а $Y_n$ имеет вид $f_n(Y)$ для некоторой функции $f_n$, то $X_n$ и $Y_n$ также независимы. Кроме того, $X_n \rightrightarrows X$, $Y_n \rightrightarrows Y$. Поэтому

$$[\mathbb{E}X] \cdot [\mathbb{E}Y] = \lim_{n\to\infty} [\mathbb{E}X_n] \cdot [\mathbb{E}Y_n] = \lim_{n\to\infty} \mathbb{E}[X_n \cdot Y_n] = \mathbb{E}[X \cdot Y].$$

Для общих неотрицательных независимых $X$ и $Y$, рассмотрим независимые ограниченные случайные величины $XI_{\{|X|<R\}}$ и $YI_{\{|Y|<R\}}$. Тогда

$$\mathbb{E}[XI_{\{|X|<R\}} \cdot YI_{\{|Y|<R\}}] = \mathbb{E}[XI_{\{|X|<R\}}] \cdot \mathbb{E}[YI_{\{|Y|<R\}}].$$

Утверждение теперь следует из леммы 0.1. □

## Дисперсия, ковариация, коэффициент корреляции и их свойства

**Определение.** *Дисперсией* случайной величины $X$ называется число $\mathbb{D}X = \mathbb{E}(X - \mathbb{E}X)^2$.

**Определение.** *Ковариацией* пары случайных величин $X, Y$ называется число $\text{cov}(X, Y) = \mathbb{E}\left((X - \mathbb{E}X)(Y - \mathbb{E}Y)\right)$.

**Определение.** *Коэффициентом корреляции* называется величина $\rho(X, Y) = \dfrac{\text{cov}(X,Y)}{\sqrt{\mathbb{D}X}\sqrt{\mathbb{D}Y}}$.

**Теорема** (Свойства дисперсии)**.**

1. *Если $\mathbb{D}X = 0$, то $X = \mathbb{E}X$ почти наверное;*
2. *Для произвольных чисел $\alpha, \beta$ верно $\mathbb{D}(\alpha X + \beta) = \alpha^2 \mathbb{D}X$;*
3. *Если $X$ и $Y$ независимы, то $\text{cov}(X, Y) = 0$ и $\mathbb{D}(X + Y) = \mathbb{D}X + \mathbb{D}Y$.*

*Доказательство.*

1. Исходя из свойства (3) мат. ожидания неотрицательной случайной величины с нулевым мат. ожиданием, $(X - \mathbb{E}X)^2 = 0 \implies X = \mathbb{E}X$ почти наверное;
2. Исходит из линейности математического ожидания;
3. Так как $\text{cov}(X, Y) = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y)$, то из независимости $X$ и $Y$ следует $\text{cov}(X, Y) = (\mathbb{E}X)(\mathbb{E}Y) - (\mathbb{E}X)(\mathbb{E}Y) = 0$. Поэтому, $\mathbb{D}(X + Y) = \mathbb{D}X + 2\text{cov}(X, Y) + \mathbb{D}Y = \mathbb{D}X + \mathbb{D}Y$. □

## Неравенство Чебышева

**Предложение.** Пусть у неотрицательной случайной величины $X$ определено математическое ожидание. Тогда $P(X \geqslant t) \leqslant \frac{\mathbb{E}X}{t}$ для каждого $t > 0$. Пусть у случайной величины $X$ конечный второй момент, т.е. $\mathbb{E}X^2 < \infty$. Тогда

$$P(|X - \mathbb{E}X| \geqslant \varepsilon) \leqslant \varepsilon^{-2}\mathbb{D}X.$$

*Доказательство.* Заметим, что $t \cdot I_{\{X \geqslant t\}} \leqslant X$, поэтому

$$tP(X \geqslant t) = \mathbb{E}[t \cdot I_{\{X \geqslant t\}}] \leqslant \mathbb{E}X.$$

Второе неравенство обосновывается рассмотрением случайной величины $|X - \mathbb{E}X|^2$ и применением первого неравенства. □