# Теория вероятностей и математическая статистика, Коллоквиум III

Версия от 05.04.2021 00:39

## Содержание

# 1. Неравенство Чебышева и закон больших чисел в слабой форме для общих случайных величин. Усиленный закон больших чисел Колмогорова (б/д). Сходимости случайных величин: почти наверное и по вероятности. Взаимосвязь сходимостей по вероятности и почти наверное.

## 1.1. Неравенство Чебышева и закон больших чисел в слабой форме для общих случайных величин.

**Теорема** (Неравенство Маркова)**.** Пусть $X$ это случайная величина и $X \geqslant 0$ почти наверное. Тогда для любого $t > 0$ выполняется

$$\mathrm{P}[X \geqslant t] \leqslant \frac{\mathrm{E}[X]}{t}.$$

*Доказательство.* Заметим, что для любого $t > 0$ выполняется $t \cdot \mathrm{I}[x \geqslant t] \leqslant X$ почти наверное (здесь I это индикатор), так как в левой части будут учтены $t \leqslant X$, с суммарным коэффициентом не больше 1.
Возьмем математическое ожидание от обеих сторон и получим то, что нас просили:

$$t \cdot \mathrm{I}[x \geqslant t] \leqslant X \iff t \cdot \mathrm{P}[x \geqslant t] \leqslant \mathrm{E}[X] \iff \mathrm{P}[x \geqslant t] \leqslant \frac{\mathrm{E}[X]}{t}.$$

■

**Теорема** (Неравенство Чебышева)**.** Пусть у случайной величины $X$ конечный второй момент, то есть $\mathrm{E}[X^2] \leqslant \infty$. Тогда

$$\mathrm{P}[|X - \mathrm{E}[X]| \geqslant \varepsilon] \leqslant \frac{\mathrm{D}[X]}{\varepsilon^2}.$$

*Доказательство.* Для доказательства рассмотрим случайную величину $Y = |X - \mathrm{E}[X]|^2$ и применим неравенство Маркова.
Для любого $\varepsilon$ выполняется

$$\mathrm{P}[Y \geqslant \varepsilon^2] \leqslant \frac{\mathrm{E}[Y]}{\varepsilon^2} \iff \mathrm{P}[|X - \mathrm{E}[X]|^2 \geqslant \varepsilon^2] \leqslant \frac{\mathrm{D}[X]}{\varepsilon^2} \iff \mathrm{P}[|X - \mathrm{E}[X]| \geqslant \varepsilon] \leqslant \frac{\mathrm{D}[X]}{\varepsilon^2}.$$

■

**Теорема** (Закон Больших Чисел в слабой форме)**.** Рассмотрим последовательность $\{X_n\}_n$ случайных независимых величин, что $\mathrm{E}[X_n^2] < \infty$ для любого $n$.
Обозначим $\mathrm{E}[X_n] = a_n$ и $\mathrm{D}[X_n] = \sigma_n^2$. Если

$$\lim_{n \to \infty} \frac{\sigma_1^2 + \dots + \sigma_n^2}{n^2} = 0,$$

то для всякого $\varepsilon > 0$ выполняется

$$\mathrm{P}\left[\left|\frac{X_1 + \dots + X_n}{n} - \frac{a_1 + \dots + a_n}{n}\right| \geqslant \varepsilon\right] \leqslant \frac{\sigma_1^2 + \dots + \sigma_n^2}{n^2 \varepsilon^2}.$$

*Доказательство.* Рассмотрим случайную величину $X = \dfrac{X_1 + \dots + X_n}{n}$.
По линейности математического ожидания получаем

$$\mathrm{E}[X] = \frac{\mathrm{E}[X_1] + \dots + \mathrm{E}[X_n]}{n} = \frac{a_1 + \dots + a_n}{n}.$$

Теперь необходимо найди дисперсию случайной величины $X$:

- Константа из дисперсии выносится с возведением в квадрат, поэтому

$$\mathrm{D}[X] = \mathrm{D}\left[\frac{X_1 + \dots + X_n}{n}\right] = \frac{\mathrm{D}[X_1 + \dots + X_n]}{n^2}.$$

- Так как $\{X_n\}_n$ это последовательность **независимых** случайных величин, дисперсия суммы может быть раскрыта как сумма дисперсий:

$$\mathrm{D}[X] = \frac{\mathrm{D}[X_1 + \dots + X_n]}{n^2} = \frac{\mathrm{D}[X_1] + \dots + \mathrm{D}[X_n]}{n^2} = \frac{\sigma_1^2 + \dots + \sigma_n^2}{n^2}.$$

Воспользуемся неравенством Чебышева для случайной величины $X$ и подставим найденное математическое ожидания и дисперсию:

$$\mathrm{P}[|X - \mathrm{E}[X]| \geqslant \varepsilon] \leqslant \frac{\mathrm{D}[X]}{\varepsilon^2} \iff \mathrm{P}\left[\left|\frac{X_1 + \cdots + X_n}{n} - \frac{a_1 + \cdots + a_n}{n}\right| \geqslant \varepsilon\right] \leqslant \frac{\sigma_1^2 + \cdots + \sigma_n^2}{n^2\varepsilon^2}$$

■

Закон больших чисел удобно применять, когда $X_n$ это независимые одинаково распределенные случайные величины (с конечным вторым моментом). В частности это означает, что у всех величин одно и то же математическое ожидание и одна и та же математическая дисперсия: $\mathrm{E}[X_n] = a$ и $\mathrm{D}[X_n] = \sigma^2$.
Тогда дисперсия среднего арифметического $\dfrac{\mathrm{D}[X_1] + \cdots + \mathrm{D}[X_n]}{n^2} = \dfrac{\sigma^2}{n}$ стремится к нулю и получаем

$$\mathrm{P}\left[\left|\frac{X_1 + \cdots + X_n}{n} - a\right| \geqslant \varepsilon\right] \to 0.$$

То есть в каком-то смысле среднее арифметическое приближается к математическому ожиданию.

### 1.2. Усиленный закон больших чисел Колмогорова (б/д).

**Теорема** (Усиленный закон больших чисел Колмогорова)**.** Пусть $\{X_n\}_n$ — это последовательность независимых одинаково распределенных случайных величин, у которых есть математическое ожидание и пусть $\mathrm{E}[X_n] = a$.
Тогда

$$\mathrm{P}\left[\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = a\right] = 1.$$

Заметьте, что мы не требуем наличия второго момента, в отличие ЗБЧ в слабой форме. Также, эта сходимость более сильная, так как предел находится внутри условия вероятности, это будет объяснено позже.

### 1.3. Сходимости случайных величин: почти наверное и по вероятности.

**Определение.** Последовательность случайных величин $X_n$ сходится к случайной величине $X$ **по вероятности**, если для любого $\varepsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}[|X_n - X| \geqslant \varepsilon] = 0.$$

Записывают в следующем виде: $X_n \xrightarrow{P} X$.

**Определение.** Последовательность случайных величин $X_n$ сходится к случайной величине $X$ **почти наверное**, если

$$\mathrm{P}[\lim_{n\to\infty} X_n = X] = 1.$$

Записывают в следующем виде: $X_n \xrightarrow{\text{п. н.}} X$.

То есть в законе больших чисел в слабой форме речь идет о сходимости по вероятности, а в усиленном законе больших чисел Колмогорова — о сходимости почти наверное.
Из сходимости почти наверное следует сходимость по вероятности, поэтому усиленный закон больших чисел называется усиленным.

### 1.4. Взаимосвязь сходимостей по вероятности и почти наверное.

**Теорема.** Если последовательность случайных величин $X_n$ сходится к $X$ почти наверное, то $X_n$ сходится к $X$ и по вероятности.

*Доказательство.* Хотим доказать. что $\mathrm{P}[|X_n - X| > \varepsilon] \to 0$, что равносильно $\mathrm{P}[|X_n - X| < \varepsilon] \to 1$, что мы и будем заказывать.
Переформулируем выражение «множество исходов, что для любого $\varepsilon > 0$ существует $N$, что для любого $n > N$ выполняется $|X_n - X| < \varepsilon$» с помощью множеств:

$$\bigcup_N \bigcap_{n=N+1}^{\infty} \{w : |X_n - X| < \varepsilon\}.$$

Но это множество включает в себя множество исходов, для которых $\lim X_n = X$:

$$\bigcup_N \bigcap_{n=N+1}^{\infty} \{w : |X_n - X| < \varepsilon\} \supseteq \{w : \lim X_n = X\}.$$

Но по условию $\mathrm{P}[\lim X_n = X] = 1$, поэтому

$$\mathrm{P}\left[\bigcup_N \bigcap_{n=N+1}^{\infty} \{w : |X_n - X| < \varepsilon\}\right] = 1.$$

Обозначим $B_N = \bigcap\limits_{n=N+1}^{\infty} \{w : |X_n - X| < \varepsilon\}$. Тогда

$$B_{N+2} \supseteq B_{N+1} \supseteq B_N \supseteq \dots \supseteq B_1,$$

так как чем больше номер множества, тем из меньшего числа пересечения оно состоит.
Из второго модуля про вероятность вложенных событий мы знаем (теорема о непрерывности вероятностных мер), что

$$\mathrm{P}\left[\bigcup_{N=1}^{\infty} B_N\right] = \lim_{N\to\infty} \mathrm{P}[B_N].$$

Но мы уже доказали, что $\mathrm{P}\left[\bigcup\limits_{N=1}^{\infty} B_N\right] = 1$, тогда

$$\mathrm{P}\left[\bigcap_{n=N+1}^{\infty} \{w : |X_n - X| < \varepsilon\}\right] \xrightarrow[N\to\infty]{} 1,$$

Заметим, что вероятность одного множества событий не меньше вероятности пересечения, поэтому о лемме о двух миллиционерах:

$$\mathrm{P}[\{w : |X_n - X| < \varepsilon\}] \to 1.$$

■

## 2. Сходимость случайных величин по распределению. Лемма о достаточном условии сходимости ожиданий функций из заданного семейства от последовательности случайных величин (лемма 2 из лекции 2). Эквивалентное описание сходимости по распределению.

### 2.1. Сходимость случайных величин по распределению

**Определение.** Последовательность случайных величин $X_n$ сходится к случайной величине $X$ **по распределению**, если $\lim\limits_{n\to\infty} F_{X_n}(x) = F_X(x)$ в каждой точке $x$, в которой непрерывна функция $F_X$.

В математических обозначениях:

$$X_n \xrightarrow{d} X \;\Leftrightarrow\; F_{X_n}(x) \to F_X(x) \;\; \forall x \text{ т.ч. } \; F_X \text{ непр. в т. } x$$

Точек разрыва у (монотонной!) функции распределения не более чем счетное число (как и попарно не пересекающихся интервалов на прямой).

### 2.2. Лемма о достаточном условии сходимости ожиданий функций из заданного семейства от последовательности случайных величин (лемма 2 из лекции 2)

**Лемма.** Пусть $\{X_n\}_{n=0}^{\infty}$ — последовательность случайных величин; $\mathcal{F} := \{f\}$ и $\mathcal{G} := \{g\}$ — системы функций на $\mathbb{R}$. Пусть также

1) $\forall f \in \mathcal{F} \;\; \mathbb{E} f(X_n) \to \mathbb{E} f(X_0)$;

2) $\forall g \in \mathcal{G}\ \forall \varepsilon > 0 \quad f_\varepsilon \in \mathcal{F}$: $\mathbb{E}\,|g(X_n) - f_\varepsilon(X_n)| \le \varepsilon \ \ \forall n \in \mathbb{N} \cup \{0\} \ \ (X_0 = X)$

$\Rightarrow \ \forall g \in \mathcal{G} : \ \mathbb{E}\, g(X_n) \to \mathbb{E}\, g(X_0)$

*Доказательство.* В силу условий

$$|\mathbb{E}\, g(X_n) - \mathbb{E}\, g(X_0)| = \mathbb{E}|(g(X_n) - f_{\varepsilon/3}(X_n)) + (f_{\varepsilon/3}(X_0) - g(X_0)) + (f_{\varepsilon/3}(X_n) - f_{\varepsilon/3}(X_0))| \le$$

$$\le \underbrace{|g(X_n) - f_{\varepsilon/3}(X_n)|}_{\le\varepsilon/3} + \underbrace{|f_{\varepsilon/3}(X_0) - g(X_0)|}_{\le\varepsilon/3} + \underbrace{|f_{\varepsilon/3}(X_n) - f_{\varepsilon/3}(X_0)|}_{\le\varepsilon/3} \le \varepsilon,$$

что и требовалось. ■

### 2.3. Эквивалентное описание сходимости по распределению

**Теорема.** Последовательность случайных величин $X_n$ сходится по распределению к $X$ тогда и только тогда, когда

$$\forall g : \mathbb{R} \to \mathbb{R} \text{ — непр. и огр. : } \lim_{n\to\infty} \mathbb{E}\, g(X_n) = \mathbb{E}\, g(X) \qquad \left(F_{X_n} \xrightarrow{d} F_X \ \Leftrightarrow \ \mathbb{E}\, g(X_n) \to \mathbb{E}\, g(X)\right)$$

*Доказательство.* $\Leftarrow$ Пусть $t$ — точка непрерывности $F_X$. Заметим, что $\underset{P(X\le x)}{F_X(x) = \mathbb{E}I_{(-\infty;x]}(x)}$.

Для всякого $\delta > 0$ определим непрерывные и ограниченные функции

$$g_\delta(x) = \begin{cases} 1, & x < t - \delta, \\ \delta^{-1}(t - x), & t - \delta \le x \le t, \\ 0, & x > t. \end{cases} \qquad h_\delta(x) = \begin{cases} 1, & x < t, \\ \delta^{-1}(t + \delta - x), & t \le x \le t + \delta, \\ 0, & x + \delta > t. \end{cases}$$

При этом

$$I_{(-\infty;x-\delta]}(t) \le g_\delta(t) \le I_{(-\infty;x]}(t) \le h_\delta(t) \le I_{(-\infty;x+\delta]}(t),$$

следовательно,

$$\mathbb{E}\, g_\delta(X_n) \le F_{X_n}(t) \le \mathbb{E}\, h_\delta(X_n).$$

Устремляя $n \to \infty$, получаем

$$\mathbb{E}\, g_\delta(X) \le \varliminf_{n\to\infty} F_{X_n}(t) \le \varlimsup_{n\to\infty} F_{X_n}(t) \le \mathbb{E}\, h_\delta(X).$$

$$F_X(t - \delta) = \mathbb{E}I_{(-\infty;t-\delta]}(X), \ I_{(-\infty;t-\delta]}(X) \le g_\delta(X) \ \Rightarrow \ F_X(t - \delta) \le \mathbb{E}\, g_\delta(X),$$

аналогично $\mathbb{E}h_\delta(X) \le F_X(t + \delta)$.

При $\delta \to 0$ приходим к равенству $\lim\limits_{n\to\infty} F_{X_n}(t) = F_X(t)$.

$\Rightarrow$ Известно, что $F_{X_n} \to F_X$ в каждой точке непрерывности, т.е. $\mathbb{E}I_{(-\infty;x]}(X_n) \to \mathbb{E}I_{(-\infty;x]}(X)$.
Заметим, что $I_{(-\infty;b]}(t) - I_{(-\infty;a]}(t) = I_{(a;b]}(t)$ и обозначим $f_x(t) = I_{(-\infty;x]}(t)$.
В силу линейности предела $\mathbb{E}(f_b(X_n) - f_a(X_n)) \to \mathbb{E}(f_b(X) - f_a(X))$.
Обозначим $f(t) = \sum\limits_{j=1}^{N} c_j I_{(a_j,b_j]}(t)$, где $(a_j, b_j]$ — точки непрерывности $F_X$.
Т.к. $f(t) = \sum\limits_{j=1}^{N} c_j (f_b(t) - f_a(t))$, по линейности предела имеем $\mathbb{E}\, f(X_n) \to \mathbb{E}\, f(X)$.

Обозначим $\mathcal{F}$ систему функций $f$. По лемме 2.3 достаточно показать, что $\mathbb{E}\, g(X_n) \to \mathbb{E}\, f(X_n)$ $\forall n \in \mathbb{N} \cup \{0\}$, т.е.

$$\forall \text{ непр. огр. } g\ \forall \varepsilon > 0\ \exists f \in \mathcal{F} : \mathbb{E}|g(X_n) - f_\varepsilon(X_n)| \le \varepsilon \ \ \forall n \in \mathbb{N} \cup \{0\} \ (X_0 = X).$$

Пусть $g$ — произвольная непре. огр. функция на $\mathbb{R}$ и пусть $\varepsilon > 0$ фиксирован.

$$\begin{cases} \lim\limits_{x\to-\infty} F_X(x) & = & 0, \\ \lim\limits_{x\to+\infty} F_X(x) & = & 1 \end{cases} \Rightarrow \ \exists A : \begin{cases} F_X(-A) & < & \varepsilon, \\ 1 - F_X(A) & < & \varepsilon \end{cases}$$

Так как у монотонной функции счетное число точек разрыва, без ограничения общности будем считать, что $A$ — точка непрерывности $F_X$. Тогда (в силу $F_{X_n} \to F_X$)

$$\exists n_0 : \forall n > n_0 \; F_{X_n}(-A) < 2\varepsilon, \; 1 - F_{X_n}(A) < 2\varepsilon$$

Увеличив $A$, можно считать, что последнее верно для всех $n$ (перешли от $n_0(A)$ к $A(n)$). Таким образом

$$\mathbb{E}\left|I_{\{-A \le X_n \le A\}} g(X_n) - g(X_n)\right| = \mathbb{E}|\overline{I_{\{-A \le X_n \le A\}}} g(X_n)| \le \sup|g| \cdot \underset{=F_{x_n}(-A)+(1-F_{X_n}(A))}{\underset{\|}{P(|X_n| > A)}} \le \sup|g| \cdot 4\varepsilon$$

Та же оценка верна и для предельной случайной величины $X$.
Непрерывная на отрезке $[-A; A]$ функция $g$ равномерно непрерывна на этом отрезке ($\forall \varepsilon > 0 \; \exists \delta : |g(x) - g(x + \delta)| < \varepsilon$), следовательно, ее можно приблизить кусочно-постоянной (ступенчатой) функцией $f_\varepsilon$:

$$|g(x) - f_\varepsilon(x)| \le \varepsilon \; \forall x \in [-A; A].$$

Также, не ограничивая общности, можно считать, что точки разрыва $f_\varepsilon$ — точки непрерывности $F_X$, т.е. $f_\varepsilon \in \mathcal{F}$ (т.к. подходящих точек на отрезке $[x_i; x_i + \delta]$ — континуум, а точек разрыва счетно).
Пусть $f_\varepsilon(x) = 0$ при $x \notin [-A; A]$. Тогда

$$\mathbb{E}\left|I_{\{-A \le X_n \le A\}} g(X_n) - f_\varepsilon(X_n)\right| = \mathbb{E}\left[|g(X_n) - f_\varepsilon(X_n)| \, I_{\{-A \le X_n \le A\}}\right] \le \varepsilon$$

Та же оценка верна и для предельной случайной величины $X$. Таким образом

$$\mathbb{E}\,|g(X_n) - f_\varepsilon(X_n)| \le 2M\varepsilon + \varepsilon.$$

Теорема доказана. ■

## 3. Абсолютная непрерывность математического ожидания. Теорема Лебега о мажорируемой сходимости. Подстановка сходящейся по вероятности последовательности случайных величин в непрерывную функцию. Взаимосвязь сходимостей по вероятности и по распределению.

### 3.1. Абсолютная непрерывность математического ожидания.

**Предложение.** Имеем случайную величину $Y \geqslant 0$ п.н., матожидание конечно.
Хотим показать $\forall \varepsilon \exists \delta > 0 : P(A) \leqslant \delta \Rightarrow \mathbb{E}\,(Y I_A) < \varepsilon$. Словами: если множество $A$ маленькое, то матожидание случайной величины, которая много где ноль, тоже маленькое.

*Доказательство.* Так как $Y \geqslant 0, \mathbb{E}Y = \mathbb{E}Y_+ = \sup\{\mathbb{E}U : U \leqslant Y_+, \text{огр.}\}$.
Фиксируем $U$ так, чтобы $\mathbb{E}U \leqslant \mathbb{E}Y_+ \leqslant \mathbb{E}U + \dfrac{\varepsilon}{2}$ — будем пользоваться тем, что $U$ ограничена, то есть $\exists R \in \text{const} : R \geqslant U$.

$$\mathbb{E}\,(Y I_A) = \mathbb{E}\,(Y_+ I_A) = \mathbb{E}\,((Y_+ - U)\, I_A) + \mathbb{E}\,(U I_A) \leqslant \mathbb{E}\,(Y_+ - U) + R P(A) \leqslant \frac{\varepsilon}{2} + R\delta$$

Выбрав $\delta = \dfrac{\varepsilon}{2R}$ получаем то что требовали. ■

Абсолютная непрерывность потому что на самом деле можно было взять $|Y|$ и работать с ним, но мы решили просто сказать что он неотрицателен удобства ради.

### 3.2. Теорема Лебега о мажорируемой сходимости.

**Теорема.** Если

- $X_n \xrightarrow{P} X$ (или $X_n \xrightarrow{\text{п. н.}} X$ потому что следует);
- существует случайная величина $Y$ такая что $|X_n| \leqslant Y$ п. н.; $|X| \leqslant Y$ п. н.,

То $\mathbb{E}X_n \to \mathbb{E}X$

*Доказательство.*

$$|\mathbb{E}X_n - \mathbb{E}X| \leqslant \mathbb{E}|X_n - X| = \mathbb{E}\left[|X_n - X| I_{\{|X_n - X| \leqslant \varepsilon\}}\right] + \mathbb{E}\left[|X_n - X| I_{\{|X_n - X| > \varepsilon\}}\right]$$

$\mathbb{E}\left[|X_n - X| I_{\{|X_n - X| \leqslant \varepsilon\}}\right] \leqslant \varepsilon$
$\mathbb{E}\left[|X_n - X| I_{\{|X_n - X| > \varepsilon\}}\right] \leqslant \mathbb{E}\left[2Y I_{\{|X_n - X| > \varepsilon\}}\right]$. Так как $X_n \xrightarrow{P} X \implies P\left(|X_n - X| > \varepsilon\right) \to 0$.
По Абсолютной непрерывности математического ожидания: $\forall \varepsilon \exists \delta : P(A) < \delta \implies \mathbb{E}\left[Y I_A\right] < \varepsilon$. Здесь $P(|X_n - X| > \varepsilon) < \delta$, поэтому применимо.
Получается $|\mathbb{E}X_n - \mathbb{E}X| \leqslant \mathbb{E}|X_n - X| = \ldots \leqslant 3\varepsilon$ ■

### 3.3. Подстановка сходящейся по вероятности последовательности случайных величин в непрерывную функцию.

Это лекция 1, если что

**Утверждение 0.1.** $X_n \xrightarrow{P} X$, $g : \mathbb{R} \to \mathbb{R}$ — *непрерывная* $\implies g(X_n) \xrightarrow{P} g(X)$

*Доказательство.* Для фиксированного $R$, $g$ равномерно непрерывна на отрезке $[-R, R]$, то есть

$$\forall \varepsilon > 0 \exists \delta > 0 : x, y \in [-R, R], |x - y| < \delta \implies |g(x) - g(y)| < \varepsilon$$

Рассмотрим множество

$$\begin{aligned}
&\{|g(X_n) - g(X)| \geqslant \varepsilon\} \subseteq \\
&\subseteq \{|g(X_n) - g(X)| \geqslant \varepsilon, X_n, X \in [-R, R]\} \cup \{|g(X_n) - g(X)| \geqslant \varepsilon, X_n \notin [-R, R]\} \cup \{|g(X_n) - g(X)| \geqslant \varepsilon, X \notin [-R, R]\}
\end{aligned}$$

$$P\left(|g(X_n) - g(X)| \geqslant \varepsilon\right) \leqslant P\left(|g(X_n) - g(X)| \geqslant \varepsilon, X_n, X \in [-R, R]\right) + P(|X_n| > R) + P(|X| > R)$$

По условию равномерной непрерывности $P\left(|g(X_n) - g(X)| \geqslant \varepsilon, X_n, X \in [-R, R]\right) \leqslant P\left(|X_n - X| \geqslant \delta\right)$ — иначе условие выполнялось бы, и разница между образами была бы меньше эпсилона.
Заметим, что $\{|X_n - X + X| > R\} \subseteq \left\{|X_n - X| > \dfrac{R}{2}\right\} \cup \left\{|X| > \dfrac{R}{2}\right\}$, тогда $P(|X_n| > R) \leqslant P\left(|X_n - X| > \dfrac{R}{2}\right) + P\left(|X| > \dfrac{R}{2}\right)$
Получается

$$P\left(|g(X_n) - g(X)| \geqslant \varepsilon\right) \leqslant P\left(|X_n - X| \geqslant \delta\right) + P\left(|X_n - X| > \frac{R}{2}\right) + P\left(|X| > \frac{R}{2}\right) + P(|X| > R)$$

Взяв большие $R$ получаем $P\left(|X| > \dfrac{R}{2}\right) + P\left(|X| > R\right) \to 0$ очев.

Взяв большие $n$ из-за сходимости по вероятности получаем $P\left(|X_n - X| \geqslant \delta\right) + P\left(|X_n - X| > \dfrac{R}{2}\right) \to 0$.
Получается

$$0 \leqslant \liminf P(|g(X_n) - g(X)| \geqslant \varepsilon) \leqslant \limsup P(|g(X_n) - g(X)| \geqslant \varepsilon) \leqslant 0$$

то есть $g(X_n) \xrightarrow{P} g(X)$

■

### 3.4. Взаимосвязь сходимостей по вероятности и по распределению.

**Следствие.** $X_n \xrightarrow{P} X \implies X_n \xrightarrow{d} X$

*Доказательство.* Нужно доказать, что для любой ограниченной непрерывной $g$ верно $\mathbb{E}g(X_n) \to \mathbb{E}g(X)$
В силу ограниченности имеем $|g(t)| \leqslant M \forall t \implies g(X_n) \leqslant M, g(X) \leqslant M$
По предыдущему пункту $g(X_n) \xrightarrow{P} g(X)$
Введем случайную величину $Y = M$ и применим Лебега (оба условия выполняются), тогда $\mathbb{E}X_n \to \mathbb{E}X$, то есть $X_n \xrightarrow{d} X$

■

# 4. Характеристические функции: определение и свойства. Вычисление характеристической функции нормальной случайной величины. Производные характеристических функций.

## 4.1. Характеристические функции: определение и свойства.

**Определение.** Пусть $X$ это случайная величина. Тогда характеристическая функция случайной величины $X$ это

$$\varphi_X(t) := \mathrm{E}[e^{itX}] = \mathrm{E}[\cos(t \cdot X)] + i \cdot \mathrm{E}[\sin(t \cdot X)].$$

**Теорема** (Свойства характеристических функций)**.** У характеристической функции есть следующие свойства:

1. Для любой случайной величины $X$ выполняется $\varphi_X(0) = 1$.
2. Для любой случайной величины и любого $t \in \mathbb{R}$ выполняется $|\varphi_X(t)| \leqslant 1$.
3. Для чисел $a$ и $b$ выполняется
$$\varphi_{aX+b}(t) = e^{itb} \cdot \varphi_X(at).$$
4. Если $X_n$ это последовательность **независимых** случайных величин, то
$$\varphi_{X_1+X_2+\dots+X_n}(t) = \varphi_{X_1}(t) \cdot \varphi_{X_2}(t) \cdot \dots \cdot \varphi_{X_n}(t).$$

*Доказательство.* Для доказательства будем пользоваться следующей формулой:

$$\varphi_X(t) = \mathrm{E}[\cos(t \cdot X)] + i \cdot \mathrm{E}[\sin(t \cdot X)],$$

которая следует из формулы Эйлера $e^{ix} = \cos x + i \cdot \sin x$.
Докажем свойства:

1. Для любой случайной величины $X$ выполняется $\varphi_X(0) = 1$.

   Проверяется подстановкой:
$$\varphi_X(0) = \mathrm{E}[e^{i \cdot 0 \cdot X}] = \mathrm{E}[e^0] = \mathrm{E}[1] = 1.$$

2. Для любой случайной величины и любого $t \in \mathbb{R}$ выполняется $|\varphi_X(t)| \leqslant 1$.

   Рассмотрим случайную величину $Y$. Знаем, что ее дисперсия неотрицательна, то есть $\mathrm{D}[Y] = \mathrm{E}[Y^2] - (\mathrm{E}[Y])^2 \geqslant 0$, откуда следует, что для любой случайной величины $Y$ справедливо $\mathrm{E}[Y^2] \geqslant (\mathrm{E}[Y])^2$.

   Значение характеристической функции это комплексное число. Квадрат модуля комплексного числа это сумма квадратов его мнимой и действительной частей:
$$|\varphi_X(t)|^2 = (\mathrm{E}[\cos(t \cdot X)])^2 + (\mathrm{E}[\sin(t \cdot X)])^2.$$

   С помощью знаний о $\mathrm{E}[Y^2] \geqslant (\mathrm{E}[Y])^2$ оценим квадрат модуля характеристической функции:
$$|\varphi_X(t)|^2 = (\mathrm{E}[\cos(t \cdot X)])^2 + (\mathrm{E}[\sin(t \cdot X)])^2 \leqslant \mathrm{E}[\cos^2(t \cdot X)] + \mathrm{E}[\sin^2(t \cdot X)] = \mathrm{E}[\cos^2(t \cdot X) + \sin^2(t \cdot X)] = \mathrm{E}[1] = 1.$$

3. Для чисел $a$ и $b$ выполняется
$$\varphi_{aX+b}(t) = e^{itb} \cdot \varphi_X(at).$$

   Заметим, что если $y$ это некоторое число, то $\mathrm{E}[y \cdot X] = y \cdot \mathrm{E}[X]$ по линейности математического ожидания.

   Запишем по определению:
$$\varphi_{aX+b}(t) = \mathrm{E}[e^{it \cdot aX + it \cdot b}] = \mathrm{E}[e^{it \cdot aX} \cdot e^{it \cdot b}] = e^{it \cdot b} \cdot \mathrm{E}[e^{it \cdot aX}] = e^{it \cdot b} \cdot \varphi_{aX}(t).$$

4. Если $X_n$ это последовательность **независимых** случайных величин, то

$$\varphi_{X_1+X_2+\dots+X_n}(t) = \varphi_{X_1}(t)\cdot\varphi_{X_2}(t)\cdot\dots\cdot\varphi_{X_n}(t).$$

Пусть $Y_n = e^{i\cdot t\cdot X_n}$. Тогда $Y_1,\dots Y_n$ это последовательность независимых случайных величин (в силу независимости $X_n$) и $\mathrm{E}[Y_1\cdot\dots\cdot Y_n] = \mathrm{E}[Y_1]\cdot\dots\cdot\mathrm{E}[Y_n]$.

Запишем по определению:

$$\varphi_{X_1+X_2+\dots+X_n}(t) = \mathrm{E}[e^{itX_1+\dots+itX_n}] = \mathrm{E}[e^{itX_1}\cdot\dots\cdot e^{itX_n}] = \mathrm{E}[Y_1\cdot\dots\cdot Y_n] = \mathrm{E}[Y_1]\cdot\dots\cdot\mathrm{E}[Y_n] = \varphi_{X_1}(t)\cdot\dots\cdot\varphi_{X_n}(t).$$

■

### 4.2. Вычисление характеристической функции нормальной случайной величины.

Хотим вычислить $\varphi_\xi(t)$, где $\xi\sim\mathcal{N}(0,1)$.
Запишем по определению:

$$\varphi_\xi(t) = \mathrm{E}[e^{it\xi}] = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}\cos(tx)\exp\left[-x^2/2\right]\mathrm{d}x + \frac{i}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}\sin(tx)\exp\left[-x^2/2\right]\mathrm{d}x.$$

Заметим, что второе слагаемое $\frac{i}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}\sin(tx)\exp\left[-x^2/2\right]\mathrm{d}x$ равно нулю, так как это интеграл нечетной функции по симметричному промежутку. Тогда

$$\varphi_\xi(t) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}\cos(tx)\exp\left[-x^2/2\right]\mathrm{d}x.$$

Возьмем производную по $t$ (считаем, что она берется):

$$\begin{aligned}\varphi'_\xi(t) &= -\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}x\cdot\sin(tx)\exp\left[-x^2/2\right]\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}\sin(tx)\,\mathrm{d}\big(\exp\left[-x^2/2\right]\big)\\ &= \frac{1}{\sqrt{2\pi}}\sin(tx)\exp\left[-x^2/2\right]\Big|_{-\infty}^{+\infty} - \frac{t}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}\cos(tx)\exp\left[-x^2/2\right]\mathrm{d}x\\ &= 0 - \frac{t}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}\cos(tx)\exp\left[-x^2/2\right]\mathrm{d}x = -t\cdot\varphi_\xi(t).\end{aligned}$$

Пришли к дифференциальному уравнению:

$$\varphi'_\xi(t) = -t\cdot\varphi_\xi(t) \implies \frac{\varphi'_\xi(t)}{\varphi_\xi(t)} = -t.$$

Интегрируем обе части:

$$\int\frac{\mathrm{d}(\varphi_\xi(t))}{\varphi_\xi(t)} = \ln|\varphi_\xi(t)| + C = \int -t\mathrm{d}t = -\frac{t^2}{2}.$$

Теперь берем экспоненту от обеих частей:

$$\varphi_\xi(t) = C'\cdot\exp\left[-t^2/2\right],$$

где $C'$ это некоторая константа.
Про характеристическую функцию мы знаем, что $\varphi_\xi(0)=1$. Тогда

$$\varphi_\xi(0) = 1 = C'\cdot\exp[0] = C',$$

откуда находим $C'=1$.
Тогда характеристическая функция стандартной нормальной величины имеет следующий вид:

$$\varphi_\xi(t) = \exp\left[-t^2/2\right].$$

### 4.3. Производные характеристических функций.

**Теорема.** Пусть $X$ это случайная величина с конечным $k$-ым моментом ($\mathrm{E}[|X|^k] < \infty$). Тогда $\varphi_X$ $k$ раз дифференцируема и
$$\varphi_X^{(k)}(0) = i^k \cdot \mathrm{E}[X^k].$$

*Доказательство.* Докажем для $k = 1$, для остальных порядков аналогично.
Мы хотим найти производную:
$$\lim_{h_n \to 0} \frac{\varphi_X(t+h_n) - \varphi_X(t)}{h_n} = \lim_{h_n \to 0} \frac{1}{h_n} \cdot \left(\mathrm{E}[e^{i(t+h_n)X}] - \mathrm{E}[e^{itX}]\right) = \lim_{h_n \to 0} E\left[\frac{e^{i(t+h_n)X} - e^{itX}}{h_n}\right] =: \lim_{h_n \to 0} E[g_n],$$
то есть обозначили $g_n = \dfrac{e^{i(t+h_n)X} - e^{itX}}{h_n}$.
Поймем, что мы знаем про функцию $g_n$:

- У нее есть поточечный предел:
$$\lim_{n \to \infty} g_n(X) = \left(e^{itX}\right)'_t = iXe^{itX}.$$
- Надо как-то оценить $|g_n|$.

  Знаем, что модуль комплексной экспоненты равен 1, то есть $|e^{itX}| = 1$. Тогда
$$|g_n(X)| = \left|\frac{e^{itX} \cdot (e^{ih_nX} - 1)}{h_n}\right| = |e^{itX}| \cdot \left|\frac{e^{ih_nX} - 1}{h_n}\right| = \left|\frac{e^{ih_nX} - 1}{h_n}\right| = \left|\frac{e^{ih_nX} - e^{i\cdot 0\cdot X}}{h_n}\right| = \left(e^{itX}\right)'_t(\xi) = \left|iXe^{i\xi X}\right|$$
  для некоторого $\xi \in (0; h_n)$.

  Предпоследний переход выполнен по теореме Лагранжа, которая гласит следующее:
$$\exists \xi \in (a;b): \quad \frac{f'(b) - f'(a)}{b-a} = f'(\xi).$$

  Опять же воспользуемся тем, что модуль комплексной экспоненты равен 1:
$$|g_n(X)| = \left|iXe^{i\xi X}\right| = |i| \cdot |X| \cdot \left|e^{i\xi X}\right| = 1 \cdot |X| \cdot 1 = |X|.$$

Мы получили, что

- $|g_n(X)| \leqslant |X|$ и $\mathrm{E}[|X|] < \infty$ (для этого и нужна конечность моментов);
- $g_n(X) \xrightarrow{\text{п. н.}} i \cdot X \cdot e^{itX}$.

Тогда по теореме Лебега предел ожиданий есть ожидание предела:
$$\lim_{n \to \infty} \mathrm{E}[g_n(X)] = \mathrm{E}[i \cdot X \cdot e^{itX}].$$

Возвращаемся в самое начало:
$$\varphi_X(t)' = \lim_{n \to \infty} \mathrm{E}[g_n] = i \cdot \mathrm{E}[X \cdot e^{itX}].$$

■

## 5. Переформулировка сходимости по распределению в терминах характеристических функций. Однозначность задания распределения случайной величины ее характеристической функцией. Центральная предельная теорема.

### 5.1. Переформулировка сходимости по распределению в терминах характеристических функций.

**Теорема.** Последовательность $X_n$ сходится по распределению к $X$ ($X_n \xrightarrow{d} X$) $\Leftrightarrow \lim_{n \to \infty} \varphi_{X_n}(t) = \varphi_X(t)\ \forall t \in \mathbb{R}$.

*Доказательство.*

$\Rightarrow$

$\varphi_{X_n}(t) = \mathbb{E}[\cos tX_n] + i\mathbb{E}[\sin tX_n]$, при этом функции $x \mapsto \cos tx$ и $x \mapsto \sin tx$ - непрерывные и ограниченные.

Мы знаем, что $X_n \xrightarrow{d} X \Leftrightarrow \forall g$ - непрерывной ограниченной $\mathbb{E}g(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}g(X)$ *(см. билет 2)*.

Тогда возьмем в качестве $g(x)$ функцию $x \mapsto \cos tx$: $\mathbb{E}[\cos tX_n] \xrightarrow[n\to\infty]{} \mathbb{E}[\cos tX]$.

Теперь возьмем в качестве $g(x)$ функцию $x \mapsto \sin tx$: $\mathbb{E}[\sin tX_n] \xrightarrow[n\to\infty]{} \mathbb{E}[\sin tX]$.

Таким образом, получаем необходимое равенство:

$$\lim_{n\to\infty} \varphi_{X_n}(t) = \lim_{n\to\infty} (\mathbb{E}[\cos tX_n] + i\mathbb{E}[\sin tX_n]) = \lim_{n\to\infty} (\mathbb{E}[\cos tX_n]) + \lim_{n\to\infty} (i\mathbb{E}[\sin tX_n]) =$$
$$= \lim_{n\to\infty} (\mathbb{E}[\cos tX_n]) + i \lim_{n\to\infty} (\mathbb{E}[\sin tX_n]) = \mathbb{E}[\cos tX] + i\mathbb{E}[\sin tX] = \varphi_X(t)$$

$\Leftarrow$

Докажем в предположении ограниченности вторых моментов, то есть $\mathbb{E}|X_n|^2 \leqslant C < \infty \ \forall n \in \mathbb{N}$, и, соответственно, $\mathbb{E}|X|^2 \leqslant C < \infty$, где $C = const$. (Доказываемое равенство мы хотим применять в ЦПТ, где ограниченность вторых моментов выполняется, поэтому для наших нужд этого достаточно. Однако на самом деле и общий факт верен)

Нам известно, что $\varphi_{X_n}(t) = \mathbb{E}[\cos tX_n] + i\mathbb{E}[\sin tX_n] \xrightarrow[n\to\infty]{} \varphi_X(t) = \mathbb{E}[\cos tX] + i\mathbb{E}[\sin tX] \ \forall t \in \mathbb{R}$. Понятно, что если комплексные числа сходятся, то вещественная часть сходится к вещественной и мнимая - к мнимой. $\mathbb{E}[\cos tX_n] \xrightarrow[n\to\infty]{} \mathbb{E}[\cos tX]$ и $\mathbb{E}[\sin tX_n] \xrightarrow[n\to\infty]{} \mathbb{E}[\sin tX] \ \forall t \in \mathbb{R}$.

По линейности будут также сходиться и всевозможные комбинации

$$\forall t_1 \dots t_N : \frac{a_0}{2} + \sum_{k=1}^{N} a_k \cos t_k x + b_k \sin t_k x = f(x) \Rightarrow \mathbb{E}f(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}f(X)$$

<u>Хотим</u>, чтобы такая сходимость была выполнена для каждой непрерывной ограниченной функции. То есть $\mathbb{E}g(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}g(X) \ \forall g$ - непр. огр., что эквивалентно сходимости по распределению.

Мы попадаем в ситуацию леммы *(см. билет 2)*: $\mathcal{F} = \{f\}$, $\mathcal{G} = \{g\}$

1. $\forall f \in \mathcal{F} \ \mathbb{E}f(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}f(X)$
2. $\forall g \in \mathcal{G} \ \forall \varepsilon > 0 \ \exists f_\varepsilon$: $\mathbb{E}|g(X_n) - f_\varepsilon(X_n)| < \varepsilon \ \forall n$ и $\mathbb{E}|g(X) - f_\varepsilon(X)| < \varepsilon$

$\Rightarrow \forall g \in \mathcal{G}$: $\mathbb{E}g(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}g(X)$.

В нашем случае $\mathcal{F} = \left\{\frac{a_0}{2} + \sum_{k=1}^{N} a_k \cos t_k x + b_k \sin t_k x = f(x)\right\}$, $\mathcal{G}$ - класс непрерывных ограниченных функций. Первый пункт леммы выполняется, осталось доказать выполнение второго.

<u>Известно</u> $\eta(\cdot)$ - непрерывная на $\mathbb{R}$ и периодическая с периодом $2t$, тогда

$$\forall \varepsilon > 0 \ \exists a_0, a_1 \dots a_N, b_1 \dots b_N\colon \sup_{x\in\mathbb{R}} \left|\eta(x) - \left(\frac{a_0}{2} + \sum_{k=1}^{N} a_k \cos \frac{\pi k}{T} x + b_k \sin \frac{\pi k}{T} x\right)\right| < \varepsilon, \text{ по теореме Вейерштрасса}$$

(знаем из матанализа, любая непрерывная периодическая функция равномерно приближается тригонометрическим многочленом)

Запишем неравенство Чебышёва для $X_n$ и $X$ (можем это сделать в силу ограниченности второго момента, $\mathbb{E}|X_n|^2 \leqslant C < \infty \ \forall n \in \mathbb{N}$ и $\mathbb{E}|X|^2 \leqslant C < \infty$):

$$P(|X_n| \geqslant A) \leqslant \frac{C}{A^2}$$

$$P(|X| \geqslant A) \leqslant \frac{C}{A^2}$$

Для достаточно большого $A$ будет выполнено:

$$P(|X_n| > A) < \varepsilon$$

$$P(|X| > A) < \varepsilon$$

Теперь введем непрерывную ограниченную периодическую функцию $g_\varepsilon$. Она совпадает с непрерывной ограниченной $g$ на отрезке $[-A, A]$, равна нулю на концах отрезка $[-A-1, A+1]$, а вне отрезка $[-A-1, A+1]$ продолжена как периодическая с периодом $T = 2A+2$. Очевидно, что $|g_\varepsilon(x)| \leqslant \sup |g|$.

Проведем оценку:

$$\mathbb{E}|g(X_n) - g_\varepsilon(X_n)| = \Big[\text{так как } g_\varepsilon(x) = g(x) \text{ на } [-A, A]\Big] = \mathbb{E}\big[|g(X_n) - g_\varepsilon(X_n)| \cdot I_{|X_n| \geqslant A}\big] \leqslant 2 \sup |g| \cdot \varepsilon$$

Получается, мы приблизили функцию $g$ непрерывной периодической функцией $g_\varepsilon$, которую в свою очередь можно приблизить равномерным тригонометрическим многочленом. Для $g_\varepsilon$ $\exists f \in \mathcal{F}$: $\sup\limits_{x\in\mathbb{R}} |g_\varepsilon(x) - f_\varepsilon(x)| < \varepsilon \Rightarrow \mathbb{E}|g_\varepsilon(X_n) - f_\varepsilon(X_n)| < \varepsilon \Rightarrow \mathbb{E}|g(X_n) - f_\varepsilon(X_n)| < \varepsilon + 2\sup|g| \cdot \varepsilon$ - константа, умноженная на $\varepsilon$.

Доказали выполнение второго пункта леммы, поэтому $\forall g \in \mathcal{G}$: $\mathbb{E}g(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}g(X) \Leftrightarrow X_n \xrightarrow{d} X$.

■

### 5.2. Однозначность задания распределения случайной величины ее характеристической функцией.

**Теорема.** Если у двух случайных величин совпадают характеристические функции, то эти величины имеют одинаковые распределения ($\varphi_X(t) = \varphi_Y(t)$ $\forall t \in \mathbb{R} \Rightarrow F_X = F_Y$ $(\mu_X = \mu_Y)$).

*Доказательство.*
Пусть $\varphi_X(t) = \varphi_Y(t)$ $\forall t \in \mathbb{R}$. Рассмотрим последовательность случайных величин равных $X$.
$X_n := X$, тогда $\varphi_{X_n}(t) = \varphi_X(t) = \varphi_Y(t) \xrightarrow[n\to\infty]{} \varphi_Y(t)$ $\forall t \in \mathbb{R}$.
Значит $X_n \xrightarrow{d} Y$, согласно переформулировке сходимости по распределению в терминах характеристических функций (доказывали выше). При этом $X_n := X$, поэтому $X \xrightarrow{d} Y \Leftrightarrow F_X(x) = F_Y(x)$ $\forall x$ - т. непрерывности $F_Y$ (равносильность по определению сходимости по распределению).
Проверим, что происходит в точках разрыва. Заметим, что у монотонных функций (а функция распределения случайной величины монотонна) не более чем счетное число точек разрыва. Тогда к каждой точке разрыва функции $F_Y$ сходится справа некоторая последовательность точек непрерывности этой функции (в которых $F_Y$ и $F_X$ совпадают).
$x_0$ - точка разрыва $F_Y$ $\exists x_n$ - последовательность т. непрерывности $F_Y$: $x_n \geqslant x_0$ и $x_n \xrightarrow[n\to\infty]{} x_0$.
При этом $F_X(x_n) = F_Y(x_n)$ и $F_X(x_n) \xrightarrow[n\to\infty]{} F_X(x_0)$, $F_Y(x_n) \xrightarrow[n\to\infty]{} F_Y(x_0)$ в силу непрерывности функции распределения справа $\Rightarrow F_X(x_0) = F_Y(x_0)$ - в точках разрыва функции распределения совпадают.
Получили, что действительно $\varphi_X(t) = \varphi_Y(t)$ $\forall t \in \mathbb{R} \Rightarrow F_X(x) = F_Y(x)$ $\forall x$. ■

### 5.3. Центральная предельная теорема.

**Теорема.** Пусть $\{X_n\}$ - последовательность независимых одинаково распределенных случайных величин, причем $\mathbb{E}X_1 = a$ и $\mathbb{D}X_1 = \sigma^2$. Тогда для всех $t$:

$$\lim_{n\to\infty} P\Big(\frac{X_1 + \ldots + X_n - na}{\sqrt{n\sigma^2}} \leqslant t\Big) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{t} e^{-\frac{x^2}{2}} dx$$

или, равносильно:

$$\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} \xrightarrow{d} Z, \text{ где } Z \sim \mathcal{N}(0,1) \text{ – стандартная нормальная случайная величина, а } S_n := X_1 + \ldots + X_n.$$

*Доказательство.*
Перейдем от случайных величин $X_j$ к центрированным случайным величинам $X_j' = (X_j - a)$. Математическое ожидание полученных случайных величин равно 0, $\mathbb{E}X_j' = 0$. Дисперсия при сдвиге не изменится, $\mathbb{D}X_j' = \sigma^2$. Тогда

$$\frac{X_1 + \ldots + X_n - na}{\sqrt{n\sigma^2}} = \frac{X_1' + \ldots + X_n'}{\sqrt{n\sigma^2}} \text{ и}$$

$$\frac{X_1+\ldots+X_n-na}{\sqrt{n\sigma^2}}\xrightarrow{d} Z\sim\mathcal{N}(0,1)\Leftrightarrow\frac{X_1'+\ldots+X_n'}{\sqrt{n\sigma^2}}\xrightarrow{d} Z\sim\mathcal{N}(0,1)$$

Вычислим характеристическую функцию случайной величины $\dfrac{X_1'+\ldots+X_n'}{\sqrt{n\sigma^2}}$.

$$\varphi_{\frac{X_1'+\ldots+X_n'}{\sqrt{n\sigma^2}}}(t)=\Big[\text{так как } X_1,\ldots,X_n \text{ независимые}\Big]=\varphi_{\frac{X_1'}{\sqrt{n\sigma^2}}}(t)\cdot\ldots\cdot\varphi_{\frac{X_n'}{\sqrt{n\sigma^2}}}(t)=$$

$$=\varphi_{X_1'}(\frac{t}{\sqrt{n\sigma^2}})\cdot\ldots\cdot\varphi_{X_n'}(\frac{t}{\sqrt{n\sigma^2}})=\Big(\varphi_{X_1'}(\frac{t}{\sqrt{n\sigma^2}})\Big)^n$$

Заметим, что

$$\varphi_{X_1'}(0)=1,\ \varphi_{X_1'}'(0)=i\mathbb{E}X_1'=0,\ \varphi_{X_1'}''(0)=-\mathbb{E}(X_1')^2=-\mathbb{D}X_1'=-\sigma^2$$

По формуле Тейлора

$$\varphi_{X_1'}(\frac{t}{\sqrt{n\sigma^2}})=\varphi_{X_1'}(0)+\frac{t}{\sqrt{n\sigma^2}}\varphi_{X_1'}'(0)+\frac{1}{2}(\frac{t}{\sqrt{n\sigma^2}})^2\varphi_{X_1'}''(0)+o\Big((\frac{t}{\sqrt{n\sigma^2}})^2\Big)=$$

$$=1-\frac{\sigma^2}{2}\cdot\frac{t^2}{n\sigma^2}+o(\frac{1}{n})=1-\frac{t^2}{2n}+o(\frac{1}{n})$$

Тогда

$$\Big(\varphi_{X_1'}(\frac{t}{\sqrt{n\sigma^2}})\Big)^n=\Big(1-\frac{t^2}{2n}+o(\frac{1}{n})\Big)^n=e^{n\ln\left(1-\frac{t^2}{2n}+o(\frac{1}{n})\right)}=\Big[\text{логарифм раскладываем по формуле Тейлора}\Big]=$$

$$=e^{n\left(-\frac{t^2}{2n}+o(\frac{1}{n})\right)}=e^{-\frac{t^2}{2}+o(1)}\xrightarrow[n\to\infty]{}e^{-\frac{t^2}{2}}$$

Остается заметить, что $e^{-\frac{t^2}{2}}$ - характеристическая функция стандартной нормальной случайной величины $Z\sim\mathcal{N}(0,1)$, $\varphi_{\frac{S_n-\mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}}}(t)\xrightarrow[n\to\infty]{}e^{-\frac{t^2}{2}}=\varphi_Z(t)\Rightarrow\dfrac{S_n-\mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}}\xrightarrow{d} Z\sim\mathcal{N}(0,1)$. ■

# 6. Подстановка сходящейся по распределению последовательности случайных величин в непрерывную функцию. Сходимость суммы и произведения сходящихся по распределению последовательностей случайных величин в случае, когда одна из предельных случайных величин постоянная. Примеры применения: выборочная дисперсия и взаимосвязь с ЦПТ. Теорема о сходимости последовательности вида $\dfrac{f(a+h_nX_n)-f(a)}{h_n}$ для сходящейся по распределению последовательности $X_n$. Взаимосвязь с ЦПТ.

## 6.1. Подстановка сходящейся по распределению последовательности случайных величин в непрерывную функцию.

**Теорема.** Если последовательность случайных величин $X_n$ сходится по распределению к X, то для всякой непрерывной функции f случайные величины $f(X_n)$ сходятся по распределению к $f(X)$.

*Доказательство.*

Из лекции 2 мы знаем, что

$$X_n\xrightarrow{d} X\Leftrightarrow\forall g\ \mathbb{E}g(X_n)\xrightarrow[n\to\infty]{}\mathbb{E}g(X),\text{ где g – непрерывная, ограниченная функция}$$

$g\circ f:=h$ – непрерыная функция (т.к. композиция непрерывных функция), ограниченная(т.к. g ограниченная)

$$\mathbb{E}g(f(X_n))=\mathbb{E}h(X_n),\ \mathbb{E}g(f(X))=\mathbb{E}h(X)$$

Значит из утверждения выше

$$X_n\xrightarrow{d} X\Rightarrow\mathbb{E}h(X_n)\xrightarrow[n\to\infty]{}\mathbb{E}h(X)$$

Снова применяем утверждение

$$\mathbb{E}g(f(X_n))\xrightarrow[n\to\infty]{}\mathbb{E}g(f(X))\Rightarrow f(X_n)\xrightarrow{d} f(X)$$

■

## 6.2. Сходимость суммы и произведения сходящихся по распределению последовательностей случайных величин в случае, когда одна из предельных случайных величин постоянная.

**Лемма.** Пусть X, Y, Z случайные величины. Тогда $\forall t \in \mathbb{R}\ \forall \varepsilon > 0$ выполнено

$$P(X+Z \leqslant t-\varepsilon) - P(|Y-Z| \geqslant \varepsilon) \leqslant P(X+Y \leqslant t) \leqslant P(X+Z \leqslant t+\varepsilon) + P(|Y-Z| \geqslant \varepsilon)$$

*Доказательство.*

$$P(X+Y \leqslant t) \leqslant P(X+Y \leqslant t, |Y-Z| \leqslant \varepsilon) + P(X+Y \leqslant t, |Y-Z| \geqslant \varepsilon) \leqslant P(X+Y \leqslant t, |Y-Z| \leqslant \varepsilon) + P(|Y-Z| \geqslant \varepsilon)$$

Расскроем модуль

$$-\varepsilon \leqslant Y - Z \Rightarrow Z - \varepsilon \leqslant Y$$

Подставим вместо Y $Z - \varepsilon$

Событие $X+Y \leqslant t \cap |Y-Z| \geqslant \varepsilon$ вложено в событие $X+Z-\varepsilon \leqslant t$

$$\leqslant P(X+Z-\varepsilon \leqslant t) + P(|Y-Z| \geqslant \varepsilon)$$

Ищем другую оценку

Заменим в получившемся неравенстве Y на Z, Z на Y

$$P(X+Z \leqslant t) \leqslant P(X+Y-\varepsilon \leqslant t) + P(|Z-Y| \geqslant \varepsilon) =$$

$$= P(X+Y \leqslant t+\varepsilon) + P(|Y-Z| \geqslant \varepsilon)$$

Обозначим $t + \varepsilon := t$

$$P(X+Z \leqslant t-\varepsilon) \leqslant P(X+Y \leqslant t) + P(|Y-Z| \geqslant \varepsilon)$$

$$P(X+Y \leqslant t) \geqslant P(X+Z \leqslant t-\varepsilon) - P(|Y-Z| \geqslant \varepsilon)$$

■

**Теорема.** Если $X_n \xrightarrow{d} X$ и $Y_n \xrightarrow{d} C = const$ то

$$X_n + Y_n \xrightarrow{d} X + C$$

$$X_n \cdot Y_n \xrightarrow{d} X \cdot C$$

*Доказательство.* Вспомним доказательство того что

$$Y_n \xrightarrow{d} C = const \Rightarrow Y_n \xrightarrow{p} C$$

$$\lim_{n\to\infty} P(|X_n - C| \geqslant \varepsilon) = \lim_{n\to\infty} P(X_n - C \geqslant \varepsilon\ or\ -X_n + C \geqslant \varepsilon) \leqslant$$
$$\lim_{n\to\infty} P(X_n - C \geqslant \varepsilon) + \lim_{n\to\infty} P(X_n \leqslant C - \varepsilon) =$$
$$= 1 - F_{X_n}(\varepsilon + C) + F_{X_n}(C - \varepsilon) \underbrace{=}_{n\to\infty} 0$$

Используем лемму

$$P(X_n + C \leqslant t-\varepsilon) - P(|Y_n - C| \geqslant \varepsilon) \leqslant P(X_n + Y_n \leqslant t) \leqslant P(X_n + C \leqslant t+\varepsilon) + P(|Y_n - C| \geqslant \varepsilon)$$

$$F_{X_n}(t-\varepsilon-C) - P(|Y_n - C| \geqslant \varepsilon) \leqslant F_{X_n+Y_n}(t) \leqslant F_{X_n}(t+\varepsilon-C) + P(|Y_n - C| \geqslant \varepsilon)$$

1) $n \to \infty$

Заметим, что мы всегда можем выбрать точки $t-\varepsilon-C, t+\varepsilon-C$ в которых функция $F_X$ непрерывна, т.к. точек разрыва счетное количество, а $\varepsilon$ континуальная переменная.

Т.к. $Y_n \xrightarrow{p} C \Leftrightarrow_{def} \lim\limits_{n\to\infty} P(|Y_n - C| \geqslant \varepsilon) = 0$

$$F_X(t-\varepsilon-C) \leqslant \underline{lim}_{n\to\infty} F_{X_n+Y_n}(t) \leqslant \overline{lim}_{n\to\infty} F_{X_n+Y_n}(t) \leqslant F_X(t+\varepsilon-C)$$

2) $\varepsilon \to 0$

Заметим, что t - C точка непрерывности функции $F_X$ тогда и только тогда, когда t точка непрерывности функции $F_{X+C}$.

$$F_X(t-C) \leqslant \underline{\lim}_{n\to\infty} F_{X_n+Y_n}(t) \leqslant \overline{\lim}_{n\to\infty} F_{X_n+Y_n}(t) \leqslant F_X(t-C)$$

Так как слева и справа у нас одно и тоже значение $\Rightarrow \exists \lim\limits_{n\to\infty} F_{X_n+Y_n}(t) = F_X(t-C) = F_{X+C}(t)$

1) C = 0

$$\{|X_n \cdot Y_n| \geqslant \varepsilon\} \subset \{|X_n| > R\} \cup \{|Y_n| \geqslant \frac{\varepsilon}{R}\}$$

$$P(|X_n \cdot Y_n| \geqslant \varepsilon) \leqslant P(|X_n| > R) + P(|Y_n| \geqslant \frac{\varepsilon}{R})$$

$$P(|X_n| \geqslant R) = P(|X_n| \geqslant R) + P|(X_n| \leqslant -R) \leqslant P(|X_n| > \frac{R}{2}) + F_{X_n}(-R) = 1 - F_{X_n}(\frac{R}{2}) + F_{X_n}(-R)$$

$$P(|X_n \cdot Y_n| \geqslant \varepsilon) \leqslant P(|X_n| > R) + P(|Y_n| \geqslant \frac{\varepsilon}{R}) \leqslant 1 - F_{X_n}(\frac{R}{2}) + F_{X_n}(-R) + \underbrace{P(|Y_n - C(=0)| \geqslant \frac{\varepsilon}{R})}_{\xrightarrow{n\to\infty} 0(\text{т.к. сх-сть по вер.})}$$

a) $n \to \infty$

$$\overline{\lim}_{n\to\infty} P(|X_n \cdot Y_n \geqslant \varepsilon) \leqslant 1 - F_X(\frac{R}{2}) + F_X(-R) + 0$$

b) $R \to \infty$

R – точка непрерывности $F_X$

$$0 \leqslant \underline{\lim}_{n\to\infty} P(|X_n \cdot Y_n| \geqslant \varepsilon) \leqslant \overline{\lim}_{n\to\infty} P(|X_n \cdot Y_n| \geqslant \varepsilon) \leqslant 1 - F_X(\frac{R}{2}) + F_X(-R) \leqslant 0$$

$$\Rightarrow X_n \cdot Y_n \xrightarrow{p} 0 \Rightarrow_{\text{Лекция 1}} X_n \cdot Y_n \xrightarrow{d} 0$$

2) Общий случай

$$X_n Y_n = X_n(Y_n - C) + X_n C$$

$$X_n(Y_n - C) \xrightarrow{d} 0 \text{ по 1)}$$

$$CX_n \xrightarrow{d} CX$$

$$CX + 0 \xrightarrow{d} CX \text{ сумму разбирали выше}$$

■

### 6.3. Примеры применения: выборочная дисперсия и взаимосвязь с ЦПТ.

**Пример 1**(Выборочная дисперсия)

Пусть задана последовательность независимых и одинаково распределенных случайных величин $X_j$, причем $\mathbb{E}X_j = a$ и $\mathbb{D}Xj = \sigma^2$. Тогда последовательность случайных величин

$$s_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \overline{X_n})^2, \text{ где } \overline{X_n} = \frac{X_1 + \dots + X_n}{n}$$

Проверим это

$$\overline{X_n} \xrightarrow{p} a(\text{ЗБЧ})$$

$$s_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \overline{X_n})^2 = \frac{n}{n-1}\frac{1}{n}\sum_{j=1}^{n} X_j^2 + \frac{1}{n-1}\left(-2\underbrace{\sum_{j=1}^{n} X_j}_{n\overline{X_n}} \cdot \overline{X_n} + n\overline{X_n}^2\right) = \frac{n}{n-1}\left(\frac{1}{n}\sum_{j=1}^{n} X_j^2 - \overline{X_n}^2\right) (*)$$

$$\frac{1}{n}\sum_{j=1}^{n} X_j^2 \xrightarrow{p} \mathbb{E}X_1^2(\text{ЗБЧ})$$

$$\overline{X_n}^2 \xrightarrow{p} (\mathbb{E}X_1)^2$$

$$\frac{n}{n-1} \to 1$$

$$(*) \xrightarrow{p} \mathbb{D}X_1 = \sigma^2$$

$$\mathbb{E}s_n^2 = \frac{n}{n-1}\left(\frac{1}{n}\cdot n\cdot \mathbb{E}X_1^2 - \mathbb{E}(\overline{X_n})^2\right) =$$

$$\mathbb{E}(\overline{X_n})^2 = \mathbb{E}(\overline{X_n} - a + a)^2 = \mathbb{E}(\overline{X_n} - a)^2 + a^2 - 2a\underbrace{\mathbb{E}(\overline{X_n} - a)}_{0} = a^2 + \mathbb{D}\overline{X_n} = a^2 + \frac{1}{n^2}\mathbb{D}(X_1 + \ldots + X_n) = a^2 + \frac{\sigma^2}{n}$$

$$= \frac{n}{n-1}(\sigma^2 + a^2 - a^2 - \frac{\sigma^2}{n}) = \sigma^2$$

**Пример 2**(Взаимосвязь с ЦПТ)

Обозначения сохранятется с прошлого примера

Хотип показать, что

$$\frac{\sqrt{n}(\overline{X_n} - a)}{\sqrt{s_n^2}} \to Z \sim \mathcal{N}(0,1)$$

$$\frac{\sqrt{n}(\overline{X_n} - a)}{\sqrt{s_n^2}} = \frac{\sqrt{n}(\overline{X_n} - a)}{\sigma}\cdot\sqrt{\frac{\sigma^2}{s_n^2}}$$

$$\frac{\sqrt{n}(\overline{X_n} - a)}{\sigma} \to Z \sim \mathcal{N}(0,1)\text{из лекции 4}$$

$$\sqrt{\frac{\sigma^2}{s_n^2}} \xrightarrow{p} \sqrt{\frac{\sigma^2}{\sigma^2}} = 1(\text{Обсуждали выше})$$

Значит

$$\frac{\sqrt{n}(\overline{X_n} - a)}{\sqrt{s_n^2}} \to Z\cdot 1 \xrightarrow{d} Z \sim \mathcal{N}(0,1)$$

## 6.4. Теорема о сходимости последовательности вида $\frac{f(a+h_nX_n)-f(a)}{h_n}$ для сходящейся по распределению последовательности $X_n$.

**Теорема.** Пусть$a, h_n \in \mathbb{R}, h_n \to 0$и $f$ непрерывная на $\mathbb{R}$ и дифференцируемая в точке a функция. Если последовательность случайных величин $X_n \xrightarrow{d} X$, то

$$\frac{f(a+h_nX_n)-f(a)}{h_n} \xrightarrow{d} f^{'}(a)X$$

*Доказательство.* Введем функция

$$g(x) = \begin{cases} \frac{f(a+x)-f(a)}{x} & x \neq 0 \\ f^{'}(a) & x = 0 \end{cases}$$

g – непрерывная

$$h_n \xrightarrow{d} 0$$

$$X_n \xrightarrow{d} X$$

$$h_nX_n \xrightarrow{d} 0(\text{ теорема про произведения}) \Rightarrow$$

$$g(h_nX_n) \xrightarrow{d} g(0)(\text{первая теорема в билете6}) = f^{'}(a)$$

$$\frac{f(a+h_nX_n)-f(a)}{h_n} = X_n\cdot g(h_nX_n) = X_n\cdot\frac{f(a+h_nX_n)-f(a)}{h_nX_n} \xrightarrow{d} f^{'}(a)X(\text{ теорема про произведения})$$

■

### 6.5. Взаимосвязь с ЦПТ.

**Пример**

Обозначения сохранятется с прошлого примера

Пусть задана последовательность независимых и одинаково распределенных случайных величин $X_j$, причем $\mathbb{E}X_j = a$ и $\mathbb{D}X_j = \sigma^2 > 0$. Если f дифференцируемая функция, то

$$\sqrt{n}(f(\overline{X_n}) - f(a)) \xrightarrow{d} Y \sim \mathcal{N}(0, q^2),\ q = \sigma f^{'}(a)$$

Докажем это

Введем

$$Z_n = \frac{\sqrt{n}(\overline{X_n} - a)}{\sigma} \xrightarrow{d} Z \sim \mathcal{N}(0,1)$$

Тогда

$$\frac{\sqrt{n}(f(\overline{X_n}) - f(a))}{\sigma} = \frac{f(a + \frac{\sigma}{\sqrt{n}}Z_n) - f(a)}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{d} f^{'}(a)Z$$

$$\sqrt{n}(f(\overline{X_n}) - f(a)) = \sigma \cdot \frac{\sqrt{n}(f(\overline{X_n}) - f(a))}{\sigma} \xrightarrow{d} \sigma f^{'}(a)Z \sim \mathcal{N}(0, q^2)$$

## 7. Неравенство типа Хефдинга-Чернова. Пример применения.

### 7.1. Неравенство типа Хедфинга-Чернова

**Теорема** (Неравенство Хедфинга-чернова)**.** Пусть случайные величины $X_1, \dots, X_n$ независимы и $a_j \leqslant X_j \leqslant b_j$. Тогда для случайной величины $S_n :- X_1 + \dots + X_n$ и для каждого $t > 0$ выполнено

$$\mathrm{P}(S_n - \mathbb{E}S_n \geqslant t) \leqslant 2exp(-\frac{t^2}{4\sum_{j=1}^n (b_j - a_j)^2})$$

*Доказательство.* Пусть $Y_j = X_j - \mathbb{E}X_j$. Тогда $|Y_j| \leqslant b_j - a_j$, т.к. $X_j \in [a_j, b_j]$ и $\mathbb{E} \in [a_j, b_j]$. Заметим, что для каждого $\lambda > 0$

$$\mathrm{P}(\sum_{j=1}^n Y_j \geqslant t) = \mathrm{P}(e^{\lambda \sum_{j=1}^n Y_j} \geqslant e^{\lambda t}) \leqslant e^{-\lambda t}\mathbb{E}e^{\lambda \sum_{j=1}^n Y_j} = e^{-\lambda t}\prod_{j=1}^n \mathbb{E}e^{\lambda Y_j}$$

Оценим каждое ожидание из произведения:

$$\mathbb{E}e^{\lambda Y_j} = 1 + \lambda \mathbb{E}Y_j + \frac{1}{2}\lambda^2\mathbb{E}Y_j^2 + \sum_{k=3}^\infty \frac{1}{k!}\lambda^k\mathbb{E}Y_j^k \leqslant 1 + \frac{1}{2}\lambda^2(b_j - a_j)^2 + \sum_{k=3}^\infty \frac{1}{k!}\lambda^k(b_j - a_j)^k,$$

здесь мы использовали $\mathbb{E}Y_j = 0$. Докажем, что при $R > 0$ выполнена оценка

$$1 + \frac{1}{2}R^2 + \sum_{k=3}^\infty \frac{1}{k!}R^k \leqslant e^{R^2}$$

Действительно, если $R > 1$, то

$$1+\frac{1}{2}R^2+\sum_{k=3}^\infty \frac{1}{k!}R^k = 1+\frac{1}{2}R^2+\sum_{m=2}^\infty \frac{1}{m!}R^{2m}[\frac{m!}{(2m-1)!}R^{-1}+\frac{m!}{(2m)!}] \leqslant 1+\frac{1}{2}R^2+\sum_{m=2}^\infty \frac{1}{m!}R^{2m}[\frac{2}{m+1}] \leqslant 1+R^2+\sum_{m=2}^\infty \frac{1}{m!}R^{2m} = e^{R^2}.$$

если же $R \leqslant 1$, то

$$1 + \frac{1}{2}R^2 + \sum_{k=3}^\infty \frac{1}{k!}R^k \leqslant 1 + \frac{1}{2}R^2 + \sum_{k=3}^\infty \frac{1}{2^{k-1}}R^2 = 1 + R^2 \leqslant e^{R^2}.$$

Таким образом,

$$\mathrm{P}(\sum_{j=1}^n Y_j \geqslant t) \leqslant 2exp(-\lambda t + \lambda^2 \sum_{j=1}^n (b_j - a_j)^2).$$

Взяв $\lambda = \dfrac{t}{2\sum_{j=1}^{n}(b_j - a_j)^2}$, получим оценку

$$\mathrm{P}(S_n - \mathbb{E}S_n \geqslant t) \leqslant exp(-\frac{t^2}{4\sum_{j=1}^{n}(b_j - a_j)^2}).$$

Аналогично, рассматривая случайные величины $X_j^{'} :- -X_j$ получаем оценку

$$\mathrm{P}(-S_n + \mathbb{E}S_n \geqslant t) \leqslant exp(-\frac{t^2}{4\sum_{j=1}^{n}(b_j - a_j)^2}).$$

объединяя полученные неравенства получаем оценку из формулировки теоремы ■

**Теорема** (следствие)**.** Пусть $X_j$ $Bern(p)$−− набор независимых Бернуллевских случайных величин, $S_n = X_1 + \cdots + X_n$, тогда

$$\mathrm{P}(|\frac{S_n}{n} - p| \geqslant t) \leqslant 2e^{-\frac{nt^2}{4}}$$

### 7.2. Пример применения

**Пример**

Пусть в ящике какое-то кол-во черных и белых шаров. Каким должен быть размер выборки, чтобы оценить долю белых шаров с малой погрешностью? Пусть $\xi_j$ − − бернуллевская случайная величина, равная 1, если шар белого цвета и 0, если цвет черный. Мы хотим оценить вероятность успеха p. По нер-ву выше

$$\mathrm{P}(|\frac{S_n}{n} - p| \geqslant t) \leqslant 2e^{-\frac{nt^2}{4}} < \varepsilon$$

Тогда при размере выборки $n = O(\dfrac{log\varepsilon^{-1}}{t^2})$ выборочное среднее приближает реальную долю белых шаров с точностью t с вероятностью более $1 - \varepsilon$.

## 8. Многомерная характеристическая функция. Сходимость по распределению последовательности случайных векторов. Эквивалентное описание сходимости по распределению через сходимость характеристических функций (без доказательства). Независимость случайных величин в терминах характеристической функции совместного распределения. Матрица ковариаций, смысл задаваемой ей билинейной формы, ее изменение при линейных преобразованиях. Многомерная ЦПТ.

### 8.1. Многомерная характеристическая функция.

**Обозначение 1.** $\langle x, y\rangle = x_1 y_1 + \cdots + x_m y_m$, где $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$, $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$

**Определение 1.** Характеристическая функция случайного вектора $X = (X_1, \ldots, X_m)$ определяется равенством

$$\varphi_X(t) = \mathbb{E}(e^{i\langle X,t\rangle})$$

### 8.2. Сходимость по распределению последовательности случайных векторов.

**Определение 2.** Последовательность случайных векторов $X^n = (X_1^n, \ldots, X_m^n)$ **сходится по распределению** к случайному вектору $X = (X_1, \ldots, X_m)$, если для каждой непрерывной, ограниченной функции $g : \mathbb{R}^m \to \mathbb{R}$ выполнено $\mathbb{E}g(X^n) \to \mathbb{E}g(X)$ (обозначение $X^n \xrightarrow{d} X$).

### 8.3. Эквивалентное описание сходимости по распределению через сходимость характеристических функций (без доказательства).

**Теорема 0.2.** ***Без доказательства***

*Последовательность случайных векторов $X^n$ сходится по распределению к случайному вектору $X$ тогда и только тогда, когда $\varphi_{X^n}(y) \to \varphi_X(y)$ для каждого $y \in \mathbb{R}^m$.*

**Следствие. Без доказательства**
Если $\varphi_X = \varphi_Y$, то векторы X и Y имеют одинаковые распределения.

### 8.4. Независимость случайных величин в терминах характеристической функции совместного распределения.

**Теорема 0.3.** *Случайные величины $X_1, \dots, X_m$ независимы тогда и только тогда, когда*

$$\varphi_X(y_1, \dots, y_m) = \varphi_{X_1}(y_1) \cdot \dots \cdot \varphi_{X_m}(y_m) \ \forall y \in \mathbb{R}^m$$

*где $X = (X_1, \dots, X_m)$*

*Доказательство.*
$\Rightarrow$

$$\varphi_X(y_1, \dots, y_m) = \mathbb{E}e^{i(X_1y_1 + \dots + X_my_m)} = \mathbb{E}\prod_{i=1}^{m} e^{iX_iy_i} \underbrace{=}_{\text{незав.}} \prod_{i=1}^{m} \mathbb{E}e^{iX_iy_i} = \prod_{i=1}^{m} \varphi_{X_i}(y_i)$$

$\Leftarrow$
Зададим случаный вектор Y

- $Y = (Y_1, \dots, Y_m)$ – независимые компоненты
- $F_Y(x_1, \dots, x_m) := F_{X_1}(x_1) \cdot \dots \cdot F_{X_m}(x_m)$, т.е. $Y_j$ имеет такое же распределение как и $X_j$

Почему мы можем задать такой вектор?

- Произведение функций распределения – функция распределения
- По любой функции распределения можно построить случайный вектор
- У этого вектора компоненты независимы, т.к. функция совместного распределения распалась в произведение.

$$\varphi_Y(y) = \varphi_{Y_1}(y_1) \cdot \dots \cdot \varphi_{Y_m}(y_m) =$$

Т.к. независимость компоненты; но если непонятно, то можно посмотреть выше как это расписывается

$$= \varphi_{X_1}(y_1) \cdot \dots \cdot \varphi_{X_m}(y_m) =$$

Т.к. $Y_j$ сходится по распределению к $X_j$, то хар.фукнии тоже сходятся(Лекция 3, теорема 5)

$$= \varphi_X(y) \text{ (см.условие)}$$

Получили

$$\varphi_X(y) = \varphi_Y(y) \ \forall y \Rightarrow F_x = F_y \text{(см. следствие выше)}$$

Если совпадают функции распределения, то и свойства независимости совпадают. Значит компоненты X тоже независимы. ■

### 8.5. Матрица ковариаций, смысл задаваемой ей билинейной формы, ее изменение при линейных преобразованиях.

**Определение 3.** Пусть $X = (X_1, \dots, X_m)$ случайный вектор. Матрица $R_X$ с компонентами $r_{kj} := cov(X_k, X_j)$ называется ковариационной матрицей вектора X.

**Теорема 0.4.** *Симметричная неотрицательно определенная матрица $R$ является ковариационной матрицей случайного вектора $X$ тогда и только тогда, когда*

$$\langle Rx, y\rangle = cov(\langle x, X\rangle, \langle y, X\rangle) = \mathbb{E}(\langle x, X - a\rangle, \langle y, X - a\rangle)$$

*, где $a = (a_1, \dots, a_m)$ вектор средних, т.е. $a = \mathbb{E}X_j$*

*Доказательство.*

$$e_k = (0, \dots 0, \underbrace{1}_{k}, \dots 0)$$

$$e_j = (0, \dots 0, \underbrace{1}_{j}, \dots 0)$$

$$x = \sum_k x_k e_k;\ y = \sum j_j e_j$$

$$\langle \sum_k x_k R_x e_k, \sum_j y_j e_j \rangle = \sum_k \sum_j \underbrace{\langle R_x e_k, e_j \rangle}_{(def) cov(x_k, x_j)} x_k y_j (*)$$

$$cov(x_k, x_j) = cov(\langle x, e_k \rangle, \langle x, e_j \rangle)$$

$$(*) = \sum_k \sum_j (cov(\langle x, e_k \rangle, \langle y, e_j \rangle)) x_k y_j = \sum_k \sum_j (cov(\langle x, x_k e_k \rangle, \langle x, y_j e_j \rangle)) =$$

$$= \sum_k (cov(\langle x, x_k e_k \rangle, \sum_j \langle x, y_j e_j \rangle)) = cov(\langle x, x \rangle, \langle x, y \rangle)$$

■

**Теорема 0.5.** *Пусть $X$ – случайный вектор с ковариационной матрицей, тогда случайный вектор $AX + b$ имеет ковариационную матрицу $ARA^*$*

*Доказательство.*

$$Y = AX + b$$

$$\langle R_y u, v \rangle = cov(\langle AX, u \rangle + \langle b, u \rangle, \langle AX, v \rangle + \langle b, v \rangle) =_* cov(\langle AX, u \rangle, \langle AX, v \rangle) = cov(\langle X, A^* u \rangle, \langle X, A^* v \rangle) =_{**}$$

* – сдвиг на константу на ковариацию не влиятет

** – см. теорему выше

$$= \langle R_x A^* u, A^* v \rangle = \langle A R_x A^* u, v \rangle$$

$$R_y = A R_x A^*$$

■

### 8.6. Многомерная ЦПТ.

**Теорема 0.6.** *Пусть случайные вектора $X^n = (X_1^n, \dots, X_m^n)$ независимы, одинаково распределены и имеют конечные $a_j = \mathbb{E} X_j^n, r_{k,j} = cov(X_k^1, X_j^1)$*
*Тогда последовательность случайных векторов $Y^n = (Y_1^n, \dots, Y_m^n)$ с компонентами*

$$Y_j^n = \frac{X_j^1 + \ldots + X_j^n - n a_j}{\sqrt{n}}$$

*сходится по распределению к вектору $Z$, характеристическая функция, которого имеет вид*

$$\varphi_Z(y) = e^{-\frac{1}{2} \langle R_y, y \rangle}, R = r_{k,j}$$

*Доказательство.*
Фиксируем $y \in \mathbb{R}^m$
Рассмотри последовательность случайных величин

$$\xi_n := \frac{\langle X^1, y \rangle + \ldots + \langle X^n, y \rangle - n \langle a, y \rangle}{\sqrt{n}} = \langle Y^n, y \rangle$$

- $\{X^i, y\}$ независимы, одинаковы распределенные
- $\mathbb{E}(\langle X^1, y \rangle) = \langle a^1, y \rangle$

Значит по одномерной цпт

$$\xi_n \xrightarrow{d} Z_y \sim \mathcal{N}(0, \mathbb{D}\langle X^1, y\rangle)$$

$$\varphi_{\xi_n}(t) \to \varphi_{Z_y} = e^{-\frac{1}{2}t^2\mathbb{D}\langle X^1, y\rangle}$$

Заметим что

$$\varphi_{Y^n}(y) = \mathbb{E}e^{i\langle Y^n, y\rangle}$$

$$\varphi_{\langle Y^n, y\rangle}(1) = \mathbb{E}e^{i\cdot 1\cdot\langle Y^n, y\rangle}$$

$$\Rightarrow \varphi_{Y^n}(y) = \varphi_{\langle Y^n, y\rangle}(1) = \varphi_{\xi_n}(1) \to_* \varphi_{Z_y}(1) = e^{-\frac{1}{2}\mathbb{D}\langle X^1, y\rangle}$$

$*$ – т.к. $\xi_n \xrightarrow{d} Z_y$

$$\mathbb{D}\langle X^1, y\rangle = cov(\langle X^1, y\rangle, \langle X^1, y\rangle) = \langle R_Y, y\rangle$$

Получили

$$\varphi_{Y^n} \to e^{-\frac{1}{2}\langle R_Y, y\rangle} \Rightarrow Y^n \xrightarrow{d} Z$$

■

# 9. Многомерное нормальное распределение. Свойства нормального вектора: линейный образ нормального распределения нормален, характеризация через одномерные распределения, значение параметров нормального вектора, равносильность независимости и некоррелированности компонент. Представление нормального вектора, как линейный образ стандартного нормального вектора, ортогонализация. Плотность нормального вектора.

## 9.1. Многомерное нормальное распределение.

**Определение** Случайный вектор X имеет нормальное распределение или является гауссовским, если $\varphi_x(y) = E[exp(i < X, y >)] = e^{-\frac{1}{2}<Ry,y>+i<a,y>}$.
Где $a = (a_1, ..., a_m) \in R^m, R-$ симметричная неотрицательно определенная $m \times m$ матрица. Далее пишем $X \sim N(a, R)$.

## 9.2. Свойства нормального вектора: линейный образ нормального распределения нормален, характеризация через одномерные распределения, значение параметров нормального вектора, равносильность независимости и некоррелированности компонент.

**Предложение 1.** $X \sim N(a, R)$, то вектор $AX + b \sim N(Aa + b, ARA^*)$
*Доказательство.* Доказывается простой подстановкой по определению:
$\varphi_{AX+b}(y) = E[exp(i < AX + b, y >)] = e^{i<b,y>}E[exp(i < X, A^*y >)] = e^{i<b,y>+i<a,A^*y>-\frac{1}{2}(RA^*y,A^*y)}$
Остается заметить, что $< RA^*y, A^*y >=< ARA^*y, y >$, а также $i < b, y > +i < a, A^*y >= i < Aa + b, y >$.

**Теорема 1.** Вектор X имеет нормальное распределение тогда и только тогда, когда для каждого вектора y случайная величина $< y, X >$ имеет нормальное распределение.
*Доказательство.* Так как по условию $X \sim N(a, R)$, то:

$$\varphi_{<X,y>}(t) = E[exp(it < X, y >)] = e^{-\frac{1}{2}t^2<Ry,y>+it<a,y>}$$

Иными словами, $< X, y >\sim N(< a, y >, < Ry, y >$, поскольку $< a, y >= E[< X, y >], < Ry, y >= D[< X, y >]$.
Докажем обратно:

$$\varphi_X(y) = Ee^{i<X,y>} = \varphi_{<X,y>}(1) = e^{-\frac{1}{2}D[<X,y>]+e[<X,y>]} = e^{-\frac{1}{2}t^2<Ry,y>+it<a,y>}$$

Не забываем, что R - ковариационная матрица X, a - вектор средних.

**Следствие 1.** Если $X \sim N(a, R)$, то R - ковариационная матрица X, a - вектор средних.

**Следствие 2.** Если вектор $(X_1, X_2)$ имеет нормальное распределение и $cov(X_1, X_2) = 0$, то величины $X_1$ и $X_2$ независимы.

*Доказательство* Если ковариация равна 0, то у нас независимость дисперсий:

$$\varphi_{(X_1,X_2)}(y_1, y_2) = e^{-\frac{1}{2}(y_1^2 DX_1 + y_2^2 DX_2 + i(y_1 EX_1 + y_2 EX_2)} = \varphi_{X_1}(y_1)\varphi_{X_2}(y_2)$$

### 9.3. Представление нормального вектора, как линейный образ стандартного нормального вектора, ортогонализация.

**Следствие 3.** Если X $\sim$ N (a, R), то найдется такая матрица A, что X = AZ + a, где $Z = (Z_1, ..., Z_k)$ и случайные величины Zj независимы и имеют стандартное нормальное распределение, $AA^* = R$.

*Доказательство.* Перейдем к случайным векторам $X_j' = X_j - a_j$. Нам надо найти ортонормированный базис в линейном пространстве $span(X_1', ..., X_m')$ со скалярным произведением $(X, Y) = E[XY]$. Для этого будем использовать метод Грамма-Шмидта. После него мы получаем случайные величины $(Z_1, ..., Z_k) = Z$, что линейно выражаются через $X_1', ...X_m'$. Т.е. в частности вектор Z - нормальный и $E[Z_j] = 0$, кроме того система является ортонормированным базисом в $span(X_1', ..., X_m')$. То, что это базис означает X = AZ. Ортономированность означает $cov(Z_k, Z_j) = E[Z_k, Z_j) = 0, D[Z_j] = E[Z_j^2] = 1$. Поэтому случайные величины $Z_j \sim N(0, 1)$ и независимы. Равенство $AA^* = R$ следует из того, как меняется матрица при линейных отображениях.

### 9.4. Плотность нормального вектора.

**Теорема 2.** Если $X \sim N(a, R)$ и $detR \neq 0$, то случайны вектор X имеет плотность:

$$\rho(x) = \frac{1}{(2\pi)^{\frac{m}{2}}\sqrt{detR}} e^{-\frac{1}{2}<R^{-1}(x-a), x-a>}$$

*Доказательство.* Поскольку можем представить X как $AZ + a$, где A матрица квадртная и невырожденная(иначе R $= AA^*$ вырождена), то:

$P(X \in B) = P(AZ + a \in B) = \dfrac{1}{(\pi)^{\frac{m}{2}}} \displaystyle\int\limits_{Ax+a\in B} e^{-\frac{1}{2}|x|^2} dx = \dfrac{1}{(\pi)^{\frac{m}{2}} detA} \displaystyle\int\limits_{B} e^{-\frac{1}{2}|A^{-1}(y-a)|^2} dy$

Осталось заметить, что $|A^{-1}(y - a)|^2 =< A^{-1}(y - a), A^{-1}(y - a) >=< (AA)^{-1}(y - a), (y - a) >$ и $(detA)^2 = detR$.

## 10. Условное математическое ожидание в дискретном случае относительно разбиения и относительно случайной величины. Основные свойства: линейность, монотонность, формула полной вероятности, условное ожидание величины, независимой с разбиением, вынесение случайной величины из под знака условного ожидания. Эквивалентное определение условного математического ожидания и геометрическая интерпретация.

### 10.1. Условное математическое ожидание в дискретном случае относительно разбиения и относительно случайной величины.

**Определение 4.** Величину $\Lambda$ называют условным математическим ожиданием $X$ относительно разбиения $\boldsymbol{\beta}$ и обозначают через $\mathbb{E}(X|\boldsymbol{\beta})$.

**Определение 5.** Рассмотрим случай, когда разбиение $\boldsymbol{\beta}$ появляется посредством некоторой случайной величины $Y = \displaystyle\sum_{k=1}^{n} y_k I_{B_k}$, где $y_k$ - различные числа и $P(B_k) > 0$. В этом случае $B_k = \{\omega : Y(\omega) = y_k\}$ и условное математическое ожидание $\mathbb{E}(X|\boldsymbol{\beta})$ обозначают символом $\mathbb{E}(X|Y)$ и называют условным математическим ожиданием случайной величины X относительно случайной величины Y.

## 10.2. Основные свойства: линейность, монотонность, формула полной вероятности, условное ожидание величины, независимой с разбиением, вынесение случайной величины из под знака условного ожидания.

**Теорема 0.7.** *Имеют место следующие свойства условного математического ожидания:*
*(i) (линейность)* $\mathbb{E}(\alpha X + \beta Y|\boldsymbol{\beta}) = \alpha\mathbb{E}(X|\boldsymbol{\beta}) + \beta\mathbb{E}(Y|\boldsymbol{\beta})$,
*(ii) (монотонность)* $X \leqslant Y$ *п.н.* $\implies \mathbb{E}(X|\boldsymbol{\beta}) \leqslant \mathbb{E}(Y|\boldsymbol{\beta})$,
*(iii) (аналог формулы полной вероятности)* $\mathbb{E}(\mathbb{E}(X|\boldsymbol{\beta})) = \mathbb{E}X$,
*(iv) (независимость) если случайная величина* $X$ *не зависит от разбиения* $\boldsymbol{\beta}$*, т.е. случайные величины* $X$ *и* $I_{B_k}$ *независимы для каждого* $k$*, то* $\mathbb{E}(X|\boldsymbol{\beta}) = \mathbb{E}X$.
*(v) для всякой случайной величины* $Z = \sum\limits_{k=1}^{n} c_k I_{B_k}$ *выполнено* $\mathbb{E}(ZX|\boldsymbol{\beta}) = Z\mathbb{E}(X|\boldsymbol{\beta})$.

*Доказательство.* Доказательство. Свойства (i) и (ii) следуют из того, что они верны для $\mathbb{E}(X|B_k)$ для каждого $k$ (т.к. они верны для математического ожидания относительно произвольной вероятностной меры).
Свойство (iii) проверяется непосредственной подстановкой в определение: $\mathbb{E}(\mathbb{E}(X|\boldsymbol{\beta})) = \mathbb{E}\left(\sum\limits_{k=1}^{n} I_{B_k}\dfrac{\mathbb{E}(XI_{B_k})}{P(B_k)}\right) = \sum\limits_{k=1}^{n}\mathbb{E}(XI_{B_k}) = \mathbb{E}X$.
Обоснуем пункт (iv). Так как $X$ и $I_{B_k}$ независимы, то $\mathbb{E}(X|B_k) = \dfrac{\mathbb{E}(XI_{B_k})}{P(B_k)} = \dfrac{\mathbb{E}X\mathbb{E}I_{B_k}}{P(B_k)} = \mathbb{E}X$.
Следовательно, $\mathbb{E}(X|\boldsymbol{\beta}) = \sum\limits_{k=1}^{n} I_{B_k}\mathbb{E}(X|B_k) = \sum\limits_{k=1}^{n} I_{B_k}\mathbb{E}X = \mathbb{E}X$.
Для обоснования (v) достаточно заметить, что $\mathbb{E}(XZ|B_k) = \dfrac{\mathbb{E}(XZI_{B_k})}{P(B_k)} = c_k\dfrac{\mathbb{E}(XI_{B_k})}{P(B_k)} = c_k\mathbb{E}(X|B_k)$.
ч.т.д. ■

**Теорема 0.8.** *В случае, когда мы рассматриваем условное ожидание относительно случайной величины, свойства следует формулировать так:*
*(i) (линейность)* $\mathbb{E}(\alpha X + \beta Y|Z) = \alpha\mathbb{E}(X|Z) + \beta\mathbb{E}(Y|Z)$,
*(ii) (монотонность)* $X \leqslant Y$ *п.н.* $\implies \mathbb{E}(X|Z) \leqslant \mathbb{E}(Y|Z)$,
*(iii) (аналог формулы полной вероятности)* $\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}X$,
*(iv) (независимость) если случайные величины* $X$ *и* $Y$ *независимы, то* $\mathbb{E}(X|Y) = \mathbb{E}X$.
*(v) для всякой случайной величины* $Z = g(Y)$ *выполнено* $\mathbb{E}(ZX|Y) = Z\mathbb{E}(X|Y)$.

## 10.3. Эквивалентное определение условного математического ожидания и геометрическая интерпретация.

Для условного математического ожидания выполнено $\mathbb{E}(g(Y)X) = \mathbb{E}(g(Y)\mathbb{E}(X|Y))$ для произвольной функции $g$. Кроме того, если для какой-то случайной величины вида $Z = f(Y)$ выполнено $\mathbb{E}(g(Y)X) = \mathbb{E}(g(Y)Z)$ для произвольной функции $g$, то $Z = \mathbb{E}(X|Y)$ п.н.

*Доказательство.* По уже доказанному $\mathbb{E}(g(Y)\mathbb{E}(X|Y)) = \mathbb{E}(\mathbb{E}(g(Y)X|Y))) = \mathbb{E}(g(Y)X)$. Наоборот, если $Z = f(Y)$ и обладает указанным свойством, то $\mathbb{E}(g(Y)\mathbb{E}(X|Y)) = \mathbb{E}(g(Y)Z)$ для произвольной $g$. Т.к. $\mathbb{E}(X|Y)$ также имеет вид $h(Y)$, то, взяв $g = f - h$, получаем $\mathbb{E}|\mathbb{E}(X|Y) - Z|^2 = 0$, что даёт равенство $Z = \mathbb{E}(X|Y)$ почти наверное. ■

**Предложение.** Условное математическое ожидание $\mathbb{E}(X|Y)$ среди всех случайных величин вида $g(Y)$ является лучшим среднеквадратическим приближением для $X$, т.е. $\underset{Z:Z=g(Y)}{min}\mathbb{E}|X - Z|^2 = \mathbb{E}|X - \mathbb{E}(X|Y)|^2$.

*Доказательство.* Пусть $Z = g(Y)$. Так как по предыдущей лемме $\mathbb{E}[(X - \mathbb{E}(X|Y))(\mathbb{E}(X|Y) - Z) = 0]$, то $\mathbb{E}|X - Z|^2 = \mathbb{E}|(Z - \mathbb{E}(X|Y)) + (\mathbb{E}(X|Y) - Z)|^2 = \mathbb{E}|X - \mathbb{E}(X|Y)|^2 + \mathbb{E}|\mathbb{E}(X|Y) - Z|^2 \geqslant \mathbb{E}|X - \mathbb{E}(X|Y)|^2$.
ч.т.д. ■

Таким образом, с геометрической точки зрения условное математическое ожидание является проекцией X на пространство случайных величин вида $g(Y)$ и полностью характеризуется тем свойством, что вектор $X - \mathbb{E}(X|Y)$ ортогонален указанному пространству, что записывается с помощью равенства $\mathbb{E}(Xg(Y)) = \mathbb{E}(\mathbb{E}(X|Y)g(Y))$ для произвольной случайной величины $g(Y)$.

# 11. Условное математическое ожидание в общем случае: определение и свойства. Формула для вычисления условного математического ожидания при известной плотности совместного распределения, условная плотность. Аналог фомрулы Байеса.

## 11.1. Условное математическое ожидание в общем случае: определение и свойства.

Случайная величина вида $f(Y)$ называется **условным математическим ожиданием** случайной величины X(обладающей математическм ожиданием) относительно случайной велечины Y и обозначается как $E[X|Y]$, если:

$$E(Xg(Y)) = E(E(X|Y)g(Y))$$

для всех ограниченных случайных величин $g(Y)$. Любые две случайные величины, удолетворяющие этому определению почти наверное совпадают.
Функцию f(y) обозначают как $E(X|Y = y)$ и трактуют как условное математическое ожидание X при условии Y = y. Надо иметь в виду, что именно f(Y) определено однозначно, но не f. Однако различные функции f совпадают почти наверное относительно распределения $m_y$. Если Y имеет положительную непрерывную плотность, то различные функции f совпадают почти всюду. В дальнейшем, если мы пишем $E(X|Y = y)$, то мы имеем в виду $E(X|Y) = f(Y)$.

**Предложение 1.** Сформулированные ранее свойства (i) - (v) условных математических ожиданий для дискретных величин остаются верными и в общем случае.
*Доказательство.* Линейность ясна из определения линейности и линейности математического ожидания.
Для доказательства монотонности достаточно в силу линейности показать, что изи $X \geqslant 0$ следует $E(X|Y) \geqslant 0$ почти наверное. Для этого в определении положим $g(Y) = 1 - sign(E(X|Y)) \geqslant 0$. Тогда $E(X|Y) - |E(X|Y)| \leqslant 0$, но:

$$E[E(X|Y) - |E(X|Y)|] = E[X(1 - sign(E(X|Y)))] \geqslant 0$$

Значит $E(X|Y) - |E(X|Y)| = 0$.
Равенство $E(E(X|Y)) = EX$ является частным случаем определения($g(Y) = 1$).
Если $X$ и $Y$ независимы, то $E(Xg(Y)) = [EX] \cdot [Eg(Y)] = E([EX] \cdot [Eg(Y)])$.
Если $Z = h(y)$(с ограниченной h), то подстановкой в определение проверяется, что $ZE(X|Y)$ является условным математическим ожиданием ZX относительно Y. Случай для общей функции h получается с помощью предельного перехода.

## 11.2. Формула для вычисления условного математического ожидания при известной плотности совместного распределения, условная плотность.

**Предложение 2.** Предположим, что распределение $(X, Y)$ задано совместной плотностью $\rho(x,y)_{X,Y}$. Тогда

$$E[(X, Y)|Y = y] = \int\limits_{-\infty}^{\infty} (x, y)\frac{\rho(x,y)_{X,Y}}{\rho(y)_Y}dx$$

*Доказательство.* Имеет место цепочка неравенств:
$$E[(X, Y)g(Y)] = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (x, y)g(y)\rho(x,y)_{X,Y}dxdy =$$
$$= \int\limits_{-\infty}^{\infty} g(y)(\int\limits_{-\infty}^{\infty} (x, y)\frac{\rho(x,y)_{X,Y}}{\rho_Y(y)}dx)\rho_Y(y)dy$$
Функцию $\rho_{X|Y}(x|y) = \dfrac{\rho_{X,Y}(x, y)}{\rho_Y(y)}$ называют условной плотностью X относительно Y(условимся, что она равно 0 в

точках y, в которых плотность $\rho_y(y) = 0$). Таким образом верны равенства:

$$E(X|Y = y) = \int\limits_{-\infty}^{\infty} x\rho_{X|Y}(x|y)dx, \rho_{X,Y}(x,y) = \rho_{X|Y}(x|y)\rho_Y(y)$$

Последнее из которых является знакомым нам аналогом $P(A \cap B) = P(A|B)P(B)$.

### 11.3. Аналог фомрулы Байеса.

Пусть X и Y - такие случайные величины, что существует измеримая функция $\rho(x|y)$, для которой выполнено:

$$P(X \in B|Y = y) = \int\limits_B \rho(x|y)dx$$

В этом случае

$$E(h(X)|Y = y) = \int\limits_R h(x)\rho(x|y)dx$$

Заметим, что для произвольной ораниченной функции h

$$Eh(X) = E(E(h(X)|Y)) = \int\limits_R h(x)E\rho(x|Y)dx$$

Тем самым $\rho_X(x) =)E\rho(x|Y)$. Для произвольных ограниченных функция $f, g$ выполнено:

$$E[f(X)E(g(Y)|X))] = E[f(X)g(Y)] = Eg(Y)E(f(X)|Y)$$

Левая часть тождества равна

$$\int\limits_R f(x)E[g(y)|X = x)\rho_X(x)dx$$

А правая

$$\int\limits_R f(x)E[g(y)|\rho_X(x))dx$$

В силу произвольности f получаем следующую формулу Байеса:

$$E(g(Y)|X = x) = \frac{E[g(Y)\rho(x|Y)}{E\rho(x|Y)}dx$$

Теперь пусть Y принимает значения 0 и 1 с вероятность p и q соотвественно. Тогда:
$P(Y = 0|X = x) = \dfrac{p\rho(x|0)}{p\rho(x|0) + q\rho(x|1)}, P(Y = 1|X = x) = \dfrac{q\rho(x|1)}{p\rho(x|0) + q\rho(x|1)}$