

# Теория вероятностей и математическая статистика, Коллоквиум III

Версия от 04.04.2021 16:50

## 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. . . . .	3
1.1.	Неравенство Чебышева и закон больших чисел в слабой форме для общих случайных величин. .	3
1.2.	Усиленный закон больших чисел Колмогорова (б/д). . . . .	4
1.3.	Сходимости случайных величин: почти наверное и по вероятности. . . . .	4
1.4.	Взаимосвязь сходимостей по вероятности и почти наверное. . . . .	4
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. . . . .	5
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. . . . .	5
4.	Характеристические функции: определение и свойства. Вычисление характеристической функции нормальной случайной величины. Производные характеристических функций. . . . .	5
4.1.	Характеристические функции: определение и свойства. . . . .	5
4.2.	Вычисление характеристической функции нормальной случайной величины. . . . .	7
4.3.	Производные характеристических функций. . . . .	7
5.	Переформулировка сходимости по распределению в терминах характеристических функций. Однозначность задания распределения случайной величины ее характеристической функцией. Центральная предельная теорема. . . . .	8
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 $\frac{f(a + h_n X_n) - f(a)}{h_n}$ для сходящейся по распределению последовательности $X_n$ . Взаимосвязь с ЦПТ. . .	8
6.1.	Подстановка сходящейся по распределению последовательности случайных величин в непрерывную функцию. . . . .	8
6.2.	Сходимость суммы и произведения сходящихся по распределению последовательностей случайных величин в случае, когда одна из предельных случайных величин постоянная. . . . .	9
6.3.	Примеры применения: выборочная дисперсия и взаимосвязь с ЦПТ. . . . .	11
6.4.	Теорема о сходимости последовательности вида $\frac{f(a + h_n X_n) - f(a)}{h_n}$ для сходящейся по распределению последовательности $X_n$ . . . . .	12
6.5.	Взаимосвязь с ЦПТ. . . . .	12
7.	Неравенство типа Хефдинга-Чернова. Пример применения. . . . .	12
7.1.	Неравенство типа Хефдинга-Чернова . . . . .	12
7.2.	Пример применения . . . . .	13

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. . . . .	14
8.1.	Многомерная характеристическая функция. . . . .	14
8.2.	Сходимость по распределению последовательности случайных векторов. . . . .	14
8.3.	Эквивалентное описание сходимости по распределению через сходимость характеристических функций (без доказательства). . . . .	14
8.4.	Независимость случайных величин в терминах характеристической функции совместного распределения. . . . .	14
8.5.	Матрица ковариаций, смысл задаваемой ей билинейной формы, ее изменение при линейных преобразованиях. . . . .	15
8.6.	Многомерная ЦПТ. . . . .	16
9.	Многомерное нормальное распределение. Свойства нормального вектора: линейный образ нормального распределения нормален, характеристика 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. . . . .	16
9.1.	Многомерное нормальное распределение. . . . .	16
9.2.	Свойства нормального вектора: линейный образ нормального распределения нормален, характеристика через одномерные распределения, значение параметров нормального вектора, равносильность независимости и некоррелированности компонент. . . . .	17
9.3.	Представление нормального вектора, как линейный образ стандартного нормального вектора, ортогонализация. . . . .	17
9.4.	Плотность нормального вектора. . . . .	17
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. . . . .	18
10.1.	Условное математическое ожидание в дискретном случае относительно разбиения и относительно случайной величины. . . . .	18
10.2.	Основные свойства: линейность, монотонность, формула полной вероятности, условное ожидание величины, независимой с разбиением, вынесение случайной величины из под знака условного ожидания. . . . .	18
10.3.	Эквивалентное определение условного математического ожидания и геометрическая интерпретация. . . . .	19
11.	Условное математическое ожидание в общем случае: определение и свойства. Формула для вычисления условного математического ожидания при известной плотности совместного распределения, условная плотность. Аналог формулы Байеса. . . . .	19

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  $X \geq 0$  почти наверное. Тогда для любого  $t > 0$  выполняется

$$P[X \geq t] \leq \frac{E[X]}{t}.$$

*Доказательство.* Заметим, что для любого  $t > 0$  выполняется  $t \cdot I[x \geq t] \leq X$  почти наверное (здесь  $I$  это индикатор), так как в левой части будут учтены  $t \leq X$ , с суммарным коэффициентом не больше 1.

Возьмем математическое ожидание от обеих сторон и получим то, что нас просили:

$$t \cdot I[x \geq t] \leq X \iff t \cdot P[x \geq t] \leq E[X] \iff P[x \geq t] \leq \frac{E[X]}{t}.$$

■

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

$$P[|X - E[X]| \geq \varepsilon] \leq \frac{D[X]}{\varepsilon^2}.$$

*Доказательство.* Для доказательства рассмотрим случайную величину  $Y = |X - E[X]|^2$  и применим неравенство Маркова.

Для любого  $\varepsilon$  выполняется

$$P[Y \geq \varepsilon^2] \leq \frac{E[Y]}{\varepsilon^2} \iff P[|X - E[X]|^2 \geq \varepsilon^2] \leq \frac{D[X]}{\varepsilon^2} \iff P[|X - E[X]| \geq \varepsilon] \leq \frac{D[X]}{\varepsilon^2}.$$

■

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

Обозначим  $E[X_n] = a_n$  и  $D[X_n] = \sigma_n^2$ . Если

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

то для всякого  $\varepsilon > 0$  выполняется

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

*Доказательство.* Рассмотрим случайную величину  $X = \frac{X_1 + \dots + X_n}{n}$ .

По линейности математического ожидания получаем

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

Теперь необходимо найти дисперсию случайной величины  $X$ :

- Константа из дисперсии выносится с возведением в квадрат, поэтому

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

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

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

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

$$P[|X - E[X]| \geq \varepsilon] \leq \frac{D[X]}{\varepsilon^2} \iff P\left[\left|\frac{X_1 + \dots + X_n}{n} - \frac{a_1 + \dots + a_n}{n}\right| \geq \varepsilon\right] \leq \frac{\sigma_1^2 + \dots + \sigma_n^2}{n^2 \varepsilon^2}$$

■

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 ожидания и одна и та же математическая дисперсия:  $E[X_n] = a$  и  $D[X_n] = \sigma^2$ .

Тогда дисперсия среднего арифметического  $\frac{D[X_1] + \dots + D[X_n]}{n^2} = \frac{\sigma^2}{n}$  стремится к нулю и получаем

$$P\left[\left|\frac{X_1 + \dots + X_n}{n} - a\right| \geq \varepsilon\right] \rightarrow 0.$$

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 ожидания и пусть  $E[X_n] = a$ .

Тогда

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

Заметьте, что мы не требуем наличия второго момента, в отличие ЗБЧ в слабой форме. Также, эта сходимости 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  $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P[|X_n - X| \geq \varepsilon] = 0.$$

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

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

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

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 сходимости 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  $P[|X_n - X| > \varepsilon] \rightarrow 0$ , что равносильно  $P[|X_n - X| < \varepsilon] \rightarrow 1$ , что мы и будем доказывать.

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  $P[\lim X_n = X] = 1$ , поэтому

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

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

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

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

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

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

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

■

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.**
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.**
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

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

**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  $|\varphi_X(t)| \leq 1$ .

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:

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

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:

$$\varphi_X(0) = \mathbb{E}[e^{i \cdot 0 \cdot X}] = \mathbb{E}[e^0] = \mathbb{E}[1] = 1.$$

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

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

Значение характеристической функции это комплексное число. Квадрат модуля комплексного числа это сумма квадратов его мнимой и действительной частей:

$$|\varphi_X(t)|^2 = (\mathbb{E}[\cos(t \cdot X)])^2 + (\mathbb{E}[\sin(t \cdot X)])^2.$$

С помощью знаний о  $\mathbb{E}[Y^2] \geq (\mathbb{E}[Y])^2$  оценим квадрат модуля характеристической функции:

$$|\varphi_X(t)|^2 = (\mathbb{E}[\cos(t \cdot X)])^2 + (\mathbb{E}[\sin(t \cdot X)])^2 \leq \mathbb{E}[\cos^2(t \cdot X)] + \mathbb{E}[\sin^2(t \cdot X)] = \mathbb{E}[\cos^2(t \cdot X) + \sin^2(t \cdot X)] = \mathbb{E}[1] = 1.$$

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  $\mathbb{E}[y \cdot X] = y \cdot \mathbb{E}[X]$  по линейности математического ожидания.

Запишем по определению:

$$\varphi_{aX+b}(t) = \mathbb{E}[e^{it \cdot aX + it \cdot b}] = \mathbb{E}[e^{it \cdot aX} \cdot e^{it \cdot b}] = e^{it \cdot b} \cdot \mathbb{E}[e^{it \cdot aX}] = e^{it \cdot b} \cdot \varphi_{aX}(t).$$

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  $\mathbb{E}[Y_1 \cdot \dots \cdot Y_n] = \mathbb{E}[Y_1] \cdot \dots \cdot \mathbb{E}[Y_n]$ .

Запишем по определению:

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

■

## 4.2. Вычисление характеристической функции нормальной случайной величины.

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

Запишем по определению:

$$\varphi_\xi(t) = \mathbb{E}[e^{it\xi}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \cos(tx) \exp[-x^2/2] dx + \frac{i}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \sin(tx) \exp[-x^2/2] dx.$$

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

$$\varphi_\xi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \cos(tx) \exp[-x^2/2] dx.$$

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

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

Пришли к дифференциальному уравнению:

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

Интегрируем обе части:

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

Теперь берем экспоненту от обеих частей:

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

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:

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

## 4.3. Производные характеристических функций.

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

$$\varphi_X^{(k)}(0) = i^k \cdot \mathbb{E}[X^k].$$

*Доказательство.* Докажем для  $k = 1$ , для остальных порядков аналогично.

Мы хотим найти производную:

$$\lim_{h_n \rightarrow 0} \frac{\varphi_X(t + h_n) - \varphi_X(t)}{h_n} = \lim_{h_n \rightarrow 0} \frac{1}{h_n} \cdot \left( \mathbb{E}[e^{i(t+h_n)X}] - \mathbb{E}[e^{itX}] \right) = \lim_{h_n \rightarrow 0} E \left[ \frac{e^{i(t+h_n)X} - e^{itX}}{h_n} \right] =: \lim_{h_n \rightarrow 0} E[g_n],$$

то есть обозначили  $g_n = \frac{e^{i(t+h_n)X} - e^{itX}}{h_n}$ .

Поймем, что мы знаем про функцию  $g_n$ :

- У нее есть поточечный предел:

$$\lim_{n \rightarrow \infty} g_n(X) = (e^{itX})'_t = iX e^{itX}.$$

- Надо как-то оценить  $|g_n|$ .

Знаем, что модуль комплексной экспоненты равен 1, то есть  $|e^{itX}| = 1$ . Тогда

$$|g_n(X)| = \left| \frac{e^{itX} \cdot (e^{ih_n X} - 1)}{h_n} \right| = |e^{itX}| \cdot \left| \frac{e^{ih_n X} - 1}{h_n} \right| = \left| \frac{e^{ih_n X} - 1}{h_n} \right| = \left| \frac{e^{ih_n X} - e^{i \cdot 0 \cdot X}}{h_n} \right| = (e^{itX})'_t(\xi) = |iX e^{i\xi X}|$$

для некоторого  $\xi \in (0; h_n)$ .

Предпоследний переход выполнен по теореме Лагранжа, которая гласит следующее:

$$\exists \xi \in (a; b) : \frac{f'(b) - f'(a)}{b - a} = f'(\xi).$$

Опять же воспользуемся тем, что модуль комплексной экспоненты равен 1:

$$|g_n(X)| = |iX e^{i\xi X}| = |i| \cdot |X| \cdot |e^{i\xi X}| = 1 \cdot |X| \cdot 1 = |X|.$$

Мы получили, что

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

Тогда по теореме Лебега предел ожиданий есть ожидание предела:

$$\lim_{n \rightarrow \infty} E[g_n(X)] = E[i \cdot X \cdot e^{itX}].$$

Возвращаемся в самое начало:

$$\varphi_X(t)' = \lim_{n \rightarrow \infty} E[g_n] = i \cdot E[X \cdot e^{itX}].$$

■

5. **Переформулировка сходимости по распределению в терминах характеристических функций.** Однозначность задания распределения случайной величины ее характеристической функцией. Центральная предельная теорема.
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  $\frac{f(a + h_n X_n) - f(a)}{h_n}$  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

$$X_n \xrightarrow{d} X \Leftrightarrow \forall g \in \mathcal{G} \quad \mathbb{E}g(X_n) \xrightarrow{n \rightarrow \infty} \mathbb{E}g(X), \text{ где } g - \text{непрерывная, ограниченная функция}$$



$g \circ f := h$  – непрерывная 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

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

Снова применяем утверждение

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

■

## 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

$$P(X + Z \leq t - \varepsilon) - P(|Y - Z| \geq \varepsilon) \leq P(X + Y \leq t) \leq P(X + Z \leq t + \varepsilon) + P(|Y - Z| \geq \varepsilon)$$

*Доказательство.*

$$P(X + Y \leq t) \leq P(X + Y \leq t, |Y - Z| \leq \varepsilon) + P(X + Y \leq t, |Y - Z| \geq \varepsilon) \leq P(X + Y \leq t, |Y - Z| \leq \varepsilon) + P(|Y - Z| \geq \varepsilon)$$

Расскроем модуль

$$-\varepsilon \leq Y - Z \Rightarrow Z - \varepsilon \leq Y$$

Подставим вместо  $Y$   $Z - \varepsilon$

Событие  $X + Y \leq t \cap |Y - Z| \geq \varepsilon$  вложено в событие  $X + Z - \varepsilon \leq t$

$$\leq P(X + Z - \varepsilon \leq t) + P(|Y - Z| \geq \varepsilon)$$

Ищем другую оценку

Заменим в получившемся неравенстве  $Y$  на  $Z$ ,  $Z$  на  $Y$

$$\begin{aligned} P(X + Z \leq t) &\leq P(X + Y - \varepsilon \leq t) + P(|Z - Y| \geq \varepsilon) = \\ &= P(X + Y \leq t + \varepsilon) + P(|Y - Z| \geq \varepsilon) \end{aligned}$$

Обозначим  $t + \varepsilon := t$

$$P(X + Z \leq t - \varepsilon) \leq P(X + Y \leq t) + P(|Y - Z| \geq \varepsilon)$$

$$P(X + Y \leq t) \geq P(X + Z \leq t - \varepsilon) - P(|Y - Z| \geq \varepsilon)$$

■

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

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

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

*Доказательство.* Вспомним доказательство того что

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

$$\begin{aligned} \lim_{n \rightarrow \infty} P(|X_n - C| \geq \varepsilon) &= \lim_{n \rightarrow \infty} P(X_n - C \geq \varepsilon \text{ or } -X_n + C \geq \varepsilon) \leq \\ &= \lim_{n \rightarrow \infty} P(X_n - C \geq \varepsilon) + \lim_{n \rightarrow \infty} P(X_n \leq C - \varepsilon) = \\ &= 1 - F_{X_n}(\varepsilon + C) + F_{X_n}(C - \varepsilon) \underbrace{\xrightarrow{n \rightarrow \infty}}_0 = 0 \end{aligned}$$

Используем лемму

$$P(X_n + C \leq t - \varepsilon) - P(|Y_n - C| \geq \varepsilon) \leq P(X_n + Y_n \leq t) \leq P(X_n + C \leq t + \varepsilon) + P(|Y_n - C| \geq \varepsilon)$$

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

1)  $n \rightarrow \infty$

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.  $Y_n \xrightarrow{p} C \Leftrightarrow_{def} \lim_{n \rightarrow \infty} P(|Y_n - C| \geq \varepsilon) = 0$

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

2)  $\varepsilon \rightarrow 0$

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

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

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

1)  $C = 0$

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

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

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

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

a)  $n \rightarrow \infty$

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

b)  $R \rightarrow \infty$

$R$  – точка непрерывности  $F_X$

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

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

2) Общий случай

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

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

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

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

■

### 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  $\mathbb{D}X_j = \sigma^2$ . 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

$$\begin{aligned} \overline{X}_n &\xrightarrow{p} a \text{ (ЗБЧ)} \\ s_n^2 &= \frac{1}{n-1} \sum_{j=1}^n (X_j - \overline{X}_n)^2 = \frac{n}{n-1} \frac{1}{n} \sum_{j=1}^n X_j^2 + \frac{1}{n-1} \left( -2 \underbrace{\sum_{j=1}^n X_j \cdot \overline{X}_n}_{n \overline{X}_n^2} + n \overline{X}_n^2 \right) = \frac{n}{n-1} \left( \frac{1}{n} \sum_{j=1}^n X_j^2 - \overline{X}_n^2 \right) (*) \\ \frac{1}{n} \sum_{j=1}^n X_j^2 &\xrightarrow{p} \mathbb{E}X_1^2 \text{ (ЗБЧ)} \\ \overline{X}_n^2 &\xrightarrow{p} (\mathbb{E}X_1)^2 \\ \frac{n}{n-1} &\rightarrow 1 \\ (*) &\xrightarrow{p} \mathbb{D}X_1 = \sigma^2 \\ \mathbb{E}s_n^2 &= \frac{n}{n-1} \left( \frac{1}{n} \cdot n \cdot \mathbb{E}X_1^2 - \mathbb{E}(\overline{X}_n)^2 \right) = \\ \mathbb{E}(\overline{X}_n)^2 &= \mathbb{E}(\overline{X}_n - a + a)^2 = \mathbb{E}(\overline{X}_n - a)^2 + a^2 - 2a \underbrace{\mathbb{E}(\overline{X}_n - a)}_0 = a^2 + \mathbb{D}\overline{X}_n = a^2 + \frac{1}{n^2} \mathbb{D}(X_1 + \dots + X_n) = a^2 + \frac{\sigma^2}{n} \\ &= \frac{n}{n-1} \left( \sigma^2 + a^2 - a^2 - \frac{\sigma^2}{n} \right) = \sigma^2 \end{aligned}$$

#### Пример 2 (Взаимосвязь с ЦПТ)

Обозначения сохраняются с прошлого примера

Хотим показать, что

$$\begin{aligned} \frac{\sqrt{n}(\overline{X}_n - a)}{\sqrt{s_n^2}} &\rightarrow Z \sim \mathcal{N}(0, 1) \\ \frac{\sqrt{n}(\overline{X}_n - a)}{\sqrt{s_n^2}} &= \frac{\sqrt{n}(\overline{X}_n - a)}{\sigma} \cdot \sqrt{\frac{\sigma^2}{s_n^2}} \\ \frac{\sqrt{n}(\overline{X}_n - a)}{\sigma} &\rightarrow Z \sim \mathcal{N}(0, 1) \text{ из лекции 4} \\ \sqrt{\frac{\sigma^2}{s_n^2}} &\xrightarrow{p} \sqrt{\frac{\sigma^2}{\sigma^2}} = 1 \text{ (Обсуждали выше)} \end{aligned}$$

Значит

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

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

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

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

*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

$$h_n \xrightarrow{d} 0$$

$$X_n \xrightarrow{d} X$$

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

$$g(h_n X_n) \xrightarrow{d} g(0) \text{ (первая теорема в билете 6)} = f'(a)$$

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

■

## 6.5. Взаимосвязь с ЦПТ.

### Пример

Обозначения сохраняются 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 Хедфинга-Чернова. Пример применения.

### 7.1. Неравенство типа Хедфинга-Чернова

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

$$P(S_n - \mathbb{E}S_n \geq t) \leq 2 \exp\left(-\frac{t^2}{4 \sum_{j=1}^n (b_j - a_j)^2}\right)$$

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

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

Оценим каждое ожидание из произведения:

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

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

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

Действительно, если  $R > 1$ , то

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

если же  $R \leq 1$ , то

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

Таким образом,

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

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

$$P(S_n - \mathbb{E} S_n \geq t) \leq \exp\left(-\frac{t^2}{4 \sum_{j=1}^n (b_j - a_j)^2}\right).$$

Аналогично, рассматривая случайные величины  $X_j' := -X_j$  получаем оценку

$$P(-S_n + \mathbb{E} S_n \geq t) \leq \exp\left(-\frac{t^2}{4 \sum_{j=1}^n (b_j - a_j)^2}\right).$$

объединяя полученные неравенства получаем оценку из формулировки теоремы ■

**Теорема (следствие).** Пусть  $X_j \text{ Bern}(p)$  — набор независимых Бернулевских случайных величин,  $S_n = X_1 + \dots + X_n$ , тогда

$$P\left(\left|\frac{S_n}{n} - p\right| \geq t\right) \leq 2e^{-\frac{nt^2}{4}}$$

## 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  $\xi_j$  — бернулевская случайная величина, равная 1, если шар белого цвета и 0, если цвет черный. Мы хотим оценить вероятность успеха  $p$ . По нер-ву выше

$$P\left(\left|\frac{S_n}{n} - p\right| \geq t\right) \leq 2e^{-\frac{nt^2}{4}} < \varepsilon$$

Тогда при размере выборки  $n = O\left(\frac{\log \varepsilon^{-1}}{t^2}\right)$  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.**  $\langle x, y \rangle = x_1 y_1 + \dots + x_m y_m$ , где  $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ ,  $y = (y_1, \dots, y_m) \in \mathbb{R}^m$

**Определение 1.** Характеристическая функция случайного вектора  $X = (X_1, \dots, X_m)$  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  $X^n = (X_1^n, \dots, X_m^n)$  сходится по распределению к случайному вектору  $X = (X_1, \dots, X_m)$ , если для каждой непрерывной, ограниченной функции  $g : \mathbb{R}^m \rightarrow \mathbb{R}$  выполнено  $\mathbb{E}g(X^n) \rightarrow \mathbb{E}g(X)$  (обозначение  $X^n \xrightarrow{d} X$ ).

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

**Теорема 0.1. Без доказательства**

Последовательность случайных векторов  $X^n$  сходится по распределению к случайному вектору  $X$  тогда и только тогда, когда  $\varphi_{X^n}(y) \rightarrow \varphi_X(y)$  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 0.2.** Случайные величины  $X_1, \dots, X_m$  независимы тогда и только тогда, когда

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

где  $X = (X_1, \dots, X_m)$

*Доказательство.*

$\Rightarrow$

$$\varphi_X(y_1, \dots, y_m) = \mathbb{E}e^{i(X_1 y_1 + \dots + X_m y_m)} = \mathbb{E} \prod_{i=1}^m e^{iX_i y_i} \underset{\text{незав. } i=1}{=} \prod_{i=1}^m \mathbb{E}e^{iX_i y_i} = \prod_{i=1}^m \varphi_{X_i}(y_i)$$

$\Leftarrow$

Зададим случайный 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.функции тоже сходятся (Лекция 3, теорема 5)

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

Получили

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

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  $r_{kj} := \text{cov}(X_k, X_j)$  называется ковариационной матрицей вектора  $X$ .

**Теорема 0.3.** Симметричная неотрицательно определенная матрица  $R$  является ковариационной матрицей случайного вектора  $X$  тогда и только тогда, когда

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

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

*Доказательство.*

$$\begin{aligned} e_k &= (0, \dots, 0, \underbrace{1}_k, \dots, 0) \\ e_j &= (0, \dots, 0, \underbrace{1}_j, \dots, 0) \\ x &= \sum_k x_k e_k; \quad y = \sum_j y_j e_j \\ \langle \sum_k x_k R e_k, \sum_j y_j e_j \rangle &= \sum_k \sum_j \underbrace{\langle R e_k, e_j \rangle}_{(def) \text{cov}(x_k, x_j)} x_k y_j (*) \\ \text{cov}(x_k, x_j) &= \text{cov}(\langle x, e_k \rangle, \langle x, e_j \rangle) \\ (*) &= \sum_k \sum_j (\text{cov}(\langle x, e_k \rangle, \langle y, e_j \rangle)) x_k y_j = \sum_k \sum_j (\text{cov}(\langle x, x_k e_k \rangle, \langle x, y_j e_j \rangle)) = \\ &= \sum_k (\text{cov}(\langle x, x_k e_k \rangle, \sum_j \langle x, y_j e_j \rangle)) = \text{cov}(\langle x, x \rangle, \langle x, y \rangle) \end{aligned}$$

**Теорема 0.4.** Пусть  $X$  – случайный вектор с ковариационной матрицей, тогда случайный вектор  $AX + b$  имеет ковариационную матрицу  $ARA^*$

*Доказательство.*

$$Y = AX + b$$

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

\* – сдвиг на константу на ковариацию не влияет

\*\* – 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 0.5.** Пусть случайные вектора  $X^n = (X_1^n, \dots, X_m^n)$  независимы, одинаково распределены и имеют конечные  $a_j = \mathbb{E}X_j^n, r_{k,j} = \text{cov}(X_k^1, X_j^1)$

Тогда последовательность случайных векторов  $Y^n = (Y_1^n, \dots, Y_m^n)$  с компонентами

$$Y_j^n = \frac{X_j^1 + \dots + X_j^n - na_j}{\sqrt{n}}$$

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  $y \in \mathbb{R}^m$

Рассмотрим последовательность случайных величин

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

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

Значит по одномерной цпт

$$\begin{aligned} \xi_n &\xrightarrow{d} Z_y \sim \mathcal{N}(0, \mathbb{D}\langle X^1, y \rangle) \\ \varphi_{\xi_n}(t) &\rightarrow \varphi_{Z_y} = e^{-\frac{1}{2}t^2 \mathbb{D}\langle X^1, y \rangle} \end{aligned}$$

Заметим что

$$\begin{aligned} \varphi_{Y^n}(y) &= \mathbb{E}e^{i\langle Y^n, y \rangle} \\ \varphi_{\langle Y^n, y \rangle}(1) &= \mathbb{E}e^{i \cdot 1 \cdot \langle Y^n, y \rangle} \\ \Rightarrow \varphi_{Y^n}(y) &= \varphi_{\langle Y^n, y \rangle}(1) = \varphi_{\xi_n}(1) \rightarrow_* \varphi_{Z_y}(1) = e^{-\frac{1}{2}\mathbb{D}\langle X^1, y \rangle} \end{aligned}$$

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

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

Получили

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

■

**9. Многомерное нормальное распределение. Свойства нормального вектора: линейный образ нормального распределения нормален, характеристика 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  $\varphi_x(y) = E[\exp(i \langle X, y \rangle)] = e^{-\frac{1}{2}\langle R y, y \rangle + i \langle a, y \rangle}$ .

Где  $a = (a_1, \dots, a_m) \in \mathbb{R}^m$ ,  $R$  – симметричная неотрицательно определенная  $m \times m$  матрица. Далее пишем  $X \sim N(a, R)$ .



## 9.2. Свойства нормального вектора: линейный образ нормального распределения нормален, характеристика 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:

$$\varphi_{AX+b}(y) = E[\exp(i \langle AX + b, y \rangle)] = e^{i \langle b, y \rangle} E[\exp(i \langle X, A^*y \rangle)] = e^{i \langle b, y \rangle + i \langle a, A^*y \rangle - \frac{1}{2} \langle RA^*y, A^*y \rangle}$$

Остается заметить, что  $\langle RA^*y, A^*y \rangle = \langle ARA^*y, y \rangle$ , а также  $i \langle b, y \rangle + i \langle a, A^*y \rangle = i \langle Aa + b, y \rangle$ .

**Теорема 1.** Вектор  $X$  имеет нормальное распределение тогда и только тогда, когда для каждого вектора  $y$  случайная величина  $\langle y, X \rangle$  имеет нормальное распределение.

*Доказательство.* Так как по условию  $X \sim N(a, R)$ , то:

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

Иными словами,  $\langle X, y \rangle \sim N(\langle a, y \rangle, \langle Ry, y \rangle)$ , поскольку  $\langle a, y \rangle = E[\langle X, y \rangle]$ ,  $\langle Ry, y \rangle = D[\langle X, y \rangle]$ .

Докажем обратно:

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

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  $Z = (Z_1, \dots, Z_k)$  и случайные величины  $Z_j$  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  $\text{span}(X'_1, \dots, X'_m)$  со скалярным произведением  $(X, Y) = E[XY]$ . Для этого будем использовать метод Грамма-Шмидта. После него мы получаем случайные величины  $(Z_1, \dots, Z_k) = Z$ , что линейно выражаются через  $X'_1, \dots, X'_m$ . Т.е. в частности вектор  $Z$  - нормальный и  $E[Z_j] = 0$ , кроме того система является ортонормированным базисом в  $\text{span}(X'_1, \dots, X'_m)$ . 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  $\det R \neq 0$ , то случайный вектор  $X$  имеет плотность:

$$\rho(x) = \frac{1}{(2\pi)^{\frac{m}{2}} \sqrt{\det R}} e^{-\frac{1}{2} \langle R^{-1}(x-a), x-a \rangle}$$

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

$$P(X \in B) = P(AZ + a \in B) = \frac{1}{(\pi)^{\frac{m}{2}}} \int_{Ax+a \in B} e^{-\frac{1}{2}|x|^2} dx = \frac{1}{(\pi)^{\frac{m}{2}} \det A} \int_B e^{-\frac{1}{2}|A^{-1}(y-a)|^2} dy$$

Осталось заметить, что  $|A^{-1}(y-a)|^2 = \langle A^{-1}(y-a), A^{-1}(y-a) \rangle = \langle (AA)^{-1}(y-a), (y-a) \rangle = \langle (det A)^{-2} (y-a), (y-a) \rangle = (det A)^{-2} \langle y-a, y-a \rangle = (det A)^{-2} |y-a|^2$ .

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

**Определение 5.** Рассмотрим случай, когда разбиение  $\beta$  появляется посредством некоторой случайной величины  $Y = \sum_{k=1}^n y_k I_{B_k}$ , где  $y_k$  - различные числа и  $P(B_k) > 0$ . В этом случае  $B_k = \{\omega : Y(\omega) = y_k\}$  и условное математическое ожидание  $\mathbb{E}(X|\beta)$  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 0.6.** Имеют место следующие свойства условного математического ожидания:

- (i) (линейность)  $\mathbb{E}(\alpha X + \beta Y|\beta) = \alpha \mathbb{E}(X|\beta) + \beta \mathbb{E}(Y|\beta)$ ,
- (ii) (монотонность)  $X \leq Y$  п.н.  $\implies \mathbb{E}(X|\beta) \leq \mathbb{E}(Y|\beta)$ ,
- (iii) (аналог формулы полной вероятности)  $\mathbb{E}(\mathbb{E}(X|\beta)) = \mathbb{E}X$ ,
- (iv) (независимость) если случайная величина  $X$  не зависит от разбиения  $\beta$ , т.е. случайные величины  $X$  и  $I_{B_k}$  независимы для каждого  $k$ , то  $\mathbb{E}(X|\beta) = \mathbb{E}X$ .
- (v) для всякой случайной величины  $Z = \sum_{k=1}^n c_k I_{B_k}$  выполнено  $\mathbb{E}(ZX|\beta) = Z \mathbb{E}(X|\beta)$ .

*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:  $\mathbb{E}(\mathbb{E}(X|\beta)) = \mathbb{E}\left(\sum_{k=1}^n I_{B_k} \frac{\mathbb{E}(X I_{B_k})}{P(B_k)}\right) = \sum_{k=1}^n \mathbb{E}(X I_{B_k}) = \mathbb{E}X$ .

Обоснуем пункт (iv). Так как  $X$  и  $I_{B_k}$  независимы, то  $\mathbb{E}(X I_{B_k}) = \frac{\mathbb{E}(X I_{B_k})}{P(B_k)} = \frac{\mathbb{E}X \mathbb{E}I_{B_k}}{P(B_k)} = \mathbb{E}X$ .

Следовательно,  $\mathbb{E}(X|\beta) = \sum_{k=1}^n I_{B_k} \mathbb{E}(X|B_k) = \sum_{k=1}^n I_{B_k} \mathbb{E}X = \mathbb{E}X$ .

Для обоснования (v) достаточно заметить, что  $\mathbb{E}(X Z|B_k) = \frac{\mathbb{E}(X Z I_{B_k})}{P(B_k)} = c_k \frac{\mathbb{E}(X I_{B_k})}{P(B_k)} = c_k \mathbb{E}(X|B_k)$ .

ч.т.д. ■

**Теорема 0.7.** 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)  $X \leq Y$  п.н.  $\implies \mathbb{E}(X|Z) \leq \mathbb{E}(Y|Z)$ ,
- (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.  $\min_{Z: Z=g(Y)} \mathbb{E}|X - Z|^2 = \mathbb{E}|X - \mathbb{E}(X|Y)|^2$ .

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

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 формулы Байеса.