

# Математическая статистика.

Андрей Тищенко @AndrewTGk

2024/2025

Лекция 10 января

## Преамбула

*Статистика.* Мнения о появлении этого слова:

1. Статистиками в Германии назывались люди, собирающие данные о населении и передающие их государству.
2. В определённый день в Венеции народ выстраивался для выплаты налогов (строго фиксированных, в зависимости от рода действий). Государство собирало данные обо всём населении. Это происходило до появления статистиков в Германии, поэтому мы будем считать, что статистика пошла из Венеции.

*Задача статистики* — по результатам наблюдений построить вероятностную модель наблюдаемой случайной величины.

## Основные определения

### Определение

Однородной выборкой объёма  $n$  называется случайный вектор  $X = (X_1, \dots, X_n)$ , компоненты которого являются независимыми и одинаково распределёнными. Элементы вектора  $X$  называются элементами выборки.

### Определение

Если элементы выборки имеют распределение  $F_\xi(x)$ , то говорят, что выборка соответствует распределению  $F_\xi(x)$  или порождена случайной величиной  $\xi$  с распределением  $F_\xi(x)$ .

### Определение

Детерминированный вектор  $x = (x_1, \dots, x_n)$ , компоненты которого  $x_i$  являются реализациями соответствующих случайных величин  $X_i$  ( $i = \overline{1, n}$ ), называется реализацией выборки.

### Уточнение

Если  $X$  — однородная выборка объёма  $n$ , то его реализацией будет вектор  $x$ , каждый элемент  $x_i$  которого является значением соответствующей ему случайной величины (элемента выборки)  $X_i$ .

### Определение

Выборочным пространством называется множество всех возможных реализаций выборки

$$X = (X_1, \dots, X_n)$$

## Пример

У вектора  $X = (X_1, \dots, X_{10})$  каждый элемент  $X_i$  которой порождён случайной величиной  $\xi \sim U(0, 1)$ , выборочным пространством является  $\mathbb{R}^{10}$  (так как  $X_i$  может принять любое значение на  $\mathbb{R}$ )

## Определение

Обозначим  $x_{(i)}$  —  $i$ -ый по возрастанию элемент, тогда будет справедливо:

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

Обозначим  $X_{(k)}$  случайную величину, реализация которой при каждой реализации  $x$  выборки  $X$  принимает значение  $x_{(k)}$ . Тогда последовательность  $X_{(1)}, \dots, X_{(n)}$  называется вариационным рядом выборки.

## Определение

Случайная величина  $X_{(k)}$  называется  $k$ -ой порядковой статистикой выборки.

## Определение

Случайные величины  $X_{(1)}, X_{(n)}$  называются экстремальными порядковыми статистиками.

## Определение

Порядковая статистика  $X_{([n \cdot p])}$  называется выборочной квантилью уровня  $p$ , где  $p \in [0, 1]$

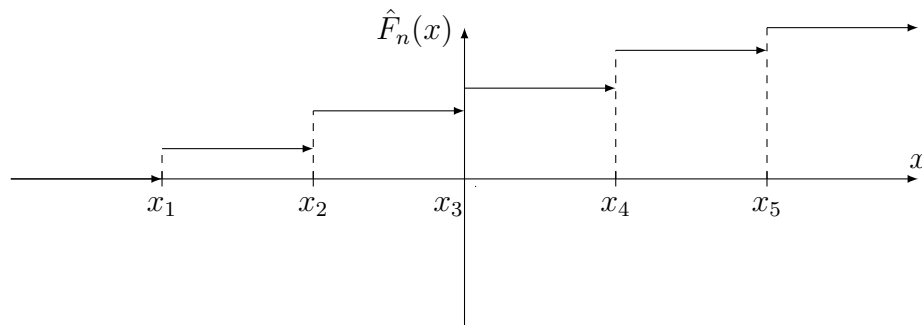
## Определение

Пусть каждый элемент выборки  $X$  объёма  $n$  имеет распределение  $F_\xi(x)$ . Эмпирической функцией распределения такой выборки называется

$$\hat{F}_n(x) = \frac{1}{n} \sum_{k=1}^n I(X_k \leq x)$$

$I$  — индикаторная функция.  $I = \begin{cases} 1, & \text{если аргумент верен} \\ 0, & \text{иначе} \end{cases}$

Пусть  $x_1, \dots, x_n$  — реализация выборки  $X_1, \dots, X_n$



Свойства  $\hat{F}_n(x)$

$$1. \forall x \in \mathbb{R} \quad E\hat{F}_n(x) = E\left(\frac{1}{n} \sum_{k=1}^n I(X_k \leq x)\right) = \frac{1}{n} \sum_{k=1}^n EI(X_k \leq x) = P(X_1 \leq x) = F_\xi(x)$$

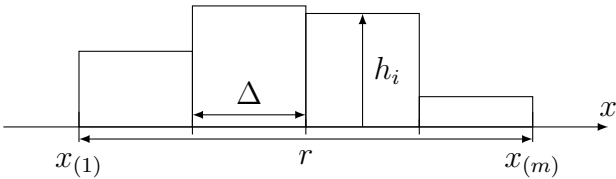
2. По усиленному закону больших чисел (УЗБЧ)

$$\forall x \in \mathbb{R} \quad \hat{F}_n(x) = \frac{1}{n} \sum_{k=1}^n I(X_k \leq x) \xrightarrow[n \rightarrow \infty]{\text{п. н.}} EI(X_k \leq x) = F_\xi(x)$$

# Гистограмма

Разбить  $\mathbb{R}$  на  $(m + 2)$  непересекающихся интервала. Рассматриваются  $x_{(1)}, \dots, x_{(m)}$

Название	Обозначение	Формула
Количество интервалов	$m$	—
Размах выборки	$r$	$r = x_{(m)} - x_{(1)}$
Ширина интервала	$\Delta$	$\Delta = \frac{r}{m}$
Количество попаданий на $i$ -ый интервал	$\nu_i$	—
Частота попаданий на $i$ -ый интервал	$h_i$	$h_i = \frac{\nu_i}{\Delta}$



Лекция 17 января

## Определение

Пусть  $X_1, \dots, X_n \sim F(x, \theta)$ .  $k$ -ым начальным выборочным моментом называется

$$\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^n x_i^k, \quad k \in \mathbb{N}$$

Выборочным средним называется:

$$\hat{\mu}_1 = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

## Определение

$k$ -ым центральным выборочным моментом называется

$$\hat{\nu}_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k, \quad k = 2, 3, \dots$$

$$\hat{\nu}_2 = S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \text{ называется выборочной дисперсией}$$

Пусть  $(x_1, y_1), \dots, (x_n, y_n)$  соответствует распределению  $F(x, y, \theta)$

## Определение

Выборочной ковариацией называется

$$\hat{K}_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

## Определение

Выборочным коэффициентом корреляции называется

$$\hat{\rho}_{xy} = \frac{\hat{K}_{xy}}{\sqrt{S_x^2 S_y^2}}$$

## Свойства выборочных моментов

1.  $E\hat{\mu}_k = E\left(\frac{1}{n} \sum_{i=1}^n X_i^k\right) = \frac{1}{n} \sum_{i=1}^n EX_i^k = EX_1^k = \mu_k$
2.  $E\bar{X} = m_x$
3.  $\mathcal{D}\hat{\mu}_k = \mathcal{D}\left(\frac{1}{n} \sum_{i=1}^n x_i^k\right) = \frac{1}{n^2} \sum_{i=1}^n \mathcal{D}X_i^k = \frac{1}{n} \mathcal{D}X_1^k = \frac{1}{n} (EX_1^{2k} - (EX_1^k)^2) = \frac{1}{n}(\mu_{2k} - \mu_k^2)$
4.  $\mathcal{D}\bar{x} = \frac{\sigma_{x_1}^2}{n}$
5. По УЗБЧ

$$\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^n x_i^k \xrightarrow[n \rightarrow \infty]{\text{п. н.}} E\hat{\mu}_k = \mu_k$$

$$\hat{\nu}_k \xrightarrow[n \rightarrow \infty]{\text{п. н.}} \nu_k$$

6. По ЦПТ

$$\frac{\hat{\mu}_k - \mu_k}{\sqrt{\frac{\mu_{2k} - \mu_k^2}{n}}} \xrightarrow[n \rightarrow \infty]{d} U, \quad U \sim N(0, 1)$$

$$\frac{\sqrt{n}(\bar{x} - m_{x_1})}{\sigma} \xrightarrow[n \rightarrow \infty]{d} U$$

7.  $ES^2 = E\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right) = \frac{n-1}{n} \sigma^2$
8.  $E\hat{K}_{xy} = \frac{n-1}{n} \text{cov}(x, y)$

## Определение

Оценкой  $\hat{\theta}$  параметра  $\theta$ , называется функция:

$$\hat{\theta} = T(x_1, \dots, x_n), \text{ не зависящая от } \theta$$

Например, отвратительная оценка среднего роста людей в аудитории.

$$\hat{m} = \frac{2x_2 + 5x_5 + 10x_{10}}{3}$$

## Определение

Оценка  $\hat{\theta}$  называется несмещённой, если  $E\hat{\theta} = \theta$  для любых возможных значений этого параметра.

## Определение

Оценка  $\hat{\theta}(x_1, \dots, x_n)$  называется асимптотически несмещённой оценкой  $\theta$ , если

$$\lim_{n \rightarrow \infty} E\hat{\theta}(x_1, \dots, x_n) = \theta$$

$$\lim_{n \rightarrow \infty} ES^2 = \lim_{n \rightarrow \infty} \frac{n-1}{n} \sigma^2 = \sigma^2$$

## Определение

Несмещённой выборочной (или исправленной) выборочной дисперсией называется

$$\tilde{S}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Оценки

$$\hat{m}_1 = \frac{x_1 + x_2 + x_3}{3}$$

$$\hat{m}_2 = \frac{\sum_{i=1}^{10} x_i}{10}$$

$$\hat{m}_3 = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Являются несмещёнными.

## Определение

Оценка  $\hat{\theta}(x_1, \dots, x_n)$  называется:

Состоятельной оценкой  $\theta$ , если

$$\hat{\theta}(x_1, \dots, x_n) \xrightarrow[n \rightarrow \infty]{p} \theta$$

Сильно состоятельной оценкой, если

$$\hat{\theta}(x_1, \dots, x_n) \xrightarrow[n \rightarrow \infty]{\text{п. н.}} \theta$$

## Определение

Пусть  $\hat{\theta}$  — несмещённая оценка параметра  $\theta$ . Если  $\mathcal{D}\hat{\theta} \leq \mathcal{D}\theta^*$ , где  $\theta^*$  — любая несмещённая оценка параметра  $\theta$ . Тогда  $\hat{\theta}$  называется эффeктивной оценкой параметра  $\theta$ .

### $R$ -эффeктивные оценки

Рассматриваем выборку  $X_1, \dots, X_n \sim f(x, \theta)$ ,  $\theta \in \Theta \subseteq \mathbb{R}^1$ . Назовём модель  $(S, f(x, \theta))$  регулярной, если она удовлетворяет следующим условиям:

1.  $\forall x \in S$  функция  $f(x, \theta) = f(x_1, \dots, x_n, \theta) > 0$  и дифференцируема по  $\theta$ .

$$2. \begin{cases} \frac{\delta}{\delta\theta} \int_S f(x, \theta) dx = \int_S \frac{\delta}{\delta\theta} f(x, \theta) dx \\ \frac{\delta}{\delta\theta} \int_S T(x) f(x, \theta) dx = \int_S \frac{\delta}{\delta\theta} T(x) f(x, \theta) dx \end{cases}$$

Пусть  $\hat{\theta} = T(x) = T(x_1, \dots, x_n)$  — несмещённая оценка параметра  $\theta$ :

$$\int_S \frac{\delta}{\delta\theta} f(x, \theta) dx = \frac{\delta}{\delta\theta} \int_S f(x, \theta) dx = \frac{\delta}{\delta\theta} 1 = 0$$

$$\int_S \frac{\delta}{\delta\theta} T(x) f(x, \theta) dx = \frac{\delta}{\delta\theta} \int_S T(x) f(x, \theta) dx = \frac{\delta}{\delta\theta} ET(x) = \frac{\delta}{\delta\theta} \theta = 1$$

## Определение

Информацией Фишера о параметре  $\theta$ , содержащейся в выборке  $X_1, \dots, X_n$  называется величина

$$I_n(\theta) = E \left( \frac{\delta \ln(f(x, \theta))}{\delta\theta} \right)^2 = \int_S \left( \frac{\delta \ln(f(x, \theta))}{\delta\theta} \right)^2 f(x, \theta) dx$$

## Неравенство Рао-Крамера

Если  $S$ ,  $f(x, \theta)$  — регулярная модель и  $\hat{\theta}$  — несмещённая оценка  $\theta$ , то

$$\mathcal{D}(\hat{\theta}) \geq \frac{1}{I_n(\theta)}$$

### Доказательство

Выпишем некоторые равенства (пригодятся в доказательстве):

$$\int_S \frac{\delta}{\delta\theta} f(x, \theta) dx = \int_S \frac{\delta f(x, \theta)}{\delta\theta} \frac{f(x, \theta)}{f(x, \theta)} dx \stackrel{*}{=} \int_S \frac{\delta \ln f(x, \theta)}{\delta\theta} f(x, \theta) dx = 0$$

Пояснение  $\stackrel{*}{=}$ . Логарифм — сложная функция. По правилу дифференцирования сложной функции:

$$\frac{\delta \ln f(x, \theta)}{\delta\theta} = \frac{1}{f(x, \theta)} \cdot \frac{\delta f(x, \theta)}{\delta\theta}$$
$$\int_S \frac{\delta}{\delta\theta} T(x) f(x, \theta) dx = \int_S T(x) \frac{\delta}{\delta\theta} f(x, \theta) \frac{f(x, \theta)}{f(x, \theta)} dx = \int_S T(x) \frac{\delta \ln f(x, \theta)}{\delta\theta} f(x, \theta) dx = 1$$

Чуть преобразуем последнее полученное равенство:

$$\int_S T(x) \frac{\delta \ln f(x, \theta)}{\delta\theta} f(x, \theta) dx = \int_S T(x) \frac{\delta \ln f(x, \theta)}{\delta\theta} f(x, \theta) dx - \underbrace{\theta \int_S \frac{\delta \ln f(x, \theta)}{\delta\theta} f(x, \theta) dx}_{=0} =$$
$$= \int_S (T(x) - \theta) \frac{\delta \ln f(x, \theta)}{\delta\theta} f(x, \theta) dx = 1 \Rightarrow 1 = 1^2 = \left( \int_S (T(x) - \theta) \frac{\delta \ln f(x, \theta)}{\delta\theta} f(x, \theta) dx \right)^2$$

Далее нам понадобится неравенство Коши-Буняковского, которое выглядит так:

$$\left( \int \varphi_1(x) \varphi_2(x) dx \right)^2 \leq \int \varphi_1^2(x) dx \int \varphi_2^2(x) dx$$

Подгоним полученное равенство ( $f(x, \theta) > 0 \Rightarrow f(x, \theta) = \sqrt{f(x, \theta)^2}$ ):

$$\left( \int_S (T(x) - \theta) \frac{\delta \ln f(x, \theta)}{\delta\theta} f(x, \theta) dx \right)^2 = \left( \int_S \underbrace{(T(x) - \theta) \sqrt{f(x, \theta)}}_{\varphi_1(x)} \cdot \underbrace{\frac{\delta \ln f(x, \theta)}{\delta\theta} \sqrt{f(x, \theta)}}_{\varphi_2(x)} dx \right)^2 = 1$$

И применим неравенство Коши-Буняковского:

$$1 = \left( \int_S \underbrace{(T(x) - \theta) \sqrt{f(x, \theta)}}_{\varphi_1(x)} \cdot \underbrace{\frac{\delta \ln f(x, \theta)}{\delta\theta} \sqrt{f(x, \theta)}}_{\varphi_2(x)} dx \right)^2 \leq$$
$$\leq \int_S \left( (T(x) - \theta) \sqrt{f(x, \theta)} \right)^2 dx \cdot \int_S \left( \frac{\delta \ln f(x, \theta)}{\delta\theta} \sqrt{f(x, \theta)} \right)^2 dx =$$
$$= \underbrace{\int_S (T(x) - \theta)^2 f(x, \theta) dx}_{=\mathcal{D}\hat{\theta}} \cdot \underbrace{\int_S \left( \frac{\delta \ln f(x, \theta)}{\delta\theta} \right)^2 f(x, \theta) dx}_{=I_n(\theta)}$$

Получаем:

$$1 \leq \mathcal{D}(\theta) \cdot I_n(\theta) \Rightarrow \mathcal{D}(\theta) \geq \frac{1}{I_n(\theta)}$$

## Определение

Оценка  $\hat{\theta}$  называется R-эффективной, если  $E\hat{\theta} = \theta$  и  $\mathcal{D}\hat{\theta} = \frac{1}{I_n(\theta)}$

Лекция 24 января

### Замечание 1

$$I_n(\theta) = \mathcal{D} \left( \frac{\delta \ln f(x, \theta)}{\delta \theta} \right)$$

### Замечание 2

$$I_n(\theta) = nI_1(\theta)$$

$$f(x, \theta) = f(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f(x_i, \theta)$$

$$\begin{aligned} E \left( \frac{\delta \ln f(x, \theta)}{\delta \theta} \right)^2 &= E \left( \sum_{i=1}^n \frac{\delta \ln f(x_i, \theta)}{\delta \theta} \right)^2 = \sum_{i \neq j} E \left( \frac{\delta \ln f(x_i, \theta)}{\delta \theta} \cdot \frac{\delta \ln f(x_j, \theta)}{\delta \theta} \right) + n E \left( \frac{\delta \ln f(x_1, \theta)}{\delta \theta} \right)^2 = \\ &= \sum_{i \neq j} \left( \underbrace{E \left( \frac{\delta \ln f(x_i, \theta)}{\delta \theta} \right)}_{=0} \cdot \underbrace{E \left( \frac{\delta \ln f(x_j, \theta)}{\delta \theta} \right)}_{=0} \right) + n E \left( \frac{\delta \ln f(x_1, \theta)}{\delta \theta} \right)^2 = n E \left( \frac{\delta \ln f(x_1, \theta)}{\delta \theta} \right)^2 = n I_1(\theta) \end{aligned}$$

### Замечание 3

Пример:  $X_1, \dots, X_n \sim N(\theta, \sigma^2)$

Рассмотрим оценку  $\hat{\theta} = \bar{X}$ , её дисперсия  $\mathcal{D}\bar{X} = \frac{\sigma^2}{n}$ . Посчитаем информацию Фишера:

$$\begin{aligned} I_1(\theta) &= E \left( \frac{\delta \ln f(x, \theta)}{\delta \theta} \right)^2 = E \left( \frac{\delta}{\delta \theta} \ln \left( \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\theta)^2}{2\sigma^2}} \right) \right)^2 = E \left( \frac{\delta}{\delta \theta} \ln \left( \frac{1}{\sqrt{2\pi}\sigma} - \frac{(x-\theta)^2}{2\sigma^2} \right) \right)^2 = E \left( \frac{x-\theta}{\sigma^2} \right)^2 = \\ &= \frac{1}{\sigma^4} E(x - \theta)^2 = \frac{\sigma^2}{\sigma^4} = \frac{1}{\sigma^2} \Rightarrow I_n(\theta) = \frac{n}{\sigma^2} \end{aligned}$$

Знаем, что  $\mathcal{D}\hat{\theta} \geq \frac{1}{nI_1(\theta)} = \frac{\sigma^2}{n} = \mathcal{D}(\bar{X}) \Rightarrow$  оценка  $\hat{\theta} = \bar{X}$  является R-эффективной.

Критерий эффективности  $X_1, \dots, X_n \sim F_\xi(x, \theta)$ ,  $\theta \in \Theta \subset \mathbb{R}^1$  выполнены условия регулярности, то есть

$$\int T(x) \frac{\delta f(x, \theta)}{\delta \theta} dx = \frac{\delta}{\delta \theta} \int T(x) f(x, \theta) dx = E\hat{\theta}$$

## Определение

Функцией вклада выборки  $X_1, \dots, X_n$  называется

$$U(x, \theta) = \sum_{i=1}^n \frac{\delta \ln f(x_i, \theta)}{\delta \theta}$$

Пусть  $0 < U(x, \theta) < \infty$ .

$\hat{\theta} = T(x_1, \dots, x_n)$  — R-эффективная оценка  $\theta \Leftrightarrow \hat{\theta} - \theta = a(\theta)U(x, \theta)$ , где  $a(\theta) = \mathcal{D}\hat{\theta}$

*Доказательство  $\Rightarrow$ :*

Пусть  $\hat{\theta} - \theta = a(\theta)U(x, \theta) \Rightarrow \hat{\theta}$  — R-эффективная оценка  $\theta$ .

Посчитаем математическое ожидание частей равенства:

$$E(\hat{\theta} - \theta) = a(\theta)EU(x, \theta) = a(\theta) \int \frac{\delta \ln f(x, \theta)}{\delta \theta} f(x, \theta) dx = 0$$

Посчитаем дисперсию частей:

$$\mathcal{D}(\hat{\theta} - \theta) = a^2(\theta)\mathcal{D}U(x, \theta) = \underbrace{a^2(\theta)}_{=(\mathcal{D}(\hat{\theta}))^2} I_n(\theta) \Rightarrow \mathcal{D}(\hat{\theta}) = (\mathcal{D}(\hat{\theta}))^2 I_n(\theta) \Rightarrow 1 = \mathcal{D}(\theta) I_n(\theta)$$

Значит оценка является R-эффективной.

*Доказательство*  $\Leftarrow$ :

Пусть  $\hat{\theta}$  — R-эффективная оценка  $\Rightarrow \hat{\theta} - \theta = a(\theta)U(x, \theta)$ . Хотим доказать, что  $\rho(\hat{\theta}, U(x, \theta)) = 1$ .

Для подсчёта корреляции нужно посчитать ковариацию:

$$\text{cov}(\hat{\theta}, U(x, \theta)) = E(\hat{\theta} - \theta)U(x, \theta) = E\hat{\theta}U(x, \theta) - \underbrace{\theta EU(x, \theta)}_{=0} =$$

$$= \int_S T(x)U(x, \theta)f(x, \theta) dx = \int_S T(x) \frac{\delta \ln f(x, \theta)}{\delta \theta} f(x, \theta) dx = 1$$

Так как  $\hat{\theta}$  — R-эффективная оценка, то  $\mathcal{D}\hat{\theta} = \frac{1}{I_n(\theta)}$ . Зная, что  $\mathcal{D}U(x, \theta) = I_n(\theta)$ , тогда:

$$\rho(\hat{\theta}, U(x, \theta)) = \frac{\text{cov}(\hat{\theta}, U(x, \theta))}{\sqrt{\mathcal{D}\hat{\theta}\mathcal{D}U(x, \theta)}} = \frac{1}{\sqrt{\frac{I_n(\theta)}{I_n(\theta)}}} = 1 \Rightarrow$$

$$\Rightarrow \hat{\theta} = c_1 + c_2 U(x, \theta)$$

$E\hat{\theta} = c_1 + Ec_2 U(x, \theta) = c_1 + 0 = \theta$ , так как оценка эффективная

$$\mathcal{D}\hat{\theta} = c_2^2 I_n(\theta) = \frac{1}{I_n(\theta)} \Rightarrow c_2^2 = \frac{1}{I_n^2(\theta)} \Rightarrow c_2 = \frac{1}{I_n(\theta)} = \mathcal{D}\hat{\theta} = a(\theta).$$

Итак,  $\hat{\theta} = \theta + a(\theta)U(x, \theta) \Rightarrow \hat{\theta} - \theta = U(x, \theta)$ .

## Метод моментов

$X_1, \dots, X_n \sim F_\xi(x, \theta), \theta \in \Theta \subset R^k$

$$\exists \mu_j < \infty, j = \overline{1, k} \quad \underbrace{\mu_j}_{=\mu_j(\theta)} = E\xi^j = \int_{-\infty}^{+\infty} x^j f(x, \theta) dx = 1$$

Тогда можно получить систему уравнений:

$$\begin{cases} \hat{\mu}_1 = \mu_1(\theta) \\ \vdots \\ \hat{\mu}_n = \mu_n(\theta) \end{cases} \quad (1)$$

Если система уравнений (1) однозначно разрешима относительно  $\theta_1, \dots, \theta_k$ , то решения  $\hat{\theta}_1, \dots, \hat{\theta}_k$  называется равной  $\theta_1, \dots, \theta_k$  по методу моментов.

## Пример

$X_1, \dots, X_n \sim N(\theta_1, \theta_2^2), \theta = (\theta_1, \theta_2^2)$ , тогда:

$$\begin{cases} \hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^n x_i = \theta_1 \Rightarrow \hat{\theta}_1 = \bar{X} \\ \hat{\mu}_2 = \frac{1}{n} \sum_{i=1}^n x_i^2 = \theta_2^2 + \theta_1^2, (E\xi^2 = \mathcal{D}\xi + (E\xi)^2) \Rightarrow \hat{\theta}_2^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{X}^2 \end{cases}$$

## Метод максимального правдоподобия (ММП)

### Определение

Функцией правдоподобия называется функция

$$L(x_1, \dots, x_n, \theta) = \begin{cases} \prod_{i=1}^n f(x_i, \theta), & \text{если } \xi \text{ — непрерывная случайная величина} \\ \prod_{i=1}^n P(\xi = x_i, \theta), & \text{если } \xi \text{ — дискретная случайная величина} \end{cases}$$



## Определение

Реализацией оценки максимального правдоподобия (ОМП) называется значение  $\hat{\theta} \in \Theta$ , такое что:

$$\hat{\theta} = \operatorname{argmax} L(x_1, \dots, x_n, \theta), \text{ где } \theta \in \Theta$$

Для нахождения точки максимума нужно взять частные производные по всем составляющим  $\theta$  от функции правдоподобия. Однако считать производную произведения нам впадлу, поэтому мы введём следующую вещь:

## Определение

Функция  $\ln L(x_1, \dots, x_n, \theta)$  называется логарифмической функцией правдоподобия.

Итак, получаем систему уравнений:

$$\begin{cases} \frac{\delta \ln L(x_1, \dots, x_n, \theta)}{\delta \theta_1} = 0 \\ \vdots \\ \frac{\delta \ln L(x_1, \dots, x_n, \theta)}{\delta \theta_k} = 0 \end{cases}$$

Логарифм монотонный, поэтому его  $\operatorname{argmax}$  совпадёт с  $\operatorname{argmax}$  функции  $L(x_1, \dots, x_n, \theta)$  (НАУКА!).

## Пример

Для Гауссовской величины  $N(\theta_1, \theta_2^2)$ :

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x_i - \theta_1)^2}{2\theta_2^2}} = \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\theta_2}\right)^n e^{-\frac{(x - \theta_1)^2}{2\theta_2^2}}$$

Логарифмируем:

$$\ln L(x_1, \dots, x_n, \theta) = \ln \left(\frac{1}{\sqrt{2\pi}}\right)^n - n \ln \theta_2 - \frac{\sum_{i=1}^n (x_i - \theta_1)^2}{2\theta_2^2}$$

Возьмём частные производные:

$$\begin{cases} \frac{\delta \ln L(x_1, \dots, x_n, \theta)}{\delta \theta_1} = \frac{\sum_{i=1}^n (x_i - \hat{\theta}_1)}{\hat{\theta}_2^2} \\ \frac{\delta \ln L(x_1, \dots, x_n, \theta)}{\delta \theta_2} = -\frac{n}{\hat{\theta}_2} + \frac{\sum_{i=1}^n (x_i - \hat{\theta}_1)^2}{\hat{\theta}_2^3} \end{cases}$$

Посчитаем  $\theta_1, \theta_2$ :

$$\begin{cases} \sum_{i=1}^n (x_i - \hat{\theta}_1) = 0 \Rightarrow \hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^n x_i = \bar{X} \\ -n\hat{\theta}_2^2 + \sum_{i=1}^n (x_i - \bar{x})^2 = 0 \Rightarrow \hat{\theta}_2^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \end{cases}$$

**Лекция 31 января.**

## Робастные оценки

От слова robust.

## Определение

Пусть оценка  $\hat{\theta}_n$  построена по выборке  $X_1, \dots, X_n$ . Затем добавлено наблюдение  $x$  и построена оценка  $\hat{\theta}_{n+1}$ , тогда кривой чувствительности, изучающей влияние наблюдения  $x$  на оценку  $\hat{\theta}$  называется функция:

$$SC_n(x) = \frac{\hat{\theta}_{n+1} - \hat{\theta}_n}{\frac{1}{n+1}} = (n+1) (\hat{\theta}_{n+1} - \hat{\theta}_n)$$

## Определение

Оценка  $\hat{\theta}$  называется  $B$ -робастной, если  $SC_n(x)$  ограничена.

## Пример

Пусть  $\hat{\theta} = \bar{X}$

$$SC_n(x) = (n+1) \left( \frac{1}{n+1} \left( \sum_{i=1}^n (x_i) + x \right) - \frac{1}{n} \sum_{i=1}^n x_i \right) = \sum_{i=1}^n x_i + x - \left( \sum_{i=1}^n x_i + \frac{1}{n} \sum_{i=1}^n x_i \right) = x - \bar{X}$$

Это линейная функция от  $x$ , то есть кривая чувствительности неограничена.

Пусть  $\hat{\theta} = \hat{\mu}$  (выборочная медиана)

$$\hat{\mu} = \begin{cases} X_{(k+1)}, & n = 2k + 1 \\ \frac{X_{(k)} + X_{(k+1)}}{2}, & n = 2k \end{cases}$$

## Определение

Пороговой точкой (BP)  $\varepsilon_n^*$  оценки  $\hat{\theta}$ , построенной на выборке  $X_1, \dots, X_n$  называется:

$$\varepsilon_n^* = \frac{1}{n} \max \left\{ m : \max_{i_1, \dots, i_m} \sup_{y_1, \dots, y_m} |\hat{\theta}(z_1, \dots, z_m)| < \infty \right\}$$

Где выборка  $z_1, \dots, z_m$  получена заменой значений  $X_{i_1}, \dots, X_{i_m}$  на произвольные значения  $y_1, \dots, y_m$

## Доверительные интервалы

### Определение

Пусть для  $X_1, \dots, X_n \sim F(x, \theta)$ ,  $\theta \in \Theta \subset \mathbb{R}^1$  построены статистики  $T_1(x_1, \dots, x_n)$  и  $T_2(x_1, \dots, x_n)$ , такие что

$$\begin{cases} T_1(x) < T_2(x) \\ P(T_1(x) < \theta < T_2(x)) = 1 - \alpha, \quad 0 < \alpha < 1 \end{cases}$$

Тогда интервал  $(T_1(x), T_2(x))$  называется доверительным интервалом уровня надёжности (доверия)  $1 - \alpha$  параметра  $\theta$ .

### Определение

Случайная функция  $G(x_1, \dots, x_n, \theta) = G(x, \theta)$  называется центральной (опорной) статистикой, если

1.  $G(x, \theta)$  непрерывна и монотонна по  $\theta$
2.  $F_G(x)$  не зависит от  $\theta$

Односторонние доверительные интервалы:

$$P(G(x, \theta) < Z_{1-\alpha}) = 1 - \alpha$$

$$P(Z_\alpha < G(x, \theta)) = 1 - \alpha$$

Квантили не зависят от  $\theta$ , с их помощью можно выразить односторонние доверительные интервалы.

Центральным доверительным интервалом будет:

$$P(Z_{\frac{\alpha}{2}} < G(x, \theta) < Z_{1-\frac{\alpha}{2}}) = 1 - \alpha$$

### Определение

Пусть случайные величины  $\xi_1, \dots, \xi_m \sim N(0, 1)$  и независимы.

Тогда случайная величина  $\eta = \sum_{i=1}^m \xi_i^2 \sim \chi^2(m)$  (удовлетворяет распределению хи-квадрат ( $\chi^2$ ) с  $m$  степенями свободы).

## Определение

Пусть  $\xi_0, \xi_1, \dots, \xi_m \sim N(0, 1)$  и независимы.

Тогда случайная величина  $\zeta = \frac{\xi_0}{\sqrt{\frac{1}{m} \sum_{i=1}^m \xi_i^2}} \sim t(m)$  (распределение Стьюдента с  $m$  степенями свободы)

## Определение

Пусть случайная величина  $\xi_1 \sim \chi^2(m)$ ,  $\xi_2 \sim \chi^2(n)$  и  $\xi_1$  и  $\xi_2$  — независимы. Тогда случайная величина  $F = \frac{\frac{1}{m}\xi_1}{\frac{1}{n}\xi_2} \sim F(m, n)$  (распределение Фишера со степенями свободы  $n, m$ )

## Теорема Фишера

Пусть  $X_1, \dots, X_n$  порождены случайной величиной  $X \sim N(m, \sigma^2)$ , тогда:

1.  $\frac{nS^2}{\sigma^2} = \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{\sigma} \right)^2 \sim \chi^2(n-1)$  (так как мы знаем  $\bar{X}$ , и все наблюдения, а по  $n-1$  наблюдению и  $\bar{X}$  можно восстановить последнее наблюдение)
2.  $\bar{X}$  и  $S^2$  — независимые случайные величины.

## Пример 1

$X_1, \dots, X_n \sim N(\theta, \sigma^2)$ ,  $\sigma^2$  — известно. Построить доверительный интервал для  $\theta$

$$\hat{\theta} = \bar{X} \sim N\left(\theta, \frac{\sigma^2}{n}\right)$$

$$\frac{\sqrt{n}(\bar{X} - \theta)}{\sigma} \sim N(0, 1)$$

$$P\left(Z_{\frac{\alpha}{2}} < \frac{\sqrt{n}(\bar{X} - \theta)}{\sigma} < Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

Поскольку по середине стоит стандартное гауссовское распределение:  $Z_{\frac{\alpha}{2}} = -Z_{1-\frac{\alpha}{2}}$

$$P\left(-\frac{Z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}} - \bar{X} < -\theta < \frac{Z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}} - \bar{X}\right) = 1 - \alpha$$

$$P\left(\bar{X} - \frac{Z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}} < \theta < \bar{X} + \frac{Z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Итак, доверительный интервал:  $\left(\bar{X} - \frac{Z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}}, \bar{X} + \frac{Z_{1-\frac{\alpha}{2}}\sigma}{\sqrt{n}}\right)$

## Пример 2

$X_1, \dots, X_n \sim N(m, \theta_2^2)$ . Построить доверительный интервал для  $\theta_2^2$

$$\sum_{i=1}^n \left( \frac{x_i - m}{\theta_2} \right)^2 \sim \chi^2(n)$$

$$P\left(\chi_{n, \frac{\alpha}{2}}^2 < \frac{\sum_{i=1}^n (x_i - m)^2}{\theta_2^2} < \chi_{n, 1-\frac{\alpha}{2}}^2\right) = 1 - \alpha$$

$$P\left(\frac{\sum_{i=1}^n (x_i - m)^2}{\chi_{n, 1-\frac{\alpha}{2}}^2} < \theta_2^2 < \frac{\sum_{i=1}^n (x_i - m)^2}{\chi_{n, \frac{\alpha}{2}}^2}\right) = 1 - \alpha$$

Здесь  $\chi_{n, \alpha}^2$  — квантиль уровня  $\alpha$  распределения  $\chi^2(n)$

### Пример 3

Если нам неизвестны оба параметра  $N(\theta_1, \theta_2^2)$ . Заменяем  $m$  на  $\bar{X}$ : Доверительный интервал для  $\theta_2$ :

$$P\left(\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{\chi_{n, 1-\frac{\alpha}{2}}^2} < \theta_2^2 < \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{\chi_{n, \frac{\alpha}{2}}^2}\right) = 1 - \alpha$$

Доверительный интервал для  $\theta_1$ :

$$\frac{\sqrt{n} \left( \frac{\bar{X} - \theta}{\sigma} \right)}{\sqrt{\frac{1}{n-1} \sum \left( \frac{(x_i - \bar{X})}{\sigma} \right)^2}} = \frac{\sqrt{n}(\bar{X} - \theta_1)}{\tilde{S}} \sim t(n-1)$$

Обозначим  $t_{n, \alpha}$  квантиль уровня  $\alpha$  распределения  $t(n)$ , заметим, что  $t_{n, 1-\alpha} = t_{n, 1-\frac{\alpha}{2}}$

$$P(t_{n, 1-\frac{\alpha}{2}} < \frac{\sqrt{n}(\bar{X} - \theta_1)}{\tilde{S}} < t_{n, \frac{\alpha}{2}}) = 1 - \alpha$$

$$P\left(\bar{X} - \frac{\tilde{S} \cdot t_{n, 1-\frac{\alpha}{2}}}{\sqrt{n}} < \theta_1 < \bar{X} + \frac{\tilde{S} \cdot t_{n, 1-\frac{\alpha}{2}}}{\sqrt{n}}\right) = 1 - \alpha$$