

# Теория вероятностей и математическая статистика—2

Винер Даниил [@danya\\_vin](#)

Версия от 19 февраля 2025 г.

## Содержание

<b>1</b>	<b>Закон больших чисел. Центральная предельная теорема</b>	<b>2</b>
1.1	Закон больших чисел в форме Бернулли . . . . .	2
1.2	Центральная предельная теорема . . . . .	2
1.3	Теорема Муавра-Лапласа . . . . .	2
1.4	Неравенство Берри-Эссена . . . . .	3
<b>2</b>	<b>Многомерное нормальное распределение—1</b>	<b>4</b>
2.1	Одномерное нормальное распределение . . . . .	4
2.2	Многомерное нормальное распределение . . . . .	4
2.3	Свойства многомерного нормального распределения . . . . .	4
2.4	Условное нормальное распределение . . . . .	5
<b>3</b>	<b>Многомерное нормальное распределение—2</b>	<b>6</b>
3.1	Условное нормальное распределение . . . . .	6
3.2	Многомерная центральная предельная теорема . . . . .	6
<b>4</b>	<b>ТВА</b>	<b>7</b>
<b>5</b>	<b>Введение в математическую статистику</b>	<b>8</b>
<b>6</b>	<b>Оптимальные оценки, характеристики выборок</b>	<b>11</b>
6.1	Свойства оптимальных оценок . . . . .	11
6.2	Способы организации выборки . . . . .	11
6.3	Характеристики выборки . . . . .	12

# 1 Закон больших чисел. Центральная предельная теорема

## 1.1 Закон больших чисел в форме Бернулли

Пусть имеются некоторые случайные величины  $\xi_i = \begin{cases} 1, & p \\ 0, & 1-p \end{cases}$ , где  $p$  — вероятность, что какое-то событие произошло. Тогда  $\mathbb{E}[\xi] = p$ ,  $\mathbb{D}[\xi] = p(1-p) \leq \frac{1}{4}$

**Теорема.** Пусть  $\hat{p} = \frac{1}{n} \sum_{i=1}^n \xi_i$  — доля успехов в  $n$  испытаниях Бернулли, тогда  $\hat{p} \xrightarrow{p} p$

**Доказательство.** Распишем по неравенству Чебышёва:

$$\mathbb{P}(|\hat{p} - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0$$

### Пример

Пусть 87% новорожденных доживают до 50 лет. Тогда  $p = 0,87$  — вероятность дожить до 50. Рассмотрим  $n = 1000$  новорожденных

Определим с какой вероятностью данная случайная величина отклонится от своего математического ожидания не более, чем на 0,04 —  $\mathbb{P}(|\hat{p} - 0,87| \leq 0,04)$ . По Чебышёву:

$$\mathbb{P}(|\hat{p} - p| \leq 0,04) \geq 1 - \frac{\mathbb{D}[\hat{p}]}{(0,04)^2} = 1 - \frac{0,87 \cdot 0,13}{0,0016 \cdot 1000} = 0,929$$

## 1.2 Центральная предельная теорема

Рассмотрим сумму независимых одинаково распределенных случайных величин:

$$S_n = \xi_1 + \dots + \xi_n,$$

при этом существует  $\mathbb{D}[\xi_i] \leq c$ ,  $\mathbb{E}[\xi_i] = \mu$ ,  $\mathbb{D}[\xi_n] = \sigma^2$

Тогда,  $Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} Z$ , где  $Z \sim \mathcal{N}(0; 1)$  — имеет стандартное нормальное распределение

Функция плотности:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

## 1.3 Теорема Муавра-Лапласа

**Теорема.** Имеется  $\xi_i = \begin{cases} 1, & p \\ 0, & 1-p \end{cases}$ .  $S_n = \sum \xi_i$  — число успехов в  $n$  испытаниях. Тогда

$$Z_n = \frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} Z \sim \mathcal{N}(0; 1)$$

### Пример

Проходит суд над Бенджамином Споком. Из 300 человек 90 — женщины, которые симпатизируют Споку, при этом 12 присяжных будут судить Спока. Требуется определить мог ли отбор присяжных быть случайным.

Число успехов в данном случае — число женщин среди 300 присяжных. Будем считать, что  $p = 0.5$ , то есть половина женщин.

$$\mathbb{P}\left(\frac{S_{300} - 150}{\sqrt{0.5 \cdot 0.5 \cdot 300}} \leq \frac{90 - 150}{\sqrt{75}}\right) \simeq \Phi(-6.93) \simeq 2.3 \cdot 10^{-12}$$

Значит, практически невозможно случайным образом выбрать 90 или меньше женщин среди 300 присяжных при справедливом распределении, то есть отбор был предвзятым

## 1.4 Неравенство Берри-Эссена

$$|F_n - \Phi| \leq \frac{C_0 \cdot \mathbb{E}[|\xi_1 - \mu|^3]}{\sigma^3 \sqrt{n}}, \text{ где } \begin{cases} F_n - \text{функция распределения стандартизированной СВ} \\ C_0 - \text{константа} \\ \mathbb{E}[|\xi_1 - \mu|^3] - \text{третий абсолютный центральный момент} \end{cases}$$

### Пример

Пусть имеется  $n = 1000$  заключенных договоров страхования с 1 января на 1 год. С вероятностью  $p = 0.05$  произойдет страховой случай, выплаты по каждому договору — 2000 у.е.  $R$  — резерв страховой компании

Требуется определить какой должен быть размер резерва, чтобы страховая компания выполнила свои обязательства с вероятностью 0.99

$$S_n = 2000(\xi_1 + \dots + \xi_n), \xi_i \sim Bi(p = 0.05)$$

$$\mathbb{P}(S_n \leq R) = \mathbb{P}\left(\frac{\sum \xi_i - 0.05 \cdot 1000}{\sqrt{1000 \cdot 0.05 \cdot 0.95}} \leq \frac{\frac{R}{2000} - 0.05 \cdot 1000}{\sqrt{1000 \cdot 0.05 \cdot 0.95}}\right) \geq 0.99$$

Значит, требуется найти квантиль уровня 0.99. Он равен 2.33, тогда

$$\frac{\frac{R}{2000} - 0.05 \cdot 1000}{\sqrt{1000 \cdot 0.05 \cdot 0.95}} = 2.33 \implies R = 132117$$

То есть, для покрытия 99% страховых случаев у страховой компании резерв должен быть размером 132117 у.е. Напротив, для покрытия всех случаев  $R = 2000000$

## 2 Многомерное нормальное распределение—1

### 2.1 Одномерное нормальное распределение

**Определение.** Случайная величина имеет нормальное распределение  $X \sim \mathcal{N}(\mu, \sigma^2)$ , если функция плотности равна

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

### 2.2 Многомерное нормальное распределение

**Определение.** Пусть случайные величины  $z_1, \dots, z_n$  независимы и  $\sim N(0, 1)$ . Тогда  $z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$  имеет многомерное нормальное распределение  $N(0, I)$ , где  $I$  — единичная матрица

Функция плотности:

$$f_Z(z) = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2} \sum_{i=1}^n z_i^2} = \frac{1}{(\sqrt{2\pi})^n} e^{-0.5 Z^T Z}$$

**Примечание.** Пусть  $Z \sim N(0, I)$ ,  $A \in \text{Mat}_{k \times n}$  — матрица полного ранга и  $k < n$ , то есть  $\text{rank} A = k$ . Тогда

$$\begin{aligned} Y &= AZ + b \sim N(b, AA^T) \\ f_Y(y) &= \frac{1}{|\det A|} f_Z(A^{-1}(y-b)) \\ &= \frac{1}{(\sqrt{2\pi})^n |\det A|} e^{-0.5(y-b)^T (A^{-1})^T A^{-1}(y-b)} \\ &\text{пусть } AA^T = C \\ &= \frac{1}{(\sqrt{2\pi})^n} \frac{1}{\sqrt{|C|}} e^{-0.5(y-b)^T C^{-1}(y-b)} \end{aligned}$$

**Определение.** Случайная величина  $Y \sim N(b, C)$ , если

$$f_Y(y) = \frac{1}{(\sqrt{2\pi})^n} \frac{1}{\sqrt{|C|}} e^{-0.5(y-b)^T C^{-1}(y-b)}$$

**Определение.** Случайный вектор  $Y \sim N(0, C)$ , если  $\forall a_1, \dots, a_n$

$$a_1 Y_1 + a_2 Y_2 + \dots + a_n Y_n$$

либо  $N(0, \cdot)$  либо const

### 2.3 Свойства многомерного нормального распределения

Пусть  $Y \sim N(b, C)$

1.  $\mathbb{E}[Y] = b, \text{cov}(Y) = C$

**Доказательство.**  $Y = AZ + b, Z \sim N(0, I)$

$$\text{cov}(Y) = \mathbb{E}[(AZ + b - \mathbb{E}[AZ + b])(AZ + b - \mathbb{E}[AZ + b])^T] = A \text{cov} Z A^T = AA^T = C$$

2. Любое линейное невырожденное преобразование многомерного нормального дает многомерный нормальный вектор

$$\forall B, a : BY + a \sim N(Bb + a, BCB^T)$$

3.  $\forall$  подвектор нормального вектора нормален

4. Если  $Y \sim N(b, D)$ , то его компоненты независимы

**Примечание.** Некоррелированность = независимость

**Доказательство.**

$$\begin{aligned} f_Y(y) &= \frac{1}{(\sqrt{2\pi})^n} e^{-0.5(y-b)^T D^{-1}(y-b)} \\ &= \frac{1}{(\sqrt{2\pi})^n} e^{-0.5 \sum \left( \frac{y_i - b_i}{\sigma_i} \right)^2} \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-0.5 \left( \frac{y_i - b_i}{\sigma_i} \right)^2} \end{aligned}$$

**Пример.**  $Y_1 \sim N(0, 1)$ ,  $\lambda = \begin{cases} 1, & p = 0.5 \\ -1, & p = 0.5 \end{cases}$ ,  $Y_2 = 2Y_1$

$$\begin{aligned} \mathbb{P}(Y_2 \leq y) &= \mathbb{P}(Y_1 \leq y | \alpha = 1) \cdot \mathbb{P}(\alpha = 1) + \mathbb{P}(-Y_1 \leq y | \alpha = -1) \cdot \frac{1}{2} \\ &= \Phi(y) \end{aligned}$$

$cov(Y_1, Y_2) = cov(Y_1, 2Y_1) = \mathbb{E}[\alpha Y_1^2] - \mathbb{E}[Y] \mathbb{E}[\alpha Y_1] = 0$ . То есть они не коррелированы

## 2.4 Условное нормальное распределение

Имеется случайный вектор  $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$ , пишут  $\Phi_2(z_1, z_2; \rho)$

Допустим, что  $z_1$  фиксирован, тогда  $z_2 | z_1 = z \sim N(\rho z, 1 - \rho^2)$

$z_2 = \rho z_1 + u$ , где  $z_1$  и  $u$  независимы и  $u \sim N(., .)$

## 3 Многомерное нормальное распределение—2

### 3.1 Условное нормальное распределение

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right), \text{ пишут } \Phi_2(z_1, z_2; \rho)$$

Допустим, что  $z_1$  фиксирован, тогда  $z_2|z_1 = z \sim N(\rho z, 1 - \rho^2)$

**Утверждение.**  $z_2 = \rho z_1 + u$ , где  $z_1$  и  $u$  независимы и  $(u, z_1) \sim N(., .)$

**Доказательство.**  $u = z_1 - \rho z_1 \implies (z_1, u) = (z_1, z_2 - \rho z_1) = \begin{pmatrix} 1 & 0 \\ -\rho & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 - \rho^2 \end{pmatrix} \right)$

$$\text{cov}(z_1, u) = A \cdot \text{cov}(z_1, z_2) \cdot A^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 - \rho^2 \end{pmatrix} \quad \square$$

**Доказательство свойства.**  $\mathbb{E}[z_2] = \rho \mathbb{E}[z_1] + \mathbb{E}[u] = 0$ ,  $\mathbb{D}[z_2] = \rho^2 \mathbb{D}[z_1] + \mathbb{D}[u] = \rho^2 + 1 - \rho^2 = 1$

$$(z_2|z_1 = z) = \rho z + u \implies \begin{cases} \mathbb{E}[z_2|z_1 = z] = \rho z + \mathbb{E}[u] = \rho z \\ \mathbb{D}[z_2|z_1 = z] = \mathbb{D}[u] = 1 - \rho^2 \end{cases} \quad \square$$

**Примечание.** Пусть вектор  $Y$  такой, что  $AY \sim N(., .)$  (многомерное нормальное), меньшей размерности, чем  $Y$ , тогда говорят, что  $Y$  имеет обобщенное нормальное распределение

**Примечание.** Двумерная Гауссова копула представима в виде  $\Phi_2(\Phi^{-1}(F_1(u_1)), \Phi^{-1}(F_2(u_2)); \rho)$

### 3.2 Многомерная центральная предельная теорема

**Теорема.** Пусть  $\xi_1^{(1)}, \dots, \xi_n^{(n)}$  — последовательность независимых одинаково распределенных случайных векторов, у каждого из которых  $\mathbb{E}[\xi^{(k)}] = b \forall k$ ,  $\text{cov}(\xi^{(k)}) = c$ ,  $\det C > 0$ .

Обозначим  $S_n = \xi_1^{(1)} + \dots + \xi_n^{(n)}$  — вектор частичных сумм. Тогда, при  $n \rightarrow \infty$  последовательность  $\eta^{(n)}$ , где  $\eta^{(n)} = \frac{S_n - nb}{\sqrt{n}}$  сходится по распределению к вектору  $\eta \sim N(\vec{0}, C)$

4 TBA

## 5 Введение в математическую статистику

- Имеется  $n$  независимых случайных величин  $X_1, \dots, X_n$ , которые имеют одинаковые функции распределения:  $F_{X_1}(x) = \dots = F_{X_n}(x) = F(x)$
- Пусть функция распределения  $F(x)$  зависит от некоторого вектора неизвестных параметров  $\theta = (\theta_1, \dots, \theta_r)$
- $F(x) = F(x; \theta)$ , где  $x$  — переменная, а  $\theta$  — вектор неизвестных параметров
- $\Theta$  — множество допустимых значений вектора  $\theta$

**Пример.** Если  $X_i \sim N(\mu, \sigma^2)$ , то  $\theta = (\mu, \sigma^2) \in (-\infty, +\infty) \times (0; +\infty)$

**Определение.** Случайной выборкой объема  $n$  наблюдений из распределения с функцией распределения  $F(x; \theta)$  называется случайный вектор  $X = (X_1, \dots, X_n)$ , компоненты которого удовлетворяют следующим условиям

- случайные величины  $X_1, \dots, X_n$  — независимы
- случайные величины  $X_1, \dots, X_n$  имеют одну и ту же функцию распределения  $F(x; \theta)$ :

$$F_{X_1}(x; \theta) = \dots = F_{X_n}(x; \theta) = F(x; \theta)$$

**Примечание.** Продифференцировав эти равенства, получаем, что все функции плотностей распределения равны

**Примечание.** Если все величины  $X_1, \dots, X_n$  дискретны, то они должны иметь одинаковые таблицы распределения

**Примечание.** При  $i \neq j$ :  $\text{cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i] \cdot \mathbb{E}[X_j] = 0$ , так как  $X_i$  и  $X_j$  независимы

Имеются случайные величины  $X_1(\omega), \dots, X_n(\omega)$ . Пусть произошел вероятностный эксперимент, в результате которого реализовался исход  $\omega_0 \in \Omega$ . То есть

$$X_1(\omega_0), \dots, X_n(\omega_0)$$

Тогда, вектор  $x = (X_1(\omega_0), \dots, X_n(\omega_0))$  называется *реализацией случайной выборки*

**Пример.**  $\Omega = \{a, b, c, d, e, f, g, h\}$ ,  $\mathcal{F}$  = все подмножества  $\Omega$

$\omega$	a	b	c	d	e	f	g	h
$\mathbb{P}(\omega)$	$p^3$	$p^2q$	$p^2q$	$pq^2$	$p^2q$	$pq^2$	$pq^2$	$q^3$
$X_1(\omega)$	1	1	1	1	0	0	0	0
$X_2(\omega)$	1	1	0	0	1	1	0	0
$X_3(\omega)$	1	0	1	0	1	0	1	0

- $\mathbb{P}(X_1 = 1) = p^3 + p^2q + p^2q + pq^2 = p(p^2 + pq + pq + q^2) = p \implies X_1 \sim Be(p)$
- $\mathbb{P}(X_2 = 1) = p^3 + p^2q + p^2q + pq^2 = p(p^2 + pq + pq + q^2) = p \implies X_2 \sim Be(p)$
- $\mathbb{P}(X_3 = 1) = p^3 + p^2q + p^2q + pq^2 = p(p^2 + pq + pq + q^2) = p \implies X_3 \sim Be(p)$

$$\begin{aligned} \mathbb{P}(\{X_1 = 1\} \cap \{X_2 = 1\} \cap \{X_3 = 1\}) &= p^3 \\ \mathbb{P}(X_1 = 1) \cdot \mathbb{P}(X_2 = 1) \cdot \mathbb{P}(X_3 = 1) &= p^3 \end{aligned}$$

Рассуждая аналогично, перебираем оставшиеся 7 случаев и получаем, что  $X_1, X_2, X_3$  — независимы

Пусть  $\omega_0 = c$ , тогда  $(X_1(\omega_0), X_2(\omega_0), X_3(\omega_0)) = (1, 0, 1)$

- Пусть  $X = (X_1, \dots, X_n)$  — случайная выборка
  - Для каждого  $\omega \in \Omega$  расположим числа  $X_1(\omega), \dots, X_n(\omega)$  в порядке возрастания
  - Получим набор чисел  $X_{(1)}(\omega) \leq \dots \leq X_{(n)}(\omega)$ , где  $(i)$  означает уже отсортированный номер
- При этом  $X_{(1)}(\omega) = \min(X_1(\omega), \dots, X_n(\omega))$ , а  $X_{(n)}(\omega) = \max(X_1(\omega), \dots, X_n(\omega))$



**Определение.** Набор случайных величин  $X_{(1)}, \dots, X_{(n)}$  называется вариационным рядом

**Определение.**  $\bar{X} := \frac{X_1 + \dots + X_n}{n}$  — выборочное среднее

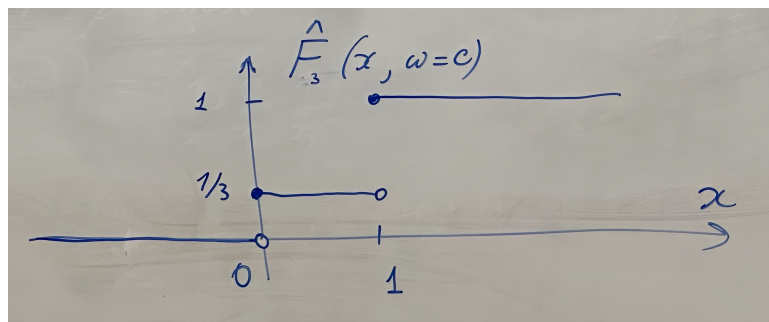
**Определение.**  $s^2 := \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$  — выборочная дисперсия

**Определение.**  $\hat{\sigma}^2 := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$  — исправленная выборочная дисперсия

**Определение.**

$$\begin{aligned} \hat{F}_n(x) &= \frac{\text{число элементов случайной выборки, которые нестрого меньше } x}{n} \\ &= \frac{\#\{i \in \{1, \dots, n\} : X_i \leq x\}}{n} \end{aligned}$$

**Пример.** Рассмотрим  $\hat{F}_3(x; \omega = c)$  и выборку  $(1, 0, 1)$ . Тогда график будет выглядеть так



**Утверждение.** Пусть  $X = (X_1, \dots, X_n)$  — случайная выборка, компоненты которой имеют конечное матожидание. Тогда

$$\mathbb{E}[\bar{X}] = \mathbb{E}[X_i]$$

**Утверждение.** Пусть  $X = (X_1, \dots, X_n)$  — случайная выборка, компоненты которой имеют конечные дисперсии. Тогда

$$\mathbb{D}[\bar{X}] = \frac{\mathbb{D}[X_i]}{n}$$

**Доказательство.**

$$\begin{aligned} \mathbb{D}[\bar{X}] &= \mathbb{D}\left[\frac{1}{n} \sum_{i=1}^n X_i\right] \\ &= \frac{1}{n^2} \mathbb{D}\left[\sum_{i=1}^n X_i\right] \\ &= \frac{1}{n^2} \sum_{i=1}^n \mathbb{D}[X_i] \\ &= \frac{1}{n^2} n \mathbb{D}[X_i] \\ &= \frac{\mathbb{D}[X_i]}{n} \end{aligned}$$

**Определение.** Пусть  $X = (X_1, \dots, X_n)$  — случайная выборка из распределения с функцией распределения  $F(x; \theta)$ , где  $\theta$  — неизвестный параметр

Оценкой параметра  $\theta$  называется случайная величина  $\hat{\theta}$ , которая является произвольной борелевской функцией от элементов случайной выборки, то есть

$$\hat{\theta} = g(X_1, \dots, X_n),$$

где  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  — произвольная борелевская функция

**Пример.**  $X = (X_1, \dots, X_n)$  — случайная выборка из распределения Бернулли с параметром  $p \in (0; 1)$

$$\hat{p} = \frac{X_1 + \dots + X_n}{n} \text{ — оценка параметра } p$$

**Определение.** Оценка  $\hat{\theta}$  неизвестного параметра  $\theta \in \Theta$  называется *несмещенной*, если

$$\mathbb{E}[\hat{\theta}] = \theta \quad \forall \theta \in \Theta$$

**Пример.**  $X = (X_1, \dots, X_n)$  — случайная выборка из распределения Бернулли с параметром  $p \in (0; 1)$ . Определим, является ли  $\hat{p} = \frac{X_1 + \dots + X_n}{n}$  несмещенной оценкой для параметра  $p$

$$\begin{aligned} \mathbb{E}[\hat{p}] &= \frac{\mathbb{E}[X_1 + \dots + X_n]}{n} \\ &= \frac{\mathbb{E}[X_1] + \dots + \mathbb{E}[X_n]}{n} \\ &= \frac{np}{p} \\ &= p \end{aligned}$$

То есть  $\hat{p}$  — несмещенная оценка

**Определение.** Говорят, что последовательность случайных величин  $(X_n)_{n=1}^{\infty}$  сходится по вероятности к случайной величине  $X$ , если

$$\forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$$

Обозначение:  $X_n \xrightarrow{\mathbb{P}} X$  при  $n \rightarrow \infty$  ИЛИ  $p \lim_{n \rightarrow \infty} X_n = x$

**Определение.** Оценка  $\hat{\theta}_n$  называется *состоятельной оценкой* неизвестного параметра  $\theta \in \Theta$ , если

$$\forall \theta \in \Theta \quad \hat{\theta}_n \xrightarrow{\mathbb{P}} \theta \text{ при } n \rightarrow \infty$$

**Пример.** Проверим, что  $\hat{p} = \frac{X_1 + \dots + X_n}{n}$  является состоятельной. По ЗБЧ:

$$\hat{p} = \frac{X_1 + \dots + X_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}[X_i] = p$$

Значит,  $\hat{p}$  — состоятельная оценка

## 6 Оптимальные оценки, характеристики выборок

### 6.1 Свойства оптимальных оценок

**Определение.** Оценкой для параметра будет некоторая функция  $\hat{\Theta} = T(X_1, \dots, X_n)$

**Определение.** Оценка *состоятельна*, если  $\hat{\Theta} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \Theta$

**Определение.** Оценка *желательно* должна обладать свойством *несмещенности*, то есть  $\mathbb{E} [\hat{\Theta}] = \Theta$

**Пример.**  $X_1, \dots, X_n$ . Построим две оценки

1.  $\hat{\mu}_1 = \frac{1}{4} (X_1, \dots, X_4)$
2.  $\hat{\mu}_2 = \frac{1}{8} X_1 + \frac{1}{8} X_2 + \frac{1}{4} X_3 + \frac{1}{2} X_4$

Первая оценка лучше как минимум потому, что в ней веса каждого наблюдения равны

**Определение.** Среднеквадратичная ошибка оценки — это

$$\begin{aligned} \mathbb{E} [\hat{\Theta} - \Theta]^2 \\ \mathbb{E} [\hat{\Theta} - \Theta]^2 &= \mathbb{E} \left[ \left( \hat{\Theta} - \mathbb{E} [\hat{\Theta}] \right) + \underbrace{\left( \mathbb{E} [\hat{\Theta}] - \Theta \right)}_{b(\Theta) - \text{смещение}} \right]^2 \\ &= \underbrace{\mathbb{E} [\hat{\Theta} - \mathbb{E} [\hat{\Theta}]]^2}_{\mathbb{D}[\hat{\Theta}]} + b^2(\Theta) + 2 \underbrace{\mathbb{E} [\hat{\Theta} - \mathbb{E} [\hat{\Theta}]]}_{=0} \cdot b(\Theta) \\ &= \underbrace{\mathbb{E} [\hat{\Theta} - \mathbb{E} [\hat{\Theta}]]^2}_{\mathbb{D}[\hat{\Theta}]} + b^2(\Theta) \end{aligned}$$

**Определение.**  $\hat{\Theta}_1$  эффективнее, чем  $\hat{\Theta}_2$ , если

$$\begin{aligned} \mathbb{E} [\hat{\Theta}_1 - \Theta]^2 &\leq \mathbb{E} [\hat{\Theta}_2 - \Theta]^2 \text{ причем } \exists \Theta_0 : \\ \mathbb{E} [\hat{\Theta}_1 - \Theta_0]^2 &< \mathbb{E} [\hat{\Theta}_2 - \Theta_0]^2 \end{aligned}$$

**Определение.** Оценка  $\hat{\Theta}$  называется эффективной (=оптимальной, *the best*) в классе  $K$ , если  $\forall \bar{\Theta} \in K$

$$\begin{aligned} \mathbb{E} [\hat{\Theta} - \Theta]^2 &\leq \mathbb{E} [\bar{\Theta} - \Theta]^2 \text{ причем } \exists \Theta_0 : \\ \mathbb{E} [\hat{\Theta} - \Theta_0]^2 &< \mathbb{E} [\bar{\Theta} - \Theta_0]^2 \end{aligned}$$

Класс  $K$  обладает смещением  $b(\Theta)$

**Определение.** Оценка  $\hat{\Theta}$  называется эффективной в классе несмещенных оценок, если  $\forall \bar{\Theta} \in K$

$$\begin{aligned} \mathbb{D} [\hat{\Theta}] &\leq \mathbb{D} [\bar{\Theta}], \text{ причем } \exists \Theta_0 : \\ \mathbb{D} [\hat{\Theta}] &< \mathbb{D} [\bar{\Theta}] \text{ для } \Theta_0 \end{aligned}$$

### 6.2 Способы организации выборки

**Определение.** Генеральная совокупность — выборка, которая в теории может быть доступна, но на практике вряд ли ее можно собрать. Например, точное население РФ

1. Простой случайный отбор — практически нереализуемо
2. Отбор с помощью механистической процедуры, не влияющий на случайность  
Например, на автоматическом производстве надо посчитать процент брака. Тогда, для выборки можно ограничиться промежутком времени 14:20-14:30
3. Стратифицированные выборки — выборка производится из страт, то есть берут людей из всех социальных слоев населения
4. Комбинирование

Таким образом, важно следить, чтобы выборка обладала всеми или большинством свойств генеральной совокупности

### 6.3 Характеристики выборки

1. Вариационный ряд — наблюдения, упорядоченные по возрастанию

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)} \text{ — случайные величины}$$

При этом,  $X_{(1)}$  и  $X_{(n)}$  — экстремальные статистики

2. Гистограмма и выборочная функция распределения

**Определение.** Выборочная (эмпирическая) функция распределения — это

$$\hat{F}_n(x) = \frac{\sum_{i=1}^n I\{X_i \leq x\}}{n},$$

где  $I\{X_i \leq x\}$  — индикаторная функция, показывающая сколько наблюдений меньше либо равны какому-то  $x$

**Определение.**  $\hat{p}_i = \frac{m(\Delta i)}{|\Delta i| \cdot n}$ , где  $m(\Delta i)$  — число наблюдений, попавших в интервал  $\Delta i$ ,  $|\Delta i|$  — длина интервала,  $n$  — число наблюдений

Гистограмма — тоже своего рода функция плотности