

Теория вероятностей и математическая статистика—2

Винер Даниил [@danya_vin](#)

Версия от 14 января 2025 г.

Содержание

1	Закон больших чисел. Центральная предельная теорема	2
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

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 $R = 2000000$