§8. Закон больших чисел и центральная предельная теорема

ПАВЕЛ Е. РЯБОВ[1]

[1]Финансовый университет при Правительстве Российской Федерации

PERyabov@fa.ru

14 мая 2024 г.

§8. Закон больших чисел и центральная предельная теорема

План лекции

- Неравенства Маркова и Чебышёва
- Различные формы закона больших чисел
- Центральная предельная теорема
- Неравенство Берри – Эссеена

§8. Закон больших чисел и центральная предельная теорема

План лекции

- Неравенства Маркова и Чебышёва
- Различные формы закона больших чисел
- Центральная предельная теорема
- Неравенство Берри – Эссеена

§8. Закон больших чисел и центральная предельная теорема

План лекции

- Неравенства Маркова и Чебышёва
- Различные формы закона больших чисел
- Центральная предельная теорема
- Неравенство Берри – Эссеена

§8. Закон больших чисел и центральная предельная теорема

План лекции

- Неравенства Маркова и Чебышёва
- Различные формы закона больших чисел
- Центральная предельная теорема
- Неравенство Берри – Эссеена

Пафнутий Львович Чебышёв и Андрей Андреевич Марков

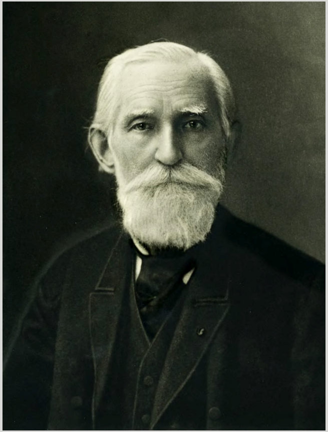

Рис.1. *Пафнутий Львович Чебышёв (4 (16) мая 1821, Окатово, Калужская губерния, Российская империя — 26 ноября (8 декабря) 1894, Санкт-Петербург, Российская империя)*

Рис.2. *Андрей Андреевич Марков (2 (14) июня 1856, Рязань — 20 июля 1922, Петроград)*

Следующее неравенство часто называют собственно неравенством Чебышёва, хотя в такой форме оно появилось впервые, видимо, в работах Андрея Андреевича Маркова (см. «Исчисление вероятностей», 1913 г.).

Лемма (неравенство Маркова)

Если X — неотрицательная случайная величина, то для любого $a > 0$ справедливо неравенство

$$\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}(X)}{a}. \tag{1}$$

Доказательство.

Рассмотрим новую случайную величину

$$Y = \begin{cases} a, \text{ если } X \geqslant a, \\ 0, \text{ если } X < a. \end{cases}$$

Тогда $X \geqslant Y$ (Почему?). Откуда, $\mathbb{E}(X) \geqslant \mathbb{E}(Y)$. Но $\mathbb{E}(Y) = a \cdot \mathbb{P}(X \geqslant a) + 0 \cdot \mathbb{P}(X < a) = a \cdot \mathbb{P}(X \geqslant a) \leqslant \mathbb{E}(X)$. Следовательно, $\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}(X)}{a}$. $\square$

В 1853 г. И.Бьенеме (I.Bienaymè) и в 1866 г., независимо от него, П.Л. Чебышёв прямыми методами доказали неравенство, которое будет удобно получить как следствие неравенства Маркова.

Следствие (неравенство Чебышёва-Бъенеме)

Если $\mathbb{E}\big(X^2\big) < \infty$, *то для любого* $\varepsilon > 0$ *справедливо неравенство*

$$\mathbb{P}(|X - \mathbb{E}(X)| \geqslant \varepsilon) \leqslant \frac{\mathrm{Var}(X)}{\varepsilon^2}. \tag{2}$$

Доказательство.

Воспользуемся неравенством Маркова (1) для $Y = [X - \mathbb{E}(X)]^2$.

$$\mathbb{P}(|X - \mathbb{E}(X)| \geqslant \varepsilon) = \mathbb{P}\left([X - \mathbb{E}(X)]^2 \geqslant \varepsilon^2\right) \leqslant \frac{\mathbb{E}\big([X-\mathbb{E}(X)]^2\big)}{\varepsilon^2} = \frac{\mathrm{Var}(X)}{\varepsilon^2}.$$

$\square$

В качестве следствия получаем так называемое «правило трёх сигм».

Следствие (общее правило 3σ)

Для любой случайной величины, для которой $\mathbb{E}\left[X^2\right] < \infty$, *выполняется неравенство*

$$\mathbb{P}(|X - \mathbb{E}(X)| < 3\sigma) \geqslant \tfrac{8}{9} = 0,889. \tag{3}$$

Доказательство.

Перейдём в неравенстве (2) к противоположному событию.

$$\mathbb{P}(|X - \mathbb{E}(X)| \geqslant \varepsilon) = 1 - \mathbb{P}(|X - \mathbb{E}(X)| < \varepsilon) \leqslant \frac{\mathbf{Var}(X)}{\varepsilon^2}.$$

Следовательно,

$$\boxed{\mathbb{P}(|X - \mathbb{E}(X)| < \varepsilon) \geqslant 1 - \frac{\mathbf{Var}(X)}{\varepsilon^2}}. \tag{4}$$

Полагая $\varepsilon = 3\sigma$ в неравенстве (4), получаем

$$\mathbb{P}(|X - \mathbb{E}(X)| < 3\sigma) \geqslant 1 - \frac{\sigma^2}{9\sigma^2} = 1 - \frac{1}{9} = \frac{8}{9}.$$

Замечание

Для каждого распределения указанная в неравенстве (3) величина вероятности своя, например, если $X \sim \mathcal{N}(\mu, \sigma^2)$, то $\mathbb{P}(|X - \mathbb{E}(X)| < 3\sigma) = 0,997 > \frac{8}{9}$.

Упражнение

Найдите $\mathbb{P}(|X - \mathbb{E}(X)| < 3\sigma)$, если случайная величина X имеет

1. равномерное распределение на отрезке $[a, b]$;
2. показательное распределение с параметром $\lambda > 0$;
3. распределение Бернулли с параметром $p = \frac{1}{2}$.

Различные формы закона больших чисел

Теорема (закон больших чисел в форме Чебышёва)

Пусть $X_1, X_2, \ldots$ — бесконечная последовательность независимых случайных величин с одним и тем же математическим ожиданием μ, $\mathbb{E}(X_k) = \mu$, и дисперсии которых ограничены в совокупности, т.е. $\mathrm{Var}(X_k) < C$ для любого $k = 1, 2, \ldots$. Тогда для любого $\varepsilon > 0$ имеет место сходимость по вероятности

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{X_1+X_2+\ldots+X_n}{n} - \mu\right| < \varepsilon\right) = 1. \tag{5}$$

Эта теорема служит обоснованием правила среднего арифметического в теории измерений. ЗБЧ утверждает, что среднее арифметическое большого числа случайных слагаемых «стабилизируется» с ростом этого числа. Как сильно каждая случайная величина не отклонялась от своего среднего значения, при суммировании эти отклонения «взаимно гасятся», так что среднее арифметическое приближается к постоянной величине.

Доказательство.

Обозначим через $S_n = X_1 + X_2 + \ldots + X_n$ сумму первых n случайных величин $X_1, X_2, \ldots, X_n$, а через $\frac{S_n}{n} = \frac{X_1+X_2+\ldots+X_n}{n}$ — их среднее арифметическое. Тогда

$$\mathbb{E}\left(\frac{S_n}{n}\right) = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(X_k) = \frac{1}{n}\cdot n\mathbb{E}(X_1) = \mu.$$

А поскольку $X_1, X_2, \ldots, X_n$ — независимые случайные величины, то

$$\mathbf{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\sum_{k=1}^{n}\mathbf{Var}(X_k) \leqslant \frac{1}{n^2}\cdot nC = \frac{C}{n}.$$

В силу неравенства Чебышёва для любого $\varepsilon > 0$ имеем

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| < \varepsilon\right) \geqslant 1 - \frac{\mathbf{Var}\left(\frac{S_n}{n}\right)}{\varepsilon^2},$$

или

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right) \geqslant 1 - \frac{C}{n\varepsilon^2}.$$

Таким образом, для любого $\varepsilon > 0$ выполняется двойное неравенство

$$1 - \frac{C}{n\varepsilon^2} \leqslant \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right) \leqslant 1.$$

Тогда при $n \to \infty$ и получаем сходимость по вероятности (5). □

Замечание

Мы не только доказали сходимость по вероятности, но и получили оценку для вероятности отклонения среднего арифметического числа независимых и одинаково распределенных случайных величин от $\mathbb{E}(X_1)$ не более, чем на заданное число $\varepsilon > 0$:

$$\mathbb{P}\left(\left|\frac{X_1+X_2+\ldots+X_n}{n} - \mathbb{E}(X_1)\right| < \varepsilon\right) \geqslant 1 - \frac{\mathrm{Var}(X_1)}{n\varepsilon^2}. \tag{6}$$

Упражнение

Пусть $X_1, X_2, \ldots$ — бесконечная последовательность случайных величин, дисперсии которых ограничены одной и той же постоянной C, а ковариации любых случайных величин X_i и X_j $(i \neq j)$, не являющихся соседними в последовательности, равны нулю. Удовлетворяет ли эта последовательность ЗБЧ? Ответ обоснуйте.

Получим в качестве следствия из ЗБЧ Чебышёва закон больших чисел Я. Бернулли (1713). В отличие от доказанного через полтора столетия ЗБЧ Чебышёва, описывающего предельное поведение среднего арифметического случайных величин с *произвольными* распределениями, ЗБЧ Бернулли — утверждение только для *схемы Бернулли*.

Теорема (закон больших чисел в форме Бернулли)

Пусть производится n независимых испытаний, в каждом из которых с вероятностью p может наступить некоторое событие A. Пусть случайная величина $k(A)$ — число наступлений события A в n испытаниях. Тогда для любого $\varepsilon > 0$ имеет место сходимость по вероятности

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{k(A)}{n} - p\right| < \varepsilon\right) = 1. \tag{7}$$

При этом справедливо неравенство

$$\mathbb{P}\left(\left|\frac{k(A)}{n} - p\right| < \varepsilon\right) \geqslant 1 - \frac{p(1-p)}{n\varepsilon^2}. \tag{8}$$

Доказательство.

Представим случайную величину $k(A)$ — число наступлений события A в n независимых испытаниях в виде $k(A) = I_1 + I_2 + \ldots + I_n$, где $I_k = I(A)$ — индикатор события A в k испытании, т.е. случайная величина, которая принимает значение 1, если при k-ом испытании событие A происходит с вероятностью p, и значение 0 — в противном случае. Тогда

$$\mathbb{E}(I_k) = 1 \cdot p = p, \mathrm{Var}(I_k) = p(1-p)$$

для любого $k = 1, \ldots, n$. Таким образом, все условия теоремы Чебышёва выполнены, так что для среднего арифметического случайных величин $I_1, I_2, \ldots, I_n$, т.е. для $\frac{k(A)}{n}$, имеет место сходимость по вероятности (7). □

Центральная предельная теорема

Сформулируем центральную предельную теорему только для последовательности ***независимых и одинаково распределенных*** случайных величин.

Теорема

Пусть $\{X_k, k \geqslant 1\}$ – последовательность ***одинаково распределенных и независимых*** случайных величин, причём $\mathbb{E}(X_k) = \mu < \infty$ и $\mathrm{Var}(X_k) = \sigma^2 < \infty$, и пусть $S_n = \sum_{k=1}^{n} X_k$. Тогда при неограниченном увеличении n закон распределения ***центрированной и нормированной*** случайной величины $S_n' = \frac{S_n - n\mu}{\sigma\sqrt{n}}$ приближается к стандартному нормальному закону распределения, т.е. для любого $x \in \mathbb{R}$ имеет место сходимость по распределению

$$\lim_{n\to\infty} F_{S_n'}(x) = \lim_{n\to\infty} \mathbb{P}\left(S_n' \leqslant x\right) = \lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leqslant x\right) = \Phi(x), \qquad (9)$$

где $\Phi(x) = \frac{1}{2} + \Phi_0(x)$ — функция стандартного нормального закона распределения, а $\Phi_0(x)$ — функция Лапласа.

Их этой теоремы следует, что для промежутка Δ любого вида предел вероятности попадания нормированной частичной суммы в Δ существует и

$$\lim_{n\to\infty} \mathbb{P}(S'_n \in \Delta) = \mathbb{P}(Z \in \Delta),$$

где $Z \sim \mathcal{N}(0, 1)$ – стандартная нормальная случайная величина. В частности, для промежутка $\Delta = (a, b)$ или $\Delta = [a, b]$ имеем

$$\lim_{n\to\infty} \mathbb{P}(S'_n \in \Delta) = \Phi_0(b) - \Phi_0(a),$$

где $\Phi_0(x)$ – функция Лапласа.

Пример

Для независимых случайных величин $X_1, X_2, \ldots,$ принимающих с равной вероятностью значения $1, 4$ и 7, найдите предел $\lim\limits_{n\to\infty} \mathbb{P}(X_1 + \ldots + X_n < 4n + \sqrt{n})$.

Решение. Сначала найдем математическое ожидание $\mathbb{E}(X_i)$ и дисперсию $\mathrm{Var}(X_i)$: $\mathbb{E}(X_i) = 4$, $\mathrm{Var}(X_i) = 6$. Тогда искомый предел равен

$$\lim_{n\to\infty} \mathbb{P}(X_1 + \ldots + X_n < 4n + \sqrt{n}) =$$
$$= \lim_{n\to\infty} \mathbb{P}\left(\frac{X_1 + \ldots + X_n - 4n}{\sqrt{n}\sqrt{6}} < \frac{1}{\sqrt{6}}\right) =$$
$$= \frac{1}{2} + \Phi_0\left(\frac{1}{\sqrt{6}}\right) = 0.5 + \Phi_0(0{,}40825) = 0{,}65845.$$

Ответ: $0{,}65845$.

Пример

Для независимых, распределенных по геометрическому закону случайных величин $X_1, X_2, \ldots$, *найдите предел* $\lim\limits_{n\to\infty} \mathbb{P}(X_1 + \ldots + X_n > 6n + \sqrt{3n})$, *если известно, что* $\mathbb{E}(X_i) = 6$.

Решение. Для случайной величины X_i, распределенной по геометрическому закону, дисперсия равна 30. Следовательно, искомый предел равен

$$\lim_{n\to\infty} \mathbb{P}(X_1 + \ldots + X_n > 6n + \sqrt{3n}) =$$
$$= \lim_{n\to\infty} \mathbb{P}\left(\frac{X_1 + \ldots + X_n - 6n}{\sqrt{n}\sqrt{30}} > \frac{1}{\sqrt{10}}\right) =$$
$$= \frac{1}{2} - \Phi_0\left(\frac{1}{\sqrt{10}}\right) = 0.5 - \Phi_0(0.31623) = 0,37591.$$

Ответ: $0,37591$.

Пример

Для независимых случайных величин $X_1, X_2, \ldots$, равномерно распределенных на отрезке $[3, 12]$, найдите предел $\lim\limits_{n\to\infty} \mathbb{P}\left(X_1 + \ldots + X_n > \frac{15}{2}n + \sqrt{n}\right)$.

Решение. Для равномерно распределенной на отрезке $[3, 12]$ случайной величины X_i математическое ожидание и дисперсия соответственно равны $\frac{15}{2}$ и $\frac{27}{4}$. Поэтому искомый предел равен

$$\lim_{n\to\infty} \mathbb{P}\left(X_1 + \ldots + X_n > \frac{15}{2}n + \sqrt{n}\right) =$$

$$= \lim_{n\to\infty} \mathbb{P}\left(\frac{X_1 + \ldots + X_n - \frac{15}{2}n}{\sqrt{n} \cdot \frac{3\sqrt{3}}{2}} > \frac{2}{3\sqrt{3}}\right) =$$

$$= \frac{1}{2} - \Phi_0\left(\frac{2}{3\sqrt{3}}\right) = 0.5 - \Phi_0(0.3849) = 0,35016.$$

Ответ: $0,35016$.

Замечание

Центральной предельной теоремой пользуются для приближенного вычисления вероятностей, связанных с суммами большого числа независимых и одинаково распределенных случайных величин. При этом распределение центрированной и нормированной случайной величины $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ заменяют на стандартное нормальное распределение, т.е. $\frac{S_n - n\mu}{\sigma\sqrt{n}} \approx \mathcal{N}(0,1)$. В этом случае $S_n \approx \mathcal{N}(n\mu; n\sigma^2)$. Насколько велика ошибка при такой замене (погрешность приближения)?

Неравенство Берри – Эссеена

Следующий результат позволяет оценить погрешность приближения в центральной предельной теореме.

Теорема (неравенство Берри – Эссеена)

В условиях центральной предельной теоремы для любого $x \in \mathbb{R}$ (то есть равномерно по x)

$$\left|\mathbb{P}\left(\frac{S_n - n\mathbb{E}X_1}{\sqrt{n \operatorname{Var} X_1}} < x\right) - \Phi(x)\right| \leqslant C \cdot \frac{\mathbb{E}|X_1 - \mathbb{E}X_1|^3}{\sqrt{n}\left(\sqrt{\operatorname{Var} X_1}\right)^3}. \qquad (10)$$

Замечание

Про постоянную C известно, что

а) в общем случае C не превышает $0{,}7655$ (И.С. Шиганов),

б) погрешность приближения наиболее велика, если слагаемые X_i имеют распределение Бернулли, и C в этом случае не меньше, чем $\frac{\sqrt{10}+3}{6\sqrt{2\pi}} \approx 0{,}4097$ (C.G.Esseen, Б.А. Рогозин);

в) как показывают расчеты, можно смело брать в качестве C число $0,4$ — даже для слагаемых с распределением Бернулли, особенно при малых n, когда и это значение постоянной оказывается слишком грубой оценкой.

Подробный обзор можно найти в монографии *В.М. Золотарёва* «Современная теория суммирования независимых случайных величин», стр. 264–291.

Получим в качестве следствия из центральной предельной теоремы предельную теорему Муавра–Лапласа (P.S. Laplace, 1812, A. de Moivre, 1730). Подобно закону больших чисел в форме Бернулли, предельная теорема Муавра–Лапласа — утверждение только для ***схемы Бернулли***.

Теорема (предельная теорема Муавра–Лапласа)

Пусть производится n независимых испытаний, в каждом из которых некоторое событие A наступает с одной и той же вероятностью $p = \mathbb{P}(A)$, и пусть $k(A)$ — число наступлений события A в n испытаниях. Тогда для любого $x \in \mathbb{R}$ имеет место сходимость по распределению

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{k(A)-np}{\sqrt{np(1-p)}} < x\right) = \Phi(x). \tag{11}$$

Иначе говоря, справедлива ***интегральная формула Муавра–Лапласа***

$$\mathbb{P}(k_1 \leqslant k(A) \leqslant k_2) \approx \Phi_0(x_2) - \Phi_0(x_1), \qquad (12)$$

где $x_1 = \frac{k_1 - np}{\sqrt{np(1-p)}}$, $x_2 = \frac{k_2 - np}{\sqrt{np(1-p)}}$. Причём разница между левой и правой частями приближенного равенства в (12) согласно неравенству Берри–Эссеена не превышает величины $C \cdot \frac{p^2+(1-p)^2}{\sqrt{np(1-p)}}$.

Доказательство.

Представим случайную величину $k(A)$ — число наступлений события A в n независимых испытаниях в виде $k(A) = I_1 + I_2 + \ldots + I_n$, где $I_k = I(A)$ — индикатор события A в k испытании, т.е. случайная величина, которая принимает значение 1, если при k-ом испытании событие A происходит с вероятностью p, и значение 0 — в противном случае. Тогда $\mathbb{E}(I_k) = 1 \cdot p = p$, $\mathrm{Var}(I_k) = p(1-p)$ для любого $k = 1, \ldots, n$. Осталось воспользоваться центральной предельной теоремой. □

Пример

Монета подбрасывается $10\,000$ *раз. Используя интегральную формулу Муавра–Лапласа, найдите вероятность того, что частота выпадения герба отличается от вероятности* $\frac{1}{2}$ *более чем на одну сотую, а также оцените погрешность приближения.*

Решение. Требуется найти $\mathbb{P}\left(\left|\frac{k(A)}{n} - \frac{1}{2}\right| > 0,01\right)$, где $n = 10\,000$, $k(A) = \sum_{k=1}^{n} X_k = S_n$ — число выпадений герба, а X_k — независимые случайные величины, имеющие одно и то же распределение Бернулли с параметром $p = \frac{1}{2}$, так что $\mathbb{E}(X_k) = \frac{1}{2}$ и $\sqrt{\operatorname{Var} X_k} = \frac{1}{2}$. Согласно центральной предельной теореме или предельной теореме Муавра–Лапласа, последовательность $\frac{k(A)-np}{\sqrt{np(1-p)}} = \frac{S_n - np}{\sqrt{np(1-p)}}$ сходится по распределению к стандартному нормальному распределению. Используя интегральную формулу Муавра–Лапласа (12), получаем

$$\mathbb{P}\left(\left|\frac{k(A)}{n} - \frac{1}{2}\right| > 0,01\right) = 1 - \mathbb{P}\left(0,49n \leqslant k(A) \leqslant 0,51n\right) \boxed{\approx} 1 - 2\Phi_0(2) = 0,0456.$$

Погрешность приближения между левой и правой частями приближенного равенства $\boxed{\approx}$ при $n = 10\,000$ и $p = \frac{1}{2}$ не превышает величины

$$C \cdot \frac{p^2 + (1-p)^2}{\sqrt{np(1-p)}} \leqslant 0,4 \cdot \frac{1}{100} = 0,004.$$

Поэтому искомая вероятность $\mathbb{P}\left(\left|\frac{k(A)}{n} - \frac{1}{2}\right| > 0,01\right)$ не больше, чем $0,0456 + 0,004 = 0,0496$.