

Законы больших чисел. Неравенство Хёфдинга. Семинар 11. 13 ноября 2018 г.

Подготовил: Горбунов Э.

Источники: [НатанТВ, Гл. 8], [Ширяев, Гл. 1, §5, Гл. 2 §10, Гл. 4, §3], [Боровков, Гл. 5, §1, Гл. 8 §1, 3, 7], [Гнеденко, Гл. 6, §27-30]

Ключевые слова: ЗАКОН БОЛЬШИХ ЧИСЕЛ В ФОРМЕ ЧЕБЫШЁВА, ТЕОРЕМА МАРКОВА, ЗАКОН БОЛЬШИХ ЧИСЕЛ В ФОРМЕ ХИНЧИНА, НЕОБХОДИМОЕ И ДОСТАТОЧНОЕ УСЛОВИЕ ДЛЯ ВЫПОЛНЕНИЯ ЗАКОНА БОЛЬШИХ ЧИСЕЛ, УСИЛЕННЫЙ ЗАКОН БОЛЬШИХ ЧИСЕЛ, НЕРАВЕНСТВО ХЁФДИНГА, СУБГАУССОВСКИЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Законы больших чисел

Определение 1. Для последовательности случайных величин $\{\xi_n\}_{n=1}^\infty$, определённых на одном вероятностном пространстве и имеющих конечные математические ожидания $\mathbb{E}\xi_n < \infty$, **выполнен закон больших чисел**, если

$$\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n \mathbb{E}\xi_k \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0.$$

Теорема 1. (ЗБЧ в форме Чебышёва). Пусть $\{\xi\}_{n=1}^\infty$ — последовательность независимых случайных величин, имеющих конечные дисперсии, ограниченные одной и той же постоянной C :

$$\forall n \in \mathbb{N} \hookrightarrow \mathbb{D}\xi_n \leq C.$$

Тогда для $\{\xi\}_{n=1}^\infty$ выполнен закон больших чисел.

Доказательство. Рассмотрим произвольное $\varepsilon > 0$. Тогда из неравенства Чебышёва следует, что

$$\mathbb{P} \left\{ \left| \frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n \mathbb{E}\xi_k \right| \geq \varepsilon \right\} \leq \frac{\mathbb{D} \left[\frac{1}{n} \sum_{k=1}^n \xi_k \right]}{\varepsilon^2} = \frac{\sum_{k=1}^n \mathbb{D}\xi_k}{n^2 \varepsilon^2} \leq \frac{C}{n \varepsilon^2} \xrightarrow[n \rightarrow \infty]{} 0.$$

□

Замечание 1. Заметим, что из доказательства ЗБЧ в форме Чебышёва следует, что достаточно потребовать $\frac{1}{n^2} \mathbb{D} \left[\sum_{k=1}^n \xi_k \right] \xrightarrow[n \rightarrow \infty]{} 0$, чтобы получить тот же результат. Отсюда мы мгновенно получаем, что верна следующая теорема.

Теорема 2. (Теорема Маркова). Если последовательность случайных величин $\{\xi\}_{n=1}^\infty$ такова, что существуют $\mathbb{D}\xi_n < \infty, n \geq 1$ и

$$\frac{1}{n^2} \mathbb{D} \left[\sum_{k=1}^n \xi_k \right] \xrightarrow[n \rightarrow \infty]{} 0,$$

то $\{\xi_n\}_{n=1}^\infty$ удовлетворяет закону больших чисел.

Отметим, что в ЗБЧ в форме Чебышёва и в теореме Маркова мы не требуем независимости от последовательности случайных величин, однако требуем существования дисперсий. Другой подход, основанный на методе производящих функций, позволяет отказаться от существования дисперсий, однако требует независимости и одинаковой распределённости случайных величин, образующих последовательность.

Теорема 3. (ЗБЧ в форме Хинчина). Пусть $\{\xi_n\}_{n=1}^\infty$ — независимые одинаково распределённые случайные величины с конечными математическими ожиданиями $\mathbb{E}\xi_n = m$. Тогда $\{\xi_n\}_{n=1}^\infty$ удовлетворяет закону больших чисел, т. е.

$$\overline{\xi_n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} m,$$

где $\overline{\xi_n} \stackrel{\text{def}}{=} \frac{1}{n} \sum_{k=1}^n \xi_k$.

Доказательство. Рассмотрим характеристическую функцию $\overline{\xi_n}$:

$$\varphi_{\overline{\xi_n}}(t) = \left(\varphi_{\xi_1} \left(\frac{t}{n} \right) \right)^n = \left(1 + \frac{it\mathbb{E}\xi_1}{n} + o\left(\frac{t}{n}\right) \right)^n \xrightarrow[n \rightarrow \infty]{} e^{it\mathbb{E}\xi_1} = \varphi_{\mathbb{E}\xi_1}(t).$$

Значит, по теореме Леви о непрерывности

$$\overline{\xi_n} \xrightarrow[n \rightarrow \infty]{d} \mathbb{E}\xi_1.$$

Так как $\mathbb{E}\xi_1 = m$ — константа, то

$$\overline{\xi_n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} m.$$

□

Пример 1. Существования математического ожидания является существенным условием (даже если рассмотреть более общее понятие о законе больших чисел; в исходном нашем определении требуется существование математических ожиданий, поэтому ЗБЧ не может выполняться в определённом ранее смысле для последовательности случайных величин, не имеющих математических ожиданий). Рассмотрим последовательность таких случайных величин $\{\xi_n\}_{n=1}^\infty$, что $\{\xi_n\}_{n=1}^\infty$ независимы и $\xi_n \sim \text{Ca}(0, 1)$. Характеристическая функция ξ_n вычисляется по формуле:

$$\varphi_{\xi_n}(t) = \mathbb{E}[e^{it\xi_n}] = \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{itx} \cdot \frac{1}{x^2 + 1} dx = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\cos(tx)}{x^2 + 1} dx = e^{-|t|},$$

где последнее равенство — *интеграл Лапласа* (он вычисляется при помощи свойств прямого и обратного преобразований Фурье; подробности см., например, в [Иванов Г. Е., Лекции по математическому анализу, часть 2, глава 17, §6, замечание после примера 1](#)). Тогда

$$\varphi_{\overline{\xi_n}}(t) = \left(\varphi_{\xi_n} \left(\frac{t}{n} \right) \right)^n = \left(e^{-|t|/n} \right)^n = e^{-|t|} = \varphi_{\xi_1}(t),$$

а значит, $\overline{\xi_n} \sim \text{Ca}(0, 1)$ для любого n . Отсюда следует, что не существует такой последовательности чисел a_1, a_2, \dots , что $\overline{\xi_n} - a_n$ не сходится по вероятности к нулю (это и есть обобщённое понятие о законе больших чисел; если взять $a_n = \frac{1}{n} \sum_{k=1}^n \mathbb{E}\xi_k$, то получим исходное определение).

Упражнение 1. Рассмотрим последовательность попарно некоррелированных случайных величин $\{\xi_n\}$ таких, что

$$\xi_n = \begin{cases} \sqrt{n}, & \text{с вероятностью } \frac{1}{2n}, \\ 0, & \text{с вероятностью } 1 - \frac{1}{n}, \\ -\sqrt{n}, & \text{с вероятностью } \frac{1}{2n}. \end{cases}$$

Выполнен ли для неё закон больших чисел?

Решение. Покажем, что для указанной последовательности выполнен закон больших чисел. Во-первых,

$$\mathbb{E}\xi_n = \sqrt{n} \cdot \frac{1}{2n} - \sqrt{n} \cdot \frac{1}{2n} = 0,$$

и

$$\mathbb{D}\xi_n = \mathbb{E}[\xi_n^2] - (\mathbb{E}\xi_n)^2 = 2 \cdot n \cdot \frac{1}{2n} = 1.$$

Тогда по теореме Чебышёва получаем, что для $\{\xi_n\}_{n=1}^\infty$ выполнен закон больших чисел.

□

Теперь сформулируем необходимо и достаточное условие выполнения закона больших чисел.

Теорема 4. Пусть задана последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ с конечными математическими ожиданиями. Тогда для последовательности $\{\xi_n\}_{n=1}^{\infty}$ выполнен закон больших чисел тогда и только тогда, когда

$$\mathbb{E} \left[\frac{(\xi_n - \mathbb{E}[\xi_n])^2}{1 + (\xi_n - \mathbb{E}[\xi_n])^2} \right] \xrightarrow{n \rightarrow \infty} 0.$$

Доказательство. Достаточно короткое и простое доказательство можно прочитать в [Гнеденко, Гл. 6, §29]. \square

Упражнение 2. Пусть последовательность $\{\xi_n\}_{n=1}^{\infty}$ независимых случайных величин такова, что

$$\mathbb{P}\{\xi_n = n^\alpha\} = \mathbb{P}\{\xi_n = -n^\alpha\} = \frac{1}{2}.$$

При каких α для заданной последовательности выполняется закон больших чисел?

Решение. Данное упражнение можно попытаться решить при помощи критерия, но мы воспользуемся сначала теоремой Маркова, а затем методом характеристических функций. Заметим, что

$$\frac{1}{n^2} \mathbb{D} \left[\sum_{k=1}^n \xi_k \right] = \frac{1}{n^2} \sum_{k=1}^n \mathbb{D} \xi_k = \frac{1}{n^2} \sum_{k=1}^n (\mathbb{E}[\xi_k^2] - (\mathbb{E} \xi_k)^2) = \frac{1}{n^2} \sum_{k=1}^n k^{2\alpha} \leq \frac{1}{n^2} \cdot n \cdot n^{2\alpha} = n^{2\alpha-1} \xrightarrow[n \rightarrow \infty]{\text{при } \alpha < \frac{1}{2}} 0,$$

а значит, по теореме Маркова для последовательности $\{\xi_n\}_{n=1}^{\infty}$ выполнен ЗБЧ при $\alpha < \frac{1}{2}$. Покажем теперь при помощи метода характеристических функций, что при $\alpha \geq \frac{1}{2}$ ЗБЧ не выполнен. Заметим, что

$$\overline{\xi_n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0 \iff \overline{\xi_n} \xrightarrow[n \rightarrow \infty]{d} 0 \iff \varphi_{\overline{\xi_n}}(t) \xrightarrow[n \rightarrow \infty]{} \varphi_0(t) = 1, \forall t \in \mathbb{R}.$$

Кроме того,

$$\varphi_{\xi_n} = \frac{1}{2} e^{itn^\alpha} + \frac{1}{2} e^{-itn^\alpha} = \cos(tn^\alpha),$$

откуда в силу независимости $\{\xi_n\}_{n=1}^{\infty}$

$$\varphi_{\overline{\xi_n}}(t) = \prod_{k=1}^n \cos\left(\frac{tk^\alpha}{n}\right).$$

Пусть $\alpha \geq \frac{1}{2}$. Выберем $t \in (0, \frac{\pi}{2})$. Тогда

$$\begin{aligned} |\varphi_{\overline{\xi_n}}(t)| &= \prod_{k=1}^n |\cos(t \frac{k^\alpha}{n})| \leq \prod_{\frac{1}{2}(\frac{n}{2})^{1/(2\alpha)} \leq k < (\frac{n}{2})^{1/(2\alpha)}} |\cos(t \frac{k^\alpha}{n})| \\ &\lesssim \left(\cos\left(t \cdot \frac{1}{2^\alpha} \cdot \frac{1}{\sqrt{n}} \cdot \frac{1}{\sqrt{2}}\right) \right)^{\frac{1}{2} \cdot (\frac{n}{2})^{1/(2\alpha)}} \\ &= \left(\left(1 - \frac{t^2}{2^{2\alpha+1}n} + o\left(\frac{1}{n}\right) \right)^{-\frac{n \cdot 2^{2\alpha+1}}{t^2}} \right)^{-\frac{t^2}{2^{4\alpha+2}} \cdot n^{\frac{1}{2\alpha}-1}} \xrightarrow[n \rightarrow \infty]{} \begin{cases} 0, & \text{при } \alpha > \frac{1}{2}, \\ e^{-\frac{t^2}{2^{4\alpha+2}}}, & \text{при } \alpha = \frac{1}{2} \end{cases} < 1, \end{aligned}$$

где второе неравенство следует из того, что $f(t) = \cos t$ убывает на $(0, \frac{\pi}{2})$. Значит, $\varphi_{\overline{\xi_n}}(t) \not\rightarrow 1$ при $n \rightarrow \infty$. \square

Теперь определим более сильное свойство последовательности случайных величин.

Определение 2. Для последовательности случайных величин $\{\xi_n\}_{n=1}^{\infty}$, определённых на одном вероятностном пространстве и имеющих конечные математические ожидания $\mathbb{E} \xi_n < \infty$, **усиленный закон больших чисел**, если

$$\frac{1}{n} \sum_{k=1}^n \xi_k - \frac{1}{n} \sum_{k=1}^n \mathbb{E} \xi_k \xrightarrow[n \rightarrow \infty]{\text{п.н.}} 0.$$

Оказывается, что для последовательности независимых одинаково распределённых случайных величин существование математического ожидания является необходимым и достаточным условием для выполнения усиленного закона больших чисел, что было доказано А. Н. Колмогоровым.

Теорема 5. (Теорема Колмогорова об УЗБЧ). Существование математического ожидания является необходимым и достаточным условием для применимости усиленного закона больших чисел к последовательности независимых одинаково распределённых случайных величин.

Доказательство. Доказательство можно прочитать в [Гнеденко, Гл. 6, §30]. \square

Пример 2. Пусть проводится некоторый эксперимент много раз, и в результате каждого эксперимента независимо может произойти событие A , то есть мы имеем дело с последовательностью независимых событий $\{A_n\}_{n=1}^{\infty}$, $\mathbb{P}\{A_n\} = p$. Из усиленного закона больших чисел следует, что $\frac{1}{n} \sum_{k=1}^n \mathbb{I}_{A_k} \xrightarrow[n \rightarrow \infty]{\text{п.н.}} p$, т. е. выполняется условие, которое мы рассматривали как статистическое определение вероятности.

Скорость сходимости в законе больших чисел

До сих пор нас не интересовала скорость сходимости в законе больших чисел, хотя для практических целей её очень полезно знать. Все рассмотренные нами результаты либо гарантировали полиномиальную скорость сходимости (ЗБЧ в форме Чебышёва гарантирует сходимость со скоростью $O(n^{-1})$), либо вообще не устанавливали скорость сходимости. Далее мы зададимся вопросом, при каких условиях можно гарантировать экспоненциальную скорость сходимости.

Определение 3. Случайная величина ξ с математическим ожиданием $\mathbb{E}\xi = m < \infty$, удовлетворяющая условию

$$\mathbb{E}[e^{\lambda\xi}] \leq e^{\lambda m + \frac{\lambda^2 \sigma^2}{2}}, \quad \forall \lambda \in \mathbb{R}$$

для некоторого $\sigma^2 > 0$, называется **субгауссовской с параметром σ^2** .

Замечание 2. В частности, гауссовская случайная величина является субгауссовской. Кроме того, как мы покажем далее, любая ограниченная с вероятностью 1 случайная величина является субгауссовской. Более подробно про субгауссовские случайные величины можно почитать в книге Романа Вершинина, [High-Dimensional Probability: An Introduction with Applications in Data Science](#). Кроме того, тем, кто серьёзно планирует заниматься анализом данных и погружаться в математику вокруг него, эта книга *может быть полезна*.

Лемма 1. (Лемма Хёфдинга). Пусть случайная величина ξ такова, что $\mathbb{P}\{\xi \in [a, b]\} = 1$ и $\mathbb{E}\xi = 0$. Тогда для любого $\lambda \in \mathbb{R}$ выполнено неравенство:

$$\mathbb{E}[e^{\lambda\xi}] \leq e^{\frac{\lambda^2(b-a)^2}{8}}$$

Доказательство. Так как $e^{\lambda x}$ является выпуклой функцией от x , то

$$e^{\lambda x} \leq \frac{b-x}{b-a} e^{\lambda a} + \frac{x-a}{b-a} e^{\lambda b}, \quad \forall x \in [a, b].$$

Отсюда следует, что

$$\mathbb{E}[e^{\lambda\xi}] \leq \frac{b - \mathbb{E}\xi}{b-a} e^{\lambda a} + \frac{\mathbb{E}\xi - a}{b-a} e^{\lambda b} = \frac{b}{b-a} e^{\lambda a} - \frac{a}{b-a} e^{\lambda b}.$$

Пусть $h = \lambda(b-a)$, $p = -\frac{a}{b-a}$ и $L(h) = -hp + \ln(1-p+pe^h)$. Тогда (проверьте это, раскрыв скобки)

$$\frac{b}{b-a} e^{\lambda a} - \frac{a}{b-a} e^{\lambda b} = e^{L(h)}.$$

Посчитаем первую и вторую производные:

$$L'(h) = -p + \frac{pe^h}{1-p+pe^h}, \quad L'(0) = 0,$$

$$L''(h) = \frac{pe^h(1-p)}{(1-p+pe^h)^2} = \frac{p(1-p)}{\left((1-p)e^{-\frac{h}{2}} + pe^{\frac{h}{2}}\right)^2}.$$

Покажем, что $L''(h) \leq \frac{1}{4}$ для всех h . Рассмотрим отдельно знаменатель: $g(h) = \left((1-p)e^{-\frac{h}{2}} + pe^{\frac{h}{2}}\right)^2 = (1-p)^2 e^{-h} + 2p(1-p) + p^2 e^h$. Функция $g(h)$ — выпуклая, т. к. равна сумме выпуклых функций. Найдём стационарную точку функции g :

$$g'(h) = -(1-p)^2 e^{-h} + p^2 e^h = 0 \iff e^h = \frac{1-p}{p}.$$

Значит, $g(h) \geq g\left(\ln\left(\frac{1-p}{p}\right)\right) = 4p(1-p)$, откуда

$$L''(h) = \frac{p(1-p)}{g(h)} \leq \frac{p(1-p)}{4p(1-p)} \leq \frac{1}{4}.$$

Тогда, используя формулу Тейлора с остаточным членом в форме Лагранжа, получим

$$L(h) = \frac{h^2}{2} L''(h_1) \leq \frac{h^2}{8} = \frac{1}{8} \lambda^2 (b-a)^2.$$

Следовательно,

$$\mathbb{E}[e^{\lambda\xi}] \leq e^{\frac{\lambda^2(b-a)^2}{8}}.$$

□

Следствие 1. Любая ограниченная с вероятностью 1 случайная величина ξ , т. е. такая, что $\mathbb{P}\{\xi \in [a, b]\} = 1$, является субгауссовской с $\sigma^2 = \frac{(b-a)^2}{4}$. Действительно, достаточно применить лемму Хёфдинга для случайной величины $\eta = \xi - \mathbb{E}\xi$. В частности, бернуллиевская случайная величина является субгауссовской с $\sigma^2 = \frac{1}{4}$.

Пример 3. Показать, что случайная величина $\xi \sim \text{Poisson}(\mu)$ не является субгауссовской.

Доказательство. Действительно,

$$\mathbb{E}[e^{\lambda\xi}] = \sum_{n=0}^{\infty} \frac{e^{\lambda n} \mu^n}{n!} \cdot e^{-\mu} = \exp(\mu(e^\lambda - 1)).$$

Для любых σ^2 неравенство $\exp(\mu(e^\lambda - 1)) \leq e^{\lambda\mu + \frac{\lambda^2\sigma^2}{2}}$ не будет выполнено для достаточно больших λ , т. к. $\lim_{\lambda \rightarrow +\infty} \frac{\mu(e^\lambda - 1)}{\lambda\mu + \frac{\lambda^2\sigma^2}{2}} = +\infty$. □

Для субгауссовских случайных величин скорость сходимости в законе больших чисел экспоненциальна, что устанавливает следующая теорема.

Теорема 6. (Неравенство Хёфдинга). Пусть дана последовательность независимых одинаково распределённых субгауссовских случайных величин $\{\xi_n\}_{n=1}^{\infty}$ с параметром σ^2 и математическим ожиданием $\mathbb{E}\xi_n = m$. Тогда для любого $t > 0$:

$$\mathbb{P}\left\{\frac{1}{n} \sum_{k=1}^n (\xi_k - m) > t\right\} \leq 2e^{-\frac{nt^2}{2\sigma^2}}.$$

Доказательство. Применим неравенство Чернова для случайной величины $\sum_{k=1}^n (\xi_k - m)$, а затем воспользуемся леммой Хёфдинга:

$$\mathbb{P}\left\{\frac{1}{n} \sum_{k=1}^n (\xi_k - m) > t\right\} \leq \inf_{\lambda > 0} \frac{\mathbb{E}\left[e^{\lambda \sum_{k=1}^n (\xi_k - m)}\right]}{e^{\lambda nt}} = \inf_{\lambda > 0} \frac{\prod_{k=1}^n \mathbb{E}[e^{\lambda(\xi_k - m)}]}{e^{\lambda nt}} \leq \inf_{\lambda > 0} e^{\frac{n\lambda^2\sigma^2}{2} - \lambda nt} = e^{-\frac{nt^2}{2\sigma^2}},$$

где последний переход следует из того, что минимальное значение параболы $\frac{n\lambda^2\sigma^2}{2} - \lambda nt$ (относительно λ) равно $-\frac{nt^2}{2\sigma^2}$. Заметим, что в силу произвольности выбора λ в определении субгауссовской случайной величины мы имеем следующее неравенство

$$\mathbb{E} [e^{-\lambda\xi_n}] \leq e^{-\lambda m + \frac{\lambda^2\sigma^2}{2}},$$

поэтому мы можем аналогичными действиями получить, что

$$\mathbb{P} \left\{ \frac{1}{n} \sum_{k=1}^n (\xi_k - m) < -t \right\} = \mathbb{P} \left\{ \frac{1}{n} \sum_{k=1}^n (m - \xi_k) > t \right\} \leq \dots \leq e^{-\frac{nt^2}{2\sigma^2}},$$

откуда следует, что

$$\mathbb{P} \left\{ \frac{1}{n} \sum_{k=1}^n (\xi_k - m) > t \right\} \leq 2e^{-\frac{nt^2}{2\sigma^2}}.$$

□

Следствие 2. Пусть $\{\xi_n\}_{n=1}^\infty$ — последовательность независимых одинаково распределённых случайных величин с распределением $\text{Be}(p)$. Тогда для любого $t > 0$

$$\mathbb{P} \left\{ |\bar{\xi}_n - p| > t \right\} \leq 2e^{-2nt^2}.$$