Законы больших чисел. Неравенство Хёффдинга. 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 Хёффдинга, 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

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

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:

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

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

Замечание 1. Заметим, что из доказательства ЗБЧ в форме Чебышёва следует, что достаточно потребовать $\frac{1}{n^2}\mathbb{D}\left[\sum\limits_{k=1}^{n}\xi_k\right] \xrightarrow[n\to\infty]{} 0$, 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 $\mathbb{D}\xi_n < \infty, n \geqslant 1$ и

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

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.

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

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

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 выполнятся 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 $\xi_n \sim \mathrm{Ca}(0,1)$. Характеристическая функция ξ_n вычисляется по формуле:

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

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

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

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

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

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

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?

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 выполнен закон больших чисел**, если

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

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.

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 $\frac{1}{n}\sum\limits_{k=1}^{n} \mathbb{I}_{A_k} \xrightarrow[n\to\infty]{\text{п.н.}} p$, 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 $O\left(n^{-1}\right)$), 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

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

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

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

Следствие 1. Любая ограниченная с вероятностью 1 случайная величина ξ, т. е. такая, что $\mathbb{P}\left\{\xi \in [a,b]\right\} = 1$, является субгауссовской с $\sigma^2 = \frac{(b-a)^2}{4}$. Действительно, достаточно применить лемму Хёффдинга 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 $\xi \sim \mathrm{Poisson}(\mu)$ не является субгауссовской.

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

Теорема 6. (Неравенство Хёффдинга). 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 $t > 0$:

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

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

$$\mathbb{P}\left\{|\overline{\xi_n} - p| > t\right\} \leqslant 2e^{-2nt^2}.$$