Сходимость последовательностей случайных величин. Закон больших чисел

октябрь–декабрь 2023 г.

Курс лекций «Теория вероятностей»

Оглавление

Сходимость последовательностей случайных величин

Мы будем изучать две предельные теоремы: закон больших чисел и центральную предельную теорему. Поскольку в них речь идет о сходимости последовательности случайных величин, следует сначала прояснить, что это значит.

Пусть на вероятностном пространстве задана последовательность $\{\xi_n, n = 1, 2, \dots\}$ случайных величин и ξ — случайная величина.

$\xi_1, \xi_2, \xi_3, \dots \dots$

Сходимость почти наверное

Определение (сходимость почти наверное)
Последовательность $\{\xi_n, n = 1, 2, \dots\}$ случайных величин *сходится почти наверное* к случайной величине ξ, если

$$\lim_{n\to\infty} \xi_n(\omega) = \xi(\omega) \quad \text{почти для всех } \omega \in \Omega,$$

т. е. $P\{\omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\} = 1$, обозначение $\lim_{n\to\infty} \xi_n = \xi$ (п. н., a. s.).

almost surely

Сходимость по вероятности

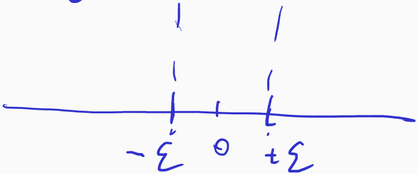
Определение (сходимость по вероятности)
Последовательность $\{\xi_n, n = 1, 2, \dots\}$ случайных величин *сходится по вероятности* к случайной величине ξ, если

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} P(|\xi_n - \xi| > \varepsilon) = 0,$$

обозначение $p\lim_{n\to\infty} \xi_n = \xi$.

$\xi_n \xrightarrow{P} \xi \Leftrightarrow \eta_n \xrightarrow{P} 0$

$\eta_n = \xi_n - \xi$



Связь сходимостей почти всюду и по вероятности

Можно показать, что

- из сходимости почти наверное вытекает сходимость по вероятности;
- если $p\lim_{n\to\infty} \xi_n = \xi$, то существует подпоследовательность $\{\xi_{n_k}, k = 1, 2, \dots\}$, такая что $\lim_{k\to\infty} \xi_{n_k} = \xi$ п. н.

$$S_n = \frac{X_1 + \ldots + X_n}{n}$$

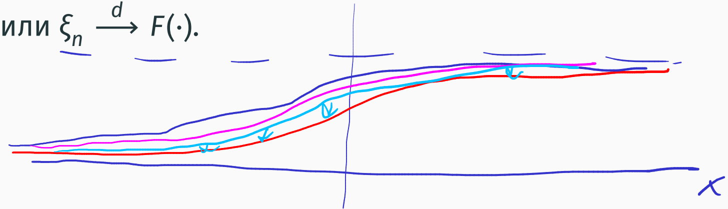
Сходимость по распределению

Определение (сходимость по распределению)
Последовательность $\{\xi_n, n = 1, 2, \dots\}$ случайных величин *сходится по распределению* к случайной величине ξ, если

$\forall A \quad P(\xi_n \in A) \to P(\xi \in A)$

$$\lim_{n\to\infty} F_n(x) = F(x)$$

в каждой точке непрерывности функции $F(x)$, где $F_n(x)$, $F(x)$ — функции распределения случайных величин ξ_n, ξ соответственно. Обозначение $\xi_n \xrightarrow{d} \xi$ или $\xi_n \xrightarrow{d} F(\cdot)$.



Сходимость по распределению: пример

Почему есть требование о сходимости только в точках непрерывности функции распределения предельной случайной величины?

Пусть $Z \sim N(0, 1)$ — стандартная нормальная случайная величина, и пусть $\xi_n = Z/n$. Тогда не вызывает сомнения тот факт, что последовательность $\{\xi_n, n = 1, 2, \dots\}$ сходится к нулю в любом разумном смысле.

Пусть $x < 0$. Тогда $F_n(x) = P(Z/n \leqslant x) = \Phi(nx) \to 0$ при $n \to \infty$.

Если $x > 0$, то $F_n(x) = \Phi(nx) \to 1$ при $n \to \infty$.

В то же время в силу симметрии распределения каждой случайной величины ξ_n выполнено равенство $F_n(0) = 1/2$.

Сходимость по распределению: продолжение примера

Значит,

$$\lim_{n\to\infty} F_n(x) = G(x) = \begin{cases} 0, & x < 0, \\ \cancel{1/2,} & \cancel{x = 0,} \\ 1, & x > 0. \end{cases}$$

точка разрыва

Функция $G(\cdot)$ даже не является функцией распределения, т. к. в точке 0 она не является непрерывной справа. А все дело в том, что в нуле функция распределения предельной случайной величины (т. е. нуля) разрывна.

Связь сходимости по распределению с другими видами сходимости

Можно показать, что

- из сходимости по вероятности вытекает сходимость по распределению;
- из сходимости почти наверное вытекает сходимость по распределению.

Сходимость по распределению к константе

Предложение
Пусть $\{\xi_n, n = 1, 2, \dots\}$ — *последовательность случайных величин, по распределению сходящихся к постоянной случайной величине* c. *Тогда* $p\lim_{n\to\infty} \xi_n = c$, *т. е. последовательность сходится и по вероятности.*

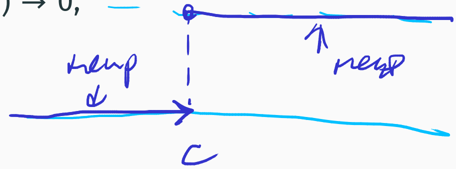
Доказательство.
Так как $\xi_n \xrightarrow{d} c$, то при любом $x > c$ последовательность $F_n(x) \to F_c(x) = 1$, а при любом $x < c$ последовательность $F_n(x) \to F_c(x) = 0$. Пусть $\varepsilon > 0$. Тогда

$$\boxed{P(|\xi_n - c| \geqslant \varepsilon)} = 1 - P(|\xi_n - c| < \varepsilon) = 1 - P(c - \varepsilon < \xi_n < c + \varepsilon) = 1 - \left(F_n(c + \varepsilon-) - F(c - \varepsilon)\right)$$

непр. слева

$$\leqslant 1 - F_n(c + \varepsilon) + F(c - \varepsilon) \to 0,$$

так как $F_n(c + \varepsilon) \to 1$ и $F(c - \varepsilon) \to 0$. ∎



Теорема (Манн – Вальд)
Пусть $\{\xi_n, n = 1, 2, \dots\}$ *— последовательность случайных величин, и пусть* $g(\cdot)$ *— непрерывная в* $\mathbb{R}$ *функция. Тогда*

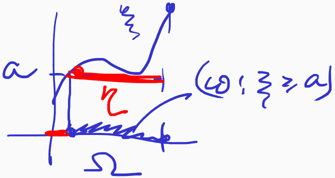
1. *если* $\xi_n \xrightarrow{\text{a.s.}} \xi$, *то* $g(\xi_n) \xrightarrow{\text{a.s.}} g(\xi)$;
2. *если* $\xi_n \xrightarrow{p} \xi$, *то* $g(\xi_n) \xrightarrow{p} g(\xi)$;
3. *если* $\xi_n \xrightarrow{d} \xi$, *то* $g(\xi_n) \xrightarrow{d} g(\xi)$.

Спойлер: придётся применять в ДЗ 7.

Неравенство Маркова

Предложение (неравенство Маркова)
Пусть $\xi \geqslant 0$ — неотрицательная случайная величина и $a > 0$. Тогда

$$P(\xi > a) \leq \quad P(\xi \geqslant a) \leqslant \frac{E(\xi)}{a}.$$



Доказательство.
Рассмотрим случайную величину η, равную 0, если $\xi < a$, и равную a, если $\xi \geqslant a$. Так как $\eta \leqslant \xi$, то $E(\eta) \leqslant E(\xi)$. И так как $E(\eta) = aP(\xi \geqslant a)$, то $aP(\xi \geqslant a) \leqslant E(\xi)$, откуда и следует искомое неравенство. □

Если ξ имеет плотность, то можно так:

$$E(\xi) = \int_0^{+\infty} xf(x)dx \geqslant \int_a^{+\infty} xf(x)dx \geqslant a\int_a^{+\infty} f(x)dx = aP(\xi \geqslant a).$$

Неравенство Маркова: пример

Пример

Пусть число деталей, которые фабрика производит за неделю, является случайной величиной с матожиданием 50. Оцените вероятность того, что выпуск за неделю превысит 75.

По неравенству Маркова

$$\cancel{P(X > 75) \leqslant \frac{E(X)}{75} = \frac{50}{75} = \frac{2}{3}.}$$

$$P(X > 75) = P(X \geq 76) \leq \frac{E(X)}{76} = \frac{50}{76} = \frac{25}{38} \approx 0.6578 < \frac{2}{3} = 0.6667$$

Следствие (неравенство Чебышева)
Пусть X — случайная величина и $\varepsilon > 0$. *Тогда*

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2}.$$

Доказательство.
Обозначим $\xi = (X - E(X))^2 \leqslant 0$. Тогда из неравенства Маркова

$$P(|X - E(X)| \geq \varepsilon) = P(\xi \geq \varepsilon^2) \leq \frac{E(\xi)}{\varepsilon^2} = \frac{\mathrm{Var}(X)}{\varepsilon^2}.$$

□

Часто неравенство Чебышева используют в следующей форме:

$$P(|X - m_X| \geq k\sigma_X) \leq \frac{\cancel{\sigma_X^2}}{k^2\cancel{\sigma_X^2}} = \frac{1}{k^2}.$$

$\sigma = 5$

Пример
Пусть число деталей, которые фабрика производит за неделю, является случайной величиной с матожиданием 50 и дисперсией 25. Оцените вероятность того, что выпуск за неделю будет находиться между 40 и 60.

По неравенству Чебышева

$$P(40 < X < 60) = 1 - \underbrace{P(|X - 50| \geqslant 10)} \geqslant 1 - \frac{\overset{25}{\mathrm{Var}(X)}}{10^2} = 1 - \frac{1}{4} = \frac{3}{4}.$$

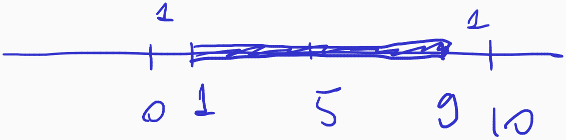
$40 < 50 < 60$ $\qquad \pm 2\sigma$

$10 \qquad 10 = 2\sigma$

Пусть X распределена равномерно на отрезке $[0, 10]$. Тогда $E(X) = 5$ и $\mathrm{Var}(X) = 10^2/12 = 25/3$. Из неравенства Чебышева следует, что

$$\boxed{P(|X - 5| \geqslant 4)} \leqslant \frac{25}{3 \cdot 4^2} \approx 0.52,$$

в то время, как истинное значение этой вероятности равно 0.2.



Точность неравенства Чебышева: нормальное распределение

Пусть X подчиняется нормальному распределению $N(\mu, \sigma^2)$. Из неравенства Чебышева следует, что $P(|X-\mu| \geqslant 3\sigma) \leq \frac{1}{9} \approx 0.11$

$$P(|X-\mu| \geqslant 2\sigma) \leqslant \frac{\sigma^2}{(2\sigma)^2} = \frac{1}{4} = 0.25,$$

настоящая < 0.005

в то время, как истинное значение этой вероятности равно

$$P(|X-\mu| \geqslant 2\sigma) = P\left(\left|\frac{X-\mu}{\sigma}\right| \geqslant 2\right) = 2(1-\Phi(2)) \approx \underline{\underline{0.0456}}.$$

Неравенства Маркова и Чебышева обычно дают очень неточные оценки. Но ослабить их нельзя (существуют распределения, для которых эти неравенства превращаются в равенства).

Теорема (закон больших чисел)
Пусть $\{\xi_n, n = 1, 2, \dots\}$ — *последовательность независимых одинаково распределенных случайных величин,*

$$E(\xi_n) = m, \quad \cancel{\operatorname{Var}(\xi_n) = \sigma^2}.$$

Тогда

$$p \lim_{n\to\infty} \frac{\xi_1 + \dots + \xi_n}{n} = m.$$

Доказательство.
Обозначим $X_n = \frac{\xi_1+\dots+\xi_n}{n} - m$. Тогда $E(X_n) = 0$, $\operatorname{Var}(X_n) = \sigma^2/n$. Из неравенства Чебышева следует, что

$$P(|X_n| \geqslant \varepsilon) \leqslant \frac{\sigma^2}{n\varepsilon^2} \to 0, \quad n \to \infty,$$

что и означает сходимость X_n к нулю и сходимость $\frac{\xi_1+\dots+\xi_n}{n}$ к m по вероятности. □

Усиленный закон больших чисел

Оказывается, используя более продвинутые методы, можно доказать, что в законе больших чисел выполняется сходимость почти наверное:

Теорема (усиленный закон больших чисел)
Пусть $\{\xi_n, n = 1, 2, \dots\}$ — *последовательность независимых одинаково распределенных случайных величин,*

$$E(\xi_n) = m, \quad \cancel{\mathrm{Var}(\xi_n) = \sigma^2}.$$

Тогда

$$\lim_{n\to\infty} \frac{\xi_1 + \dots + \xi_n}{n} = m, \quad п.\,н.$$

При этом можно даже отказаться от условия $E(\xi_i^2) < +\infty$.

Закон больших чисел: метод Монте-Карло

Пусть $\{\xi_n, n = 1, 2, \dots\}$ — последовательность независимых случайных величин, равномерно распределенных на $[0, 1]$, и пусть $g(x)$ — непрерывная на $[0, 1]$ функция. Тогда последовательность $\{g(\xi_n), n = 1, 2, \dots\}$ является последовательностью независимых одинаково распределенных величин с конечным вторым моментом, и по закону больших чисел

$f(x) \equiv 1$

$$\lim_{n\to\infty} \frac{g(\xi_1) + \dots + g(\xi_n)}{n} = E(g(\xi_i)) = \int_0^1 g(x)dx.$$

Этот метод вычисления интеграла называется методом Монте-Карло.

Неравенство Иенсена

Предложение (неравенство Иенсена)
Пусть $g(x)$ — выпуклая функция, ξ — случайная величина. Тогда $E(g(\xi)) \geqslant g(E(\xi))$, если оба математических ожидания существуют.

Доказательство.
Докажем для гладкой функции $g(x)$. Обозначим $\mu = E(\xi)$. Для выпуклой функции $g(x)$ справедливо неравенство

$$g(x) \geqslant g(\mu) + g'(\mu)(x - \mu)$$

(график выпуклой функции лежит выше ее касательной). Поэтому

$$g(\xi) \geqslant g(\mu) + g'(\mu)(\xi - \mu),$$

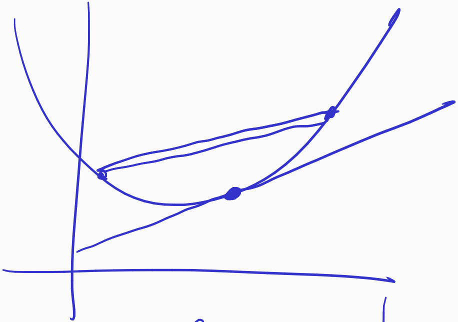
откуда

$$E(g(\xi)) \geqslant E(g(\mu) + g'(\mu)(\xi - \mu)) = E(g(\mu)) + g'(\mu)E(\xi - \mu) = g(\mu) = g(E(\xi)).$$

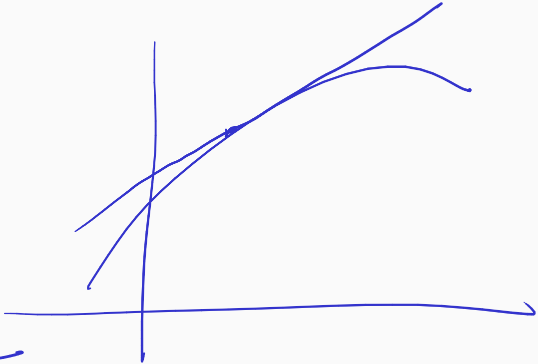
□

Для заметок

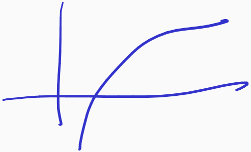
Выпуклое



Вогнутое



$\ln x$



Неравенство Иенсена: продолжение

Предложение (неравенство Иенсена для вогнутой функции)
Пусть $g(x)$ — вогнутая функция, ξ — случайная величина. Тогда $E(g(\xi)) \leqslant g(E(\xi))$, если оба математических ожидания существуют.

Предложение (неравенство Иенсена для строго выпуклой/вогнутой функции)
Пусть $g(x)$ — строго выпуклая (строго вогнутая) функция, ξ — невырожденная случайная величина. Тогда $E(g(\xi)) > g(E(\xi))$ $(E(g(\xi)) < g(E(\xi)))$, если оба математических ожидания существуют.

Неравенство Иенсена: неожиданное приложение

Доказать неравенство между средним арифметическим и средним геометрическим

$$\frac{x_1 + \dots + x_n}{n} \geqslant \sqrt[n]{x_1 \cdot \ldots \cdot x_n}. = (x_1 \cdot \ldots \cdot x_n)^{\frac{1}{n}}$$

Прологарифмируем левую и правую части:

$$\ln \frac{x_1 + \dots + x_n}{n}, \geq \frac{\ln x_1 + \dots + \ln x_n}{n}.$$

Рассмотрим дискретную случайную величину ξ, принимающую значения $x_1, \ldots, x_n$ с вероятностями $1/n$ каждое. Тогда, поскольку $\ln x$ вогнутая, то согласно неравенству Иенсена

$$\ln E(\xi) = \ln \frac{x_1 + \dots + x_n}{n} \geqslant E(\ln \xi) = \frac{\ln x_1 + \dots + \ln x_n}{n}.$$

(Из строгой вогнутости $\ln x$ можно сделать вывод о том, что если $x_1, \ldots, x_n$ не все одинаковые, то неравенство строгое.)

Литература

[1] Sh. Ross. *A First Course in Probability*. 9th ed. London: Pearson, 2014.

[2] Б. В. Гнеденко. *Курс теории вероятностей*. М.: Издательство УРСС, 2005.

[3] С. В. Головань, П. К. Катышев, А. А. Пересецкий. *Сборник задач по курсу теории вероятностей*. М.: Издательский дом «Дело» РАНХиГС, 2021.

Для заметок