Сходимости

По вероятности

Определение 2.1. $X_n \to X$ *по вероятности*, если для любого $\varepsilon > 0$ $\lim\limits_{n\to\infty} \mathbf{P}(|X_n - X| > \varepsilon) = 0$.

Средне квадратическая

Определение 2.2. $X_n \to X$ *в среднем квадратическом*, если $\lim\limits_{n\to\infty} \mathrm{M}\{|X_n - X|^2\} = 0$.

Почти наверное

Определение 2.3. $X_n \to X$ *почти наверное (с вероятностью 1)*, если $\mathbf{P}\left(\omega \in \Omega\colon \lim\limits_{n\to\infty} X_n(\omega) = X(\omega)\right) = 1$.

Свойства сходящихся последовательностей

Перечислим некоторые свойства сходящихся последовательностей.

1) Если $X_n \xrightarrow{\text{п.н.}} X$ или $X_n \xrightarrow{\text{с.к.}} X$, $n \to \infty$, то $X_n \xrightarrow{\text{P}} X$, $n \to \infty$.

2) Если $X_n \xrightarrow{\text{п.н.}} X$, $Y_n \xrightarrow{\text{п.н.}} Y$, $n \to \infty$, $a, b = \text{const}$, тогда $aX_n + bY_n \xrightarrow{\text{п.н.}} aX + bY$, $n \to \infty$.

3) Пусть $g(x)$ — произвольная борелевская функция, заданная на прямой $\mathbb{R}^1$, а A — множество точек разрыва функции $g(x)$. Если $X_n \xrightarrow{\text{п.н.}} X$, $n \to \infty$, причем $\mathbf{P}(X \in A) = 0$, то $g(X_n) \xrightarrow{\text{п.н.}} g(X)$, $n \to \infty$.

4) Если $X_n \sim \mathcal{N}(m_n; D_n)$ и $X_n \xrightarrow{\text{с.к.}} X$, $n \to \infty$, то $X \sim \mathcal{N}(m_X; D_X)$, где $m_X = \lim\limits_{n\to\infty} m_n$, $D_X = \lim\limits_{n\to\infty} D_n$, причем пределы существуют и конечны.

5) Свойства 2) и 3) справедливы для сходимости по вероятности, а свойство 2) — для с.к.-сходимости.

Неравенство Чебышева

Теорема 2.1. *Для любых $\varepsilon > 0$ выполнено неравенство Чебышева*

$$\mathbf{P}(|\xi_n| > \varepsilon) \leqslant \frac{\mathrm{M}\{\xi_n^2\}}{\varepsilon^2}. \tag{2.1}$$

Борель-Кантелли

Теорема 2.2 (Борель-Кантелли). *Пусть $\{A_1, A_2, \ldots, A_n, \ldots\}$ — бесконечная последовательность случайных событий, а B =*

$= \bigcap_{n \geqslant 1} \bigcup_{k \geqslant n} A_k$ — *событие, состоящее в том, что произойдет бесконечно много событий* $\{A_i\}$. *Тогда*

1) *если* $\sum_{k=1}^{\infty} \mathbf{P}(A_k) < \infty$, *то* $\mathbf{P}(B) = 0$;

2) *если* $\{A_1, A_2, \ldots, A_n, \ldots\}$ *независимы и* $\sum_{k=1}^{\infty} \mathbf{P}(A_k) = \infty$, *то* $\mathbf{P}(B) = 1$.

Пример 2.1. Пусть $X_n \xrightarrow{\text{p}} X$, $Y_n \xrightarrow{\text{p}} Y$, $n \to \infty$. Показать, что $X_n + Y_n \xrightarrow{\text{p}} X + Y$, $n \to \infty$.

Центральная предельная теорема

Сходимость по распределению

Определение 3.1. Последовательность $\{X_n,\ n = 1, 2, \ldots\}$ *сходится по распределению* к случайной величине X, если для любого B такого, что $\mathbf{P}(X \in \partial B) = 0$, выполнено

$$\mathbf{P}(X_n \in B) \to \mathbf{P}(X \in B), \quad n \to \infty. \tag{3.1}$$

Слабая сходимость

Сходимость по распределению, которую также называют *слабой сходимостью*, будем обозначать

$$X_n \xrightarrow{\text{d}} X, \quad n \to \infty. \tag{3.2}$$

Теорема

Пусть $F_n(x)$ и $F_X(x)$ — функции распределения, соответственно, X_n и X.

Теорема 3.1. *Следующие три утверждения эквивалентны:*

1) $X_n \xrightarrow{\text{d}} X$, $n \to \infty$;

2) $F_n(x) \to F_X(x)$, $n \to \infty$ *для любой точки непрерывности функции* $F_X(x)$;

3) $\mathbf{M}\{g(X_n)\} \to \mathbf{M}\{g(X)\}$, $n \to \infty$ *для любой непрерывной и ограниченной на* $\mathbb{R}^1$ *функции* $g(x)$.

Теорема 3.2. *Если* $X_n \xrightarrow{\text{p}} X$, $n \to \infty$, *то* $X_n \xrightarrow{\text{d}} X$, $n \to \infty$.

Связь сходимостей

Таким образом, слабая сходимость случайной последовательности следует из сходимости по вероятности, а, следовательно, из сходимости почти наверное и сходимости в среднем квадратическом.

В одном важном частном случае сходимость по распределению и сходимость по вероятности эквивалентны: если $X_n \xrightarrow{\mathrm{d}} a$, $n \to \infty$, где $a = \text{const}$, то $X_n \xrightarrow{\mathrm{p}} a$, $n \to \infty$.

Сходимость, характеристические функции

Теорема 3.3. *Пусть $\Psi_n(\lambda)$ и $\Psi(\lambda)$ — характеристические функции, соответственно, X_n и X. Пусть также для любого $\lambda \in \mathbb{R}^1$*

$$\Psi_n(\lambda) \to \Psi(\lambda), \quad n \to \infty.$$

Тогда $X_n \xrightarrow{\mathrm{d}} X$, $n \to \infty$.

Асимптотически нормальная последовательность

Определение 3.2. Последовательность $\{X_n,\, n = 1, 2, \ldots\}$ называется *асимптотически нормальной* с параметрами $(m; \sigma^2)$, если

$$X_n \xrightarrow{\mathrm{d}} X, \quad n \to \infty, \text{ где } X \sim \mathcal{N}(m; \sigma^2). \tag{3.3}$$

Из определения 3.2 и теоремы 3.1 следует, что для любого $x \in \mathbb{R}^1$

$$F_n(x) \to \Phi\left(\frac{x - m}{\sigma}\right), \quad n \to \infty. \tag{3.4}$$

Стандартизированная сумма

Пусть теперь $\{\xi_k,\, k = 1, 2, \ldots\}$ — последовательность независимых случайных величин с параметрами

$$\mathbf{M}\{\xi_k\} = a; \quad \mathbf{D}\{\xi_k\} = \sigma^2 > 0.$$

Рассмотрим случайную *последовательность сумм* этих величин:

$$X_n = \sum_{k=1}^{n} \xi_k, \quad n = 1, 2, \ldots \tag{3.5}$$

Очевидно, $\mathbf{M}\{X_n\} = na$, $\mathbf{D}\{X_n\} = n\sigma^2$.

Введем *стандартизованную сумму*

$$\widetilde{X}_n = \frac{X_n - na}{\sigma\sqrt{n}}, \quad n = 1, 2, \ldots \tag{3.6}$$

Нетрудно проверить, что $\mathbf{M}\{\widetilde{X}_n\} = 0$, $\mathbf{D}\{\widetilde{X}_n\} = 1$.

Центральная предельная теорема для одинаково распределенных слагаемых

Теорема 3.4 (ЦПТ для одинаково распределенных слагаемых). *Пусть случайные величины* $\{\xi_k,\ k = 1, 2, \ldots\}$ *одинаково распределены, тогда последовательность* $\{\widetilde{X}_n,\ n = 1, 2, \ldots\}$ *асимптотически нормальна с параметрами* $(0; 1)$.

Следствие 3.1. *Для любых чисел* $a \leqslant b$ *выполнено*

$$\mathbf{P}(a \leqslant \widetilde{X}_n \leqslant b) \to \Phi(b) - \Phi(a), \quad n \to \infty. \tag{3.7}$$

Интегральная теорема Муавра-Лапласа

Следствие 3.2 (Интегральная теорема Муавра-Лапласа). *Пусть* X_n *— число успехов в серии из* n *испытаний Бернулли, а* p *— вероятность успеха в одном испытании. Тогда при* $n \to \infty$

$$\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathrm{d}} X \sim \mathcal{N}(0; 1). \tag{3.8}$$

Если слагаемые $\{\xi_k,\ k = 1, 2, \ldots\}$ распределены не одинаково, то утверждение, аналогичное ЦПТ, останется в силе при некоторых дополнительных ограничениях.

Теорема Ляпунова

Если слагаемые $\{\xi_k,\ k = 1, 2, \ldots\}$ распределены не одинаково, то утверждение, аналогичное ЦПТ, останется в силе при некоторых дополнительных ограничениях.

Пусть $\mathbf{M}\{\xi_k\} = a_k$, $\mathbf{D}\{\xi_k\} = \sigma_k^2$, $\mathbf{M}\{|\xi_k - a_k|^3\} = c_k^3$. Обозначим $A_n = \mathbf{M}\{X_n\} = \sum_{k=1}^{n} a_k$, $D_n^2 = \mathbf{D}\{X_n\} = \sum_{k=1}^{n} \sigma_k^2$, $C_n^3 = \sum_{k=1}^{n} c_k^3$.

Теорема 3.5 (Ляпунов). *Пусть* A_n, D_n, C_n *конечны при всех* $n \geqslant 1$, *причем* $\frac{C_n}{D_n} \to 0$, $n \to \infty$. *Тогда последовательность* $\{\widetilde{X}_n,\ n = 1, 2, \ldots\}$, *где* $\widetilde{X}_n = \dfrac{X_n - A_n}{D_n}$, *асимптотически нормальна с параметрами* $(0; 1)$.

Закон больших чисел

Законом больших чисел называется совокупность утверждений о поведении последовательности $\{\overline{X}_n,\ n = 1, 2, \ldots\}$ выборочных средних при $n \to \infty$.

Теорема 4.1. *Если* $\{X_k,\ k = 1, 2, \ldots\}$ *одинаково распределены и* $\mathbf{M}\{X_k\} = a$ *конечно, то* $\overline{X}_n \xrightarrow{\text{п.н.}} a$, $n \to \infty$.

Теорема 4.1 называется *“Усиленный закон больших чисел А.Н. Колмогорова”*.

Выборка и ее основные характеристики

Однородная выборка объема n

Определение 5.1. Совокупность $\{X_k, k = 1, \ldots, n\}$ независимых случайных величин, имеющих одинаковые функции распределения $F_{X_k}(x) = F(x)$, называется *однородной выборкой объема* n, соответствующей функции распределения $F(x)$.

Гауссовская выборка

Определение 5.2. Выборка $\{X_k, k = 1, \ldots, n\}$ называется *гауссовской*, если Z_n — n-мерный гауссовский вектор.

Неоднородная выборка

Определение 5.3. Выборка Z_n называется *неоднородной*, если законы распределения $F_{X_k}(x)$ ее элементов неодинаковы.

Реализация выборки

Определение 5.4. *Реализацией выборки* Z_n называется неслучайный вектор $z_n = \{x_1, \ldots, x_n\}^\top$, компонентами которого являются реализации соответствующих элементов выборки.

Порядковая статистика

Определение 5.5. СВ $X_{(k)}$, реализацией которой для каждой z_n является число $x_{(k)}$, называется *k-й порядковой статистикой*, $k = 1, \ldots, n$. Случайный вектор $Z_{(n)} = \{X_{(1)}, \ldots, X_{(n)}\}^\top$ называется *вариационным рядом* выборки.

СВ $X_{(1)}$ и $X_{(n)}$ (т.е. крайние элементы вариационного ряда) называются *экстремальными статистиками*.

Статистика

Определение 5.6. СВ $Y = \varphi(X_1, \ldots, X_n)$, где $\varphi(x_1, \ldots, x_n)$ — произвольная (борелевская) функция на $\mathbb{R}^n$, называется *статистикой*.

Основные понятия

Выборочное среднее

1) $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ называется *выборочным средним*.

Выборочная дисперсия

2) $\overline{S}_n^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \overline{X}_n)^2$ называется *выборочной дисперсией.*

Выборочный начальный момент

3) $\overline{\nu}_r(n) = \frac{1}{n}\sum_{k=1}^{n}(X_k)^r$, $r = 1, 2, \ldots$, называется *выборочным начальным моментом* r-го порядка.

Выборочный центральный момент r-ого порядка

4) $\overline{\mu}_r(n) = \frac{1}{n}\sum_{k=1}^{n}(X_k - \overline{X}_n)^r$, $r = 1, 2, \ldots$, называется *выборочным центральным моментом* r-го порядка.

Теорема

Теорема 5.1. *При неограниченном увеличении объема выборки n выборочные моменты $\overline{\nu}_r(n)$ и $\overline{\mu}_r(n)$, $r = 1, 2, \ldots$ почти наверное сходятся к теоретическим моментам ν_r и μ_r соответственно.*

Следствие 5.1. *Если $m_X = \mathrm{M}\{X_k\}$ существует, то $\overline{X}_n \xrightarrow{\text{п.н.}} m_X$, $n \to \infty$. Если ν_2 существует, то $\overline{S}_n^2 \xrightarrow{\text{п.н.}} D_X$, $n \to \infty$.*

Определение 5.8. Случайная функция $\widehat{F}_n(x) = \frac{M_n(x)}{n}$, $x \in \mathbb{R}^1$, называется *выборочной (эмпирической) функцией распределения* СВ X.

При достаточно больших n функция $\widehat{F}_n(x)$ весьма точно аппроксимирует функцию распределения $F(x)$, которой соответствует выборка, о чем свидетельствуют следующие утверждения.

Теорема 5.3 (Гливенко-Кантелли). *$\widehat{F}_n(x)$ сходится к $F(x)$ почти наверное равномерно по x при $n \to \infty$, т.е.*

$$\sup_{x \in \mathbb{R}^1} |\widehat{F}_n(x) - F(x)| \xrightarrow{\text{п.н.}} 0, \quad n \to \infty.$$

Теорема 5.4. *При любом $x \in \mathbb{R}^1$ последовательность $\{\widehat{F}_n(x), n = 1, 2, \ldots\}$ асимптотически нормальна:*

$$\sqrt{n}\left(\widehat{F}_n(x) - F(x)\right) \xrightarrow{\mathrm{d}} \xi \sim \mathcal{N}\left(0; F(x)(1 - F(x))\right), \quad n \to \infty.$$

Пусть двумерная выборка $\{(X_k, Y_k), k = 1, \ldots, n\}$ порождена случайным вектором $\xi = \{X, Y\}^\top$. Обозначим через $k_{XY} = \mathrm{M}\{(X - m_X)(Y - m_Y)\} = \mathrm{M}\{XY\} - m_X m_Y$ ковариацию случайных величин X и Y.

Определение 5.9. Статистика $\widehat{k}_{XY}(n) = \frac{1}{n}\sum_{k=1}^{n} X_k Y_k - \overline{X}_n \overline{Y}_n$ называется *выборочной ковариацией* случайных величин X и Y.

Теорема 5.5. *Если СВ X и Y имеют конечные дисперсии, то*

1) $\mathrm{M}\left\{\widehat{k}_{XY}(n)\right\} = \dfrac{n-1}{n} k_{XY}$;

2) $\widehat{k}_{XY}(n) \xrightarrow{\text{п.н.}} k_{XY}$, $n \to \infty$;

3) *если* $\mathrm{M}\{|X|^4 + |Y|^4\} < \infty$, *то*

$$\sqrt{n}\left(\widehat{k}_{XY}(n) - k_{XY}\right) \xrightarrow{\mathrm{d}} \eta \sim \mathcal{N}\left(0; \mu_{22} - k_{XY}^2\right), \quad n \to \infty,$$

где $\mu_{22} = \mathrm{M}\{(X - m_X)^2 (Y - m_Y)^2\}$.

Точечные оценки и их свойства