

Seminar 1: Виды сходимостей случайных векторов

Carlos Buitrago

Table of contents

0.1	Связь с разложением Тейлора	3
1	Дельта-метод	5
2	Предельные Теоремы	6
2.1	Пример: многомерная ЦПТ	6

Пусть $\xi, \xi_1, \xi_2, \dots$ — случайные векторы со значениями в $\mathbb{R}^m$,

$$\xi_n = (\xi_n^{(1)}, \dots, \xi_n^{(m)})^\top, \quad \xi = (\xi^{(1)}, \dots, \xi^{(m)})^\top.$$

Сходимость почти наверное

Последовательность случайных векторов $\{\xi_n\}_{n \geq 1}$ называется **сходящейся почти наверное** к случайному вектору ξ , если

$$\mathbb{P} \left(\omega : \lim_{n \rightarrow \infty} \xi_n(\omega) = \xi(\omega) \right) = 1.$$

Обозначение:

$$\xi_n \xrightarrow{\text{п.н.}} \xi.$$

Эквивалентно, для любой нормы $\|\cdot\|$ на $\mathbb{R}^m$,

$$\mathbb{P} \left(\lim_{n \rightarrow \infty} \|\xi_n - \xi\| = 0 \right) = 1.$$

Сходимость по вероятности

Последовательность случайных векторов $\{\xi_n\}_{n \geq 1}$ называется **сходящейся по вероятности** к случайному вектору ξ , если для любого $\varepsilon > 0$

$$\mathbb{P} (\|\xi_n - \xi\| \geq \varepsilon) \longrightarrow 0, \quad n \rightarrow \infty.$$

Обозначение:

$$\xi_n \xrightarrow{\mathbb{P}} \xi.$$

Сходимость в L^p

Пусть $p \geq 1$. Последовательность случайных векторов $\{\xi_n\}_{n \geq 1}$ называется **сходящейся в L^p** к случайному вектору ξ , если

$$\mathbb{E}\|\xi_n - \xi\|^p \longrightarrow 0, \quad n \rightarrow \infty.$$

Обозначение:

$$\xi_n \xrightarrow{L^p} \xi.$$

Сходимость по распределению

Последовательность случайных векторов $\{\xi_n\}_{n \geq 1}$ называется **сходящейся по распределению** к случайному вектору ξ , если для любой ограниченной непрерывной функции $f: \mathbb{R}^m \rightarrow \mathbb{R}$, выполнено

$$\mathbb{E}f(\xi_n) \longrightarrow \mathbb{E}f(\xi).$$

Обозначение:

$$\xi_n \xrightarrow{d} \xi.$$

Сходимость по распределению является свойством распределений случайных векторов. В отличие от сходимости почти наверное, по вероятности или в L^p , случайные векторы ξ_n и ξ не обязаны быть определены на одном и том же вероятностном пространстве.

Теорема Александрова для случайных векторов

Сходимость по распределению $\xi_n \xrightarrow{d} \xi$ имеет место тогда и только тогда, когда

$$F_{\xi_n}(x) \longrightarrow F_{\xi}(x)$$

для любой точки $x \in \mathbb{R}^m$, в которой функция распределения F_{ξ} непрерывна.

Теорема: Связь между типами сходимости

Для случайных векторов справедливы импликации

$$\begin{array}{ccccc} \xi_n \xrightarrow{\text{п.н.}} \xi & & \xi_n \xrightarrow{\mathbb{P}} \xi & \implies & \xi_n \xrightarrow{d} \xi \\ \xi_n \xrightarrow{L^p} \xi & \implies & \xi_n \xrightarrow{\mathbb{P}} \xi & \implies & \xi_n \xrightarrow{d} \xi \end{array}$$

Обратные импликации в общем случае неверны.

Асимптотические обозначения $o_{\mathbb{P}}$ и $O_{\mathbb{P}}$

В асимптотической статистике удобно использовать вероятностные аналоги классических обозначений $o(\cdot)$ и $O(\cdot)$.

Говорят, что $X_n = o_{\mathbb{P}}(1)$, если

$$X_n \xrightarrow{\mathbb{P}} 0.$$

Более общо, $X_n = o_{\mathbb{P}}(a_n)$, если

$$\frac{X_n}{a_n} \xrightarrow{\mathbb{P}} 0.$$

Последовательность случайных величин $\{X_n\}$ называется **ограниченной по вероятности**, и пишут $X_n = O_{\mathbb{P}}(1)$, если для любого $\varepsilon > 0$ существует $M > 0$ такое, что

$$\limsup_{n \rightarrow \infty} \mathbb{P}(|X_n| > M) \leq \varepsilon.$$

Более общо, $X_n = O_{\mathbb{P}}(a_n)$, если

$$\frac{X_n}{a_n} = O_{\mathbb{P}}(1).$$

Например, если $\sqrt{n}(T_n - \theta) \xrightarrow{d} Z$, то

$$T_n - \theta = O_{\mathbb{P}}(n^{-1/2}).$$

0.1 Связь с разложением Тейлора

Пусть

$$T_n - \theta = O_{\mathbb{P}}(n^{-1/2}).$$

Если функция h дважды дифференцируема в окрестности θ , то

$$h(T_n) = h(\theta) + h'(\theta)(T_n - \theta) + O_{\mathbb{P}}((T_n - \theta)^2).$$

Поскольку

$$(T_n - \theta)^2 = O_{\mathbb{P}}(n^{-1}),$$

получаем

$$h(T_n) = h(\theta) + h'(\theta)(T_n - \theta) + O_{\mathbb{P}}(n^{-1}).$$

Именно такая запись часто используется в асимптотической теории статистических оценок.

Теорема: Сходимость случайного вектора и его компонент

Пусть

$$\xi_n = (\xi_n^{(1)}, \dots, \xi_n^{(m)})^{\top}, \quad \xi = (\xi^{(1)}, \dots, \xi^{(m)})^{\top}.$$

Тогда справедливо

$$\xi_n \xrightarrow{\text{п.н.}} \xi \iff \xi_n^{(j)} \xrightarrow{\text{п.н.}} \xi^{(j)} \quad \forall j = 1, \dots, m.$$

$$\xi_n \xrightarrow{\mathbb{P}} \xi \iff \xi_n^{(j)} \xrightarrow{\mathbb{P}} \xi^{(j)} \quad \forall j = 1, \dots, m.$$

$$\xi_n \xrightarrow{L^p} \xi \iff \xi_n^{(j)} \xrightarrow{L^p} \xi^{(j)} \quad \forall j = 1, \dots, m.$$

Замечание о сходимости по распределению

Для сходимости по распределению покомпонатная сходимость уже недостаточна. То есть

$$\xi_n \xrightarrow{d} \xi \implies \xi_n^{(j)} \xrightarrow{d} \xi^{(j)}, \quad j = 1, \dots, m.$$

Однако обратная импликация в общем случае неверна.

Причина состоит в том, что покомпонатная сходимость определяет только предельные маргинальные распределения компонент. Она не содержит полной информации о зависимости между компонентами и, следовательно, не определяет совместное предельное распределение.

Теорема: Критерий Крамера–Уолда

Пусть ξ_n, ξ — случайные векторы в $\mathbb{R}^m$. Тогда

$$\xi_n \xrightarrow{d} \xi$$

тогда и только тогда, когда для любого $t \in \mathbb{R}^m$

$$t^\top \xi_n \xrightarrow{d} t^\top \xi.$$

Критерий Крамера–Уолда показывает, что многомерная слабая сходимость полностью определяется поведением всех одномерных линейных проекций

$$t^\top \xi_n.$$

Этот результат является одним из основных инструментов доказательства многомерной центральной предельной теоремы.

теорема о наследовании сходимости

Пусть

$$h : \mathbb{R}^m \rightarrow \mathbb{R}^s$$

непрерывна почти всюду относительно распределения ξ .

Тогда

$$\xi_n \xrightarrow{\text{п.н.}} \xi \implies h(\xi_n) \xrightarrow{\text{п.н.}} h(\xi),$$

$$\xi_n \xrightarrow{\mathbb{P}} \xi \implies h(\xi_n) \xrightarrow{\mathbb{P}} h(\xi),$$

$$\xi_n \xrightarrow{d} \xi \implies h(\xi_n) \xrightarrow{d} h(\xi).$$

Лемма Slutsky

Пусть

$$\xi_n \xrightarrow{d} \xi, \quad \eta_n \xrightarrow{d} c,$$

где c — детерминированная константа. Тогда

$$(\xi_n, \eta_n) \xrightarrow{d} (\xi, c).$$

1 Дельта-метод

Одномерный случай

Пусть

$$\xi_n \xrightarrow{d} \xi, \quad b_n \rightarrow 0, \quad b_n \neq 0,$$

и пусть $h : \mathbb{R} \rightarrow \mathbb{R}$ — функция, дифференцируемая в точке $a \in \mathbb{R}$.
Тогда

$$\frac{h(a + \xi_n b_n) - h(a)}{b_n} \xrightarrow{d} \xi h'(a).$$

В частности, если $\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, \sigma^2)$, то

$$\sqrt{n}(h(T_n) - h(\theta)) \xrightarrow{d} N(0, [h'(\theta)]^2 \sigma^2).$$

Многомерный случай

Пусть $\xi_n \xrightarrow{d} \xi$, где ξ_n, ξ — случайные векторы размерности m , и пусть

$$h : \mathbb{R}^m \rightarrow \mathbb{R}$$

— функция, дифференцируемая в точке $a \in \mathbb{R}^m$. Пусть также

$$b_n \rightarrow 0, \quad b_n \neq 0.$$

Тогда

$$\frac{h(a + \xi_n b_n) - h(a)}{b_n} \xrightarrow{d} \langle \xi, \nabla h(a) \rangle.$$

В частности, если $\sqrt{n}(T_n - \theta) \xrightarrow{d} N_m(0, \Sigma)$, то

$$\sqrt{n}(h(T_n) - h(\theta)) \xrightarrow{d} N(0, \nabla h(\theta)^\top \Sigma \nabla h(\theta)).$$

Дельта-метод второго порядка: одномерный случай

Пусть

$$\xi_n \xrightarrow{d} \xi, \quad b_n \rightarrow 0, \quad b_n \neq 0,$$

и пусть $h : \mathbb{R} \rightarrow \mathbb{R}$ — функция, дважды дифференцируемая в точке $a \in \mathbb{R}$.
Предположим, что $h'(a) = 0$. Тогда

$$\frac{h(a + \xi_n b_n) - h(a)}{b_n^2} \xrightarrow{d} \frac{1}{2} \xi^2 h''(a).$$

2 Предельные Теоремы

Многомерный УЗБЧ

Пусть $X_1, X_2, \dots$ — независимые одинаково распределенные случайные векторы в $\mathbb{R}^m$ такие, что $\mathbb{E}\|X_1\| < \infty$. Тогда

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\text{п.н.}} \mathbb{E}X_1.$$

Доказательство следует непосредственно из одномерного усиленного закона больших чисел, примененного к каждой компоненте:

$$\frac{1}{n} \sum_{i=1}^n X_i^{(j)} \xrightarrow{\text{п.н.}} \mathbb{E}X_1^{(j)}, \quad j = 1, \dots, m.$$

После этого используется эквивалентность покоординатной и векторной сходимости почти наверное.

Многомерная ЦПТ

Пусть $X_1, X_2, \dots$ — независимые одинаково распределенные случайные векторы в $\mathbb{R}^m$ такие, что $\mathbb{E}X_1 = \mu$ и $\text{Cov}(X_1) = \Sigma$, где все вторые моменты конечны. Тогда

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - \mu) \xrightarrow{d} N_m(0, \Sigma).$$

Если ковариационная матрица Σ положительно определена, то существует матрица $\Sigma^{-1/2}$. Из многомерной ЦПТ следует

$$\sqrt{n} \Sigma^{-1/2} (\bar{X}_n - \mu) \xrightarrow{d} N_m(0, I_m).$$

2.1 Пример: многомерная ЦПТ

Рассмотрим случайный вектор $X = (X_1, X_2)^\top$ с независимыми компонентами

$$X_1 \sim \text{Exp}(1), \quad X_2 \sim \text{Exp}(1).$$

Тогда распределение X существенно не является нормальным.

Однако многомерная ЦПТ утверждает, что

$$\sqrt{n} (\bar{X}_n - \mathbb{E}X) \xrightarrow{d} N_2(0, I_2).$$

Посмотрим, как меняется форма распределения при росте n .

