Содержание

1 Виды сходимости случайных векторов и связи между ними

Пусть $\xi, \{\xi_n\}_{n=1}^{\infty}$ – случайные векторы размерности.

Определение 1.1. Сходимость почти наверное:

$$\xi_n \overset{\text{п.н.}}{\to} \xi \Leftrightarrow P(\xi_n \to \xi) = 1$$

Определение 1.2. Сходимость по вероятности:

$$\xi_n \overset{\mathrm{P}}{\to} \xi \Leftrightarrow \forall \varepsilon > 0 : \lim_{n\to\infty} P(\|\xi_n - \xi\|_2 > \varepsilon) = 0$$

Определение 1.3. Сходимость в L_p (в среднем):

$$\xi_n \overset{L_p}{\to} \xi \Leftrightarrow \lim_{n\to\infty} \mathbb{E}(\|\xi_n - \xi\|_p)^p = 0$$

Определение 1.4. Сходимость по распределению:

$$\xi_n \overset{\mathrm{d}}{\to} \xi \Leftrightarrow \forall f \in \mathrm{BC}(\mathbb{R}^m) : \mathbb{E}f(\xi_n) \overset{n\to\infty}{\to} \mathbb{E}f(\xi)$$

Утверждение 1.1. *Связь между сходимостями:*

1. *п.н.* $\Rightarrow P$
2. $L_p \Rightarrow P$
3. $P \Rightarrow d$

Утверждение 1.2. $\xi_n \overset{d}{\to} const \Rightarrow \xi_n \overset{P}{\to} const$

Утверждение 1.3. *Связь между сходимостью векторов и сходимостью их компонент:*

1. $$\xi_n \overset{\text{п.н.}}{\to} \xi \Leftrightarrow \forall i : \xi_n^{(i)} \overset{\text{п.н.}}{\to} \xi^{(i)}$$

2. $$\xi_n \overset{P}{\to} \xi \Leftrightarrow \forall i : \xi_n^{(i)} \overset{P}{\to} \xi^{(i)}$$

3. $$\xi_n \overset{L_p}{\to} \xi \Leftrightarrow \forall i : \xi_n^{(i)} \overset{L_p}{\to} \xi^{(i)}$$

4. $$\xi_n \overset{d}{\to} \xi \Rightarrow \forall i : \xi_n^{(i)} \overset{d}{\to} \xi^{(i)}$$

Доказательство. 1.

$$\cap_{i=1}^m \{\xi_n^{(i)} \to \xi^{(i)}\} = \{\xi_n \to \xi\} \subset \{\xi_n^{(i)} \to \xi^{(i)}\}$$

Тогда для $\Rightarrow$ используем включение и свойство меры:

$$1 = P(\{\xi_n \to \xi\}) \leqslant P(\{\xi_n^{(i)} \to \xi^{(i)}\})$$

А для $\Leftarrow$:

$$1 = P(\cap_{i=1}^m \{\xi_n^{(i)} \to \xi^{(i)}\}) = P(\{\xi_n \to \xi\})$$

2. Для $\Rightarrow$:

$$\{|\xi_n^{(i)} - \xi^{(i)}| > \varepsilon\} \subset \{\|\xi_n - \xi\|_2 > \varepsilon\}$$

   А для $\Leftarrow$:

$$\{\|\xi_n - \xi\|_2 > \varepsilon\} \subset \bigcup_{i=1}^{m} \left\{|\xi_n^{(i)} - \xi^{(i)}| > \frac{\varepsilon}{\sqrt{m}}\right\}$$

3. Заметим, что

$$\forall i: \lim_{n\to\infty} \mathbb{E}|\xi_n^{(i)} - \xi^{(i)}|^p = 0 \Leftrightarrow \lim_{n\to\infty} \mathbb{E}\|\xi_n - \xi\|_p^p = \lim_{n\to\infty} \mathbb{E}\sum_{i=1}^{n} |\xi_n^{(i)} - \xi^{(i)}|^p = 0$$

4. Для $\Rightarrow$ в качестве f возьмём функцию-проектор.

$\square$

Теорема 1.1. *О наследовании сходимостей.*

Пусть $\xi, \{\xi_n\}_{n=1}^{\infty}$ – случайные векторы в $\mathbb{R}^m$, причём $\exists B \in \mathcal{B}(\mathbb{R}^m): P(\xi \in B) = 1$ и $h: \mathbb{R}^m \to \mathbb{R}^k$ непрерывна в каждой точке множества B. Тогда

$$\xi_n \overset{\text{п.н.}, P, d}{\to} \xi \Rightarrow h(\xi_n) \overset{\text{п.н.}, P, d}{\to} h(\xi) \tag{1}$$

Доказательство.

- Случай п.н.:

$$P(h(\xi_n) \to h(\xi)) \geqslant P(h(\xi_n) \to h(\xi), \xi \in B) \geqslant P(\xi_n \to \xi, \xi \in B) = 1$$

- Случай P:

  Пусть $h(\xi_n) \overset{P}{\not\to} h(\xi) \Rightarrow$:

$$\exists \varepsilon_0, \delta_0, \{n_k\}_{k=1}^{\infty}: P(\|h(\xi_{n_k}) - h(\xi)\| > \varepsilon_0) \geqslant \delta_0$$

  Но из неё мы можем выбрать $\{\xi_{n_{k_m}}\}_{m=1}^{\infty}$, сходящуюся почти всюду (по прошлому семестру), но тогда мы получили противоречие с предыдущим пунктом доказательства.

- Докажем для непрерывных h:

  Тогда

$$\forall f \in BC(\mathbb{R}^k): f(h(x)) \in BC(\mathbb{R}^m)$$

  Значит мы можем взять $f \circ h$ в качестве функции из определения сходимости по вероятности и получить требуемое.

$\square$

2 Закон больших чисел, усиленный закон больших чисел...

Теорема 2.1. *ЗБЧ.*

Пусть $\{\xi_n\}_{n=1}^{\infty}$ – попарно некорелированные вектора и $\sup_{n,i} \mathbb{V}\xi_n^{(i)} \leqslant C$. Тогда

$$\frac{s_n - \mathbb{E}s_n}{n} \overset{P}{\to} 0$$

где $\{s_n\}_{n=1}^{\infty} = \{\sum_{i=1}^{n} \xi_i\}_{n=1}^{\infty}$

Теорема 2.2. *УЗБЧ.*

Пусть $\{\xi_n\}_{n=1}^{\infty}$ *– независимые одинаково распределённые, причём* $\mathbb{E}\xi < +\infty$. *Тогда*

$$\frac{s_n}{n} \xrightarrow{п.н.} \mathbb{E}\xi$$

Теорема 2.3. *ЦПТ.*

Пусть $\{\xi_n\}_{n=1}^{\infty}$ *– независимые одинаково распределённые, причём* $\exists$ *ковариационные матрица* $\mathbb{V}\xi$. *Тогда*

$$\sqrt{n}\left(\frac{s_n}{n} - \mathbb{E}\xi\right) \xrightarrow{d} \mathcal{N}(0, \mathbb{V}\xi)$$

Лемма 2.1. *Лемма Слуцкого.*

Пусть $\xi_n \xrightarrow{d} \xi$ *и* $\eta_n \xrightarrow{d} c\,(const)$. *Тогда*

$$\xi_n + \eta_n \xrightarrow{d} \xi + \eta; \quad \xi_n \cdot \eta_n \xrightarrow{d} \xi \cdot c$$

Доказательство. По некому утверждению без доказательства, будет верно

$$\begin{pmatrix} \xi_n \\ \eta_n \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \xi \\ c \end{pmatrix}$$

Тогда, применив теорему о наследовании сходимостей с функциями $+, \cdot$ всё получится. $\square$

Пример. Применение леммы Слуцкого.

Пусть $\xi_n \xrightarrow{d} \xi$ – последовательность случайных величин и $H : \mathbb{R} \to \mathbb{R}$ – дифференцируемая в точке a и $b_n \to 0$, причём $b_n \neq 0$. Тогда

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

Доказательство. Введём

$$h(x) := \begin{cases} \frac{H(a+x)-H(a)}{x}, x \neq 0 \\ H'(a), x = 0 \end{cases}$$

Тогда h непрерывна в 0.

По лемме Слуцкого:

$$b_n \xi_n \xrightarrow{d} 0$$

По теореме о наследовании сходимости

$$h(b_n \xi_n) \xrightarrow{d} h(0) = H'(a) \Rightarrow \frac{H(a + \xi_n b_n) - H(a)}{b_n} = h(b_n \xi_n)\xi_n \xrightarrow{d} H'(a)\xi$$

$\square$

Теорема 2.4. *Обобщение на многомерный случай.*

Пусть $\xi_n \xrightarrow{d} \xi$ *в* $\mathbb{R}^m$, *и* $H : \mathbb{R}^m \to \mathbb{R}^s$, *у которой в точке* $a \in \mathbb{R}^m \exists$ *матрица частных производных* $H'(x) = \left(\frac{\partial H_i}{\partial x_j}\right)_{i=1,\,j=1}^{s,\,m}$, *а также числовая последовательность* $b_n \to 0, b_n \neq 0$. *Тогда*

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

3 Вероятно-статистическая модель...

Пусть $(\Omega, \mathcal{F})$ и $(E, \mathcal{E})$ – измеримые пространства.

Определение 3.1. Если $\xi : \Omega \to E$ такова, что

$$\forall B \in \mathcal{E} : \xi^{-1}(B) \in \mathcal{F}$$

то ξ называется **случайным элементом**.

Если $(E, \mathcal{E}) = (\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$, то ξ называется **случайным вектором**.

Более того, если $m = 1$, то ξ называется **случайном величиной**.

Определение 3.2. **Распределением** случайного элемента ξ называется мера P_ξ на $\mathcal{E}$, такая что $P_\xi(B) = P(\xi \in B)$

Определение 3.3. **Выборочное** пространство $\mathcal{X}$ – множество всевозможных исходов одного эксперимента (обычно $\mathbb{R}^m$).

$\mathcal{B}_\mathcal{X}$ – σ-алгебра на $\mathcal{X}$ будем считать Барелевской.

Замечание. Построим модель эксперимента, как случайной величины.

Пусть

$$\forall x \in \mathcal{X} : X(x) = X$$

получим отображение $X : \mathcal{X} \to \mathcal{X}$, которое является случайным элементом на вероятностном пространстве $(\mathcal{X}, \mathcal{B}_\mathcal{X}, P)$ и имеет распределение $P_X = P$

Замечание. Построим модель n независимых повторений нашего эксперимента.

Рассмотрим $\mathcal{X}^n = \mathcal{X} \times \cdots \times \mathcal{X}$ и $\mathcal{B}_\mathcal{X}^n = \mathcal{B}(\mathcal{X}^n) = \sigma(B_1 \times \cdots \times B_n), B_i \in \mathcal{B}_\mathcal{X}$, а $P^n = P \otimes \cdots \otimes P$ – мера на $(\mathcal{X}^n, \mathcal{B}_\mathcal{X}^n)$, такая что $P^n(B_1 \times \cdots \times B_n) = P(B_1) * \cdots * P(B_n)$.

Для этого рассмотрим тождественное отображение $X : \mathcal{X}^n \to \mathcal{X}^n$. Его i-я компонента X_i является случайным вектором с распределением P, причём $X_1, \cdots, X_n$ независимы в совокупности.

Определение 3.4. Совокупность $X = (X_1, \cdots, X_n)$ независимых одинаково распределённых случайных величин (или векторов) с распределением P называется **выборкой** размера n из распределения P.

Также выборку X иногда будем называть **наблюдением**.

Замечание. Для бесконечных выборок определим $\mathcal{X}^\infty = \mathcal{X} \times \mathcal{X} \times \cdots$ и $\mathcal{B}_\mathcal{X}^\infty = \sigma(\{B_1 \times \cdots \times B_n \times \mathcal{X} \times \mathcal{X} \times \mathcal{X} \times \cdots\}_{n=1}^\infty)$, а меру $P^\infty(B_1 \times \cdots \times B_n \times \mathcal{X} \times \cdots) = P(B_1) * \cdots * P(B_n)$, такая мера существует и единственна.

Аналогично предыдущим пунктам определяем **бесконечную серию экспериментов**.

Определение 3.5. Тройка $(\mathcal{X}, \mathcal{B}_\mathcal{X}, \mathcal{P})$ называется вероятностно-статистической моделью.

Замечание. Пусть $X_1, \cdots, X_n$ – случайные величины (или векторы), и $X_1(\omega) = x_1, \cdots, X_n(\omega) = x_n$ – их значения, называются **реализацией выборки**.

Задачей статистики является сделать вывод о неизвестном распределении по реализации выборки.

Определение 3.6. Вероятно-статистическая модель $(\mathcal{X}, \mathcal{B}_\mathcal{X}, \mathcal{P})$ называется **параметрической**, если семейство $\mathcal{P}$ параметризованно, то есть

$$\mathcal{P} = \{P_\theta, \theta \in \Theta\}$$

обычно $\Theta \subset \mathbb{R}^m$.

4 Эмпирическое распределение и эмпирическая функция распределения

Определение 4.1. Для $\forall B \in \mathcal{B}(\mathbb{R}^m)$ положим

$$P_n^*(B) := \frac{\sum_{i=1}^n \mathbb{I}\{X_i \in B\}}{n}$$

распределение P_n^* называется **эмпирическим распределением**, построенным по выборке $X_1, \cdots, X_n$.

Это случайное распределение (зависит от ω)

Определение 4.2. Функция $F_n^*(x) = \frac{\sum_{i=1}^n \mathbb{I}\{X_i \leqslant x\}}{n}$ называется **эмпирической функцией распределения**.

Теорема 4.1. *Гливенко-Кантелли.*

Пусть $X_1, \cdots, X_n$ – независимые случайные величины с функцией распределения $F(x)$. Тогда

$$D_n = \sup_{x \in \mathbb{R}} |F_n^*(x) - F(x)| \xrightarrow{п.н.} 0$$

Доказательство. Почему D_n – случайная величина?

F непрерывна справа, и $\forall \omega :\ F_n^*$ также непрерывна справа $\Rightarrow$

$$D_n(\omega) = \sup_{x \in \mathbb{R}} |F_n^*(x) - F(x)| = \sup_{x \in \mathbb{Q}} |F_n^*(x) - F(x)|$$

Значит D_n является случайной совокупностью случайных величин $\Rightarrow$ D_n – случайная величина.

Фиксируем $N \in \mathbb{N}$, тогда $\forall k \in \{1, \cdots, N-1\}$ положим

$$X_{N,K} = \inf\{x \in \mathbb{R} \mid F(x) \geqslant \frac{K}{N}\}$$

Заметим, что это число конечно, а также определим $X_{N,0} = -\infty, X_{N,N} = +\infty$.

Если $x \in [X_{N,K}, X_{N,K+1}) \Rightarrow$

$$\begin{aligned} F_n^*(x) - F(x) \leqslant F_n^*(X_{N,K+1} - 0) - F(X_{N,K}) = \\ F_n^*(X_{N,K+1} - 0) - F(X_{N,K+1} - 0) + F(X_{N,K+1} - 0) - F(X_{N,K}) \leqslant \\ F_n^*(X_{N,K+1} - 0) - F(X_{N,K+1} - 0) + \frac{1}{N} \end{aligned}$$

Последний переход получили благодаря тому, что $F(X_{N,K+1} - 0)$ – отсуп чуть влево, от нижней границы значения, где $F(x) \geqslant \frac{K+1}{N}$, значит там $\leqslant \frac{K+1}{N}$. Ну а $F(X_{N,K})$ по определению $\geqslant \frac{K}{N}$.

Аналогично $F_n^*(x) - F(x) \geqslant F_n^*(X_{N,K}) - F(X_{N,K}) - \frac{1}{N}$. Тогда

$$|F_N^*(x) - F(x)| \leqslant \max(|F_n^*(X_{N,K+1} - 0) - F(X_{N,K+1} - 0)|, |F_n^*(X_{N,K}) - F(X_{N,K})|) + \frac{1}{N}$$

Но тогда супремум по всей прямой

$$\sup_{x \in \mathbb{R}} |F_N^*(x) - F(x)| \leqslant \max_{0 \leqslant K \leqslant N-1} \max(|F_n^*(X_{N,K+1} - 0) - F(X_{N,K+1} - 0)|, |F_n^*(X_{N,K}) - F(X_{N,K})|) + \frac{1}{N}$$

Из какого-то вспомогательного утверждения следует, что $F_n^*(y-0) = P_n^*((-\infty, y)) \to P_X((-\infty, y)) = F(y-0)$.

Теперь для ε фиксируем $\frac{1}{N} < \varepsilon \Rightarrow$

$$\overline{\lim}_n \sup_{x\in\mathbb{R}} |F_N^*(x) - F(x)| \overset{\text{п.н.}}{<} \varepsilon$$

В силу произвольности ε получаем требуемое. □

5 Статистики и оценки

Определение 5.1. Пусть $(\mathcal{X}, \mathcal{B}_\mathcal{X}, \mathcal{P})$ – вероятно-статистическая модель, X – наблюдение, $(E, \mathcal{E})$ – измеримое пространство, и $S: \mathcal{X} \to E$ – измеримое отображение. Тогда $S(x)$ называется **статистикой**.

Определение 5.2. Пусть X – наблюдение в параметрической модели $(\mathcal{X}, \mathcal{B}_\mathcal{X}, \{P_\theta\}_{\theta\in\Theta})$ и $S(X)$ – статистика со значениями в Θ. Тогда $S(X)$ называется **оценкой** неизвестного параметра Θ.

Пример. Пусть $X = (X_1, \cdots, X_n)$ – выборка из распределения в $\mathbb{R}^n$.

1. Если $g(x)$ – борелевская функция, то

$$\overline{g(X)} = \frac{1}{n}\sum_{i=1}^n g(X_i)$$

называется выборочной характеристикой функции $g(x)$. Например $\overline{X} = \frac{\sum X_i}{n}$ – выборочное среднее. $\overline{X^k} = \frac{1}{n}\sum_{i=1}^n X_i^k$ – выборочный момент k-го порядка.

2. Функции от выборочных квантилей:

$$S(X) = h(\overline{g_1(X)}, \cdots, \overline{g_k(X)})$$

где h – борелевская.

Например, $s^2 = \overline{X^2} - (\overline{X})^2$ – выборочная дисперсия. $M_k = \frac{1}{n}\sum(X_i - \overline{X})^k$ – выборочный центральный момент k-го порядка.

3. Порядковые статистики:

$$X_{(1)} = \min(X_1, \cdots, X_n)$$
$$X_{(2)} \text{ – второй элемент в отсортированной выборке}$$
$$X_{(n)} = \max(X_1, \cdots, X_n)$$

вектор $(X_{(1)}, \cdots, X_{(n)})$ называется **вариационным рядом**.

Пусть $X = (X_1, \cdots, X_n)$ – выборка из неизвестного распределения $P \in \{P_\theta, \theta \in \Theta\}, \Theta \subset \mathbb{R}^k$.

Определение 5.3. Оценка $\theta^*(X)$ называется **несмещённой** оценкой параметра θ, если

$$\forall\theta \in \Theta: \ \mathbb{E}_\theta \theta^*(X) = \theta$$

где $\mathbb{E}_\theta$ – матожидание в случае, когда элементы выборки имеют распределение P_θ.

Определение 5.4. Оценка $\theta^*_n(X_1, \cdots, X_n)$ (а точнее последовательность оценок) называется **состоятельной**, если

$$\forall\theta \in \Theta :\ \theta^*(X) \xrightarrow{P_\theta} \theta$$

и называется **сильно состоятельной** если

$$\forall\theta \in \Theta :\ \theta^*(X) \xrightarrow{P_\theta\text{- п. н.}} \theta$$

Определение 5.5. Оценка $\theta^*(X_1, \cdots, X_n)$ называется **асимптотически нормальной** оценкой θ, если

$$\forall\theta \in \Theta :\ \sqrt{n}(\theta^*_n - \theta) \xrightarrow{d_\theta} \mathcal{N}(0, \sigma^2(\theta))$$

Утверждение 5.1. *Пусть $T(X)$ – асимптотически нормальная оценка для $\tau(\theta)$. Тогда $T(X)$ – состоятельная оценка для $\tau(\theta)$.*

Доказательство. Используя лемму Слуцкого, получаем

$$\frac{1}{\sqrt{n}} \cdot \sqrt{n}(T_n - \tau(\theta)) \xrightarrow{d_\theta} 0$$

Но мы знаем, что из сходимости по распределению к константе следует сходимость по мере. □

Утверждение 5.2. *Из сильной состоятельности и асимптотической нормальности оценки следует её состоятельность.*

Доказательство. Следствие из сильной состоятельности автоматически следует из связи сходимостей.

Следствите из асимптотической нормальности было доказано в предыдущем утверждении. □

6 О наследовании состоятельностей

Утверждение 6.1. *Наследование состоятельности и сильной состоятельности при взятии непрерывной функции.*

Пусть $\theta^*_n(X)$ – сильно состоятельная (состоятельная) оценка θ. Если $\tau : \mathbb{R}^k \to \mathbb{R}^s$ непрерывна на $\Theta \subset \mathbb{R}^k$, то $\tau(\theta^*_n)$ – сильно состоятельная (состоятельная) оценка $\tau(\theta)$.

Доказательство. Смотри доказательство теоремы о наследовании сходимости. □

Лемма 6.1. *О наследовании асимптотической нормальности.*

Пусть $\theta^*_n(X)$ – асимптотически нормальная оценка $\theta \in \Theta$ с асимптотической дисперсией $\sigma^2(\theta)$ и числовая функция $T : \mathbb{R} \to \mathbb{R}$ дифференцируема в $\forall\theta \in \Theta$. Тогда $T(\theta^*_n)$ – асимптотически нормальная оценка $T(\theta)$ с асимптотической дисперсией $\sigma^2(\theta)(T'(\theta))^2$

Доказательство. Фиксируем $\theta, \xi_n := \sqrt{n}(\theta^*_n(X) - \theta) \xrightarrow{d_\theta} \xi \sim \mathcal{N}(0, \sigma^2(\theta)), b_n := \frac{1}{\sqrt{n}} \to 0$.

Вспомним дельта метод, взяв

$$a = \theta, h = T \Rightarrow \frac{T(\theta + \xi_n b_n) - T(\theta)}{b_n} \xrightarrow{d_\theta} T'(\theta)\xi \Rightarrow \sqrt{n}(T(\theta^*_n) - T(\theta)) \xrightarrow{d_\theta} T'(\theta)\mathcal{N}(0, \sigma^2(\theta))$$

□

7 Метод подстановки и метод моментов

Определение 7.1. Пусть в параметрическом семействе $\{P_\theta, \theta \in \Theta\}$ для некоторой функции G выполнено:

$$\forall \theta \in \Theta: \ \theta = G(P_\theta)$$

Тогда оценкой по **методу подстановки** называется $\theta^*(X_1, \cdots, X_n) = G(P_n^*)$

Пусть $X_1, \cdots, X_n$ – выборка из $P \in \{P_\theta, \theta \in \Theta\}, \Theta \subset \mathbb{R}^k$. Рассмотрим барелевские функции $g_1(x), \cdots, g_k(x)$ со значениями в $\mathbb{R}$.

Пусть $m_1(\theta) = \mathbb{E}_\theta g_i(X_1)$ конечно при $1 \leqslant i \leqslant k$.

Определение 7.2. Если $\exists!$ решение системы

$$\begin{cases} m_1(\theta) = \overline{g_1(X)} \\ \dots \\ m_k(\theta) = \overline{g_k(X)} \end{cases}$$

Тогда оценкой по **методу моментов** называется $\theta^* = m^{-1}(\overline{g})$, где

$$m(\theta) := \begin{pmatrix} m_1(\theta) \\ \vdots \\ m_k(\theta) \end{pmatrix}; \quad \overline{g} = \begin{pmatrix} \frac{\sum_{i=1}^n g_1(X_i)}{n} \\ \vdots \\ \frac{\sum_{i=1}^n g_k(X_i)}{n} \end{pmatrix}$$

Стандартные **пробные функции**: $g_i(X) = X^i$ (i-й момент).

Замечание. О связи методов.

Заметим, что

$$\theta = m^{-1} \begin{pmatrix} \int_{\mathcal{X}} g_1(x) dP_\theta(x) \\ \vdots \\ \int_{\mathcal{X}} g_k(x) dP_\theta(x) \end{pmatrix} = G(P_\theta)$$

Тогда по методу подстановки получим

$$\theta_n^* = m^{-1} \begin{pmatrix} \int_{\mathcal{X}} g_1(x) dP_n^*(x) \\ \vdots \\ \int_{\mathcal{X}} g_k(x) dP_n^*(x) \end{pmatrix} = G(P_n^*)$$

Таким образом, метод моментов – это частный случай метода подстановки.

Теорема 7.1. *Сильная состоятельной оценки методом моментов.*

Если m биективна и функцию m^{-1} можно доопределить до функции, заданной на всём $\mathbb{R}^k$ и непрерывной в каждой точке множества $m(\Theta)$ тогда оценка по методу моментов является сильно состоятельной оценкой параметра θ.

Доказательство. Фиксируем θ, по УЗБЧ знаем, что

$$\overline{g} \overset{P_\theta \text{ п.н.}}{\to} m(\theta)$$

Используя теорему о наследовании сходимости, навесим m^{-1}:

$$\theta_n^* = m^{-1}(\overline{g}) \overset{P_\theta \text{ п.н.}}{\to} m^{-1}(m(\theta)) = \theta$$

□

Теорема 7.2. *Асимптотическая нормальность ОММ.*

Если в условиях предыдущей теоремы m^{-1} дифференцируема на $m(\Theta)$ и $\forall i \leqslant k$: $\mathbb{E}_\theta g_i^2(X_1) < +\infty$. Тогда ОММ θ_n^* является асимптотически нормальной оценкой параметра θ.

Доказательство. По ЦПТ:

$$\sqrt{n}(\overline{g} - m(\theta)) \xrightarrow{d_\theta} \mathcal{N}(0, \Sigma)$$

Применяем многомерный дельта-метод и получаем требуемое. □

8 Квантили и выборочные квантили

Определение 8.1. Пусть P – распределение вероятности на $\mathbb{R}$. Пусть $p \in (0,1)$. **p-квантилью** распределения P называют

$$z_p = \inf\{x \in \mathbb{R} \mid F(x) \geqslant p\}$$

Определение 8.2. Пусть $X_1, \cdots, X_n$ – выборка, статистика

$$z_{n,p} = \begin{cases} X_{(\lceil np \rceil)}, np \notin \mathbb{Z} \\ X_{(np)}, np \in \mathbb{Z} \end{cases}$$

называется **выборочной p-квантилью**.

Теорема 8.1. *О выборочной квантили.*

Пусть $X_1, \cdots, X_n$ – выборка из распределения P с плотностью $f(x)$. Пусть z_p – это p-квантиль распределения P, причём $f(x)$ непрерывно дифференцируема в окрестности z_p, причём $f(z_p) > 0$. Тогда

$$\sqrt{n}(z_{n,p} - z_p) \xrightarrow{d} \mathcal{N}\left(0, \frac{p(1-p)}{f^2(z_p)}\right)$$

Доказательство. Пусть $k := \lceil np \rceil$.

Из соображений комбинаторики, заметим, что

$$P(X_{(k)} \leqslant x) = \sum_{m=k}^{n} C_n^m F^m(x)(1-F(x))^{n-m}$$

Засчёт свойств биномиальных коэффициентов, после дифференцирования выражения выше, получим

$$p_{X_{(k)}}(x) = nC_{n-1}^{k-1} F^{k-1}(x)(1-F(x))^{n-k} f(x)$$

Введём

$$\eta_n = (z_{n,p} - z_p)\sqrt{\frac{nf^2(z_p)}{p(1-p)}}$$

Плотность такого линейного преобразования легко считается

$$p_{\eta_n}(x) = \sqrt{\frac{p(1-p)}{nf^2(z_p)}} p_{X_{(k)}}(t_n(x))$$

где $t_n(x) = z_p + \frac{x}{f(z_p)}\sqrt{\frac{p(1-p)}{n}}$

Откуда это взялось? Вспомним, как меняется плотность при линейном преобразовании:

$$p_{a\xi+b}(x) = F'_{a\xi+b}(x) = P'(a\xi + b \leqslant x) = P'(\xi \leqslant \frac{x-b}{a}) = F'_{\xi}(\frac{x-b}{a}) = \frac{1}{a}p_{\xi}(\frac{x-b}{a})$$

Раскроем $p_{X_{(k)}}$ по формуле, которую получили в начале доказательства и разложим полученную плотность η_n в следующее произведение:

$$p_{\eta_n}(x) = A_1(n)A_2(n)A_3(n)$$

где

$$A_1(n) = \sqrt{npq}C_{n-1}^{k-1}p^{k-1}q^{n-k}$$

$$A_2(n) = \frac{f(t_n(x))}{f(z_p)}$$

$$A_3(n) = \left(\frac{F(t_n(x))}{p}\right)^{k-1}\left(\frac{1-F(t_n(x))}{q}\right)^{n-k}$$

Осталось заметить, что

$$A_1(n) \to \frac{1}{\sqrt{2\pi}}; \quad A_2(n) \to 1;$$

Для $A_3(n)$ немного сложнее, разложим $F(t_n(x))$ в ряд Тейлора в окрестности z_p. (так как $t_n(x) \to z_p$):

$$F(t_n(x)) = F(z_p) + (t_n - z_p)F'(z_p) + \frac{1}{2}(t_n - z_p)^2F''(z_p) + o(t_n - z_p)^2$$

Давайте упростим это выражение, раскрыв t_n и применив свойство квантиля $F(z_p) = p$:

$$F(t_n(x)) = p + x\sqrt{\frac{pq}{n}} + \frac{1}{2}\frac{x^2pq}{n} \cdot \frac{f'(z_p)}{f^2(z_p)} + o(\frac{1}{n}), n \to +\infty$$

Теперь должны расписать приближение $\ln\left(\frac{F(t_n(x))}{p}\right)$, используя формулу $\ln(1+x) = x - \frac{x^2}{2} + o(x^3)$, причём в квадрате нам нужен будет только $x\sqrt{\frac{pq}{n}}$:

$$\ln\left(\frac{F(t_n(x))}{p}\right) = x\sqrt{\frac{q}{pn}} + \frac{1}{2}\frac{x^2q}{n}\frac{f'(z_p)}{f^2(z_p)} + o\left(\frac{1}{n}\right) - \frac{x^2}{2}\frac{q}{np}$$

Аналогично разложив для $\ln\left(\frac{1-F(t_n(x))}{q}\right)$, получим

$$\ln A_3(n) \to -\frac{x^2}{2}$$

Таким образом, $p_{\eta_n(x)} \to \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ и эта сходимость равномерна на $\forall[-N, N]$.

Используя теорему из теории вероятностей,

$$\eta_n \xrightarrow{d} \mathcal{N}(0,1)$$

$\square$

Определение 8.3. Медианой распределения P называется $\frac{1}{2}$ квантиль.

Выборочной медианой называется

$$\hat{\mu} = \begin{cases} X_{(k)}, n = 2k+1 \\ \frac{X_{(k)} + X_{(k+1)}}{2}, n = 2k \end{cases}$$

Теорема 8.2. *О выборочной медиане.*

В условиях теоремы о выборочной квантили:

$$\sqrt{n}(\hat{\mu} - z_{\frac{1}{2}}) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{4f^2(z_{\frac{1}{2}})}\right)$$