Решения задач по математической статистике

КОНСТАНТИН ЗЮБИН

Задача 1аб

Пусть $X_1, \ldots, X_n \sim \mathcal{R}[0,\theta]$ — независимые случайные величины. Докажем, что оценки $\sqrt[k]{(k+1)\overline{X^k}}$ асимптотически нормальны для θ.

Прежде проверим, что оценка (последовательность оценок) $(k+1)\overline{X^k}$ асимптотически нормальна для θ и вычислим её асимптотическую дисперсию. Согласно вычислениям в задаче 4 первой домашней работы $\mathrm{E}_\theta(k+1)X_1^k = \theta^k$ и $\mathrm{D}_\theta(k+1)X_1^k = \frac{\theta^{2k}k^2}{2k+1}$. Тогда по центральной предельной теореме

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(k+1)X_i^k - \theta^k\right) \xrightarrow[n\to+\infty]{d} \eta \sim \mathcal{N}(0, \tfrac{\theta^{2k}k^2}{2k+1}).$$

Пусть $f(x) = \sqrt[k]{x}$. Тогда $f'(x) = \frac{1}{k\sqrt[k]{x^{k-1}}}$ не равно 0 при $\theta > 0$. По лемме об асимптотической нормальности для функции $f(x)$ получаем

$$\sqrt{n}\left(\sqrt[k]{\frac{1}{n}\sum_{i=1}^{n}(k+1)X_i^k} - \theta\right) \xrightarrow[n\to+\infty]{d} \tilde{\eta} \sim \mathcal{N}(0, \tfrac{\theta^{2k}k^2}{2k+1}\cdot(\tfrac{1}{k\theta^{k-1}})^2).$$

Асимптотическая дисперсия равна

$$\tfrac{\theta^{2k}k^2}{2k+1}\cdot(\tfrac{1}{k\theta^{k-1}})^2 = \tfrac{\theta^2}{2k+1}.$$

Поскольку для $m > k$ выполнено неравенство $\frac{\theta^2}{2m+1} < \frac{\theta^2}{2k+1}$, то среди оценок такого вида нет оценки с наименьшей (при фиксированном θ) асимптотической дисперсией.

Задача 2а

Пусть $X_1, \ldots, X_n \sim \mathcal{R}[0,\theta]$ — независимые случайные величины. Проверим, что оценка $X_{(n)} = \max(X_1, \ldots, X_n)$ состоятельна для θ. Согласно вычислениям, выполненным ранее, имеем формулу

$$F_{X_{(n)},\theta}(x) = \begin{cases} 0, & x < 0; \\ \frac{x^n}{\theta^n}, & x \in [0,\theta]; \\ 1, & x > \theta. \end{cases}$$

Поэтому для малых $\varepsilon > 0$ имеем

$$\mathrm{P}_\theta(|X_{(n)} - \theta| > \varepsilon) = 1 - F_{X_{(n)}}(\theta+\varepsilon) + F_{X_{(n)}}(\theta-\varepsilon) = 1 - 1 + \tfrac{(\theta-\varepsilon)^n}{\theta^n} = \left(1 - \tfrac{\varepsilon}{\theta}\right)^n.$$

Последнее выражение стремится к 0 при $n \to +\infty$, так как $|1 - \frac{\varepsilon}{\theta}| < 1$. Следовательно, последовательность $X_{(n)}$ сходится по вероятности к θ и, по определению, состоятельна для θ.

Задача 2б

Найдём функцию распределения величины $n(\theta - X_{(n)})$:

$$F_{n(\theta-X_{(n)}),\theta}(x) = \mathrm{P}_\theta(n(\theta - X_{(n)}) \leqslant x) = \mathrm{P}_\theta(X_{(n)} \geqslant \theta - \tfrac{x}{n}) = 1 - F_{X_{(n)},\theta}(\theta - \tfrac{x}{n}) =$$

$$= \begin{cases} 0, & x < 0; \\ 1 - (1 - \frac{x}{n\theta})^n, & x \in [0, n\theta]; \\ 1, & x > n\theta. \end{cases}$$

При $n \to +\infty$ данная последовательность функций распределений сходится к функции экспоненциального распределению

$$F(x) = \begin{cases} 0, & x < 0; \\ 1 - e^{-\frac{x}{\theta}}, & x \geqslant 0. \end{cases}$$

Таким образом, $n(\theta - X_{(n)}) \xrightarrow[n\to+\infty]{d} \xi \sim \mathrm{E}(\frac{1}{\theta})$.

Задача 3а

Пусть независимые случайные величины $X_1, \ldots, X_n \sim \mathcal{N}(\theta, 1)$ имеют стандартное нормальное распределение. Случайная величина $S_n = \sum\limits_{i=1}^{n} X_i$, как свёртка n независимых нормально распределённых случайных величин, имеет распределение $\mathcal{N}(n\theta, n)$. Тогда случайная величина $\overline{X} = \frac{1}{n} S_n$ имеет распределение $\mathcal{N}(\theta, \frac{1}{n})$.

Задача 3б

Докажем, что при $\theta \neq 0$ последовательность случайных величин $\overline{X}^2$ является асимптотически нормальной для θ^2.

По центральной предельной теореме имеем сходимость

$$\sqrt{n}(\overline{X} - \theta) \xrightarrow[n\to+\infty]{d} \eta \sim \mathcal{N}(0, 1).$$

Положим $f(x) = x^2$. При $\theta \neq 0$ производная $f'(x) = 2x$ не равна 0, поэтому по лемме об асимптотической нормальности имеем

$$\sqrt{n}(\overline{X}^2 - \theta^2) \xrightarrow[n\to+\infty]{d} \tilde{\eta} \sim \mathcal{N}(0, (2\theta)^2).$$

Задача 3в

В случае $\theta = 0$ рассмотрим функцию распределения случайной величины $\sqrt{n}\overline{X}^2$ при $n \to +\infty$:

$$F_{\sqrt{n}\overline{X}^2}(x) = \iint\limits_{x_1 x_2 \leqslant \frac{x}{\sqrt{n}}} \frac{n}{2\pi} e^{-\frac{n(x_1^2 + x_2^2)}{2}} dx_1 dx_2 = \iint\limits_{u_1 u_2 \leqslant \sqrt{n} x} \frac{1}{2\pi} e^{-\frac{u_1^2 + u_2^2}{2}} du_1 du_2.$$

Для $x > 0$ имеем

$$\iint\limits_{u_1 u_2 \leqslant \sqrt{n} x, u_1^2 + u_2^2 \leqslant n} \frac{1}{2\pi} e^{-\frac{u_1^2 + u_2^2}{2}} du_1 du_2 \leqslant \iint\limits_{u_1 u_2 \leqslant \sqrt{n} x} \frac{1}{2\pi} e^{-\frac{u_1^2 + u_2^2}{2}} du_1 du_2 \leqslant 1.$$

Поскольку интеграл сходится абсолютно на всей плоскости, то при $x > 0$ выражение в левой части неравенства при $n \to +\infty$ стремится к 1, поэтому при $x > 0$ выполнено

$$\lim_{n\to+\infty} F_{\sqrt{n}\overline{X}^2}(x) = 1.$$

При $x < 0$ имеем

$$0 \leqslant \iint\limits_{u_1 u_2 \leqslant \sqrt{n} x} \tfrac{1}{2\pi} e^{-\frac{u_1^2+u_2^2}{2}} du_1 du_2 = \iint\limits_{-u_1 u_2 \geqslant -\sqrt{n} x} \tfrac{1}{2\pi} e^{-\frac{u_1^2+u_2^2}{2}} du_1 du_2 \leqslant \iint\limits_{u_1^2+u_2^2 \geqslant -\sqrt{n} x} \tfrac{1}{2\pi} e^{-\frac{u_1^2+u_2^2}{2}} du_1 du_2.$$

Правая часть двойного неравенства стремится к 0 при $n \to +\infty$, поэтому для $x < 0$ имеем

$$\lim_{n \to +\infty} F_{\sqrt{n}\overline{X}^2}(x) = 0.$$

Таким образом, функции распределения $F_{\sqrt{n}\overline{X}^2}$ сходятся к функции распределения тождественно нулевой случайной величины и $\sqrt{n}\overline{X}^2 \underset{n\to+\infty}{\xrightarrow{d}} 0$.

Задача 4а

Пусть даны независимые случайные величины $X_1, \ldots, X_n \sim \mathrm{Be}(\theta)$, где $\theta \in (0, 1)$. Поскольку $\mathrm{E}\, X_i = \theta$ и $\mathrm{D}\, X_i = \theta(1-\theta)$, то по центральной предельной теореме имеем

$$\sqrt{n}(\overline{X} - \theta) \underset{n\to+\infty}{\xrightarrow{d}} \eta \sim \mathcal{N}(0, \theta(1-\theta)).$$

Положим $f(x) = e^x$. Тогда производная $f'(x) = e^x$ нигде не равна 0 и по лемме об асимптотической нормальности

$$\sqrt{n}(e^{\overline{X}} - e^{\theta}) \underset{n\to+\infty}{\xrightarrow{d}} \tilde{\eta} \sim \mathcal{N}(0, \theta(1-\theta) \cdot e^{2\theta}).$$

Асимптотическая дисперсия равна $\theta(1-\theta)e^{4\theta^2}$.

Задача 4б

Пусть даны независимые случайные величины $X_1, \ldots, X_n \sim \mathrm{Poiss}(\theta)$, где $\theta > 0$. Поскольку $\mathrm{E}\, X_i = \theta$ и $\mathrm{D}\, X_i = \theta$, то по центральной предельной теореме имеем

$$\sqrt{n}(\overline{X} - \theta) \underset{n\to+\infty}{\xrightarrow{d}} \eta \sim \mathcal{N}(0, \theta).$$

Положим $f(x) = x^3$. Тогда производная $f'(x) = 3x^2$ не равна 0 при $\theta > 0$. По лемме об асимптотической нормальности

$$\sqrt{n}(\overline{X}^3 - \theta^3) \underset{n\to+\infty}{\xrightarrow{d}} \tilde{\eta} \sim \mathcal{N}(0, \theta \cdot (3\theta^2)^2).$$

Асимптотическая дисперсия равна $9\theta^5$.

Задача 4в

Пусть даны независимые случайные величины $X_1, \ldots, X_n \sim \mathrm{Geom}(\theta)$, где $\theta \in (0, 1)$. Поскольку $\mathrm{E}\, X_i = \frac{1}{\theta}$ и $\mathrm{D}\, X_i = \frac{1-\theta}{\theta^2}$, то по центральной предельной теореме имеем

$$\sqrt{n}(\overline{X} - \tfrac{1}{\theta}) \underset{n\to+\infty}{\xrightarrow{d}} \eta \sim \mathcal{N}(0, \tfrac{1-\theta}{\theta^2}).$$

Положим $f(x) = \frac{1}{x^2}$. Тогда производная $f'(x) = \frac{-2}{x^3}$ не равна 0 при $\theta \in (0, 1)$. По лемме об асимптотической нормальности

$$\sqrt{n}(\overline{X}^{-2} - \theta^2) \underset{n\to+\infty}{\xrightarrow{d}} \tilde{\eta} \sim \mathcal{N}(0, \tfrac{1-\theta}{\theta^2} \cdot (-2\theta^3)^2).$$

Асимптотическая дисперсия равна $4(1-\theta)\theta^4$.