Примеры возможных задач к семинару по лекции 4

Задача 1. *Пусть* $(X_1, \ldots, X_n)$ — *выборка, порожденная случайной величиной* X *с конечной дисперсией* $\mathrm{D}\{X\}$. *Доказать, что оценка*

$$S_n^{\textit{испр}} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2,$$

где $\overline{X}_n = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$, *является состоятельной оценкой для дисперсии* $\mathrm{D}\{X\}$.

Доказательство. Воспользуемся представлением исправленной выборочной дисперсии через выборочные начальные моменты:

$$S_n^{\text{испр}} = \frac{n}{n-1}\left(\overline{X^2} - (\overline{X}_n)^2\right), \quad \text{где } \overline{X^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

По теореме 3 из лекции 3 (состоятельность выборочных начальных моментов) имеем:

$$\overline{X^2} \xrightarrow[n\to\infty]{P} \mathrm{M}\big\{X^2\big\}, \qquad \overline{X}_n \xrightarrow[n\to\infty]{P} \mathrm{M}\{X\}.$$

Функция $f(u,v) = u - v^2$ непрерывна, поэтому по теореме Слуцкого (лекция 3)

$$\overline{X^2} - (\overline{X}_n)^2 \xrightarrow{P} \mathrm{M}\big\{X^2\big\} - (\mathrm{M}\{X\})^2 = \mathrm{D}\{X\}.$$

Множитель $\frac{n}{n-1} \to 1$, следовательно, по свойствам сходимости по вероятности

$$S_n^{\text{испр}} \xrightarrow{P} \mathrm{D}\{X\},$$

что и означает состоятельность оценки. □

Задача 2. *Пусть выборка* $X_1, \ldots, X_n$ *подчиняется показательному распределению с параметром* $\lambda > 0$ *(т.е. плотность* $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$*). Показать, что оценка*

$$\widehat{\lambda}_n = \sqrt{\frac{2n}{\sum\limits_{i=1}^{n} X_i^2}}$$

является состоятельной оценкой для параметра λ.

Доказательство. Рассмотрим статистику $T_n = \frac{1}{2n}\sum_{i=1}^n X_i^2$. Для показательного распределения известны моменты:

$$\mathrm{M}\{X_i^2\} = \frac{2}{\lambda^2}, \qquad \mathrm{D}\{X_i^2\} = \frac{20}{\lambda^4}$$

(поскольку $\mathrm{M}\{X_i^4\} = 24/\lambda^4$). Тогда

$$\mathrm{M}\{T_n\} = \frac{1}{2n}\cdot n \cdot \frac{2}{\lambda^2} = \frac{1}{\lambda^2}, \qquad \mathrm{D}\{T_n\} = \frac{1}{4n^2}\cdot n\cdot\frac{20}{\lambda^4} = \frac{5}{n\lambda^4} \xrightarrow[n\to\infty]{} 0.$$

По достаточным условиям состоятельности (теорема 1 лекции 3) T_n является состоятельной оценкой для $1/\lambda^2$. Функция $g(t) = 1/\sqrt{t}$ непрерывна при $t > 0$, поэтому по теореме Слуцкого

$$\widehat{\lambda}_n = \frac{1}{\sqrt{T_n}} \xrightarrow{P} \frac{1}{\sqrt{1/\lambda^2}} = \lambda.$$

Альтернативно, можно сразу применить закон больших чисел к $\frac{1}{n}\sum X_i^2$. □

Задача 3. *Выборка $X_1, \ldots, X_n$ порождена случайной величиной X, имеющей равномерное распределение на отрезке $[0, \theta]$, $\theta > 0$. Показать, что оценка*

$$\widehat{\theta} = 2\overline{X}_n, \quad \text{где } \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i,$$

является несмещенной и состоятельной оценкой для θ.

Доказательство. Для равномерного распределения $U[0,\theta]$ имеем $\mathrm{M}\{X\} = \theta/2$. Тогда

$$\mathrm{M}\left\{\widehat{\theta}\right\} = 2\mathrm{M}\{\overline{X}_n\} = 2\cdot\frac{\theta}{2} = \theta,$$

что доказывает несмещённость (см. лекцию 4, утверждение 1).

Далее, дисперсия $\mathrm{D}\{X\} = \theta^2/12$, поэтому

$$\mathrm{D}\left\{\widehat{\theta}\right\} = 4\mathrm{D}\{\overline{X}_n\} = 4\cdot\frac{\mathrm{D}\{X\}}{n} = \frac{4}{n}\cdot\frac{\theta^2}{12} = \frac{\theta^2}{3n} \xrightarrow[n\to\infty]{} 0.$$

Так как оценка несмещена и её дисперсия стремится к нулю, по теореме 1 лекции 3 она состоятельна. Можно также заметить, что $\overline{X}_n \xrightarrow{P} \theta/2$ по закону больших чисел, и тогда $\widehat{\theta} = 2\overline{X}_n \xrightarrow{P} \theta$ по теореме Слуцкого. □

Задача 4. *Пусть выборка $X_1, \dots, X_n$ порождена случайной величиной X, имеющей равномерное распределение на отрезке $[a, b]$, где a и b неизвестны. Рассмотрим оценки*

$$a_n = X_{(1)}, \qquad b_n = X_{(n)},$$

где $X_{(1)}$ и $X_{(n)}$ — соответствующие порядковые статистики (минимальный и максимальный элементы выборки). Найти $\mathrm{M}\{a_n\}$, $\mathrm{D}\{a_n\}$, $\mathrm{M}\{b_n\}$, $\mathrm{D}\{b_n\}$ и показать, что оценки a_n и b_n являются асимптотически несмещенными и состоятельными.

Доказательство. Введём нормированную величину $Y = \frac{X-a}{b-a}$. Тогда $Y \sim U[0, 1]$, и $X = a + (b - a)Y$. Для порядковых статистик:

$$X_{(1)} = a + (b-a)Y_{(1)}, \qquad X_{(n)} = a + (b-a)Y_{(n)},$$

где $Y_{(1)}$ и $Y_{(n)}$ — соответствующие порядковые статистики из $U[0, 1]$. Их распределения известны (см. пример в лекции 3):

$$f_{Y_{(1)}}(y) = n(1-y)^{n-1},\ 0 \le y \le 1, \qquad f_{Y_{(n)}}(y) = ny^{n-1},\ 0 \le y \le 1.$$

Вычислим моменты $Y_{(1)}$ и $Y_{(n)}$:

$$\begin{aligned}
\mathrm{M}\{Y_{(1)}\} &= \int_0^1 y\, n(1-y)^{n-1} dy = \frac{1}{n+1}, \\
\mathrm{M}\{Y_{(1)}^2\} &= \int_0^1 y^2 n(1-y)^{n-1} dy = \frac{2}{(n+1)(n+2)}, \\
\mathrm{D}\{Y_{(1)}\} &= \frac{2}{(n+1)(n+2)} - \frac{1}{(n+1)^2} = \frac{n}{(n+1)^2(n+2)}, \\
\mathrm{M}\{Y_{(n)}\} &= \int_0^1 y\, ny^{n-1} dy = \frac{n}{n+1}, \\
\mathrm{M}\{Y_{(n)}^2\} &= \int_0^1 y^2 ny^{n-1} dy = \frac{n}{n+2}, \\
\mathrm{D}\{Y_{(n)}\} &= \frac{n}{n+2} - \frac{n^2}{(n+1)^2} = \frac{n}{(n+1)^2(n+2)}.
\end{aligned}$$

Тогда для исходных оценок:

$$\begin{aligned}
&\mathrm{M}\{a_n\} = a + (b-a)\mathrm{M}\big\{Y_{(1)}\big\} = a + \frac{b-a}{n+1} = \frac{an+b}{n+1},\\
&\mathrm{D}\{a_n\} = (b-a)^2\mathrm{D}\big\{Y_{(1)}\big\} = \frac{(b-a)^2 n}{(n+1)^2(n+2)},\\
&\mathrm{M}\{b_n\} = a + (b-a)\mathrm{M}\big\{Y_{(n)}\big\} = a + \frac{(b-a)n}{n+1} = \frac{a+bn}{n+1},\\
&\mathrm{D}\{b_n\} = (b-a)^2\mathrm{D}\big\{Y_{(n)}\big\} = \frac{(b-a)^2 n}{(n+1)^2(n+2)}.
\end{aligned}$$

При $n \to \infty$ получаем

$$\mathrm{M}\{a_n\} \to a, \qquad \mathrm{M}\{b_n\} \to b, \qquad \mathrm{D}\{a_n\} \to 0, \qquad \mathrm{D}\{b_n\} \to 0.$$

Таким образом, оценки являются асимптотически несмещёнными, и по теореме 1 лекции 3 (достаточные условия состоятельности) они состоятельны. Можно также доказать состоятельность непосредственно: для любого $\varepsilon > 0$

$$P(|a_n - a| \geq \varepsilon) = P(X_{(1)} \geq a + \varepsilon) = \left(1 - \frac{\varepsilon}{b-a}\right)^n \xrightarrow[n\to\infty]{} 0,$$

и аналогично для b_n. □

Справочный материал из лекций

Лекция 3

Определение 1. Оценка $\hat{\theta}_n$ называется *несмещённой*, если $\mathrm{M}\Big\{\hat{\theta}_n\Big\} = \theta$. Если $\lim\limits_{n\to\infty} \mathrm{M}\Big\{\hat{\theta}_n\Big\} = \theta$, то оценка называется *асимптотически несмещённой.*

Определение 2. Оценка $\hat{\theta}_n$ параметра θ называется *состоятельной*, если

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} P\{|\hat{\theta}_n - \theta| < \varepsilon\} = 1,$$

т.е. $\hat{\theta}_n \xrightarrow{P} \theta$ при $n \to \infty$.

Теорема 1 (Достаточные условия состоятельности)**.** *Пусть для оценки $\hat{\theta}_n$ параметра θ выполнены условия:*

1. $\lim\limits_{n\to\infty} \mathrm{M}\Big\{\hat{\theta}_n\Big\} = \theta$ *(асимптотическая несмещённость);*

2. $\lim\limits_{n\to\infty} \mathrm{D}\Big\{\hat{\theta}_n\Big\} = 0$*;*

тогда $\hat{\theta}_n$ *— состоятельная оценка для* θ*.*

Теорема 2 (Состоятельность выборочных начальных моментов)**.** *Если существует теоретический начальный момент* ν_{2m}*, то выборочный начальный момент*

$$\bar{\nu}_k(n) = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

является состоятельной оценкой для ν_k *при любом* $k \le m$*.*

Теорема (Слуцкого)**.** *Пусть имеется* k *последовательностей случайных величин:* $Y_n^{(1)} \xrightarrow{P} a_1, \ \ldots, \ Y_n^{(k)} \xrightarrow{P} a_k$*. Пусть функция* $f(y_1,\ldots,y_k)$ *непрерывна в точке* $(a_1,\ldots,a_k)$*. Тогда*

$$Z_n = f\big(Y_n^{(1)},\ldots,Y_n^{(k)}\big) \xrightarrow{P} f(a_1,\ldots,a_k).$$

Пример 1. Порядковые статистики для равномерного распределения *Пусть* $X \sim U[0,\theta]$*,* $\theta > 0$*, и* $(X_1,\ldots,X_n)$ *— выборка. Обозначим* $X_{(n)} = \max\{X_1,\ldots,X_n\}$*. Тогда функция распределения* $X_{(n)}$ *равна*

$$F_{(n)}(x) = \left(\frac{x}{\theta}\right)^n, \quad 0 \le x \le \theta,$$

а плотность

$$f_{(n)}(x) = n\left(\frac{x}{\theta}\right)^{n-1}\cdot\frac{1}{\theta}, \quad 0 \le x \le \theta.$$

Математическое ожидание:

$$\mathrm{M}\big\{X_{(n)}\big\} = \int_0^\theta x \cdot n\frac{x^{n-1}}{\theta^n}\,dx = \frac{n}{n+1}\,\theta.$$

Аналогично для $X_{(1)} = \min\{X_1,\ldots,X_n\}$*:*

$$F_{(1)}(x) = 1 - \left(1-\frac{x}{\theta}\right)^n, \quad f_{(1)}(x) = n\left(1-\frac{x}{\theta}\right)^{n-1}\frac{1}{\theta},$$

$$\mathrm{M}\big\{X_{(1)}\big\} = \frac{\theta}{n+1}.$$

Для нормированной величины $Y = X/\theta \sim U[0,1]$ *получаем:*

$$\mathrm{M}\big\{Y_{(1)}\big\} = \frac{1}{n+1}, \quad \mathrm{M}\big\{Y_{(n)}\big\} = \frac{n}{n+1},$$

$$\mathrm{D}\big\{Y_{(1)}\big\} = \mathrm{D}\big\{Y_{(n)}\big\} = \frac{n}{(n+1)^2(n+2)}.$$

Лекция 4

Задача 5 (Свойства выборочного среднего)**.** *Для выборки* $(X_1, \ldots, X_n)$ *из распределения с конечным вторым моментом:*

- $\overline{X}_n$ *— несмещённая оценка для* $\mathrm{M}\{X\}$*:* $\mathrm{M}\{\overline{X}_n\} = \mathrm{M}\{X\}$.
- $\overline{X}_n$ *— состоятельная оценка для* $\mathrm{M}\{X\}$*.*
- $\mathrm{D}\{\overline{X}_n\} = \dfrac{\mathrm{D}\{X\}}{n}$.