

Математична статистика. Точкові оцінки.
Незміщеність, слушність, та ефективність оцінок. Нерівність
Крамера-Рао

Оцінка $\hat{\theta} = \hat{\theta}_n = \hat{\theta}(X_1, \dots, X_n)$ називається консистентною (КО) (слушною; рос. – состоятельной), якщо $\hat{\theta}_n \xrightarrow{P} \theta$, $n \rightarrow \infty$, тобто, якщо $\forall \varepsilon > 0 \quad P\left\{\left|\hat{\theta}_n - \theta\right| < \varepsilon\right\} \xrightarrow{n \rightarrow \infty} 1$, де $X_1, \dots, X_n$ – незалежні спостереження, θ – справжнє значення параметра розподілу, $\hat{\theta}$ – оцінка невідомого значення θ .

Теорема 7.1. $M\hat{\theta}_n \xrightarrow{n \rightarrow \infty} \theta$, $D\hat{\theta}_n \xrightarrow{n \rightarrow \infty} 0 \Rightarrow \hat{\theta}_n$ – консистентна оцінка.

$\hat{\theta}_n$ – незміщена оцінка (незсунена – НО), якщо $M\hat{\theta}_n = \theta$, інакше $\hat{\theta}_n$ називається зміщеною, а $\Delta = M\hat{\theta}_n - \theta$ називається зсувом. Очевидно, що $(\hat{\theta}_n - \Delta)$ буде незміщеною оцінкою.

Теорема 7.2. $M_\theta \left| \frac{\partial}{\partial \theta} \ln L(X, \theta) \right| < \infty$, $M_\theta \left| \frac{\partial^2}{\partial \theta^2} \ln L(X, \theta) \right| < \infty$, $M_\theta \left| \frac{\partial}{\partial \theta} \ln L(X, \theta) \right|^2 < \infty$, $\hat{\theta}_n$ – незміщена оцінка для θ , $D\hat{\theta}_n < \infty$, $\forall \theta \int_{R^n} \left| \hat{\theta}_n \frac{\partial}{\partial \theta} L(t, \theta) \right| dt < \infty \Rightarrow$

$$\Rightarrow \boxed{D\hat{\theta}_n \geq \frac{1}{i_n(\theta)}} \quad (\text{нерівність Крамера-Рао}).$$

Тут $i_n(\theta) = M_\theta \left(\frac{\partial}{\partial \theta} \ln L(X, \theta) \right)^2$ – кількість інформації за Фішером, що міститься у вибірці $X = (X_1, \dots, X_n)$ про параметр θ , де $L(t, \theta)$ – щільність вибіркового вектора $X = (X_1, \dots, X_n)$ (у неперервному випадку), або $L(x, \theta) = P\{X = x\}$ (у дискретному випадку) (X – випадковий вектор, x – вибірка, яку спостерігаємо).

Якщо в Теоремі 7.2 досягається рівність, то $\hat{\theta}_n$ називається ефективною оцінкою (ЕО).

Якщо ж $\lim_{n \rightarrow \infty} \frac{1}{i_n(\theta) D\hat{\theta}_n} = 1$, то $\hat{\theta}_n$ називається асимптотично ефективною оцінкою (АЕО).

Задача 7.1 $\hat{\theta}_n = \bar{m} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ – оцінка для середнього m у генеральній сукупності із скінченною дисперсією σ^2 . Довести: а) консистентність, б) незміщеність оцінки.

Задача 7.2 $(X_1, \dots, X_n)$ – незалежна вибірка з генеральної сукупності із скінченим математичним сподіванням m та невідомою дисперсією σ^2 .

а) Довести, що вибіркова дисперсія $\hat{D} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ є зміщеною оцінкою для невідомої дисперсії; б) знайти зсув цієї оцінки; в) показати, що $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ є незміщеною оцінкою для невідомої дисперсії.

Задача 7.3 $(X_1, \dots, X_n)$ – незалежна вибірка з генеральної сукупності із відомим математичним сподіванням m та невідомою дисперсією σ^2 . Довести, що вибіркова дисперсія $S_0^2 = \frac{1}{n} \sum_{i=1}^n (X_i - m)^2$ є незміщеною оцінкою для невідомої дисперсії.

Задача 7.4 Є дві вибірки об'ємів n_1 та n_2 з генеральної сукупності з середнім m та дисперсією σ^2 . Нехай $\bar{X}_1, \bar{X}_2, S_1^2, S_2^2$ – незміщені оцінки середніх та дисперсій в цих вибірках. Довести, що об'єднані оцінки, що мають вид

$$\bar{X} = \frac{n_1 \bar{X}_1 + n_2 \bar{X}_2}{n_1 + n_2} \quad \text{та} \quad S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

є а) незсуненими та б) консистентними оцінками для m та σ^2 відповідно.

Задача 7.5 $i_1(\theta) = M_\theta \left(\frac{\partial}{\partial \theta} \ln f(X_1, \theta) \right)^2$ – кількість інформації за Фішером, що міститься у одному спостереженні, $i_n(\theta) = M_\theta \left(\frac{\partial}{\partial \theta} \ln f(\bar{X}, \theta) \right)^2$ – кількість інформації за Фішером, що міститься у n спостереженнях. Якщо спостереження незалежні, то $i_n(\theta) = n i_1(\theta)$. Довести це.

Задача 7.6 Довести, що $i_1(\theta) = -M_\theta \left(\frac{\partial^2}{\partial \theta^2} \ln f(X_1, \theta) \right)$.

Задача 7.7 $X_1 \sim N(\theta, \sigma^2)$. Знайти $i_1(\theta)$.

Задача 7.8 $(X_1, \dots, X_n)$ – незалежна вибірка з генеральної сукупності з розподілом $N(m, \sigma^2)$.

Довести, що вибіркове середнє $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ є ефективною оцінкою для m .

Домашнє завдання № 7

- Довести, що оцінки $\hat{D} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ та $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ є консистентними оцінками для невідомої дисперсії.
- Довести, що вибіркове середнє, пораховане для вибірки з генеральної сукупності з розподілом Пуассона, є а) незміщеною, б) консистентною та в) ефективною оцінкою для невідомого математичного сподівання.
- $X_1 \sim N(m, \theta^2)$. Знайти $i_1(\theta)$ та $i_1(\theta^2)$.

Додаткові задачі № 7

- Довести, що $S_0^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2$ є консистентною (слухною) оцінкою невідомої дисперсії $D\xi = \sigma^2$ при відомому середньому m .
- Довести, що $S_0^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2$ є ефективною оцінкою невідомої дисперсії $D\xi = \sigma^2$ при відомому середньому m для гауссівського розподілу $N(m, \sigma^2)$.