Заняття №11

Вибіркові моменти

Означення 1. Нехай задано вибірку з генеральної сукупності $\xi_1, \xi_2, ..., \xi_n$ з функцією розподілу $F_\xi(x, \theta)$, де θ - невідомий параметр розподілу. Тоді **оцінкою параметра** $\hat{\theta}$ будемо називати функціонал від вибірки, який можна використовувати замість невідомого параметра

$$\hat{\theta} = g(\xi_1, \xi_2, ..., \xi_n).$$

Якщо замість вибірки $\xi_1, \xi_2, ..., \xi_n$ в цей функціонал підставити її реалізацію $x_1, x_2, ..., x_n$, то одержимо число $\hat{\theta} = g(x_1, x_2, ..., x_n)$, яке називається **реалізацією оцінки**.

Означення 2. Оцінка $\hat{\theta}$ називається **незміщеною**, якщо її математичне сподівання дорівнює істинному значенню параметра

$$M\hat{\theta} = Mg(\xi_1, \xi_2, ..., \xi_n) = \theta.$$

Означення 3. **Вибірковим моментом k-го порядку** A_k будемо називати наступну функцію від вибірки

$$A_k = \frac{\sum_{i=1}^{n} \xi_i^k}{n}.$$

Означення 4. **Вибірковим середнім** $\overline{\xi}$ будемо називати вибірковий момент першого порядку

$$\overline{\xi} = A_1 = \frac{\sum_{i=1}^{n} \xi_i}{n}.$$

При цьому реалізацією вибіркового середнього буде число

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

Твердження 1. Нехай задано вибірку $\xi_1, \xi_2, ..., \xi_n$ з генеральної сукупності, таку що $M\xi_i = a$. Тоді вибіркове середнє є незміщеною оцінкою математичного сподівання $M\overline{\xi} = a$.

Доведення.

Введемо позначення $M\xi_i = a, i = 1, 2, ..., n$. Тоді

$$M\overline{\xi} = M\frac{1}{n}\sum_{i=1}^{n}\xi_i = \frac{1}{n}\sum_{i=1}^{n}M\xi_i = \frac{na}{n} = a.$$

(Саме для таких дій потрібно розрізняти поняття вибірки і її реалізації, оскільки $M\overline{x} = \overline{x}$)

Означення 5. **Вибірковим центральним моментом** k **-го порядку** μ_k будемо називати наступний функціонал від вибірки

$$\mu_k = \frac{\sum_{i=1}^{n} (\xi_i - \overline{\xi})^k}{n}.$$

Означення 6. **Вибірковою дисперсією** S^2 будемо називати вибірковий центральний момент другого порядку

$$S^2 = \mu_2 = \frac{\sum_{i=1}^{n} (\xi_i - \overline{\xi})^2}{n}.$$

Слід зауважити, що S^2 є зміщеною оцінкою дисперсії, тому на практиці, в якості оцінки дисперсії частіше використовують наступні функціонали:

1) $\hat{S}^2 = \mu_2 = \dfrac{\sum_{i=1}^{n} (\xi_i - \overline{\xi})^2}{n-1}$;

2) $S^2 = \mu_2 = \dfrac{\sum_{i=1}^{n} (\xi_i - a)^2}{n}$, де $M\xi_i = a, i = 1, 2, ..., n$.

Ці два функціонали є незміщеними оцінками дисперсії.

У подальшому, будемо вважати, що у нас задано вибірку з генеральної сукупності з такими характеристиками $M\xi_i = a, D\xi_i = \sigma^2, i = 1, 2, ..., n$.

Задача 1. Нехай задано вибірку з генеральної сукупності $\xi_1, \xi_2, ..., \xi_n$, для якої відомо, що $D\xi_i = \sigma^2, i = 1, 2, ..., n$. Потрібно знайти $D\overline{\xi}$.

$$D\overline{\xi} = D\frac{1}{n}\sum_{i=1}^{n} \xi_i = \frac{1}{n^2}\sum_{i=1}^{n} D\xi_i = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Задача 2. Спростити вираз $\sum_{i=1}^{n} (\xi_i - a)(\overline{\xi} - a)$.

$$\sum_{i=1}^{n}(\xi_i - a)(\overline{\xi} - a) = (\overline{\xi} - a)\sum_{i=1}^{n}(\xi_i - a) = (\overline{\xi} - a)(n\overline{\xi} - na) = n(\overline{\xi} - a)^2.$$

Задача 3. Довести, що

$$\frac{1}{n}\sum_{i=1}^{n}\left(\xi_i - \overline{\xi}\right)^2 \le \frac{1}{n}\sum_{i=1}^{n}\left(\xi_i - a\right)^2, \; a \in \mathbb{R}.$$

$$\sum_{i=1}^{n}\left(\xi_i - \overline{\xi}\right)^2 = \sum_{i=1}^{n}\left((\xi_i - a) - (\overline{\xi} - a)\right)^2 = \sum_{i=1}^{n}\left(\xi_i - a\right)^2 - 2\sum_{i=1}^{n}\left(\xi_i - a\right)\left(\overline{\xi} - a\right) +$$

$$+\sum_{i=1}^{n}\left(\overline{\xi} - a\right)^2 = \sum_{i=1}^{n}\left(\xi_i - a\right)^2 - 2n\left(\overline{\xi} - a\right)^2 + n\left(\overline{\xi} - a\right)^2 =$$

$$= \sum_{i=1}^{n}\left(\xi_i - a\right)^2 - n\left(\overline{\xi} - a\right)^2 \le \sum_{i=1}^{n}\left(\xi_i - a\right)^2.$$

Ефективні оцінки. Нерівність Крамера-Рао.

Означення. Нехай M - клас усіх незміщених оцінок. Тоді оцінка $\hat{\theta}^*$ називається оптимальною, якщо $D\hat{\theta}^* = \min\limits_{\hat{\theta} \in M} D\hat{\theta}$.

Означення. Нехай $L(x,\theta)$ - щільність розподілу вибіркового вектора $(\xi_1, \xi_2, ..., \xi_n),\; x \in R^n$ (або ймовірність у дискретному випадку). Для вибірки з генеральної сукупності цю функцію можна подати у такому вигляді $L(x,\theta) = \prod\limits_{k=1}^{n} p(x_k, \theta)$, де $p(x_k, \theta)$ - щільність ξ_k у точці x_k (або ймовірність у дискретному випадку). Функцію $L(x,\theta)$, визначену таким чином, будемо називати **функцією вірогідності**.

Лема: Нехай функція $L(x,\theta)$ двічі диференційована по θ і

$$M\left|\frac{\partial}{\partial\theta}\ln L(\xi,\theta)\right| < \infty,\; M\left|\frac{\partial^2}{\partial\theta^2}\ln L(\xi,\theta)\right| < \infty,\; M\left|\frac{\partial}{\partial\theta}\ln L(\xi,\theta)\right|^2 < \infty.$$

Тоді $M\dfrac{\partial}{\partial\theta}\ln L(\xi,\theta) = 0, \quad M\left(\dfrac{\partial}{\partial\theta}\ln L(\xi,\theta)\right)^2 = -M\dfrac{\partial^2}{\partial\theta^2}\ln L(\xi,\theta)$.

Означення. Функцію $I(\theta) = M\left(\frac{\partial}{\partial\theta}\ln L(\xi,\theta)\right)^2 = -M\,\frac{\partial^2}{\partial\theta^2}\ln L(\xi,\theta)$ називають кількістю інформації по Фішеру.

Наслідки з нерівності Крамера-Рао:

1) Якщо оцінка $\hat{\theta}$ незміщена $M\hat{\theta} = \theta$, тоді $D\hat{\theta} \ge I^{-1}(\theta)$.

2) Якщо оцінка $\hat{\theta}$ зміщена $M\hat{\theta} = \theta + b(\theta)$, тоді $D\hat{\theta} \ge \frac{(1+b'(\theta))^2}{I(\theta)}$, де $b(\theta)$ - деякий зсув.

Означення. Якщо нижня межа в нерівностях 1) або 2) досягається, то відповідна оцінка називається **ефективною**.

Ефективна, незміщена оцінка є оптимальною.

Приклад 1.

Задано вибірку з генеральної сукупності з нормальним розподілом і відомою дисперсією

$$f(x,\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\theta)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

Показати, що $\hat{\theta} = \frac{\sum_{i=1}^{n}\xi_i}{n}$ буде ефективною оцінкою параметра θ.

1. Визначаємо, чи є оцінка незміщеною:

$M\hat{\theta} = M\,\frac{\sum_{i=1}^{n}\xi_i}{n} = \theta$. Отже ми будемо працювати за першим наслідком з нерівності Крамера-Рао.

2. Шукаємо дисперсію оцінки

$$D\hat{\theta} = D\frac{1}{n}\sum_{i=1}^{n}\xi_i = \frac{1}{n^2}\sum_{i=1}^{n}D\xi_i = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

3. Шукаємо кількість інформації по Фішеру

$$L(x,\theta) = \prod_{i=1}^{n} f(x_i,\theta) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^{n}(x_i-\theta)^2}{2\sigma^2}}.$$

Логарифмуємо $\ln L(x,\theta) = -\frac{n}{2}\ln\left(2\pi\sigma^2\right) - \frac{\sum_{i=1}^{n}(x_i-\theta)^2}{2\sigma^2}$.

$$\frac{\partial \ln L(x,\theta)}{\partial \theta} = \frac{\sum_{i=1}^{n} x_i}{\sigma^2} - \frac{n\theta}{\sigma^2}.$$

$$I(\theta) = -M\,\frac{\partial^2}{\partial \theta^2} \ln L(\xi,\theta) = \frac{n}{\sigma^2}.$$

4. Перевіряємо, чи досягається нижня межа в нерівності 1).

Бачимо, що $D\hat{\theta} = \left(I(\theta)\right)^{-1}$. Отже, оцінка є ефективною.

Приклад 2.

Нехай задано вибірку з генеральної сукупності $\xi_1, \xi_2, ..., \xi_n$ з розподілом Пуассона

$P\left(\xi = k\right) = \frac{\theta^k}{k!} e^{-\theta}, \theta > 0, k = 0,1,...$. Показати, що $\hat{\theta} = \frac{\sum_{i=1}^{n} \xi_i}{n}$ буде ефективною оцінкою параметра θ.

1. Визначаємо, чи є оцінка незміщеною: $M\hat{\theta} = M\,\frac{\sum_{i=1}^{n} \xi_i}{n} = \theta$.

2. Шукаємо дисперсію оцінки $D\hat{\theta} = D\frac{1}{n}\sum_{i=1}^{n} \xi_i = \frac{1}{n^2}\sum_{i=1}^{n} D\xi_i = \frac{n\theta}{n^2} = \frac{\theta}{n}$.

3. Шукаємо кількість інформації по Фішеру

$$L(x,\theta) = \prod_{k=1}^{n} P(\xi = x_k) = \prod_{k=1}^{n} \frac{\theta^{x_k}}{x_k!} e^{-\theta} = \frac{\theta^{\sum_{k=1}^{n} x_k}}{\prod_{k=1}^{n} x_k!} e^{-n\theta}.$$

$L(x,\theta) = \prod_{k=1}^{n} P(\xi = x_k)$ - цей запис стандартний для всіх дискретних розподілів.

$$\ln L(x,\theta) = \sum_{k=1}^{n} x_k \ln\theta - n\theta - \sum_{k=1}^{n} \ln x_k!$$

$$\frac{\partial}{\partial \theta} \ln L(x,\theta) = \frac{\sum_{k=1}^{n} x_k - n\theta}{\theta}.$$

Тепер шукаємо кількість інформації по Фішеру і не забуваємо перейти від x_k до ξ_k.

$$I(\theta) = -M\,\frac{\partial^2}{\partial \theta^2} \ln L(\xi,\theta) = M\,\frac{\sum_{k=1}^{n} \xi_k}{\theta^2} = \frac{n\theta}{\theta^2} = \frac{n}{\theta}.$$

4. Перевіряємо, чи досягається нижня межа в нерівності 1).

Бачимо, що $D\hat{\theta} = \left(I(\theta)\right)^{-1}$. Отже, оцінка є ефективною.

Приклад 3.

Нехай задано вибірку з генеральної сукупності $\xi_1, \xi_2, ..., \xi_n$ з показниковим розподілом

$$f(x,\theta) = \begin{cases} \theta e^{-\theta x}, \; x \ge 0, \\ 0, \; x < 0. \end{cases}$$

Перевірити чи буде $\hat{\theta} = \dfrac{n}{\sum\limits_{i=1}^{n} \xi_i}$ ефективною оцінкою параметра θ.

Допоміжна інформація:

а) Випадкова величина $\gamma_n = \sum\limits_{k=1}^{n} \xi_k$ має розподіл Ерланга, щільність якого має наступний вигляд $f_{\gamma_n}(x) = \dfrac{\theta^n x^{n-1}}{(n-1)!} e^{-\theta x}, \; x \ge 0.$

б) Гамма функція і її властивості.

$$\Gamma(\alpha) = \int\limits_0^\infty x^{\alpha-1} e^{-x} dx;$$

$$\Gamma(\alpha+1) = \alpha \Gamma(\alpha); \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

1. Визначаємо, чи є оцінка незміщеною:

$$M\hat{\theta} = M\frac{n}{\gamma_n} = \int\limits_0^\infty \frac{n}{x} \cdot \frac{\theta^n x^{n-1}}{(n-1)!} e^{-\theta x} dx = \begin{vmatrix} t = \theta x \\ dt = \theta dx \end{vmatrix} = \frac{n\theta}{(n-1)!} \int\limits_0^\infty t^{n-2} e^{-t} dt =$$

$$= \frac{n\theta}{(n-1)!} \Gamma(n-1) = \frac{n\theta (n-2)!}{(n-1)!} = \frac{n\theta}{n-1} = \theta + \frac{\theta}{n-1}.$$

Бачимо, що оцінка зміщена і $b(\theta) = \dfrac{\theta}{n-1}$.

2. Шукаємо дисперсію оцінки

$$M\hat{\theta}^2 = M\left(\frac{n}{\gamma_n}\right)^2 = \int\limits_0^\infty \frac{n^2}{x^2} \cdot \frac{\theta^n x^{n-1}}{(n-1)!} e^{-\theta x} dx = \begin{vmatrix} t = \theta x \\ dt = \theta dx \end{vmatrix} = \frac{n^2\theta^2}{(n-1)!} \int\limits_0^\infty t^{n-3} e^{-t} dt =$$

$$= \frac{n^2\theta^2}{(n-1)!} \Gamma(n-2) = \frac{n^2\theta^2 (n-3)!}{(n-1)!} = \frac{n^2\theta^2}{(n-1)(n-2)};$$

$$D\hat{\theta} = M\hat{\theta}^2 - \left(M\hat{\theta}\right)^2 = \frac{n^2\theta^2}{(n-1)(n-2)} - \frac{n^2\theta^2}{(n-1)^2} = \frac{n^2\theta^2}{(n-1)^2(n-2)}.$$

3. Шукаємо кількість інформації по Фішеру

$$L(x,\theta)=\prod_{i=1}^{n} f(x_i,\theta)=\theta^n e^{-\theta\sum x_i},\Rightarrow \ln L(x,\theta)=n\ln\theta-\sum_{i=1}^{n} x_i,$$

$$\frac{\partial}{\partial\theta}\ln L(x,\theta)=\frac{n}{\theta}-\sum_{i=1}^{n} x_i,\Rightarrow \frac{\partial^2}{\partial\theta^2}\ln L(x,\theta)=-\frac{n}{\theta^2};$$

$$I(\theta)=-M\,\frac{\partial^2}{\partial\theta^2}\ln L(\xi,\theta)=\frac{n}{\theta^2}.$$

4. Перевіряємо, чи досягається нижня межа в нерівності 2).
Бачимо, що

$$\frac{\left(1+b'(\theta)\right)^2}{I(\theta)}=\frac{\left(1+\dfrac{1}{n+1}\right)^2\theta^2}{n}=\frac{n\theta^2}{(n-1)^2}.$$

Оскільки $\dfrac{n}{n-2}>1,\Rightarrow D\hat{\theta}>\dfrac{\left(1+b'(\theta)\right)^2}{I(\theta)}$.

Робимо висновок, що $\hat{\theta}$ не буде ефективною оцінкою параметра θ.

Завдання для самостійної роботи

1) Нехай $S^2=\dfrac{\sum_{i=1}^{n}(\xi_i-\overline{\xi})^2}{n}$ - вибіркова дисперсія.

Потрібно показати, що $MS^2=\dfrac{n-1}{n}\sigma^2$.

2) Нехай задано вибірку з генеральної сукупності $\xi_1,\xi_2,\dots,\xi_n$ з розподілом Паскаля

$P\left(\xi=k\right)=\dfrac{\theta^k}{\left(1+\theta\right)^{k+1}}$, $\theta>0$, $k=0,1,\dots$. Показати, що $\hat{\theta}=\dfrac{\sum_{i=1}^{n}\xi_i}{n}$ буде ефективною оцінкою параметра θ.