Лекция 2 17.01

Выборочные моменты

$X_1 \dots X_n \sim F(x, \theta)$

Опр k-ым начальным выборочным моментом называется $\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$, $k = 1, 2, \dots$

$\hat{\mu}_1 = \overline{X} = \frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ называется выборочное среднее

Опр k-ым центральным выборочным моментом называется $\hat{\nu}_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k$ $k = 2, 3 \dots$

$\hat{\nu}_2 = S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ называется выборочная дисперсия

Пусть $(X_1, Y_1), \dots (X_n, Y_n)$ соответствует распределению $F(x, y, \theta)$

Опр $\hat{K}_{xy} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})$ называется выборочной ковариацией

Опр $\hat{\rho}_{xy} = \frac{\hat{K}_{xy}}{\sqrt{S_x^2 S_y^2}}$ называется выборочным коэф-том корреляции

Св-ва выборочных моментов

1. $E\hat{\mu}_k = E\left(\frac{1}{n}\sum_{i=1}^{n}X_i^k\right) = \frac{1}{n}\sum_{i=1}^{n} EX_i^k = EX_1^k = \mu_k$

$E\overline{X} = m_{X_1}$

2. $D\hat{\mu}_k = D\left(\frac{1}{n}\sum_{i=1}^{n} X_i^k\right) = \frac{1}{n^2}\sum_{i=1}^{n} DX_i^k = \frac{1}{n} D X_1^k = \frac{1}{n}\left(EX_i^{2k} - (EX_1^k)^2\right) = \frac{1}{n}\left(\mu_{2k} - \mu_k^2\right)$

$D\overline{X} = \frac{\sigma^2_{X_1}}{n}$

3. по УЗБЧ

$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow[n\to\infty]{п.н.} E\hat{\mu}_k = \mu_k$

4. по УЗБЧ

$\hat{\nu}_k \xrightarrow{п.н.} \nu_k$

5. по ЦПТ $\frac{\hat{\mu}_k - \mu_k}{\sqrt{\frac{\mu_{2k} - \mu_k^2}{n}}} \xrightarrow[n\to\infty]{d} U, U \sim N(0; 1)$

$\frac{\sqrt{n}\,(\overline{X} - m_{X_1})}{\sigma} \xrightarrow[n\to\infty]{d} U$

6. $ES^2 = E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\right) = \underbrace{\frac{n-1}{n}}_{\text{несмещенная}}\sigma^2$

7. $E\hat{K}_{xy} = \frac{n-1}{n}\operatorname{cov}(X, Y)$

Опр Оценкой $\hat{\theta}$ параметра θ называется ф-ия $\hat{\theta} = T(X_1, \dots, X_n)$, не зависящая от θ

$\hat{m}_1 = \frac{X_1 + X_2 + X_3}{3}$ $\quad E\left(\frac{X_1 + X_2 + X_3}{3}\right) = \frac{1}{3}\cdot 3m = m$ несмещенная

Опр Оценка $\hat{\theta}$ называется несмещенной оценкой параметра θ, если $E\hat{\theta} = \theta$, для $\forall\, \theta \in \Theta$ → множество возможных значений для параметров

Опр Оценка $\hat\theta(x_1,\dots,x_n)$ назыв. асимптотически несмещенной оценкой θ, если $\lim\limits_{n\to\infty} E\hat\theta(x_1,\dots,x_n)=\theta$

$\lim\limits_{n\to\infty} ES^2 = \lim\limits_{n\to\infty}\frac{n-1}{n}\sigma^2=\sigma^2$

$\tilde{S}^2=\frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\bar{x})^2$ – несмещенная („исправленная") выборочная дисперсия

Опр Оценка $\hat\theta(x_1,\dots,x_n)$ называется состоятельной оценкой θ, если $\hat\theta(x_1,\dots,x_n)\xrightarrow[n\to\infty]{P}\theta$

оценка называется сильно состоятельной, если $\hat\theta(x_1,\dots,x_n)\xrightarrow[n\to\infty]{п.н.}\theta$

Опр пусть $\hat\theta$ – несмещ. оценка параметра θ

Если $D\hat\theta\le D\theta^*$, где θ^* – любая несмещенная оценка пар-ра θ, то $\hat\theta$ называется эффективной оценкой параметра θ.

R-эффективные оценки

$x_1\dots x_n\sim f(x,\theta)$, $\theta\in\Theta\in R^1$

назовем модель $(S, f(x,\theta))$ регулярной, если она удовлетворяет следующим условиям:

1) ф-ия $f(x,\theta)=f(x_1,\dots,x_n,\theta)>0$ для $\forall x\in S$ и дифф-ма по θ

2) $\frac{\partial}{\partial\theta}\int\limits_S f(x,\theta)dx=\int\limits_S\frac{\partial}{\partial\theta}f(x,\theta)dx$

$\frac{\partial}{\partial\theta}\int\limits_S T(x)f(x,\theta)dx=\int\limits_S\frac{\partial}{\partial\theta}T(x)f(x,\theta)dx$

пусть $T(x)=T(x_1,\dots,x_n)$ – несмещ. оценка пар-ра θ

вычтем из 2 строчки 1

$$0=\int\limits_S\frac{\partial}{\partial\theta}f(x,\theta)dx=\int\frac{\partial f(x,\theta)}{\partial\theta}\frac{f(x,\theta)}{f(x,\theta)}dx=\int\limits_S\frac{\partial\ln f(x,\theta)}{\partial\theta}f(x,\theta)dx$$

$$1=\int\frac{\partial}{\partial\theta}T(x)f(x,\theta)dx=\int T(x)\frac{\partial}{\partial\theta}f(x,\theta)\cdot\frac{f(x,\theta)}{f(x,\theta)}dx=\int T(x)\frac{\partial\ln f(x,\theta)}{\partial\theta}f(x,\theta)dx \quad *$$

Опр Информация Фишера о параметре θ, содержащейся в выборке $x_1,\dots,x_n$, называется величина

$$I_n(\theta)=E\left(\frac{\partial\ln f(X,\theta)}{\partial\theta}\right)^2=\int\limits_S\left(\frac{\partial\ln f(x,\theta)}{\partial\theta}\right)^2 f(x,\theta)dx$$

Теорема (неравенство Рао–Крамера)

Если $(S,f(x,\theta))$ – регулярная модель и $\hat\theta$ - несмещ. оценка θ, то $D\hat\theta\ge\frac{1}{I_n(\theta)}$

неравенство Коши–Буняковского

$$\left(\int\varphi_1(x)\varphi_2(x)dx\right)^2\le\int\varphi_1^2(x)dx\int\varphi_2^2(x)dx$$

$$*\ 1=\int(T(x)-\theta)\frac{\partial\ln f(x,\theta)}{\partial\theta}f(x,\theta)dx\le\underbrace{\int\limits_S(T(x)-\theta)^2f(x,\theta)dx}_{D\hat\theta}\underbrace{\int\limits_S\left(\frac{\partial\ln f(x,\theta)}{\partial\theta}\right)^2f(x,\theta)dx}_{I_n(\theta)}=$$

применим КБ

$\varphi_1(x)=(T(x)-\theta)\sqrt{f(x,\theta)}$

$\varphi_2(x)=\frac{\partial\ln f(x,\theta)}{\partial\theta}\sqrt{f(x,\theta)}$

$$=\int\limits_S\left(\frac{\partial\ln f(x,\theta)}{\partial\theta}\right)^2f(x,\theta)dx=I_n(\theta)$$

Опр Оценка $\hat\theta$ называется R-эффективной, если $E\hat\theta$ и $D\hat\theta=\frac{1}{I_n(\theta)}$