26.05.2014 ТВМС - лек.

б) оценки с минимална дисперсия.

$m \longrightarrow \underbrace{\overline{X}, X_1, X_2}, \frac{X_1+X_2}{2}, \ldots$

вс. са неизместени оценки за m.

Коя от тях да предпочетем?

Дисперсията отчита „точността"

$$DX_1 = DX_2 = DX$$

$$D\frac{X_1+X_2}{2} = \frac{1}{4}D(X_1+X_2) = \frac{1}{4}(DX_1+DX_2) =$$

$$= \frac{1}{4}\cdot 2DX = \frac{DX}{2}$$

$$D\overline{X} = D\frac{X_1+X_2+\ldots+X_n}{n} = \frac{1}{n^2}(DX_1+DX_2+\ldots+DX_n) =$$

$$= \frac{DX}{n}$$ най „качествена" оценка

по-малка дисперсия = най-точна оценка

Лема: Ако t е неизместена оценка с минимална дисперсия за θ, то тя $\exists$!.

$t = t(x_1, \ldots, x_n)$ - ф-ция на извадката.

Д-во: Нека t_1 и t_2 са неизместени оценки с мин. дисперсия

$\Rightarrow Et_1 = Et_2 = \theta$,

но и $Dt_1 = Dt_2 = d > 0$

Разгл. $t_3 = \frac{t_1+t_2}{2}$.

$Et_3 = \theta \Rightarrow t_3$ също е неизм. оценка

$Dt_3 = D\frac{t_1+t_2}{2} = \frac{1}{4}(D(t_1+t_2))$

Но $t_i = t_i(x_1, \dots, x_n)$ $i=1,2 \Rightarrow$ те са зависими

$\Rightarrow$ прилагаме общата ф-ла за

дисперсия на сумата:

$$Dt_3 = \frac{1}{4} D(t_1+t_2) = \frac{Dt_1 + Dt_2 + 2\operatorname{cov}(t_1,t_2)}{4}$$

$$= \frac{d + \operatorname{cov}(t_1,t_2)}{2}.$$

Но $d \le Dt_3$

$\Rightarrow \operatorname{cov}(t_1,t_2) \ge d$

Но от нер. на Коши-Буняковски:

$\operatorname{cov}(t_1,t_2) \le \sqrt{Dt_1 \cdot Dt_2} = d$

$\Rightarrow \operatorname{cov}(t_1,t_2) = d$

$\Rightarrow$ в нер. на Коши се достига "$=$"

т.е. м/у величините има детерминирана линейна зависимост.

$\Rightarrow t_1 = at_2 + b$, a, b са числа.

$D(t_1) = D(at_2 + b)$

$d = D(at_2) = a^2 \cdot D(t_2) = a^2 \cdot d$

$\Rightarrow a = \pm 1$

$a = -1 \Rightarrow \operatorname{cov}(t_1,t_2) = -d \Rightarrow a \neq -1$

$\Rightarrow a = 1$

$\Rightarrow t_1 = t_2 + b$

$Et_1 = E(t_2 + b)$

$\theta = Et_2 + b = \theta + b$

$\Rightarrow b = 0$

$\Rightarrow t_1 = t_2$.

Неравенство на Рао-Крамер

$f(x_1, x_2, \_\_\_, x_n | \theta)$ - съвместна вероятностна плътност на извадката, ако неизвестният параметър е $= \theta$.

t е неизместена оценка за детерминираната ф. $\tau(\theta)$.

$$Dt \geq \frac{(\tau'(\theta))^2}{\underbrace{D\frac{\partial}{\partial\theta} \ln f(x_1, x_2, \ldots, x_n | \theta)}}$$

нар. се информационно количество на Фишер.

При t оценка на θ

$$\Rightarrow Dt \geq \frac{1}{D\frac{\partial}{\partial\theta} \ldots}$$ - обратна пропорционалност

По-голямо информац. количество

$\Rightarrow$ по-голям знам. $\Rightarrow$ по-точна оценка!