

02.06.2014г.

ТЗМС - лек.

Неравенство на Рао-Крамер

t = неизв. оценка за $z(\theta)$ $t = (x_1, \dots, x_n)$
 $f(x_1, x_2, \dots, x_n | \theta)$ - съвм. вероятностна
 плътност на извадката $x_1, x_2, \dots, x_n$,
 ако стойността на неизв. параметър е θ .

$$Dt \geq \frac{(z'(\theta))^2}{E \left(\frac{\partial}{\partial \theta} \ln f(x_1, x_2, \dots, x_n | \theta) \right)}$$

информационно количество
на функцията.

До-во: $E t = z(\theta) \quad \frac{d}{d\theta} \left| \quad z(\theta) = \int t(x_1, \dots, x_n) \cdot f(x_1, \dots, x_n | \theta) dx_1 dx_2 \dots dx_n \right.$
 $\mathbb{R}^n$

$$\frac{d}{d\theta} \left| \quad 1 = \int f(x_1, \dots, x_n | \theta) dx_1 dx_2 \dots dx_n \right.$$

Диференциране по θ

$$z'(\theta) = \int_{\mathbb{R}^n} t(x_1, \dots, x_n) \frac{\partial}{\partial \theta} f(x_1, \dots, x_n | \theta) dx_1 dx_2 \dots dx_n$$

$$0 = \int_{\mathbb{R}^n} \frac{\partial}{\partial \theta} f(x_1, \dots, x_n | \theta) dx_1 \dots dx_n$$

$$z(\theta) = \int_{\mathbb{R}^n} t(x_1, \dots, x_n) \frac{\frac{\partial}{\partial \theta} f(x_1, \dots, x_n | \theta)}{f(x_1, \dots, x_n | \theta)} \cdot f(x_1, \dots, x_n | \theta) dx_1 \dots dx_n$$

$$D = \int_{\mathbb{R}^n} \frac{\frac{\partial}{\partial \theta} f(x_1, \dots, x_n | \theta)}{f(x_1, \dots, x_n | \theta)} \cdot f(x_1, \dots, x_n | \theta) dx_1 \dots dx_n$$

$$= \frac{\partial}{\partial \theta} \ln f(x_1, \dots, x_n | \theta)$$

$$z'(\theta) = E \left(t(x_1, \dots, x_n) \frac{\partial}{\partial \theta} \ln f(x_1, \dots, x_n | \theta) \right)$$

$$0 = E \left(\frac{\partial}{\partial \theta} \ln f(x_1, \dots, x_n | \theta) \right) \quad (A.2)$$

$$z'(\theta) = E \left(\underbrace{(t(x_1, x_2, \dots, x_n) - z(\theta))}_{\text{z}} \cdot \underbrace{\frac{\partial}{\partial \theta} \ln f(x_1, \dots, x_n | \theta)}_{\text{y}} \right) \leq$$

$\downarrow E(t) = E(z(\theta))$ заради това, че t е неизм. сума за $z(\theta)$.

$$E |z \cdot y| \leq \sqrt{E(z^2) \cdot E(y^2)}$$

↳ неравенство на Коши

$$\leq \sqrt{D(t(x_1, \dots, x_n)) \cdot D \left(\frac{\partial}{\partial \theta} \ln f(x_1, \dots, x_n | \theta) \right)}$$

Визгаме $\sqrt{2}$ и дълим и по това

Случай на независима извадка

$$f(x_1, \dots, x_n | \theta) = \prod_{i=1}^n f(x_i | \theta)$$



вероятностна плътност
на генералната съвкупност: X

$$\frac{\partial}{\partial \theta} \ln f(x_1, x_2, \dots, x_n | \theta) = \sum_{i=1}^n \frac{\partial}{\partial \theta} \ln f(x_i | \theta)$$

$$D \left(\frac{\partial}{\partial \theta} \ln f(x_1, \dots, x_n | \theta) \right) = D \left(\sum_{i=1}^n \frac{\partial}{\partial \theta} \ln f(x_i | \theta) \right) =$$

$$= \sum_{i=1}^n D \left(\frac{\partial}{\partial \theta} \ln f(x_i | \theta) \right) \stackrel{\text{об. 2}}{=} n \cdot D \left(\frac{\partial}{\partial \theta} \ln f(x | \theta) \right)$$

$$D t \geq \frac{(t'(\theta))^2}{n \cdot D \left(\frac{\partial}{\partial \theta} \ln f(x | \theta) \right)}$$

Д.) t - ефективна оценка, ако в
неравенството на Ф.К. се достига "=".

Всяка ефективна оценка е с минимална дисперсия

Пример на ефективна оценка:

$$X \sim N(\theta, 1)$$

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

→ ефективна оц.

$$D \bar{X} = \frac{1}{n}$$

$$f(x | \theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}}$$

$$\ln f(x | \theta) = -\frac{1}{2} \ln(2\pi) - \frac{(x-\theta)^2}{2} \quad \left| \frac{\partial}{\partial \theta} \right|$$

$$\frac{\partial}{\partial \theta} \ln f(x | \theta) = x - \theta$$

$$\int_{\mathcal{D}} \int_{\mathcal{D}} \text{pdf}(x|\theta) = \int (x-\theta) = 1$$

Статистически хипотези

(генерална съвкупност)

θ - неизв. параметър на з.с X

$x_1, \dots, x_n$ - извадка (случайна точка в $\mathbb{R}^n$)!

$\theta = ?$

1. Угед на метода:

θ_0 - число, $\theta \in$ възможна стойност на θ .

(възможно е $\theta = \theta_0$)

$\theta \neq \theta_0$ също е възможно

$H_0: \theta = \theta_0$ = нулева (основна) хипотеза

$H_1: \theta \neq \theta_0$ алтернативна (контра) хипотеза;

Въз основа на извадката кое от хипотезите да изберем?

$H_0: \theta \leq \theta_0$

или

$H_0: \theta \geq \theta_0$

фирма - производство

$H_1: \theta > \theta_0$

$H_1: \theta < \theta_0$

неизвестност

п.к.картаво;

2. Как да приемем кое хипотеза да приемем?

- решаващо правило:

избира се област $K \subset \mathbb{R}^n$. (критична област)

Ако $(x_1, x_2, \dots, x_n) \in K$ → отхвърляме H_0 ,

В противен сл. → приемаме H_0 .

3. Анализ на грешките

а) Ако $(x_1, \dots, x_n) \in K$, а $g_a \in \text{Зарна } H_0$.

→ грешка от първи род.

б) Ако $(x_1, \dots, x_n) \notin K$, а $g_a \in \text{Зарна } H_1$.

→ грешка от втори род.

Действително верна хипотеза	Приета за верна	
	H_0	H_1
H_0	Правилно <u>1-α</u>	Грешка от I род <u>α</u>
H_1	Грешка от II род <u>$1-\beta$</u>	Правилно <u>1-β</u> β -ние

Вероятност за гр. I род:

$$P((x_1, \dots, x_n) \in K | H_0) = \alpha \rightarrow \text{тило на значимост}$$

II род:

$$P((x_1, \dots, x_n) \notin K | H_1) \rightarrow \min$$

Да изберем такава кр. област K :
грешките да са малки.

Фиксиране грешката от I род и искане да
минимизираме грешката от II род.
оказва се, че тази гр. има β -ние

β -нието е нар. оптимална критична
област.

$$\beta(x_1, \dots, x_n \in K | H_1) = \beta$$

$1 - \beta$ = мощност на теста;

4. Лема на Нейман - Пирсън:

н.к. от вероятностните функции на
непр. параметр θ :

$$\theta = \{ \theta_0, \theta_1 \} \quad \left| \begin{array}{l} H_0: \theta = \theta_0 \\ H_1: \theta = \theta_1 \end{array} \right.$$

$$f(x_1, \dots, x_n | \theta = \theta_0) = f_0(x_1, \dots, x_n)$$

$$f(x_1, \dots, x_n | \theta = \theta_1) = f_1(x_1, \dots, x_n)$$

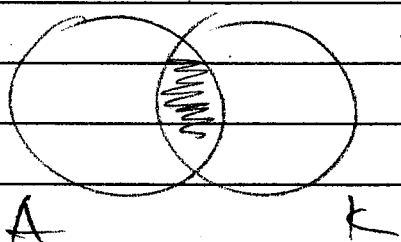
$$\rightarrow K = \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n : \frac{f_1(x_1, \dots, x_n)}{f_0(x_1, \dots, x_n)} \geq C_\alpha \right\}$$

C зависи от нивото на значимост

Д-во: Винаги 2 обратни: $K \subset A$:

$$P((x_1, \dots, x_n) \in A | H_0) = \alpha = P((x_1, x_2, \dots, x_n) \in K | H_0)$$

$$\int_A f_0(x_1, \dots, x_n) dx_1 dx_2 \dots dx_n = \int_K f_0(x_1, \dots, x_n) dx_1 \dots dx_n$$



$$\int - \int$$

$A \setminus K \quad K \setminus A$

$$\text{УгЕ. док. че: } \int_{\substack{K \\ (K \setminus A)}} f_1(x_1, \dots, x_n) dx_1 \dots dx_n \geq \int_{\substack{A \\ (K \setminus K)}} f_1(x_1, \dots, x_n) dx_1 \dots dx_n$$

$$\int_{K \setminus A} f_1(x_1, \dots, x_n) dx_1 \dots dx_n \geq C_\alpha \int_{K \setminus A} f_0(x_1, \dots, x_n) dx_1 \dots dx_n =$$

$$= C_\alpha \int_{\substack{A \\ (K \setminus K)}} f_0(x_1, \dots, x_n) dx_1 \dots dx_n >$$

$$> \int_{\substack{A \\ (K \setminus K)}} f_1(x_1, \dots, x_n) dx_1 \dots dx_n$$

→ док. Топсетомо.

Дп: $X \sim N(\theta, 1)$
 наг. оз. дуча.

$$H_0: \theta = 0$$

$$H_1: \theta = 1$$

$$f(x_1, \dots, x_n | \theta) = (2\pi)^{-\frac{n}{2}} \cdot e^{-\frac{1}{2} \sum_{i=1}^n (x_i - \theta)^2}$$

$$f_1(x_1, \dots, x_n) = (2\pi)^{-\frac{n}{2}} \cdot e^{-\frac{1}{2} \sum_{i=1}^n (x_i - 1)^2}$$

$$f_0(x_1, \dots, x_n) = (2\pi)^{-\frac{n}{2}} \cdot e^{-\frac{1}{2} \sum_{i=1}^n x_i^2}$$

$$\frac{f_1(x_1, \dots, x_n)}{f_0(x_1, \dots, x_n)} = e^{-\frac{1}{2} \sum_{i=1}^n ((x_i - 1)^2 - x_i^2)} =$$

$$= e^{\sum_{i=1}^n (x_i - \frac{n}{2})} = e^{n \left(\bar{x} - \frac{1}{2} \right)} \geq C_\alpha$$

ср. аритм.

Потрапување:

$$K = \{ (x_1, \dots, x_n) : \bar{x} > c_\alpha' \}$$

Правилно $\Rightarrow$ $\bar{x} > c_\alpha'$

$$P(\bar{x} > c_\alpha') = \alpha$$

при верно H_0 .
и така напредок
 c_α'
