Διάλεξη 3η
15/03/2019

ΣΤΑΤΙΣΤΙΚΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ

Αμερόληπτος ΑΟΕΔ εκτιμητής
- i) Αμερόληπτος
- ii) Έχει την ελάχιστη διακύμανση μεταξύ όλων των άλλων αμερόληπτων εκτιμητών.

Ερώτημα: Πώς μπορώ να κατασκευάσω ΑΟΕΔ εκτιμητή;
↳ **Απάντηση:** 1ος τρόπος: Ανισότητα Cramer-Rao

Ανισότητα Cramer-Rao: Έστω τ.δ. $X_1, X_2, \ldots, X_n$ από πληθυσμό με κατανομή $f(x,\theta)$
$\theta \in \Theta \subseteq \mathbb{R}$ ($\mathbb{R}^m$). Έστω $f(\underset{\sim}{x},\theta)$ η από κοινού κατανομή του $\underset{\sim}{x} = (x_1, \ldots, x_n)'$ η οποία είναι $f(\underset{\sim}{x},\theta) = \prod_{i=1}^{n} f(x_i,\theta)$.

Συνθήκες κανονικότητας:

I_1) Ο παραμετρικός χώρος Θ είναι ανοιχτό σύνολο

I_2) Το $S = \{(x_1,\ldots,x_n) : f(\underset{\sim}{x},\theta) > 0\}$ δεν εξαρτάται από την παράμετρο θ.

I_3) Η παράγωγος της $\frac{\partial}{\partial\theta} f(\underset{\sim}{x},\theta)$ υπάρχει και είναι πεπερασμένη

I_4) $\int \frac{\partial}{\partial\theta} f(\underset{\sim}{x},\theta)\, d\underset{\sim}{x} = \frac{\partial}{\partial\underset{\sim}{\theta}} \int f(\underset{\sim}{x},\theta)\, d\underset{\sim}{x} \left(= \frac{\partial}{\partial\underset{\sim}{\theta}} 1 = 0 \right)$

I_5) Αν $T = T(X_1,\ldots,X_n)$ είναι μια στατιστική συνάρτηση

$$\int T(\underset{\sim}{x}) \frac{\partial}{\partial\theta} f(\underset{\sim}{x},\theta)\, d\underset{\sim}{x} = \frac{\partial}{\partial\theta} \int T(\underset{\sim}{x}) f(\underset{\sim}{x},\theta)\, d\underset{\sim}{x}$$

Ι6) Η ποσότητα $I^F_{\underset{\sim}{x}}(\theta) = \int \left(\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x}, \theta) \right)^2 f(\underset{\sim}{x}, \theta) d\underset{\sim}{x}$

υπάρχει και $0 < I^F(\theta) < +\infty$.

Θεώρημα: (Ανισότητα Cramer-Rao)

Έστω ότι οι συνθήκες κανονικότητας ισχύουν.
Αν $T = T(\underset{\sim}{x})$ είναι αμερόληπτη εκτιμήτρια της $g(\theta)$
τότε $Var(T(\underset{\sim}{x})) \geq \frac{[g'(\theta)]^2}{I^F_{\underset{\sim}{x}}(\theta)}$

Απόδ.

$$E\left(\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x}, \theta)\right) = \int \left(\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x}, \theta)\right) f(\underset{\sim}{x}, \theta) d\underset{\sim}{x}$$

$$= \int \frac{\frac{\partial}{\partial\theta} f(\underset{\sim}{x}, \theta)}{\cancel{f(\underset{\sim}{x}, \theta)}} \cancel{f(\underset{\sim}{x}, \theta)} dx \overset{(I_4)}{=} 0$$

$$Var\left(\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x}, \theta)\right) = E\left(\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x}, \theta)\right)^2 - \left(E\left(\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x}, \theta)\right)\right)^2$$

Επ-ορισμό $I^F_{\underset{\sim}{x}}(\theta)$.

$$Cov\left(T(\underset{\sim}{x}), \frac{\partial}{\partial\theta} \log f(\underset{\sim}{x}, \theta)\right) \overset{ορισμό}{=} E\left(T(\underset{\sim}{x}) \frac{\partial}{\partial\theta} \log f(\underset{\sim}{x}, \theta)\right) -$$

$$- E(T(\underset{\sim}{x})) \times \underbrace{E\left(\cancel{\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x}, \theta)}\right)}_{\to 0} =$$

$$= \int T(\underset{\sim}{x}) \left(\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x}, \theta)\right) f(\underset{\sim}{x}, \theta) d\underset{\sim}{x}$$

$$= \int T(\underset{\sim}{x}) \frac{\frac{\partial}{\partial\theta} f(\underset{\sim}{x}, \theta)}{f(\underset{\sim}{x}, \theta)} f(\underset{\sim}{x}, \theta) d\underset{\sim}{x} \overset{(I_5)}{=} \frac{\partial}{\partial\theta} \int T(\underset{\sim}{x}) f(\underset{\sim}{x}, \theta) d\underset{\sim}{x}$$

$$= \frac{\partial}{\partial\theta} g(\theta) = g'(\theta).$$

Cauchy-Schwartz:

$$Cov^2\left(T(\underset{\sim}{x}), \frac{\partial}{\partial\theta}\log f(\underset{\sim}{x},\theta)\right) \le Var\left(T(\underset{\sim}{x})\right) Var\left(\frac{\partial}{\partial\theta}\log f(\underset{\sim}{x},\theta)\right)$$

$$\Rightarrow [g'(\theta)]^2 \le Var(T(\underset{\sim}{x}))\, I^F_{\underset{\sim}{x}}(\theta) \Rightarrow Var(T(\underset{\sim}{x})) \ge \frac{[g'(\theta)]^2}{I^F_{\underset{\sim}{x}}(\theta)}$$

Παρατηρήσεις:

① $K_{\text{ΦΕΙR}} = \dfrac{(g'(\theta))^2}{I^F_{\underset{\sim}{x}}(\theta)}$.

② Το $I^F_{\underset{\sim}{x}}(\theta)$ ονομάζεται μέτρο πληροφορίας του Fisher. Εκφράζει το ποσό της πληροφορίας που περιέχεται στο τ.δ. για την παράμετρο θ.

③ $$I^F_{\underset{\sim}{x}}(\theta) = E\left(\frac{\partial}{\partial\theta}\log f(\underset{\sim}{x},\theta)\right)^2 = Var\left(\frac{\partial}{\partial\theta}\log f(\underset{\sim}{x},\theta)\right)$$
$$= Var\left(\frac{\partial}{\partial\theta}\log \prod_{i=1}^{n} f(x_i,\theta)\right)$$
$$= Var\left(\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(x_i,\theta)\right)$$

$$\underset{x_1, x_2, \dots, x_n}{\overset{\text{ανεξαρτησία}}{=}} \sum_{i=1}^{n} Var\left(\frac{\partial}{\partial\theta}\log f(x_i,\theta)\right)$$

$$= \sum_{i=1}^{n}\left\{ E\left(\frac{\partial}{\partial\theta}\log f(x_i,\theta)\right)^2 - \left(E\left(\frac{\partial}{\partial\theta}\log f(x_i,\theta)\right)\right)^2 \right\}$$

$$= \sum_{i=1}^{n} E\left(\frac{\partial}{\partial\theta}\log f(x_i,\theta)\right)^2 = \sum_{i=1}^{n} I^F_{x_i}(\theta) \underset{\substack{f(x_i,\theta) = \\ f(x,\theta),\,\forall i}}{\overset{x_i \text{ ισόνομες}}{=}} n \cdot I^F_x(\theta)$$

$$\boxed{I^F_{\underset{\sim}{x}}(\theta) = n \cdot I^F_x(\theta)}$$

$$I_{\underline{x}}^{F}(\theta) = \frac{1}{\theta(1-\theta)}$$

<u>άλλος τρόπος:</u> $\underbrace{I^{F}_{x}(\theta)} = E\left(\frac{\partial^2}{\partial\theta^2} \log P(x,\theta)\right)$

$$\frac{\partial^2}{\partial\theta^2} \log P(x,\theta) = \frac{\partial^2}{\partial\theta^2}\left(x\log\theta + (1-x)\log(1-\theta)\right)$$

$$= \frac{\partial}{\partial\theta}\left(\frac{x}{\theta} - \frac{1-x}{1-\theta}\right)$$

$$= \frac{x}{\theta^2} + \frac{1-x}{(1-\theta)^2}$$

$$= -E\left(\frac{x}{\theta^2} + \frac{1-x}{(1-\theta)^2}\right) = \frac{E(x)}{\theta^2} + \frac{1-E(x)}{(1-\theta)^2} = \frac{\theta}{\theta^2} + \frac{(1-\theta)}{(1-\theta)^2} = \dots$$

$$K\Phi_{C-R} = \frac{[g^{f}(\theta)]^2}{n\, I^{F}_{x}(\theta)} = \frac{1}{n[1/\theta(1-\theta)]} \Rightarrow K\Phi_{C-R} = \frac{\theta(1-\theta)}{n}$$

Δοκιμάζω την $T = T(\underline{x}) = \bar{x}$.

$$E(T) = E(\bar{x}) = E\left(\tfrac{1}{n}\sum x_i\right) = \tfrac{1}{n}\sum E(x_i) = \tfrac{1}{n}\sum\theta = \tfrac{1}{n}n\theta = \theta.$$

$$Var(T) = Var(\bar{x}) = Var\left(\tfrac{1}{n}\sum x_i\right) = \tfrac{1}{n^2}\sum Var(x_i) = \tfrac{1}{n^2}\sum\theta(1-\theta)$$

$$= \tfrac{1}{n^2} n\theta(1-\theta) = \frac{\theta(1-\theta)}{n} = K\Phi_{C-R}$$

$\bar{x}$: ΑΟΕΔ θ.

4) Απαιτούμενα του ΚΦ.

$$K\Phi_{C-R} = \frac{[g^{f}(\theta)]^2}{n\, I^{F}_{x}(\theta)}$$

5) $I^{F}_{\underline{x}}(\theta) = \underbrace{E\left(\frac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right)^2 = -E\left(\frac{\partial^2}{\partial\theta^2}\log f(x,\theta)\right)}_{\text{χρήσιμη σχέση.}}$

Πως εφαρμόζεται η ανισότητα C-R;

1ο Βήμα: Υπολογισμός $ΚΦ_{C-R}$.

2ο Βήμα: Εύρεση μιας στατιστικής συνάρτησης.

$$T = T(\underline{X}) \text{ τ.ω } \begin{cases} α)\ E(T) = g(\theta) \\ β)\ Var(T) = ΚΦ_{C-R}. \end{cases}$$

Παράδειγμα 1: Έστω τ.δ. $X_1, \dots, X_n$ από πληθυσμό με κατανομή Poisson (θ). $P(x,\theta) = \dfrac{e^{-\theta}\theta^x}{x!}$, $\theta > 0$, $x = 0,1,2,\dots$

Να βρεθεί ΑΟΕΔ της θ.

Απάντηση

Κάτω φράγμα

$$ΚΦ_{C-R} = \frac{[g'(\theta)]^2}{n I_X^F(\theta)} = \frac{1}{n I_X^F(\theta)},$$

$$I_X^F(\theta) = E\left(\frac{\partial}{\partial\theta}\log P(x,\theta)\right)^2 = -E\left(\frac{\partial^2}{\partial\theta^2}\log P(x,\theta)\right) \text{ : σχέση (A)}$$

- $\log P(x,\theta) = -\theta + x\log\theta - \log x!$
- $\dfrac{\partial}{\partial\theta}\log P(x,\theta) = -1 + \dfrac{x}{\theta}$.

Σχέση (A): $$I_X^F(\theta) = E\left(\frac{x}{\theta} - 1\right) = \frac{1}{\theta^2}E(X-\theta)^2 = \frac{1}{\theta^2}E(X - E(X))^2 = \frac{1}{\theta^2}Var(X) = \frac{\theta}{\theta^2} = \frac{1}{\theta}$$

$$E(X-\theta)^2 = E(X^2) - 2\theta E(X) + \theta^2 = Var(X) + (EX)^2 - 2\theta E(X) + \theta^2$$