16/3/2017

$X_1,\dots,X_n \to f(x,\theta),\ \theta \in \Theta \subseteq \mathbb{R}$

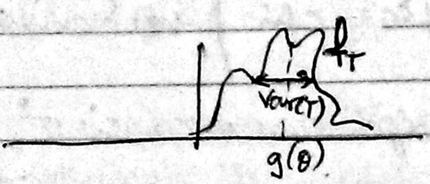
Πρόβλημα: Εύρεση $T(X_1,\dots,X_n) \rightsquigarrow g(\theta)$
↑ εκτιμητής

$ΜΤΣ(T,\theta) = \underset{\text{διακύμανση}}{Var(T)} + Μεροληψία^2$

Αμεροληψία: T: αμερόληπτος αν $E(T)=g(\theta)$

ΑΟΕΔ εκτιμητής: Ο T θα λέγεται ΑΟΕΔ εκτιμητής $g(\theta)$

① T: αμερόληπτος της $g(\theta)$



② Αν $Var(T)$ είναι η ελάχιστη μεταξύ των διακυμάνσεων ΟΛΩΝ των αμερόληπτων εκτιμητών $g(\theta)$.

Θεώρημα μοναδικότητας ΑΟΕΔ
Αν υπάρχει ο ΑΟΕΔ εκτιμητής $T(X_1,\dots,X_n)$ της $g(\theta)$, $\theta\in\Theta$, τότε ο $T(X_1,\dots,X_n)$ είναι μοναδικός

Ερώτημα: Πως μπορώ να βρω ΑΟΕΔ εκτιμητές;

$\boxed{\text{1η Απάντηση:}}$ → Μια μέθοδος εύρεσης ΑΟΕΔ από εφαρμογή ανισότητα Cramer-Rao

Συμβολισμοί - Συνθήκες κανονικότητας
Έστω $X_1,\dots,X_n$ τ.δ. από πληθυσμό $f(x,\theta)$, $\theta\in\Theta\subseteq\mathbb{R}$
Συμβολίζω με $f(\underset{\sim}{x},\theta)$ την από κοινού κατανομή του τυχαίου διανύσματος $(X_1,X_2,\dots,X_n)$. Τότε: Επειδή $X_1,\dots,X_n$ είναι ανεξάρτητες + ισόνομες

$$f(\underset{\sim}{x},\theta) = \prod_{i=1}^{n} f(x_i,\theta)$$

Συνθήκες Κανονικότητας: Θεωρώ ότι:

1. Ο παραμετρικός χώρος Θ είναι ανοικτό $\subseteq \mathbb{R}$
2. Το σύνολο θετικότητας της $f(x, \theta)$ $S = \{ \underset{\sim}{x} = (x_1, \dots, x_n) : f(\underset{\sim}{x}, \theta) > 0 \}$ ανεξάρτητο της θ.
3. $\forall \underset{\sim}{x} \in S$ και $\theta \in \Theta$ η $\frac{\partial}{\partial \theta} \log f(\underset{\sim}{x}, \theta)$ υπάρχει και είναι πεπερασμένη.
4. $\int_{\mathbb{R}^n} \frac{\partial}{\partial \theta} f(\underset{\sim}{x}, \theta) dx = \frac{\partial}{\partial \theta} \int_{\mathbb{R}^n} f(\underset{\sim}{x}, \theta) dx \left(= \frac{\partial}{\partial \theta} 1 = 0 \right)$
5. $\int_{\mathbb{R}^n} T(\underset{\sim}{x}) \frac{\partial}{\partial \theta} f(\underset{\sim}{x}, \theta) d\underset{\sim}{x} = \frac{\partial}{\partial \theta} \int_{\mathbb{R}^n} T(\underset{\sim}{x}) f(\underset{\sim}{x}, \theta) dx.$
6. $0 < I_x(\theta) < +\infty$ όπου $I_x(\theta) = E\left[\frac{\partial}{\partial \theta} \log f(x, \theta) \right]^2$

ΘΕΩΡΗΜΑ

(Ανισότητα Cramer-Rao)

Έστω ότι οι συνθήκες κανονικότητας ισχύουν. Αν $T = T(x_1, \dots, x_n)$ είναι αμερόληπτος εκτιμητής της $g(\theta)$ τότε:

$$Var(T) \geq \frac{[g'(\theta)]^2}{I_x(\theta)}, \ \forall \theta \in \Theta$$

ΑΠΟΔΕΙΞΗ:

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

$$= \int \frac{\frac{\partial}{\partial \theta} f(\underset{\sim}{x}, \theta)}{f(\underset{\sim}{x}, \theta)} f(\underset{\sim}{x}, \theta) \partial \underset{\sim}{x} = \int \frac{\partial}{\partial \theta} f(\underset{\sim}{x}, \theta) \partial \underset{\sim}{x} = \frac{\partial}{\partial \theta} \int f(\underset{\sim}{x}, \theta) \partial \underset{\sim}{x} = \frac{\partial}{\partial \theta} 1 = 0 \quad (1)$$

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

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

$$Cov\left(\frac{\partial}{\partial \theta} \log f(\underset{\sim}{x}, \theta), T(\underset{\sim}{x}) \right) \quad \frac{Cov(W, S)}{E(WS) - E(W)E(S)}$$

$$= E\left[\left(\frac{\partial}{\partial \theta} \log f(\underset{\sim}{x}, \theta) \right) T(\underset{\sim}{x}) \right] - \cancel{E\left[\frac{\partial}{\partial \theta} \log f(\underset{\sim}{x}, \theta) \right]} E(T(\underset{\sim}{x}))$$

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

$$= \int T(\underset{\sim}{x}) \frac{\partial}{\partial\theta} f(\underset{\sim}{x},\theta)\, dx \overset{(5)}{=} \frac{\partial}{\partial\theta} \int T(\underset{\sim}{x}) f(\underset{\sim}{x},\theta)\, d\underset{\sim}{x} = \frac{\partial}{\partial\theta} E(T) = \frac{\partial}{\partial\theta} g(\theta) = g'(\theta)$$

$$\Rightarrow Cov\left(\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x},\theta), T(\underset{\sim}{x}) \right) = g'(\theta) \quad (3)$$

Από την ανισότητα των Cauchy Schwarz και (3)

$$\left(Cov^2(W,S) \le Var(W)\, Var(S) \right)$$

$$[g'(\theta)]^2 = Cov^2\left(\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x},\theta), T(\underset{\sim}{x}) \right) \le Var\left(\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x},\theta) \right) Var(T(\underset{\sim}{x})) =$$

$$\overset{(2)}{=} I_x(\theta) \cdot Var(T(\underset{\sim}{x})) \Rightarrow Var(T(\underset{\sim}{x})) \ge \frac{[g'(\theta)]^2}{I_x(\theta)}$$

<u>ΠΑΡΑΤΗΡΗΣΗ</u>: ① $K\Phi_{C-R} = \dfrac{[g'(\theta)]^2}{I_x(\theta)}$

ΚΦ: κάτω φράγμα
C-R: Cramer-Rao

② Η $I_x(\theta) = E\left[\frac{\partial}{\partial\theta} \log f(\underset{\sim}{x},\theta) \right]^2$ ονομάζεται <u>μέτρο πληροφορίας του Fisher</u> και (διαισθητικά) ποσοτικοποιεί την πληροφορία που περιέχεται στο τ.δ. $x_1,\dots,x_n$ για την άγνωστη παράμετρο

Αποδεικνύονται: → ① $I_x(\theta) = n I_x(\theta) \overset{\text{φ}}{=} n E\left[\frac{\partial}{\partial\theta} \log f(x,\theta) \right]^2$
(με πράξεις)

→ ② $I_x(\theta) = E\left[\frac{\partial}{\partial\theta} \log f(x,\theta) \right]^2 = -E\left[\frac{\partial^2}{\partial\theta^2} \log f(x,\theta) \right]$

③ $\boxed{K\Phi_{C-R} = \dfrac{[g'(\theta)]^2}{n I_x(\theta)}}$

④ Οι συνθήκες κανονικότητας ισχύουν τουλάχιστον για $B(n,\theta)$, $P(\lambda)$, $N(\mu,\sigma^2)$

ΕΦΑΡΜΟΓΗ ΑΝΙΣΟΤΗΤΑΣ Cramer-Rao

Βήμα 1 : Υπολογισμός $KΦ_{C-R} = \frac{[g'(\theta)]^2}{nI_x(\theta)}$

Βήμα 2 : Προσπαθούμε να βρούμε αμερόληπτο εκτιμητή $T=T(x_1,\dots,x_n)$ της $g(\theta)$ τέτοιον ώστε $Var(T)=KΦ_{C-R}$. Αν μπορώ να βρω ένα τέτοιο εκτιμητή, αυτός είναι ο ΑΟΕΔ της $g(\theta)$ (και είναι μοναδικός)

ΠΑΡΑΔΕΙΓΜΑ 1 Έστω Τ.Δ. $(x_1,\dots,x_n)$ από πληθυσμό με Poisson $P(\theta)$ με

$$p(x,\theta)=\frac{e^{-\theta}\theta^x}{x!}, \quad x=0,1,2,\dots, \quad \theta>0.$$

Να βρεθεί ΑΟΕΔ της θ.

$$KΦ_{C-R}=\frac{[g'(\theta)]^2}{nI_x(\theta)} \qquad g(\theta)=\theta, \quad g'(\theta)=1$$

$$I_x(\theta)\overset{\text{ορ}}{=}E\left(\frac{\partial}{\partial\theta}\log P(x,\theta)\right)^2=-E\left(\frac{\partial^2}{\partial\theta}\log P(x,\theta)\right)$$

$$\log P(x,\theta)=\log\frac{e^{-\theta}\theta^x}{x!}=-\theta+x\log\theta-\log x!$$

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

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

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

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

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

Επομένως $KΦ_{C-R}=\frac{[g'(\theta)]^2}{nI_x(\theta)}=\frac{1}{\frac{n}{\theta}}=\frac{\theta}{n}$

Πρέπει να βρούμε έναν εκτιμητή που να είναι αμερόληπτος με διακύμανση ίση με $KΦ_{C-R}$

Γνωρίζω ότι ο $\bar{X}$ αμερόληπτος της μέσης τιμής οποιασδήποτε πληθυσμού άρα και της Poisson (θ)

Επομένως $\bar{X}$ αμερ. της $\theta = E(x)$ $\quad Var(\bar{x}) = \frac{\sigma^2}{n}$

Αρκεί η $Var(\bar{X}) = K\Phi_{C-R}$

$$Var(\bar{X}) = Var\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n^2} Var\left(\sum X_i\right)$$

$$\underset{\text{ανεξ.}}{\overset{X_i}{=}} \frac{1}{n^2}\sum Var(X_i) = \frac{1}{n^2}\sum \theta = \frac{1}{n^2} n\theta = \frac{\theta}{n}$$

□

<u>Παράδειγμα 2:</u> Έστω τυχαίο δείγμα (τ.δ.) $X_1, \dots, X_n$ από πληθυσμό με κατανομή $B(n=1,p)$, $0<p<1$. Να βρεθεί ΑΟΕΔ για p.

$$B(1,p) \to p(x,p) = p^x(1-p)^{1-x}$$

> Όπου log ο νεπέριος δηλαδή ln

$$K\Phi_{C-R} = \frac{[g'(p)]^2}{n I_X(p)}, \quad g(p) = p$$

$$\log p(x,p) = \log\left[p^x(1-p)^{1-x}\right] = x\log p + (1-x)\log(1-p)$$

$$\frac{\partial}{\partial p}\log p(x,p) = \frac{x}{p} - \frac{1-x}{1-p}$$

$$\frac{\partial^2}{\partial p^2}\log p(x,p) = -\frac{x}{p^2} - \frac{1-x}{(1-p)^2}$$

$$I_X(p) = -E\left(\frac{\partial^2}{\partial p^2}\log p(x,p)\right) = \frac{E(x)}{p^2} + \frac{1-E(x)}{(1-p)^2} = \frac{p}{p^2} + \frac{1-p}{(1-p)^2} = \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)}$$

Επομένως $K\Phi_{C-R} = \frac{1}{n I_X(p)} = \frac{p(1-p)}{n}$

Αναζητώ αμερόληπτο της p με διασπορά $= K\Phi_{C-R}$

Ο $\bar{X}$ αμερόληπτος της p

και

$$Var(\bar{x}) = \frac{1}{n^2}\sum Var(X_i) = \frac{1}{n^2}\sum p(1-p) = \frac{p(1-p)}{n} = K\Phi_{C-R}$$

Άρα $\bar{X}$ ΑΟΕΔ του p

□

Ερώτημα: Ποια η μορφή ενός αμερόληπτου εκτιμητή ώστε η διακύμανσή του να πετυχαίνει το $K\Phi_{C-R}$

ΠΡΟΤΑΣΗ Έστω οι συνθήκες κανονικότητας ικανοποιούνται και ο $U=U(\underline{x})=U(x_1,\dots,x_n)$ είναι αμερόληπτος της $g(\theta)$. Τότε:

$$Var(U)=K\Phi_{C-R} \Leftrightarrow U=g(\theta)+a(\theta)\,W, \text{ όπου } W=W(x_1,\dots,x_n) \overset{\text{op}}{=} \frac{\partial}{\partial\theta}\log f(\underline{x},\theta) \underset{f(\underline{x},\theta)=\prod f(x_i;\theta)}{=} \sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(x_i,\theta)$$

ΑΠΟΔΕΙΞΗ

"$\Leftarrow$" Έστω $U=g(\theta)+a(\theta)W$. Αφού U γραμμική συνάρτηση της W $\rho(U,W)=\pm 1$ Άρα $\rho^2(U,W)=1$

$$\text{Αλλά } \rho(U,W)\overset{\text{op.}}{=}\frac{Cov(U,W)}{\sqrt{Var U}\sqrt{Var(W)}}$$

$$\text{Άρα } \rho^2(U,W)=1 \Rightarrow Cov^2(U,W)=Var(U)\,Var(W)$$

$$\Rightarrow Cov^2\left(U,\frac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right)=Var(U)\,Var\left(\frac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right)$$

Από απόδειξη ανισότητας Cramer-Rao

$$Cov^2\left(U,\frac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right)=[g'(\theta)]^2$$

και

$$Var\left(\frac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right)=I_{\underline{x}}(\theta)$$

$$\text{Άρα } [g'(\theta)]^2=Var(U)\,I_x(\theta)$$

$$\Rightarrow Var(U)=\frac{[g'(\theta)]^2}{I_x(\theta)}=K\Phi_{C-R}$$

"$\Rightarrow$" Αποδεικνύεται ακολουθώντας αντίστροφα βήματα. $\square$

Παρατηρηση: Για ένα αμερόληπτο εκτιμητή U της $g(\theta)$ το $K\Phi_{C-R}$ πετυχαίνεται εαν-ν

$$\sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(x_i,\theta) = W = K(\theta,n)[U(\underset{\sim}{x}) - g(\theta)]$$

Δηλαδή Αν U αμερόληπτος της $g(\theta)$ και

$$W = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(x_i,\theta) = K(\theta,n)[U - g(\theta)]$$ τότε η $Var(U) = K\Phi_{C-R}$

και επομένως U ΑΟΕΔ της $g(\theta)$

Παραδειγμα: 1 (Συνέχεια) Έστω τ.δ. $X_1,\dots,X_n$ από Poisson(θ),

$$p(x,\theta) = \frac{e^{-\theta}\theta^x}{x!}$$

Ζητείται ΑΟΕΔ της θ.

$$W = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(x_i,\theta) = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log \frac{e^{-\theta}\theta^{x_i}}{x_i!} =$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial\theta}(-\theta + x_i \log\theta - \log x_i!) = \sum_{i=1}^{n}\left(-1 + \frac{x_i}{\theta}\right) = \frac{1}{\theta}\sum(x_i - \theta) = \frac{1}{\theta}(\sum x_i - n\theta) =$$

$$= \frac{1}{\theta}(n\bar{x} - n\theta) = \frac{n}{\theta}(\bar{x} - \theta) \quad (*)$$

Αφού το αμερόληπτο $\bar{X}$ της θ ικανοποιεί $(*)$ είναι ο ΑΟΕΔ της θ.

$\square$

Παραδειγμα: 2 (Συνέχεια). Έστω τ.δ. $X_1,\dots,X_n$ από $B(n=1,p)$.

Να βρεθεί ΑΟΕΔ της p.

$$p(x,p) = p^x(1-p)^{1-x}$$

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

$$= \sum_{i=1}^{n} \frac{\partial}{\partial p}\left[x_i \log p + (1-x_i)\log(1-p)\right] = \sum_{i=1}^{n}\left(\frac{x_i}{p} - \frac{1-x_i}{1-p}\right) =$$

$$= \sum_{i=1}^{n} \frac{x_i - px_i - p + px_i}{p(1-p)} = \sum_{i=1}^{n} \frac{x_i - p}{p(1-p)} = \frac{1}{p(1-p)} \sum (x_i - p) =$$

$$= \frac{1}{p(1-p)}(n\bar{x} - np) = \frac{n}{p(1-p)}(\bar{x} - p)$$

$\bar{X}$: ΑΟΕΔ εκτιμητής της p.

□

Παραδειγμα:

Έστω τ.δ. $X_1, \ldots, X_n$ από πληθυσμό με κατανομή $N(\mu, \theta)$, μ = γνωστό.

$$\left(f(x,\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x-\mu)^2}{2\theta}} \right)$$

Να βρεθεί ΑΟΕΔ της θ με την (i) Ανισότητα Cramer-Rao

(ii) Με την Παρατήρηση.