• $X_1,\dots,X_n \to f(x,\theta),\ \theta\in\Theta$

σ.σ. $T=T(x_1,\dots,x_n)$

$MT\Sigma(T, g(\theta)) = Var(T(x)) - (E(T(x)) - g(\theta))^2$ (*)

Αμερόληπτος Εκτιμητής:

Ορισμός: ένας εκτιμητής $T=T(\underline{x})=T(x_1,\dots,x_n)$ λέγεται αμερόληπτος της $g(\theta)$, $\theta\in\Theta$, αν και μόνο αν $E(T(x))=g(\theta)$, $\theta\in\Theta$. Στην αντίθετη περίπτωση ο $T=T(\underline{x})$ λέγεται μεροληπτικός.

Διαισθητικά: Έστω T_1, T_2 αμερόληπτοι της $g(\theta)$ τότε $E(T_1(x))$

Ορισμός: Αν $T=T(\underline{x})$ ένας εκτιμητής της $g(\theta)$ η μεροληψία του εκτιμητή T συμβολίζεται με $b(T,\theta)$ και ορίζεται: $b(T,\theta)=E[T(x)]-g(\theta)$, $\theta\in\Theta$.

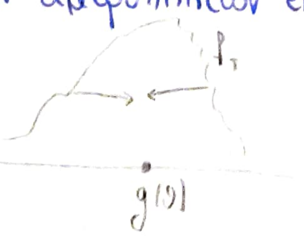
Πρόταση: Αν ο εκτιμητής $T=T(\underline{x})=T(x_1,\dots,x_n)$ είναι αμερόληπτος της $g(\theta)$ $(E(T(x))=g(\theta))$ τότε $MT\Sigma(T,g(\theta))=Var(T)$

Απόδ: Προφανής από την (*) και $E(T(\underline{x}))=g(\theta)$

Αμερόληπτος ομοιόμορφα ελάχιστης διακύμανσης εκτιμητής (ΑΟΕΔ)

Ορισμός: Έστω τ.δ. $x=(x_1,\dots,x_n)$ από πληθυσμό με κατανομή $f(x,\theta)$, $\theta\in\Theta$ και έστω $T=T(\underline{x})=T(x_1,\dots,x_n)$ ένας εκτιμητής της $g(\theta)$. Ο T λέγεται ΑΟΕΔ της $g(\theta)$ αν: (i) Είναι αμερόληπτος της $g(\theta)$, και (ii) έχει την μικρότερη διακύμανση μεταξύ όλων των άλλων αμερόληπτων εκτιμητών της $g(\theta)$, $\theta\in\Theta$.

Διαισθητικά:



Θεώρημα Μοναδικότητας ΑΟΕΔ εκτιμητή: Αν υπάρχει ΑΟΕΔ εκτιμητής $T=T(\underline{x})$ της $g(\theta)$ ο T είναι μοναδικός.

Παράδειγμα: Έστω τ.δ. $x_1, x_2, \dots, x_n$ από πληθυσμό με κατανομή με μέση τιμή μ και διασπορά σ^2 ($E(x_i)=\mu$, $Var(x_i)=\sigma^2$, $i=1,\dots,n$). Έστω ο δειγματικός μέσος $\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i$ και η δειγματική διασπορά $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$. Να δειχθεί ότι:

(α) $\bar{X}$ αμερόληπτος εκτιμητής του μ.

(β) $Var(\bar{x})=\sigma^2/n$

(γ) S^2 αμερόληπτος εκτιμητής του σ^2.

Απόδ.

(α) $E(\bar{x})=E\left(\frac{1}{n}\sum x_i\right)=\frac{1}{n}\sum E(x_i)=\frac{1}{n}\sum\mu=\frac{1}{n}n\mu=\mu$

(β) $Var(\bar{x})=Var\left(\frac{1}{n}\sum x_i\right)=\frac{1}{n^2}Var\left(\sum x_i\right)\xrightarrow[\text{(από το πρώτο μάθημα)}]{x_i \text{ ανεξ.}}\frac{1}{n^2}\sum Var\, x_i=\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2=\frac{1}{n^2}n\sigma^2=\sigma^2/n$

(γ) $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\xrightarrow{\text{πράξεις}}\frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right)$

$$E(S^2)=E\left[\frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right)\right]=\frac{1}{n-1}\left\{\sum_{i=1}^{n}E(x_i^2)-nE(\bar{x}^2)\right\}\xrightarrow{Var(W)=EW^2-(EW)^2}$$

$$=\frac{1}{n-1}\sum_{i=1}^{n}\left\{Var(x_i)+(E(x_i))^2\right\}-n\left[Var(\bar{x})+(E\bar{x})^2\right]\Big\}=\frac{1}{n-1}\left\{\sum_{i=1}^{n}\left[\sigma^2+\mu^2\right]-n\left[\frac{\sigma^2}{n}+\mu^2\right]\right\}=$$

$$=\frac{1}{n-1}\left\{n\sigma^2+\cancel{n\mu^2}-\sigma^2-\cancel{n\mu^2}\right\}=\frac{(\cancel{n-1})\sigma^2}{\cancel{n-1}}=\sigma^2.$$

Ερώτημα: Πως κατασκευάζω ΑΟΕΔ εκτιμητή;

1η Μέθοδος: Εφαρμογή Ανισότητας Cramer-Rao.

Συνθήκες κανονικότητας: Έστω τ.δ. $x_1,\dots,x_n$ από πληθυσμό με κατανομή $f(x,\theta)$, $\theta\in\Theta\subseteq\mathbb{R}$.

Έστω $f(\underline{x},\theta)$ η από κοινού κατανομή των $\underline{x}=(x_1,\dots,x_n)$ και επειδή x_i ανεξάρτητες $f(\underline{x},\theta)=\prod_{i=1}^{n}f(x_i,\theta)$.

$\theta\in\Theta\subseteq\mathbb{R}$. Οι συνθήκες κανονικότητας είναι

(Σ1) Ο Θ είναι ένα ανοικτό υποσύνολο του $\mathbb{R}$.

(Σ2) Το σύνολο $S_\theta=\{(x_1,\dots,x_n)=\underline{x}\in\mathbb{R}: f(\underline{x},\theta)>0\}$ είναι ανεξάρτητο της παραμέτρου θ.

(I_3) Η $\frac{\partial}{\partial\theta}\log f(\underline{x},\theta)$ υπάρχει και είναι πεπερασμένη, $\underline{x}\in S_n$, $\theta\in\Theta$.

(I_4) $\int \frac{\partial}{\partial\theta} f(\underline{x},\theta)\,d\underline{x} \overset{\text{από κάτω}}{=} \frac{\partial}{\partial\theta}\int f(\underline{x},\theta)\,d\underline{x}\ (=0)$, $\forall\theta\in\Theta$

(I_5) $\int T(\underline{x})\frac{\partial}{\partial\theta} f(\underline{x},\theta)\,d\underline{x} = \frac{\partial}{\partial\theta}\int T(\underline{x}) f(\underline{x},\theta)\,d\underline{x}$, $T(\underline{x})$ στατιστική συνάρτηση.

(I_6) Αν $I_{\underline{x}}^{F}(\theta) = E\left[\frac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right]^2$, $\forall\theta\in\Theta$ τότε $0\leq I_{\underline{x}}^{F}(\theta) < +\infty$

Θεώρημα (Ανισότητα Cramer-Rao) Έστω ότι οι συνθήκες κανονικότητας ικανοποιούνται.
Αν $T = T(\underline{x}) = T(x_1,\dots,x_n)$ είναι αμερόληπτος της $g(\theta)$ τότε $Var(T) \geq \dfrac{[g'(\theta)]^2}{I_{\underline{x}}^{F}(\theta)}$, $\theta\in\Theta$

ίσως θέμα για τα διαστήματα

Απόδ.

$$E\left(\frac{\partial}{\partial\theta}\log(f(\underline{x},\theta))\right) = \int\left(\frac{\partial}{\partial\theta}\log f(x,\theta)\right)f(\underline{x},\theta)\,dx = \int\frac{\partial}{\partial\theta}f(x,\theta)\,d\underline{x} \overset{(I_4)}{=\!=} \frac{\partial}{\partial\theta}\int f(\underline{x},\theta)\,dx = \frac{\partial}{\partial\theta}1 \overset{(1)}{=} \underset{\theta\in\Theta}{0}$$

$$Var\left(\frac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right) = E\left(\frac{\partial}{\partial\theta}\log f(x,\theta)\right)^2 - \left\{E\left(\frac{\partial}{\partial\theta}\log f(x,\theta)\right)\right\}^2 \overset{(I_6),(1)}{=\!=} I_{\underline{x}}^{F}(\theta) - 0^2 = I_{\underline{x}}^{F}(\theta),\ \theta\in\Theta \quad (2)$$

$$Cov\left(\frac{\partial}{\partial\theta}\log f(\underline{x},\theta),\ T(\underline{x})\right) \overset{cov(w,z)=E(wz)-E(w)E(z)}{=\!=\!=} E\left[T(\underline{x})\frac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right] - E[T(\underline{x})]\,E\left[\frac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right]$$

$$\overset{(1)}{=\!=} E\left[T(x)\frac{\partial}{\partial\theta}\log f(x,\theta)\right] = \int\left(T(x)\frac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right)f(\underline{x},\theta)\,d\underline{x} = \int T(\underline{x})\frac{\frac{\partial}{\partial\theta}f(\underline{x},\theta)}{f(\underline{x},\theta)}f(\underline{x},\theta)\,d\underline{x}$$

είναι αμερόληπτη της $g(\theta)$

$$= \int T(\underline{x})\frac{\partial}{\partial\theta}f(\underline{x},\theta)\,dx \overset{(I_5)}{=\!=} \frac{\partial}{\partial\theta}\int T(\underline{x})f(\underline{x},\theta)\,d\underline{x} = \frac{\partial}{\partial\theta}E[T(\underline{x})] = \frac{\partial}{\partial\theta}g(\theta) = g'(\theta) \quad (3)$$

Ισχύει: $cov^2(z,w) \leq Var(z)Var(w)$

Από την (3): $cov^2\left(\frac{\partial}{\partial\theta}\log f(\underline{x},\theta),\ T(\underline{x})\right) \leq Var\left(\frac{\partial}{\partial\theta}\log f(x,\theta)\right).Var(T(\underline{x}))$

$$\overset{(3),(2)}{\Longrightarrow} [g'(\theta)]^2 \leq I_{\underline{x}}^{F}(\theta).Var(T(\underline{x})) \Rightarrow Var(T(\underline{x})) \geq \frac{[g'(\theta)]^2}{I_{\underline{x}}^{F}(\theta)}$$

αντιστρέφω το $I_{\underline{x}}^{F}(\theta)$ κι έτσι γίνεται μεγαλύτερο

Παρατηρήσεις

(α) $ΚΦ_{C-R} = \dfrac{[g'(\theta)]^2}{I_{\underline{x}}^F(\theta)}$

κάτω φράγμα της ανισότητας Cramer Rao

(β) Το $I_{\underline{x}}^F(\theta)$ ονομάζεται μέτρο πληροφορίας του Fischer και εκφράζει την πληροφορία που υπάρχει στα δεδομένα $\underline{x}=(x_1,\dots,x_n)$ για την άγνωστη παράμετρο θ.

(γ) Αν $x_1,\dots,x_n$ ανεξάρτητες και ισόνομες με κοινή κατανομή $f(x,\theta)$, τότε αποδεικνύεται ότι $I_{\underline{x}}^F(\theta) = n\, I_x^F(\theta)$ όπως $I_x^F(\theta) = E\left[\dfrac{\partial}{\partial\theta}\log f(x,\theta)\right]^2$

(δ) Αποδεικνύεται $I_x^F(\theta) = E\left[\dfrac{\partial}{\partial\theta}\log f(x,\theta)\right]^2 = -E\left[\dfrac{\partial^2}{\partial\theta^2}\log f(x,\theta)\right]$

(ε) $ΚΦ_{C-R} = \dfrac{[g'(\theta)]^2}{n I_x^F(\theta)}$

Για να εφαρμόσω την μέθοδο Cramer-Rao ακολουθώ τα εξής βήματα

Βήμα 1: Υπολογίζω το $ΚΦ_{C-R}$

Βήμα 2: Προσπαθώ να βρω αμερόληπτο $T=T(\underline{x})$ της $g(\theta)$ με $Var(T) = ΚΦ_{C-R}$

Τότε ο T είναι ΑΟΕΔ.

Παράδειγμα: Έστω τ.δ. $x_1,\dots,x_n$ από πληθυσμό με κατανομή $P(\theta)$ και σ.π.: $p(x,\theta) = \dfrac{e^{-\theta}\theta^x}{x!}$, $\theta>0$, $x=0,1,\dots$

Να κατασκευαστεί ΑΟΕΔ εκτιμητής της $g(\theta)=\theta$.

Απόδ.

Βήμα 1: $ΚΦ_{C-R} = \dfrac{[g'(\theta)]^2}{n I_x^F(\theta)} = \dfrac{1}{n I_x^F(\theta)}$

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

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

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

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

$$I_X^F(\theta) = -E\left(\frac{\partial^2}{\partial\theta^2}\log p(x,\theta)\right) = -E\left(-\frac{x}{\theta^2}\right) = \frac{E(x)}{\theta^2} = \frac{\theta}{\theta^2} = \frac{1}{\theta} > 0$$ (γενικά πρέπει να είναι ≥ 0) — από τυπολόγιο

Άρα $\kappa\phi_{C-R} = \frac{1}{n(1/\theta)} = \frac{\theta}{n}$

Βήμα 2: Αρκεί να βρω $T = T(\underline{x})$ αμερόληπτο του θ και με $Var(T) = \kappa\phi_{C-R} = \frac{\theta}{n}$.

Πως θα βρω τέτοιον T; $\rightarrow \bar{X}$ συνήθως $\rightarrow$ Δοκιμάζω!

Δοκιμάζω το $\bar{x}$ γιατί ξέρω ότι εκτιμά αμερόληπτα την μέση τιμή του πληθυσμού που για την Poisson είναι θ.

Θεωρώ τον $T = T(\underline{X}) = \bar{x}$. Προφανώς ο $\bar{x}$ αμερόληπτος της $g(\theta) = \theta$ και $Var(\bar{x}) = \frac{1}{n^2}\sum Var(x_i) =$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\theta = \frac{1}{n^2}n\theta = \frac{\theta}{n} \Rightarrow Var(\bar{x}) = \frac{\theta}{n} = \kappa\phi_{C-R}.$$

Άρα $\bar{x}$ ΑΟΕΔ

Άσκηση: Έστω τ.δ. $x_1,\dots,x_n$ από πληθυσμό με κατανομή Bernoulli με σ.π.π.

$p(x,\theta) = \theta^x(1-\theta)^{1-x}$, $x = 0, 1$, $0 < \theta < 1$. Να κατασκευαστεί ΑΟΕΔ της $g(\theta) = \theta$.

Υπενθύμιση: Bernoulli $= B(1,\theta)$.