23/3/2017

(μια σπό αυτές π.ο. ανεξάρτητο του θ)

Αν η $f(x,\theta)$ (σ.π.π. ή σ.π) ικανοποιεί τις συνθήκες ομαλότητας και $T(x)$ είναι αμερόληπτος εκτιμητής της $g(\theta)$ (που ικανοποιεί μια επιπλέον συνθήκη)
Τότε $T(x)$ είναι Α.Ο.Ε.Δ της $g(\theta)$ αν η διακύμανση είναι ίση με το

Κ.Φ. $C\text{-}R = \frac{(g'(\theta))^2}{I_x(\theta)}$, $I_x(\theta) = nI_{(\theta)} = E\left[\frac{\partial}{\partial\theta}\ln f(x,\theta)\right]^2$

ΕΚΘΕΤΙΚΗ ΟΙΚΟΓΕΝΕΙΑ ΚΑΤΑΝΟΜΩΝ: Είναι μια ευρεία οικογένεια κατανομών η οποία περιλαμβάνει πολλές κλασσικές κατανομές και ορίζεται ως εξής:

ΟΡΙΣΜΟΣ: Η κατανομή του τ.δ. $X=(X_1,\dots,X_n)$ $n\geq 1$ ανήκει στη μονοπαραμετρική εκθετική οικογένεια κατανομών αν η σ.π.π ή σ.π του $\underset{\sim}{X}$, $f(\underset{\sim}{x},\theta)$:
$\theta \in \Theta \subseteq \mathbb{R}$ αν

$$f(\underset{\sim}{x},\theta) = e^{A(\theta)+B(\underset{\sim}{x})+C(\theta)D(\underset{\sim}{x})}, \quad x\in S$$

με S ανεξάρτητο του θ.

ή ισοδύναμα

$$f(\underset{\sim}{x},\theta) = h(\underset{\sim}{x})\,\omega(\theta)\, e^{C(\theta)D(\underset{\sim}{x})}, \quad x\in S$$

με $h(\underset{\sim}{x}) = e^{B(x)}$, $A(\theta) = \omega(\theta)$

ΠΡΟΤΑΣΗ: Εάν η οικογένεια κατανομών του X είναι η Ε.Ο.Κ. και $C(\theta)$ έχει συνεχή μη μηδενική παράγωγο στο Θ, τότε ισχύουν οι συνθήκες ομαλότητας (σελ. 15) Το αντίστροφο δεν ισχύει

ΠΑΡΑΤΗΡΗΣΗ: Έχουμε δείξει ότι η ισότητα στη Cramer-Rao επιτυγχάνεται όταν

$$\frac{\partial \ln f(\underset{\sim}{x},\theta)}{\partial\theta} = k(\theta,n)\left[T(\underset{\sim}{x})-g(\theta)\right]$$

Πλέον επιλέγω εκτιμητή που ικανοποιεί αυτή τη σχέση

Παραδείγματα

① $(x_1, x_2, \dots, x_n)$ είναι τυχαίο δείγμα από πληθυσμό με $N(\mu, \theta)$, μ γνωστό.
Να βρεθεί ΑΟΕΔ εκτιμητής της θ.

1ος τρόπος: $f(x,\theta) = \dfrac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2\theta}(x-\mu)^2}$, $x \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\theta > 0$

$$f(\underline{x}, \theta) = \prod_{i=1}^{n} f(x_i, \theta) = \underbrace{(2\pi\theta)^{-n/2}}_{w(\theta)} e^{\underbrace{\left(-\frac{1}{2\theta}\right)}_{C(\theta)} \underbrace{\sum_{i=1}^{n} (x_i-\mu)^2}_{D(x)}}$$

$$= w(\theta)\, e^{C(\theta) D(x)}$$

$$K\Phi_{C-R} = \frac{(g'(\theta))^2}{nI_x(\theta)} = \frac{(\theta')^2}{nI_x(\theta)} = \frac{1}{nI_x(\theta)} = \boxed{\frac{2\theta^2}{n}}$$

Βρίσκαμε το $K\Phi_{C-R}$

$$I_x(\theta) = E\left(\frac{\partial}{\partial\theta} \ln f(x,\theta)\right)^2$$

$$\ln f(x,\theta) = -\frac{1}{2}\ln(2\pi\theta) - \frac{1}{2\theta}(x-\mu)^2$$

$$\frac{\partial}{\partial\theta} \ln f(x,\theta) = -\frac{1}{2\theta} + \frac{1}{2\theta^2}(x-\mu)^2$$

E: αναμενόμενη τιμή

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

$$\frac{\partial^2}{\partial\theta^2} = \frac{1}{2\theta^2} - \frac{(x-\mu)^2}{\theta^3}$$

$$I_x(\theta) = -E\left[\frac{1}{2\theta^2} - \frac{(x-\mu)^2}{\theta^3}\right] = -\frac{1}{2\theta^2} + \frac{E(x-\mu)^2}{\theta^3} = -\frac{1}{2\theta^2} + \frac{Var X}{\theta^3} = \frac{1}{2\theta^2}$$

$$E\left[\sum_{i=1}^{n} (x_i-\mu)^2\right] = \sum_{i=1}^{n} E(x_i-\mu)^2 = \sum_{i=1}^{n} Var X_i = \sum_{i=1}^{n} \theta = n\theta$$

$$\boxed{E\,\frac{\sum (x_i-\mu)^2}{n} = \theta}$$

$$Var\ \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$$

$$X_i \sim N(\mu,\theta) \Rightarrow \frac{X_i-\mu}{\sqrt{\theta}} \sim N(0,1)$$

$$\frac{(X_i-\mu)^2}{\theta} \sim X_1^2 \Rightarrow \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\theta} \sim X_n^2$$

$$Var\left[\frac{\sum(x_i-\mu)^2}{\theta}\right] = 2n$$

Άρα $Var\sum(x_i-\mu)^2 = 2n\theta$

Όταν μια σταθερά βγαίνει έξω από τη διακύμανση υψώνεται στο τετράγωνο

$$Var\ \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} = \frac{1}{n^2}2n\theta^2 = \frac{2\theta^2}{n}$$

$2^{ος}$ τρόπος: Γνωρίζω ότι η ισότητα στο C-R επιτυγχάνεται εάν:

$$\frac{\partial}{\partial\theta}\ln f(\underline{x},\theta) = k(\theta,n)\left(T(\underline{x})-g(\theta)\right)$$

$$\prod_{i=1}^{n} f(x_i,\theta) = f(\underline{x},\theta) = (2\pi\theta)^{-n/2}\, e^{-\frac{1}{2\theta}\sum(x_i-\mu)^2}$$

$$\ln f(\underline{x},\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\theta - \frac{1}{2\theta}\sum(x_i-\mu)^2$$

$$\frac{\partial}{\partial\theta}\ln f(\underline{x},\theta) = -\frac{n}{2}\frac{1}{\theta} + \frac{1}{2\theta^2}\boxed{\sum(x_i-\mu)^2}$$

$$= \frac{n}{2\theta^2}\left[\frac{\sum(x_i-\mu)^2}{n} - \theta\right]$$

(2) Ελέγξτε εάν η στατιστική συνάρτηση $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ είναι αμερόληπτος εκτιμητής του θ και να υπολογίσετε τη διακύμανσή της. Τι παρατηρείτε;

Γνωρίζω ότι $\frac{(n-1)S^2}{\sigma^2} \sim X_{n-1}^2$

Άρα εδώ $\frac{(n-1)S^2}{\theta} \sim \chi^2_{n-1}$

$$E\left[\frac{(n-1)S^2}{\theta}\right] = (n-1) \Rightarrow E(S^2) = \theta$$

$$Var\left[\frac{(n-1)S^2}{\theta}\right] = 2(n-1) \Rightarrow Var S^2 = \frac{2\theta^2}{n-1}$$

$$T = \frac{\Sigma(x_i-\mu)^2}{n}$$

$$Var T = \frac{2\theta^2}{n}$$

$Var T < Var S^2$ όπως αναμενόταν

ΟΡΙΣΜΟΣ (ΑΠΟΤΕΛΕΣΜΑΤΙΚΟΣ ΕΚΤΙΜΗΤΗΣ)

Ο εκτιμητής $U(x)$ της $g(\theta)$ λέγεται αποτελεσματικός εάν ο $U(x)$ είναι ο ΑΟΕΔ της $g(\theta)$.

ΟΡΙΣΜΟΣ (ΣΧΕΤΙΚΗ ΑΠΟΤΕΛΕΣΜΑΤΙΚΟΤΗΤΑ)

Έστω $U(x)$ αποτελεσματικός εκτιμητής της $g(\theta)$. Έστω $T(x)$ αμερόληπτος εκτιμητής της $g(\theta)$ με $Var T(x) < +\infty$. Η σχετική αποτελεσματικότητα του $T(x)$ ορίζεται ως

$$0 < \frac{Var\, U(x)}{Var\, T(x)} < 1$$

ΠΑΡΑΔΕΙΓΜΑ

Έστω X_1, X_2, X_3 από την Poisson(θ). Νδο $\bar{X}$ αποτελεσματικός της θ και να βρεθεί η σχετική αποτελεσματικότητα του $T = \frac{X_1 + 2X_2 + 3X_3}{6}$ ως προς $U = \bar{X}$, $f(x,\theta) = \frac{\theta^x e^{-\theta}}{x!}$, $x = 0, 1, 2, \ldots$

$$f(\underline{x},\theta) = \frac{\theta^{\Sigma x_i} e^{-n\theta}}{\prod_{i=1}^{n} x_i!} = \underset{e^{-n\theta}}{w(\theta)}\, \underset{\frac{1}{\prod_{i=1}^{n} x_i!}}{h(x)}\, e^{\ln \theta^{\Sigma x_i}} = w(\theta) h(x) e^{\Sigma x_i \ln\theta}$$

$$\frac{\partial}{\partial\theta} \ln f(\underline{x},\theta) = k(\theta,n)\left(U(\underline{x}) - g(\theta)\right)$$

$$\ln f(\underline{x},\theta) = \Sigma x_i \ln\theta - n\theta - \ln\prod x_i$$

$$\frac{\partial}{\partial\theta} \ln f(\underline{x},\theta) = \frac{\sum_{i=1}^{n} x_i}{\theta} - n$$

$$= \frac{n}{\theta}\left[\frac{\Sigma x_i}{n} - \theta\right]$$

$$Var \frac{\Sigma X_i}{n} = \frac{1}{n^2} \Sigma Var X_i = \frac{1}{n^2} \Sigma \theta = \frac{n\theta}{n^2} = \frac{\theta}{n} = \frac{\theta}{3}$$

$$Var T = Var \frac{X_1 + 2X_2 + 3X_3}{6}$$

$$= \frac{1}{36}(\theta + 4\theta + 9\theta) = \frac{7\theta}{18}$$

$$\Sigma\chi\varepsilon\tau.\ A\pi o\tau\varepsilon\lambda = \frac{\theta/3}{7\theta/18} = 85\%$$

③ Έστω τ.μ. $X \sim N(\theta, 1)$

$$f(x,\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}}, \quad x \in \mathbb{R},\ \theta \in \mathbb{R}$$

Νδο ανήκει στην εκθετική οικογένεια κατανομών (Ε.Ο.Κ.)

$$f(x,\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2 - 2x\theta + \theta^2}{2}} = \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}_{h(x)} \underbrace{e^{-\frac{\theta^2}{2}}}_{W(\theta)} \underbrace{e^{\frac{2x\theta}{2}}}_{e^{C(\theta)D(x)}}$$

④ Έστω x ακολουθεί διωνυμική με $X \sim B(n,\theta)$

$$f(x,\theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, \quad x = 0,1,\dots,n, \quad n \text{ γνωστό}$$

$$= \underbrace{\binom{n}{x}}_{h(x)} \underbrace{\left(\frac{\theta}{1-\theta}\right)^x}_{e^{\ln\left(\frac{\theta}{1-\theta}\right)^x}} \underbrace{(1-\theta)^n}_{W(\theta)}$$

$$= h(x)\, e^{x \ln \frac{\theta}{1-\theta}}\, W(\theta)$$

⑤ $X_1, X_2, X_3, \dots, X_n$ τ.δ. από την ομοιόμορφη $U(0,\theta)$

$$f(x,\theta)=\frac{1}{\theta}, \quad 0<x<\theta.$$

Το πεδίο ορισμού εξαρτάται από το θ. Άρα δεν ανήκει στην Ε.Ο.Κ.

Ούτε η Cauchy ανήκει

ΕΠΑΡΚΕΙΑ: Επιθυμητό θα ήταν να εκφράσουμε το τ.δ. ως ένα βαθμωτό τυχαίο μέγεθος, με τρόπο τέτοιο ώστε να μη χάνει καμία πληροφορία που περιέχεται σε αυτό και αφορά το χαρακτηριστικό μέγεθος του πληθυσμού που μελετάμε. Το σκοπό αυτό επιτελεί η επαρκής στατιστική συνάρτηση.

Διαισθητικά: $T(\underline{x})$ επαρκής για την θ αν Ποσό πληροφορίας $(x_1, \dots, x_n)$ = Ποσό πληροφορίας $T(\underline{x}) = T(x_1, \dots, x_n)$

ΟΡΙΣΜΟΣ: Έστω $x_1, .., x_n$ τ.δ. από έναν πληθυσμό με σ.π.π. ή σ.π. $f(x,\theta)$. Η στατιστική συνάρτηση $T(\underline{x})$ λέγεται επαρκής για το $\theta \in \Theta$ αν η δεσμευμένη κατανομή του X δοθέντος $T=t$ δεν εξαρτάται από το θ για κάθε τιμή t της T για την οποία μπορεί να οριστεί η δεσμευμένη κατανομή.

ΠΑΡΑΔΕΙΓΜΑ

Έστω $x_1, \dots, x_n$ τ.δ. από τη διωνυμική $B(1,\theta)$, $f(x,\theta)=\theta^{x}(1-\theta)^{1-x}$, $x=0,1$

$T = \sum_{i=1}^{3} X_i$

τότε η στατιστική συνάρτηση είναι επαρκής.

τιμές

δεσμευμένες πιθανότητες $P(\underset{\sim}{X}=x \mid T=t)$

$\underset{\sim}{X} = (x_1, x_2, x_3)$

0	(0,0,0)	1
1	(1,0,0)	1/3
	(0,1,0)	1/3
	(0,0,1)	1/3
2	(1,1,0)	1/3
	(0,1,1)	1/3
	(1,0,1)	1/3
3	(1,1,1)	1/3
		1

Παραγοντικό Θεώρημα Neyman-Fisher

Έστω το $\underline{X}$ έχει πυκνότητα $f(x,\theta)$, $\theta \in \Theta$ τότε η σ.σ. $T(x)$ είναι επαρκής αν-ν υπάρχουν συναρτήσεις $q(\cdot,\cdot)$ και $h(\cdot,\cdot)$ τ.ώ.

$$f(x,\theta) = q(T(x),\theta)\, h(x) \quad \forall x, \forall \theta$$

Παράδειγμα

$X_1, \dots, X_K$ τ.δ. από τη διωνυμική με $B(n,\theta)$, $0<\theta<1$.

$$f(\underline{x},\theta) = \prod_{i=1}^{K} f(x_i;\theta) = \prod_{i=1}^{K} \binom{n}{x_i} \theta^{x_i} (1-\theta)^{n-x_i}$$

$$= \prod_{i=1}^{K} \binom{n}{x_i} \theta^{\Sigma x_i} (1-\theta)^{nK} (1-\theta)^{-\Sigma x_i}$$

$$= \prod_{i=1}^{n} \binom{n}{x_i} \underbrace{(1-\theta)^{nK} \left(\frac{\theta}{1-\theta}\right)^{\Sigma x_i}}_{q(\Sigma x_i,\theta)}$$

Παράδειγμα

$X_1, X_2, \dots, X_n$ τ.δ. από την ομοιόμορφη στο $U(0,\theta)$. $f(x_i/\theta) = \frac{1}{\theta}$, $0<x_i<\theta$

$$f(\underline{x},\theta) = \prod_{i=1}^{n} \frac{1}{\theta} = \frac{1}{\theta^n} I_{(x_{(n)},+\infty)}(\theta)\, I_{(0,x_{(n)})}(x_{(1)})$$

πρέπει να ισχύουν αυτά:
$$\begin{cases} 0<x_1<\theta \\ 0<x_2\,\theta \\ 0<x_n<\theta \end{cases}$$

$$0 < \underbrace{\min x_i}_{\downarrow} < \underbrace{\max x_i}_{\downarrow} < \theta$$

$$0 < X_{(1)} < X_{(n)} < \theta$$

$$I_A = \begin{cases} 1 & x\in A \\ 0 & \text{αλλού} \end{cases}$$