ΤΡΙΤΗ 9 ΜΑΡΤΙΟΥ 2021 - ΣΤΑΤΙΣΤΙΚΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ

Παράδειγμα (Bernoulli) → από το προηγούμενο μάθημα $\equiv B(1,\theta)$

τ.δ. $X_1,\dots,X_n$ με σ.π. $p(x,\theta)=\theta^x(1-\theta)^{1-x}$, $0<\theta<1$, $x=0,1$. ΑΟΕΑ της θ.

Λύση

1ο Βήμα: Εύρεση κφ C-R $=\dfrac{[g'(\theta)]^2}{n I_X^F(\theta)}$, $g(\theta)=\theta$

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

$$\log p(x,\theta)=x\log\theta+(1-x)\log(1-\theta)$$

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

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

υπολογίζω μέση τιμή της κατανομής διότι ξέρουμε ότι Bernoulli $\equiv B(1,\theta)$

$$I_X^F(\theta)=\frac{E(x)}{\theta^2}+\frac{E(1-x)}{(1-\theta)^2}=\frac{\theta}{\theta^2}+\frac{1-\theta}{(1-\theta)^2}\Rightarrow I_X^F(\theta)=\frac{1}{\theta(1-\theta)}$$

Άρα $\text{κφ}_{C-R}=\dfrac{[g'(\theta)]^2}{n I_X^F(\theta)}=\dfrac{1}{\dfrac{n}{\theta(1-\theta)}}=\dfrac{\theta(1-\theta)}{n}$

2ο Βήμα: Εύρεση αμερόληπτου εκτιμητή της $g(\theta)=\theta$ με διασπορά ίση με το κφ C-R

Θεωρούμε την $T(\underline{X})=\bar{X}$ ή και άλλοι αμερόληπτοι

(αφού θ είναι η μέση τιμή του πληθυσμού και $\bar{x}$ εκτιμά τη μέση τιμή του πληθυσμού, επιλέγω αυτό)

Παρατηρώ ότι $E(\bar{X})=\theta$ $\left(E(\bar{X})=E\left(\frac{1}{n}\sum x_i\right)=\frac{1}{n}\sum E(x_i)=\frac{1}{n}\sum\theta=\frac{1}{n}n\theta=\theta\right)$

Βρήκα αμερόληπτο. Αν $Var\bar{X}=\text{κφ}_{C-R}$ είναι ο ΑΟΕΑ.

$$Var\bar{X}=Var\left(\frac{1}{n}\sum x_i\right)=\frac{1}{n^2}\sum Var(x_i)=\frac{1}{n^2}\sum\theta(1-\theta)=\frac{1}{n^2}n\theta(1-\theta)=\frac{\theta(1-\theta)}{n}=\text{κφ}_{C-R}$$

Άρα, ο $\bar{X}$ είναι ΑΟΕΑ της θ.

Παρατήρηση (Συνθήκες κανονικότητας): (Υποθέτουμε γενικά ότι ισχύουν) Ικανοποιούνται για τις γνωστές κατανομές (και για εκθετικό τύπο μοντέλα)

Σελ 107-111 (Πιπεράγγελ) για να δεις κάποια πράγματα

Γενικός κανόνας: Αν με ενδιαφέρει μια ποσότητα που σχετίζεται με το μ (μέση τιμή), τότε ξεκινάω με το $\bar{x}$. Αν με ενδιαφέρει μια ποσότητα που σχετίζεται με το σ (διακύμανση), τότε ξεκινάω με το S^2.

Πρόταση: Έστω ότι οι συνθήκες κανονικότητας (ομαλότητας) ικανοποιούνται για τον αμερόληπτο εκτιμητή $U=U(\underline{x})$ της $g(\theta)$. Τότε: $Var(U) = \text{ΚΦ}_{C-R} = \dfrac{[g'(\theta)]^2}{n I_X^F(\theta)}$ εάν-ν $U = g(\theta) + a(\theta)W$

δηλ. να είναι μια γραμμική συνάρτηση της W

όπου $W = \dfrac{\partial}{\partial\theta}\log f(\underline{x},\theta) = \sum_{i=1}^{n} \dfrac{\partial}{\partial\theta}\log f(x_i,\theta)$.

Απόδ.

($\Leftarrow$) Υποθέτω ότι $U = g(\theta) + a(\theta)W$. Επειδή U γραμμική συνάρτηση της W $\Rightarrow$

$$\Rightarrow \rho(U,W) = \pm 1 \Rightarrow \rho^2(U,W) = 1 \Rightarrow \frac{cov^2(U,W)}{Var(U)Var(W)} = 1 \Rightarrow cov^2(U,W) = Var(U)Var(W) \quad (1)$$

συντελεστής συσχέτισης του Pearson

Από την απόδειξη της ανισότητας Cramer-Rao $Var(W) = Var\left(\dfrac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right) = I_{\underline{X}}^F(\theta) = n I_X^F(\theta)$ (2)

(Πάλι από την απόδειξη της ανισότητας Cramer-Rao): $Cov(U,W) = Cov\left(U, \dfrac{\partial}{\partial\theta}\log f(\underline{x},\theta)\right) = g'(\theta)$ (3)

Από (1), (2), (3)

$$[g'(\theta)]^2 = Var(U)Var(W) = Var(U)\, n I_X^F(\theta) \Rightarrow Var(U) = \frac{[g'(\theta)]^2}{n I_X^F(\theta)} = \text{ΚΦ}_{C-R}.$$

($\Rightarrow$) Αποδεικνύεται ακολουθώντας την ακριβώς αντίστροφη πορεία.

Παρατήρηση: Στην ανισότητα Cramer-Rao το ίσον επιτυγχάνεται εάν και μόνο εάν

$$\boxed{\frac{\partial}{\partial\theta}\log f(\underline{x},\theta) = \frac{\partial}{\partial\theta}\log \prod_{i=1}^{n} f(x_i,\theta) = h(\theta,n)[U(\underline{x}) - g(\theta)]}$$

για τον αμερόληπτο εκτιμητή $U(\underline{x})$ της $g(\theta)$.

μια συνάρτηση της θ και του μεγέθους του δείγματος

Παράδειγμα Έστω τ.δ. $x_1, x_2, \dots x_n$ από Poisson(ϑ). ΑΟΕΔ της ϑ.

Απόδ.

επειδή τυχ. ανεξάρτητα (τα γινόμενα των περιθωρίων)

$$\frac{\partial}{\partial\vartheta}\log f(\underline{x},\vartheta) = \frac{\partial}{\partial\vartheta}\log\prod_{i=1}^{n} f(x_i,\vartheta) = \frac{\partial}{\partial\vartheta}\log\prod_{i=1}^{n}\frac{e^{-\vartheta}\vartheta^{x_i}}{x_i!} = \frac{\partial}{\partial\vartheta}\log\frac{e^{-n\vartheta}\vartheta^{\sum x_i}}{\prod x_i!} =$$

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

$$= \underbrace{\frac{n}{\vartheta}}_{h(n,\vartheta)}\Big(\underbrace{\frac{1}{n}\sum_{i=1}^{n} x_i}_{U(\underline{x})=\bar{x}} - \underbrace{\vartheta}_{g(\vartheta)=\vartheta}\Big)$$

Άρα (με βάση την προηγούμενη πρόταση και παρατήρηση): $\bar{x}$ ΑΟΕΔ της ϑ.

Παράδειγμα Έστω $x_1, x_2, \dots, x_n$ τ.δ. από Bernoulli(ϑ). ΑΟΕΔ της ϑ.

Απόδ.

$$\frac{\partial}{\partial\vartheta}\log f(\underline{x},\vartheta) = \frac{\partial}{\partial\vartheta}\log\prod_{i=1}^{n} f(x_i,\vartheta) = \frac{\partial}{\partial\vartheta}\log\prod_{i=1}^{n}\vartheta^{x_i}(1-\vartheta)^{1-x_i} = \frac{\partial}{\partial\vartheta}\log\ \vartheta^{\sum x_i}(1-\vartheta)^{n-\sum x_i} =$$

$$= \frac{\partial}{\partial\vartheta}\left[\sum x_i\log\vartheta + (n-\sum x_i)\log(1-\vartheta)\right] = \frac{\sum x_i}{\vartheta} - \frac{n-\sum x_i}{1-\vartheta} = \frac{\sum x_i - n\vartheta}{\vartheta(1-\vartheta)} =$$

$$= \frac{1}{\vartheta(1-\vartheta)}\left[\sum x_i - n\vartheta\right] = \underbrace{\frac{n}{\vartheta(1-\vartheta)}}_{h(\vartheta,n)}\Big[\underbrace{\frac{1}{n}\sum x_i}_{U=U(\underline{x})=\bar{x}} - \underbrace{\vartheta}_{g(\vartheta)=\vartheta}\Big].$$

Άρα ΑΟΕΔ: $U = \frac{1}{n}\sum x_i = \bar{x}$

Παράδειγμα (Κανονική Κατανομή $N(\mu,\vartheta)$, μ = γνωστή, $\vartheta=\sigma^2$ = άγνωστη.)

Έστω τ.δ. $x_1, x_2, \dots, x_n$ από $N(\mu,\vartheta)$. ϑ άγνωστο. Να βρεθεί ΑΟΕΔ της ϑ.

Λύση

(Υποθέτουμε ότι ισχύουν οι συνθήκες κανονικότητας) $\to$ Μπορώ όπως να εφαρμόσω (ταυτόχρονα) εύρεση του $K_{\phi C-R} = \dfrac{[g'(\vartheta)]^2}{n I_X^f(\vartheta)}$

$g(\vartheta) = \vartheta \Rightarrow g'(\vartheta) = 1$

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

Άσκηση: Δοκιμάστε τον πρώτο τύπο

$$\log f(x,\vartheta) = \log \frac{1}{\sqrt{2\pi}\sqrt{\vartheta}} e^{-\frac{(x-\mu)^2}{2\vartheta}} = -\frac{1}{2}\log(2\pi\vartheta) - \frac{(x-\mu)^2}{2\vartheta}$$

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

$$\frac{\partial^2}{\partial\vartheta^2}\log f(x,\vartheta) = \frac{1}{2\vartheta^2} - \frac{(x-\mu)^2}{\vartheta^3}$$

Άρα $I_X^F(\vartheta) = -E\left[\frac{\partial^2}{\partial\vartheta^2}\log f(x,\vartheta)\right] = -E\left(\frac{1}{2\vartheta^2} - \frac{(x-\mu)^2}{\vartheta^3}\right) = -\frac{1}{2\vartheta^2} + \frac{E(x-\mu)^2}{\vartheta^3} =$

$$= -\frac{1}{2\vartheta^2} + \frac{E[(x-E(x))]^2}{\vartheta^3} = -\frac{1}{2\vartheta^2} + \frac{\overbrace{Var(x)}^{=\vartheta \text{ από εκφώνηση}}}{\vartheta^3} = -\frac{1}{2\vartheta^2} + \frac{\vartheta}{\vartheta^3} = \frac{1}{2\vartheta^2}$$

Δηλ. $I_X^F(\vartheta) = \frac{1}{2\vartheta^2}$

Άρα $\text{κφ}_{C-R} = \frac{1}{\frac{n}{2\vartheta^2}} \Rightarrow \text{κφ}_{C-R} = \frac{2\vartheta^2}{n}$

Εύρεση αμερόληπτου εκτιμητή της ϑ που η διακύμανσή του είναι ίση με κφ_{C-R}.

Ξεκινάω με την $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})$ που εκτιμά αμερόληπτα την σ^2 που είναι η ϑ.

Επειδή γνωρίζω την μ, θεωρώ $\boxed{S_*^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)^2}$

Η S_*^2 είναι αμερόληπτη της ϑ. Πράγματι: $E(S_*^2) = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)^2\right] = \frac{1}{n-1}\sum_{i=1}^{n}E(x_i-\mu)^2 =$

$= \frac{1}{n-1}\sum_{i=1}^{n}E(x_i-\underbrace{E(x_i)}_{=\mu})^2 = \frac{1}{n-1}\sum_{i=1}^{n}Var(x_i) = \frac{1}{n-1}\sum_{i=1}^{n}\vartheta = \frac{n}{n-1}\vartheta$ → Δεν είναι αμερόληπτος της ϑ

Θα πρέπει να διορθώσουμε ώστε να γίνει αμερόληπτος

Παρατηρώ ότι: $E\left(\frac{n-1}{n}S_*^2\right) = \frac{n-1}{n}E(S_*^2) = \frac{n-1}{n}\frac{n}{n-1}\vartheta = \vartheta.$

Άρα ο $T = \frac{n-1}{n}S_*^2 = \frac{\cancel{n-1}}{n}\frac{1}{\cancel{n-1}}\sum_{i=1}^{n}(x_i-\mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$: αμερόληπτος της ϑ

Αρκεί ν.δ.ο $Var(T) = \text{κάτω φράγμα C-R} = \frac{2\theta^2}{n}$

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

$$x_i \sim N(\mu,\theta) \Rightarrow \frac{x_i-\mu}{\sqrt{\theta}} \sim N(0,1) \Rightarrow \frac{(x_i-\mu)^2}{\theta} \sim N^2(0,1) = \chi^2_1$$

Άρα $Var\left[\frac{(x_i-\mu)^2}{\theta}\right] = Var(\chi^2_1) \overset{Var(\chi^2_\nu)=2\nu}{=\!=\!=} 2\cdot 1 = 2 \Rightarrow \frac{1}{\theta^2} Var(x_i-\mu)^2 = 2 \Rightarrow Var(x_i-\mu)^2 = 2\theta^2 \ \forall\, i=1,\dots,n$

Επομένως $Var(T) = \frac{1}{n^2}\sum_{i=1}^{n} Var(x_i-\mu)^2 = \frac{1}{n^2}\sum_{i=1}^{n} 2\theta^2 = \frac{1}{n^2}\, n\, 2\theta^2 = \frac{2\theta^2}{n} = \text{κφ C-R}$

Άρα ο $T = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$ είναι ο ΑΟΕΔ της θ.

<u>2ος τρόπος</u> Εφαρμογή της Παρατήρησης

$$\frac{\partial}{\partial\theta}\log f(\underline{x},\theta) = k(\theta,n)\,[\overset{\text{ΑΟΕΔ}}{U(\underline{x})} - g(\theta)]$$

$$\frac{\partial}{\partial\theta}\log f(\underline{x},\theta) = \frac{\partial}{\partial\theta}\log\prod_{i=1}^{n} f(x_i,\theta) = \frac{\partial}{\partial\theta}\log\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\theta}}e^{-\frac{(x_i-\mu)^2}{2\theta}} = \frac{\partial}{\partial\theta}\left(-\frac{n}{2}\log(2\pi\theta) - \frac{1}{2\theta}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

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

$$= \frac{1}{2\theta^2}\left[\sum_{i=1}^{n}(x_i-\mu)^2 - n\theta\right] = \underbrace{\frac{n}{2\theta^2}}_{k(\theta,n)}\Big[\underbrace{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2}_{U(\underline{x})} - \underbrace{\theta}_{g(\theta)}\Big]$$

(Γράφω ότι επειδή είναι στη μορφή $k(\theta,n)[U(x)-g(\theta)]$ ο $U(x)$ είναι ΑΟΕΔ της θ]

Επομένως ο $U = \frac{1}{n}\sum(x_i-\mu)^2$ είναι ΑΟΕΔ της θ.

<u>Ορισμός (Αποτελεσματικός εκτιμητής)</u>

Ο εκτιμητής $U = U(\underline{x})$ της $g(\theta)$ λέγεται αποτελεσματικός, αν ο U είναι ΑΟΕΔ της $g(\theta)$.

<u>Σχετική αποτελεσματικότητα:</u>

Έστω $U = U(\underline{x})$ αποτελεσματικός εκτιμητής της $g(\theta)$. Έστω $T = T(\underline{x})$ αμερόληπτος της $g(\theta)$ με $Var(T) < \infty$ (ότι η διακύμανση είναι πεπερασμένη). Η σχετική αποτελεσμααπόσταση του T σε σχέση με τον U ορίζεται από το πηλίκο:

Σχ. Απόσ. του T σε σχέση με U $= \dfrac{Var(U)}{Var(T)}$

Παρατήρηση: $0 \le \dfrac{Var(U)}{Var(T)} \le 1$.

Παράδειγμα: Έστω τ.δ. X_1, X_2, X_3 από κατανομή Poisson(θ). Έστω ο εκτιμητής $T = \dfrac{X_1 + 2X_2 + 3X_3}{6}$ της θ, $\theta > 0$. Να βρεθεί η σχετική αποτελεσματικότητα του T ως προς τον αποτελεσματικό εκτιμητή U της θ.

Αποδ.:

Πρώτα πρέπει να δω αν ο T είναι αμερόληπτος

$$E(T) = \frac{1}{6}E(X_1) + \frac{2}{6}E(X_2) + \frac{3}{6}E(X_3) \xrightarrow{E(X_i)=\theta} \theta.$$

Άρα έτσι ο T είναι αμερόληπτος της θ.

Πρέπει να βρω την διακύμανση του T.

$$Var(T) = \frac{1}{36}\left[Var(X_1) + 4Var(X_2) + 9Var(X_3)\right] \xrightarrow{Var(X_i)=\theta} \frac{7\theta}{18}$$

Ο ΑΟΕΔ της θ είναι ο $U = \bar{x}$ (το δείγμα προηγουμένως)

και $Var\left(\frac{1}{3}\sum X_i\right) = \frac{1}{9}\sum Var(X_i) = \frac{3\theta}{9} = \frac{\theta}{3}$.

σχ. αποτελ. $= \dfrac{Var(U)}{Var(T)} = \dfrac{\theta/3}{7\theta/18} = 0{,}8571 = 85{,}71\%$ (αυτό σημαίνει ότι ο T είναι 85,71% καλός σε σχέση με το U ή ότι ο T υπολείπεται 14,39% σε σχέση με τον άριστο/αποτελεσματικό /ΑΟΕΔ)

Μονοπαραμετρική Εκθετική Οικογένεια Κατανομών (ΜΕΟΚ)

Η μονοπαραμετρική (η παράμετρος $\theta \in \Theta \subseteq \mathbb{R}$) οικογένεια πυκνοτήτων $f(x;\theta)$ λέγεται ότι ανήκει στην ΜΕΟΚ, αν μπορεί να εκφραστεί στη μορφή: $f(x;\theta) = c(\theta)h(x)e^{Q(\theta)T(x)}$, $x \in A \subseteq \mathbb{R}$

με A ανεξάρτητο της θ και c, h, Q, T συναρτήσεις των θ και x.

$\to f(x,\theta) = \theta e^{-\theta x}$ Μέσω $c(\theta) = \theta$, $h(x) = 1$, $Q(\theta) = -\theta$, $T(x) = x$

Μπορείτε να αποδείξετε ότι

Εκθ (θ)

Poisson (θ)

$N(\theta, 1)$

$N(\theta, \sigma^2)$, σ^2: γνωστό

$N(\mu, \theta)$, μ γν.

$B(n, \theta)$

ανήκουν στην ΜΕΟΚ

Παρατήρηση: Η ΜΕΟΚ ικανοποιεί τις συνθήκες κανονικότητας της ανισότητας Cramer-Rao