Διάλεξη 5η

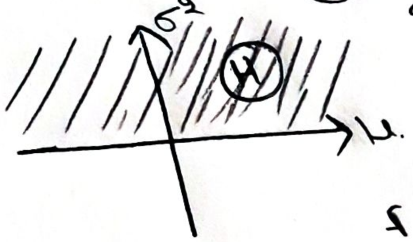
27/03/2019

ΣΤΑΤΙΣΤΙΚΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ

Παράδειγμα: Έστω τ.δ. $x_1, \dots, x_n$ από πληθυσμό με κατανομή $N(\mu, \sigma^2)$, μ, σ^2: άγνωστα. Να βρεθεί επαρκές στατιστικό.

Λύσ.

Παραμ. χώρος $\Theta = \{(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0, \mu, \sigma^2 : \text{άγνωστο}\} \subseteq \mathbb{R}^2$



Θεώρημα N-F:

$$f(\underline{x}, \mu, \sigma^2) = \prod_{i=1}^{n} f(x_i, \mu, \sigma^2) =$$

$$= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i - \mu)^2}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i^2 - 2\mu x_i + \mu^2)}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 + \frac{\mu}{\sigma^2}\sum_{i=1}^{n} x_i - \frac{n\mu^2}{2\sigma^2}}$$

$$= g(T(\underline{x}), \mu, \sigma^2)\, \underbrace{h(\underline{x})}_{1}$$

$g\left(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i^2, \mu, \sigma^2\right)$: επαρκές στατιστικό.

Από θεωρ N-F το επαρκές στατιστικό

$$T(\underline{x}) = \left(\underbrace{\sum_{i=1}^{n} x_i}_{\mu}, \underbrace{\sum_{i=1}^{n} x_i^2}_{\sigma^2}\right)$$

Παράδειγμα: Έστω τ.δ. $X_1, \ldots, X_n$ από πληθ. $U(0,\theta)$, $\theta>0$.
Να βρεθεί επαρκή στατιστικό θ.

Λύση

Θυμίζω ότι αν $X \sim U(\alpha,\beta)$ τότε $f(x) = \begin{cases} \frac{1}{\beta-\alpha}, & \alpha<x<\beta \\ 0, & \text{αλλού.} \end{cases}$

άρα $X \sim U(0,\theta)$ τότε $f(x) = \begin{cases} \frac{1}{\theta}, & 0<x<\theta \\ 0, & \text{αλλού.} \end{cases}$

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

$$= \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0,\theta)}(x_i)$$

$$\prod_{i=1}^{n} I_{(0,\theta)}(x_i) = \begin{cases} 1, & 0<x_i<\theta, \ \forall i \\ 0, & \exists \text{ τουλ. ένα } i : x_i \notin (0,\theta) \end{cases} = \begin{cases} 1, & 0<\underset{\min x_i}{x_{(1)}}<\underset{\max x_i}{x_{(n)}}<\theta \\ 0, & \text{αλλού.} \end{cases}$$

$$= I_{(0,x_{(n)})}(x_{(1)}) \cdot I_{(0,\theta)}(x_{(n)})$$

Άρα $f(\underline{x},\theta) = \frac{1}{\theta^n} I_{(0,x_{(n)})}(x_{(1)}) \cdot I_{(0,\theta)}(x_{(n)})$

$$= g(x_{(n)},\theta)\, h(\underline{x})$$

$$g(x_{(n)},\theta) = \frac{1}{\theta^n} I_{(0,\theta)}(x_{(n)})$$

$$g(\underline{x}) = I_{(0,x_{(n)})}(x_{(1)})$$

Παράδειγμα: Έστω τ.δ. $X_1,\dots,X_n$ από την $U(\theta,\theta+1)$, $\theta>0$. Να βρεθεί επαρκές στατιστικό για την θ.

Λύση

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

$\forall i,\ \theta < x_i < \theta+1$

$\theta < x_{(1)} < x_{(2)} < \dots < x_{(n)} < \theta+1$

$$=\prod_{i=1}^{n} I_{(\theta,\theta+1)}(x_i)=I_{(\theta,x(n))}(x(1))\,I_{(\theta,\theta+1)}(x(n))$$

$$=g(x(1),x(n),\theta)\,h(\underline{x}).$$

όπου $h(\underline{x})=1$, $g(x(1),x(n),\theta)=I_{(\theta,x(n))}(x(1))\,I_{(\theta,\theta+1)}(x(n))$

Άρα $(x(1),x(n))$ επαρκές για την θ.

$(\theta,$ ← πλησιάζεται από $x(1)$

$(\ ,\theta)$ ← πλησιάζεται από $x(n)$

Παρατηρήσεις στην επάρκεια:

1) Τετριμμένα το τ.δ. $X_1,\dots,X_n$ είναι πάντα επαρκές.

$$f(\underline{x},\theta)=g(x_1,\dots,x_n,\theta)\,h(\underline{x})$$

όπου $h(\underline{x})=1$, $g(x_1,\dots,x_n,\theta)=f(\underline{x},\theta)$

2) Αν T_1 είναι επαρκές στατιστικό και ω μία "1-1" συνάρτηση τότε και το $T_2=\omega(T_1)$ είναι επαρκές.

Απόδ.

Αφού το T_1 επαρκές $\exists\, g$ και h.

$$f(\underline{x},\theta)=g(T_1(\underline{x}),\theta)h(\underline{x})=g(\omega^{-1}(T_2),\theta)h(\underline{x})$$

$$=g^*(T_2,\theta)\cdot h(\underline{x}) \longrightarrow T_2:\ \text{επαρκές}$$

3) Αν $T=T(\underline{x})$ είναι ένας μετασχηματισμός του τ.δ. $X_1,\dots,X_n$ τότε αποδ. ότι $I^F_{\underline{x}}(\theta) \geq I^F_{T(\underline{x})}(\theta)$ με ισότητα αν-ν το $T(\underline{x})$: επαρκές.

Βελτίωση εκτιμητή με χρήση της επάρκειας:

Κάθε εκτιμήτρια μπορεί να βελτιωθεί (με κριτήριο τη διακύμανση) με χρήση ενός επαρκούς στατιστικού.

Θεώρημα: (Rao-Blackwell)

Έστω τ.δ. $X_1,\dots,X_n$ από πληθυσμό με κατανομή $f(x,\theta), \theta\in\Theta$. Έστω $T=T(\underline{x})$ επαρκής σ.σ. και έστω $S=S(\underline{x})$ ένας εκτιμητής της $g(\theta)$ τότε αν $S^*=E(S|T)$

i) $E(S^*)=E(S)$, $\forall\theta\in\Theta$ (οπότε αν ο S είναι αμερόληπτος και ο S^* είναι αμερόληπτος)

ii) $Var(S^*)\leq Var(S)$ (Δεν μου τον κάνει ΑΟΕΔ, όμως μου τον βελτιώνει)

Αποδ.

Αφού το T επαρκές η κατανομή του $\underline{X}|T$ ανεξ. της θ. Άρα $S(\underline{x})|T$ ανεξ. της θ.

Άρα και η $E(S|T)$ ανεξ. της θ.

Άρα το $S^*=E(S|T)$ ανεξ. της θ.

Η απόδειξη στηρίζεται

$$E(E(X_1|X_2))=E(X_1) \quad (*)$$

$$Var(X_1)=E(Var(X_1|X_2))+Var(E(X_1|X_2)) \quad (**)$$

i) $E(S^*)=E(E(S|T)) \overset{(*)}{=} E(S)$

ii)
$$\left.\begin{aligned} Var(S) &= E(Var(S|T))+Var(E(S|T)) \\ &= Var(S^*)+E(Var(S|T)) \end{aligned}\right\} \Rightarrow \boxed{Var(S^*)\leq Var(S)}$$

Αλλά $Var(S|T)\geq 0$. Άρα $E(Var(S|T))\geq 0$

Το θεώρημα Rao-Blackwell είναι το πρώτο βήμα βελτίωσης εκτιμητή με την έννοια της επάρκειας. Αν ενισχυθούν οι υποθέσεις με την έννοια της πληρότητας που μπορούμε να οδηγηθούμε σε ΑΟΕΔ εκτιμητή.

Ορισμός: Η σσ. $T(\underline{x})=T$ λέγεται πλήρης σ.σ. αν η σχέση $E(\phi(T))=0$, $\forall \theta \in \Theta$ συνεπάγεται ότι $\phi(T)=0$, δηλ. $\phi(t)=0$ $\forall$ τιμή t της σσ. T.

Θεώρημα: (Lehmann - Scheffe)

Έστω τ.δ $X_1,\dots,X_n$ από πληθυσμό με κατανομή της $f(x,\theta)$, $\theta\in\Theta$. Έστω $T=T(X_1,\dots,X_n)$ μια πλήρης και επαρκής σσ. και έστω $S=S(T)$ ένας αμερόληπτος εκτιμητής της $g(\theta)$ που είναι συνάρτηση του επαρκούς και πλήρους T. Τότε η S είναι ΑΟΕΔ εκτιμητής της $g(\theta)$ και μάλιστα μοναδικός.

Απόδ

Έστω $S^*=S^*(T)$ μια άλλη σ.σ. του επαρκούς και πλήρους που είναι αμερόληπτη της $g(\theta)$.

Αφού S αμερόληπτη $E(S)=g(\theta)$
-//- S^* -//- $E(S^*)=g(\theta)$

$$\left.\begin{aligned} E(S)&=g(\theta)\\ E(S^*)&=g(\theta)\end{aligned}\right\} \Rightarrow E(S)-E(S^*)=0 \Rightarrow E(S-S^*)=0 \Rightarrow E((S-S^*)(T))=0 \xRightarrow[\text{του } T]{\text{πληρότ.}}$$

$(S-S^*)(T)=0 \Rightarrow S(T)=S^*(T)$

Άρα $\exists$ ένας και μόνο αμερόληπτος εκτιμητής της $g(\theta)$ που είναι συνάρτηση του επαρκούς και πλήρους.

Έστω S^{**} ένας άλλος αμερόληπτος εκτιμητής της $g(\theta)$ που δεν είναι κατ' ανάγκη συνάρτηση του T.
Θεωρώ την βελτίωση του S^{*} μέσω του θεωρ. Rao-Blackwell, δηλαδή θεωρώ την $E(S^{**}/T)$
Επειδή (i) $E(E(S^{**}|T)) = E(S^{**}) = g(\theta)$
και επειδή $E(S^{**}|T)$ είναι συνάρτηση του επαρκούς πλήρους και αμερόληπτου της $g(\theta)$ θα πρέπει
$E(S^{**}|T) = S$
Από το θεωρ. Rao-Blackwell:

$Var(S) \leq Var(S^{**})$

Άρα ο S έχει την μικρότερη διακύμανση μεταξύ όλων των αμερ. εκτιμητών της $g(\theta)$. Άρα ΑΟΕΔ.

Εφαρμογή: Θεώρημα L-S για εύρεση ΑΟΕΔ.

Βήμα 1: Εύρεση του επαρκούς και πλήρους στατιστικού T.

Βήμα 2: Εύρεση συνάρτησης του T που είναι αμερόληπτη της $g(\theta)$.

Παράδειγμα: Έστω τ.δ. $X_1, \dots, X_n$ από κατανομή Bernoulli με παράμετρο θ, $0<\theta<1$, (δηλ. $B(1,\theta)$). Να βρεθούν ΑΟΕΔ των $\theta, \theta^2, \theta(1-\theta)$.

Λύση

Εύρεση επαρκούς και πλήρους.
Το επαρκές το βρήκαμε και είναι $T = \sum_{i=1}^{n} x_i$
Για την απόδειξη της πληρότητας αρκεί νδο η συνεπαγωγή $E(\phi(T)) = 0, \forall \theta \in \Theta = (0,1) \Rightarrow \phi(t) = 0, \forall t$.

Έστω $\underbrace{E(\phi(T))=0}$, $\forall \theta \in \Theta$

Για να δράσω την $E(\phi(T))$ χρειάζομαι την κατανομή του T. Θέλω την κατανομή του $T=\sum_{i=1}^{n} X_i$.

Από το 1ο λήμμα μετην μεθόδου της ροπογεννήτριας το $T \sim B(n,\theta)$, η $p_T(t)=\binom{n}{t}\theta^t(1-\theta)^{n-t}$, $t=0,\dots,n$

Επομένως $E(\phi(T))=0 \Rightarrow \sum_{t=0}^{n}\phi(t)p_T(t)=0$, $\forall \theta \in \Theta$

$$\Rightarrow \sum_{t=0}^{n}\phi(t)\binom{n}{t}\theta^t(1-\theta)^{n-t}=0, \forall \theta.$$

$$\Rightarrow \sum_{t=0}^{n}\phi(t)\binom{n}{t}\left(\frac{\theta}{1-\theta}\right)^t(1-\theta)^n=0, \forall \theta.$$

$$\Rightarrow \underbrace{\sum_{t=0}^{n}\phi(t)\binom{n}{t}p^t}=0, \quad p=\frac{\theta}{1-\theta}, \forall p>0.$$

πολυώνυμο n-οστού βαθμού του οποίου είναι $=0$, $\forall p>0$ έχει περισσότερες n ρίζες. Άρα είναι το μηδενικό πολυώνυμο, άρα $\phi(t)\binom{n}{t}=0$, $\forall t \Rightarrow \phi(t)=0$, $\forall t$.

Το $T=\sum_{i=1}^{n} X_i$ είναι και πλήρες

Για ΑΟΕΔ της θ αρκεί να βρω συνάρτηση του T που να είναι αμερ. της θ.

__Δοκιμάζω:__ $E(T) \xlongequal[T\sim B(n,\theta)]{\text{αφού}} n\theta \Rightarrow \frac{1}{n}E(T)=\theta \Rightarrow E\left(\frac{1}{n}\cdot T\right)=\theta$

$\Rightarrow E\left(\frac{T}{n}\right)=\theta \Rightarrow \frac{T}{n}$ είναι συνάρτηση του T και αμερ. της θ.

Άρα $\frac{1}{n}T=\frac{1}{n}\sum X_i=\bar{X}$: ΑΟΕΔ της θ.

8

$Var(T) \overset{T\sim B(n,\theta)}{=} n\theta(1-\theta) \Rightarrow$

$E(T^2)-(ET)^2 = n\theta - n\theta^2 \Rightarrow E(T^2)-(n\theta)^2 = n\theta - n\theta^2$

$\Rightarrow E(T^2) = n\theta + n^2\theta^2 - n\theta^2$

$\Rightarrow E(T^2) - n\theta = n^2\theta^2 - n\theta^2$

$\Rightarrow E(T^2) - E(T) = \theta^2 n(n-1)$

$\Rightarrow E(T^2 - T) = \theta^2 n(n-1)$

$\Rightarrow E\left(\frac{T^2-T}{n(n-1)}\right) = \theta^2.$ $\Rightarrow$ Άρα $\frac{T^2-T}{n(n-1)}$ ΑΟΕΔ της θ^2.

$\theta(1-\theta) = \theta - \theta^2.$

Θεωρώ $\frac{1}{n}T - \frac{T^2-T}{n(n-1)}$

Είναι $E\left(\frac{1}{n}T - \frac{T^2-T}{n(n-1)}\right) = E\left(\frac{1}{n}T\right) - E\left(\frac{T^2-T}{n(n-1)}\right) = \theta(1-\theta)$

$= \theta - \theta^2.$

Άρα $\frac{1}{n}T - \frac{T^2-T}{n(n-1)}$ ΑΟΕΔ $\theta(1-\theta)$.