Επάρκεια

$\boxed{\text{Παραγοντικό Θεώρημα Neyman - Fisher}}$

$f(\underline{x},\vartheta) = g[T(\underline{x}),\vartheta]\,h(\underline{x})$ τότε $T(x)$ επαρκές.

Παράδειγμα

1) Έστω τ.δ $X_1,\dots,X_n$ από Poisson (ϑ) $\quad p(x,\vartheta) = \dfrac{e^{-\vartheta}\vartheta^{x}}{x!}$ $x=0,1,2,\dots$, $\vartheta>0$

Να βρεθεί επαρκές στατιστικό.

Λύση

$$f(\underline{x},\vartheta) = \prod_{i=1}^{n} p_X(x_i,\vartheta) = \prod_{i=1}^{n}\frac{e^{-\vartheta}\vartheta^{x_i}}{x_i!} = \frac{e^{-n\vartheta}\,\vartheta^{\Sigma x_i}}{\prod_{i=1}^{n} x_i!} = g[T(\underline{x}),\vartheta]\,h(\underline{x})$$

όπου $h(\underline{x}) = \dfrac{1}{\prod_{i=1}^{n} x_i!}$, $g[T(\underline{x}),\vartheta] = e^{-n\vartheta}\,\vartheta^{\sum_{i=1}^{n} x_i}$

Άρα $T(\underline{x}) = \sum_{i=1}^{n} x_i$ επαρκές.

2) Έστω $X_1,\dots,X_n$ τ.δ από $U(0,\vartheta)$, $\vartheta>0$. Να βρεθεί το επαρκές στατιστικό

Λύση

Η σ.π.π της $U(0,\vartheta)$ είναι $f(x,\vartheta) = \begin{cases} \frac{1}{\vartheta} & , 0<x<\vartheta \\ 0 & , \text{αλλού.} \end{cases}$

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

όπου $I_A(x) = \begin{cases} 1 & x\in A \\ 0 & x\notin A \end{cases}$ (χαρακτηριστική ή δείκτρια συνάρτηση)

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

$$\prod_{i=1}^{n} I_{(0,\theta)}(x_i) = \begin{cases} 1 & , 0 < x_i < \theta_i \ \forall i = 1, \dots n \\ 0 & , \exists \text{ τουλ. ένα } i \text{ με } x_i \notin (0,\theta) \end{cases} =$$

$$= \begin{cases} 1 & , 0 < x_i < \theta \ \forall i \\ 0 & , \text{διαφορετικά} \end{cases} = \begin{cases} 1 & , 0 < X_{(1)} \le x_{(n)} < \theta \\ 0 & , \text{διαφ} \end{cases} =$$

(όπου $X_{(1)} = \min\{x_1, \dots, x_n\}$ και $X_{(n)} = \max\{x_1, \dots, x_n\}$)

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

Άρα $f(\underline{x},\theta) = \frac{1}{\theta^n} \underbrace{I_{(0,x_{(n)})}(x_{(1)})}_{\text{συνάρτηση μόνο για των } x_1, \dots, x_n} I_{(0,\theta)}(x_{(n)})$

$f(\underline{x},\theta) = [T(\underline{x}), \theta]\, h(\underline{x})$

όπου $h(\underline{x}) = I_{(0,x_{(n)})}(x_{(1)})$ και $g[T(\underline{x}), \theta] = \frac{1}{\theta^n} I_{(0,\theta)}(x_{(n)})$

Άρα $T(\underline{x}) = X_{(n)} = \max\{x_1, \dots, x_n\}$ επαρκές.

3) Έστω $X_1, \dots, X_n$ τ.δ από $U(\theta, \theta+1)$, $\theta > 0$. Να βρεθεί επαρκές.

<u>Λύση</u>

Η σ.π.π της $U(\theta,\theta+1)$ είναι $f(x,\theta) = \begin{cases} \frac{1}{\theta+1-\theta} = 1 & , \theta < x < \theta+1 \\ 0 & , \text{αλλού} \end{cases}$

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

$$= \begin{cases} 1, & \theta < x_i < \theta+1 \quad \forall i \\ 0, & \text{αλλού} \end{cases} = \begin{cases} 1, & \theta < x_{(1)} \leq x_{(n)} < \theta+1 \\ 0, & \text{αλλού} \end{cases} =$$

$$= I_{(\theta, x_{(n)})}(x_{(1)}) \cdot I_{(\theta,\theta+1)}(x_{(n)}) = g[T(\underline{x}),\theta]\, h(\underline{x})$$

όπου $h(\underline{x}) = 1$ και $g[T(\underline{x}),\theta] = I_{(\theta,x_{(n)})}(x_{(1)}) \cdot I_{(\theta,\theta+1)}(x_{(n)})$

με $T(\underline{x}) = (x_{(1)}, x_{(n)})$ (διδιάστατη συνάρτηση)

Άρα το $(X_{(1)}, X_{(n)})$ είναι επαρκές για θ.

<u>Παρατήρηση</u> (στην επάρκεια)

1. Το ίδιο το τυχαίο δείγμα $X_1,\dots,X_n$ είναι επαρκές.

$$\left.\begin{aligned} f(\underline{x},\theta) &= g(x_1,\dots,x_n,\theta)\, h(\underline{x}) \\ \text{με } h(\underline{x}) &= 1 \quad g(x_1,\dots,x_n,\theta) \end{aligned}\right\} = f(x_1,\dots,x_n,\theta)$$

2. Έστω $T_1(x_1,\dots,x_n)$ επαρκές και έστω $T_2(x_1,\dots,x_n) = K[T_1(x_1,\dots,x_n)]$ με K να είναι 1-1 συνάρτηση. Τότε το T_2 είναι επαρκές

$\boxed{\text{Απόδειξη}}$

Αφού T_1 επαρκές $f(\underline{x},\theta) = g[T_1(\underline{x}),\theta]\, h(\underline{x}) =$

$= [K^{-1}(T_2(\underline{x})), \theta] \cdot h(\underline{x})$

ή $g^* [T_2(\underline{x}), \theta] \cdot h(\bar{x})$

με g^* συνάρτηση της g και της K^{-1}

Άρα T_2 επαρκές.

3. Αν $\underline{X}$ τ.δ και $T(\underline{X})$ στατιστική συνάρτηση τότε $I_{\underline{X}}(\theta) \geq I_{T(\underline{X})}(\theta)$ με την ισότητα να ισχύει <u>ανν</u> $T(\underline{X})$ επαρκές.

Χρήση της επάρκειας για την βελτίωση (ως προς την διακύμανση) εκτιμητών

$\boxed{\text{ΘΕΩΡΗΜΑ}}$ Rao - Blackwell

Έστω $T(\underline{X})$ επαρκές στατιστικό και $S(\underline{X})$ ένας εκτιμητής της $g(\theta)$

Έστω ακόμη $S^*(\underline{X}) = E(S/T)$. Τότε

α. $E(S) = E(S^*)$ και έτσι αν ο S είναι αμερόληπτος της $g(\theta)$ και ο S^* είναι αμερόληπτος της $g(\theta)$

β. $Var(S^*) \leq Var(S)$

(όπου $E(S/T) = \int s f_{S/T} \, ds$)

$\boxed{\text{Απόδειξη}}$

Η απόδειξη βασίζεται σε 2 σχέσεις:

Αν X_1, X_2 τυχαίες μεταβλητές τότε i) $E[E(X_1/X_2)] = E(X_1)$

ii) $Var(X_1) = E(Var(X_1/X_2)) + Var(E(X_1/X_2))$

α. $E(S^*) = E(E(S/T)) \overset{(i)}{=} E(S)$

Αν ο S αμερόληπτος της $g(\theta)$ τότε $\left.\begin{array}{l} E(S) = g(\theta) \\ \text{Επειδή } E(S^*) = E(S) \end{array}\right\} \Rightarrow E(S^*) = g(\theta)$

Άρα S^* αμερόληπτος της $g(\theta)$

β. $Var(S) \overset{iii)}{=} E[Var(S/T)] + Var[E(S/T)] = E[Var(S/T)] + Var(S^*)$

Άρα αρκεί $E[Var(S/T)] \geq 0$ που ισχύει αφού $Var(S/T) \geq 0$ ως διακύμανση. $\Rightarrow$

$\Rightarrow Var(S^*) \leq Var(S)$

Ερώτημα: Μπορώ να βελτιώσω ~~εκτιμ~~ ακόμη περισσότερο έναν εκτιμητή ώστε να γίνει ΑΟΕΔ?

Απάντηση: ΝΑΙ. αν το T εκτός από επαρκές έχει ακόμη μια ιδιότητα· είναι πλήρες

ΟΡΙΣΜΟΣ (πληρότητας)

Η στατιστική συνάρτηση $T(\underline{x})$ λέγεται πλήρης αν η σχέση $E[\varphi(T)] = 0$ $\forall \theta \in \Theta$ συνεπάγεται ότι $\varphi(T) = 0$ δηλαδή $\varphi(t) = 0$ $\forall$ δυνατή τιμή t της σ.σ.

Ερμηνεία ορισμού

Η μοναδική αμερόληπτη εκτιμήτρια του 0 είναι η μηδενική συνάρτηση

Το Θεώρημα των Lehmann-Scheffe που ακολουθεί διατυπώνει ότι αν το επαρκές T είναι και πλήρες τότε η Rao-Blackwell βελτίωση οδηγεί σε ΑΟΕΔ.

$\boxed{\text{ΘΕΩΡΗΜΑ}}$ (Lehmann - Scheffe)

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

$\boxed{\text{Απόδειξη}}$

Έστω $S^* = S^*(T)$ ένας άλλος αμερόληπτος εκτιμητής της $g(\theta)$

$$\left.\begin{array}{l} \text{Τότε} \quad E(S^*) = g(\theta) \\ \text{Αλλά} \quad E(S) = g(\theta) \end{array}\right\} \Rightarrow E(S^*) - E(S) = g(\theta) - g(\theta) = 0 \Rightarrow E(S^* - S) = 0 \Rightarrow$$

$$\Rightarrow E[S(T) - S^*(T)] = 0 \Rightarrow E[(S - S^*)(T)] = 0 \xrightarrow[\text{πλήρες}]{T} (S - S^*)(t) = 0 \ \forall t \Rightarrow$$

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

Άρα μόνο ένας αμερόληπτος της $g(\theta)$ υπάρχει που να είναι συνάρτηση του επαρκούς και πλήρους T

Έστω S^{**} αμερόληπτος της $g(\theta)$ ο οποίος δεν είναι κατ' ανάγκη συνάρτηση του T. Θεωρώ τον $E(S^{**}/T)$

(1) $E[E(S^{**}/T)] \stackrel{(i)}{=} E(S^{**}) = g(\theta)$

Άρα ο $E(S^{**}/T)$ που προφανώς είναι συνάρτηση του T αμερόληπτος της $g(\theta)$. Αλλά μόνο ένας αμερόληπτος της $g(\theta)$ υπάρχει που να είναι συνάρτηση του επαρκούς και πλήρους T.

Άρα $E(S^{**}/T) \equiv S(T)$ ⊕

Από το Θεώρημα των Rao- Blackwell

$$Var\left(E(S^{**}|T)\right) \le Var(S^{**})$$

$$\xRightarrow{(+)} Var(S(T)) \le Var(S^{**})$$

Άρα $S=S(T)$ ΑΟΕΔ

2ος τρόπος εύρεσης ΑΟΕΔ εκτιμητών μέσω του Θεωρήματος των Lehmann-Scheffe

Βήμα 1: Βρίσκω το επαρκές και πλήρες T

Βήμα 2: Προσπαθώ να βρω συνάρτηση S του επαρκούς και πλήρους T που είναι αμερόληπτη της $g(\theta)$

Τότε S είναι ΑΟΕΔ

Παράδειγμα

1) Έστω τ.δ $X_1, \dots, X_n$ από πληθυσμό με κατανομή Bernoulli δηλ. $B(1,\theta)$, $0<\theta<1$. Να βρεθούν ΑΟΕΔ της θ και θ^2

Λύση

Η $B(1,\theta)$ $p_X(x,\theta) = \theta^x (1-\theta)^{1-x}$, $0<\theta<1$, $x=0,1$.

Βήμα 1: εύρεση επαρκούς και πλήρους.

$$p(\underline{x},\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i} =$$

$$= g(T(\underline{x}),\theta)\, h(\underline{x}) \quad \text{όπου } h(x)=1 \text{ και } g(T(\underline{x}),\theta) = \theta^{\Sigma x_i}(1-\theta)^{1-\Sigma x_i}$$

Άρα $T(\underline{x}) = \Sigma x_i$ επαρκές.

Το $T = T(x_1, \dots, x_n) = \sum_{i=1}^{n} x_i$ επαρκές

Πληρότητα: Αρκεί $E[\varphi(T)] = 0 \ \forall \theta \in (0,1) \Rightarrow \varphi(t) = 0 \ \forall t$

$$E[\varphi(T)] = \begin{cases} \sum_t \varphi(t) p_T(t) & , T \text{ διακριτή} \\ \int_{-\infty}^{+\infty} \varphi(t) f_T(t)\, dt & , T \text{ συνεχής.} \end{cases}$$

Ποια η κατανομή του $T = \sum_{i=1}^{n} x_i$?

Με την μέθοδο της ροπογεννήτριας το $T = \sum_{i=1}^{n} x_i \sim B(n, \theta)$

$$E[\varphi(T)] = 0 \Rightarrow \sum_{t=0}^{n} \varphi(t)\, p_T(t) = 0 \Rightarrow \sum_{t=0}^{n} \varphi(t) \binom{n}{t} \theta^t (1-\theta)^{n-t} = 0 \ \forall \theta \in (0,1)$$

$$\Rightarrow \sum_{t=0}^{n} \varphi(t) \binom{n}{t} (1-\theta)^n \left(\frac{\theta}{1-\theta}\right)^t = 0 \ \forall \theta \in (0,1) \xRightarrow{p = \frac{\theta}{1-\theta}}$$

$$\Rightarrow \sum_{t=0}^{n} \varphi(t) \binom{n}{t} p^t = 0 \ \forall \theta \in (0,1)$$

(Παρατηρώ ότι είναι πολυώνυμο n-οστού βαθμού ως προς p)

Το άθροισμα είναι ένα πολυώνυμο (ως προς p) n-οστού βαθμού. Επειδή είναι ίσο με 0 $\forall \theta$ άρα $\forall p$ έχει άπειρες ρίζες. Αυτό όμως δεν μπορεί να συμβαίνει εκτός και αν είναι το μηδενικό πολυώνυμο δηλαδή οι συντελεστές να είναι ίσοι με μηδέν. (Αφού κάθε πολυώνυμο n-οστού βαθμού έχει το πολύ n ρίζες)

Άρα $\varphi(t) \binom{n}{t} = 0 \ \forall t \Rightarrow \varphi(t) = \ \forall t$

Άρα το $T(\underline{X}) = \sum_{i=1}^{n} x_i$ είναι επαρκές και πλήρες στατιστικό.

<u>Βήμα 2:</u> Αρκεί να βρω συνάρτηση του T που να είναι αμερόληπτη της παραμέτρου θ.

$$T = \sum X_i \sim B(n,\theta)$$

Δοκιμάζω $E(T) = n\theta \Rightarrow \frac{1}{n}E(T) = \theta \Rightarrow E\left(\frac{1}{n}T\right) = \theta$

$\Rightarrow E\left(\frac{1}{n}\sum X_i\right) = \theta \Rightarrow E(\bar{X}) = \theta$

Συγκεντρωτικά $\bar{X}$ συνάρτηση επαρκούς και πλήρους και αμερόληπτη θ. Άρα $\bar{X}$ ΑΟΕΔ θ!

<u>Για το θ^2:</u> $Var(T) = n\theta(1-\theta)$

και $Var(T) = E(T^2) - (ET)^2$

Δοκιμάζω $E(T^2) = Var(T) + (ET)^2 = n\theta(1-\theta) + (n\theta)^2 \Rightarrow$

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

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

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

Συγκεντρωτικά ο εκτιμητής $\frac{T^2-T}{n(n-1)}$ είναι συνάρτηση επαρκούς και πλήρους αμερόληπτη θ^2. Άρα $\frac{T^2-T}{n(n-1)}$ ΑΟΕΔ θ^2!