Επανάληψη:

$X_1,\dots,X_n \to f(x,\theta),\ \theta\in\Theta\subseteq\mathbb{R}$

$T=T(X_1,\dots,X_n)$ ← στατιστικό

Αμεροληψία + ελάχιστη διακύμανση ⟿ ΑΟΕΔ

$\uparrow$

$K\Phi_{C\text{-}R}$

Επάρκεια - Επαρκής Στατιστικού

Διαισθητικά: Επαρκές Στατιστικό T

Πληροφορία$(x_1,\dots,x_n) \equiv$ Πληρ$(T=T(x_1,\dots,x_n))$

ΟΡΙΣΜΟΣ ΕΠΑΡΚΕΙΑΣ: $T=T(x_1,\dots,x_n)$ επαρκής (-ές) αν η κατανομή $\underset{\sim}{X}\,|\,T$ ανεξάρτητη της θ

ΠΑΡΑΓΟΝΤΙΚΟ ΘΕΩΡΗΜΑ NEYMAN-FISHER:

Το $T=T(x_1,\dots,x_n)$ επαρκές αν-ν $f(\underset{\sim}{x},\theta)=q(T(\underset{\sim}{x}),\theta)\,h(\underset{\sim}{x})$

ΠΑΡΑΔΕΙΓΜΑ: Έστω $x_1,\dots,x_n$ τ.δ. από Poisson(θ), $\theta>0$. Να βρεθεί επαρκές στατιστικό

$$p(\underset{\sim}{x},\theta)=\prod_{i=1}^{n}p(x_i,\theta)=\prod_{i=1}^{n}\frac{e^{-\theta}\theta^{x_i}}{x_i!}=\frac{e^{-n\theta}\theta^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i!}=\underbrace{e^{-\theta n}\,\theta^{\sum_{i=1}^{n}x_i}}_{q\left(\sum_{i=1}^{n}x_i,\,\theta\right)}\ \underbrace{\frac{1}{\prod_{i=1}^{n}(x_i)}}_{h(\underset{\sim}{x})}$$

$$\Downarrow$$

$$T(\underset{\sim}{x})=\sum x_i \text{ επαρκές}$$

ΠΑΡΑΔΕΙΓΜΑ: Έστω $x_1,\dots,x_n$ τ.δ. από πληθυσμό $N(\mu,\sigma^2)$. Να βρεθεί επαρκές στατιστικό στις ακόλουθες περιπτώσεις

(I) Όταν $\sigma^2=$ γνωστό και $\mu=\theta$ άγνωστο

(II) Όταν $\mu=$ γνωστό και $\sigma^2=\theta$ άγνωστο

(III) Όταν μ,σ^2 άγνωστες:

$$f(x,\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad \sigma^2>0,\ \mu\in\mathbb{R}$$

(I) $f(\underset{\sim}{x},\theta) = \prod_{i=1}^{n} f(x_i,\theta) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i-\theta)^2} = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2} =$

$= -\|- \, e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i + \frac{\theta}{\sigma^2}\sum x_i - \frac{\theta^2}{2\sigma^2}}$

$= q(T(\underset{\sim}{x}),\theta)\, h(\underset{\sim}{x})$

όπου $h(\underset{\sim}{x}) = e^{-\frac{1}{2\sigma^2}\sum x_i^2}$ και $q(T(\underset{\sim}{x}),\theta) = e^{\frac{\theta}{\sigma^2}\sum x_i}\, e^{-\frac{\theta^2}{2\sigma^2}}$

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

$= q(T(\underset{\sim}{x}),\theta)\, h(\underset{\sim}{x})$

$h(\underset{\sim}{x}) = 1$

$q(T(\underset{\sim}{x}),\theta) = \frac{1}{(2\pi\theta)^{n/2}} e^{-\frac{1}{2\theta}\sum_{i=1}^{n}(x_i-\mu)^2}$

$T(\underset{\sim}{x}) = \sum_{i=1}^{n}(x_i-\mu)^2$

Επαρκής

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

$= \frac{1}{(\sigma\sqrt{2\pi})^{n}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 + \frac{\mu}{\sigma^2}\sum_{i=1}^{n} x_i - \frac{n\mu^2}{2\sigma^2}} = q(T_1(\underset{\sim}{x}), T_2(\underset{\sim}{x}), \mu, \sigma^2)\, h(\underset{\sim}{x})$ όπου $h(\underset{\sim}{x}) = 1$

$q(T_1(\underset{\sim}{x}), T_2(\underset{\sim}{x}), \mu, \sigma^2) = \frac{1}{(\sigma\sqrt{2\pi})^{n}} e^{-\frac{1}{2\sigma^2}\sum x_i^2 + \frac{\mu}{2\sigma^2}\sum x_i}\, e^{-\frac{n\mu^2}{2\sigma^2}}$

$(T_1(\underset{\sim}{x}), T_2(\underset{\sim}{x})) = \left(\sum_{i=1}^{n} x_i,\ \sum_{i=1}^{n} x_i^2\right)$ Επαρκής στατιστικό (μ,σ^2)

ΠΑΡΑΔΕΙΓΜΑ: Έστω $x_1, \dots, x_n$ τ.δ από πληθυσμό $U(\theta, \theta+1)$, $\theta > 0$ θ άγνωστο.
Να βρεθεί επαρκές στατιστικό

Υπενθύμιση: W $W \sim U(a, \beta)$ αν $f_W(w) = \begin{cases} \frac{1}{\beta - a}, & a < w < \beta \\ 0, & \text{αλλού} \end{cases}$

Στην περίπτωσή μας $f(x, \theta) = \begin{cases} \frac{1}{\theta+1-\theta}, & \theta < x < \theta+1 \\ 0, & \text{αλλού} \end{cases}$

$$f(x, \theta) = \begin{cases} 1, & \theta < x < \theta+1 \\ 0, & \text{αλλού} \end{cases}$$

$$f(\underset{\sim}{x}, \theta) = \prod_{i=1}^{n} f(x_i, \theta) = \prod_{i=1}^{n} 1 \, I_{(\theta, \theta+1)}(x_i) \quad \text{όπου} \quad I_A(w) \overset{\text{ορ}}{=} \begin{cases} 1, & w \in A \\ 0, & w \notin A \end{cases}$$

$$= \prod_{i=1}^{n} I_{(\theta, \theta+1)}(x_i) = \begin{cases} 1, & \theta < x_i < \theta+1, \ \forall i \\ 0, & \text{αλλού} \end{cases}$$

$$\theta < x_i < \theta+1 \ \ \forall i \longrightarrow \theta < \min x_i < \max x_i < \theta+1 \Rightarrow$$
$$\Rightarrow \theta < x_{(1)} \leq x_{(n)} < \theta+1.$$

$$f(\underset{\sim}{x}, \theta) = \begin{cases} 1, & \theta < x_{(1)} \leq x_{(n)} < \theta+1 \\ 0, & \text{αλλού} \end{cases}$$

$$= I_{(\theta, \theta+1)}(x_{(n)}) \cdot I_{(\theta, x_{(n)})}(x_{(1)}) = q(T_1(\underset{\sim}{x}), T_2(\underset{\sim}{x}), \theta) h(\underset{\sim}{x})$$

όπου $h(\underset{\sim}{x}) = 1$

$$q(T_1(\underset{\sim}{x}), T_2(\underset{\sim}{x}), \theta) = I_{(\theta, \theta+1)}(x_{(n)}) I_{(\theta, x_{(n)})}(x_{(1)})$$

Άρα $(x_{(1)}, X_{(n)})$ επαρκές για θ

ΠΑΡΑΤΗΡΗΣΕΙΣ: ① Το τδ. $x_1, \dots, x_n$ είναι ένα πετριμμένο επαρκές στατιστικό

Πράγματι $f(\underset{\sim}{x}, \theta) = q(T(\underset{\sim}{x}), \theta) h(\underset{\sim}{x})$

με $h(\underset{\sim}{x}) = 1$ και $T = T(\underset{\sim}{x}) = (x_1, \dots, x_n)$

② Έστω T_1 επαρκές και W μια 1-1 συνάρτηση. Τότε $T_2 = W(T_1)$ είναι επαρκές

Απόδειξη: Αφού T_1 επαρκές

$$f(\underset{\sim}{x}, \theta) = q(T_1(\underset{\sim}{x}), \theta) h(\underset{\sim}{x}) \xrightarrow[\exists\, W^{-1}]{W \text{ είναι } 1\text{-}1}$$

$$\Rightarrow f(\underset{\sim}{x}, \theta) = q(W^{-1}(T_2), \theta) h(\underset{\sim}{x}) = q^*(T_2(x, \theta)) h(\underset{\sim}{x})$$

$$\Downarrow$$

T_2 επαρκές

③ Ισχύει $I_{\underset{\sim}{x}}(\theta) \geq I_{T(\underset{\sim}{x})}(\theta)$ με ισότητα εάν-ν $T(\underset{\sim}{x})$ επαρκές

όπου $I_{\underset{\sim}{x}}(\theta) = n I_x(\theta)$

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

$$I_{T(\underset{\sim}{x})}(\theta) \overset{\text{ορ.}}{=} E\left(\frac{\partial}{\partial \theta} \log f_T(T, \theta)\right)^2 = -E\left(\frac{\partial^2}{\partial \theta^2} \log f_T(T, \theta)\right)$$

ΑΞΙΟΠΟΙΗΣΗ ΕΠΑΡΚΕΙΑΣ ΣΤΗ ΒΕΛΤΙΩΣΗ ΕΚΤΙΜΗΤΩΝ

ΘΕΩΡΗΜΑ RAO-BLACKWELL: Έστω $T(\underset{\sim}{x})$ επαρκές στατιστικό και έστω $S(\underset{\sim}{x})$ είναι εκτιμητής της $g(\theta)$. Ορίζω $S^* = E(S/T)$

① $E(S^*) = E(S)$. Άρα αν S αμερόληπτος της $g(\theta)$, τότε S^* αμερ. $g(\theta)$

② $Var(S^*) \leq Var(S)$

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

③ $ΜΤΣ(S^*, \theta) \leq ΜΤΣ(S, \theta)$

Απόδειξη:

Από Θεωρία Πιθανοτήτων: Αν X_1 και X_2 είναι δύο τ.μ. ΤΟΤΕ:

ⓐ $E(E(X_1|X_2)) = E(X_1)$

ⓑ $Var(X_1) = E[Var(X_1|X_2)] + Var(E(X_1|X_2))$

Αποδ. 1: $E(S^*) = E[E(S|T)] \overset{(a)}{=} E(S)$

Αποδ. 2 $Var(S) \overset{(b)}{=} E[Var(S|T)] + Var[E(S|T)]$
$= E[Var(S|T)] + Var(S^*)$

Πάντα $Var \geq 0$
Αν $h(x) \geq 0$ τότε $E(h(x)) \geq 0$

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

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

Αποδ.3 Άμεση από τα 1 και 2 και $MTS(T, g(\theta)) = Var(T) + [E(T) - g(\theta)]^2$

Πρακτική Αξία Θ. R-B
έγκειται στη βελτίωση εκτιμητών μέσω ενός επαρκούς T.
Πως;
Αν S εκτιμητής και $g(\theta)$ + T επαρκής στατιστικό
ΤΟΤΕ
ο $S^* \overset{op}{=} E(S|T)$ είναι ο βελτιωμένος εκτιμητής αφού ισχύει 1, 2, 3.

Παρ' ότι το Θ. R-B βελτιώνει ένα εκτιμητή, ο S^* που λέγεται βελτίωση R-B του S, δεν είναι ΑΟΕΔ.

Μπορώ να κάνω τον S^* ΑΟΕΔ;
Απάντηση: ΝΑΙ
Αν ικανοποιείται η έννοια της πληρότητας.

ΟΡΙΣΜΟΣ (ΠΛΗΡΟΤΗΤΑΣ): Η στατιστική συνάρτηση $T=T(\underset{\sim}{X})$ λέγεται πλήρης αν η σχέση $E[\varphi(T)]=0,\ \forall\ \theta\in\Theta$ συνεπάγεται ότι $\varphi(T)=0$, δηλ. $\varphi(t)=0$ $\forall\ t$ τιμής του T

ΔΙΑΙΣΘΗΤΙΚΑ: Αν η μόνη αμερόληπτη εκτιμήτρια του 0 είναι η μηδενική συνάρτηση.

ΘΕΩΡΗΜΑ LEHMAN - SCHEFFE: Έστω τ.δ. $X_1,\dots,X_n$ από πληθυσμό με κατανομή $f(x,\theta)$ $\theta\in\Theta\subseteq\mathbb{R}$. Έστω $T=T(X_1,\dots,X_n)$ είναι επαρκής και πλήρης. 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)

ΑΠΟΔΕΙΞΗ:

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

$$\left.\begin{array}{l}\text{Τότε: } E(S^*)=g(\theta)\\[2em] \text{Αλλά: } E(S)=g(\theta)\end{array}\right\}\Rightarrow E(S^*)=E(S)\Rightarrow E(S^*-S)=0\Rightarrow$$

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

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

$$\Big\Downarrow \text{ πλήρης}$$

$$S^*(t)=S(t)\ \forall t$$

Άρα ο αμερόληπτος εκτιμητής της $g(\theta)$ που είναι συνάρτηση του επαρκούς και πλήρους είναι μοναδικός.

Έστω S^{**} ένας αμερόληπτος της $g(\theta)$ που δεν είναι κατ' ανάγκη συνάρτηση του επαρκούς και πλήρους T. Θεωρώ την Rao-Blackwell βελτίωση του S^{**} δηλ. θεωρώ τον εκτιμητή $E(S^{**}/T)$

Ισχύει: $E[E(S^{**}|T)]=E(S^{**})=g(\theta)$

Άρα η Rao-Blackwell βελτίωση $E(S^{**}|T)$ είναι αμεροληπτη.

Αποδείξαμε όμως ότι ΜΟΝΟ ένας αμερόληπτος υπάρχει που να είναι συνάρτηση του επαρκούς και πλήρους.

Άρα $E(S^{**}|T)=S$

Από (β) θεωρ. Rao-Blackwell
Ισχύει: $Var(\underbrace{E(S^{**}|T)}_{S}) \leq Var(S^{**})$

$\Downarrow$

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

$\Downarrow$

S ΑΟΕΔ αφού είναι αμερολ. της $g(\theta)$ και η διακύμ. του $S \leq$ από τη διακ. οποιουδήποτε άλλου αμερ. S^{**}

ΕΦΑΡΜΟΓΗ Θ. LEHMANN-SCHEFFE ΣΤΗΝ ΕΥΡΕΣΗ ΑΟΕΔ ΕΚΤΙΜΗΤΩΝ

Βήμα 1: Βρίσκω το επαρκές και πλήρες T

Βήμα 2: Προσπαθούμε να βρούμε συνάρτηση του T που να είναι αμερόληπτος εκτιμητής της $g(\theta)$
ΤΟΤΕ
η συνάρτηση αυτή του επαρκούς και πλήρους είναι ΑΟΕΔ.