Έστω τ.δ. $X_1,\dots,X_n$ από πληθ. $N(\mu,\overset{\sigma^2}{\theta})$, μ: γνωστό

Ζητώ ΑΟΕΔ της $g(\theta)=\theta$

__Λύση__ (Α' τρόπος)

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

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

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

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

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

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

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

$$= -\frac{1}{2\theta^2}+\frac{1}{\theta^3}E\big((X-E(X))\big)^2 = -\frac{1}{2\theta^2}+\frac{1}{\theta^3}Var(X) = -\frac{1}{2\theta^2}+\frac{1}{\theta^3}\theta$$

Έτσι το μέτρο πληροφορίας του Fisher είναι: $I_x(\theta)=\frac{1}{2\theta^2}$

Άρα $K\Phi_{CR} = \frac{[g'(\theta)]^2}{nI_x(\theta)} = \frac{1}{n/2\theta^2} = \frac{2\theta^2}{n}$

2) Δοκιμάζω το $S^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar X)^2$, $S^{*2}=\frac{1}{n-1}\sum_{i=1}^n(X_i-\mu)^2$

$E(S^{*2}) = \frac{1}{n-1}\sum_{i=1}^n E(X_i-\mu)^2 = \frac{1}{n-1}\sum_{i=1}^n Var(X_i) = \frac{1}{n-1}\sum_{i=1}^n\theta = \frac{n}{n-1}\theta$

άρα ο S^{*2} δεν είναι αμερόλ. εκτιμητής, παίρνω:

$E(\frac{n-1}{n}S^2)=\theta$ άρα $\frac{n-1}{n}S^{*2} = \cancel{\frac{n-1}{n}\frac{n}{n-1}}\Sigma(X_i-\mu)^2$

Άρα θεωρώ εκτιμητή: $T=\frac{1}{n}\sum_{i=1}^n(X_i-\mu)^2$

Ο T είναι αμερόληπτος της θ, αρκεί ν.δ.ο. $Var(T)=K\Phi_{CR}$

$Var(T) = Var\left(\frac{1}{n}\sum_{i=1}^n(X_i-\mu)^2\right) = \frac{1}{n^2}Var\left(\sum_{i=1}^n(X_i-\mu)^2\right) \underset{\text{ανεξ}}{\underline{X_i,X_n}}$

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

$X_i\sim N(\mu,\theta)\ \forall i \Rightarrow \frac{X_i-\mu}{\sqrt\theta}\sim N(0,1)\ \forall i \Rightarrow \frac{(X_i-\mu)^2}{\theta}\sim N^2(0,1)\equiv\chi_1^2$

άρα $Var\frac{(X_i-\mu)^2}{\theta}\equiv Var(\chi_n^2)$. Ισχύει: $E(\chi_\nu^2)=\nu$ και $Var(\chi_\nu^2)=2\nu$

άρα $Var\frac{(X_i-\mu)^2}{\theta} = Var(\chi_1^2)=2 \Rightarrow \frac{1}{\theta^2}Var(X_i-\mu)^2=2$

$\Rightarrow Var(X_i-\mu)^2=2\theta^2\ \forall i$

και άρα ο T είναι ΑΟΕΔ εκτιμητής

Β' Τρόπος

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

$$\log \prod_{i=1}^{n} f(x_i,\theta) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\theta - \frac{1}{2\theta}\Sigma(x_i-\mu)^2$$

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

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

Άρα $T(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$ ΑΟΕΔ του θ

Αποτελεσματικός Εκτιμητής - Σχετική Αποτελεσματικότητα

Ορισμός

Ο εκτιμητής $U(\underline{X})$ της $g(\theta)$ θα λέγεται αποτελεσματικός αν ο $U(\underline{X})$ είναι ο ΑΟΕΔ της $g(\theta)$

Ορισμός

Έστω $U(\underline{X})$ αποτελεσματικός εκτιμητής της $g(\theta)$.
Αν $T(\underline{X})$ αμερόληπτος της $g(\theta)$ η σχετική αποτελεσματικότητα του $T(\underline{X})$ (ως προς τον $U(\underline{X})$) ορίζεται: $\frac{Var(U(\underline{X}))}{Var(T(\underline{X}))} \in [0,1]$

Πχ

Έστω ένα τ.δ. X_1, X_2, X_3 από πληθ. Poisson (θ)
Ν.δ.ο. ο $U = \bar{X}$ είναι αποτελεσματικός και να υπολογισθεί η σχετική αποτελεσματικότητα του $T = \frac{X_1+2X_2+3X_3}{6}$ ως προς U

Λύση

Αποδείξαμε ότι $U = \bar{X}$ είναι ΑΟΕΔ, άρα ο U είναι αποτελ/κός.

$$E(T) = E\left(\frac{X_1+2X_2+3X_3}{6}\right) = \frac{1}{6}E(X_1) + \frac{2}{6}E(X_2) + \frac{3}{6}E(X_3)$$

$$\underset{E(X_i)=\theta}{\overset{X_i\sim Poisson(\theta)}{=}} \frac{1}{6}\theta + \frac{2}{6}\theta + \frac{3}{6}\theta = \theta \Rightarrow T: \text{αμερ. της } \theta$$

$Var(U) = \frac{\theta}{3}$. Αποδείξαμε ότι $Var(\bar{X}) = \frac{\text{Διακύμανση πληθ}}{n}$

$$Var(T) = Var\left(\frac{X_1+2X_2+3X_3}{6}\right) = \frac{1}{36}Var(X_1) + \frac{4}{36}Var(X_2) + \frac{9}{36}Var(X_3)$$
$$= \frac{1}{36}\theta + \frac{4}{36}\theta + \frac{9}{36}\theta = \frac{17}{18}\theta$$

άρα $\frac{Var(U(\underset{\sim}{X}))}{Var(T(\underset{\sim}{X}))} = \frac{\theta/3}{70/28} = 0,85 = 85\%$

$\Rightarrow$ ο T υπολείπεται κατά 15% της ποιότητας της U

Εκθετική Οικογένεια Κατανομών (ΕΟΚ)

Ορισμός

Έστω τυχ. διάν. $\underset{\sim}{X} = (X_1, \dots, X_n)^T$, $n \geq 1$.
Η κατανομή του $\underset{\sim}{X}$ ανήκει στην εκθετική οικογένεια κατανομών αν

1) Το $S = \{(x_1, \dots, x_n) : f(x_1, \dots, x_n) > 0\}$ δεν εξαρτάται απ' το θ
2) Αν η κατανομή του $X = (X_1, \dots, X_n)$ είναι της μορφής
$f(\underset{\sim}{x}, \theta) = e^{A(\theta) + B(\underset{\sim}{x}) + C(\theta)D(\underset{\sim}{x})}$, $\underset{\sim}{x} \in S$

Παρατήρηση

Αν $n = 1$, έστω η τ.μ. X. Η κατανομή της X ανήκει στην ΕΟΚ αν $S = \{x : f(x) > 0\}$ ανεξ. της θ και
$f(x, \theta) = e^{A(\theta) + B(x) + C(\theta)D(x)}$

Μέλη της ΕΟΚ

$N(\theta, 1)$, $N(\mu, \theta^2)$, $B(n, \theta)$, $Εκθ(\theta)$, $Poisson(\theta)$

Όχι μέλος της ΕΟΚ

$U(0, \theta)$, Cauchy: $\frac{1}{\pi(1+(x-\theta)^2)}$, $x \in \mathbb{R}$, $\theta \in \mathbb{R}$

$\boxed{\text{Για την Poisson βλέπουμε}}$

$P(\theta)$: $p_X(x) = \frac{e^{-\theta}\theta^x}{x!}$, $x = 0, 1, 2, \dots$
$p_X(x) > 0 \quad \forall\, x = 0, 1, 2, \dots$
$S = \{0, 1, 2, \dots\}$ ανεξ. θ
$p_X(x) = \frac{e^{-\theta}\theta^x}{x!} = e^{-\theta - \log x! + x\log\theta} = e^{A(\theta) + B(x) + C(\theta)D(x)}$

όπου $A(\theta) = -\theta$, $C(\theta) = \log\theta$ άρα είναι ΕΟΚ
$B(x) = -\log x!$, $D(x) = x$

Πρόταση

Αν η κατανομή του $\underset{\sim}{X}$ ανήκει στην ΕΟΚ τότε οι συνθήκες τις ανισότητας των Cramer-Rao ικανοποιούνται

Επάρκεια - Επαρκές Στατιστικό

Στατιστικός: τ.δ. $X_1, \ldots, X_n$ από πληθ. $f(x,\theta)$, $\theta \in \Theta$

$T = T(\underset{\sim}{X})$, $T^* = T^*(\underset{\sim}{X})$

Διαισθητικά: Επαρκές στατιστικό T είναι μία συνάρτηση των τ.δ. $X_1, \ldots, X_n$ που περιέχει πληροφορία για το θ όση και τα αρχικά δεδομένα $X_1, \ldots, X_n$

Ορισμός (Επαρκούς στατιστικού)

Έστω τ.δ. $X_1, \ldots, X_n$ από πληθυσμό με κατανομή $f(x,\theta)$, $\theta \in \Theta$
Η στατιστική συνάρτηση $T = T(X_1, \ldots, X_n)$ θα λέγεται επαρκής για την κατανομή $f(x,\theta)$ (ή επαρκής για την θ) αν η δεσμευμένη κατανομή του $\underset{\sim}{X} | T(\underset{\sim}{x}) = t$ είναι ανεξάρτητη της παραμέτρου θ, $\forall \theta \in \Theta$

Ερώτημα

Πως μπορώ να κατασκευάσω επαρκής στατιστικό?
↳ Neyman-Fisher ⇝ Θεώρημα

Θεώρημα (Neyman-Fisher)

Έστω τ.δ. $X_1, \ldots, X_n$ από πληθ. $f(x,\theta)$, $\theta \in \Theta$
Η στατ. συνάρτηση $T = T(X_1, \ldots, X_n)$ είναι επαρκής αν-ν υπάρχουν συναρτήσεις $g(\cdot)$ και $h(\cdot)$ τέτοιες ώστε:
$f(\underset{\sim}{x},\theta) = g(T(x_1, \ldots, x_n), \theta) h(x_1, \ldots, x_n)$

Απόδειξη:

Είναι δύσκολη γι' αυτό και δεν την κάναμε καν

Ασκ

Έστω τ.δ. $X_1,\dots,X_n$ από πληθ. $B(n,\vartheta)$, $0<\vartheta<1$. Να βρεθεί επαρκές

Λύση

$$f(\underline{x},\vartheta)=\prod_{i=1}^{n}P_{X_i}(x_i,\vartheta)=\prod_{i=1}^{n}\binom{n}{x_i}\vartheta^{x_i}(1-\vartheta)^{n-x_i}=\left(\prod_{i=1}^{n}\binom{n}{x_i}\right)\vartheta^{\Sigma x_i}(1-\vartheta)^{n-\Sigma x_i}$$

$$=\vartheta^{\Sigma x_i}(1-\vartheta)^{-\Sigma x_i}(1-\vartheta)^{n}\left(\prod_{i=1}^{n}\binom{n}{x_i}\right)=\left(\frac{\vartheta}{1-\vartheta}\right)^{\Sigma x_i}(1-\vartheta)^{n}\left(\prod_{i=1}^{n}\binom{n}{x_i}\right)$$

$=g[T(\underline{x}),\vartheta]\,h(\underline{x})$, όπου $h(\underline{x})=\prod_{i=1}^{n}\binom{n}{x_i}$ και

$$g(T(\underline{x}),\vartheta)=(1-\vartheta)^{n}\left(\frac{\vartheta}{1-\vartheta}\right)^{\Sigma x_i}$$

Άρα N-F το $T(\underline{X})=\sum_{i=1}^{n}X_i$ επαρκής

Ασκ

Έστω τ.δ. $X_1,\dots,X_n$ από πληθ. $N(\mu,\sigma^2)$

Να βρεθεί επαρκές στατιστικό στις ακόλουθες περιπτώσεις:

(I) σ^2 γνωστό, $\mu=\vartheta$ άγνωστο

(II) μ γνωστό, $\sigma^2=\vartheta$ άγνωστο

(III) μ,σ^2 άγνωστα

Λύση

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

$$=\frac{1}{(\sigma\sqrt{2\pi})^n}e^{-\frac{1}{2\sigma^2}(\Sigma x_i^2-2\vartheta\Sigma x_i+n\vartheta^2)}$$

$$=\frac{1}{(\sigma\sqrt{2\pi})^n}e^{-\frac{1}{2\sigma^2}\Sigma x_i^2+\frac{\vartheta}{\sigma^2}\Sigma x_i-\frac{n\vartheta^2}{2\sigma^2}}=g[T(\underline{x}),\vartheta]\,h(\underline{x})$$

όπου $h(\underline{x})=\frac{1}{(\sigma\sqrt{2\pi})^n}e^{-\frac{1}{2\sigma^2}\Sigma x_i^2}$ και

$$g[T(\underline{x}),\vartheta]=e^{\frac{\vartheta}{\sigma^2}\Sigma x_i-\frac{n\vartheta^2}{2\sigma^2}}=e^{\frac{\vartheta}{\sigma^2}T(\underline{x})-\frac{n\vartheta^2}{2\sigma^2}}$$

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

(II) $f(\underline{x},\vartheta)=\prod_{i=1}^{n} f(\underline{x},\vartheta)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\vartheta}}e^{-\frac{1}{2\vartheta}(x_i-\mu)^2}$

$$=\frac{1}{(2\pi\vartheta)^{n/2}}e^{-\frac{1}{2\vartheta}\sum_{i=1}^{n}(x_i-\mu)^2}=g[T(\underline{x}),\vartheta]h(\underline{x})$$

όπου $h(\underline{x})=1$ και $g[T(\underline{x}),\vartheta]=\frac{1}{(2\pi\vartheta)^{n/2}}e^{-\frac{1}{2\vartheta}\sum(x_i-\mu)^2}$

Άρα $T(\underline{x})=\sum(x_i-\mu)^2$ επαρκές

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

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

$=g[T(\underline{x}),\vartheta]h(\underline{x})$ όπου $h(\underline{x})=1$ και

$$g[T(\underline{x}),\mu,\sigma^2]=\frac{1}{(\sigma\sqrt{2\pi})^n}e^{-\frac{1}{2\sigma^2}\sum x_i^2+\frac{\mu}{\sigma^2}\sum x_i-\frac{n\mu^2}{2\sigma^2}}$$

$=g((T_1(x_1,\ldots,x_n),T_2(x_1,\ldots,x_n),\mu,\sigma^2)$

με $T_1(x_1,\ldots,x_n)=\sum x_i^2$ και

$T_2(x_1,\ldots,x_n)=\sum x_i$

Άρα το $(T_1(\underline{x}),T_2(\underline{x}))=(\sum_{i=1}^{n}x_i^2,\sum_{i=1}^{n}x_i)$ είναι

επαρκές για το (μ,σ^2)