ΑΣΚΗΣΕΙΣ ΣΤΗΝ ΕΚΤΙΜΗΤΙΚΗ

1) Έστω $X_1,\ldots, X_n$ και $Y_1,\ldots,Y_m$ ανεξάρτητα τ.μ. από πληθυσμούς με μέση τιμή θ και γνωστές διασπορές σ_1^2 και σ_2^2. Δείξτε ότι για $c \in [0,1]$ η $U = c\overline{X} + (1-c)\overline{Y}$ είναι αμερόληπτη εκτιμήτρια της παραμέτρου θ και βρείτε το c για το οποίο η διασπορά της U είναι ελάχιστη.

Λύση:

Έστω $U = c\overline{X} + (1-c)\overline{Y}$, $c \in [0,1]$. Τότε:

$$E(U) = cE(\overline{X})+(1-c)E(\overline{Y}) = cE(\frac{1}{n}\sum_{i=1}^{n}X_i)+(1-c)E(\frac{1}{m}\sum_{i=1}^{m}Y_i) = c\frac{1}{n}n\theta+(1-c)\frac{1}{m}m\theta = \theta\ ,$$

δηλαδή η U είναι αμερόληπτη εκτιμήτρια της θ.

Ακόμα:

$$Var(U) = c^2Var(\overline{X})+(1-c)^2Var(\overline{Y}) = c^2Var(\frac{1}{n}\sum_{i=1}^{n}X_i)+(1-c)^2Var(\frac{1}{m}\sum_{i=1}^{m}Y_i)\ .$$

Αλλά X_i, Y_j (i=1,..n; j=1,..,m) ανεξάρτητα, άρα και ασυσχέτιστα, οπότε η τελευταία έκφραση παίρνει την μορφή:

$$Var(U)=c^2\frac{1}{n^2}n\sigma_1^2+(1-c)^2\frac{1}{m^2}m\sigma_2^2 = g(c)\ .$$

Η τελευταία συνάρτηση ελαχιστοποιείται στο σημείο που μηδενίζεται η παράγωγος της, δηλαδή:

$$g'(c)=2c\frac{\sigma_1^2}{n}-2(1-c)\frac{\sigma_2^2}{m}=0 \Rightarrow c^* = \frac{\frac{\sigma_2^2}{m}}{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}\ .$$

2) Έστω $X_1,\ldots, X_n$ τ.δ. από την $U(\theta,2\theta)$, $\theta \in \mathbb{R}$.

i) Να βρεθεί η σ.π.π. της $Y = \max\{X_1,\ldots, X_n\}$.

ii) Δείξτε ότι η $T=\frac{n+1}{2n+1}Y$ είναι αμερόληπτη εκτιμήτρια της θ.

Λύση:

i) Θυμίζουμε ότι όταν μια τ.μ.X ακολουθεί την U(α,β) έχει συνάρτηση κατανομής:

$$F(x)=\begin{cases} 0, & x < \alpha \\ \frac{x-\alpha}{\beta-\alpha}, & \alpha \le x < \beta \\ 1, & \beta \le x \end{cases}$$

Ενδιαφερόμαστε για την συνάρτηση κατανομής της τ.μ Υ.

$$F(y) = P(Y \le y) = P(\max\{X_1,\ldots,X_n\} \le y) = P(X_1 \le y,\ldots,X_n \le y) \overset{X_i \text{ ανεξάρτητα}}{=} \prod_{i=1}^{n} P(X_i \le y) \overset{X_i \sim U(\theta,2\theta)}{=} \left(\frac{y-\theta}{\theta}\right)^n.$$

Άρα η σ.π.π. της Υ είναι:

$$f(y) = F^{'}(y) = n\left(\frac{y-\theta}{\theta}\right)^{n-1}\frac{1}{\theta} = n\frac{(y-\theta)^{n-1}}{\theta^n},\ y\in(\theta,2\theta).$$

ii) Είναι:

$$E(Y) = \int_\theta^{2\theta} n\frac{y(y-\theta)^{n-1}}{\theta^n}dy = \frac{y(y-\theta)^n}{\theta^n}\Big|_\theta^{2\theta} - \int_\theta^{2\theta}\frac{(y-\theta)^n}{\theta^n}dy = 2\theta-\frac{\theta}{n+1} = \frac{2n+1}{n+1}\theta.$$

Άρα $E(\frac{n+1}{2n+1}Y) = \theta$, οπότε η $T=\frac{n+1}{2n+1}Y$ είναι αμερόληπτη εκτιμήτρια της θ.

3) Έστω $X_1,\ldots, X_n$ τ.δ. από πληθυσμό με σ.π.π $f(x) = a^2xe^{-ax}$, $x>0$.

i) Να βρεθεί η Ε.Μ.Π. της α.
ii) Να βρεθεί η Ε.Μ.Π. της μέσης τιμής του πληθυσμού.

Λύση:

ι) Καταρχήν παρατηρούμε ότι η συγκεκριμένη σ.π.π. είναι η G(a, p=2). Η συνάρτηση πιθανοφάνειας είναι:

$$L(a) = \prod_{i=1}^{n} f(x_i) = a^{2n}\left(\prod_{i=1}^{n} x_i\right)e^{-a\sum_{i=1}^{n}x_i} \Rightarrow l(a) = 2n\cdot\ln(a) + \ln(\sum_{i=1}^{n} x_i) - a\sum_{i=1}^{n} x_i .$$

Άρα η Ε.Μ.Π. της α προκύπτει από την λύση της εξίσωσης:

$$l^{'}(a) = 0 \Rightarrow \frac{2n}{a} - \sum_{i=1}^{n} X_i = 0 \Rightarrow \hat{a} = \frac{2n}{\sum_{i=1}^{n} X_i}.$$

ii) Γνωρίζουμε ότι η μέση τιμή μιας G(a,p) ισούται με $\mu = \frac{p}{a}$. Στην συγκεκριμένη οπότε περίπτωση $\mu = \frac{2}{a}$. Με χρήση του Θεωρήματος 9.1 η Ε.Μ.Π τότε της μ θα είναι:

$$\hat{\mu} = \frac{2}{\hat{a}} = \frac{2}{\frac{2n}{\sum_{i=1}^{n} X_i}} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}.$$

4) Έστω $X_1,\dots,X_n$ τ.δ. από την G(a,p). Να προσδιοριστούν οι εκτιμήτριες των a, p με την μέθοδο των ροπών.

Λύση:

Γνωρίζουμε ότι η σ.π.π. μιας G(a,p) είναι η $f(x)=\dfrac{a^p}{\Gamma(p)}x^{p-1}e^{-ax}$.

Οι δύο πρώτες δειγματικές ροπές περί την αρχή είναι:

$$m_1' = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X} \quad \text{και} \quad m_2' = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \overline{X^2}.$$

Οι αντίστοιχες δύο πρώτες ροπές του πληθυσμού είναι:

$$\mu_1' = E(X) = \frac{p}{a} \quad \text{και} \quad \mu_2' = E(X^2) = \sigma^2 + \mu^2 = \frac{p(p+1)}{a^2}.$$

Άρα έχουμε το σύστημα:

$$\begin{cases} \dfrac{p}{a} = \overline{X} \\ \dfrac{p(p+1)}{a^2} = \overline{X^2} \end{cases} \Rightarrow \begin{cases} p' = \dfrac{\left(\overline{X}\right)^2}{\overline{X^2} - \left(\overline{X^2}\right)^2} \\ a' = \dfrac{\overline{X}}{\overline{X^2} - \left(\overline{X^2}\right)^2} \end{cases}.$$

5) Έστω $X_1,\dots,X_n$ τ.δ. από την U(a,b).

i) Να προσδιοριστούν οι εκτιμήτριες των a, b με την μέθοδο των ροπών.

ii) Να προσδιοριστούν οι Ε.Μ.Π των a, b.

Λύση:

Γνωρίζουμε ότι η σ.π.π. μιας U(a,b) είναι η $f(x)=\begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \text{διαφορετικά} \end{cases}$.

ι) Οι δύο πρώτες δειγματικές ροπές περί την αρχή είναι:

$$m_1' = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X} \quad \text{και} \quad m_2' = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \overline{X^2}.$$

Οι αντίστοιχες δύο πρώτες ροπές του πληθυσμού είναι:

$$\mu_1' = E(X) = \frac{a+b}{2} \quad \text{και} \quad \mu_2' = E(X^2) = \sigma^2 + \mu^2 = \frac{a^2+ab+b^2}{3}.$$

Άρα έχουμε το σύστημα:

$$\left\{\begin{array}{l}\dfrac{a+b}{2}=\overline{X}\\[2mm] \dfrac{a^2+ab+b^2}{3}=\overline{X^2}\end{array}\right. \Rightarrow \left\{\begin{array}{l} a^{'}=\overline{X}-\sqrt{3}\,\left(\overline{X^2}\text{-}\overline{X}^2\right)^{1/2}\\[2mm] b^{'}=\overline{X}+\sqrt{3}\,\left(\overline{X^2}\text{-}\overline{X}^2\right)^{1/2}\end{array}\right. .$$

ii) Η συνάρτηση πιθανοφάνειας είναι:

$$L(a,b)=\prod_{i=1}^{n}\frac{1}{b\text{-}a}\,I_{[a,b]}(x_i)=\left(\frac{1}{b\text{-}a}\right)^n\prod_{i=1}^{n}I_{[a,b]}(x_i)=\left(\frac{1}{b\text{-}a}\right)^n\prod_{i=1}^{n}\{I_{(-\infty,b]}(x_i)I_{[a,+\infty]}(x_i)\}$$

$$=\left(\frac{1}{b\text{-}a}\right)^n I_{(-\infty,b]}(\max x_i)I_{[a,+\infty]}(\min x_i).$$

Η συνάρτηση αυτή δεν διαφορίζεται παντού ως προς a και b, για να μεγιστοποιηθεί όμως είναι φανερό ότι θα πρέπει να ελαχιστοποιηθεί η διαφορά b-a με την προϋπόθεση ότι ισχύει $\max x_i \le b$ και $\min x_i \ge \alpha$.

Επομένως οι Ε.Μ.Π. των a, b είναι: $\hat{a}=\min x_i$ και $\hat{b}=\max x_i$.

6) Το πλάτος ενός παλμού είναι τ.μ. $X \sim N(\mu, 4)$. Στην έξοδο του μηχανήματος μπορούμε να παρατηρήσουμε αν το X υπερβαίνει την τιμή 40 ή όχι. Αν σε 100 παρατηρήσεις το X υπερέβη την τιμή αυτή 80 φορές, ποια είναι η Ε.Μ.Π. της μέσης τιμής μ;

Λύση:

Έστω $Y_i=\begin{cases}1, & x_i>40\\ 0, & x_i\le 40\end{cases}$. Γνωρίζουμε ότι $\hat{p}=\hat{P}(Y_i=1)=0.8$ και $(1\text{-}\hat{p})=\hat{P}(Y_i=0)=0.2$.

Τότε:

$$p=P(x>40)=P\left(\frac{X\text{-}\mu}{2}>\frac{40\text{-}\mu}{2}\right)=1-\Phi\left(\frac{40\text{-}\mu}{2}\right)\Rightarrow\Phi\left(\frac{40\text{-}\mu}{2}\right)=1-p.$$

Αν $\hat{p}$ η Ε.Μ.Π του p και $\hat{\mu}$ η Ε.Μ.Π. του μ τότε:

$$\Phi\left(\frac{40\text{-}\hat{\mu}}{2}\right)=1-\hat{p}=0.20\Rightarrow\hat{\mu}=41.68.$$

7) Ο αριθμός των σωματιδίων α που εκπέμπονται από ραδιενεργό πηγή σε χρόνο t(sec) ακολουθεί την P(λ), δηλαδή $P(X=k)=e^{-\lambda t}\dfrac{(\lambda t)^k}{k!}$. Για την εκτίμηση του λ καταγράφηκε ο αριθμός ν των πηγών που εξέπεμψαν ένα τουλάχιστον σωματίδιο σε χρόνο 1sec από 30 πηγές. Να προσδιοριστεί η Ε.Μ.Π της λ για ν = 20.

Λύση:

Έστω $Y_i = \begin{cases} 1, & x_i \geq 1 \\ 0, & x_i = 0 \end{cases}$.

Τότε $p = P(Y_i = 1) = P(X_i \geq 1) = 1 - P(X_i = 0) = 1 - e^{-\lambda}$.

Αν $\hat{p}$ η Ε.Μ.Π του p και $\hat{\lambda}$ η Ε.Μ.Π. του λ τότε:
$\hat{p} = 1 - e^{-\hat{\lambda}} \Rightarrow \frac{20}{30} = 1 - e^{-\hat{\lambda}} \Rightarrow \hat{\lambda} = \log(3)$.

8) Έστω $X_1,\ldots, X_n$ τ.δ. από την B(N,p) με N γνωστό. Να προσδιοριστεί η Ε.Μ.Π της p.

Λύση:

Η συνάρτηση πιθανοφάνειας είναι:

$$L(p) = \prod_{i=1}^{n} \{\binom{N}{x_i} p^{x_i} (1\text{-}p)^{N\text{-}x_i}\} = \prod_{i=1}^{n} \binom{N}{x_i}\, p^{\sum_{i=1}^{n} x_i} (1\text{-}p)^{(nN - \sum_{i=1}^{n} x_i)}$$
$$\Rightarrow l(p) = \log\left(\prod_{i=1}^{n} \binom{N}{x_i}\right) + \sum_{i=1}^{n} x_i \log(p) - (nN - \sum_{i=1}^{n} x_i)\log(1\text{-}p).$$

Άρα η Ε.Μ.Π. της p προκύπτει από την λύση της εξίσωσης:

$$l'(p) = 0 \Rightarrow \frac{\sum_{i=1}^{n} x_i}{p} - \frac{nN - \sum_{i=1}^{n} x_i}{1\text{-}p} = 0 \Rightarrow \hat{p} = \frac{\sum_{i=1}^{n} x_i}{nN} = \frac{\overline{X}}{N}.$$

Σημείωση: $\hat{p} = \frac{\overline{X}}{N} = \frac{k}{m}$, όπου $k = \sum_{i=1}^{n} x_i$ είναι ο συνολικός αριθμός επιτυχιών και $m = n \cdot N$ ο συνολικός αριθμός δοκιμών Bernoulli.

9) Έστω $X_1,\ldots, X_n$ τ.δ. από την $N(\mu, \sigma^2)$.
  i) Να βρεθεί η Ε.Μ.Π της μ όταν σ^2 γνωστό.
  ii) Να βρεθεί η Ε.Μ.Π της σ^2 όταν μ γνωστό.
  iii) Να βρεθούν οι Ε.Μ.Π. των μ, σ^2.

Λύση:

i) Η συνάρτηση πιθανοφάνειας είναι:

$$L(\mu)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}=(2\pi\sigma^2)^{-n/2}\cdot e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}\Rightarrow l(\mu)=-\frac{n}{2}\log(2\pi\sigma^2)-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}.$$

Άρα η Ε.Μ.Π. της μ προκύπτει από την λύση της εξίσωσης:

$$l'(\mu)=0\Rightarrow\frac{1}{2\sigma^2}\sum_{i=1}^{n}2(x_i-\mu)=0\Rightarrow\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)=0\Rightarrow\frac{n}{\sigma^2}(\bar{X}-\mu)=0\Rightarrow\hat{\mu}=\bar{X}.$$

ii) Η συνάρτηση πιθανοφάνειας είναι:

$$L(\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}=(2\pi\sigma^2)^{-n/2}\cdot e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}\Rightarrow l(\mu)=-\frac{n}{2}\log(2\pi\sigma^2)-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}.$$

Άρα η Ε.Μ.Π. της σ^2 προκύπτει από την λύση της εξίσωσης:

$$l'(\sigma^2)=0\Rightarrow-\frac{n}{2}\cdot\frac{2\pi}{2\pi\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(X_i-\mu)^2=0\Rightarrow-n+\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}=0\Rightarrow\widehat{\sigma^2}=\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{n}.$$

iii) Η συνάρτηση πιθανοφάνειας είναι:

$$L(\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}=(2\pi\sigma^2)^{-n/2}\cdot e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}\Rightarrow l(\mu)=-\frac{n}{2}\log(2\pi\sigma^2)-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}.$$

Άρα οι Ε.Μ.Π. των μ, σ^2 προκύπτουν από την λύση του συστήματος:

$$\begin{cases}\dfrac{\partial l(\mu,\sigma^2)}{\partial\mu}=0\\[2ex]\dfrac{\partial l(\mu,\sigma^2)}{\partial\sigma^2}=0\end{cases}\Rightarrow\begin{cases}\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)=0\\[2ex]-n+\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}=0\end{cases}\Rightarrow\begin{cases}\hat{\mu}=\bar{X}\\[2ex]\widehat{\sigma^2}=\dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}\end{cases}.$$

10) Έστω $X_1,\dots,X_n$ τ.δ. από πληθυσμό με συνάρτηση κατανομής

$$F(x;\theta)=1-(1+x^2)^{-\theta},\ x>0,\ \theta>0.$$

Να βρεθεί η Ε.Μ.Π της θ.

<u>Λύση:</u>

Εύκολα μπορούμε να υπολογίσουμε την σ.π.π. της παραπάνω κατανομής.

$$f(x;\theta)=F'(x;\theta)=2x\cdot\theta(1+x^2)^{-(\theta+1)},\ x>0.$$

Η συνάρτηση πιθανοφάνειας είναι:

$$L(\theta) = \prod_{i=1}^{n}\{\theta(1+x_i^2)^{-(\theta+1)} \cdot 2x_i\} = (2\theta)^n \frac{\prod_{i=1}^{n} x_i}{\prod_{i=1}^{n}(1+x_i^2)^{-(\theta+1)}}$$

$$\Rightarrow l(\theta) = n\log(2\theta) + \sum_{i=1}^{n}\log(x_i) - (\theta+1)\sum_{i=1}^{n}\log(1+x_i^2).$$

Άρα η Ε.Μ.Π. της θ προκύπτει από την λύση της εξίσωσης:

$$l^{'}(\theta) = 0 \Rightarrow \frac{n}{\theta} - \sum_{i=1}^{n}\log(1+x_i^2) = 0 \Rightarrow \hat{\theta} = \frac{\sum_{i=1}^{n}\log(1+x_i^2)}{n}.$$

Μιας και η συγκεκριμένη κατανομή δεν είναι κάποια από τις γνωστές κατανομές που έχουμε μάθει για να είμαστε σίγουροι ότι η παραπάνω λύση είναι μέγιστο θα πρέπει να δείξουμε ότι η δεύτερη παράγωγος είναι αρνητική για $\theta = \hat{\theta}$. Πράγματι:

$$l^{''}(\theta) = -\frac{n}{\theta} < 0 \ \forall \ \theta > 0.$$