

Έστω η κατανομή έως πλήθους θ . Θ μία άγνωστη παράμετρος του δείγματος που θέλουμε να εκτιμήσουμε

- Εκτιμήτρια συνάρτηση: Οποιαδήποτε συνάρτηση με την οποία θέλουμε να εκτιμήσουμε την θ , τέτοια ώστε:
 - Ορισμένη στο δείγμα
 - Ανεξάρτητη της παραμέτρου θ

Βασικά κριτήρια αξιολόγησης εκτιμητριών συναρτήσεων:

- αμεροληψία: $E(T) = \theta$ (εγκυρότητα)
- αποτελεσματικότητα: μικρή $Var(T)$ (αν $Var(T_1) < Var(T_2) \Rightarrow T_1$ αποτελεσματικότερη) (αξιοπιστία)
- συνέπεια: Αν $T_n \xrightarrow{P} \theta$ καθώς $n \rightarrow \infty$
 $\hookrightarrow$ Αν T_n αμεροληπτική (ή ασυμμετρικά αμεροληπτική) και $Var(T_n) \rightarrow 0$ καθώς $n \rightarrow \infty \Rightarrow T_n$ συνεπής

Μέθοδοι εύρεσης εκτιμητριών:

- Μέθοδος ροών
- Μέθοδος μέγιστης πιθανοφάνειας
- Μέθοδος ελαχίστων τετραγώνων $\rightarrow$ Παλινδρόμηση

Άσκηση 1

Wednesday, December 13, 2023 12:44 PM

Έστω αμερόληπτες εκτιμήτριες συναρτήσεις T_1 και T_2 για μια άγνωστη παράμετρο θ . Παίρνουμε την εκτιμήτρια $T = \lambda T_1 + (1 - \lambda)T_2$, όπου λ μια σταθερά. Είναι η T αμερόληπτη;

Αμερόληπτη εκτιμήτρια συνάρτηση $T \Rightarrow E T = \theta$ χρησιμοποιείται

$$\begin{aligned} E T_1 &= \theta \\ E T_2 &= \theta \\ E T &= E[\lambda T_1 + (1 - \lambda)T_2] = \lambda E T_1 + (1 - \lambda) E T_2 = \\ &= \lambda \theta + (1 - \lambda) \theta \Rightarrow E T = \theta \Rightarrow \\ &\Rightarrow T \text{ αμερόληπτη} \end{aligned}$$

Άσκηση 2

Wednesday, December 13, 2023 12:44 PM

Έστω τυχαίο δείγμα $X_1, X_2, \dots, X_n$ από κατανομή με άγνωστη παράμετρο θ , και έστω η ακόλουθη στατιστική συνάρτηση: $T = \sum_{i=1}^n \lambda_i X_i$, όπου λ_i σταθερές ($i = 1, 2, \dots, n$). Τι πρέπει να ισχύει για τα λ_i ώστε η συνάρτηση T να είναι αμερόληπτη εκτιμήτρια για την παράμετρο θ ;

$$\begin{aligned}
 ET = \theta &\Rightarrow E\left[\sum_{i=1}^n \lambda_i X_i\right] = \theta \quad \xrightarrow{\text{γραμμικότητα}} \sum_{i=1}^n \lambda_i EX_i = \theta \Rightarrow \sum_{i=1}^n \lambda_i \mu = \theta \Rightarrow \\
 &\Rightarrow \mu \sum_{i=1}^n \lambda_i = \theta \Rightarrow \boxed{\sum_{i=1}^n \lambda_i = \frac{\theta}{\mu}}
 \end{aligned}$$

Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα από πληθυσμό που ακολουθεί κατανομή με $\mu = \theta$ και $\sigma^2 = 2\theta$.

Ν.δ.ο. οι εκτιμήτριες $T_1 = \frac{X_1 + X_2}{2}$, $T_2 = \frac{3X_1 + 2X_2}{5}$ είναι αμερόληπτες ως προς την παράμετρο θ και να βρεθεί η πιο αποτελεσματική.

$$ET_1 = E\left[\frac{X_1 + X_2}{2}\right] = \frac{EX_1 + EX_2}{2} = \frac{\theta + \theta}{2} = \theta \Rightarrow \text{αμερόληπτη}$$

$$ET_2 = E\left[\frac{3X_1 + 2X_2}{5}\right] = \frac{3EX_1 + 2EX_2}{5} = \frac{3\theta + 2\theta}{5} = \theta \Rightarrow \text{αμερόληπτη}$$

$$\text{Var}T_1 = \text{Var}\left[\frac{X_1 + X_2}{2}\right] = \frac{1}{2^2} (\text{Var}X_1 + \text{Var}X_2) = \frac{1}{4} (2\theta + 2\theta) = \theta$$

$$\text{Var}T_2 = \text{Var}\left[\frac{3X_1 + 2X_2}{5}\right] = \frac{1}{5^2} (3^2 \text{Var}X_1 + 2^2 \text{Var}X_2) = \frac{1}{25} (9 \cdot 2\theta + 4 \cdot 2\theta) \Rightarrow$$

$$\Rightarrow \text{Var}T_2 = \frac{26}{25} \theta > \theta = \text{Var}T_1 \Rightarrow T_1 \text{ αποτελεσματικότερη}$$

Εξεργ: Ποια T_3 είναι η πιο αποτελεσματική;

$$T_3 = \bar{X} = \frac{X_1 + \dots + X_n}{n}$$

$$\text{Var}T_3 = \frac{1}{n^2} (\text{Var}X_1 + \dots + \text{Var}X_n) = \frac{1}{n^2} \cdot (n \cdot 2\theta) = \frac{2\theta}{n}$$

Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα από πληθυσμό με οποιαδήποτε κατανομή και $p = P(X \leq a)$. Να δειχθεί ότι η στατιστική συνάρτηση $T(X_1, X_2, \dots, X_n) = \frac{1}{n} [\text{αριθμός των } X_k, k = 1, 2, \dots, n \text{ με } X_k \leq a]$ είναι αμερόληπτη εκτιμήτρια της παραμέτρου p ;

Έστω Y_k ε.ρ. δι. του

$$Y_k = \begin{cases} 1, & X_k \leq a \\ 0, & X_k > a \end{cases} \quad \begin{matrix} \nearrow EY_k = p \\ P(Y_k = 1) = p \Rightarrow P(Y_k = 0) = 1 - p \end{matrix}$$

Y ε.ρ. έκδοια ώστε $Y = \sum_{i=1}^n Y_i \rightarrow$ δείχνει πόσα $X_k \leq a$

$$\text{Άρα, } T(X_1, \dots, X_n) = \frac{1}{n} Y \Rightarrow ET(X_1, \dots, X_n) = E\left[\frac{Y}{n}\right] = \frac{EY}{n} =$$

$$= \frac{1}{n} E\left[\sum_{i=1}^n Y_i\right] = \frac{1}{n} \sum_{i=1}^n EY_i = \frac{1}{n} \sum_{i=1}^n p = \frac{1}{n} n \cdot p \Rightarrow ET(X_1, \dots, X_n) = p \Rightarrow$$

$\Rightarrow T(X_1, \dots, X_n)$ είναι αμερόληπτη

Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα από πληθυσμό με συνάρτηση πυκνότητας πιθανότητας $f(x) = \theta x^{\theta-1}$, $0 < x < 1$, $\theta >$

0. Να βρεθεί η εκτιμήτρια της παραμέτρου θ :

(α) Με τη μέθοδο των ροπών,

(β) Με τη μέθοδο μέγιστης πιθανοφάνειας (MLE).

α) Μέθοδος ροπών

1ο βήμα: Έχουμε μία άγνωστη παράμετρο, άρα θα χρησιμοποιήσουμε
εν $1n$ ροπή, δηλαδή τη μέση τιμή:

$$EX = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x \theta x^{\theta-1} dx = \int_0^1 \theta x^{\theta} dx = \theta \int_0^1 x^{\theta} dx = \theta \left[\frac{x^{\theta+1}}{\theta+1} \right]_0^1 =$$

$$= \theta \left(\frac{1^{\theta+1}}{\theta+1} - \frac{0^{\theta+1}}{\theta+1} \right) \Rightarrow \boxed{EX = \frac{\theta}{\theta+1}}$$

2ο βήμα: Εισάγουμε $EX = \bar{X} \Rightarrow \frac{\theta}{\theta+1} = \bar{X} \Rightarrow$

$$\Rightarrow \theta = \theta \bar{X} + \bar{X} \Rightarrow \theta(1 - \bar{X}) = \bar{X} \Rightarrow \boxed{\theta = \frac{\bar{X}}{1 - \bar{X}}}$$

β) Μέθοδος μέγιστης Πιθανοφάνειας

$$L(\theta) = f(x_1, x_2, \dots, x_n; \theta) = f(x_1; \theta) \cdot f(x_2; \theta) \cdot \dots \cdot f(x_n; \theta) \Rightarrow$$

$$\log L(\theta) = \log [f(x_1; \theta) \cdot f(x_2; \theta) \cdot \dots \cdot f(x_n; \theta)] = \sum_{i=1}^n \log f(x_i; \theta) =$$

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

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

$$\frac{d \log L(\theta)}{d \theta} = \frac{d (n \log \theta + (\theta-1) \sum_{i=1}^n \log x_i)}{d \theta} = \frac{n}{\theta} + \sum_{i=1}^n \log x_i$$

$$\left. \frac{d \log L(\theta)}{d \theta} \right|_{\theta = \hat{\theta}} = 0 \Rightarrow \frac{n}{\hat{\theta}} + \sum_{i=1}^n \log x_i = 0 \Rightarrow \boxed{\hat{\theta} = - \frac{n}{\sum_{i=1}^n \log x_i}} \quad (1)$$

$$\left. \frac{d \log L(\theta)}{d\theta} \right|_{\theta=\hat{\theta}} = 0 \Rightarrow \frac{n}{\hat{\theta}} + \sum_{i=1}^n \log x_i = 0 \Rightarrow \boxed{\hat{\theta} = - \frac{n}{\sum_{i=1}^n \log x_i}} \quad (1)$$

Θέλουμε να δείξουμε ότι $\frac{d^2 \log L(\theta)}{d\theta^2} < 0$

$$\frac{d^2 \log L(\theta)}{d\theta^2} = \frac{d \left(\frac{n}{\theta} + \sum_{i=1}^n \log x_i \right)}{d\theta} = -\frac{n}{\theta^2} < 0$$

Άρα η $L(\theta)$ έχει μέγιστο για $\boxed{\hat{\theta} = - \frac{n}{\sum_{i=1}^n \log x_i}}$