Statistiques

Édouard Lumet – IR2020

Exercice

a,b réel, 0 < a < b
x_1 _____ x_n échantillon de n v.a.r. iid
$X_i \sim \mathcal{U}[0, b]$
La fonction $sign(x) = \begin{cases} +1 \; si \; x \geq 0 \\ -1 \; si \; x < 0 \end{cases}$
Soit $\{Y_j\}_{j=1-n}$ $Y_j = sign(X_j - a)$

Question 1. Quelle est la loi des Y_j, calculer $E[Y_j]$ et $Var[Y_j]$
Question 2. Existe-t-il un k tel que â soit un estimateur de a sous biais de la forme

$$\hat{a} = k \sum_{j=1}^{n} Y_j$$

Réponse 1. Réflexe lorsqu'on me demande une loi ? → Continu ou discret
Chaque Y_j est à valeurs dans {-1,1}, Y_j est une var discrète.

$$P[Y_j = 1] = P[sign(X_j - a) \geq 0] = P[X_j \geq a] \; avec \; X_j \sim \mathcal{U}[0, b]$$

$$= \int_0^b \frac{1}{b} dx_j = \frac{1}{b} [x]_a^b = \frac{b-a}{b}$$

$$P[Y_j = -1] = P[sign(X_j - a) < 0] = P[X_j < a] = 1 - P[Y_j = 1] = 1 - \frac{b-a}{b} = \frac{a}{b}$$

$$E[Y_j] = 1P[Y_j = 1] - 1P[Y_j = -1] = 1.\frac{b-a}{b} - 1.\frac{a}{b} = \frac{b-2a}{b}$$

$$Var[Y_j] = E[Y_j^2] - (E[Y_j])^2 = \frac{4a}{b} - \frac{4a^2}{b^2}$$

Réponse 2.

$$E[\hat{a}] = a \; (inconnue \; à \; estimer)$$

$$E[fct(Y_1 \dots Y_n)] = a$$

$$E[\hat{a}] = E\left[k \sum_{j=1}^{n} Y_j\right] = kE\left[\sum_{j=1}^{n} Y_j\right] = k \sum_{j=1}^{n} E[Y_j] = k \sum_{j=1}^{n} \left(\frac{b-2a}{b}\right) = k.n\left(\frac{b-2a}{b}\right)$$

Si â est sous biais alors E[â]=a

$$k.n\left(\frac{b-2a}{b}\right) \Rightarrow k = \frac{ba}{n(b-2a)}$$

$\hat{a} = \frac{ba}{n(b-2a)} \sum_{j=1}^{n} Y_j$ est-il un estimateur de a (inconnu) ? NON !
NB : un estimateur â de a ne peut pas dépendre de a !!

Exercice

$X_1 \dots X_n$ n v.a.r. iid

$X_i \sim \mathcal{N}(m, m^2)$ $avec\ m > 0$

On rappelle que $\left| \begin{array}{c} si\ Z \sim \mathcal{N}(m, \sigma^2) \\ E[(Z-m)^4 = 3\sigma^4 \\ E[Z^4] = 3\sigma^2 + 6m^2\sigma^2 + m^4 \end{array} \right.$

Indication (théorème de Fisher) :

$T_1 \dots T_n$ n v.a.r. iid $T_i \sim \mathcal{N}(m, \sigma^2)$

$\bar{T} = \frac{1}{n}\sum_{i=1}^{n} T_i \quad \bar{T} \sim \mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$

Question 1. Soit $\theta = m^2 \quad X_i \sim \mathcal{N}(\sqrt{\theta}, \theta)$ $avec\ m > 0$

On considère 2 estimateurs de θ

$$\widehat{\theta_1} = fct_1(X_1 \dots X_n) = \frac{1}{2n}\sum_{j=1}^{n} X_j^2$$

$$\widehat{\theta_2} = fct_2(X_1 \dots X_n) = (\bar{X})^2 = \left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)^2$$

$\widehat{\theta_1}$ et $\widehat{\theta_2}$ sont-ils de bons estimateurs de θ ?

$\widehat{\theta_1}$ est-il sans biais ?

$$E\left[\widehat{\theta_1}\right] = \frac{1}{2n}\sum_{j=1}^{n} E\left[X_j^2\right] = \frac{1}{2n}\sum_{j=1}^{n} 2\theta = \frac{1}{2n} n2\theta = \theta$$

Oui, $\widehat{\theta_1}$ est un estimateur sans biais de θ.

ET CETERA