
Révisions - séance 5 (extrait essec II 2009)

Premier problème d'estimation

Dans ce premier problème, on dispose d'une seule observation notée X .
On suppose que X admet pour densité f_θ définie sur $\mathbb{R}$ par :

$$f_\theta(x) = \begin{cases} \frac{k+1}{\theta^{k+1}} x^k & \text{si } x \in [0, \theta] \\ 0 & \text{sinon} \end{cases}$$

où k est un entier naturel non nul et θ un paramètre réel inconnu strictement positif que l'on souhaite estimer.

On dit qu'un estimateur Z de θ est *sans biais* si $E(Z) = \theta$.

1) Montrer que f_θ est bien une densité de probabilité.

2) Calculer $E(X)$.

3) Déterminer λ_0 réel dépendant uniquement de k tel que $\lambda_0 X$ soit un estimateur sans biais de θ .

4) Calculer $V(X)$.

On définit le *risque quadratique* de T , estimateur de θ , par :

$$R(T, \theta) = E((T - \theta)^2).$$

Plus petit est le risque quadratique, meilleur est l'estimateur.

5) Montrer que pour tout estimateur T de θ , on a :

$$R(T, \theta) = (E(T) - \theta)^2 + V(T).$$

6) A l'aide de l'égalité précédente, donner la valeur de $R(\lambda_0 X, \theta)$.

Le but de la fin de cette partie I est de déterminer un estimateur de θ ayant un plus petit risque quadratique que celui de $\lambda_0 X$.

7) En revenant à la définition du risque quadratique, montrer que pour tout λ réel

$$R(\lambda X, \theta) = \theta^2 Q(\lambda)$$

où Q est un polynôme de degré 2 dont les coefficients ne dépendent que de k .

8) Montrer que la fonction $\lambda \mapsto Q(\lambda)$ atteint son minimum en un unique réel noté λ^* que l'on exprimera en fonction de k .

9) Conclure sur le but recherché.

Second problème d'estimation

Dans ce second problème d'estimation, on dispose de n observations indépendantes ($n \geq 2$) notées $X_1, \dots, X_n$ de même loi de Poisson de paramètre $\theta > 0$ inconnu.

On souhaite estimer le paramètre $\exp(-\theta)$.

On définit pour tout i élément de $\{1, \dots, n\}$ la variable aléatoire Y_i par :

$$Y_i = \begin{cases} 1 & \text{si } X_i = 0 \\ 0 & \text{sinon} \end{cases}$$

Puis on note

$$\overline{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$$

1) Pour tout i élément de $\{1, \dots, n\}$, donner la loi de Y_i .

2) Donner la loi de $\sum_{i=1}^n Y_i$, puis montrer que $E(\overline{Y}_n) = \exp(-\theta)$.

$\overline{Y}_n$ est donc un estimateur sans biais de $\exp(-\theta)$.

3) Calculer $V(\overline{Y}_n)$.

Pour tout k élément de $\{1, \dots, n\}$ on définit $S_k = \sum_{i=1}^k X_i$.

4) Rappeler sans démonstration la loi de S_k pour tout k élément de $\{1, \dots, n\}$.

On définit jusqu'à la fin de cette partie II pour tout j entier naturel

$$\varphi(j) = P_{(S_n=j)}(X_1 = 0)$$

5) Montrer que pour tout j entier naturel

$$\varphi(j) = \left(1 - \frac{1}{n}\right)^j$$

On a donc $\varphi(j)$ indépendant du paramètre θ inconnu.

D'après la question II.5, on peut définir l'estimateur

$$\varphi(S_n) = \left(1 - \frac{1}{n}\right)^{S_n}$$

6) Montrer que $\varphi(S_n)$ admet une espérance et que $E(\varphi(S_n)) = \exp(-\theta)$.

$\varphi(S_n)$ est donc un estimateur sans biais de $\exp(-\theta)$.

7) Montrer que $\varphi(S_n)$ admet une variance vérifiant

$$V(\varphi(S_n)) = \exp(-2\theta) \left(\exp\left(\frac{\theta}{n}\right) - 1 \right)$$

8) On souhaite comparer les performances de $\overline{Y_n}$ et $\varphi(S_n)$ en tant qu'estimateurs de $\exp(-\theta)$.

a) En utilisant l'inégalité des accroissements finis, démontrer que

$$1 \leq \frac{\exp(\theta) - 1}{\theta} \leq \exp(\theta)$$

b) Soit la fonction $h : [0, 1] \rightarrow \mathbb{R}$ définie par

$$h(t) = t \exp(\theta) + (1 - t) - \exp(t\theta)$$

pour tout $t \in [0, 1]$. Etudier les variations de h .

c) En déduire que

$$\exp\left(\frac{\theta}{n}\right) \leq \frac{\exp(\theta)}{n} + \frac{n-1}{n}$$

puis l'inégalité

$$V(\varphi(S_n)) \leq V(\overline{Y_n})$$

d) Comparer les risques quadratiques de $\overline{Y_n}$ et $\varphi(S_n)$.