DS8

Devoir surveillé du 04/03/2022

Durée : 4h

La qualité de la rédaction, la clarté et la précision des raisonnements interviendront pour une part importante dans l'appréciation des copies. Les résultats doivent être encadrés.

La calculatrice n'est pas autorisée.

Problème 1.

Soit a un réel strictement positif.
On considère dans toute la suite du problème une suite $(X_n)_{n\geqslant 1}$ de variables aléatoires mutuellement indépendantes et identiquement distribuées, toutes définies sur un même espace probabilisé $(\Omega, \mathscr{A}, P)$, et suivant toutes la loi uniforme sur l'intervalle $[0, a]$. L'objectif de ce problème est d'étudier puis de comparer deux estimateurs de a.

Les parties 1 et 2 de ce problème sont indépendantes.

Partie 1 : Estimateur du maximum de vraisemblance

On note pour tout $n \geqslant 1, V_n = \max(X_1, \ldots, X_n)$, appelé *estimateur de a du maximum de vraisemblance.*

1. (a) On rappelle qu'en `Scilab`, l'instruction `grand(n,m,'unf',a,b)` permet d'obtenir une matrice à `n` lignes et `m` colonnes, où chaque coefficient simule une loi uniforme sur l'intervalle [`a`,`b`].

   Écrire une fonction d'en-tête `function V=simV(n,a)` prenant en entrée un entier naturel non nul `n` et un réel `a` strictement positif, et qui renvoie une réalisation de V_n.

   (b) On a tracé ci-dessous cinq réalisations mutuellement indépendantes de $(V_1, V_2, \ldots, V_{100})$, dans le cas où $a = 1$.

   À partir de ce graphique, que peut-on conjecturer sur l'estimateur V_n ?

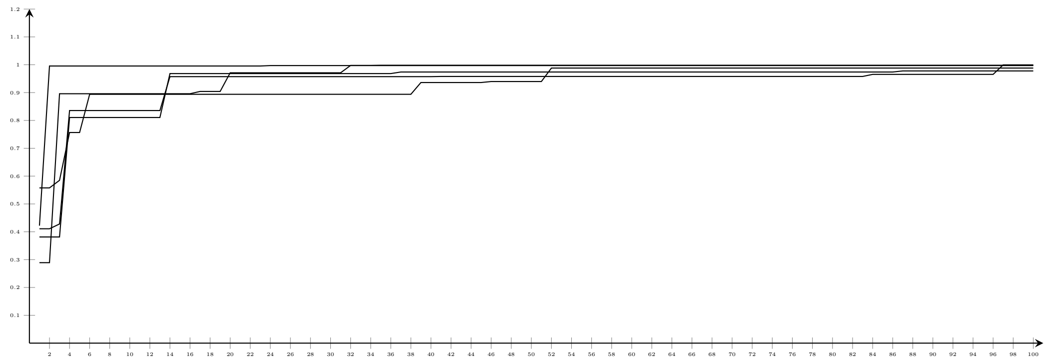

FIGURE N°2 - Cinq évolutions de $(V_1, V_2, \ldots, V_{100})$ pour $a = 1$

2. Soit $n \in \mathbb{N}^*$.

   (a) Rappeler l'expression de la fonction de répartition de X_1, suivant la loi uniforme $\mathscr{U}([0, a])$.

   (b) Déterminer la fonction de répartition F_n de V_n.

   (c) En déduire que V_n est une variable aléatoire à densité et donner une densité de V_n.

3. Soit $n \in \mathbb{N}^*$. Justifier que V_n admet une espérance et déterminer l'espérance de V_n.

   L'estimateur V_n est-il sans biais ?

4. Soit $\varepsilon > 0$ et soit $n \in \mathbb{N}^*$. Exprimer $P(|V_n - a| \geqslant \varepsilon)$ en fonction de F_n, de a et de ε.
   L'estimateur V_n est-il convergent ?
5. Soit $n \in \mathbb{N}^*$.
   (a) Montrer que V_n admet un moment d'ordre 2, que l'on déterminera.
   (b) Montrer que le risque quadratique de V_n vaut $\frac{2a^2}{(n+1)(n+2)}$.
   Quel résultat précédemment établi cela permet-il de retrouver ?

Partie 2 : Méthode des moments

Pour un entier $n \geqslant 1$, on note $\overline{X}_n$ la moyenne empirique de l'échantillon $(X_1, \ldots, X_n)$, c'est-à-dire :

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

On note $M_n = 2\overline{X}_n$, appelé estimateur de a par la méthode des moments.

6. Écrire une fonction d'en-tête `function y=simM(n,a)` qui, prenant en entrée un entier naturel non nul `n` et le réel `a`> 0, renvoie une réalisation de la variable aléatoire M_n.
7. Déterminer l'espérance et la variance de $\overline{X}_n$. En déduire que M_n est un estimateur sans biais.
8. Déterminer le risque quadratique de M_n. Cet estimateur est-il convergent?
9. Comparer le risque quadratique de M_n a celui de V_n, obtenu a la question 7.(b).
   Commenter ce résultat à l'aide de la figure 3 ci-dessous :

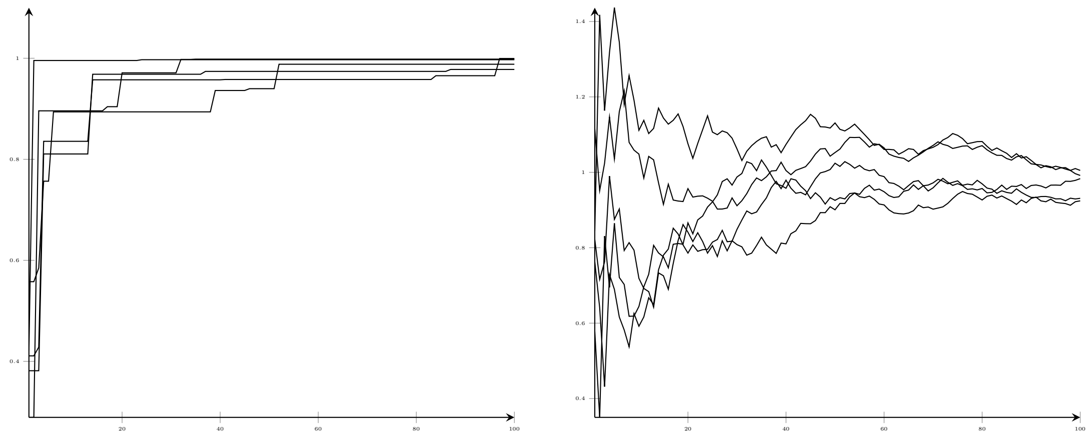

Figure 3 - Cinq évolutions de $(V_1, V_2, \ldots, V_{100})$ (à gauche) et de $(M_1, M_2, \ldots, M_{100})$ (à droite) pour $a = 1$

Partie 3 : Consistance de ces estimateurs

Dans les parties précédentes, nous avons montré que (V_n) convergeait « plus vite » vers a que (M_n). Nous allons maintenant étudier la sensibilité de ces estimateurs à une perturbation, en supposant que la première mesure (X_1) est erronée.
Nous supposons donc toujours que les variables aléatoires X_i sont mutuellement indépendantes, mais nous supposons maintenant que :

- X_1 suit la loi uniforme sur $[0, 2a]$;
- si $i \geqslant 2$, X_i suit la loi uniforme sur $[0, a]$ (comme précédemment).

On considère toujours, pour tout entier $n \geqslant 1$: $V_n = \max(X_1, \ldots, X_n)$ et $M_n = 2\overline{X}_n = \frac{2}{n}(X_1 + \cdots + X_n)$.

10. (a) Soit $n \in \mathbb{N}^*$. Pour tout réel t de $]a, 2a]$, montrer que : $P(V_n \leqslant t) = \frac{t}{2a}$.
    (b) Pour $n \in \mathbb{N}^*$, déterminer la fonction de répartition de V_n.
    La suite de variables aléatoires $(V_n)_{n \geqslant 1}$ converge-t-elle en loi ?
    (c) Calculer $P\left(V_n > \frac{3}{2}a\right)$.
    L'estimateur V_n est-il toujours convergent ?
11. On pose pour tout entier naturel n supérieur ou égal à 2 : $M'_n = \frac{2}{n-1}(X_2 + \cdots + X_n)$.
    On rappelle que la suite $(M'_n)_{n \geqslant 2}$ converge en probabilité vers a.
    (a) Pour tout entier naturel n supérieur ou égal à 2, exprimer M_n en fonction de X_1, M'_n et n.
    (b) En déduire que pour tout entier naturel n supérieur ou égal à 2 :
    $$|M_n - a| \leqslant \frac{3a}{n} + |M'_n - a|.$$
    (c) Soit $\varepsilon > 0$ et soit n_0 un entier naturel supérieur ou égal à 2 tel que $\frac{3a}{n_0} < \varepsilon$.
    Pour tout entier n vérifiant $n \geqslant n_0$, comparer les événements $[\,|M'_n - a| < \varepsilon]$ et $[\,|M_n - a| < 2\varepsilon]$.
    (d) La suite de variables aléatoires $(M_n)_{n \geqslant 2}$ converge-t-elle en probabilité vers a ?
12. Commenter les résultats de cette partie à partir des parties précédentes.

Problème 2.

Dans tout le problème, n est un entier naturel supérieur ou égal à 2, et E est un espace euclidien de dimension n dont le produit scalaire est noté $\langle \cdot, \cdot \rangle$ et la norme associée est notée $\| \cdot \|$. On note id_E l'application identique de E, et $\tilde{0}$ l'application nulle de E.

Si F est un sous-espace vectoriel de E, on note $F^{\perp}$ le sous-espace vectoriel supplémentaire orthogonal de F dans E. Le projecteur de E sur F parallèlement à $F^{\perp}$ est appelé projecteur orthogonal sur F.

Pour tout endomorphisme f de E et toute valeur propre λ de f, on note $E_{\lambda}(f)$ le sous-espace propre de f associé à la valeur propre λ.

Partie I : Inverse généralisé d'un endomorphisme symétrique

On considère un endomorphisme symétrique f de E, c'est-à-dire un endomorphisme f tel que :

$$\forall (x, y) \in E^2, \quad \langle f(x), y \rangle = \langle x, f(y) \rangle.$$

On suppose de plus que f est non inversible et non nul.

1. Montrer que 0 est valeur propre de f et que f admet au moins une valeur propre non nulle.
2. (a) Soient λ et μ deux valeurs propres de f. Montrer, pour tout vecteur x de $E_{\lambda}(f)$ et pour tout vecteur y de $E_{\mu}(f)$:
   $$\lambda \langle x, y \rangle = \mu \langle x, y \rangle.$$
   (b) En déduire que les sous-espaces propres de f sont deux à deux orthogonaux.
3. Montrer que les sous-espaces vectoriels $\mathrm{Im}(f)$ et $\mathrm{Ker}(f)$ sont supplémentaires orthogonaux dans E.

On suppose que f admet exactement $k + 1$ valeurs propres deux à deux distinctes $\lambda_0, \lambda_1, \ldots, \lambda_k$ avec $k \geqslant 1$, $\lambda_0 = 0$ et $0 < |\lambda_1| \leqslant \cdots \leqslant |\lambda_k|$.

Pour tout entier naturel j inférieur ou égal à k, on note p_j le projecteur orthogonal sur $E_{\lambda_j}(f)$.

4. Soit x un vecteur de E.
   (a) Montrer qu'il existe un unique $(k+1)$-uplet $(x_0, x_1 \dots, x_k)$ de $E_0(f) \times E_{\lambda_1}(f) \times \dots \times E_{\lambda_k}(f)$ tel que $x = x_0 + x_1 + \dots + x_k$.
   (b) Pour tout entier naturel j inférieur ou égal à k, montrer : $p_j(x) = x_j$.

   Ainsi, la relation suivante est clairement vérifiée :
$$\mathrm{id}_E = p_0 + p_1 + \dots + p_k$$
5. (a) Établir, pour tout couple (i, j) d'entiers naturels inférieurs ou égaux à k :
$$i \neq j \quad \Longrightarrow \quad p_i \circ p_j = \tilde{0}$$
   (b) Montrer : $\quad f = \lambda_1 p_1 + \lambda_2 p_2 + \dots + \lambda_k p_k$.
   (c) Montrer que le projecteur orthogonal p sur $\mathrm{Im}(f)$ vérifie :
$$p = p_1 + p_2 + \dots + p_k.$$

On note $f^\sharp$ l'endomorphisme de E défini par $f^\sharp = \dfrac{1}{\lambda_1} p_1 + \dfrac{1}{\lambda_2} p_2 + \dots + \dfrac{1}{\lambda_k} p_k$.
On dit que $f^\sharp$ est l'inverse généralisé de f.

6. (a) Montrer : $\quad f \circ f^\sharp = p$.
   (b) En déduire : $\quad \forall (x, y) \in E^2,\ \left(f(x) = p(y) \iff x - f^\sharp(y) \in \mathrm{Ker}(f)\right)$.
7. Soit y un vecteur de E.
   (a) Montrer : $\quad \forall x \in E,\ \left(\|f(x) - y\| = \inf\limits_{z \in E} \|f(z) - y\| \iff x - f^\sharp(y) \in \mathrm{Ker}(f)\right)$.
   (b) En déduire que $f^\sharp(y)$ est le vecteur x de E de plus petite norme vérifiant :
$$\|f(x) - y\| = \inf_{z \in E} \|f(z) - y\|.$$

Partie II : Application à un exemple

Dans cette question, E est un espace euclidien de dimension 4 et $\mathscr{B} = (e_1, e_2, e_3, e_4)$ est une base orthonormale de E. On note :
$$A = \begin{pmatrix} 3 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 3 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}.$$
Soit f l'endomorphisme de E associé à la matrice A relativement à la base $\mathscr{B}$.

8. Justifier que f est un endomorphisme symétrique non nul et non inversible.
9. Montrer que f admet exactement trois valeurs propres distinctes $\lambda_0, \lambda_1, \lambda_2$ avec $\lambda_0 < \lambda_1 < \lambda_2$.

On note p_1 le projecteur orthogonal sur $E_{\lambda_1}(f)$ et M_1 la matrice associée à p_1 relativement à la base $\mathscr{B}$.
On note p_2 le projecteur orthogonal sur $E_{\lambda_2}(f)$ et M_2 la matrice associée à p_2 relativement à la base $\mathscr{B}$.

10. Montrer : $\quad A = 2M_1 + 4M_2$.
11. (a) Montrer que $E_{\lambda_2}(f)$ est de dimension 1 et déterminer un vecteur v_2 de $E_{\lambda_2}(f)$ tel que $\|v_2\| = 1$.
    (b) Montrer : $\quad \forall x \in E, \quad p_2(x) = \langle x, v_2 \rangle v_2$.
    (c) Déterminer la matrice M_2.
12. En déduire la matrice associée à $f^\sharp$ relativement à la base $\mathcal{B}$.