

Examen du 31 janvier 2006 : durée 3 heures
(Notes de cours autorisées)

Problème

Soit X une variable aléatoire réelle de carré intégrable, de loi \mathbb{P}_X sur (Ω, \mathbb{P}) . On suppose que \mathbb{P}_X est absolument continue par rapport à la mesure de Lebesgue dans \mathbb{R} et on note F la fonction de distribution de X . Pour $N \geq 1$, on note $\mathbb{R}_{<}^N$ l'ensemble des N -uplets $x = (x_1, \dots, x_N) \in \mathbb{R}^N$ avec $x_1 < \dots < x_N$. On note D_N^X la fonction de distorsion quadratique de X , i.e.

$$D_N^X(x) = \mathbb{E} \left[\min_{i=1, \dots, N} |X - x_i|^2 \right], \quad x \in \mathbb{R}_{<}^N,$$

et on rappelle que D_N^X est différentiable sur $\mathbb{R}_{<}^N$. On note $\underline{D}_N^X = \min_{x \in \mathbb{R}_{<}^N} D_N^X(x)$ la distorsion minimale.

Pour tout $x = (x_1, \dots, x_N) \in \mathbb{R}_{<}^N$, on pose $\tilde{x}_i = \frac{x_i + x_{i+1}}{2}$, $1 \leq i \leq N-1$.

Partie I.

1) Expliciter les cellules fermées de Voronoi de $x \in \mathbb{R}_{<}^N$:

$$\bar{C}_i(x) = \left\{ u \in \mathbb{R} : |u - x_i| = \min_{j=1, \dots, N} |u - x_j| \right\}, \quad i = 1, \dots, N.$$

2) Expliciter sous forme intégrale (par rapport à \mathbb{P}_X), la distorsion $D_N^X(x)$ et son gradient $\nabla D_N^X(x)$ en fonction de \tilde{x}_i , $1 \leq i \leq N-1$.

3) a) Soit $x = (x_1, \dots, x_N) \in \mathbb{R}_{<}^N$ un quantifieur stationnaire de X , i.e. $\nabla D_N^X(x) = 0$. Montrer que x vérifie les relations

$$\begin{aligned} x_1 F(\tilde{x}_1) &= \int_{-\infty}^{\tilde{x}_1} u \mathbb{P}_X(du) \\ x_i [F(\tilde{x}_i) - F(\tilde{x}_{i-1})] &= \int_{\tilde{x}_{i-1}}^{\tilde{x}_i} u \mathbb{P}_X(du), \quad 2 \leq i \leq N-1 \\ x_N [1 - F(\tilde{x}_{N-1})] &= \int_{\tilde{x}_{N-1}}^{\infty} u \mathbb{P}_X(du). \end{aligned}$$

b) En déduire dans le cas où $N = 2$ et \mathbb{P}_X est une loi symétrique sur \mathbb{R} , i.e. $F(-v) + F(v) = 1$ pour tout $v \in \mathbb{R}$, que $x = (-\mathbb{E}|X|, \mathbb{E}|X|)$ est un quantifieur stationnaire.

Partie II.

On considère dans cette partie le cas particulier où X suit une loi exponentielle $\mathbb{P}_X \sim \mathcal{E}(1)$ de paramètre 1, i.e. $\mathbb{P}_X(du) = 1_{\mathbb{R}_+}(u)e^{-u}du$. On note $\mathbb{R}_{+, <}^N = \mathbb{R}_{<}^N \cap (\mathbb{R}_+)^N$ i.e. l'ensemble des N -uplets de $\mathbb{R}_{<}^N$ de composantes dans le support \mathbb{R}_+ de la loi de $\mathcal{E}(1)$.

4) Soit $x = (x_1, \dots, x_N) \in \mathbb{R}_{+, <}^N$ un quantifieur stationnaire de $X \sim \mathcal{E}(1)$. Montrer que

$$D_N^X(x) = x_1^2.$$

Indication. On utilisera l'expression de la distorsion D_N^X de la question 2) sous forme intégrale qu'on intégrera par parties puis qu'on simplifiera avec les relations de stationnarité de la question 3).

On considère la fonction définie sur \mathbb{R} par

$$\Phi(v) = \int_0^{v/2} ue^{-u} du, \quad v \in \mathbb{R}.$$

5) Etudier la fonction Φ et montrer que

$$\Phi(v) - \Phi(-v) \quad \begin{cases} < 0 & \text{if } v > 0 \\ > 0 & \text{if } v < 0 \end{cases}$$

6) a) Montrer qu'on peut définir de manière unique la suite $(v_k)_{k \geq 1}$ dans \mathbb{R}_+ par les relations récursives implicites, en partant de $v_0 = +\infty$:

$$\Phi(-v_{k+1}) = \Phi(v_k), \quad k \geq 0.$$

b) Montrer que $(v_k)_{k \geq 1}$ est décroissante et converge vers 0.

7) a) Soit $x = (x_1, \dots, x_N) \in \mathbb{R}_{+, <}^N$ un quantifieur stationnaire de $X \sim \mathcal{E}(1)$. On pose $\hat{v}_0 = +\infty$, $\hat{v}_i = x_{N-i+1} - x_{N-i}$, $1 \leq i \leq N-1$, $\hat{v}_N = 2x_1$. Montrer que

$$\Phi(-\hat{v}_{i+1}) = \Phi(\hat{v}_i), \quad 1 \leq i \leq N.$$

Indication. On utilisera les relations de stationnarité de la question 3) et on fera les changements de variable adéquats.

b) En déduire que l'unique quantifieur stationnaire de $X \sim \mathcal{E}(1)$ dans $\mathbb{R}_{+, <}^N$ est donné par :

$$x_i = \frac{v_N}{2} + \sum_{k=N+1-i}^{N-1} v_k, \quad 1 \leq i \leq N.$$

c) Que vaut la distorsion minimale \underline{D}_N^X ?

Partie III.

On suppose que le support de la loi de X est compact, égal à $[a, b]$. On note f la densité de X par rapport à la mesure de Lebesgue sur \mathbb{R} , supposée continue sur $[a, b]$.

8) Dans ce contexte, énoncer le théorème de Zador concernant la vitesse de convergence de \underline{D}_N^X quand N tend vers l'infini.

9) Soit $x = (x_1, \dots, x_N) \in \mathbb{R}_{<}^N \cap [a, b]^N$. En posant $x_0 = a$, $x_{N+1} = b$, vérifier que

$$\begin{aligned} D_N^X(x) &= \int_{x_0}^{x_1} (u - x_1)^2 f(u) du + \sum_{i=2}^N \int_{\tilde{x}_{i-1}}^{x_i} (u - x_i)^2 f(u) du \\ &+ \sum_{i=1}^{N-1} \int_{x_i}^{\tilde{x}_i} (u - x_i)^2 f(u) du + \int_{x_N}^{x_{N+1}} (u - x_N)^2 f(u) du. \end{aligned} \quad (1)$$

10) Montrer qu'il existe $\theta_i \in [x_i, x_{i+1}]$, $0 \leq i \leq N$ tel que

$$D_N^X(x) = f(\theta_0) \frac{(x_1 - x_0)^3}{3} + \sum_{i=1}^{N-1} f(\theta_i) \frac{(x_{i+1} - x_i)^3}{12} + f(\theta_N) \frac{(x_{N+1} - x_N)^3}{3}.$$

Indication : on utilisera le théorème des valeurs intermédiaires à chacun des termes intégraux de (1).

11) On définit la densité g sur $[a, b]$ par

$$g(u) = \frac{f^{\frac{1}{3}}(u)}{\int_a^b |f(u)|^{\frac{1}{3}} du}.$$

On choisit maintenant x_i comme le $(2i - 1)/2N$ quantile de la loi associée, i.e.

$$\int_a^{x_i} g(u) du = \frac{2i - 1}{2N}, \quad 1 \leq i \leq N. \quad (2)$$

a) Montrer qu'il existe $\eta_i \in [x_i, x_{i+1}]$, $0 \leq i \leq N$, tel que

$$D_N^X(x) = \frac{1}{12N^2} \left[\sum_{i=0}^N \frac{f(\theta_i)}{g(\eta_i)^2} (x_{i+1} - x_i) \right].$$

Indication : on utilisera le théorème des valeurs intermédiaires à g dans (2) pour estimer $x_{i+1} - x_i$, $i = 0, \dots, N$.

b) En déduire que

$$\lim_{N \rightarrow \infty} N^2 D_N^X(x) = \frac{1}{12} \left(\int_a^b |f(u)|^{\frac{1}{3}} du \right)^3.$$

c) Interpréter.