DEA de Mathématiques Appliquées
Examen de Grandes Déviations
2001-2002

Epreuve du 7 Février 2002

1 Estimateur à noyau

Soit $X_1, \ldots, X_n$ des variables aléatoires i.i.d. à valeurs dans $\mathbb{R}$ de loi F. On suppose que F est absolument continu par rapport à la mesure de Lebesgue et l'on appelle f sa densité. On suppose que la fonction f est strictement positive et est continue et bornée sur $\mathbb{R}$. Soit K une fonction positive et paire sur $\mathbb{R}$. On suppose que K est bornée et que son intégrale sur $\mathbb{R}$ vaut 1. Pour $x \in \mathbb{R}$, on considère l'estimateur à noyau $\widehat{f}_n(x)$ de $f(x)$:

$$\widehat{f}_n(x) = \frac{1}{nh_n} \sum_{j=1}^{n} K\left(\frac{X_j - x}{h_n}\right),$$

où (h_n) est une suite de nombres strictement positifs qui vérifie :

$$\lim_{n\to\infty} h_n = 0 \text{ et } \lim_{n\to\infty} nh_n = +\infty. \tag{1}$$

a) Montrer que si l'on s'intéeresse aux propriétés asymptotiques de $\widehat{f}_n(x)$ ($x \in \mathbb{R}$ est fixé), on peut supposer sans perte de généralité que $x = 0$.

b) Montrer que

$$\lim_{n\to\infty} \int_{\mathbb{R}} K(z) f(h_n z) dz = f(0).$$

En déduire que l'espérance de $\widehat{f}_n(x)$ converge vers une limite à préciser.

c) Montrer que $\widehat{f}_n(0)$ converge dans L^2 vers $f(0)$. En déduire, pour $x \in \mathbb{R}$, la limite en probabilité de la suite $\widehat{f}_n(x)$.

d) On se propose de montrer que la suite $(\widehat{f}_n(x))$ satisfait un principe de grandes déviations pour la vitesse nh_n. Pour $\tau \in \mathbb{R}$, on pose

$$\psi(\tau) = \int_{\mathbb{R}} (e^{\tau K(z)} - 1) dz. \tag{2}$$

i) Montrer que ψ est définie continue et dérivable sur $\mathbb{R}$. Montrer que ψ est convexe.

ii) Montrer que

$$\lim_{n\to\infty} \int_{\mathbb{R}} (e^{\tau K(z)} - 1) f(h_n z) dz = \psi(\tau) f(0) \quad (\tau \in \mathbb{R}).$$

En utilisant l'équivalent en 0 de $\log(1+h)$, en déduire que, pour $x \in \mathbb{R}$ fixé, $(\widehat{f}_n(x))$ satisfait un P.G.D. (pour la vitesse nh_n). Préciser la bonne fonction de taux I_x. Que vaut $I_x(y)$ pour $y < 0$. Montrer que l'on peut écrire, pour $y \in \mathbb{R}$, $I_x(y)$ sous la forme $I_x(y) = f(x) I(y/f(x))$ où I est une fonction convexe qui ne dépend plus de x que l'on précisera. Calculer cette fonction dans le cas où $K(t) = K_0(t) = \mathbb{1}_{[-2^{-1}, 2^{-1}]}(t)$, $(t \in [0,1])$.

iii) Montrer que $I_x(y) > 0$ si $y \neq f(x)$. Pour $\varepsilon > 0$ et $x \in \mathbb{R}$, A_n^ϵ est l'événement défini par

$$A_n^\varepsilon = \{|\widehat{f}_n(x) - f(x)| > \varepsilon\}.$$

Montrer que la série $\sum_{n\geq 1} \mathbb{P}(A_n^\varepsilon)$ est convergente. En déduire que la suite $(\widehat{f}_n(x))$ converge presque sûrement vers $f(x)$.

e) Dans cette question on se place dans le cas où $K = K_0$. Soit p un entier strictement positif et $x_1 < \ldots < x_p$ des réels donnés. Montrer que la suite des vecteurs aléatoires $(\widehat{f}_n(x_1), \ldots, \widehat{f}_n(x_p))^T$ satisfait un P.G.D., pour la vitesse nh_n et la bonne fonction de taux :

$$J_p(y) = I_{x_1}(y_1) + \cdots + I_{x_p}(y_p), \quad y = (y_1, \ldots, y_p)^T \in \mathbb{R}^p.$$

2 Autour du théorème de Schilder

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité. Sur cet espace on considère un mouvement brownien $w = (w_t)_{t\in[0,1]}$ et une suite i.i.d. $X = (X_n)$ à valeurs dans $[0,1]$. On suppose que X est indépendante de w et que la variance de X_1 n'est pas nulle. Pour $t \in [0,1]$, on pose $Y_t = w_{t\overline{X}_n}$ où comme toujours $\overline{X}_n$ désigne la moyenne empirique de l'échantillon $X_1 \ldots X_n$ ($n \in \mathbb{N}^*$). Le but de ce problème est de montrer que le processus $Y_n = (\sqrt{n^{-1}}Y_t + t\overline{X}_n)_{t\in[0,1]}$ satisfait un P.G.D. sur $C_0([0,1])$ muni de la norme uniforme.

a) Montrer ou admettre que la suite $(\overline{X}_n, \sqrt{n^{-1}}w)$, à valeurs dans $[0,1]\times C_0([0,1])$, satisfait un P.G.D. (à vitesse n) avec une bonne fonction de taux I à déterminer. Indications : pour la majoration, montrer d'abord que $(\overline{X}_n, \sqrt{n^{-1}}w)$ est exponentiellement tendue puis raisonner sur un compact. Pour la minoration raisonner sur un produit d'ouverts.

b) On considère l'application T de $[0,1]\times C_0([0,1])$ dans $C_0([0,1])$ qui à un couple $(x, f) \in [0,1] \times C_0([0,1])$ associe la fonction g définie pour $t \in [0,1]$ par $g(t) = f(xt) + tx$. Montrer que l'application T est continue. En déduire que Y_n satisfait un P.G.D. (à vitesse n) avec une bonne fonction de taux J. Exprimer J à partir de I.

3 Minimum

Soit $X_1, \ldots, X_n$ des variables i.i.d. à valeurs dans $\mathbb{R}^+$. On suppose que la loi de X_1 possède une densité f avec $f(x) > 0$ sur $[0, \varepsilon]$ pour un $\epsilon > 0$. On pose $Y_n = \min_{i=1,\ldots,n} X_i$.

a) Montrer que Y_n possède une densité g_n à déterminer. Montrer que Y_n converge en loi et en probabilité vers 0.

b) i) Que vaut, pour $t > 0$, $n^{-1} \log \mathbb{P}(Y_n > t)$?

ii) Pour $t > \varepsilon > 0$ on pose $I_\varepsilon =]t - \varepsilon, t + \varepsilon[$. Calculer

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(Y_n \in I_\varepsilon).$$

Soit O un ouvert de $\mathbb{R}^+$. Minorer la limite inférieure de $n^{-1} \log \mathbb{P}(Y_n \in O)$

iii) Soit F un fermé de $\mathbb{R}^+$. Utiliser i) pour majorer la limite supérieure de $n^{-1} \log \mathbb{P}(Y_n \in F)$.

c) En utilisant la question b), montrer que la suite (Y_n) satisfait un P.G.D. Quelle est la fonction de taux J. Cette fonction est-elle bonne?

d) Calculer pour $\alpha > 0$

$$\lim_{n\to\infty} \frac{1}{n} \log \int_{[0,1]^n} \exp(n\alpha(\min_{i=1,\dots,n} x_i)^2) dx_1 \cdots dx_n.$$

4 Processus M.A.

Soit (ε_n) une suite de variables aléatoires centrées de variance 1 et i.i.d. On suppose que la log-Laplace de ε_1 est finie sur $\mathbb{R}$. Soit $a_1, \dots, a_p$ des nombres réels. Pour $n \in \mathbb{N}^*$, on pose

$$X_n = \sum_{i=1}^{p} a_i \varepsilon_{n+i}.$$

Comme toujours on appelle $\overline{X}_n$ la moyenne empirique construite à partir de $X_1, \dots, X_n$.

a) Montrer que le processus (X_n) est stationnaire. Déterminer sa densité spectrale.

b) On suppose que ε_1 suit la loi normale centrée réduite. Déterminer la loi de $\overline{X}_n$. Montrer que n Var$\overline{X}_n$ converge, quand n tend vers l'infini, vers $\sum_{j=1}^{p} a_j^2$. En déduire que $\overline{X}_n$ satisfait un P.G.D. (la vitesse ici est n). Donner un equivalent pour n grand et $t > 0$ de $\mathbb{P}(\overline{X}_n > t)$.

c) On ne suppose plus que ε_1 est gaussien. Montrer que $\overline{X}_n$ satisfait un P.G.D. avec une bonne fonction de taux à déterminer.