

FACULTÉ DES SCIENCES ET DE GÉNIE

**THÈSE
PRÉSENTÉE
A L'ÉCOLE DES GRADUÉS
DE L'UNIVERSITÉ LAVAL
POUR L'OBTENTION
DU GRADE DE PHILOSOPHIAE DOCTOR (Ph. D.)**

PAR

BELKACEM ABDOUS

**ÉTUDE D'UNE CLASSE D'ESTIMATEURS À NOYAU
DE LA DENSITÉ D'UNE LOI DE PROBABILITÉ**

AVRIL 1989

© droits réservés de Belkacem Abdous 1989

A mes parents, frères et sœurs.

Résumé

Dans ce travail nous donnons un aperçu des plus intéressantes approches visant à déterminer la fenêtre optimale en estimation de la densité d'une loi de probabilité par la méthode du noyau. Nous construisons ensuite une classe d'estimateurs à noyau de la densité pour lesquels nous avons établi des conditions suffisantes de convergence uniforme presque sûre et L^1 presque sûre vers la densité à estimer f . Cette classe d'estimateurs à noyau étant assez générale, elle nous a permis d'appliquer ces résultats de convergence à des estimateurs à noyau classiques comme ceux de Deheuvels (1977-a), Shanmugam (1977), Bierens (1983), et Devroye et Wagner (1983). Elle nous a permis également, de construire une famille d'estimateurs à noyau de moyenne μ_n et de matrice de variance-covariance V_n , où μ_n est un estimateur non spécifié de la moyenne de f et V_n , à une constante multiplicative près, la matrice de variance-covariance empirique. Enfin, en simulant quelques modèles univariés connus, nous avons comparé les performances de l'estimateur à noyau de Parzen-Rosenblatt avec celles de l'estimateur à noyau de variance la variance empirique S_n^2 et de moyenne μ_n , où μ_n a été choisi comme étant la moyenne empirique \bar{X}_n ou bien la médiane \tilde{X}_n ou bien la moyenne empirique α -tronquée ($\alpha = 0.1$) ou bien l'estimateur de Gastwirth (1966).

Résumé

Dans ce travail nous donnons un aperçu des plus intéressantes approches visant à déterminer la fenêtre optimale en estimation de la densité d'une loi de probabilité par la méthode du noyau. Nous construisons ensuite une classe d'estimateurs à noyau de la densité pour lesquels nous avons établi des conditions suffisantes de convergence uniforme presque sûre et L^1 presque sûre vers la densité à estimer f . Ces résultats sont alors appliqués à des estimateurs à noyau classiques et à une famille d'estimateurs à noyau de moyenne μ_n et de matrice de variance-covariance V_n , où μ_n est un estimateur non spécifié de la moyenne de f et V_n , à une constante multiplicative près, la matrice de variance-covariance empirique. Enfin, en simulant quelques modèles univariés connus, nous avons comparé les performances de ces estimateurs à noyau.

Avant-propos

Je tiens à remercier mon directeur de recherche Monsieur Radu Theodorescu pour ses conseils, encouragements et critiques constructives qu'il m'a prodigués tout au long de ce travail. Mes remerciements vont également à Messieurs Denis Bosq, Jean-Pierre Carmichael et Louis-Paul Rivest qui ont bien voulu accepter de juger ce travail.

Je remercie l'Université Laval, son département de mathématiques et de statistique, ainsi que l'équipe "Probabilités et Statistique", pour leur sympathie et soutien financier.

Mon amie Martine Pelletier m'a été d'un grand support moral et a su faire preuve d'une grande compréhension, je l'en remercie vivement.

TABLE DES MATIÈRES

	page
Résumé	i
Avant-propos	ii
Table des matières	iii
Introduction	1
Chapitre I: CONVERGENCE UNIFORME PRESQUE SÛRE	11
I.1 Conditions suffisantes pour la convergence uniforme presque sûre de f_n	11
I.2 Applications	28
Chapitre II: CONVERGENCE L^1 PRESQUE SÛRE	32
II.1 Conditions suffisantes pour la convergence L^1 presque sûre de f_n	32
II.2 Applications	52
Chapitre III: SIMULATIONS ET RÉSULTATS NUMÉRIQUES	54
Références	67

INTRODUCTION

Durant les trente dernières années l'estimation fonctionnelle non paramétrique a connu un développement considérable. En particulier, l'estimation de la densité d'une loi de probabilité dans \mathbb{R}^k , $k \geq 1$, a fait l'objet de multiples travaux par des méthodes très différentes. De nombreux ouvrages y ont été consacrés, notamment Tapia et Thompson (1978), Prakasa Rao (1983), Devroye et Györfi (1985), Bosq et Lecoutre (1986), et Silverman (1986).

Les premiers travaux dans ce domaine remontent à ceux de Karl Pearson en 1902. Ils portaient sur les densités solutions de l'équation différentielle

$$\frac{df}{dx} = \frac{(x-a)f}{b_0 + b_1x + b_2x^2},$$

où a , b_0 , b_1 , et b_2 sont des constantes qui s'expriment en fonction des quatre premiers moments de la densité f . Il s'agissait alors d'estimer les moments de cette densité. Par la suite, de nombreuses autres méthodes ont été proposées. Celle qui a retenu le plus d'attention, de par sa facilité, est la méthode du noyau. C'est Rosenblatt (1956), reprenant l'idée de Fix et Hodges (1951), qui a introduit cette méthode. L'idée consistait à estimer la densité en un point x en comptant le nombre d'observations situées dans l'intervalle $[x-h, x+h]$. Rosenblatt a, en fait, proposé une forme plus générale donnée par

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right), \quad x \in \mathbb{R},$$

où K est une densité de probabilité sur \mathbb{R} , (h_n) est une suite de réels positifs et X_1, \dots, X_n est un échantillon issu de f . L'extension de f_n au cas multivarié a été faite par Cacoullos (1966):

$$f_n(x) = \frac{1}{nh_n^k} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right), \quad x \in \mathbb{R}^k, \quad k \geq 1,$$

où K est une densité de probabilité sur \mathbb{R}^k . Dans la littérature f_n est communément appelé estimateur de Parzen-Rosenblatt. Dans ce qui suit nous le noterons f_n^{PR} . Vint ensuite Deheuvels (1977-a), qui au lieu de prendre une matrice diagonale en a pris une pleine, et a posé

$$f_n^D(x) = \frac{1}{n |\det H_n|} \sum_{i=1}^n K(H_n^{-1}(x - X_i)), \quad x \in \mathbb{R}^k,$$

où H_n est une matrice $k \times k$ inversible.

Plusieurs modes de convergence de f_n^{PR} vers f ont été explorés. Parzen (1962) pour $k = 1$ et Cacoullos (1966) pour $k > 1$ ont établi la convergence en probabilité de

$$\sup_{x \in \mathbb{R}^k} |f_n^{PR}(x) - f(x)|,$$

et la convergence de l'erreur quadratique (MSE)

$$MSE = E(f_n^{PR}(x) - f(x))^2, \quad x \in \mathbb{R}^k.$$

Tiago de Oliveira (1963) a obtenu, pour $k = 1$, la convergence de l'erreur quadratique intégrée moyenne

$$MISE = \int_{\mathbb{R}^k} E(f_n^{PR}(x) - f(x))^2 dx.$$

Quant à Nadaraja (1965), il a démontré la convergence uniforme presque sûre de f_n^{PR} vers f . Résultat qui a été amélioré par Geffroy (1973), Deheuvels (1974) et Bertrand-Retali (1978). La convergence L^1 de f_n^{PR} a été traitée en premier, par Abou-Jaoudé (1977), une étude détaillée de ce mode de convergence est fournie par Devroye et

Györfi (1985). Les critères permettant de mesurer l'efficacité de f_n^{PR} et ayant suscité le plus d'intérêt sont l'erreur quadratique asymptotique $AMSE$ et l'erreur quadratique intégrée asymptotique $AMISE$. Dans le cas univarié, elles s'écrivent

$$AMSE = \frac{f(x)}{nh_n} \int_{\mathbb{R}} K^2(y) dy + \frac{h_n^4}{4} \left(\int_{\mathbb{R}} y^2 K(y) dy \right)^2 (f''(x))^2, \quad x \in \mathbb{R},$$

$$AMISE = \frac{1}{nh_n} \int_{\mathbb{R}} K^2(y) dy + \frac{h_n^4}{4} \left(\int_{\mathbb{R}} y^2 K(y) dy \right)^2 \int_{\mathbb{R}} f''^2(x) dx,$$

où f est supposée de classe C^2 et K est une densité de probabilité, paire, de carré intégrable, et de moment d'ordre 2 fini. Ces deux résultats ont été obtenus par Rosenblatt (1956) et Parzen (1962), puis généralisés au cas multivarié et au cas f de classe quelconque, par de nombreux auteurs. Pour plus de détails nous renvoyons au livre de Bosq et Lecoutre (1986). D'autres théorèmes limite ont été étudiés, notamment, la loi limite et la loi du logarithme itéré; à ce sujet, on pourra trouver une bibliographie détaillée dans Prakasa Rao (1983). Ce que nous pouvons retenir de toutes ces études est qu'un noyau K , convenablement normalisé par le paramètre de lissage h_n , influe peu sur la qualité de l'estimateur f_n^{PR} . D'ailleurs, Epanechnikov (1969) a trouvé le noyau optimal au sens de l' $AMISE$. A cause de sa simplicité, la majorité des auteurs optent pour le noyau normal. Quant au choix du paramètre h_n , il est crucial pour la précision de f_n^{PR} . Un h_n trop petit entraîne des fluctuations trop importantes dans f_n^{PR} , tandis qu'un h_n trop grand donne un lissage exagéré. Là encore, de nombreux auteurs se sont penchés sur ce problème sans toutefois y apporter une solution définitive. Nous donnons dans ce qui suit, un bref aperçu des méthodes les plus utilisées. Afin de simplifier l'exposé, nous relaterons ces méthodes dans le cas univarié seulement.

1. Critère asymptotique

Rosenblatt (1956) et Parzen (1962) ont montré que le paramètre de lissage donné par

$$h_n = \left(\frac{\int_{\mathbb{R}} K^2(y) dy}{\left(\int_{\mathbb{R}} y^2 K(y) dy \right)^2} \right)^{\frac{1}{5}} \frac{1}{\left(\int_{\mathbb{R}} f''^2(x) dx \right)^{\frac{1}{5}}} n^{-\frac{1}{5}}$$

$$= \alpha(K)\beta(f)n^{-\frac{1}{5}}$$

est optimal au sens de l'*AMISE*. Or ce pas dépend de la densité inconnue f . Woodroof (1970) et Nadaraja (1974) ont suggéré une procédure en deux étapes, basée sur une pré-estimation de f , voir également Bosq et Lecoutre (1986). Scott, Tapia et Thompson (1977) ont choisi d'appliquer l'algorithme de Newton à la recherche du point fixe de la relation fonctionnelle

$$h = n^{-\frac{1}{5}} \alpha(K) \hat{\beta}(h)$$

$$\text{où } \hat{\beta}(h) = \left(\int_{\mathbb{R}} (f_n^{PR})''^2(x) dx \right)^{-\frac{1}{5}}.$$

Lorsque cette équation admet plusieurs solutions, la plus grande est retenue (eventuellement $h = 0$ peut se produire aussi). Une autre approche, utilisant ce critère asymptotique, est celle de Deheuvels (1977-b). Elle consiste à remplacer $\beta(f)$ par $\frac{S}{C^{\frac{1}{5}}}$, où $S = S(X_1, \dots, X_n)$ est une statistique estimant un paramètre de dispersion de la loi à estimer et C est une constante convenablement choisie et indépendante des observations X_i , $i = 1, \dots, n$. Plus particulièrement pour la densité normale $N(\mu, \sigma)$, Deheuvels pose $\beta(f) = \frac{S_n}{(3/8\sqrt{\pi})^{\frac{1}{5}}}$ où S_n^2 est la variance empirique. Ce choix fournit un pas équivalent au pas optimal car pour $N(\mu, \sigma)$ la quantité $\beta(f)$ vaut $\frac{\sigma}{(3/8\sqrt{\pi})^{\frac{1}{5}}}$. Ensuite, Deheuvels a opté pour ce dernier choix, en disant qu'à cause de la puissance 1/5, des choix de C très différents de $3/8\sqrt{\pi}$ changent peu les valeurs du pas de lissage. On pourra trouver une amélioration de cette méthode dans Deheuvels et Hominal (1980).

2. Méthode de la validation croisée

Parmi les approches, qui font intervenir les observations pour évaluer le paramètre de lissage optimal, la validation croisée est la méthode qui a eu le plus de succès. Les travaux de Stone (1974-1977) ont largement contribué à son adoption par de nombreux auteurs. Sa technique est utilisée dans plusieurs domaines d'estimation non paramétrique. Pour des applications ultérieures, nous la présentons dans le cadre d'un estimateur à noyau plus général que f_n^{PR} i.e.

$$f_n(x, D, \tau) = \frac{1}{n} \sum_{i=1}^n K(x, X_i, \tau), \quad x \in \mathbb{R},$$

où τ est le paramètre de lissage, $D = \{X_1, \dots, X_n\}$ et K une fonction réelle définie sur \mathbb{R} .

La validation croisée est un critère basé sur la prédiction. Aussi, $f_n(\cdot, D, \tau)$ peut être aperçu comme un prédicteur de $\delta_Y(\cdot)$, la distribution dégénérée associée à la v.a. Y de loi f . Ayant fixé une fonction de perte $L(\cdot, \cdot)$, le critère optimal serait de minimiser en τ

$$\begin{aligned} \mathcal{L}^*(\tau, D) &= E_Y \{L(\delta_Y, f_n(\cdot, D, \tau) | D)\} \\ &= L^*(f, f_n(\cdot, D, \tau)). \end{aligned}$$

La solution optimale dépend de la quantité inconnue f , c'est pourquoi on utilise l'estimateur $\bar{\mathcal{L}}(\tau, D)$ de $\mathcal{L}^*(\tau, D)$, défini par

$$\bar{\mathcal{L}}(\tau, D) = \frac{1}{n} \sum_{i=1}^n L(\delta_{X_i}, f_n(\cdot, D, \tau)).$$

Donc un choix "naïf" du paramètre de lissage τ serait de prendre $\tau^0(D)$ qui minimise $\bar{\mathcal{L}}(\tau, D)$. Dans ce contexte, la validation croisée consiste à subdiviser l'échantillon D en deux sous-échantillons: un échantillon $D_i = \{X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n\}$ dit de

construction et un autre (X_i) dit de *validation*, pour $i = 1, \dots, n$. Ensuite elle évalue la performance de chaque $f_{n-1}(\cdot, D_i, \tau)$ à prédire δ_{X_i} , autrement dit il faut minimiser, en τ ,

$$\mathcal{L}(\tau, D) = \frac{1}{n} \sum_{i=1}^n L(\delta_{X_i}, f_{n-1}(\cdot, D_i, \tau)).$$

Pour la fonction de perte $L(\cdot, \cdot)$ le plus souvent on utilise soit celle de Kullback-Leibler

$$L(p, q) = \int_{\mathbb{R}} p(x) \log \{p(x)/q(x)\} dx,$$

soit la fonction de perte quadratique

$$L(p, q) = \int_{\mathbb{R}} \{p(x) - q(x)\}^2 dx.$$

A cause de δ_{X_i} , la quantité $\mathcal{L}(\tau, D)$ est infinie pour la fonction de perte Kullback-Leibler. Pour remédier à cela, au lieu d'évaluer directement la performance de $f_{n-1}(\cdot, D_i, \tau)$, il suffit de la comparer avec celle de la vraie densité f en utilisant la fonction de perte

$$H(p, q, r) = L(p, q) - L(p, r).$$

Ainsi la méthode de la validation croisée consiste à minimiser

$$\tilde{\mathcal{L}}(\tau, D) = \frac{1}{n} \sum_{i=1}^n H(\delta_{X_i}, f_{n-1}(\cdot, D_i, \tau), f(\cdot)).$$

Dans ce cas, lorsqu'il s'agit de la fonction de perte de Kullback-Leibler il suffit de minimiser en τ

$$L_1(\tau) = \frac{1}{n} \sum_{i=1}^n \log \{f(X_i)/f_{n-1}(X_i, D_i, \tau)\}$$

ou équivalamment maximiser en τ

$$L'_1(\tau) = \prod_{i=1}^n f_{n-1}(X_i, D_i, \tau),$$

Habbema, Hermans et Van den Broek (1974), et Duin (1976), en utilisant le principe du maximum de vraisemblance, ont abouti au critère $L'_1(\cdot)$. Le lien entre les deux critères $L_1(\cdot)$ et $L'_1(\cdot)$ a été fait par Bowman (1984). Enfin lorsqu'il s'agit de la fonction de perte quadratique la validation croisée fournit un pas de lissage qui minimise

$$L_2(\tau) = \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}} f_{n-1}^2(x, D_i, \tau) dx - \frac{2}{n} \sum_{i=1}^n f_{n-1}(X_i, D_i, \tau) + \frac{2}{n} \sum_{i=1}^n f(X_i) - \int_{\mathbb{R}} f^2(x) dx.$$

Les deux derniers termes sont indépendants de τ et peuvent être ignorés. Ce critère a été présenté en terme de validation croisée par Bowman (1982) et comme estimateur du risque quadratique intégré par Rudemo (1982).

Le comportement asymptotique des deux critères $L_1(\cdot)$ et $L_2(\cdot)$ a été traité séparément par plusieurs auteurs e.g. Schuster et Gregory(1981), Chow, Geman et Wu (1982), Hall (1982-1985), Stone (1984), Marron(1985), etc.

D'autres approches, non moins intéressantes, ont été proposées: procédure du "test graph", critère d'ajustement basé sur la statistique de Cramer-von Mises, méthode du $k^{\text{ième}}$ plus proche voisin etc. Pour plus de détails nous renvoyons à Bean et Tsokos (1982), Bowman (1985), Titterington (1985), et Silverman (1986).

Ce qui ressort de toutes ces approches est que le paramètre de lissage doit dépendre des observations. Ceci a incité de nombreux auteurs à proposer des variantes de f_n^{PR} , où h_n tient compte de l'information apportée par l'échantillon. A titre non exhaustif nous citons Loftsgaarden et Quesenberry (1965), Breiman, Meisel et Purcell (1977), Shanmugam (1977), Deheuvels (1977-b), Devroye et Wagner (1980) et Bierens (1983). Dans ce travail nous nous proposons de construire une généralisation de f_n^{PR} qui contient comme cas particuliers, un bon nombre de ces estimateurs.

En effet, soit l'estimateur à noyau f_n , défini par

$$f_n(x) = \frac{1}{n|\det H_n|} \sum_{i=1}^n K(H_n^{-1}(x - \phi_n(\omega, X_i))), \quad x \in \mathbb{R}^k,$$

où $\phi_n(.,.)$ est une fonction définie de $\mathbb{R}^{k \times n} \times \mathbb{R}^k$ dans \mathbb{R}^k et H_n est une matrice $k \times k$ inversible, pouvant dépendre de l'échantillon $\omega = (X_1, \dots, X_n)$.

L'estimateur de Deheuvels (1977-a), $f_n^D(x)$, est obtenu pour

$$\begin{cases} \phi_n(\omega, X_i) = X_i, & i = 1, \dots, n, \\ H_n \text{ est une matrice } k \times k \text{ inversible ne dépendant pas de } \omega. \end{cases}$$

Celui de Devroye et Wagner (1983), $f_n^{DW}(x)$, correspond à

$$\begin{cases} \phi_n(\omega, X_i) = X_i, & i = 1, \dots, n, \\ H_n = \begin{pmatrix} h_n(\omega) & & 0 \\ & \ddots & \\ 0 & & h_n(\omega) \end{pmatrix}, \end{cases}$$

où $(h_n(.))$ est une suite de réels positifs, dépendant de l'échantillon ω et telle que $h_n(.)$ est égal à $o(1)$ p.s.

Les estimateurs de Shanmugam (1977) et de Bierens (1983) font partie d'une classe d'estimateurs à noyau construits à partir de f_n de la façon suivante: Supposons

que f admet un vecteur moyen μ et une matrice de variance-covariance V . Notons μ_n un estimateur de μ et V_n , à une constante multiplicative près α_n , la matrice de variance-covariance empirique $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)(X_i - \bar{X}_n)^t$, où “ t ” désigne la transposition, et $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. Alors en plus d’estimer f , nous imposons à f_n les conditions suivantes:

$$\left\{ \begin{array}{l} f_n(\cdot) \text{ est une densité de probabilité,} \\ \int_{\mathbb{R}^k} x f_n(x) dx = \mu_n, \\ \int_{\mathbb{R}^k} (x - \mu_n)(x - \mu_n)^t f_n(x) dx = V_n. \end{array} \right. \quad (1)$$

Si la matrice H_n est symétrique et le noyau K vérifie

$$K(x) \geq 0, \quad x \in \mathbb{R}^k, \quad \int_{\mathbb{R}^k} x K(x) dx = 0, \quad \int_{\mathbb{R}^k} x x^t K(x) dx = I, \quad (2)$$

où I est la matrice identité, alors (1) est réalisé si H_n , ϕ_n , μ_n , et V_n satisfont

$$\left\{ \begin{array}{l} \frac{1}{n} \sum_{i=1}^n \phi_n(\omega, X_i) = \mu_n, \\ H_n^2 = V_n - \frac{1}{n} \sum_{i=1}^n (\phi_n(\omega, X_i) - \mu_n)(\phi_n(\omega, X_i) - \mu_n)^t. \end{array} \right. \quad (3)$$

Dans le cas où $V_n = \alpha_n S_n^2$ avec $\alpha_n \geq 1$ et $\alpha_n \rightarrow 1$ quand $n \rightarrow \infty$ et μ_n est laissé non spécifié, alors une solution de (3) est donnée par

$$\left\{ \begin{array}{l} \phi_n(\omega, X_i) = \sqrt{1 - \gamma_n^2} (X_i - \bar{X}_n) + \mu_n, \quad i = 1, \dots, n, \quad \gamma_n \in (0, 1], \\ H_n^2 = (\alpha_n - 1 + \gamma_n^2) S_n^2. \end{array} \right. \quad (4)$$

L’estimateur correspondant à (4), noté $f_n^B(\cdot)$, généralise celui de Bierens (1983) obtenu pour $\mu_n = \bar{X}_n$, $V_n = S_n^2$, et un noyau normal.

Lorsque μ_n est choisi égal à \bar{X}_n et $V_n = \alpha_n S_n^2$ avec $\alpha_n > 1$, alors une solution de (3) est

$$\begin{cases} \phi_n(\omega, X_i) = X_i, & i = 1, \dots, n, \\ H_n^2 = (\alpha_n - 1)S_n^2. \end{cases} \quad (5)$$

La solution (5) nous donne un estimateur $f_n^S(\cdot)$ qui généralise celui de Shanmugam (1977) correspondant à un noyau normal K et à $\alpha_n = 1 + \frac{n}{n-1} \frac{1}{n^{2\beta}}$ où $\beta \in (0, \frac{1}{2})$ est fixé.

Pour le choix de μ_n , dans (3), si nous n'avons aucune information sur le modèle étudié, nous prendrons $\mu_n = \bar{X}_n$. Par contre, si la loi est univariée et nous savons, par exemple, que f est symétrique alors nous opterons pour d'autres estimateurs de μ qui sont robustes et plus efficaces que \bar{X}_n . En effet, lorsque f a une queue lourde, alors, comme Huber (1972) l'a recommandé, nous prendrons l'estimateur X_n^* de Gastwirth (1966), qui est construit à partir de quelques statistiques d'ordre (par exemple $X_n^* = 0.3X_{(n/3)} + 0.4X_{(n/2)} + 0.3X_{(2n/3)}$, où $X_{(i)}$ est la statistique d'ordre de rang i). Dans le cas où la loi est approximativement normale, la moyenne α -tronquée $\bar{X}_{n\alpha}$ (avec $\alpha=0.1$ ou 0.15) semble être un bon estimateur de μ (Huber (1972)). Il existe d'autres estimateurs adaptés à des situations autres que celles citées. Pour plus de détails nous renvoyons à Gastwirth et Cohen (1970), Huber (1981) et Lehmann (1983). Pour le cas multivarié, le choix d'estimateurs autres que \bar{X}_n est plus limité. On pourra prendre la médiane spatiale (voir Gower (1974) ou Ducharme et Milasevic (1987)) ou les M -estimateurs multivariés (voir Huber (1981), chapitre 8).

Le premier chapitre de ce travail porte sur la convergence uniforme p.s. de f_n . Nous appliquerons ce résultat à f_n^B , f_n^S , f_n^D , et f_n^{DW} . Le deuxième chapitre est consacré à l'étude de la convergence L^1 p.s. de f_n . Des applications aux cas f_n^B , f_n^S , f_n^D , et f_n^{DW} sont également données. Dans le troisième et dernier chapitre, par le biais de simulations de quelques modèles connus, nous comparerons f_n^{PR} et les estimateurs f_n^B correspondants à $\mu_n = \bar{X}_n$, ou $\bar{X}_{n\alpha}$ ou X_n^* ou encore la médiane empirique \tilde{X}_n . Le choix des paramètres de lissage γ_n et h_n , dans f_n^B et f_n^{PR} respectivement, sera effectué selon les critères $L_1(\cdot)$, $L_2(\cdot)$ et le critère asymptotique de Deheuvels (1977-b).

CHAPITRE I

CONVERGENCE UNIFORME PRESQUE SÛRE

I.1 Conditions suffisantes pour la convergence uniforme presque sûre de f_n

Ce chapitre concerne l'étude de la convergence uniforme presque sûre de f_n vers f . Ce résultat sera ensuite appliqué aux estimateurs f_n^B , f_n^S , f_n^{DW} , et f_n^D . Les techniques de démonstration s'inspirent de celles utilisées par Abdous et Berlinet (1986) et Devroye et Wagner (1980). Commençons par fixer quelques notations: pour tout

$x = (x_1, \dots, x_k) \in \mathbb{R}^k$, $k \geq 1$, prenons $\|x\| = \sum_{i=1}^k |x_i|$, et pour toute matrice $k \times k$,

$A = (A_{ij})$, posons $\|A\|_\infty = \max_{i=1, \dots, k} \sum_{j=1}^k |A_{ij}|$. F désigne la fonction de répartition associée à f et ν la mesure correspondante, de même F_n est la fonction de répartition empirique et ν_n la mesure associée. Nous appellerons pavé de \mathbb{R}^k tout k -produit d'intervalles de la forme (a_i, b_i) ou $(a_i, b_i]$ ou $[a_i, b_i)$ ou $[a_i, b_i]$, avec $a_i \leq b_i$ et

$a_i, b_i \in \mathbb{R}$. Dans ce chapitre, nous supposerons que le noyau K vérifie les conditions suivantes:

(i) K est une densité de probabilité sur \mathbb{R}^k , bornée, on notera $M_1 = \sup_{y \in \mathbb{R}^k} K(y)$;

(ii) K est Riemann intégrable;

(iii) si $L(t) = \sup_{\|x\| \geq t} K(x)$, $t \in [0, \infty)$, alors $\int_0^\infty t^{k-1} L(t) dt \leq \infty$.

Pour tout $\omega = (X_1, \dots, X_n)$ fixé, l'application $\phi_n(\omega, \cdot)$ est bijective, son inverse et la valeur absolue de son Jacobien sont notés $\phi_n^{-1}(\omega, \cdot)$ et $J\phi_n(\omega, \cdot)$ respectivement. De plus nous imposons à H_n et ϕ_n les conditions suivantes

$$(iv) \quad \frac{1}{n|\det H_n|} \sum_{i=1}^n |1 - J\phi_n(\omega, X_i)| \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.};$$

$$(v) \quad \forall a > 0, \sup_{\substack{x \in \mathbb{R}^k \\ \|y\| \leq a}} \|\phi_n^{-1}(\omega, x - H_n y) - x\| \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.};$$

$$(vi) \quad \frac{n}{\log n} \left(\sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \right)^{-2} \xrightarrow{n \rightarrow \infty} \infty \quad \text{p.s.}$$

Théorème 1.1. Soit f une densité de probabilité uniformément continue. Si K , H_n , et ϕ_n vérifient (i) – (vi), alors

$$\sup_{x \in \mathbb{R}^k} |f_n(x) - f(x)| \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.} \quad (1.1)$$

La démonstration de ce résultat se fait à partir de l'inégalité

$$|f_n(x) - f(x)| \leq |f_n(x) - \tilde{f}_n(x)| + |\tilde{f}_n(x) - f_n^*(x)| + |f_n^*(x) - f(x)|, \quad (1.2)$$

où

$$\tilde{f}_n(x) = \frac{1}{n|\det H_n|} \sum_{i=1}^n J\phi_n(\omega, X_i) K(H_n^{-1}(x - \phi_n(\omega, X_i))), \quad x \in \mathbb{R}^k,$$

$$f_n^*(x) = \frac{1}{|\det H_n|} \int_{\mathbb{R}^k} J\phi_n(\omega, y) K(H_n^{-1}(x - \phi_n(\omega, y))) f(y) dy.$$

Le lemme 1.1 montre la convergence uniforme presque sûre des deux termes $|f_n(x) - \tilde{f}_n(x)|$ et $|f_n^*(x) - f(x)|$, tandis que les lemmes 1.2 et 1.3 serviront à montrer que $\sup_{x \in \mathbb{R}^k} |\tilde{f}_n(x) - f_n^*(x)| \xrightarrow[n \rightarrow \infty]{} 0$ p.s. .

Lemme 1.1. *Soit f une densité de probabilité uniformément continue. Si K , ϕ_n et H_n satisfont (i), (iv) et (v), alors*

$$(a) \quad \sup_{x \in \mathbb{R}^k} |f_n(x) - \tilde{f}_n(x)| \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{p.s.},$$

$$(b) \quad \sup_{x \in \mathbb{R}^k} |f_n^*(x) - f(x)| \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{p.s.}$$

DÉMONSTRATION:

(a) découle directement de (iv) et de l'inégalité

$$\sup_{x \in \mathbb{R}^k} |f_n(x) - \tilde{f}_n(x)| \leq \frac{M_1}{n|\det H_n|} \sum_{i=1}^n |1 - J\phi_n(\omega, X_i)|.$$

(b) En posant $z = H_n^{-1}(x - \phi_n(\omega, y))$, $f_n^*(x)$ s'écrira

$$f_n^*(x) = \int_{\mathbb{R}^k} K(z) f(\phi_n^{-1}(\omega, x - H_n z)) dz. \quad (1.3)$$

Fixons $\varepsilon > 0$, et soit $\delta(\varepsilon) > 0$ et $a(\varepsilon) > 0$ tels que

$$\|z - z'\| < \delta(\varepsilon) \implies |f(z) - f(z')| < \frac{\varepsilon}{2}, \quad (1.4)$$

et

$$\int_{\|z\| > a(\varepsilon)} K(z) dz \leq \frac{\varepsilon}{4M_2}, \quad \left(M_2 = \sup_{x \in \mathbb{R}^k} f(x) \right). \quad (1.5)$$

Pour tout $x \in \mathbb{R}^k$; (1.3)-(1.5) et (v) entraînent que

$$\begin{aligned} |f_n^*(x) - f(x)| &\leq \int_{\mathbb{R}^k} K(z) |f(\phi_n^{-1}(\omega, x - H_n z)) - f(x)| dz \\ &\leq \frac{\varepsilon}{2} + 2M_2 \int_{\Delta_n(x)} K(z) dz, \end{aligned} \quad (1.6)$$

où $\Delta_n(x) = \{z \in \mathbb{R}^k : \|\phi_n^{-1}(\omega, x - H_n z) - x\| > \delta(\varepsilon)\}$. Considérons maintenant les ensembles $\Delta_n = \bigcup_{x \in \mathbb{R}^k} \Delta_n(x)$ et Ω_0 formé de tous les échantillons ω pour lesquels (v) est vérifiée. Pour tout ω_0 fixé dans Ω_0 , on écrira $f_{n0}^*(x)$ et Δ_{n0} au lieu de $f_n^*(x)$ et Δ_n respectivement. Dans ce cas, lorsque n est assez grand nous obtenons:

$$\lambda^k \left(S(0, a(\varepsilon)) \cap \Delta_{n0} \right) = 0,$$

où λ^k est la mesure de Lebesgue sur \mathbb{R}^k et $S(0, a(\varepsilon)) = \{z \in \mathbb{R}^k : \|z\| \leq a(\varepsilon)\}$.

Par conséquent (1.6) implique que

$$\sup_{x \in \mathbb{R}^k} |f_{n0}^*(x) - f(x)| \leq \frac{\varepsilon}{2} + 2M_2 \int_{\|z\| > a(\varepsilon)} K(z) dz \leq \varepsilon.$$

Lemme 1.2. Soit K un noyau positif, satisfaisant (i) – (ii). Pour tout $\eta, \delta, \rho, > 0$, il existe une fonction

$$K^*(x) = \sum_{i=1}^N \alpha_i 1_{E_i}(x), \quad x \in \mathbb{R}^k,$$

telle que:

(a) $\alpha_1, \dots, \alpha_N$ sont des réels positifs;

(b) E_1, \dots, E_N sont des pavés disjoints dans $[-\rho, \rho]^k$;

(c) $\sup_{y \in \mathbb{R}^k} K^*(y) \leq M_1$;

(d) $|K^*(x) - K(x)| < \eta$ sur $[-\rho, \rho]^k \cap D^c$; où D vérifie:

(e) $D \subset B \subset \bigcup_{i=1}^M B_i$ où les B_i , $i = 1 \dots, M$, sont des pavés

de $[-\rho, \rho]^k$ et B est tel que $\lambda^k(B) < \delta$.

DÉMONSTRATION: K étant Riemann intégrable, il existe une partition de $[-\rho, \rho]^k$ en pavés disjoints, telle que les intégrales inférieure et supérieure de K , sur cette partition, diffèrent de moins de $\eta\delta$ (cf. Spivak (1965) chapitre 3). Notons K_1 et K_2 les fonctions correspondant à ces intégrales inférieure et supérieure respectivement. Alors

$$\left\{ x \in [-\rho, \rho]^k : K(x) - K_2(x) \geq \eta \right\} \subset \left\{ x \in [-\rho, \rho]^k : K_1(x) - K_2(x) \geq \eta \right\},$$

l'ensemble $\{x \in [-\rho, \rho]^k : K_1(x) - K_2(x) \geq \eta\}$ est l'union de pavés disjoints de mesure de Lebesgue inférieure à δ . En posant $K^*(x) = K_2(x)$ on obtient les assertions du lemme 1.2.

Dans ce qui suit nous allons reprendre quelques résultats généraux concernant l'approximation de la mesure théorique ν par la mesure empirique ν_n . Résultats que nous particulariserons par la suite à notre problème.

Soit \mathcal{R} une classe arbitraire de parties de \mathbb{R}^k ; sa complexité est mesurée par

$$m^{\mathcal{R}}(n) = \max_{(x_1, \dots, x_n) \in \mathbb{R}^{kn}} N_{\mathcal{R}}(x_1, \dots, x_n), \quad n \geq 1,$$

où $N_{\mathcal{R}}(x_1, \dots, x_n)$ désigne le nombre d'ensembles distincts dans

$$\left\{ \{x_1, \dots, x_n\} \cap B, \quad B \in \mathcal{R} \right\}.$$

S'il existe un n tel que $m^{\mathcal{R}}(n) < 2^n$, alors \mathcal{R} est dite une classe de Vapnik-Chervonenkis.

Considérons maintenant la classe \mathcal{R}_ω définie par

$$\mathcal{R}_\omega = \left\{ C_\omega : C_\omega = \{y : y = \phi_n^{-1}(\omega, H_n z), z \in C\}, \text{ où } C \in \mathcal{R} \right\}.$$

Pour tout ω fixé, $\phi_n(\omega, \cdot)$ étant un difféomorphisme sur \mathbb{R}^k , si \mathcal{R} est une classe de Vapnik-Chervonenkis alors \mathcal{R}_ω le sera aussi, avec $m^{\mathcal{R}}(r) = m^{\mathcal{R}_\omega}(r)$, $\forall r \geq 1$. Le lemme suivant donne une approximation uniforme, sur \mathcal{R} , de ν par ν_n . Ce résultat subsiste pour les classes \mathcal{R}_ω .

Lemme 1.3. *Pour tout $\varepsilon > 0$ et tout $n > 2/\varepsilon^2$, on a*

$$Pr \left(\sup_{A \in \mathcal{R}} |\nu(A) - \nu_n(A)| \geq \varepsilon \right) \leq 4m^{\mathcal{R}}(2n) \exp(-\varepsilon^2 n/8).$$

Pour la démonstration de ce lemme, voir Vapnik et Chervonenkis (1971).

Dans toute la suite \mathcal{R} désignera la classe de tous les pavés de \mathbb{R}^k . Pour cette classe, nous savons que $m^{\mathcal{R}}(n) \leq 2n^{2k}$ (i.e. \mathcal{R} est une classe de Vapnik-Chervonenkis) (voir Bosq et Lecoutre (1985)).

Démonstration du théorème 1.1 : D'après le lemme 1.1 et l'inégalité (1.2), il suffit de montrer que

$$\sup_{x \in \mathbb{R}^k} |\tilde{f}_n(x) - f_n^*(x)| \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{p.s.}$$

En effet,

$$\sup_{x \in \mathbb{R}^k} |\tilde{f}_n(x) - f_n^*(x)| \leq U_1 + U_2 + U_3,$$

où

$$U_1 = \sup_{x \in \mathbb{R}^k} \int_{\mathbb{R}^k} \frac{J\phi_n(\omega, y)}{|\det H_n|} |K(H_n^{-1}(x - \phi_n(\omega, y))) - K^*(H_n^{-1}(x - \phi_n(\omega, y)))| dF(y),$$

$$U_2 = \sup_{x \in \mathbb{R}^k} \left| \int_{\mathbb{R}^k} \frac{J\phi_n(\omega, y)}{|\det H_n|} K^*(H_n^{-1}(x - \phi_n(\omega, y))) (dF_n(y) - dF(y)) \right|,$$

$$U_3 = \sup_{x \in \mathbb{R}^k} \int_{\mathbb{R}^k} \frac{J\phi_n(\omega, y)}{|\det H_n|} |K^*(H_n^{-1}(x - \phi_n(\omega, y))) - K(H_n^{-1}(x - \phi_n(\omega, y)))| dF_n(y).$$

Considérons la partition de \mathbb{R}^k , donnée par

$$C_1 = \left\{ y : H_n^{-1}(x - \phi_n(\omega, y)) \in [-\rho, \rho]^k \right\}^c,$$

$$C_2 = \left\{ y : H_n^{-1}(x - \phi_n(\omega, y)) \in [-\rho, \rho]^k \cap D^c \right\},$$

$$C_3 = \left\{ y : H_n^{-1}(x - \phi_n(\omega, y)) \in D \right\},$$

et majorons successivement les trois termes U_1 , U_2 , et U_3 . En effet, U_1 est borné par

$$\sum_{i=1}^3 \sup_{x \in \mathbb{R}^k} \int_{C_i} \frac{J\phi_n(\omega, y)}{|\det H_n|} |K(H_n^{-1}(x - \phi_n(\omega, y))) - K^*(H_n^{-1}(x - \phi_n(\omega, y)))| dF(y). \quad (1.7)$$

Or, ayant $K^*(H_n^{-1}(x - \phi_n(\omega, y))) = 0$ pour tout $y \in C_1$ (cf. lemme 1.2. (b)), le premier terme dans (1.7), est majoré par

$$\sup_{x \in \mathbb{R}^k} \int_{C_1} \frac{J\phi_n(\omega, y)}{|\det H_n|} K(H_n^{-1}(x - \phi_n(\omega, y))) dF(y), \quad (1.8)$$

tandis que le second terme est inférieur à (cf. lemme 1.2. (d))

$$\eta M_2 \sup_{x \in \mathbb{R}^k} \int_{C_1^c} \frac{J\phi_n(\omega, y)}{|\det H_n|} dy \leq \eta M_2 \lambda^k ([-\rho, \rho]^k) \leq \eta M_2 (2\rho)^k. \quad (1.9)$$

Quant au troisième terme, nous obtenons d'après le lemme 1.2. (c, e)

$$\begin{aligned} & \sup_{x \in \mathbb{R}^k} \int_{C_3} \frac{J\phi_n(\omega, y)}{|\det H_n|} |K(H_n^{-1}(x - \phi_n(\omega, y))) - K^*(H_n^{-1}(x - \phi_n(\omega, y)))| dF(y) \\ & \leq 2M_1 \sup_{x \in \mathbb{R}^k} \int_{C_3} \frac{J\phi_n(\omega, y)}{|\det H_n|} dF(y) \\ & \leq 2M_1 M_2 \lambda^k(D) \leq 2\delta M_1 M_2. \end{aligned} \quad (1.10)$$

Considérons maintenant U_2 . Par définition de $K^*(\cdot)$, nous avons

$$U_2 = \sup_{x \in \mathbb{R}^k} \left| \sum_{i=1}^N \alpha_i \int_{E_i^*} \frac{J\phi_n(\omega, y)}{|\det H_n|} (dF_n(y) - dF(y)) \right|,$$

où $E_i^* = \{y : H_n^{-1}(x - \phi_n(\omega, y)) \in E_i\}$, $i = 1, \dots, N$, et les E_i sont définis au lemme 1.2. (b); donc

$$\begin{aligned}
U_2 &\leq NM_1 \sup_{x \in \mathbb{R}^k} \max_{i=1, \dots, N} \left| \int_{E_i^*} \frac{J\phi_n(\omega, y)}{|\det H_n|} (dF_n(y) - dF(y)) \right| \\
&\leq NM_1 \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)|.
\end{aligned} \tag{1.11}$$

Finalemment, pour U_3 , nous obtenons

$$U_3 \leq I_1 + I_2 + I_3,$$

où, pour $i = 1, 2, 3$

$$I_i = \sup_{x \in \mathbb{R}^k} \int_{C_i} \frac{J\phi_n(\omega, y)}{|\det H_n|} |K(H_n^{-1}(x - \phi_n(\omega, y))) - K^*(H_n^{-1}(x - \phi_n(\omega, y)))| dF_n(y).$$

D'après le lemme 1.2 (b), I_1 est borné par

$$\sup_{x \in \mathbb{R}^k} \int_{C_1} \frac{J\phi_n(\omega, y)}{|\det H_n|} K(H_n^{-1}(x - \phi_n(\omega, y))) dF_n(y). \tag{1.12}$$

Ensuite

$$\begin{aligned}
I_2 &\leq \eta \sup_{x \in \mathbb{R}^k} \int_{C_2} \frac{J\phi_n(\omega, y)}{|\det H_n|} dF_n(y) \leq \eta \sup_{x \in \mathbb{R}^k} \int_{C_1^c} \frac{J\phi_n(\omega, y)}{|\det H_n|} dF(y) \\
&\quad + \eta \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{x \in \mathbb{R}^k} \left| \int_{C_1^c} (dF(y) - dF_n(y)) \right| \\
&\leq \eta(2\rho)^k M_2 + \eta \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)|,
\end{aligned} \tag{1.13}$$

nous avons utilisé le lemme 1.2 (d) et le fait que $C_2 \subset C_1^c$.

Enfin, comme $D \subset \bigcup_{i=1}^M B_i$,

$$I_3 \leq 2M_2 \sup_{x \in \mathbb{R}^k} \int_{B^*} \frac{J\phi_n(\omega, y)}{|\det H_n|} dF_n(y),$$

où

$$B^* = \{y : H_n^{-1}(x - \phi_n(\omega, y)) \in B\} = \bigcup_{i=1}^M B_i^*,$$

avec

$$B_i^* = \{y : H_n^{-1}(x - \phi_n(\omega, y)) \in B_i\} \quad i = 1, \dots, M.$$

Donc

$$\begin{aligned} I_3 &\leq 2M_1 \sup_{x \in \mathbb{R}^k} \int_{B^*} \frac{J\phi_n(\omega, y)}{|\det H_n|} dF(y) \\ &\quad + 2M_1 \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{x \in \mathbb{R}^k} \left| \int_{B^*} (dF(y) - dF_n(y)) \right| \\ &\leq 2M_1 M_2 \lambda^k(B) + 2MM_1 \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)|. \end{aligned} \tag{1.14}$$

A ce stade, regroupons les étapes (1.8) à (1.14) pour obtenir

$$\begin{aligned} \sup_{x \in \mathbb{R}^k} |\tilde{f}_n(x) - f_n^*(x)| &\leq 2\eta M_2 (2\rho)^k + 4\delta M_1 M_2 \\ &\quad + \sup_{x \in \mathbb{R}^k} \int_{C_1} \frac{J\phi_n(\omega, y)}{|\det H_n|} K(H_n^{-1}(x - \phi_n(\omega, y))) (dF(y) + dF_n(y)) \\ &\quad + [NM_1 + \eta + 2MM_1] \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)|. \end{aligned} \tag{1.15}$$

Maintenant, examinons la quantité

$$Q_n = \sup_{x \in \mathbb{R}^k} \int_{C_1} \frac{J\phi_n(\omega, y)}{|\det H_n|} K(H_n^{-1}(x - \phi_n(\omega, y))) \{dF(y) + dF_n(y)\}.$$

Si K est à support compact, alors on peut toujours choisir ρ de telle sorte que $Q_n = 0$; sinon, par définition de $L(\cdot)$, nous avons

$$K(z) \leq L(\|z\|) \quad \text{pour tout } z \in \mathbb{R}^k,$$

par conséquent, $Q_n \leq Q_{n1} + Q_{n2}$ où

$$Q_{n1} = 2 \sup_{x \in \mathbb{R}^k} \int_{C_1} \frac{J\phi_n(\omega, y)}{|\det H_n|} L(\|H_n^{-1}(x - \phi_n(\omega, y))\|) dF(y),$$

$$Q_{n2} = \sup_{x \in \mathbb{R}^k} \left| \int_{C_1} \frac{J\phi_n(\omega, y)}{|\det H_n|} L(\|H_n^{-1}(x - \phi_n(\omega, y))\|) (dF_n(y) - dF(y)) \right|.$$

En majorant $dF(\cdot)$ par M_2 et en posant $z = H_n^{-1}(x - \phi_n(\omega, y))$, nous obtenons pour Q_{n1}

$$Q_{n1} \leq 2M_2 \int_{([-\rho, \rho]^k)^c} L(\|z\|) dz$$

ensuite le changement de variable $t = \|z\| = \sum_{j=1}^k |z_j|$ et $t_i = z_i$ pour $i = 2, \dots, k$, nous donne

$$Q_{n1} \leq 2^k M_2 \int_{\rho}^{\infty} t^{k-1} L(t) dt. \quad (1.16)$$

Par ailleurs, afin de majorer Q_{n2} , nous aurons besoin du résultat suivant

$$\text{si } x \in ([-\rho, \rho]^k)^c, \quad \text{alors} \quad |L(\|x\|) - L'(x)| < \frac{1}{l} L(\rho), \quad (1.17)$$

où

$$L'(x) = \sum_{j=1}^l \frac{j-1}{l} L(\rho) 1_{S_j}(x), \quad x \in \mathbb{R}^k,$$

et

$$S_j = \left\{ x : \frac{j-1}{l} L(\rho) < L(\|x\|) < \frac{j}{l} L(\rho) \right\}, \quad j = 1, \dots, l.$$

En effet, pour tout $x \in ([-\rho, \rho]^k)^c$, nous avons deux possibilités: il existe un j_0 dans $\{1, \dots, l\}$ tel que $x \in S_{j_0}$ ou bien un tel j_0 n'existe pas. Dans le premier cas, nous aurons $L'(x) = \frac{j_0-1}{l} L(\rho)$ et $\frac{j_0-1}{l} L(\rho) < L(\|x\|) \leq \frac{j_0}{l} L(\rho)$, ce qui donne (1.17); dans le second cas, nous aurons $L(\|x\|) = L'(x) = 0$, ce qui donne encore (1.17). Maintenant nous pouvons écrire

$$Q_{n2} \leq Q_{n2}^1 + Q_{n2}^2,$$

où

$$Q_{n2}^1 = \sup_{x \in \mathbb{R}^k} \int_{C_1} \frac{J\phi_n(\omega, y)}{|\det H_n|} |L'(H_n^{-1}(x - \phi_n(\omega, y)))(dF_n(y) - dF(y))|,$$

$$Q_{n2}^2 = \sup_{x \in \mathbb{R}^k} \int_{C_1} \frac{J\phi_n(\omega, y)}{|\det H_n|} |L'(H_n^{-1}(x - \phi_n(\omega, y))) - L(\|H_n^{-1}(x - \phi_n(\omega, y))\|)| \{dF_n(y) + dF(y)\}.$$

Donc (1.17) entraîne que

$$\begin{aligned} Q_{n2}^2 &\leq \frac{L(\rho)}{l} \sup_{x \in \mathbb{R}^k} \int_{C_1} \frac{J\phi_n(\omega, y)}{|\det H_n|} \{dF_n(y) + dF(y)\} \\ &\leq \frac{2L(\rho)}{l} \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right]. \end{aligned} \tag{1.18}$$

Ensuite

$$Q_{n2}^1 = \sup_{x \in \mathbb{R}^k} \left| \frac{L(\rho)}{l} \sum_{j=1}^l (j-1) \int_{S_j^* \cap C_1} \frac{J\phi_n(\omega, y)}{|\det H_n|} (dF_n(y) - dF(y)) \right|,$$

où

$$S_j^* = \{y : H_n^{-1}(x - \phi_n(\omega, y)) \in S_j\}, \quad j = 1, \dots, l.$$

Par ailleurs, si on pose

$$T_j^* = \bigcup_{i=j}^l = \left\{ y : H_n^{-1}(x - \phi_n(\omega, y)) \in \bigcup_{i=j}^l S_i \right\}, \quad j = 1, \dots, l,$$

alors

$$\begin{aligned} & \sum_{j=1}^l (j-1) \int_{S_j^*} \frac{J\phi_n(\omega, y)}{|\det H_n|} (dF_n(y) - dF(y)) \\ &= \sum_{j=2}^l (j-1) \int_{S_j^*} \frac{J\phi_n(\omega, y)}{|\det H_n|} (dF_n(y) - dF(y)) \\ &= \sum_{j=2}^l \sum_{i=j}^l \int_{S_j^*} \frac{J\phi_n(\omega, y)}{|\det H_n|} (dF_n(y) - dF(y)) \\ &= \sum_{j=2}^l \int_{T_j^*} \frac{J\phi_n(\omega, y)}{|\det H_n|} (dF_n(y) - dF(y)); \end{aligned}$$

d'où

$$\begin{aligned} Q_{n2}^1 &\leq \frac{L(\rho)}{l} \sup_{\substack{x \in \mathbb{R}^k \\ j=1, \dots, l}} \left| \int_{T_j^* \cap C_1} \frac{J\phi_n(\omega, y)}{|\det H_n|} (dF_n(y) - dF(y)) \right| \\ &\leq \frac{L(\rho)}{l} \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)|. \end{aligned} \tag{1.19}$$

Finalement, à partir de (1.15) et en vertu de (1.16), (1.18) et (1.19), nous concluons que

$$\begin{aligned}
\sup_{x \in \mathbb{R}^k} |\tilde{f}_n(x) - f_n^*(x)| &\leq 2\eta M_2 (2\rho)^k + 4\delta M_1 M_2 + 2^k M_2 \int_{\rho}^{\infty} t^{k-1} L(t) dt \\
&+ [NM_1 + \eta + 2MM_1 + 2L(\rho)] \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)| \\
&+ \frac{2L(\rho)}{l} \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right]
\end{aligned} \tag{1.20}$$

l étant arbitraire, nous pouvons le prendre égal à $(1 + l^*)$ où l^* est la partie entière de $\sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right]$; ainsi nous aurons

$$\frac{1}{l} \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \leq 1. \tag{1.21}$$

Ensuite, fixons $\varepsilon > 0$ et considérons $\gamma(\varepsilon) > 0$ tel que

$$\rho > \gamma(\varepsilon) \implies \int_{\rho}^{\infty} t^{k-1} L(t) dt < \frac{\varepsilon}{4} \frac{1}{2^k M_2} \quad \text{et} \quad L(\rho) < \frac{\varepsilon}{8}. \tag{1.22}$$

D'autre part, η et δ étant arbitraires, choisissons-les tels que

$$\delta < \frac{\varepsilon}{16M_1 M_2} \quad \text{et} \quad \eta \leq \frac{\varepsilon}{8M_2 (2\rho)^k}. \tag{1.23}$$

Par conséquent, sous (1.21)-(1.23) et (vi), l'inégalité (1.20) devient

$$\sup_{x \in \mathbb{R}^k} |\tilde{f}_n(x) - f_n^*(x)| \leq \varepsilon + \theta \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)|,$$

avec $\theta = NM_1 + \eta + 2MM_1 + L(\rho)$. ρ , δ , et η étant fixés de façon à ce que (1.22) et (1.23) soient vérifiées, le lemme 1.3 nous permet de fixer N et M , donc aussi θ . Pour

obtenir le théorème 1.1 il suffira donc de montrer que

$$\sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)| \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.}$$

En effet, en utilisant la condition (vi), définissons la suite d'entiers (n'_j) par

$$n'_1 = 1$$

$$n'_j = \inf \left\{ n : n > n'_{j-1}, \Pr \left(\bigcup_{m \geq n} \left\{ \frac{m}{\log m} \left(\sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_m(\omega, y)}{|\det H_m|} \right] \right)^{-2} \leq j \right\} \right) \leq 2^{-j} \right\}, \quad j \geq 2.$$

Posons $h'_n = j$ sur $[n'_j, n'_{j+1})$, $j \geq 2$. La suite h'_n ainsi obtenue ne dépend pas de ω et vérifie $h'_n \xrightarrow{n \rightarrow \infty} \infty$ et de plus

$$\frac{n}{\log n} \left(\sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \right)^{-2} \geq h'_n, \quad \text{p.s.},$$

car pour $n \in [n'_j, n'_{j+1})$, on a

$$\begin{aligned} & \Pr \left(\bigcup_{m \geq n} \left\{ \frac{m}{\log m} \left(\sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_m(\omega, y)}{|\det H_m|} \right] \right)^{-2} \leq h'_m \right\} \right) \\ & \leq \sum_{i=j}^{\infty} \Pr \left(\bigcup_{n'_i \leq m < n'_{i+1}} \left\{ \frac{m}{\log m} \left(\sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_m(\omega, y)}{|\det H_m|} \right] \right)^{-2} \leq i \right\} \right) \\ & \leq \sum_{i=j}^{\infty} 2^{-i} = 2^{-j+1} \xrightarrow{j \rightarrow \infty} 0. \end{aligned}$$

Par conséquent, $\forall \varepsilon > 0$

$$\begin{aligned}
& \sum_{n \geq 1} \Pr \left(\sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)| > \varepsilon \right) \\
& \leq \sum_{n \geq 1} \Pr \left(\sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)| > \varepsilon \left(\frac{h'_n \log n}{n} \right)^{1/2} \right) \\
& \leq \sum_{n: h'_n \log n < 2/\varepsilon^2} \Pr \left(\sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)| > \varepsilon \left(\frac{h'_n \log n}{n} \right)^{1/2} \right) \\
& + \sum_{n: h'_n \log n \geq 2/\varepsilon^2} \Pr \left(\sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)| > \varepsilon \left(\frac{h'_n \log n}{n} \right)^{1/2} \right) \\
& \leq \sum_{n: h'_n \log n < 2/\varepsilon^2} \Pr \left(\sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)| > \varepsilon \left(\frac{h'_n \log n}{n} \right)^{1/2} \right) \\
& + 8 \sum_{n: h'_n \log n \geq 2/\varepsilon^2} (2n)^{2k} \exp(-\varepsilon^2 h'_n \log n) < \infty.
\end{aligned}$$

Dans la dernière inégalité, nous avons utilisé le lemme 1.3 et le fait que $m^{\mathcal{R}_\omega}(n) = m^{\mathcal{R}}(n) \leq 2n^{2k}$ (voir Bosq et Lecoutre (1985)). Ceci conclut la preuve du théorème 1.1.

REMARQUE

Si, dans les conditions (iv) et (vi) du théorème 1.1, nous remplaçons l'expression *presque sûrement* par l'expression *en probabilité*, et si au lieu de (vi) nous utilisons la condition:

$$(vi') \quad n \left(\sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \right)^{-2} \xrightarrow[n \rightarrow \infty]{} \infty \text{ en probabilité,}$$

alors, en suivant les étapes de la preuve du théorème 1.1, nous aboutissons à

$$\sup_{x \in \mathbb{R}^k} |\tilde{f}_n(x) - f_n^*(x)| \leq \varepsilon + \theta \sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \sup_{A \in \mathcal{R}_\omega} |\nu(A) - \nu_n(A)|.$$

Ensuite on pourra toujours exhiber une suite de réels positifs (β_n) , telle que $\beta_n \xrightarrow[n \rightarrow \infty]{} \infty$ et

$$\sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \leq \left(\frac{\beta_n}{n} \right)^{-1/2} \quad \text{en probabilité,}$$

et appliquer le lemme 1.3 pour obtenir la convergence en probabilité vers 0 de

$$\sup_{x \in \mathbb{R}^k} |\tilde{f}_n(x) - f_n^*(x)|.$$

I.2 Applications

On s'intéresse dans cette partie, aux applications du théorème 1.1 aux estimateurs f_n^B , f_n^S , f_n^D , et f_n^{DW} (voir pages 8-10).

I.2.1 L'estimateur f_n^B

Rappelons que f_n^B est une généralisation de l'estimateur à noyau de Bierens (1983) (obtenu à partir de f_n^B pour $\mu_n = \bar{X}_n$). L'étude de Bierens a été faite dans le cadre du noyau normal seulement, pour lequel la convergence uniforme en probabilité a été établie. Le théorème suivant fournit une condition suffisante pour la convergence uniforme (sur un compact $W \subset \mathbb{R}^k$) presque sûre de f_n^B vers f .

Théorème 1.2. *Soit f une densité de probabilité uniformément continue admettant une moyenne μ et une matrice de variance-covariance V . Si le noyau K vérifie (i) – (ii), la suite (γ_n) est telle que $\gamma_n \xrightarrow{n \rightarrow \infty} 0$ et $\frac{n\gamma_n^{2k}}{\log n} \xrightarrow{n \rightarrow \infty} \infty$ et $\mu_n \xrightarrow{n \rightarrow \infty} \mu$ p.s., alors*

$$\sup_{x \in \mathbb{R}^k} |f_n^B(x) - f(x)| \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.}$$

où $W \subset \mathbb{R}^k$ est un compact.

DÉMONSTRATION: Par définition de f_n^B (voir (4) page 9),

$$J\phi_n(\omega, X_i) = \sqrt{1 - \gamma_n^{2k}} \quad i = 1, \dots, n$$

et

$$\phi_n^{-1}(\omega, x - H_n z) = \frac{x - \gamma_n S_n z + \sqrt{1 - \gamma_n^2} \bar{X}_n - \mu_n}{\sqrt{1 - \gamma_n^2}};$$

donc

$$\frac{1}{n |\det H_n|} \sum_{i=1}^n |1 - J\phi_n(\omega, X_i)| = \frac{(1 - \sqrt{1 - \gamma_n^2})^k}{\gamma_n^k |\det S_n|}.$$

Si, dans l'énoncé du théorème 1.1, nous remplaçons la condition (v) par:

$$(v') \quad \forall a > 0, \sup_{\substack{x \in \bar{W} \\ \|z\| \leq a}} \|\phi_n^{-1}(\omega, x - H_n z) - x\| \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.},$$

alors

$$\sup_{x \in \bar{W}} |f_n^B(x) - f(x)| \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s. .}$$

Il reste donc à vérifier la validité de (iv), (v'), et (vi). En effet, la condition (iv) du théorème 1.1 est réalisée, car $|\det S_n| \xrightarrow{n \rightarrow \infty} |\det V|$ p.s. et $\frac{(1 - \sqrt{1 - \gamma_n^2})^k}{\gamma_n^k} \xrightarrow{n \rightarrow \infty} 0$, puisque $\gamma_n \xrightarrow{n \rightarrow \infty} 0$. Ensuite

$$\|\phi_n^{-1}(\omega, x - H_n z) - x\| \leq \frac{1 - \sqrt{1 - \gamma_n^2}}{\sqrt{1 - \gamma_n^2}} \|x\| + \left\| \bar{X}_n - \frac{\mu_n}{\sqrt{1 - \gamma_n^2}} \right\| + \frac{\gamma_n}{\sqrt{1 - \gamma_n^2}} \|S_n z\|;$$

puisque $\gamma_n \xrightarrow{n \rightarrow \infty} 0$, $S_n \xrightarrow{n \rightarrow \infty} V$ p.s., $\bar{X}_n \xrightarrow{n \rightarrow \infty} \mu$ p.s., et $\mu_n \xrightarrow{n \rightarrow \infty} \mu$ p.s. on a la condition (v'). tout x et z Quant à la condition (vi), elle est vérifiée car

$$\left(\sup_{y \in \mathbb{R}^k} \left[\frac{J\phi_n(\omega, y)}{|\det H_n|} \right] \right)^{-2} = \frac{\gamma_n^{2k}}{(1 - \gamma_n^2)^k} |\det S_n|^2 \geq \gamma_n^{2k} |\det S_n|^2.$$

REMARQUE

Sous des conditions sur K , plus fortes que celles du théorème 1.2, et en utilisant des techniques de démonstration différentes de celles utilisées ici, nous avons (Abdous (1989)) établi la convergence uniforme (sur \mathbb{R}^k) presque sûre de l'estimateur de Bierens (1983).

1.2.2 Estimateur f_n^S

Shanmugam (1977) a proposé et a étudié f_n^S lorsque $\alpha_n = 1 + \frac{n}{n-1} \frac{1}{n^{2\beta}}$ et lorsque le noyau est normal (voir (5) page 11). Il a en fait obtenu la convergence simple en probabilité. Le théorème 1.3 donne la convergence uniforme p.s. de f_n^S pour un α_n et un noyau quelconques.

Théorème 1.3. Soit f une densité de probabilité uniformément continue de moyenne μ et de matrice de variance-covariance V . Si K vérifie (i) – (iii) et (α_n) satisfait: $\alpha_n > 1$, $\alpha_n \xrightarrow[n \rightarrow \infty]{} 1$ et $\frac{n(\alpha_n - 1)^k}{\log n} \xrightarrow[n \rightarrow \infty]{} \infty$, alors

$$\sup_{x \in \mathbb{R}^k} |f_n^S(x) - f(x)| \xrightarrow[n \rightarrow \infty]{} 0 \quad p.s.$$

DÉMONSTRATION: D'après (5), $|\det H_n| = (\alpha_n - 1)^{\frac{k}{2}} |\det S_n|$, $J\phi_n(\omega, X_i) = 1$, $i = 1, \dots, n$ et $\phi_n^{-1}(\omega, x - H_n z) = x - \sqrt{\alpha_n - 1} S_n z$. Donc, les conditions (iv) – (vi) du théorème 1.1 sont satisfaites.

1.2.3 L'estimateur f_n^D

L'estimateur de Deheuvels (1977-a) est une forme généralisée de l'estimateur de Parzen-Rosenblatt. Au lieu d'avoir une matrice avec les éléments de la diagonale égaux à h_n et des 0 ailleurs, on prend une matrice pleine H_n ne dépendant que de n . Pour f_n^D nous obtenons

Théorème 1.4. Soit f une densité de probabilité uniformément continue. Si K vérifie (i) – (iii) et H_n est telle que $\frac{n|\det H_n|^2}{\log n} \xrightarrow[n \rightarrow \infty]{} \infty$, $\|H_n\|_\infty \xrightarrow[n \rightarrow \infty]{} 0$, alors

$$\sup_{x \in \mathbb{R}^k} |f_n^D(x) - f(x)| \xrightarrow[n \rightarrow \infty]{} 0 \quad p.s.$$

DÉMONSTRATION: Par définition de f_n^D , nous avons $J\phi_n(\omega, X_i) = 1$, et $\phi_n(\omega, X_i) = X_i$ pour $i = 1, \dots, n$. Les conditions (iv) – (vi) du théorème 1.1 se déduisent aisément des hypothèses du théorème 1.4.

REMARQUE

En imposant à K des conditions plus fortes que les nôtres, Deheuvels (1977-a) a établi une condition nécessaire et suffisante pour la convergence uniforme p.s. de f_n^D .

I.2.4 L'estimateur f_n^{DW}

L'estimateur de Devroye et Wagner (1980) est en fait celui de Parzen-Rosenblatt, mais avec un pas de lissage h_n dépendant des observations. Nous avons

Théorème 1.5. *Soit f une densité de probabilité uniformément continue. Si K satisfait (i) – (iii) et $(h_n(\cdot))$ est telle que $h_n(\cdot) \xrightarrow[n \rightarrow \infty]{} 0$ p.s. et $\frac{nh_n^{2k}(\cdot)}{\log n} \xrightarrow[n \rightarrow \infty]{} \infty$ p.s., alors*

$$\sup_{x \in \mathbb{R}^k} |f_n^{DW}(x) - f(x)| \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{p.s.}$$

DÉMONSTRATION:

Les conditions du théorème 1.5 sont exactement celles du théorème 1.1 lorsque

$$\phi_n(\omega, X_i) = X_i, \quad i = 1, \dots, n \text{ et } H_n = \begin{pmatrix} h_n(\omega) & & 0 \\ & \ddots & \\ 0 & & h_n(\omega) \end{pmatrix}.$$

CHAPITRE II

CONVERGENCE L^1 PRESQUE SÛRE

II.1 Conditions suffisantes pour la convergence L^1 presque sûre de f_n

Lorsqu'on veut, par l'intermédiaire de f_n , estimer ν en vue de calculer des probabilités d'événements, ou quand on applique une transformation monotone aux axes, la distance L^1 semble être la mieux adaptée pour mesurer l'efficacité de f_n . En effet, dans le premier cas, le lemme de Scheffé (1947) lie la distance L_1 entre f_n et f , et l'écart maximal entre ν'_n et ν par la relation

$$\int_{\mathbb{R}^k} |f_n(x) - f(x)| dx = \sup_{A \in \mathcal{B}^k} |\nu(A) - \nu'_n(A)|,$$

où, \mathcal{B}^k est la tribue borélienne de \mathbb{R}^k , ν'_n et ν sont les mesures associées à f_n et f respectivement. Dans le second cas, pour toute application bijective $T: \mathbb{R}^k \rightarrow \mathbb{R}^k$, si f, g, f^*, g^* sont les densités des vecteurs aléatoires $X, Y, T(X)$ et $T(Y)$ respectivement, alors

$$\int_{\mathbb{R}^k} |f(x) - g(x)| dx = \int_{\mathbb{R}^k} |f^*(x) - g^*(x)| dx.$$

Nous nous proposons, donc d'étudier dans cette troisième partie, la convergence L^1 p.s. de f_n vers f . Cette étude comporte trois phases. Dans un premier temps, nous allons établir la convergence L^1 p.s. de f_n^D (théorème 2.1), ensuite nous étendrons ce résultat au cas où la matrice H_n est aléatoire (théorème 2.2), c'est-à-dire à l'estimateur $\tilde{f}_n^{DW}(x)$ qui généralise ceux de Deheuvels (1977-a) et Devroye et Wagner (1980). En dernier lieu, en se servant des théorèmes 2.1 et 2.2, nous établirons la convergence L^1 p.s. de f_n (théorème 2.3). Enfin nous appliquerons le théorème 2.3 aux cas particuliers f_n^B et f_n^S . Auparavant, fixons quelques notations:

Nous désignerons par C_0 l'ensemble de toutes les fonctions continues, à support compact dans \mathbb{R}^k et à valeurs réelles. L'opération $E\Delta F$ référera à la différence symétrique entre les deux ensembles E et F . Nous dirons qu'une suite de matrices $k \times k$ $A^{(n)} = (a_{ij}^{(n)})$ converge vers la matrice $k \times k$, $A = (a_{ij})$ si et seulement si $\lim_{n \rightarrow \infty} a_{ij}^{(n)} = a_{ij}$, pour tous $i, j = 1, \dots, k$. Deux matrices $k \times k$, $A = (a_{ij})$ et $B = (b_{ij})$ seront dites $A \leq B$ si et seulement si $a_{ij} \leq b_{ij}$, $i, j = 1, \dots, k$.

Théorème 2.1. *Soit K une fonction absolument intégrable, d'intégrale 1. Si H_n est une matrice $k \times k$ inversible vérifiant $\|H_n\|_\infty \xrightarrow{n \rightarrow \infty} 0$ et $n |\det H_n| \xrightarrow{n \rightarrow \infty} \infty$, alors*

$$\int_{\mathbb{R}^k} |f_n^D(x) - f(x)| dx \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.} \quad (2.1)$$

DÉMONSTRATION: Nous avons

$$\int_{\mathbb{R}^k} |f_n^D(x) - f(x)| dx \leq \int_{\mathbb{R}^k} |f_n^D(x) - \tilde{f}_n^D(x)| dx + \int_{\mathbb{R}^k} |\tilde{f}_n^D(x) - f(x)| dx, \quad (2.2)$$

où

$$\tilde{f}_n^D(x) = \frac{1}{|\det H_n|} \int_{\mathbb{R}^k} K(H_n^{-1}(x-y)) f(y) dy, \quad x \in \mathbb{R}^k;$$

donc, pour obtenir (2.1), il suffit de montrer que les deux termes à droite de (2.2) convergent p.s. vers 0 lorsque $n \rightarrow \infty$. Les deux lemmes suivants nous donnent cela.

Lemme 2.1. *Soit K une fonction absolument intégrable et d'intégrale 1. Si H_n est une matrice $k \times k$ inversible et telle que $\|H_n\|_\infty \xrightarrow{n \rightarrow \infty} 0$, alors*

$$\int_{\mathbb{R}^k} |\tilde{f}_n^D(x) - f(x)| dx \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.} \quad (2.3)$$

DÉMONSTRATION: Dans un premier temps, nous établirons (2.3) pour $f \in C_0$. En effet, soit $w(t) = \sup_{\|x-y\| \leq t} |f(x) - f(y)|$, $t \geq 0$, le module de continuité de f . Nous avons

$$\begin{aligned} \int_{\mathbb{R}^k} |\tilde{f}_n^D(x) - f(x)| dx &\leq \int_{\mathbb{R}^k} dx \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} |K(H_n^{-1}(x-y))| |f(y) - f(x)| dy \\ &\leq \int_{\mathbb{R}^k} dx \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} |K_1(H_n^{-1}(x-y))| |f(y) - f(x)| dy \\ &\quad + \int_{\mathbb{R}^k} dx \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} |K_2(H_n^{-1}(x-y))| \{f(y) + f(x)\} dy, \end{aligned}$$

où $K_1(x) = K(x)1_{S_R(0)}(x)$, $K_2(x) = K(x)1_{S_R^c(0)}(x)$, $R > 0$ une constante arbitraire et $S_R(0)$ la boule de centre 0 et de rayon R . Il est clair que

$$\int_{\mathbb{R}^k} dx \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} |K_2(H_n^{-1}(x-y))| \{f(y) + f(x)\} dy \leq 2 \int_{\mathbb{R}^k} |K_2(x)| dx,$$

et par un choix convenable de R , la quantité $\int_{\mathbb{R}^k} |K_2(x)| dx$ peut être rendue aussi petite que possible. Quant à $\int_{\mathbb{R}^k} dx \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} |K_1(H_n^{-1}(x-y))| |f(y) - f(x)| dy$, elle est égale à

$$\begin{aligned}
& \int_B dx \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} |K_1(H_n^{-1}y)| |f(y) - f(x)| dy \\
& \leq w(R \|H_n\|_\infty) \int_B dx \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} |K_1(H_n^{-1}y)| dy \\
& \leq w(R \|H_n\|_\infty) \lambda^k(B) \int_{\mathbb{R}^k} |K(x)| dx = o(1),
\end{aligned} \tag{2.4}$$

où B est un compact de \mathbb{R}^k suffisamment grand. Ceci conclue la preuve de (2.3) pour $f \in \mathcal{C}_0$. Maintenant, si f est arbitraire, on peut toujours trouver une fonction $f^* \in \mathcal{C}_0$, telle que $\int_{\mathbb{R}^k} |f(x) - f^*(x)| dx$ est aussi petite que l'on veut. Donc, ce fait et (2.4) impliquent que:

$$\begin{aligned}
\int_{\mathbb{R}^k} |\tilde{f}_n^D(x) - f(x)| dx & \leq \int_{\mathbb{R}^k} dx \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} |K(H_n^{-1}(x-y))| |f(y) - f^*(y)| dy \\
& + \int_{\mathbb{R}^k} dx \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} |K(H_n^{-1}(x-y))| |f^*(y) - f^*(x)| dy \\
& + \int_{\mathbb{R}^k} |f^*(x) - f(x)| dx \\
& \leq \left(\int_{\mathbb{R}^k} |K(x)| dx + 1 \right) \int_{\mathbb{R}^k} |f(x) - f^*(x)| dx + o(1) = o(1),
\end{aligned}$$

d'où (2.3).

Lemme 2.2. Soit K une fonction absolument intégrable et d'intégrale 1. Si H_n est une matrice $k \times k$ inversible satisfaisant $\|H_n\|_\infty \xrightarrow{n \rightarrow \infty} 0$ et $n |\det H_n| \xrightarrow{n \rightarrow \infty} \infty$, alors

$$\int_{\mathbb{R}^k} |\tilde{f}_n^D(x) - f_n^D(x)| dx \xrightarrow{n \rightarrow \infty} 0 \quad p.s. \tag{2.5}$$

DÉMONSTRATION: Pour tout $\varepsilon > 0$, il existe des constantes $M_2, \rho, N, \alpha_1, \dots, \alpha_N$ et des pavés disjoints dans \mathbb{R}^k : A_1, \dots, A_N , tels que la fonction

$$K^*(x) = \sum_{i=1}^N \alpha_i 1_{A_i}(x), \quad x \in \mathbb{R}^k,$$

satisfait

$$|K^*(x)| \leq M_2, \quad x \in \mathbb{R}^k, \quad K^*(x) = 0, \quad x \notin [-\rho, \rho]^k,$$

et

$$\int_{\mathbb{R}^k} |K(x) - K^*(x)| dx < \varepsilon.$$

Ainsi

$$\begin{aligned} & \int_{\mathbb{R}^k} |\tilde{f}_n^D(x) - f_n^D(x)| dx \\ & \leq \int_{\mathbb{R}^k} dx \left| \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} \{K(H_n^{-1}(x-y)) - K^*(H_n^{-1}(x-y))\} dF_n(y) \right| \\ & + \int_{\mathbb{R}^k} dx \left| \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} \{K^*(H_n^{-1}(x-y)) - K(H_n^{-1}(x-y))\} dF(y) \right| \\ & + \int_{\mathbb{R}^k} dx \left| \int_{\mathbb{R}^k} \frac{1}{|\det H_n|} K^*(H_n^{-1}(x-y)) \{dF_n(y) - dF(y)\} \right| \\ & \leq 2\varepsilon + \frac{1}{|\det H_n|} \int_{\mathbb{R}^k} dx \left| \sum_{i=1}^N \int_{x+H_n A_i} \{dF_n(y) - dF(y)\} \right| \\ & \leq 2\varepsilon + \frac{1}{|\det H_n|} \sum_{i=1}^N \int_{\mathbb{R}^k} |\nu_n(x + H_n A_i) - \nu(x + H_n A_i)| dx, \end{aligned}$$

où $x + H_n A = \{y : y = x + H_n z, z \in A\}$. Par conséquent, (2.5) suivra une fois que nous aurons montré que pour tout pavé fini A de \mathbb{R}^k

$$\frac{1}{|\det H_n|} \int_{\mathbb{R}^k} |\nu_n(x + H_n A) - \nu(x + H_n A)| dx \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.}$$

En effet, soit l une constante arbitraire et Ψ la partition de \mathbb{R}^k définie par

$$\prod_{i=1}^k \left[\frac{q_i - 1}{l}, \frac{q_i}{l} \right), \quad (q_1, \dots, q_k) \in Z^k,$$

où Z^k est l'ensemble des entiers relatifs.

Choisissons $A = \prod_{i=1}^k [c_i, c_i + d_i)$ avec $\min_{i=1, \dots, k} d_i \geq \frac{2}{l}$ et c_i, d_i deux réels quelconques.

Posons

$$A' = \prod_{i=1}^k \left[c_i + \frac{1}{l}, c_i + d_i - \frac{1}{l} \right),$$

et

$$C_x = x + H_n A - \bigcup_{\substack{B \in \Psi \\ B_n \subset x + H_n A}} B_n,$$

où $B_n = H_n B = \{y \in \mathbb{R}^k : y = H_n z, z \in B\}$ et $B \in \Psi$.

En écrivant, en abrégé, $H_n = (h_{ij})$, on obtient

$$x + H_n A = x + \prod_{i=1}^k \left[\sum_{j=1}^k \min \{h_{ij}c_j, h_{ij}(c_j + d_j)\}, \sum_{j=1}^k \max \{h_{ij}c_j, h_{ij}(c_j + d_j)\} \right],$$

$$x + H_n A' = x + \prod_{i=1}^k \left[\sum_{j=1}^k \min \left\{ h_{ij} \left(c_j + \frac{1}{l} \right), h_{ij} \left(c_j + d_j - \frac{1}{l} \right) \right\}, \right. \\ \left. \sum_{j=1}^k \max \left\{ h_{ij} \left(c_j + \frac{1}{l} \right), h_{ij} \left(c_j + d_j - \frac{1}{l} \right) \right\} \right],$$

et

$$B_n = \prod_{i=1}^k \left[\sum_{j=1}^k \min \left\{ \frac{q_j - 1}{l} h_{ij}, \frac{q_j}{l} h_{ij} \right\}, \sum_{j=1}^k \max \left\{ \frac{q_j - 1}{l} h_{ij}, \frac{q_j}{l} h_{ij} \right\} \right].$$

Par ailleurs, pour tout i fixé dans $\{1, \dots, k\}$ et tout $j = 1, \dots, k$, il existe q_j et q'_j dans Z tels que

$$\min \left(\frac{q_j}{l} h_{ij}, \frac{q_j - 1}{l} h_{ij} \right) \leq \frac{x_i}{k} + \max \left(h_{ij} \left(c_j + \frac{1}{l} \right), h_{ij} \left(c_j + d_j - \frac{1}{l} \right) \right) \\ \leq \max \left(\frac{q_j}{l} h_{ij}, \frac{q_j - 1}{l} h_{ij} \right) \leq \frac{x_i}{k} + \max (h_{ij}c_j, h_{ij}(c_j + d_j))$$

et

$$\frac{x_i}{k} + \min (h_{ij}c_j, h_{ij}(c_j + d_j)) \leq \min \left(\frac{q'_j}{l} h_{ij}, \frac{q'_j - 1}{l} h_{ij} \right) \\ \leq \frac{x_i}{k} + \min \left(h_{ij} \left(c_j + \frac{1}{l} \right), h_{ij} \left(c_j + d_j - \frac{1}{l} \right) \right) \leq \max \left(\frac{q'_j - 1}{l} h_{ij}, \frac{q'_j}{l} h_{ij} \right).$$

D'où

$$x + H_n A' \subset \bigcup_{\substack{B \in \Psi \\ B_n \subset x + H_n A}} B_n$$

et par conséquent,

$$C_x \subset x + H_n(A - A').$$

Cette dernière inclusion nous permet d'avoir

$$\begin{aligned} & \frac{1}{|\det H_n|} \int_{\mathbb{R}^k} |\nu_n(x + H_n A) - \nu(x + H_n A)| dx \\ & \leq \frac{1}{|\det H_n|} \int_{\mathbb{R}^k} \left\{ \nu_n(x + H_n(A - A')) + \nu(x + H_n(A - A')) \right. \\ & \quad \left. + \sum_{\substack{B \in \Psi \\ B_n \subset x + H_n A}} |\nu_n(B_n) - \nu(B_n)| \right\} dx \end{aligned} \quad (2.6)$$

Ensuite, une application du théorème de Fubini nous donne

$$\begin{aligned} \frac{1}{|\det H_n|} \int_{\mathbb{R}^k} \nu(x + H_n(A - A')) dx &= \frac{1}{|\det H_n|} \int_{\mathbb{R}^k} dx \int_{x + H_n(A - A')} f(y) dy \\ &= \frac{1}{|\det H_n|} \int_{\mathbb{R}^k} f(y) \int_{y - H_n(A - A')} dx \\ &= \lambda^k(A - A'), \end{aligned}$$

de même

$$\frac{1}{|\det H_n|} \int_{\mathbb{R}^k} \nu_n(x + H_n(A - A')) dx = \lambda^k(A - A').$$

D'autre part, pour un choix convenable de la constante l , nous obtenons

$$\begin{aligned} \lambda^k(A - A') &= \prod_{i=1}^k d_i - \prod_{i=1}^k \left(d_i - \frac{2}{l}\right) = \lambda^k(A) \left(1 - \prod_{i=1}^k \left(1 - \frac{2}{ld_i}\right)\right) \\ &\leq \lambda^k(A) \sum_{i=1}^k \frac{2}{ld_i} < \varepsilon, \quad \text{pour tout } \varepsilon > 0. \end{aligned}$$

Donc

$$\frac{1}{|\det H_n|} \int_{\mathbb{R}^k} \{ \nu(x + H_n(A - A')) + \nu_n(x + H_n(A - A')) \} dx \leq 2\varepsilon. \quad (2.7)$$

Reste donc à majorer le terme

$$D = \frac{1}{|\det H_n|} \int_{\mathbb{R}^k} \sum_{\substack{B \in \Psi \\ B_n \subset x + H_n A}} |\nu_n(B_n) - \nu(B_n)| dx.$$

En effet, pour tout $R > 0$, nous avons

$$\begin{aligned} D &= \frac{1}{|\det H_n|} \int_{\mathbb{R}^k} \left\{ \sum_{\substack{B \in \Psi \\ B_n \subset x + H_n A \\ B_n \cap S_R(0) \neq \emptyset}} |\nu_n(B_n) - \nu(B_n)| \right. \\ &\quad \left. + \sum_{\substack{B \in \Psi \\ B_n \subset x + H_n A \\ B_n \cap S_R(0) = \emptyset}} |\nu_n(B_n) - \nu(B_n)| \right\} dx \\ &\leq \frac{1}{|\det H_n|} \left\{ \nu_n(S_R^c(0)) - \nu(S_R^c(0)) + 2\nu(S_R^c(0)) \right. \\ &\quad \left. + \sum_{\substack{B \in \Psi \\ B_n \cap S_R(0) \neq \emptyset}} |\nu_n(B_n) - \nu(B_n)| \right\} \int_{x + H_n A} dx. \quad (2.8) \end{aligned}$$

Par ailleurs, $\frac{1}{|\det H_n|} \int_{x + H_n A} dx = \lambda^k(A)$ et pour un choix judicieux de R , nous aurons $\nu(S_R^c(0)) < \varepsilon$. Finalement, en regroupant les inégalités (2.6) à (2.8), nous aboutissons à

$$\begin{aligned} &\frac{1}{|\det H_n|} \int_{\mathbb{R}^k} |\nu_n(x + H_n A) - \nu(x + H_n A)| dx \\ &\leq 2\varepsilon + \lambda^k(A) \left\{ 2\varepsilon + \nu_n(S_R^c(0)) - \nu(S_R^c(0)) \right. \\ &\quad \left. + \sum_{\substack{B \in \Psi \\ B_n \cap S_R(0) \neq \emptyset}} |\nu_n(B_n) - \nu(B_n)| \right\}. \quad (2.9) \end{aligned}$$

Ensuite, l'inégalité de Hoeffding (cf. Hoeffding (1968)), appliquée au terme $(\nu_n(S_R^c(0)) - \nu(S_R^c(0)))$, nous donne

$$\forall \varepsilon > 0, \quad P \{ \nu_n(S_R^c(0)) - \nu(S_R^c(0)) > \varepsilon \} \leq \exp(-2n\varepsilon^2). \quad (2.10)$$

Quant au terme $\sum_{\substack{B \in \Psi \\ B_n \cap S_R(0) \neq \emptyset}} |\nu_n(B_n) - \nu(B_n)|$, remarquons que du fait que Ψ est une partition de \mathbb{R}^k et que H_n définit un automorphisme sur \mathbb{R}^k , alors la famille des $B \in \Psi$, tels que $B_n \cap S_R(0) \neq \emptyset$, a au plus $(\prod_{i=1}^k (\frac{2Rl}{\sum_{j=1}^k |h_{ij}|} + 2))$ éléments. D'autre part, puisque l'inégalité $|\det H| \leq \prod_{i=1}^k (\sum_{j=1}^k h_{ij}^2)^{\frac{1}{2}}$ est valable pour toute matrice H $k \times k$ (voir Graybill (1983) p. 413), nous avons d'après l'équivalence des normes $\|\cdot\|$ et $\|\cdot\|_E$ sur \mathbb{R}^k

$$|\det H| \leq \prod_{i=1}^k (\sum_{j=1}^k h_{ij}^2)^{\frac{1}{2}} \leq a \prod_{i=1}^k (\sum_{j=1}^k |h_{ij}|),$$

où $\|x\|_E = (\sum_{j=1}^k |x_j|^2)^{\frac{1}{2}}$ et a est une constante ne dépendant que de k . Par conséquent, les hypothèses $\|H_n\|_\infty \xrightarrow{n \rightarrow \infty} 0$ et $n|\det H_n| \xrightarrow{n \rightarrow \infty} \infty$ nous permettent d'avoir

$$\prod_{i=1}^k (\frac{2Rl}{\sum_{j=1}^k |h_{ij}|} + 2) \leq a \frac{(2Rl + 2\|H_n\|_\infty)^k}{|\det H_n|} = o(n).$$

Par ailleurs, un résultat de Devroye et Györfi (1985 page 13) stipule que tout vecteur aléatoire (X_1, \dots, X_k) , suivant une loi multinomiale de paramètres (n, p_1, \dots, p_k) , vérifie

$$P \left(\sum_{i=1}^k |X_i - E(X_i)| \geq n\varepsilon \right) \leq 3 \exp \left(-\frac{n\varepsilon^2}{25} \right),$$

$$\forall \varepsilon \in (0, 1) \quad \text{et } \forall k \in \mathbb{N} : \frac{k}{n} < \frac{\varepsilon^2}{20}.$$

Donc, pour notre cas, on obtient

$$P\left(\sum_{\substack{B \in \Psi \\ B_n \cap S_R(0) \neq \emptyset}} |\nu_n(B_n) - \nu(B_n)| > \varepsilon\right) \leq 3 \exp\left(-\frac{n\varepsilon^2}{25}\right),$$

$$\forall \varepsilon > 0 \text{ et } n \text{ assez grand.} \quad (2.11)$$

La conclusion est immédiate à partir des inégalités (2.9) à (2.11).

REMARQUE 2.1

En fait , nous avons: si $\|H_n\|_\infty \xrightarrow{n \rightarrow \infty} 0$ et $n|\det H_n| \xrightarrow{n \rightarrow \infty} \infty$ alors

$$\forall \varepsilon > 0, \exists r, n_0 > 0: \forall n > n_0$$

$$P\left\{\int_{\mathbb{R}^k} |f_n^D(x) - f(x)| dx > \varepsilon\right\} \leq \exp(-rn\varepsilon^2), \quad (2.12)$$

où n_0 dépend de f et de ε . Remarquons que pour un f fixé, s'il existe deux fonctions $h_0(\varepsilon)$ et $c_0(\varepsilon)$ telles que

$$\frac{c_0(\varepsilon)}{n} < |\det H_n| \leq \|H_n\|_\infty^k < h_0(\varepsilon),$$

alors, l'inégalité (2.12) aura encore lieu.

Comme nous l'avons signalé au début, la seconde étape de cette étude consiste à établir la convergence L^1 p.s. de $\tilde{f}_n^{DW}(x)$ (i.e. l'estimateur de Deheuvels $f_n^D(x)$ correspondant à une matrice H_n aléatoire); pour se faire, nous aurons besoin des lemmes suivants.

Lemme 2.3. Soit K une densité de probabilité et (U_n) une suite de matrices $k \times k$ inversibles. Nous avons

$$\lim_{U_n \rightarrow I} \int_{\mathbb{R}^k} \left| \frac{1}{|\det U_n|} K(U_n^{-1}x) - K(x) \right| dx = 0 \quad (2.13)$$

DÉMONSTRATION: Si K est continue, alors

$$\lim_{U_n \rightarrow I} \left(\frac{1}{|\det U_n|} K(U_n^{-1}x) \right) = K(x), \quad \text{pour tout } x \in \mathbb{R}^k.$$

Donc, ayant

$$\int_{\mathbb{R}^k} \left| \frac{1}{|\det U_n|} K(U_n^{-1}x) - K(x) \right| dx = 2 \int_{\Gamma_n} \left\{ \frac{1}{|\det U_n|} K(U_n^{-1}x) - K(x) \right\} dx,$$

où $\Gamma_n = \{x \mid \frac{1}{|\det U_n|} K(U_n^{-1}x) - K(x) \geq 0\}$, le théorème de la convergence dominée permet d'obtenir (2.13). Par contre, si K est une densité de probabilité quelconque, alors, pour tout $\varepsilon > 0$, on peut exhiber une densité de probabilité continue K^* , telle que

$$\int_{\mathbb{R}^k} |K(x) - K^*(x)| dx \leq \varepsilon.$$

D'où, du fait que

$$\begin{aligned} & \int_{\mathbb{R}^k} \left| \frac{1}{|\det U_n|} K(U_n^{-1}x) - K(x) \right| dx \\ & \leq \int_{\mathbb{R}^k} \left| \frac{1}{|\det U_n|} K(U_n^{-1}x) - \frac{1}{|\det U_n|} K^*(U_n^{-1}x) \right| dx \\ & \quad + \int_{\mathbb{R}^k} \left| \frac{1}{|\det U_n|} K^*(U_n^{-1}x) - K^*(x) \right| dx + \int_{\mathbb{R}^k} |K^*(x) - K(x)| dx, \end{aligned}$$

nous aurons

$$\begin{aligned}
\limsup_{U_n \rightarrow I} \int_{\mathbb{R}^k} \left| \frac{1}{|\det U_n|} K(U_n^{-1}x) - K(x) \right| dx \\
\leq 2\varepsilon + \lim_{U_n \rightarrow I} \int_{\mathbb{R}^k} \left| \frac{1}{|\det U_n|} K^*(U_n^{-1}x) - K^*(x) \right| dx \\
= 2\varepsilon;
\end{aligned}$$

ce qui donne (2.13).

Pour tout $\delta > 0$, posons

$$\Phi(\delta I) = \sup_{(1-\delta)I < U < (1+\delta)I} \int_{\mathbb{R}^k} \left| \frac{1}{|\det U|} K(U^{-1}x) - K(x) \right| dx,$$

où U est une matrice $k \times k$ inversible et K est une densité de probabilité. Maintenant, considérons l'ensemble \mathcal{M}_n des matrices $k \times k$ inversibles U , telles que

$$\alpha_n^k \leq |\det U| \leq \|U\|_\infty^k \leq \beta_n^k,$$

où (α_n) et (β_n) sont deux suites de réels telles que $n\alpha_n^k \xrightarrow{n \rightarrow \infty} \infty$ et $\beta_n \xrightarrow{n \rightarrow \infty} 0$.

Lemme 2.4. *Soit K une densité de probabilité et $H_n(\omega)$ une matrice $k \times k$ inversible dépendant de l'échantillon $\omega = (X_1, \dots, X_n)$. Alors, pour tout $\varepsilon > 0$, il existe deux réels $r, n_0 > 0$, tels que*

$$P \left(\sup_{H_n(\omega) \in \mathcal{M}_n} \int_{\mathbb{R}^k} |\tilde{f}_n^{DW}(x) - f(x)| dx > \varepsilon \right) \leq n \exp(-rn\varepsilon^2), \quad \text{pour } n \geq n_0.$$

DÉMONSTRATION: Nous avons

$$\lim_{n \rightarrow \infty} \frac{\beta_n^k}{n \alpha_n^k} = 0 \quad \text{et} \quad \frac{\beta_n}{\alpha_n} \geq 1,$$

donc, il existe une suite de réels (θ_n) , telle que $\theta_n \xrightarrow[n \rightarrow \infty]{} 0$ et $\frac{\beta_n}{\alpha_n} = \theta_n n^{\frac{1}{k}}$. Soit δ_n la solution de l'équation $(1 + \delta_n)^n = \theta_n n^{\frac{1}{k}}$, où δ_n satisfait $\delta_n \xrightarrow[n \rightarrow \infty]{} 0$. Ensuite, soit la suite de matrices $k \times k$ inversibles $(U_{ni})_{i=0, \dots, n}$ définie par

$$U_{ni} = (1 + \delta_n)^i \alpha_n I, \quad i = 0, \dots, n.$$

A ce type de matrices, on fait correspondre la suite $(\mathcal{M}_{ni})_{i=0, \dots, n}$ des sous-ensembles de \mathcal{M}_n , tels que

$$\mathcal{M}_{ni} = \{U : \|U_{ni}\|_{\infty}^k \leq |\det U| \leq \|U\|_{\infty}^k \leq \|U_{ni+1}\|_{\infty}^k\}.$$

Pour que la dépendance de $\tilde{f}_n^{DW}(x)$ et de $H_n(\omega)$ soit explicite, nous écrirons $f_{nH_n(\omega)}(x)$ au lieu de $\tilde{f}_n^{DW}(x)$. A partir de ces considérations, nous obtenons

$$\begin{aligned} & \sup_{H_n(\omega) \in \mathcal{M}_n} \int_{\mathbb{R}^k} |f_{nH_n(\omega)}(x) - f(x)| dx \\ & \leq \sup_{j=0, \dots, n-1} \sup_{H_n(\omega) \in \mathcal{M}_{nj}} \int_{\mathbb{R}^k} |f_{nH_n(\omega)}(x) - f(x)| dx \\ & \leq \sup_{j=0, \dots, n-1} \left\{ \int_{\mathbb{R}^k} |f_{nU_{nj}}(x) - f(x)| dx \right. \\ & \quad \left. + \sup_{H_n(\omega) \in \mathcal{M}_{nj}} \int_{\mathbb{R}^k} |f_{nH_n(\omega)}(x) - f_{nU_{nj}}(x)| dx \right\}, \end{aligned} \quad (2.14)$$

or

$$\begin{aligned}
& \sup_{H_n(\omega) \in \mathcal{M}_{nj}} \int_{\mathbb{R}^k} |f_{nH_n(\omega)}(x) - f_{nU_{nj}}(x)| dx \\
& \leq \sup_{H_n(\omega) \in \mathcal{M}_{nj}} \int_{\mathbb{R}^k} \left| \frac{1}{|\det(H_n^{-1}(\omega)U_{nj})|} K(U_{nj}^{-1}H_n(\omega)x) - K(x) \right| dx \\
& \leq \sup_{H_n(\omega) \in \mathcal{M}_{nj}} \sup_{2I - U_{nj}^{-1}H_n(\omega) \leq U \leq U_{nj}^{-1}H_n(\omega)} \int_{\mathbb{R}^k} \left| \frac{1}{|\det U|} K(U^{-1}x) - K(x) \right| dx \\
& = \Phi(\delta_n I). \tag{2.15}
\end{aligned}$$

Une conséquence du lemme 2.3 est que pour tout $\varepsilon > 0$, il existe un entier $n_1 > 0$ tel que pour $n > n_1$, $\Phi(\delta_n I) < \varepsilon$. Donc, (2.14) et (2.15) entraînent que

$$P\left\{ \sup_{H_n(\omega) \in \mathcal{M}_n} \int_{\mathbb{R}^k} |f_{nH_n(\omega)}(x) - f(x)| dx > 2\varepsilon \right\} \leq \sum_{j=1}^n P\left\{ \int_{\mathbb{R}^k} |f_{nU_{nj}}(x) - f(x)| dx > \varepsilon \right\}.$$

Considérons maintenant, $h_0(\varepsilon)$ et $c_0(\varepsilon)$ tels que

$$\left(\frac{c_0(\varepsilon)}{n}\right)^k \leq |\det U_{nj}| = \|U_{nj}\|_{\infty}^k = \left((1 + \delta_n)^j \alpha_n\right)^k \leq h_0(\varepsilon), \quad j = 1, \dots, n.$$

Dans ce cas, la remarque 2.1 permet de conclure que pour n assez grand

$$\sum_{j=1}^n P\left\{ \int_{\mathbb{R}^k} |f_{nU_{nj}}(x) - f(x)| dx > \varepsilon \right\} \leq \exp(-rn\varepsilon^2).$$

Nous sommes maintenant en mesure d'établir la convergence L^1 p.s. de $\tilde{f}_n^{DW}(x)$.

Théorème 2.2. Soit f et K deux densités de probabilité. Si $H_n(\omega)$ est une matrice $k \times k$ inversible et telle que $\|H_n(\omega)\|_\infty \xrightarrow[n \rightarrow \infty]{} 0$ p.s. et $n|\det H_n(\omega)| \xrightarrow[n \rightarrow \infty]{} \infty$ p.s. Alors

$$\int_{\mathbb{R}^k} |\tilde{f}_n^{DW}(x) - f(x)| dx \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{p.s.}$$

DÉMONSTRATION: Etant donné que $\|H_n(\cdot)\|_\infty \xrightarrow[n \rightarrow \infty]{} 0$ et $n|\det H_n(\cdot)| \xrightarrow[n \rightarrow \infty]{} \infty$ p.s., on peut considérer les 2 suites d'entiers (n'_j) et (n''_j) définies par

$$\begin{aligned} n'_1 &= n''_1, \\ n'_j &= \inf \left\{ n : n > n'_{j-1}, \Pr \left(\bigcup_{m \geq n} \left\{ \|H_m(\cdot)\|_\infty > \frac{1}{j} \right\} \right) \leq 2^{-j} \right\}, \quad j \geq 2, \\ n''_j &= \inf \left\{ n : n > n''_{j-1}, \Pr \left(\bigcup_{m \geq n} \{m|\det H_m(\cdot)| \leq j\} \right) \leq 2^{-j} \right\}, \quad j \geq 2. \end{aligned}$$

A ces deux suites (n'_j) , (n''_j) on associe les deux suites de réels positifs

$$\beta_n = \frac{1}{j} \quad \text{sur } [n'_j, n'_{j+1}), \quad j \geq 2,$$

$$\alpha_n = \left(\frac{j}{n} \right)^{1/k} \quad \text{sur } [n''_j, n''_{j+1}), \quad j \geq 2.$$

Les deux suites (α_n) et (β_n) ainsi définies vérifient $n\alpha_n^k \xrightarrow[n \rightarrow \infty]{} \infty$ et $\beta_n \xrightarrow[n \rightarrow \infty]{} 0$, et de plus

$$\alpha_n^k \leq |\det H_n(\cdot)| \leq \|H_n(\cdot)\|_\infty^k \leq \beta_n^k \quad \text{p.s.,}$$

car sur $[n'_j, n'_{j+1})$

$$\begin{aligned} \Pr\left(\bigcup_{m \geq n} \left\{ \|H_m(\cdot)\|_\infty^k \geq \alpha_m^k \right\}\right) &\leq \sum_{i=j}^{\infty} \Pr\left(\bigcup_{n'_i \leq m < n'_{i+1}} \left\{ \|H_m(\cdot)\|_\infty \geq \frac{1}{i} \right\}\right) \\ &\leq \sum_{i=j}^{\infty} 2^{-i} = 2^{-j+1} \xrightarrow{j \rightarrow \infty} 0. \end{aligned}$$

Par un raisonnement analogue on a

$$|\det H_n(\cdot)| \geq \alpha_n^k \quad \text{p.s.}$$

Pour conclure reprenant l'ensemble des matrices \mathcal{M}_n (défini au lemme 2.4) associé à α_n et β_n définies ci-dessus et notons

$$A_1 = \left\{ \omega : \int_{\mathbb{R}^k} |\tilde{f}_n^{DW}(x) - f(x)| dx > \varepsilon \right\}$$

$$A_2 = \left\{ \omega : H_n(\omega) \notin \mathcal{M}_n \right\}$$

$$A_3 = \left\{ \omega : \sup_{H_n(\omega) \in \mathcal{M}_n} \int_{\mathbb{R}^k} |\tilde{f}_n^{DW}(x) - f(x)| dx > \varepsilon \right\}.$$

D'après ce qui précède $1_{A_2}(\omega) \xrightarrow{n \rightarrow \infty} 0$ p.s., et d'après le lemme 2.4 $1_{A_3}(\omega) \xrightarrow{n \rightarrow \infty} \infty$ p.s., par conséquent le théorème 2.2 découlera de l'inégalité

$$\forall \varepsilon > 0, \quad 1_{A_1} \leq 1_{A_2} + 1_{A_3}.$$

Enfin, nous en arrivons à la démonstration, proprement dite, de la convergence L^1 p.s. de f_n .

Théorème 2.3. Soit f et K deux densités de probabilité. Si $\|H_n(\cdot)\|_\infty \xrightarrow[n \rightarrow \infty]{} 0$ p.s. et $n|\det H_n(\cdot)| \xrightarrow[n \rightarrow \infty]{} \infty$ p.s., et $\phi_n(\cdot, \cdot)$ est telle que: pour tout i fixé dans $\{1, \dots, n\}$, $\|H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i))\| \xrightarrow[n \rightarrow \infty]{} 0$ p.s., alors

$$\int_{\mathbb{R}^k} |f_n(x) - f(x)| dx \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{p.s.} \quad (2.16)$$

DÉMONSTRATION: Nous avons

$$\int_{\mathbb{R}^k} |f_n(x) - f(x)| dx \leq \int_{\mathbb{R}^k} |f_n(x) - \tilde{f}_n^{DW}(x)| dx + \int_{\mathbb{R}^k} |\tilde{f}_n^{DW}(x) - f(x)| dx.$$

D'après le théorème 2.2, le terme $\int_{\mathbb{R}^k} |\tilde{f}_n^{DW}(x) - f(x)| dx \xrightarrow[n \rightarrow \infty]{} 0$ p.s. Il suffit donc de montrer que $\int_{\mathbb{R}^k} |f_n(x) - \tilde{f}_n^{DW}(x)| dx \xrightarrow[n \rightarrow \infty]{} 0$ p.s. En effet,

$$\int_{\mathbb{R}^k} |f_n(x) - \tilde{f}_n^{DW}(x)| dx =$$

$$\int_{\mathbb{R}^k} \left| K(z) - \frac{1}{n} \sum_{i=1}^n K(z + H_n^{-1}(\omega)(\phi_n(\omega, X_i) - X_i)) \right| dz. \quad (2.17)$$

Or, en reprenant la fonction K^* définie au lemme 2.2, nous obtenons à l'aide de (2.17), les inégalités suivantes

$$\int_{\mathbb{R}^k} |f_n(x) - \tilde{f}_n^{DW}(x)| dx$$

$$\leq 2\varepsilon + \int_{\mathbb{R}^k} \left| K^*(z) - \frac{1}{n} \sum_{i=1}^n K^*(z + H_n^{-1}(\omega)(\phi_n(\omega, X_i) - X_i)) \right| dz$$

$$\leq 2\varepsilon + M_2 \sum_{j=1}^N \int_{\mathbb{R}^k} \left| 1_{A_j}(z) - \frac{1}{n} \sum_{i=1}^n 1_{A_j}(z + H_n^{-1}(\omega)(\phi_n(\omega, X_i) - X_i)) \right| dz$$

$$\leq 2\varepsilon + M_2 \sum_{j=1}^N \left\{ \frac{1}{n} \sum_{i=1}^n \lambda^k(A_j \Delta(A_j + H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i)))) \right\}.$$

Le théorème 2.3 sera établi une fois que l'on aura montré que pour tout pavé fini

$$A = \prod_{i=1}^k [a_i, b_i) \text{ dans } \mathbb{R}^k$$

$$\frac{1}{n} \sum_{i=1}^n \lambda^k \{ A \Delta (A + H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i))) \} \xrightarrow{n \rightarrow \infty} 0 \quad \text{p.s.}$$

Pour cela, notons

$$I_1 = \{i \in \{1, \dots, n\}, \exists j \in \{1, \dots, k\} : |(H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i)))_j| \geq (b_j - a_j)\},$$

$$I_2 = \{i \in \{1, \dots, n\}, \forall j \in \{1, \dots, k\} : |(H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i)))_j| < (b_j - a_j)\},$$

dans ce cas,

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \lambda^k (A \Delta (A + H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i)))) \\ &= \frac{2}{n} \sum_{i=1}^n \left[\prod_{j=1}^k (b_j - a_j) - \prod_{j=1}^k \left(b_j - a_j - |(H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i)))_j| \right) \right] 1_{I_2}(i) \\ &+ 2\lambda^k(A) \frac{1}{n} \sum_{i=1}^n 1_{I_1}(i) \end{aligned} \quad (2.18)$$

et

$$\begin{aligned} & \prod_{j=1}^k (b_j - a_j) - \prod_{j=1}^k \left(b_j - a_j - |(H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i)))_j| \right) \\ &= \lambda^k(A) \left[1 - \prod_{j=1}^k \left(1 - \frac{|(H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i)))_j|}{b_j - a_j} \right) \right] \\ &\leq \lambda^k(A) \sum_{j=1}^k \frac{|(H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i)))_j|}{b_j - a_j} \\ &\leq \eta_A \|H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i))\|, \end{aligned} \quad (2.19)$$

où η_A est une constante dépendant de A . En injectant (2.19) dans (2.18), on aboutit à

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \lambda^k (A \Delta (A + H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i))) \\ & \leq 2\lambda^k(A) \frac{1}{n} \sum_{i=1}^n 1_{I_1}(i) + 2\eta_A \frac{1}{n} \sum_{i=1}^n \|H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i))\| 1_{I_2}(i), \end{aligned}$$

pour conclure, remarquons que pour n assez grand, I_1 est p.s. vide, car par hypothèse, pour tout i fixé dans $\{1, \dots, n\}$, $\|H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i))\| \xrightarrow[n \rightarrow \infty]{} 0$ p.s., tandis que

$$\frac{1}{n} \sum_{i=1}^n \|H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i))\| 1_{I_2}(i) \leq \frac{1}{n} \sum_{i=1}^n \|H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i))\|.$$

Ce dernier terme converge p.s. vers ∞ lorsque $n \rightarrow \infty$, car si une suite à double entrée $(a_{in})_{i=1, \dots, n} \ n > 1$ est telle que $a_{in} \xrightarrow[n \rightarrow \infty]{} 0$ pour tout i fixé, alors $\frac{1}{n} \sum_{i=1}^n a_{in} \xrightarrow[n \rightarrow \infty]{} 0$.

II.2 Applications

Pour établir la convergence L^1 p.s. de f_n vers f , nous avons dû la démontrer pour les estimateurs $f_n^D(x)$ et $\tilde{f}_n^{DW}(x)$. Donc, par la même occasion, nous avons obtenu la convergence L^1 p.s. des estimateurs de Parzen-Rosenblatt, de Deheuvels (1977-a) et de Devroye et Wagner (1980). Il nous reste à appliquer le théorème 2.3. aux estimateurs f_n^B et f_n^S .

II.2.1 L'estimateur f_n^B

Lorsque $\mu_n = \bar{X}_n$ et $V_n = S_n^2$, nous obtenons pour f_n^B le résultat suivant:

Théorème 2.4. *Soit f une densité de probabilité de moyenne μ et de matrice de variance-covariance V . Si K est une densité de probabilité et (γ_n) est une suite dans $(0, 1)$ et telle que $\gamma_n \xrightarrow[n \rightarrow \infty]{} 0$ et $n\gamma_n^k \xrightarrow[n \rightarrow \infty]{} \infty$, alors*

$$\int_{\mathbb{R}^k} |f_n^B(x) - f(x)| dx \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{p.s.}$$

DÉMONSTRATION: Puisque f_n^B est un cas particulier de f_n , il suffit de montrer que les conditions du théorème 2.3 sont satisfaites. En effet, on a

$$\phi_n(\omega, X_i) = \sqrt{1 - \gamma_n^2}(X_i - \bar{X}_n) + \bar{X}_n, \quad \text{et} \quad H_n(\omega) = \gamma_n S_n,$$

d'où $\|H_n(\omega)\|_\infty = \gamma_n \|S_n\|_\infty \xrightarrow[n \rightarrow \infty]{} 0$ p.s. car $\gamma_n \xrightarrow[n \rightarrow \infty]{} 0$ et $\|S_n\|_\infty \xrightarrow[n \rightarrow \infty]{} \|V\|_\infty$ p.s. De même $n|\det H_n(\omega)| = n\gamma_n^k |\det S_n| \xrightarrow[n \rightarrow \infty]{} \infty$ p.s. Par ailleurs, on a

$$\begin{aligned} \|H_n^{-1}(\omega)(X_i - \phi_n(\omega, X_i))\| &= \frac{1}{\gamma_n} \left\| S_n^{-1}(1 - \sqrt{1 - \gamma_n^2})(X_i - \bar{X}_n) \right\| \\ &\leq \frac{1 - \sqrt{1 - \gamma_n^2}}{\gamma_n} \|S_n^{-1}(X_i - \bar{X}_n)\|, \end{aligned}$$

et la loi forte des grands nombres combinée avec l'hypothèse $\gamma_n \xrightarrow[n \rightarrow \infty]{} 0$, permettent de conclure.

II.2.2 L'estimateur f_n^S

Théorème 2.5. Soit f et K deux densités de probabilité. Si (θ_n) est une suite de réels dans $(1, \infty)$ satisfaisant $\theta_n \xrightarrow[n \rightarrow \infty]{} 1$ et $n(\theta_n - 1)^{\frac{k}{2}} \xrightarrow[n \rightarrow \infty]{} \infty$, alors

$$\int_{\mathbb{R}^k} |f_n^S(x) - f(x)| dx \xrightarrow[n \rightarrow \infty]{} 0 \quad p.s.$$

DÉMONSTRATION: Sous les hypothèses du théorème 2.5 les conditions du théorème 2.3 sont automatiquement vérifiées par f_n^S .

CHAPITRE III

SIMULATIONS ET RÉSULTATS NUMÉRIQUES

Nous nous proposons, dans cette dernière partie, de comparer les performances des estimateurs univariés f_n^{PR} et f_n^B . Nous considérerons simultanément quatre formes de f_n^B (voir (4) page 9). Elles correspondent à $\alpha_n = 1$ et $\mu_n = \bar{X}_n$ ou $\bar{X}_{n\alpha}$ ou \tilde{X}_n ou X_n^* . Pour des commodités de calcul le noyau utilisé est le noyau normal standard. Les choix des pas de lissage h_n et γ_n seront effectués selon les critères $L_1(\cdot)$, $L_2(\cdot)$ et le critère asymptotique de Deheuvels (1977-b) (voir pages 3-7). Plus précisément, dans le cas de f_n^{PR} le pas de lissage de Deheuvels est donné par

$$h_n^D = \frac{1.059}{n^{1/5}} S_n.$$

A cause de la complexité du $MISE$ associé à l'estimateur f_n^B , nous ne pouvons pas procéder de la même manière que Deheuvels pour aboutir à une fenêtre γ_n asymptotiquement optimale. Néanmoins, la ressemblance entre h_n^D et $\gamma_n S_n$ incite à prendre le pas de lissage dans f_n^B comme étant

$$\gamma_n = \gamma_n^D = \frac{1.059}{n^{1/5}}, \quad \text{avec } n \geq 2.$$

Nous référerons à la méthode de Deheuvels par *DEH*. Quant aux deux critères $L_1(\cdot)$ et $L_2(\cdot)$, il suffit de remplacer $f_n(\cdot, D, \tau)$ par f_n^{PR} ou f_n^B , où dans un cas τ est remplacé par h et dans l'autre par γ . Pour f_n^{PR} , le critère $L_1(\cdot)$ consiste à minimiser en h la quantité

$$L_1^{PR}(h) = \sum_{i=1}^n \log \left\{ \sum_{\substack{j=1 \\ j \neq i}}^n \exp \left(-\frac{(X_i - X_j)^2}{2h^2} \right) \right\} - n \log h.$$

Tandis que pour f_n^B , il suffit de minimiser en γ

$$L_1^B(\gamma) = \sum_{i=1}^n \log \left\{ \sum_{\substack{j=1 \\ j \neq i}}^n \frac{1}{S_{ni}} \exp \left(-\frac{(X_i - \sqrt{1-\gamma^2}(X_j - \bar{X}_{ni}) - \mu_{ni})^2}{2\gamma^2 S_{ni}^2} \right) \right\} - n \log \gamma,$$

où S_{ni} , \bar{X}_{ni} et μ_{ni} correspondent aux estimateurs S_n , \bar{X}_n , μ_n construits à partir de l'échantillon $D_i = \{X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n\}$, $i = 1, \dots, n$. Les fenêtres fournies par le critère $L_1(\cdot)$ seront désignés par *VCKL*. Enfin, pour le critère $L_2(\cdot)$, lorsqu'il s'agit de l'estimateur f_n^{PR} , il suffit de minimiser en h

$$L_2^{PR}(h) = \frac{1}{h} \left\{ n + \sum_{\substack{i=1, j=1 \\ i \neq j}}^n \left[\frac{n-2}{n-1} \exp \left(-\frac{(X_i - X_j)^2}{4h^2} \right) - 2\sqrt{2} \exp \left(-\frac{(X_i - X_j)^2}{2h^2} \right) \right] \right\}.$$

De même pour f_n^B , il suffit de minimiser en γ

$$L_2^B(\gamma) = \frac{1}{\gamma} \left\{ \sum_{i=1}^n \frac{1}{S_{ni}} + \frac{1}{n-1} \sum_{\substack{i=1, j=1, k=1 \\ i \neq j \neq k}}^n \frac{1}{S_{ni}} \exp \left(-\frac{(1-\gamma^2)(X_j - X_k)^2}{4\gamma^2 S_{ni}^2} \right) \right. \\ \left. - 2\sqrt{2} \sum_{\substack{i=1, j=1 \\ i \neq j}}^n \frac{1}{S_{ni}} \exp \left(-\frac{(X_i - \sqrt{1-\gamma^2}(X_j - \bar{X}_{ni}) - \mu_{ni})^2}{2\gamma^2 S_{ni}^2} \right) \right\}.$$

Nous référerons aux fenêtres associées au critère $L_2(\cdot)$ par *VCRQ*.

Afin de comparer les performances de f_n^{PR} et des différents estimateurs f_n^B , nous avons simulé quelques lois connues (voir figure 1.). La première loi considérée est la loi normale standard $N(0, 1)$. Cette loi a été choisie pour ses nombreuses applications, ses propriétés de symétrie et l'existence de ses moments. Ensuite, nous avons la loi du chi-deux à 4 degrés de liberté $\chi_2(4)$. Cette loi a la particularité d'être asymétrique à droite et naturellement d'avoir des moments finis. Pour la troisième loi nous avons choisi la loi de Student à 5 degrés de liberté $t(5)$, pour sa queue lourde et l'existence de ses moments. Vient ensuite le mélange de normales $\frac{1}{2}N(-1.5, 1) + \frac{1}{2}N(1.5, 1)$, qui fournit une densité bimodale de moments finis. Enfin comme dernière loi, nous avons choisi une densité symétrique et de queue plus légère que celle de la loi normale. Elle appartient à la famille de lois indexées par le paramètre α :

$$f_\alpha(x) = \frac{1}{2\Gamma(\frac{\alpha+1}{\alpha})} \exp(-|x|^\alpha).$$

A partir de chaque loi, nous avons tiré au hasard 100 échantillons de taille 25, 50 échantillons de taille 50 et 25 échantillons de taille 100. Le temps de calcul assez grand nous a obligé à réduire le nombre d'échantillons lorsque le taille augmente. Les simulations de ces différentes lois ont été effectuées de la façon suivante:

1. Générer des variables aléatoires U_1, U_2, \dots de loi uniforme $\mathcal{U}_{(0, 1)}$, en utilisant la méthode de congruence multiplicative: $U_{n+1} = a U_n \pmod{2^m}$, où $m = 36$,

$a = 64155708247$ et $U_1 = a$ (voir Lebart, Morineau et Tabard (1977) page 245).

2. Pour chacune des lois indiquées, utiliser les transformations :

a) si U_1 et U_2 sont deux v. a. indépendantes de loi $\mathcal{U}_{(0,1)}$, alors, d'après Box et Muller, en posant

$$\begin{cases} X = \mu + \sigma\sqrt{-2 \log U_1} \cos 2\pi U_2 \\ Y = \mu + \sigma\sqrt{-2 \log U_1} \sin 2\pi U_2, \end{cases}$$

on obtient deux v. a. indépendantes de même loi $N(\mu, \sigma^2)$.

b) Si X et X_1, \dots, X_ν sont des v. a. indépendantes de même loi $N(\mu, \sigma^2)$, alors les v. a. Y et Z sont de loi respective χ_ν^2 et $t(\nu)$, où

$$\begin{cases} Y = \sum_{i=1}^{\nu} \frac{(X_i - \mu)^2}{\sigma^2} \\ Z = \frac{((X - \mu)/\sigma)^2}{\sqrt{Y/\nu}}. \end{cases}$$

c) Si X et Y sont deux v. a. indépendantes, de loi respective $N(\mu_1, \sigma_1^2)$ et $N(\mu_2, \sigma_2^2)$ et ε est un réel fixé dans $(0, 1)$, alors la v. a.

$$Z = \begin{cases} X & \text{avec probabilité} = \varepsilon \\ Y & \text{avec probabilité} = 1 - \varepsilon, \end{cases}$$

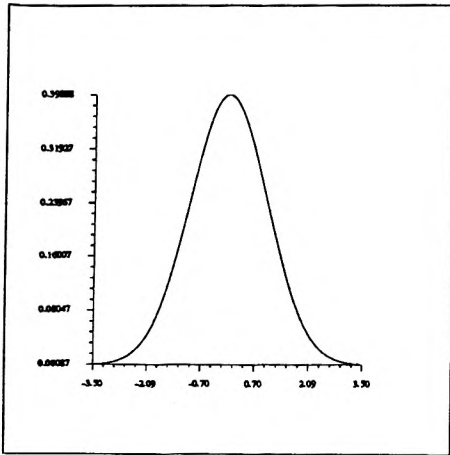
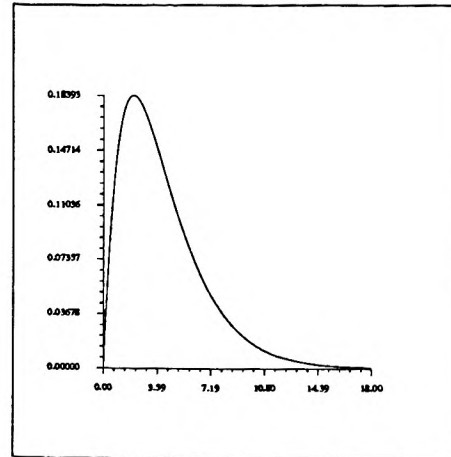
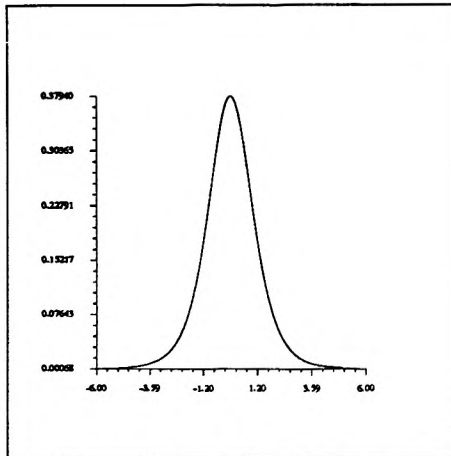
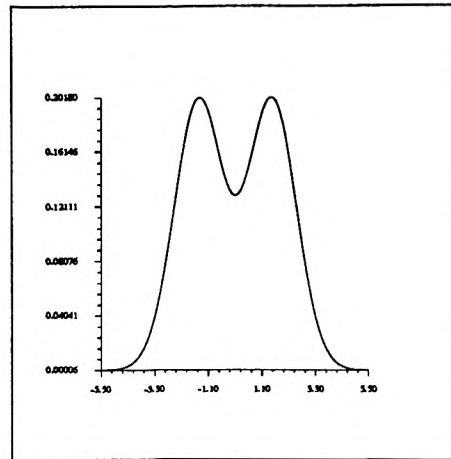
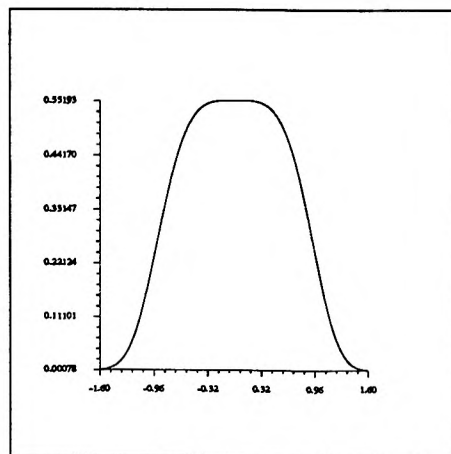
est de loi $\varepsilon N(\mu_1, \sigma_1^2) + (1 - \varepsilon)N(\mu_2, \sigma_2^2)$.

d) Enfin, on montre que si X est une v. a. de loi $\Gamma(1/\alpha, 1)$, alors la v. a. définie par

$$Y = \begin{cases} X^{\frac{1}{\alpha}} & \text{avec probabilité} = 1/2 \\ -X^{\frac{1}{\alpha}} & \text{avec probabilité} = 1/2, \end{cases}$$

est de loi $f_\alpha(\cdot)$. (Pour générer une $\Gamma(1/\alpha, 1)$, avec $\alpha > 1$, nous avons utilisé un algorithme de Ahrens & Dieter (1974) page 228).

Figure 1.

Loi Normale $N(0, 1)$ Loi du Khi-deux $\chi^2_{(4)}$ Loi de Student $t(5)$ Loi Binormale $\frac{1}{2}N(-1.5, 1) + \frac{1}{2}N(1.5, 1)$ Loi $f_4(\cdot)$

L'écart entre les différents estimateurs et f a été mesuré par la fonction de perte quadratique intégrée (ISE)

$$ISE(p, q) = \int_{\mathbb{R}} (p(x) - q(x))^2 dx.$$

Cette intégrale a été approximée par la méthode de Simpson composée i.e.

$$\int_a^b g(t) dt = \frac{b-a}{6N} \sum_{i=1}^N \left[g(a_{i-1}) + 4g\left(\frac{a_i + a_{i-1}}{2}\right) + g(a_i) \right],$$

où $a_i = a + i\left(\frac{b-a}{N}\right)$, $i = 0, \dots, N$. Le paramètre N a été choisi de telle façon que le pas d'intégration $\frac{b-a}{N}$ soit égal à 0.02. L'évaluation des ISE a été effectuée sur les intervalles: $[-3.5, 3.5]$ pour la normale $N(0, 1)$, $[0., 18.]$ pour la chi-deux $\chi_2(4)$, $[-6., 6.]$ pour la student $t(5)$, $[-5.5, 5.5]$ pour la binormale $\frac{1}{2}N(-1.5, 1.) + \frac{1}{2}N(1.5, 1.)$ et enfin $[-1.6, 1.6]$ pour la loi $f_4(\cdot)$.

A cause de l'irrégularité des fonctions $L_1^{PR}(\cdot)$, $L_1^B(\cdot)$, $L_2^{PR}(\cdot)$ et $L_2^B(\cdot)$, la recherche des fenêtres $VCKL$ et $VCRQ$ a été faite par discrétisation avec un pas de 0.01. L'intervalle considéré pour la fenêtre h_n est $[0.01, 2]$ celui de γ_n est $[0.01, 0.99]$. Les résultats des différentes simulations sont présentés dans les tables 1.1 à 5.2. Les tables 1.1, 2.1, 3.1, 4.1 et 5.1 contiennent la moyenne et l'écart-type (entre parenthèses) des ISE associés à chacune des tailles d'échantillons $n_1 = 25$, $n_1 = 50$ et $n_1 = 100$. Quant aux tables 1.2, 2.2, 3.2, 4.2 et 5.2, elles concernent la moyenne et l'écart-type des pas de lissage utilisés.

Notre but n'est pas de comparer les performances des procédés DEH , $VCKL$ et $VCRQ$, cependant on peut signaler une conclusion importante: au vu des résultats de nos simulations et de par sa facilité de mise en œuvre, la méthode de Deheuvels est la plus efficace. Des études comparatives de procédures permettant de choisir la fenêtre de lissage dans f_n^{PR} peuvent être trouvées dans Bean et Tsokos (1982), Bowman (1985), Duin (1976), Kappenman (1987) et Marron (1985). Il en ressort qu'il n'existe

pas de méthode qui soit toujours meilleure que les autres. Cependant, il semble que le critère asymptotique (basé sur la normalité) est très efficace lorsqu'il s'agit de densités unimodales, et que globalement, les choix $VCKL$ et $VCRQ$ sont efficaces (au sens du risque quadratique) pour des petites, respectivement grandes tailles d'échantillons.

Afin de comparer les performances des différents estimateurs, nous nous baserons sur les moyennes des ISE . Les écart-types indiquent que, contrairement aux pas de lissage $VCRQ$ et $VCKL$, les ISE associés aux trois méthodes DEH , $VCRQ$ et $VCKL$ varient peu. Ensuite, un examen cas par cas, des différentes lois étudiées, permet de conclure que pour la loi normale $N(0, 1)$ (table 1.1), l'estimateur le plus performant est f_n^B avec $\mu_n = \bar{X}_n$. Pour $n_1 = 25$ ou 50 la méthode DEH est meilleure que les deux autres, tandis que pour $n_1 = 100$, la méthode $VCRQ$ semble dominer. Par ailleurs, dans le cadre des estimateurs f_n^B , les méthodes $VCRQ$ et $VCKL$ fournissent des pas de lissage assez proches de 1. Ceci revient à dire qu'on estime la loi normale $N(0, 1)$ par $N(\mu_n, S_n^2)$, autrement dit, l'estimateur non paramétrique f_n^B tend à devenir un estimateur paramétrique de la normale $N(0, 1)$.

Pour la loi du khi-deux $\chi_2(4)$, on constate que pour les trois méthodes DEH , $VCRQ$, $VCKL$ et les trois tailles d'échantillons $n_1 = 25, 50, 100$, l'estimateur f_n^B , correspondant à $\mu_n = \bar{X}_n$, est l'estimateur le moins efficace. Cependant, malgré que $\bar{X}_{n\alpha}$ et X_n^* ne reflètent pas la moyenne de cette loi, les meilleures estimations sont fournies par f_n^B correspondant à ces deux estimateurs.

En ce qui concerne les lois de Student $t(5)$ (table 3.1) et le mélange de normales $\frac{1}{2}N(-1.5, 1) + \frac{1}{2}N(1.5, 1)$ (table 4.1), l'estimateur f_n^B , correspondant à $\mu_n = \bar{X}_n$ et la méthode DEH , est le plus efficace.

Enfin, pour la densité $f_\alpha(\cdot)$, qui, rappelons-le, admet une queue moins lourde que la loi normale, la table 5.1 donne des résultats qui ne vont pas dans le même sens que ceux des autres lois. Les meilleurs résultats sont fournis par f_n^{PR} avec le pas de lissage DEH . On observe également une mauvaise performance des méthodes $VCKL$ et $VCRQ$.

En conclusion, nos simulations indiquent qu'il n'existe pas d'estimateur parmi f_n^{PR} et

f_n^B qui soit meilleur dans tous les cas. Cependant, on constate qu'à l'exception de la loi $f_\alpha(\cdot)$ et pour chacune des méthodes *DEH*, *VCKL* et *VCRQ*, il existe au moins un μ_n pour lequel l'estimateur f_n^B fournit une moyenne de l'*ISE* inférieure à celle de f_n^{PR} . Cela tend à confirmer les idées de Bierens (1983) et de Hogg et Lenth (1984), selon lesquelles si on adaptait l'estimateur à noyau aux observations en lui imposant d'avoir \bar{X}_n et S_n^2 comme moyenne et variance respectivement, on obtiendrait une amélioration de l'estimation de la densité. Nos simulations montrent qu'on peut mieux faire en choisissant un estimateur μ_n de la moyenne qui ne soit pas nécessairement égal à \bar{X}_n . Le problème qui se pose maintenant est de choisir μ_n . Nous n'y avons pas apporté de solution définitive. Néanmoins, nous préconisons une procédure à deux étapes. La première étape consisterait à estimer la densité f par l'estimateur de Parzen-Rosenblatt f_n^{PR} . Ceci donnerait une idée sur la forme et le type de la loi étudiée et permettrait de choisir le μ_n adéquat parmi \bar{X}_n , $\bar{X}_{n\alpha}$, \tilde{X}_n , X_n^* ou parmi des estimateurs adaptables ou partiellement adaptables comme ceux de Hogg (1967), Prescott (1978), De Wet et Van Wyk (1979) etc. La seconde étape consisterait alors à réestimer la densité en utilisant l'estimateur f_n^B correspondant au μ_n sélectionné lors de la première étape.

TABLE 1.1 moyenne (écart-type) de l'ISE associé à la loi normale: $N(0, 1)$

Pas	Estimateur	Tailles et nombres d'échantillons					
		$n_1=25$	$n_2=100$	$n_1=50$	$n_2=50$	$n_1=100$	$n_2=25$
DEH	f_n^{PR}	0.0160 (0.0146)		0.0095 (0.0070)		0.0063 (0.0035)	
	f_n^B avec $\mu_n = \bar{X}_n$	0.0157 (0.0172)		0.0080 (0.0076)		0.0049 (0.0020)	
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0211 (0.0240)		0.0110 (0.0114)		0.0072 (0.0039)	
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0170 (0.0186)		0.0084 (0.0084)		0.0055 (0.0024)	
	f_n^B avec $\mu_n = X_n^*$	0.0211 (0.0237)		0.0101 (0.0106)		0.0062 (0.0030)	
VCKL	f_n^{PR}	0.0199 (0.0181)		0.0112 (0.0090)		0.0075 (0.0042)	
	f_n^B avec $\mu_n = \bar{X}_n$	0.0168 (0.0194)		0.0082 (0.0090)		0.0044 (0.0031)	
	f_n^B avec $\mu_n = \bar{X}_n$	0.0201 (0.0241)		0.0104 (0.0112)		0.0059 (0.0041)	
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0178 (0.0205)		0.0085 (0.0094)		0.0048 (0.0033)	
	f_n^B avec $\mu_n = X_n^*$	0.0209 (0.0239)		0.0096 (0.0104)		0.0053 (0.0033)	
VCRQ	f_n^{PR}	0.0266 (0.0346)		0.0171 (0.0195)		0.0112 (0.0077)	
	f_n^B avec $\mu_n = \bar{X}_n$	0.0229 (0.0351)		0.0103 (0.0148)		0.0028 (0.0026)	
	f_n^B avec $\mu_n = \bar{X}_n$	0.0223 (0.0309)		0.0088 (0.0081)		0.0039 (0.0034)	
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0204 (0.0328)		0.0105 (0.0146)		0.0031 (0.0029)	
	f_n^B avec $\mu_n = X_n^*$	0.0232 (0.0313)		0.0102 (0.0118)		0.0037 (0.0034)	

TABLE 1.2 moyenne (écart-type) du pas associé à la loi normale: $N(0, 1)$

Estimateur	Pas	Tailles et nombres d'échantillons					
		$n_1=25$	$n_2=100$	$n_1=50$	$n_2=50$	$n_1=100$	$n_2=25$
f_n^{PR}	DEH	0.5359 (0.0868)		0.4734 (0.0494)		0.4208 (0.0298)	
	VCKL	0.6101 (0.1778)		0.5025 (0.1280)		0.4082 (0.0971)	
	VCRQ	0.6145 (0.2182)		0.4917 (0.1873)		0.4917 (0.1873)	
f_n^B avec $\mu_n = \bar{X}_n$	DEH	0.5564 (0.0000)		0.4843 (0.0000)		0.4216 (0.0000)	
	VCKL	0.8462 (0.2371)		0.8164 (0.2605)		0.7099 (0.2749)	
	VCRQ	0.7991 (0.2790)		0.7664 (0.2730)		0.8440 (0.2250)	
f_n^B avec $\mu_n = \bar{X}_n$	DEH	0.5564 (0.0000)		0.4843 (0.0000)		0.4216 (0.0000)	
	VCKL	0.8551 (0.2148)		0.8420 (0.2471)		0.7409 (0.2643)	
	VCRQ	0.8542 (0.2283)		0.8292 (0.2161)		0.8716 (0.2273)	
f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	DEH	0.5564 (0.0000)		0.4843 (0.0000)		0.4216 (0.0000)	
	VCKL	0.8459 (0.2362)		0.8145 (0.2628)		0.7101 (0.2747)	
	VCRQ	0.8386 (0.2411)		0.7780 (0.2725)		0.8544 (0.2203)	
f_n^B avec $\mu_n = X_n^*$	DEH	0.5564 (0.0000)		0.4843 (0.0000)		0.4216 (0.0000)	
	VCKL	0.8487 (0.2232)		0.8266 (0.2566)		0.7110 (0.2740)	
	VCRQ	0.8397 (0.2440)		0.8200 (0.2419)		0.8868 (0.2167)	

TABLE 2.1 moyenne (écart-type) de l'ISE associé à la loi du khi-deux: $\chi_2(4)$

Pas	Estimateur	Tailles et nombres d'échantillons		
		$n_1=25$ $n_2=100$	$n_1=50$ $n_2=50$	$n_1=100$ $n_2=25$
DEH	f_n^{PR}	0.0091 (0.0056)	0.0064 (0.0029)	0.0043 (0.0015)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0103 (0.0071)	0.0069 (0.0036)	0.0046 (0.0017)
	f_n^B avec $\mu_n = \tilde{X}_n$	0.0099 (0.0070)	0.0057 (0.0033)	0.0041 (0.0025)
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0086 (0.0062)	0.0050 (0.0023)	0.0032 (0.0016)
	f_n^B avec $\mu_n = X_n^*$	0.0089 (0.0060)	0.0056 (0.0033)	0.0038 (0.0024)
VCKL	f_n^{PR}	0.0114 (0.0102)	0.0079 (0.0063)	0.0048 (0.0026)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0134 (0.0103)	0.0094 (0.0066)	0.0053 (0.0027)
	f_n^B avec $\mu_n = \tilde{X}_n$	0.0116 (0.0090)	0.0069 (0.0039)	0.0051 (0.0033)
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0119 (0.0095)	0.0075 (0.0053)	0.0045 (0.0028)
	f_n^B avec $\mu_n = X_n^*$	0.0104 (0.0081)	0.0068 (0.0039)	0.0049 (0.0034)
VCRQ	f_n^{PR}	0.0161 (0.0159)	0.0077 (0.0055)	0.0049 (0.0033)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0176 (0.0164)	0.0096 (0.0071)	0.0055 (0.0044)
	f_n^B avec $\mu_n = \tilde{X}_n$	0.0125 (0.0105)	0.0072 (0.0039)	0.0041 (0.0016)
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0131 (0.0099)	0.0073 (0.0045)	0.0038 (0.0021)
	f_n^B avec $\mu_n = X_n^*$	0.0112 (0.0095)	0.0065 (0.0033)	0.0039 (0.0017)

TABLE 2.2 moyenne (écart-type) du pas associé à la loi du khi-deux: $\chi_2(4)$

Estimateur	Pas	Tailles et nombres d'échantillons		
		$n_1=25$ $n_2=100$	$n_1=50$ $n_2=50$	$n_1=100$ $n_2=25$
$f_n^{PR}(\cdot)$	DEH	1.4848 (0.2914)	1.3311 (0.2100)	1.1766 (0.1494)
	VCKL	1.2953 (0.4808)	1.1345 (0.4240)	0.7951 (0.3420)
	VCRQ	1.1554 (0.4951)	0.9654 (0.2736)	0.7205 (0.2164)
f_n^B avec $\mu_n = \bar{X}_n$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)
	VCKL	0.5813 (0.2527)	0.4892 (0.1926)	0.3244 (0.1423)
	VCRQ	0.4670 (0.2240)	0.3636 (0.1640)	0.2860 (0.1036)
f_n^B avec $\mu_n = \tilde{X}_n$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)
	VCKL	0.6494 (0.2148)	0.5647 (0.1676)	0.4204 (0.1523)
	VCRQ	0.6094 (0.2049)	0.5472 (0.1809)	0.4676 (0.1097)
f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)
	VCKL	0.5999 (0.2461)	0.5156 (0.1910)	0.3623 (0.1444)
	VCRQ	0.5409 (0.2127)	0.4632 (0.1821)	0.3772 (0.0983)
f_n^B avec $\mu_n = X_n^*$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)
	VCKL	0.6524 (0.2128)	0.5760 (0.1646)	0.4138 (0.1402)
	VCRQ	0.6603 (0.1900)	0.5908 (0.1631)	0.4952 (0.0813)

TABLE 3.1 moyenne (écart-type) de l'ISE associé à la loi de student: $t(5)$

Pas	Estimateur	Tailles et nombres d'échantillons				
		$n_1=25$	$n_2=100$	$n_1=50$	$n_2=50$	$n_1=100$
DEH	f_n^{PR}	0.0154 (0.0143)	0.0080 (0.0050)	0.0058 (0.0044)		
	f_n^B avec $\mu_n = \bar{X}_n$	0.0134 (0.0153)	0.0057 (0.0035)	0.0043 (0.0038)		
	f_n^B avec $\mu_n = \tilde{X}_n$	0.0174 (0.0194)	0.0077 (0.0056)	0.0062 (0.0060)		
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0145 (0.0167)	0.0062 (0.0042)	0.0054 (0.0050)		
	f_n^B avec $\mu_n = X_n^*$	0.0168 (0.0196)	0.0070 (0.0053)	0.0057 (0.0067)		
VCKL	f_n^{PR}	0.0226 (0.0195)	0.0127 (0.0091)	0.0090 (0.0060)		
	f_n^B avec $\mu_n = \bar{X}_n$	0.0189 (0.0224)	0.0090 (0.0087)	0.0058 (0.0046)		
	f_n^B avec $\mu_n = \tilde{X}_n$	0.0192 (0.0227)	0.0099 (0.0093)	0.0069 (0.0059)		
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0189 (0.0221)	0.0087 (0.0088)	0.0065 (0.0051)		
	f_n^B avec $\mu_n = X_n^*$	0.0195 (0.0226)	0.0094 (0.0093)	0.0066 (0.0060)		
VCRQ	f_n^{PR}	0.0250 (0.0282)	0.0186 (0.0176)	0.0107 (0.0203)		
	f_n^B avec $\mu_n = \bar{X}_n$	0.0226 (0.0292)	0.0120 (0.0135)	0.0053 (0.0058)		
	f_n^B avec $\mu_n = \tilde{X}_n$	0.0210 (0.0240)	0.0111 (0.0100)	0.0058 (0.0047)		
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0225 (0.0255)	0.0115 (0.0108)	0.0059 (0.0056)		
	f_n^B avec $\mu_n = X_n^*$	0.0217 (0.0232)	0.0102 (0.0086)	0.0049 (0.0034)		

TABLE 3.2 moyenne (écart-type) du pas associé à la loi de student: $t(5)$

Estimateur	Pas	Tailles et nombres d'échantillons				
		$n_1=25$	$n_2=100$	$n_1=50$	$n_2=50$	$n_1=100$
f_n^{PR}	DEH	0.6845 (0.2610)	0.5935 (0.0788)	0.5284 (0.0581)		
	VCKL	0.7894 (0.3454)	0.7087 (0.2143)	0.6647 (0.1983)		
	VCRQ	0.6199 (0.2393)	0.4875 (0.2160)	0.4554 (0.1366)		
f_n^B avec $\mu_n = \bar{X}_n$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)		
	VCKL	0.7876 (0.2542)	0.8037 (0.2476)	0.7328 (0.2382)		
	VCRQ	0.6439 (0.2948)	0.5530 (0.2889)	0.5084 (0.2817)		
f_n^B avec $\mu_n = \tilde{X}_n$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)		
	VCKL	0.7979 (0.2360)	0.8053 (0.2410)	0.7406 (0.2339)		
	VCRQ	0.7273 (0.2692)	0.6358 (0.2666)	0.5520 (0.2810)		
f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)		
	VCKL	0.7859 (0.2557)	0.8051 (0.2466)	0.7360 (0.2365)		
	VCRQ	0.6646 (0.2915)	0.5820 (0.2788)	0.5224 (0.2849)		
f_n^B avec $\mu_n = X_n^*$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)		
	VCKL	0.7934 (0.2458)	0.7985 (0.2468)	0.7393 (0.2342)		
	VCRQ	0.7003 (0.2727)	0.6428 (0.2590)	0.5556 (0.2671)		

TABLE 4.1 moyenne (écart-type) de l'ISE associé à la loi binormale: $\frac{1}{2}N(-1.5, 1.) + \frac{1}{2}N(1.5, 1.)$

Pas	Estimateur	Tailles et nombres d'échantillons		
		$n_1=25$ $n_2=100$	$n_1=50$ $n_2=50$	$n_1=100$ $n_2=25$
DEH	f_n^{PR}	0.0157 (0.0116)	0.0091 (0.0055)	0.0055 (0.0041)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0142 (0.0080)	0.0093 (0.0052)	0.0056 (0.0022)
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0185 (0.0105)	0.0120 (0.0072)	0.0089 (0.0054)
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0146 (0.0083)	0.0096 (0.0055)	0.0059 (0.0026)
	f_n^B avec $\mu_n = X_n^*$	0.0173 (0.0100)	0.0109 (0.0068)	0.0068 (0.0034)
VCKL	f_n^{PR}	0.0190 (0.0133)	0.0112 (0.0066)	0.0066 (0.0041)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0151 (0.0130)	0.0079 (0.0060)	0.0038 (0.0030)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0171 (0.0150)	0.0102 (0.0079)	0.0064 (0.0077)
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0150 (0.0128)	0.0085 (0.0069)	0.0046 (0.0045)
	f_n^B avec $\mu_n = X_n^*$	0.0178 (0.0172)	0.0089 (0.0066)	0.0055 (0.0059)
VCRQ	f_n^{PR}	0.0295 (0.0593)	0.0170 (0.0169)	0.0094 (0.0099)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0224 (0.0304)	0.0121 (0.0131)	0.0061 (0.0072)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0208 (0.0237)	0.0106 (0.0088)	0.0060 (0.0050)
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0201 (0.0240)	0.0107 (0.0122)	0.0057 (0.0064)
	f_n^B avec $\mu_n = X_n^*$	0.0226 (0.0271)	0.0125 (0.0125)	0.0049 (0.0049)

TABLE 4.2 moyenne (écart-type) du pas associé à la loi binormale: $\frac{1}{2}N(-1.5, 1.) + \frac{1}{2}N(1.5, 1.)$

Estimateur	Pas	Tailles et nombres d'échantillons		
		$n_1=25$ $n_2=100$	$n_1=50$ $n_2=50$	$n_1=100$ $n_2=25$
f_n^{PR}	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)
	VCKL	0.7554 (0.2235)	0.6290 (0.1745)	0.5035 (0.1190)
	VCRQ	0.6693 (0.2379)	0.5042 (0.1896)	0.4293 (0.1400)
f_n^B avec $\mu_n = \bar{X}_n$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)
	VCKL	0.8370 (0.2152)	0.7472 (0.2560)	0.5881 (0.2216)
	VCRQ	0.7398 (0.2756)	0.5884 (0.2675)	0.5208 (0.2175)
f_n^B avec $\mu_n = \bar{X}_n$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)
	VCKL	0.8430 (0.2020)	0.7741 (0.2330)	0.6142 (0.1999)
	VCRQ	0.8056 (0.2463)	0.7248 (0.2599)	0.6412 (0.2301)
f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)
	VCKL	0.8381 (0.2134)	0.7506 (0.2569)	0.6073 (0.2040)
	VCRQ	0.7879 (0.2559)	0.6452 (0.2627)	0.5864 (0.2077)
f_n^B avec $\mu_n = X_n^*$	DEH	0.5564 (0.0000)	0.4843 (0.0000)	0.4216 (0.0000)
	VCKL	0.8377 (0.2136)	0.7519 (0.2454)	0.6114 (0.2019)
	VCRQ	0.7716 (0.2622)	0.6346 (0.2794)	0.6292 (0.1938)

TABLE 5.1 moyenne (écart-type) de l'ISE associé à la loi: $f_{\alpha}(\cdot)$ avec $\alpha = 4$

Pas	Estimateur	Tailles et nombres d'échantillons					
		$n_1=25$	$n_2=100$	$n_1=50$	$n_2=50$	$n_1=100$	$n_2=25$
DEH	f_n^{PR}	0.0275	(0.0190)	0.0142	(0.0081)	0.0113	(0.0076)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0283	(0.0256)	0.0167	(0.0138)	0.0116	(0.0075)
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0411	(0.0343)	0.0199	(0.0154)	0.0153	(0.0113)
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0297	(0.0265)	0.0175	(0.0143)	0.0125	(0.0086)
	f_n^B avec $\mu_n = X_n^*$	0.0292	(0.0260)	0.0170	(0.0139)	0.0117	(0.0077)
VCKL	f_n^{PR}	0.0475	(0.0479)	0.0210	(0.0164)	0.0158	(0.0106)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0480	(0.0475)	0.0218	(0.0146)	0.0154	(0.0100)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0514	(0.0391)	0.0241	(0.0160)	0.0187	(0.0126)
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0484	(0.0428)	0.0226	(0.0151)	0.0164	(0.0108)
	f_n^B avec $\mu_n = X_n^*$	0.0552	(0.0435)	0.0245	(0.0171)	0.0178	(0.0119)
VCRQ	f_n^{PR}	0.0637	(0.0945)	0.0200	(0.0200)	0.0221	(0.0356)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0649	(0.0946)	0.0324	(0.0315)	0.0189	(0.0153)
	f_n^B avec $\mu_n = \bar{X}_n$	0.0519	(0.0446)	0.0331	(0.0233)	0.0195	(0.0164)
	f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	0.0508	(0.0528)	0.0331	(0.0308)	0.0193	(0.0156)
	f_n^B avec $\mu_n = X_n^*$	0.0571	(0.0583)	0.0319	(0.0213)	0.0201	(0.0164)

TABLE 5.2 moyenne (écart-type) du pas associé à la loi: $f_{\alpha}(\cdot)$ avec $\alpha = 4$

Estimateur	Pas	Tailles et nombres d'échantillons					
		$n_1=25$	$n_2=100$	$n_1=50$	$n_2=50$	$n_1=100$	$n_2=25$
f_n^{PR}	DEH	0.5564	(0.0000)	0.4843	(0.0000)	0.4216	(0.0000)
	VCKL	0.2546	(0.0880)	0.2141	(0.0646)	0.1832	(0.0455)
	VCRQ	0.3442	(0.1519)	0.3130	(0.0879)	0.2465	(0.0861)
f_n^B avec $\mu_n = \bar{X}_n$	DEH	0.5564	(0.0000)	0.4843	(0.0000)	0.4216	(0.0000)
	VCKL	0.5865	(0.2808)	0.5697	(0.2568)	0.4159	(0.1795)
	VCRQ	0.6256	(0.3011)	0.5084	(0.2538)	0.3932	(0.1611)
f_n^B avec $\mu_n = \bar{X}_n$	DEH	0.5564	(0.0000)	0.4843	(0.0000)	0.4216	(0.0000)
	VCKL	0.5905	(0.2474)	0.5582	(0.2434)	0.4146	(0.1805)
	VCRQ	0.6564	(0.2614)	0.5478	(0.2545)	0.4196	(0.1545)
f_n^B avec $\mu_n = \bar{X}_{n\alpha}$	DEH	0.5564	(0.0000)	0.4843	(0.0000)	0.4216	(0.0000)
	VCKL	0.5783	(0.2772)	0.5612	(0.2510)	0.4133	(0.1794)
	VCRQ	0.6389	(0.2816)	0.5098	(0.2536)	0.3928	(0.1623)
f_n^B avec $\mu_n = X_n^*$	DEH	0.5564	(0.0000)	0.4843	(0.0000)	0.4216	(0.0000)
	VCKL	0.5821	(0.2689)	0.5507	(0.2363)	0.4131	(0.1762)
	VCRQ	0.6488	(0.2627)	0.5436	(0.2435)	0.3948	(0.1633)

REFERENCES

- B. Abdous (1989), *Strong uniform consistency of kernel probability density estimators based on sample moments*, Prob. Letters and Statist. (sous presse).
- B. Abdous et A. Berlinet (1986), *Convergence uniforme presque sûre d'une classe d'estimateurs à noyau, pour la densité*, C. R. Acad. Sci. Paris **303**, 761-764.
- S. Abou-Jaoudé (1977), *Sur la convergence L^1 et L^∞ de certains estimateurs d'une densité de probabilité*, Doctorat d'Etat. Université de Paris VI.
- J. H. Ahrens et U. Dieter (1974), *Computer methods for sampling from Gamma, Beta, Poisson and Binomial Distributions*, Computing **12**, 223-246.
- S. Bean et C. P. Tsokos (1982), *Bandwidth selection procedures for kernel density estimates*, Commun. Statist. Theor. Meth. (11)**9**, 1045-1069.
- M. Bertrand-Retali (1978), *Convergence uniforme d'un estimateur de la densité par la méthode du noyau*, Rev. Roumaine Math. Pures et Appliquées **23**, 361-385.
- H. J. Bierens (1983), *Sample moments integrating normal kernel estimators of multivariate density and regression functions*, Sankhya B **45**, 160-192.
- D. Bosq et J. P. Lecoutre (1985), "Théorie de l'estimation fonctionnelle," Economica, Paris.

- A. W. Bowman (1982), *A comparative study of some kernel-based nonparametric density estimates*, Manchester-Sheffield School of Probability and Statistics Research report 84/AWB/1.
- A. W. Bowman (1984), *An alternative method of cross-validation for the smoothing of density estimates*, *Biometrika* **71**, 353-360.
- A. W. Bowman (1985), *A comparative study of some kernel-based non-parametric density estimators*, *J. Statist. Comput. Simul.* **21**, 313-327.
- L. Breiman, W. Meisel et E. Purcell (1977), *Variable kernel estimates of multivariate densities*, *Technometrics* **19**, 135-144.
- T. Cacoullos (1966), *Estimation of multivariate density*, *Ann. Inst. Statist. math.* **18**, 179-189.
- Y.S. Chow, S. Geman, L.D. Wu (1983), *Consistent cross-validated density estimation*, *Ann. Statist.* **11**, 25-38.
- T. De Wet et T. W. J. Van Wyk (1979), *Efficiency and robustness of Hogg's adaptive trimmed means*, *Communications in Statistics A8*, 117-128.
- T. De Wet et T. W. J. Van Wyk (1979), *Some large sample properties of Hogg's adaptive trimmed means*, *South African Statistics Journal* **13**, 53-69.
- P. Deheuvels (1974), *Conditions nécessaires et suffisantes de convergence presque sûre et uniforme presque sûre des estimateurs de la densité*, *C. R. Acad. Sci. Paris A* **278**, 1217-1220.
- P. Deheuvels (1977-a), *Estimation non paramétrique de la densité par histogrammes généralisés (II)*, *Pub. Inst. Statist. Univ. Paris XXII*, 1-23.
- P. Deheuvels (1977-b), *Estimation non paramétrique de la densité par histogrammes généralisés*, *Rev. Statist. Appl.* **25**, 5-42.

- P. Deheuvels et P. Hominal (1980), *Estimation automatique de la densité*, Rev. Statist. Appl. **28**, 25-55.
- L. P. Devroye et T. J. Wagner (1980), *The strong uniform consistency of kernel density estimates*, Mult. Anal. **V**, 59-77, Ed. Krishnaiah, North-Holland, New York.
- L. P. Devroye et L. Györfi (1985), "Nonparametric Density Estimation: The L_1 -view," Wiley, New York.
- G.R. Ducharme et P. Milasevic (1987), *Spatial median and directional data*, Biometrika **74**, 212-215.
- R. P. W. Duin (1976), *On the choice of smoothing parameters of Parzen estimators of probability density functions*, IEEE Trans. Comput. **C-25**, 1175-1179.
- V. A. Epanechnikov (1969), *Nonparametric estimation of multidimensional probability density*, Theory Prob. Appl. **14**, 153-158.
- E. Fix et J. L. Jr. Hodges (1951), *Discriminatory analysis, nonparametric discrimination: consistency property*, Report no 4, USAF School of Aviation Medicine, Texas.
- J. L. Gastwirth (1966), *On robust procedures*, J. Amer. Statist. Assoc. **61**, 929-948.
- J. L. Gastwirth et M. L. Cohen (1970), *Small sample behavior of some robust linear estimators of location*, J. Amer. Statist. Assoc. **65**, 946-973.
- J. Geffroy (1973), *Sur la convergence uniforme des estimateurs d'une densité de probabilité*, Séminaire de Statistique - I.S.U.P.
- J.C. Gower (1974), *The median centre*, Appl. Statist. **23**, 466-470.
- F. A. Graybill (1983), "Matrices with applications in Statistics," Wadsworth, Belmont, California.

- J. D. F. Habbema, J. Hermans et K. Van Den Broek (1974), *A stepwise discriminant analysis program using density estimation*, Compstat 1974, Ed. G. Bruckman,, Vienna: Physica Verlag.
- P. Hall (1982), *Limit theorems for stochastic measures of the accuracy of density estimators*, Stochastic Processes Appl. **13**, 11-25.
- P. Hall (1983), *Large sample optimality of least squares cross-validation in density estimation*, Ann. Statist. **11**, 1156-1174.
- P. Hall (1984), *Central limit theorem for integrated square error of multivariate non-parametric density estimators*, J. Multivariate Anal. **14**, 1-16.
- P. Hall (1985), *Asymptotic theory of minimum integrated square error for multivariate density estimation*, In Krishnaiah, P.R. (ed.) Proc. Sixth Internat. Sympos. Multivariate Analysis. Amsterdam: North Holland, 289-309.
- W. Hoeffding (1963), *Probability inequalities for the sum of bounded random variables*, J. Amer. Statist. Assoc. **58**, 13-30.
- R. V. Hogg (1967), *Some observations on robust estimation*, J. Amer. Statist. Assoc. **62**, 1179-1186.
- R. V. Hogg et R. V. Lenth (1984), *A review of some adaptive statistical techniques*, Commun. Statist. Theor. Meth. **13**, 1551-1579.
- P. J. Huber (1972), *Robust statistics: A review*, Ann. Math. Statist. **43**, 1041-1067.
- P. J. Huber (1981), "Robust statistics," Wiley, New York.
- R. F. Kappenman (1987), *A nonparametric data based univariate density function estimate*, Computational Statist. & Data Analysis **5**, 1-7.
- L. Lebart, A. Morineau et T. Tabard (1977), "Techniques de la description statistique," Dunod, Paris.

- E. L. Lehmann (1983), "Theory point estimation," Wiley, New York.
- F. Loftsgaarden et C. Quesenberry (1965), *A nonparametric estimate of multivariate density function*, Ann. Math. Statist. **36**, 1049-1051.
- J.S. Marron (1985), *A comparison of cross-validation techniques in density estimation*, N.C. Inst. of Statist. Mimeo Series # 1568.
- E. Nadaraja (1965), *On nonparametric estimation of density function and regression*, Theory Prob. Appl. **10**, 186-190.
- E. Nadaraja (1974), *On the integral mean square error of some nonparametric estimates of probability density*, Theory Prob. Appl. **19**, 133-141.
- E. Parzen (1962), *On estimation of probability density function and mode*, Ann. Math. Statist. **33**, 1065-1076.
- K. Pearson (1902), *On the systematic fitting of curves to observations and measurements I, II*, Biometrika, (1) 265-303 and (2) 1-23.
- B. L. S. Prakasa Rao (1983), "Nonparametric functional estimation," Academic Press, New York.
- P. Prescott (1978), *Selection of trimming proportions for robust adaptive trimmed means*, J. Amer. Statist. Assoc. **73**, 133-140.
- M. Rosenblatt (1956), *Remarks on some nonparametric estimates of a density function*, Ann. Math. Statist. **27**, 832-837.
- M. Rudemo (1982), *Empirical choice of histograms and density estimators*, Scand. J. Statist. **9**, 65-78.
- H. Scheffé (1947), *A useful convergence theorem for probability distribution*, Ann. Math. Statist. **18**, 434-438.
- E. F. Schuster et G. G. Gregory (1981), *On the nonconsistency of maximum likelihood*

- nonparametric density estimators*, In the 13th Annual Symposium on the Interface of Computer Science and Statistics, 295-298, Springer Verlag, New York.
- D. W. Scott, R. A. Tapia et J. R. Thompson (1977), *Kernel density estimation revisited*, *Nonlinear Analysis; Theory, Methods and Applications* **1**, 339-372.
- K. S. Shanmugam (1977), *On a modified form of Parzen estimator of nonparametric pattern recognition*, *Pattern Recognition* **9**, 167-170.
- B. W. Silverman (1986), "Density Estimation for Statistics and Data Analysis," Chapman and Hall, New York.
- M. Spivak (1965), "Calculus on manifolds," W. A. Benjamin, Inc., New York.
- C. J. Stone (1984), *An asymptotically optimal window selection rule for kernel density estimates*, *Ann. Statist.* **12**, 1285-1297.
- M. Stone (1974), *Cross-validatory choice and assessment of statistical predictions*, *J. R. Statist. Soc. B* **36**, 111-147.
- M. Stone (1974), *Cross-validation and multinomial prediction*, *Biometrika* **61**, 509-515.
- M. Stone (1977), *An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion*, *J. R. Statist. Soc. B* **39**, 44-47.
- M. Stone (1977), *Asymptotics for and against cross-validation*, *Biometrika* **64**, 29-35.
- R. Tapia et J. Thompson (1978), "Nonparametric Probability Density Estimation," Johns Hopkins University Press, Baltimore, Maryland.
- J. Tiago De Oliveira (1963), *Estatística de desiderades. Resultados assintoticos*, *Rev. Fac. Univ. Lisboa A* **9**, 111-206.
- D.M. Titterington (1985), *Common structure of smoothing techniques in statistics*, *Int. Statist. Rev.* **53**, 141-170.
- V.N. Vapnik et A.Y. Chervonenkis (1971), *On the uniform convergence of the relative*

frequencies of events to their probabilities, Theory Probab. Appl. **16**, 264-280.

M. Woodroffe (1967), *On the maximum deviation of the sample density*, Ann. Math. Statist. **38**, 475-481.

