N° d'ordre : 1339

55376
1986
19

THESE

présentée à

L'UNIVERSITE DES SCIENCES
ET TECHNIQUES DE LILLE FLANDRES ARTOIS

pour obtenir

LE GRADE DE DOCTEUR DE 3e CYCLE
SPECIALITE : MATHEMATIQUES APPLIQUEES

par

ABDOUS Belkacem

CONVERGENCE ET CORRECTION LOCALE D'ESTIMATEURS A NOYAUX

Membres du Jury : D. BOSQ, Président
A. BERLINET, Rapporteur (Université de Grenoble I)
C. GOURIEROUX, Examinateur
P. JACOB, Examinateur
J.P. LECOUTRE, Invité

Soutenue le 26 Juin 1986

Je remercie vivement Monsieur le Professeur Denis BOSQ qui m'a fait l'honneur d'accepter la présidence du jury de cette thèse.

Monsieur le Professeur Alain BERLINET m'a proposé le sujet de cette thèse, sa rigueur, ses critiques stimulantes, ses conseils et ses encouragements m'ont été d'une aide efficace. Qu'il trouve ici l'expression de ma profonde gratitude.

Je suis très honoré de la présence de Messieurs les Professeurs Christian GOURIEROUX, Pierre JACOB et Jean-Pierre LECOUTRE qui ont accepté de s'intéresser à ce travail et de le juger.

Mes remerciements vont aussi aux membres de l'équipe des Probabilistes et Statisticiens de l'U.F.R. de Mathématiques de Lille I qui m'ont aidé de leurs remarques et entouré de leur sympathie.

Madame Arlette LENGAIGNE a dactylographié ce travail avec beaucoup de soin et de gentillesse. Qu'elle en soit remerciée ainsi que toutes les personnes ayant pris part à la réalisation matérielle de cette thèse.

P L A N

./..

Pages

INTRODUCTION

Durant les vingt dernières années l'estimation fonctionnelle non paramétrique a connu un vaste développement, en particulier l'estimation de la densité dans $\mathbb{R}^k$, $k \geq 1$ a fait l'objet de multiples travaux par des méthodes diverses. Nous ne citerons ici que les plus utilisées et plus importantes qu'on peut répertorier en quatre catégories :

i) les histogrammes ou estimateurs à noyaux.

ii) les estimations par fonctions orthogonales.

iii) la méthode des splines.

iv) les méthodes dites de maximum de vraisemblance.

Pour une bibliographie générale et des exposés de synthèse nous renvoyons aux livres de : R.A. TAPIA et J.R. THOMPSON (1978) B.L.S. PRAKASA RAO (1983), L. DEVROYE et L. GYÖRFI (1984) et D. BOSQ et J.P. LECOUTRE (1985).

La méthode qui semble commode, simple à mettre en oeuvre, robuste et ne nécessitant pas un choix multiple de paramètres est la méthode du noyau, méthode qui a suscité le plus d'intérêt depuis l'apparition des articles de M. ROSENBLATT (1956) et de E. PARZEN (1962), qui, dans le cas univarié, consiste en ceci :

" Si $\{X_n, n \geq 1\}$ est une suite de v.a. indépendantes de même loi sur $\mathbb{R}$, possédant une densité $f(x)$, relativement à la mesure de Lebesgue sur $\mathbb{R}$, continue, alors l'estimateur à noyau de $f(x)$ s'écrit :

$$f_n(x) = \frac{1}{n} \sum_{j=1}^{n} \frac{1}{h_n} K\left(\frac{x - X_j}{h_n}\right) \qquad \text{où}$$

- $K(u)$ est généralement supposé tel que :

$$\forall u \quad K(u) \geq 0 \ , \quad \int K(u)\, du = 1 \ , \quad \sup_u |K(u)| \leq \bar{K} \ ,$$

- h_n est un paramètre réel positif, appelé généralement pas ou fenêtre ou largeur de la fenêtre " .

Le choix de K , convenablement normalisé par h_n ; influe peu sur l'efficacité de l'estimation, en pratique on se limite le plus souvent à des noyaux bornés et simples, comme le noyau unité ou triangulaire ou parabolique (Epanechnikov) ou encore le noyau normal. Par contre le choix de h_n est primordial et a une importance considérable pour l'efficacité de l'estimation, un choix de h_n trop petit implique la présence de fluctuations aléatoires trop importantes et un choix de h_n trop grand élimine les aléas mais introduit des biais de lissage trop importants. De nombreux auteurs se sont penchés sur ce problème mais malgré cela la difficulté du choix pratique et adéquat de h_n subsiste.

Pour ces raisons nous avons dressé, dans la première partie de ce travail, le bilan des plus intéressantes méthodologies qui ont été proposées pour choisir le pas optimal, nous avons mis en évidence le principe et les motivations de chacune de ces méthodes, à la fin nous avons proposé une procédure d'estimation en trois étapes, elle utilise la procédure de J.P. LECOUTRE (1975) ainsi que le conditionnement par rapport à une statistique localement exhaustive.

Un article de H.J. BIERENS (1983) est à l'origine du second chapitre, l'idée de H.J. BIERENS est de construire un estimateur à noyau strict, dans le sens qu'il est une densité et que ses moments et les moments empiriques d'ordre 1 et 2 de la densité à estimer coïncident, moyennant une légère transformation de l'estimateur de Parzen-Rosenblatt, H.J. BIERENS a construit un estimateur

f_{n_1} qui répond aux conditions :

a) f_{n_1} est une densité sur $\mathbb{R}^k$,

b) $\int_{\mathbb{R}^k} x\, f_{n_1}(x)\, dx = \bar{X} \ (= \frac{1}{n} \sum_{i=1}^{n} X_i)$,

c) $\int_{\mathbb{R}^k} x\, x^t\, f_{n_1}(x)\, dx = \frac{1}{n} \sum_{j=1}^{n} X_j\, X_j^t$,

où $(X_j)_{j=1,\dots,n}$ sont des v.a.i.i.d. issues de la loi à estimer, dont une version de la densité par rapport à la mesure de Lebesgue λ^k , $k \geq 1$, est notée f. Cet estimateur s'écrit sous la forme :

$$f_{n_1}(x) = \frac{1}{n\, \gamma_n^k\, |\mathrm{Det}\ S_n|^{1/2}} \sum_{j=1}^{n} K\left(S^{-1/2}\left(\frac{x-\sqrt{1-\gamma_n^2}\, X_j - (1-\sqrt{1-\gamma_n^2})\, \bar{X}}{\gamma_n}\right)\right)$$

où γ_n est une suite réelle dans $]0,1]$ et telle que $\gamma_n \xrightarrow[n\to\infty]{} 0$ et S_n est la matrice variance - covariance empirique : $\frac{1}{n} \sum_{j=1}^{n} (X_j-\bar{X})\ (X_j-\bar{X})^t$.
On remarquera que le paramètre de lissage est fonction de la dispersion S_n de l'échantillon $(X_i)_{i=1,\dots,n}$, comme cela est le cas pour la fenêtre proposée par P. DEHEUVELS (1977).

H.J. BIERENS a établi la convergence uniforme en probabilité de f_{n_1} vers f , dans le cas du noyau normal uniquement ; pour notre part, nous avons donné une condition suffisante de convergence uniforme presque sûre, dans le cas d'un noyau quelconque et avons proposé un choix de γ_n , pour lequel la convergence du risque quadratique a été établie. Ensuite, nous avons repris le principe de cette méthode pour construire une classe d'estimateurs à noyaux qui généralise celle de f_{n_1} , en prenant des estimateurs non précisés m_{n_1} et m_{n_2} de $\int_{\mathbb{R}^k} x\, f(x)\, dx$ resp. $\int_{\mathbb{R}^k} x\, x^t\, f(x)\, dx$, la motivation de cette démarche est que : suivant les informations dont on dispose sur le modèle à étudier, il est parfois préférable de considérer des estimateurs autres que la moyenne et variance empiriques, nous faisons allusion à la moyenne empirique α-tronquée ou α-winsorizée, à la matrice variance - covariance modifée $S'_n = \frac{n}{n-1} S_n$ ou α-tronquée etc ... Il serait également intéressant de construire un estimateur dont le mode est une estimation du mode de la loi à étudier, lorsque celui-ci existe, on peut faire la même chose pour la médiane ou tout autre paramètre estimable, nous comptons reprendre tout cela dans un prochain travail.

Actuellement, la majorité des travaux sur les estimateurs à noyaux penchent vers le choix d'une fenêtre qui dépend à la fois des observations et du point où l'on veut estimer la densité. Ces dépendances varient suivant les motivations et les raisons avancées ; mais ces estimateurs se ressemblent tous par leur forme,une généralisation de cette forme, donnée par A. BERLINET, est la suivante :

$$f_n(x) = \frac{1}{n} \sum_{i=1}^{n} |J_n(\omega,x,X_i(\omega))| \; K(\theta_n(\omega,x,X_i(\omega))$$

où $(X_i)_{i=1,\dots,n}$ est une suite de variables aléatoires définies sur un espace probabilisé $(\Omega,\mathcal{A},P)$ à valeurs dans $\mathbb{R}^k$, $k \geq 1$, indépendantes et de même loi à densité admettant une version continue, bornée f.

θ_n est une application définie sur $\Omega \times \mathbb{R}^k \times \mathbb{R}^k$ à valeurs réelles vérifiant les conditions suivantes :

- Pour P-presque tout ω, $\theta_n(\omega,x,.)$ est, quelque soit x, un homéomorphisme de $\mathbb{R}^k$, à dérivées partielles continues dont le jacobien est noté $J_n(\omega,x,.)$.

- $\theta_n(.,x,X(.))$ et $J_n(.,x,X(.))$ sont des applications mesurables.

K est une application de $\mathbb{R}^k$ dans $\mathbb{R}$, intégrable par rapport à la mesure de Lebesgue λ^k et $\int_{\mathbb{R}^k} K \, d\lambda^k = 1$.

Dans la troisième partie, sous forme d'un théorème général, nous avons établi la convergence uniforme p.s. de f_n vers f, sur des sous-ensembles de $\mathbb{R}^k$. En applications nous avons montré que cette classe recouvre un bon nombre d'estimateurs classiques : P. DEHEUVELS (1977), L. DEVROYE et T.J. WAGNER (1980), CHAI GENXIANG (1984) H.J. BIERENS (1983) etc ..., estimateurs pour lesquels nous avons donné à chaque fois une version du théorème général.

La dernière partie est essentiellement consacrée à la correction locale de l'estimateur à noyau univarié, par application de la technique du conditionnement introduite par A. BERLINET (1984-a) et (1984-b), dans un contexte plus général que l'estimation de la densité en un point. Pour éviter d'entrer dans les détails nous allons nous contenter de reprendre les idées intuitives de cette démarche : (on ne s'intéresse qu'à la densité).

Soit $\mathcal{P}$ l'ensemble des lois de probabilité Q sur $\mathbb{R}$, absolument continues par rapport à la mesure de Lebesgue, dont la densité admet une version continue f_Q, x un réel fixé ; sous la seule hypothèse $Q \in \mathcal{P}$, la connaissance de Q sur un sous-ensemble E de $\mathbb{R}$ tel que $Q(E) < 1$ ne nous permet pas de déterminer f_Q, tandis que la connaissance de sa restriction à un voisinage U de x entraîne celle de $f_Q(x)$, le paramètre $f_Q(x)$ est dit localisé par U (pour une définition précise de la localisation, nous renvoyons à D. BOSQ (1970) ou A. BERLINET (1984) page : X-07). La conséquence, de cette constatation, sur l'estimation de $f_Q(x)$ au vu d'un échantillon $X_1,\dots,X_n$, est que les observations tombant à proximité de U fourniront l'essentiel de l'information apportée par cet échantillon, d'où l'idée de considérer la statistique $\mathcal{S}$ qui retient les observations qui tombent dans U et qui censure les autres (Q est supposée connue sur U^c) cette statistique a la remarquable propriété de résumer l'information que l'on peut avoir sur la restriction de Q à U (notée $Q|_U$), i.e. ce qui manque pour connaître entièrement la loi Q, en d'autres termes, cela revient à dire que $\mathcal{S}$ est localement exhaustive ou U-exhaustive, d'une façon générale une statistique est dite U-exhaustive si elle est exhaustive au sens usuel dans le modèle $(\mathcal{R}, \mathcal{B}_{\mathbb{R}}, \mathcal{P}(Q,U^c))$ où $\mathcal{P}(Q,U^c) = \{P \in \mathcal{P} : P|_{U^c} = Q|_{U^c}\}$. Par ailleurs, d'après le théorème de RAO-BLACKWELL pour améliorer un estimateur il suffit de le conditionner par rapport à une statistique exhaustive.

En choisissant $U =]x - h_n , x + h_n[$, A. BERLINET (1984) a étudié les effets du conditionnement par rapport à la statistique $\mathcal{S} = (S_1,\ldots,S_n)$ où

$$\begin{cases} S_i = X_i & \text{si} \quad X_i \in U \\ \quad = x_o & \text{sinon avec} \quad x_o \in U^c \end{cases}$$

sur l'estimateur à noyau $f_n(x) = \dfrac{1}{nh_n} \sum_{j=1}^{n} K\left(\dfrac{x - X_j}{h_n}\right)$.

Nous avons repris cela dans le cas d'un ouvert U de la forme $]x - H_n , x + H_n[$ où H_n est une suite de réels ≥ 0, différente de h_n. Ensuite, nous avons montré la normalité asymptotique et une loi du log itéré pour cet estimateur. La convergence presque complète est également traitée. D'autre part, la technique du conditionnement repose sur la connaissance d'une bonne pré-estimation f_P de f_Q sur U^c, dans le cas où cette pré-estimation est prise égale à l'estimateur à noyau de f_P, nous avons établi la convergence en moyenne et en moyenne quadratique de l'estimateur conditionné qui en résulte. Enfin nous avons défini l'estimateur à noyau conditionné séquentiel, par le biais d'ouverts de la forme $U_i =]x - h_i, x + h_i[$ où $h_i \in \mathbb{R}_+^*$, estimateur pour lequels différents résultats de convergence ont été établis.

CHAPITRE I

CHOIX DE LA FENETRE EN ESTIMATION DE LA DENSITE PAR LA METHODE DU NOYAU.

I - RAPPELS SUR LE CALCUL DU M.I.S.E.-

Soit (X_i) $i = 1,\dots,n$ une suite de variables aléatoires indépendantes de même loi, à valeurs dans $\mathbb{R}$, dont une version de la densité par rapport à la mesure de Lebesgue est notée f. L'estimateur à noyau de f, en $x \in \mathbb{R}$, s'écrit

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

où $(h_n)_{n\geq 1}$ est une suite réelle telle que $h_n > 0$ et $h_n \xrightarrow[n\to\infty]{} 0$.

K est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, λ-intégrable : $\int K\, d\lambda = 1$.
Pour mesurer l'efficacité de f_n plusieurs critères ont été avancés, le critère le plus commode semble être celui de l'écart quadratique moyen intégré (Mean integrated square error) défini par :

$$R_n^2 = \int_{-\infty}^{+\infty} E(f_n(x) - f(x))^2\, dx = E\left(\int_{-\infty}^{+\infty} (f_n(x) - f(x))^2\, dx\right) = B_1^2 + B_2^2$$

où $$B_1^2 = \int_{-\infty}^{+\infty} (E(f_n(x)) - f(x))^2\, dx \quad \text{et} \quad B_2^2 = \int_{-\infty}^{+\infty} \operatorname{Var}\,(f_n(x))\, dx$$

d'autres mesures d'écart ont été étudiées notamment par P. BICKEL et M. ROSENBLATT (1973) :

$$\sup_x \frac{|f_n(x) - f(x)|}{f(x)} \quad , \quad \int_{-\infty}^{+\infty} \frac{(f_n(x) - f(x))^2}{f(x)}\, dx$$

par L. DEVROYE et C.S. PENROD (1984) : $\int_{-\infty}^{+\infty} |f_n(x) - f(x)| \, dx$.

Une estimation réalisant le minimum de R_n^2 sera dite estimation du MISE,

nous avons

$$E(f_n(x)) = \int_{-\infty}^{+\infty} K(y) \, f(x - h_n \, y) \, dy \quad \text{et}$$

$$\mathrm{Var}(f_n(x)) = \frac{1}{n \, h_n} \int_{-\infty}^{+\infty} K^2(y) \, f(x - h_n \, y) \, dy - \frac{1}{n} \left(\int_{-\infty}^{+\infty} K(y) \, f(x - h_n \, y) \, dy\right)^2$$

d'où si f est de classe C^2 alors :

$$B_1^2 \sim B_1^{2*} = \frac{1}{4} h_n^4 \left[\int_{\mathbb{R}} y^2 \, K(y) \, dy\right]^2 . \int_{\mathbb{R}} f''^2(x) \, dx \quad \text{et}$$

$$B_2^2 \sim B_2^{2*} = \frac{1}{n \, h_n} \int_{\mathbb{R}} K^2(y) \, dy$$

à condition que $h_n \xrightarrow[n\to\infty]{} 0$ et $n \, h_n \xrightarrow[n\to\infty]{} \infty$.

On définit un M.I.S.E. asymptotique R_n^{2*} , (A.M.I.S.E. voir P. DEHEUVELS (1977-a)

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

le pas h_n^* minimisant le critère de l'A.M.I.S.E. est alors donné par :

$$h_n^* = \alpha(K) \, \beta(f) \, . \frac{1}{n^{1/5}} \quad \text{où} \tag{0}$$

$$\alpha(K) = \left(\frac{\int_{\mathbb{R}} K^2(y) \, dy}{\left(\int_{\mathbb{R}} y^2 \, K(y) \, dy\right)^2} \right)^{1/5} \quad \text{et} \quad \beta(f) = \frac{1}{\left(\int_{\mathbb{R}} f''^2(x) \, dx\right)^{1/5}}$$

mais en pratique, une détermination exacte de h_n^* est impossible car h_n^* dépend de la quantité inconnue $\int_{\mathbb{R}} f''^2(x) \, dx$, pour surmonter ce handicap plusieurs approches ont été proposées, avant d'en citer quelques unes, nous allons évoquer un résultat très important du point de vue théorique, il est dû à G. WATSON et M. LEADBETTER (1963) :

Pour évaluer le M.I.S.E. on utilise l'égalité de Parseval i.e. :

$$R_n^2 = E\left(\int_{\mathbb{R}} (f_n(x) - f(x))^2\, dx\right) = \frac{1}{2\pi} E\left(\int_{\mathbb{R}} |\psi_{f_n}(u) - \psi_f(u)|^2\, du\right)$$

où
$$\psi_{f_n}(u) = \int_{\mathbb{R}} e^{itu} f_n(t)\, dt = \frac{1}{n}\, \psi_K(h_n u) \cdot \sum_{j=1}^{n} e^{iuX_j}$$

et
$$\psi_K(u) = \int_{\mathbb{R}} e^{itu} K(t)\, dt \ , \ \psi_f(u) = \int_{\mathbb{R}} e^{itu} f(t)\, dt$$

le fait que les X_i $i = 1,\dots,n$ soient des v.a. i.i.d. permet d'avoir :

$$R_n^2 = \frac{1}{2\pi} \left\{ \int_{-\infty}^{+\infty} \left[\frac{1}{n} + \frac{n-1}{n} |\psi_f(u)|^2\right] \left[\left|\psi_K(h_n u) - \frac{|\psi_f(u)|^2}{\frac{1}{n} + \frac{n-1}{n} |\psi_f(u)|^2}\right|^2\right] du \right.$$
$$\left. + \int_{-\infty}^{+\infty} \frac{|\psi_f(u)|^2 (1 - |\psi_f(u)|^2)}{1 + (n-1)\, |\psi_f(u)|^2}\, du \right\}$$

R_n^2 est minimal pour :

(i) $$\psi_K(h_n u) = \frac{|\psi_f(u)|^2}{\frac{1}{n} + \frac{n-1}{n} |\psi_f(u)|^2} \quad , \quad \forall u \in \mathbb{R}$$

le M.I.S.E. minimal est donc égal à :

$$R_n^{2**} = \frac{K_n(0)}{2\pi n} - \frac{1}{2\pi} \int \frac{|\psi_f(t)|^4}{1 + (n-1)\, |\psi_f(t)|^2}\, dt \quad ,$$

K_n est le noyau solution de (i). Malheureusement l'inversion de (i) n'est simple que dans quelques cas particuliers, ce qui fait que ce résultat est inexploitable en pratique.

II - QUELQUES METHODES D'APPROXIMATION DE LA FENETRE OPTIMALE.-

1°) Méthode de P. DEHEUVELS.

P. DEHEUVELS (1977-a) propose de remplacer $\beta(f)$, dans l'expression (O) page 2 , par $\frac{S}{C^{1/5}}$ où $S = S(X_1,\ldots,X_n)$ est une statistique de l'échantillon estimant un paramètre de dispersion de la loi des X_i et C est une constante convenablement choisie et indépendante des observations X_i , $i = 1,\ldots,n$, plus particulièrement pour la densité d'une loi normale $N(0,1)$, P. Deheuvels pose $\beta(f) = \frac{S_n}{(3/8\sqrt{\pi})^{1/5}}$ avec $S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i-\bar{X})^2$ et $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ ainsi comme choix final de h_n^* il prend :

$$h_n^* = \frac{1}{n^{1/5}} \alpha(K) \frac{S_n}{(3/8\sqrt{\pi})^{1/5}} .$$

Ce choix se justifie par le fait qu'il est anormal de prendre le même pas pour un échantillon où les observations sont très dispersées et un autre où elles sont regroupées. D'autre part si X_1 possède une variance σ^2 alors S_n^2 converge p.s. vers σ^2 et $h_{n_1}^*$ est équivalent à h_n^* si $C = \int_{\mathbb{R}} f''^2(x)\, dx$, pour $n \to \infty$ et des choix de C très différents de $3/8\sqrt{\pi}$ changent peu les valeurs obtenues, du fait que C est élevé à la puissance $1/5$. Il faut aussi remarquer que $h^*_{n_1}$ est optimal quand on est en présence d'une loi normale $N(m,\sigma^2)$ et quand $\sigma^2 = S^2_n$, car dans ce cas

$$\beta(f) = \int_{\mathbb{R}} f''^2(x)\, dx = \frac{3}{8\sqrt{\pi}}\, \sigma^{-5} \quad .$$

Pour des raisons de stabilité numérique il est recommandé de ne pas appliquer directement la formule $\frac{1}{n-1} \sum_{i=1}^{n} (X_i-\bar{X})^2$, mais utiliser plutôt la statistique ordonnée $X_{(1)},\ldots,X_{(n)}$. La robustesse de $h^*_{n_1}$ est également étudiée par P. DEHEUVELS (1977-a) pour les détails nous renvoyons à son article.

2°) Méthode de P. DEHEUVELS et P. HOMINAL.

P. HOMINAL et P. DEHEUVELS (1980) étudient les variations de $\int_{\mathbb{R}} f''^2_n(x)\,dx$ en fixant l'échantillon, ils considèrent un noyau symétrique, positif, borné et à support compact, d'intégrale égale à un et dont la dérivée seconde n'est pas identiquement nulle, on le notera H , ils obtiennent le résultat suivant :

Théorème 1.1.- Si on pose $g''_n(x) = \frac{1}{n\,\lambda_n^3} \sum_{i=1}^{n} H''\left(\frac{X_i-x}{\lambda_n}\right)$ où $\{\lambda_n,\ n \geq 1\}$ est une suite de réels positifs.

Alors si $\{\lambda_n\}_n$ vérifie $\lim_{n\to\infty} \lambda_n = 0$, $\lim_{n\to\infty} n\,\lambda_n^5 = 0$

on a
$$\lim_{n\to\infty} \int_{\mathbb{R}} (g''_n(x))^2\, dx \underset{p.s.}{=} \frac{1}{4} \left[\int y^2\, H''(y)\, dy\right]^2 . \int f''^2(y)\, dy \quad .$$

A partir de ce théorème, ils proposent de modifier le pas optimal h^*_n comme suit :

$$h^*_{n_2} = \left[\frac{\int K^2(y)\,dy \,.\, (\int y^2 H''(y)\,dy)^2}{4\,(\int y^2 K(y)\,dy)^2}\right]^{1/5} \cdot \frac{1}{n^{1/5}} \cdot \frac{1}{(\int (g''_n(x))^2\,dx)^{1/5}}$$

mais ce nouveau choix introduit un paramètre inconnu λ_n . Sous les hypothèses faites sur H , une étude des variations de $\int g_n''^2(x)\,dx$ les a incités à prendre la quantité :

$$\hat{\theta}_n(\lambda_n) = \frac{4}{(\int y^2 H''(y)\,dy)^2} \int g_n''^2(x)\,dx - \frac{1}{n\,\lambda_n^5} \int H''^2(y)\,dy$$

comme approximation de $\int f''^2(x)\,dx$ et à proposer le critère suivant pour le choix de λ_n :

" λ_n^* est la valeur qui réalise le maximum de $\hat{\theta}_n(\lambda)$ pour $\lambda \geq I$,

où $I = \text{Inf}\,\{\xi > 0 \,,\ \forall \eta \geq \xi \,,\ \hat{\theta}_n(\eta) > 0\}$, ou bien pour $\lambda \geq J$, avec $J = \text{Inf}\,\{\xi > 0 \,,\ \forall \eta \geq \xi \,,\ \hat{\theta}_n(\eta) \leq \hat{\theta}_n(\xi)\}$ " .

Les résultats pratiques de cette méthode sont satisfaisants aussi bien pour des échantillons de petite taille que pour ceux de grande taille.

3°) Méthode de L. BREIMAN, W. MEISEL, E. PURCELL (1977).

L'estimateur f_n présente quelques défauts, par exemple :

i) Si dans une région où f est voisine de 0 , on ne dispose que d'une seule observation, disons X_j , alors f_n aura un pic en $x = X_j$ alors qu'il est censé être proche de 0 dans toute cette région. Inversement, dans une région où f est grande, les observations vont s'y grouper ainsi l'estimateur f_n aura tendance à s'étaler.

ii) Aucun résultat asymptotique ne nous fournit un pas optimal indépendant de f et de ses dérivées.

Pour remédier à ces deux problèmes L. BREIMAN, W. MEISEL et E. PURCELL (1977) ont proposé l'estimateur suivant :

$$f_{n_1}(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\alpha_k\, d_{i_k}} K\left(\frac{X_i - x}{\alpha_k\, d_{i_k}}\right) ,$$

où - d_{i_k} = distance de X_i à son $k^{\text{ème}}$ plus proche voisin

- α_k est une constante multiplicative

- k est un entier fixé, dépendant de n.

Cet estimateur répond aux deux remarques précédentes car d'une part d_{j_k} est grand si on est dans une région où f est proche de 0 et on ne dispose que de X_j et par conséquent f_{n_1} s'étalera, l'inverse se produit si on est dans une région où f prend des grandes valeurs, d'autre part pour le choix de k et de α_k ils utilisent essentiellement le théorème suivant :

Théorème 1.2.- Soit f une densité de probabilité sur $\mathbb{R}^p$, $p \geq 1$ et (X_i) , $i = 1,\ldots,n$ un échantillon issu de f.

Soit $w_j = \exp(-\, n f(X_j)\, v(d_{j,1}))$, $j \in \overline{1,n}$, où $v(r) = \dfrac{\pi^{p/2}\, r^p}{\Gamma(1+p/2)}$ est le volume d'une boule de rayon r.

$d_{j,1}$ = distance de X_j à son plus proche voisin.

Alors les (w_j) $j \in \overline{1,n}$, sont asymptotiquement uniformes sur $]0,1[$.

Donc si $f_{n_1}(x)$ ajuste convenablement $f(x)$ alors les $\hat{w}_j = \exp(- n f_{n_1}(X_j)\, v(d_{j,1}))$ doivent suivre approximativement une loi uniforme.

Ils mesurent la dissemblance entre la répartition des $(\hat{w}_j)$ et la loi uniforme sur $\{\frac{1}{n}, \frac{2}{n}, \ldots, 1\}$ au moyen de la statistique :

$$\hat{S}_n = \sum_{j=1}^{n} (\hat{w}(j) - j/n)^2 \quad \text{avec} \quad \hat{w}(j) = \text{le } j^{\text{ème}} \text{ terme de}$$

la suite ordonnée $(\hat{w}_j)$ $j \in \overline{1,n}$, qu'ils cherchent à minimiser.

4°) Méthode de B.W. SILVERMAN (1978).

La méthode de B.W. SILVERMAN (1978) repose sur le résultat suivant :

" Si f admet une dérivée 2^{nde} uniformément continue, K vérifie certaines conditions (techniques) que nous omettrons ici, $(h_n)_n$ est choisie de telle façon que f''_n converge uniformément vers f'' quand $n \to \infty$. Alors

$$\sup_x \left[f''_n(x) - E(f''_n(x))\right] \xrightarrow[n\to\infty]{P} + \alpha \sup_x |f''(x)|$$

$$\inf_x \left[f''_n(x) - E(f''_n(x))\right] \xrightarrow[n\to\infty]{P} - \alpha \sup_x |f''(x)|$$

avec $$\alpha = \frac{1}{2} \left|\int x^2 K(x)\, dx\right| \left[\frac{\int K''^2(x)\, dx}{\int K^2(x)\, dx}\right]^{1/2}$$ " .

B.W. Silverman interprète ce résultat comme suit :

Si la fenêtre h_n est choisie de telle façon à avoir la meilleure estimation de $f(x)$ alors les fluctuations aléatoires de la dérivée seconde de f_n seront asymptotiquement au maximum de l'ordre de $\pm\, \alpha \sup |f''|$, car si on considère que $E f''_n$ est une bonne approximation de f'' alors toutes fluctuations dans f''_n sont dues au terme aléatoire $(f''_n - E f''_n)$, tandis que toute variation systématique est due à celle de $E f''_n$ i.e. de f'' .

Enfin pour déterminer le pas optimal B.W. Silverman suggère de dessiner le graphe de $f''_n(x)$ pour différentes valeurs de h_n et à partir de ces graphes déterminer h_n^* pour lequel le graphe de $f''_n(x)$ a des fluctuations ni trop marquées ni trop lissées.

Cette méthode présente l'inconvénient d'être très coûteuse lorsque le nombre d'observations n est trop élevé et de mal se généraliser pour le cas multivarié.

Pour son étude B.W. Silverman a choisi le noyau :

$$K(x) = \begin{cases} \frac{1}{4}|x|^4 - \frac{1}{2}|x|^3 + \frac{1}{2} & \text{si} \quad |x| \leq 1 \\ \frac{1}{4}|x|(2-|x|)^3 & \text{si} \quad 1 \leq |x| \leq 2 \\ 0 & \text{sinon} \end{cases}$$

qui est un noyau facile à mettre en oeuvre et dont la dérivée seconde est un polynôme de second degré, ce qui évite d'avoir une courbe avec des pics, car il est plus difficile de faire la part entre fluctuations systématiques et aléatoires avec des courbes affines par morceaux.

5°) Méthode de R.A. TAPIA et J.R. THOMPSON (1978).

La démarche heuristique de R.A. TAPIA et J.R. THOMPSON (1978) p. 67 se base sur le fait que $\int f_n''^2(x)\,dx$ converge p.s. vers $\int f''^2(x)\,dx$, sous certaines conditions sur f et K, ils ont donc remplacé $\beta(f)$ par $\beta(f_n)$ dans l'expression (0) p. 2 , du pas optimal h_n^* , à la suite de cela R.A. TAPIA et J.R. THOMPSON ont proposé deux approches qui déterminent une approximation du pas optimal h_n^*. La première consiste à appliquer la relation récursive suivante :

$$h_n^{i+1} = n^{-1/5}\, \alpha(K)\, \beta(f_n^i)$$

où $$f_n^i(x) = \frac{1}{n\, h_n^i} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n^i}\right)$$

relation pour laquelle ils n'ont précisé ni l'initialisation ni le test d'arrêt.

La seconde approche est l'application d'une méthode de Newton à la recherche du point fixe de la relation fonctionnelle :

$$h = g_n(h) \quad \text{avec} \quad g_n(h) = n^{-1/5}\, \alpha(K)\, \hat{\beta}(h)$$

où $$\hat{\beta}(h) = \left(\int f_{nh}''^{2}(x)\, dx\right)^{-1/5} \quad \text{et} \quad f_{nh}(x) = \frac{1}{n\, h} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h}\right) .$$

Méthode pour laquelle, après cinq itérations environ, ils obtiennent ; pour leurs exemples : $|h_n^{i+1} - h_n^i| < 10^{-5}$ cependant aucun résultat de convergence de h_n^i vers h_n^* n'a été démontré.

6°) *La validation croisée.*

Parmi les approches, qui font intervenir les observations pour évaluer la paramètre de lissage optimal, la validation croisée est la méthode qui a eu le plus de succès, elle a été introduite par J.D.F. HABBEMA J. HERMANS et K. VAN DEN BROEK (1974), R.P.W. DUIN (1976), les travaux de C. STONE (1974 - 1977) ont largement contribué à son adoption par de nombreux auteurs. Sa technique est utilisée dans plusieurs domaines d'estimation non-paramétrique.

Nous rappelons brièvement en quoi elle consiste dans le cas de l'estimateur à noyau :

si $f_{n-1,i}$ désigne l'estimateur à noyau de f, obtenu après avoir éliminé la $i^{\text{ème}}$ observation, <u>i.e.</u>

$$f_{n-1,i}(x) = \frac{1}{(n-1)h} \sum_{\substack{j=1 \\ j\neq i}}^{n} K\left(\frac{x - X_j}{h}\right)$$

alors la quantité $f_{n-1,i}(X_i)$ servira à mesurer l'adéquation de h comme valeur du paramètre de lissage de la façon suivante :

si la quantité $f_{n-1,i}(X_i)$ est grande alors on dira que $f_{n-1,i}$ anticipe l'observation X_i et h est un pas convenable, mais si elle est petite alors on dira que l'observation X_i est "malchanceuse" sous la densité $f_{n-1,i}$ et cela peut être attribué au mauvais choix de h. En faisant varier i de 1 à n, on obtient n mesures d'ajustement qu'on peut résumer sous forme d'une fonction en h, définie par :

$$L_n(h) = \prod_{i=1}^{n} f_{n-1,i}(X_i)$$

qu'il faut donc maximiser, le caractère heuristique et arbitraire de cette approche fait qu'elle peut diverger, car on n'est pas toujours sûr que f^i_{n-1} anticipe ou rejette l'observation X_i, par exemple dans le cas de l'histogramme

$f_{nh}(x) = \frac{1}{nh} \sum_{j=-\infty}^{+\infty} 1_{[h(j-1),hj)]}(x) \{ \sum_{i=1}^{n} 1_{[h(j-1),hj]}(X_i)\}$, si les X_i proviennent d'une loi telle que $\Delta_n = X_{(n)} - X_{(n-1)}$ diverge, où $(X_{(i)})_i$ est la statistique d'ordre associée à $(X_i)_i$ alors la validation croisée fournit un h^* pour lequel f_{nh^*} diverge, car si $X_i = \max(X_1,\ldots,X_n)$ alors pour $h < \Delta_n$ on aura $f^i_{n-1,h}(X_i) = 0$ donc $L_n(h) = 0$ par conséquent le paramètre h^* fourni par la validation croisée doit être supérieur à Δ_n

il est donc divergent.

E.F. SCHUSTER et G.G. GREGORY (1981) ont également démontré la divergence de ce procédé dans le cas du modèle exponentiel et de toute densité dont la queue est très longue.

Le lien entre cette procédure et la fonction de perte de Kullback - Leibler $I(f,g) = \int f(x) \log(f(x)/g(x))\, dx$ a été établi par A.W. BOWMAN (1980) qui a montré que maximiser $L_n(h)$ est équivalent à minimiser $L'_n(h)$ définie par :

$$L'_n(h) = \frac{1}{n} \sum_{i=1}^{n} \operatorname{Log}(f(X_i)/f_{n-1,i}(X_i))$$

cette dernière méthode est connue sous le nom de "Validation croisée de Kullbach - Leibler". On peut trouver un résultat de convergence la concernant dans A.W. BOWMAN (1980), cependant E.F. SCHUSTER et G.G. GREGORY (1981) ont montré qu'elle ne fournit pas toujours des estimations convergentes.

Enfin nous signalons que A.W. BOWMAN (1982) a montré qu'en terme de validation - croisée, la minimisation du risque quadratique intégré revient à minimiser la quantité :

$$L''_n(h) = \frac{1}{n} \sum_{i=1}^{n} \int (f_{n-1,i}(x))^2\, dx - \frac{2}{n} \sum_{i=1}^{n} f_{n-1,i}(X_i) + \frac{2}{n} \sum_{i=1}^{n} f(X_i) - \int f^2(x)\, dx \,.$$

Procédé obtenu également, par M. RUDEMO (1982) ; qui, dans le souci de remédier au problème que pose la minimisation du risque quadratique intégré $J_n(h) = \int f_n^2(x)\, dx - 2 \int f_n(x)\, f(x)\, dx + \int f^2(x)\, dx$ (<u>i.e.</u> la dépendance de $\int f_n(x)\, f(x)\, dx$ du paramètre inconnu f), a proposé de minimiser un estimateur de $J_n(h)$, l'estimateur en question est $L''_n(h)$.

Pour l'étude asymptotique de cette procédure nous renvoyons à C.J. STONE (1984), on peut également consulter Y.S. CHOW, S. GEMAN et L.D. WU (1983).

III - DEMARCHE PROPOSEE POUR ESTIMER f(x).-

La démarche que nous allons proposer utilise les résultats de P. DEHEUVELS (1977-a) ainsi que la technique du conditionnement définie au chapitre IV, de ce fait elle ne sera intéressante que lors de l'estimation en un point fixé x elle se fait en trois étapes.

i) Pré-estimation : On calcule le paramètre de lissage h_{n_1} à l'aide de :

$$h_{n_1} = \left\{ \frac{\int K^2(y)\, dy}{\left(\int y^2\, K(y)\, dy\right)^2} \right\}^{1/5} \cdot \frac{\hat{S}_n}{(3/8\sqrt{\pi})^{1/5}} \cdot \frac{1}{n^{1/5}}$$

où $\hat{S}_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_{(i)} - \bar{X}')^2$, $\bar{X}' = \frac{1}{n} \sum_{i=1}^{n} X_{(i)}$ et $(X_{(i)})\ i \in \overline{1,n}$ est l'échantillon ordonné associé à $(X_i)\ i \in \overline{1,n}$.

Comme pré-estimation de $f(x)$ on prend :

$$f_{n_1}(x) = \frac{1}{n\, h_{n_1}} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h_{n_1}}\right) .$$

ii) Correction de la fenêtre : On reprend l'expression du pas optimal au sens de l'A.M.I.S.E. et on remplace $\int f''^2(x)\, dx$ par $\int f_{n_1}''^2(x)\, dx$, on obtient ainsi un nouveau pas h_{n_2} :

$$h_{n_2} = \left\{ \frac{\int K^2(y)\, dy}{\left(\int y^2 K(y)\, dy\right)^2} \right\}^{1/5} \frac{1}{\left(\int (f''_{n_1}(x))^2\, dx\right)^{1/5}} \cdot \frac{1}{n^{1/5}} \; .$$

On pourra éventuellement reprendre l'approche de R.A. TAPIA et J.R. THOMPSON, choisir h_{n_1} comme pas initial pour cet algorithme et poser h_{n_2} égal au pas fourni par cette méthode.

Le nouvel estimateur de $f(x)$ sera donc :

$$f_{n_2}(x) = \frac{1}{n\, h_{n_2}} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_{n_2}}\right) \; .$$

iii) Correction locale de f_{n_2} *:*

A. BERLINET (1984-a, -b) a montré qu'en conditionnant l'estimateur de Parzen - Rosenblatt, par rapport à la statistique localement exhaustive S, définie par :

$$S = (S_1, \ldots, S_n) \quad \text{où} \quad S_i = \begin{cases} X_i & \text{si} \quad x_i \in U \\ x_o & \text{sinon} \end{cases} ,$$

avec $x_o \in U^c$ et U un ouvert contenant x.

Alors, par application du théorème de RAO-BLACKWELL, l'estimateur obtenu est meilleur au sens du risque quadratique (voir chapitre IV, pour plus de détails).

Comme dernière phase d'estimation on propose donc d'appliquer la technique du conditionnement à f_{n_2}, pour cela on choisira U égal à $]x - h_{n_2}, x + h_{n_2}[$ et pour approximation finale de $f(x)$ on prendra :

$$f_{n_3}(x) = \frac{1}{n\, h_{n_2}} \sum_{i=1}^{n} \{K(\frac{x - X_i}{h_{n_2}})\, 1_U(X_i) + m_n(f_{n_2})\, 1_{U_n^c}(X_i)\}$$

où
$$m_n(f_{n_2}) = \frac{1}{P_n(U^c)} \cdot \int_{U^c} K(\frac{x - y}{h_{n_2}})\, f_{n_2}(y)\, dy$$

et
$$P_n(U^c) = \int_{U^c} f_{n_2}(y)\, dy \;.$$

Remarque : La technique du conditionnement suppose la connaissance d'une pré-estimation de f sur U^c , dans notre cas, elle est égale à f_{n_2} .

CHAPITRE II

ETUDE DE L'ESTIMATEUR DE HERMAN J. BIERENS.

I - *INTRODUCTION*.-

H.J. BIERENS (1983) a introduit une nouvelle classe d'estimateurs à noyau de la densité d'une loi de probabilité sur $\mathbb{R}^k$, $k \geq 1$, l'originalité de ces estimateurs que la fenêtre usuelle h_n est fonction des observations et que ses moments d'ordre 1 et 2 coïncident avec les moments empiriques de la loi à estimer.

L'étude de H.J. BIERENS a été faite pour le noyau normal et la convergence en probabilité uniforme uniquement, l'objet de ce chapitre est de généraliser la forme de cet estimateur et d'établir la convergence p.s. pour des noyaux autres que le noyau normal. Nous allons donc rappeler en quoi consiste cet estimateur et en même temps indiquer la généralisation que nous en proposons :

Soit P une loi de probabilité définie sur $\mathbb{R}^k$, $k \geq 1$, dont une version de la densité par rapport à la mesure de Lebesgue sera notée f.
On suppose que les moments d'ordre 1 et 2 de P existent, étant donné un échantillon $(X_i)_{i=1,\dots,n}$ issu de P , l'estimateur de Parzen-Rosenblatt de f au point $x \in \mathbb{R}^k$ s'écrit :

$$f_n(x) = \frac{1}{n\, h_n^k} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n}\right)$$

où h_n est une suite de réels strictement positifs et telle que $h_n \xrightarrow[n\to\infty]{} 0$

K est une fonction de $\mathbb{R}^k$ dans $\mathbb{R}$.

Il est utile, dans beaucoup de cas, d'obtenir un estimateur strict de la densité f, c'est-à-dire imposer à f_n d'être une densité sur $\mathbb{R}^k$, ce qui est réalisé si K vérifie :

a) $K(z) \geq 0 \quad \forall z \in \mathbb{R}^k , \int_{\mathbb{R}^k} K(z)\, dz = 1$.

Puisque nous avons supposé que les moments $\mu_1 = \int_{\mathbb{R}^k} x\, f(x)\, dx$ et $\mu_2 = \int_{\mathbb{R}^k} x\, x^t\, f(x)\, dx$, existent, il est tout à fait raisonnable d'imposer à f_n d'avoir des moments d'ordre 1 et 2 qui estiment μ_1 respectivement μ_2 , dans ce but H.J. BIERENS a supposé que f_n vérifie les deux conditions suivantes :

b) $\bar{X} = \int_{\mathbb{R}^k} x\, f_n(x)\, dx$ où $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$

et c) $\frac{1}{n} \sum_{i=1}^{n} X_i X_i^t = \int_{\mathbb{R}^k} x\, x^t\, f_n(x)\, dx$

(cette dernière égalité est une égalité matricielle).

Or ces deux conditions sont incompatibles car b) entraîne que $\int_{\mathbb{R}^k} x\, K(x)\, dx = 0$ et c) implique que :

$$h_n\, \bar{X} \int_{\mathbb{R}^k} x^t\, K(x)\, dx + h_n \left(\int_{\mathbb{R}^k} x\, K(x)\, dx\right) \bar{X}^t + h_n^2 \int_{\mathbb{R}^k} x\, x^t\, K(x)\, dx = 0$$

donc si elles sont simultanément vérifiées alors $\int_{\mathbb{R}^k} x\, x^t\, K(x)\, dx = 0$, ce qui est impossible puisque la condition a) entraîne que $\int_{\mathbb{R}^k} x\, x^t\, K(x)\, dx$ est une matrice définie positive, pour remédier à cela H.J. BIERENS a introduit une légère modification sur f_n en considérant le nouvel estimateur :

$$f_{n_1}(x) = \frac{1}{n\, h_n^k} \sum_{j=1}^{n} K\left(\frac{x - (1-\beta_n)\, X_j}{h_n}\right) ,$$

où β_n est une suite de réels, telle que $\beta_n \xrightarrow[n\to\infty]{} 0$.

A partie de b) il a obtenu $\int_{\mathbb{R}^k} x\, K(x)\, dx = \frac{\beta_n}{h_n} \bar{X}$, ensuite la condition c) s'écrit :

$$\int_{\mathbb{R}} x\, x^t\, K(x)\, dx = \frac{2\,\beta_n - \beta_n^2}{h_n^2} \left(\frac{1}{n} \sum_{j=1}^{n} X_j\, X_j^t - \bar{X}\, \bar{X}^t\right) + \int_{\mathbb{R}^k} x\, K(x)\, dx \int_{\mathbb{R}^k} x^t\, K(x)\, dx$$

ainsi pour que les conditions a), b) et c) soient vérifiées, la fonction K doit être une densité, de moyenne $\frac{\beta_n}{h_n} \bar{X}$ et dont la matrice variance - covariance est égale à $\frac{2\,\beta_n - \beta_n^2}{h_n^2}\, S_n$ où $S_n = \frac{1}{n} \sum_{j=1}^{n} (X_j - \bar{X})\, (X_j - \bar{X})^t$

afin que cette matrice soit définie positive il faut choisir β_n telle que $2\,\beta_n - \beta_n^2 > 0$ <u>i.e.</u> prendre β_n dans $]0,2[$ ce qui entraîne que : $(2\,\beta_n - \beta_n^2) \in]0,1]$.

Il est clair que si on pose :

$$K(x) = \frac{K_o\left[\left(\frac{2\,\beta_n - \beta_n^2}{h_n^2}\, S_n\right)^{-1/2} \left(x - \frac{\beta_n}{h_n} \bar{X}\right)\right]}{\left|\mathrm{Det}\left(\frac{2\,\beta_n - \beta_n^2}{h_n^2}\, S_n\right)\right|^{1/2}} \qquad (1)$$

où K_o est une fonction de $\mathbb{R}^k$ dans $\mathbb{R}$ telle que :

$$K_o(x) \geq 0\ ,\ \forall x \in \mathbb{R}^k\ ,\ \int_{\mathbb{R}^k} K_o(x)\, dx = 1\ ,\ \int_{\mathbb{R}^k} x\, K_o(x)\, dx = 0\ ,$$

$$\int_{\mathbb{R}^k} x\, x^t\, K_o(x)\, dx = I$$

avec I = matrice identité de $\mathbb{R}^k$, alors les conditions a), b) et c) seront satisfaites par l'estimateur f_{n_2} obtenu après transformation de f_{n_1} à l'aide de la relation (1) i.e.

$$f_{n_2}(x) = \frac{1}{n(\sqrt{2\beta_n-\beta_n^2})^k |\mathrm{Det} S_n|^{1/2}} \cdot \sum_{j=1}^{n} K_o\left(S^{-1/2}\left(\frac{x-(1-\beta_n)X_j-\beta_n\bar{X}}{\sqrt{2\beta_n-\beta_n^2}}\right)\right)$$

on remarquera que le paramètre de lissage initial h_n a disparu et que son rôle est joué par β_n. Pour alléger les notations on posera $\gamma_n = \sqrt{2\beta_n - \beta_n^2}$ et on notera le noyau K au lieu de K_o :

$$f_{n_2}(x) = \frac{1}{n\gamma_n^k |\mathrm{Det} S_n|^{1/2}} \sum_{j=1}^{n} K\left(S_n^{-1/2}\left(\frac{x-\sqrt{1-\gamma_n^2}.X_j-(1-\sqrt{1-\gamma_n^2})\bar{X}}{\gamma_n}\right)\right)$$

avec $\gamma_n \in]0,1]$.

Suivant les informations dont on dispose sur le modèle à étudier ou pour des raisons de robustesse, il est parfois préférable de prendre des estimateurs autres que la moyenne empirique $\bar{X}$ ou la variance S_n, on choisirait, par exemple, la moyenne α-tronquée ou α-winsorizée qui sont définies respectivement par :

$$\bar{X}_\alpha = \frac{1}{n-2[n\alpha]} \sum_{j=1+[n\alpha]}^{n-[n\alpha]} X_{(j)}$$

$$\bar{X}_{n,\alpha} = \frac{1}{(1-2\alpha)n} \left(p \,.\, X_{(\ell)} + \sum_{j=\ell+1}^{n-\ell} X_{(j)} + p .\, X_{(n-\ell+1)}\right)$$

Si $\quad \ell-1 \leq (n-1)\,\alpha < n\alpha \leq \ell \quad$, avec $p = \ell - n\alpha$

$$= \frac{1}{(1-2\alpha)n}\left[(1-p)\, X_{(\ell+1)} + \sum_{j=\ell+2}^{n-\ell-1} X_{(j)} + (1-p)\, X_{n-\ell}\right]$$

où $(X_{(j)})_{j\in\overline{1,n}}$ est la statistique d'ordre, α est un réel appartenant à $]0,1/2[$ et $[\cdot]$ désigne la partie entière d'un réel, on peut également appliquer la technique du Jackknife à ces deux estimateurs, pour obtenir une estimation de la moyenne, encore plus robuste ; pour plus de détails nous renvoyons à P.J. HUBER. (1981).

Quant à la variance empirique $S_n = \frac{1}{n} \sum_{j=1}^{n} (X_j - \bar{X})(X_j - \bar{X})^t$, souvent on lui préfère $S'_n = \frac{n}{n-1} S_n$, qui dans certains cas est optimal sans biais, on peut notamment proposer la variance α-tronquée ou α-winsorizée (cf. P.J. HUBER).

Pour ces raisons nous désignons par m_{n_1} un estimateur de μ_1 et par m_{n_2} un estimateur de μ_2 et afin de donner une forme plus générale à l'estimateur de H.J. BIERENS nous introduisons la famille de fonctions auxiliaires $\{\psi_n^j , j \in \overline{1,n}\}$ définies de $\mathbb{R}^k$ dans $\mathbb{R}^k$ et considérons le nouvel estimateur f_{n_3} défini par

$$f_{n_3}(x) = \frac{1}{n\, h_n^k} \sum_{j=1}^{n} K\left(\frac{x - \psi_n^j(X_j)}{h_n}\right)$$

la condition a) reste inchangée pour cet estimateur, par contre les conditions b) et c) s'écriront respectivement :

$$\text{b')} \quad \int_{\mathbb{R}^k} x\, f_{n_3}(x)\, dx = m_{n_1}$$

$$\text{c')} \quad \int_{\mathbb{R}^k} x\, x^t\, f_{n_3}(x)\, dx = m_{n_2}$$

nous obtenons, à partir de b'),

$$(1) \qquad h_n \int_{\mathbb{R}^k} zK(z)\, dz = m_{n_1} - \bar{\psi}_n \quad \text{où} \quad \bar{\psi}_n = \frac{1}{n} \sum_{j=1}^{n} \psi_n^j(X_j)$$

ensuite c') nous donne :

$$h_n^2 \int z\, z^t\, K(z)\, dz = m_{n_2} - \frac{1}{n} \sum_{j=1}^{n} \varphi_n^j(X_j)\, (\varphi_n^j(X_j))^t - \bar{\varphi}_n.(m_{n_1} - \bar{\varphi}_n)^t - (m_{n_1} - \bar{\varphi}_n).\bar{\varphi}_n^t$$

On en déduit que : sous b') et c') K vérifie

$$(2) \qquad \int_{\mathbb{R}^k} z\, z^t\, K(z)\, dz - \int_{\mathbb{R}^k} z\, K(z)\, dz \int_{\mathbb{R}^k} z^t\, K(z)\, dz = \frac{1}{h_n^2}\, [m_{n_2} - m_{n_1} m_{n_1}^t - \varphi_n^2]$$

où

$$\varphi_n^2 = \frac{1}{n} \sum_{j=1}^{n} (\varphi_n^j(X_j) - \bar{\varphi}_n)\, (\varphi_n^j(X_j) - \bar{\varphi}_n)^t \quad .$$

Pour que (2) ait un sens il faut que la matrice $(m_{n_2} - m_{n_1} m_{n_1}^t - \varphi_n^2)$ soit définie positive, car la quantité $\{\int_{\mathbb{R}^k} z\, z^k\, K(z)\, dz - \int_{\mathbb{R}^k} z\, K(z)\, dz \int_{\mathbb{R}^k} z^t\, K(z)\, dz\}$ est la matrice variance-covariance du noyau K.

Dans le cas de l'estimateur de H.J. Bierens $(m_{n_2} - m_{n_1} m_{n_1}^t - \varphi_n^2)$ vaut $\frac{1}{h_n^2}\, \beta_n(2-\beta_n)\, S_n$, qui est p.s. une matrice définie positive, car β_n est pris dans $]0,1]$. C'est d'ailleurs, la raison qui l'a poussé à transformer l'estimateur de Parzen-Rosenblatt $f_n(x)$ en $f_{n_2}(x)$. Par ailleurs si $m_{n_1} = \bar{X}$ et $m_{n_2} = \frac{1}{n} \sum_{j=1}^{n} X_j\, X_j^t$ alors dans le cas où la matrice de variance-covariance des $(\varphi_n^j(X_j))_j$ est dominée par celle de l'échantillon, la matrice $(m_{n_2} - m_{n_1} m_{n_1}^t - \varphi_n^2)$ sera définie positive. Elle l'est également dans le cas où $\varphi_n^j(X_j) = X_j \quad \forall j \in \overline{1,n}$ et $S_n = \frac{1}{n} \sum_{j=1}^{n} X_j\, X_j^t - \bar{X}\, X^t$ est strictement dominée par $m_{n_2} - m_{n_1} m_{n_1}^t$. Enfin pour des commodités d'écriture on notera $M_n = m_{n_2} - m_{n_1} m_{n_1}^t - \varphi_n^2$, ainsi on pose

$$K(x) = \frac{1}{|\mathrm{Det}(M_n/h_n^2)|^{1/2}}\, K_o\left(\left(\frac{M_n}{h_n^2}\right)^{-1/2} \left(x - \frac{m_{n_1} - \bar{\varphi}_n}{h_n}\right)\right)$$

avec K_o une fonction de $\mathbb{R}^k$ dans $\mathbb{R}$ telle que :

$$K_o(x) \geq 0 \ , \ \int_{\mathbb{R}^k} K_o(x)\, dx = 1 \ , \ \int_{\mathbb{R}^k} x\, K_o(x)\, dx = 0 \ ,$$
$$\int_{\mathbb{R}^k} x\, x^t\, K_o(x)\, dx = I \ .$$

Alors le noyau K vérifiera les conditions a), b') et c') et f_{n_3} s'écrira

$$f_{n_3}(x) = \frac{1}{n\,|\mathrm{Det}\ M_n|^{1/2}} \sum_{j=1}^{n} K_o(M_n^{-1/2}\,(x - \psi_n^j(X_j) - m_{n_1} + \bar{\psi}_n)) \ .$$

On retrouve l'estimateur de H.J. Bierens $f_{n_2}(x)$ lorsque $\psi_n^j(X_j) = (1-\beta_n)\, X_j$ et $m_{n_1} = \bar{X}$, $m_{n_2} = \frac{1}{n} \sum_{j=1}^{n} X_j\, X_j^t$.

On se propose, dans ce qui suit, d'établir la convergence uniforme p.s. de f_{n_2} vers f, le même résultat de convergence, peut être obtenu, pour f_{n_3} moyennant des conditions restrictives sur la famille $\{\psi_n^j\ ,\ j \in \overline{1,n}\}$. Cela fera l'objet de développements ultérieurs.

II - CONVERGENCE UNIFORME PRESQUE SURE.-

Nous désignerons par :

F : la fonction de répartition associée à f.

F_n : la fonction de répartition empirique.

V_K : la variation totale de K .

$$\mu_1 = \int_{\mathbb{R}^k} x\, f(x)\, dx \ , \quad \mu_2 = \int_{\mathbb{R}^k} x\, x^t\, f(x)\, dx \ , \quad V = \mu_2 - \mu_1\, \mu_1^t$$

$$f^*_{n_2}(x) = \frac{1}{\gamma_n^k |\mathrm{Det}\ S_n|^{1/2}} \int_{\mathbb{R}^k} K\left(S_n^{-1/2}\left(\frac{x - \sqrt{1-\gamma_n^2}\, y - (1-\sqrt{1-\gamma_n^2})\bar{X}}{\gamma_n} \right)\right) f(y)\, dy \ .$$

$$\bar{f}_{n_2}(x) = \frac{1}{\gamma_n^k |\mathrm{Det}\ V|^{1/2}} \int_{\mathbb{R}^k} K\left(V^{-1/2}\left(\frac{x - \sqrt{1-\gamma_n^2}\, y - (1-\sqrt{1-\gamma_n^2})\, \mu_1}{\gamma_n} \right)\right) f(y)\, dy \ .$$

Théorème 2.1.- Si f est uniformément continue, bornée sur $\mathbb{R}^k$

K est une densité de probabilité à variation bornée et telle que

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

$(\gamma_n)_{n\geq 1}$ est une suite de réels dans $]0,1]$, telle que :

$$\gamma_n \xrightarrow[n\to\infty]{} 0 \qquad \frac{n\gamma_n^{2k}}{\mathrm{Log}\ n} \xrightarrow[n\to\infty]{} +\infty \ .$$

Alors 1) l'estimateur f_{n_2} défini par :

$$f_{n_2}(x) = \frac{1}{n\gamma_n^k |\mathrm{Det}\ S_n|^{1/2}} \sum_{j=1}^{n} K\left(S_n^{-1/2}\left(\frac{x - \sqrt{1-\gamma_n^2}\, X_j - (1-\sqrt{1-\gamma_n^2})\, \bar{X}}{\gamma_n} \right)\right)$$

est une densité de probabilité sur $\mathbb{R}^k$ telle que :

$$\int_{\mathbb{R}^k} x\, f_{n_2}(x)\, dx = \bar{X} \quad \text{et} \quad \int_{\mathbb{R}^k} x\, x^t\, f_{n_2}(x)\, dx = \frac{1}{n} \sum_{j=1}^{n} X_j\, X_j^t$$

2) $$\mathrm{Sup}_{x\in\mathbb{R}^k} |f_{n_2}(x) - f(x)| \xrightarrow[n\to\infty]{p.s.} 0$$

la démonstration de ce théorème fait intervenir les lemmes 2.1, 2.2 et 2.3 suivants :

Lemme 2.1.- Si f est uniformément continue bornée sur $\mathbb{R}^k$.

K une fonction à variation bornée sur $\mathbb{R}^k$.

$(\gamma_n)_n$ est une suite de réels dans $]0,1]$, telle que

$$\gamma_n \xrightarrow[n\to\infty]{} 0 \quad \text{et} \quad \frac{n\,\gamma_n^{2k}}{\text{Log } n} \xrightarrow[n\to\infty]{} +\infty$$

Alors
$$\sup_x |f_{n_2}(x) - f^*_{n_2}(x)| \xrightarrow[n\to\infty]{p.s.} 0 \tag{1.0}$$

Preuve : Notons M l'ensemble des matrices $(k \times k)$ définies positives dont le déterminant est plus grand que $\alpha > 0$ où α est défini par :

$$|\text{Det } V| > \alpha \tag{1.1}$$

Nous avons
$$\sup_x |f_{n_2}(x) - f^*_{n_2}(x)| \leqslant \sup_x |f_{n_2}(x) - f^*_{n_2}(x)| \; 1_M(S_n)$$
$$+ \sup_x |f_{n_2}(x) - f^*_{n_2}(x)| \; 1_{M^c}(S_n)$$

d'après la loi forte des grands nombres $S_n \xrightarrow[n\to\infty]{p.s.} V$, donc par définition de M l'événement $\{S_n \in M\}$ est p.s. réalisé à partir d'un certain rang, ainsi pour avoir (1.0) il suffit de montrer que

$$1_M(S_n) \cdot \sup_x |f_{n_2}(x) - f^*_{n_2}(x)| \xrightarrow[n\to\infty]{p.s.} 0 \tag{1.2}$$

en effet

$$\sup_x |f_{n_2}(x)-f^*_{n_2}(x)| =$$

$$\frac{1}{\gamma_n^k} \cdot \frac{1}{|\mathrm{Det} S_n|^{1/2}} \sup_x \left| \int_{\mathbb{R}^k} K\left(S_n^{-1/2}\left(\frac{x-\sqrt{1-\gamma_n^2}y-(1-\sqrt{1-\gamma_n^2})\bar{X}}{\gamma_n}\right)\right) \left(dF_n(y)-dF(y)\right)\right|$$

$$\leq \frac{1}{\gamma_n^k} \frac{1}{|\mathrm{Det} S_n|^{1/2}} \sup_x |F_n(x)-F(x)| \cdot \int_{\mathbb{R}^k} \left|dK\left(S_n^{-1/2}\left(\frac{x-\sqrt{1-\gamma_n^2}y-(1-\sqrt{1-\gamma_n^2})\bar{X}}{\gamma_n}\right)\right)\right|$$

or

$$\int_{\mathbb{R}^k} \left|dK\left(S_n^{-1/2}\left(\frac{x-\sqrt{1-\gamma_n^2}y-(1-\sqrt{1-\gamma_n^2})\bar{X}}{\gamma_n}\right)\right)\right| = V_K$$

nous avons donc : $\sup_x |f_{n_2}(x)-f^*_{n_2}(x)| \leq \dfrac{V_K}{\gamma_n^k |\mathrm{Det} S_n|^{1/2}} \cdot \sup_x |F_n(x)-F(x)|$ (1.3)

ce qui entraîne que

$$\sup_x |f_{n_2}(x)-f^*_{n_2}(x)| \; 1_M(S_n) \leq \frac{V_K}{\gamma_n^k \cdot \alpha} \sup_x |F_n(x)-F(x)| \tag{1.4}$$

or il est bien connu que $\sup_x |F_n(x)-F(x)| = O\left(\frac{\mathrm{Log}\ n}{n}\right)^{1/2}$ p.s. (1.5)

(voir D. BOSQ, (1985) chapitre II, Estimation fonctionnelle asymptotique)

<u>i.e.</u> il existe une constante $\theta > 0$ telle que

$$\sup_x |F_n(x)-F(x)| \leq \theta \cdot \left(\frac{\mathrm{Log}\ n}{n}\right)^{1/2} \quad \text{p.s.},$$

nous obtenons donc l'inégalité suivante :

$$\sup_x |f_{n_2}(x)-f^*_{n_2}(x)| \; 1\,(S_n) \leq \frac{\theta}{\alpha} V_K \cdot \left(\frac{\mathrm{Log}\ n}{n\ \gamma_n^{2k}}\right)^{1/2} \quad \text{p.s.}$$

ce qui permet de conclure, puisque par hypothèse $\dfrac{n\ \gamma_n^{2k}}{\mathrm{Log}\ n} \xrightarrow[n\to\infty]{} \infty$. ■

Remarque : On peut obtenir la convergence presque complète de $\operatorname{Sup}_x |f_{n_2}(x)-f^*_{n_2}(x)|\ 1_M(S_n)$ grâce à l'inégalité de A. DVORETZKY, J. KIEFER et J. WOLFOWITZ, (1956) :

$$\forall\lambda > 0\ ,\ \exists c > 0\ ,\ \beta \in]0,2[\ :\ \Pr\{\operatorname{Sup}_x |F_n(x)-F(x)| > \lambda/\sqrt{n}\} \leq c\, e^{-\beta\lambda^2}$$

et à la condition : $\forall\mu > 0 \ \sum_{n=1}^{\infty} e^{-\mu n\gamma_n^{2k}} < \infty$, qui est plus faible que la condition $n\,\gamma_n^{2k}/\operatorname{Log} n \xrightarrow[n\to\infty]{} +\infty$.

Lemme 2.2.- Si f est uniformément continue bornée sur $\mathbb{R}^k$, K est une densité de probabilité sur $\mathbb{R}^k$ et $(\gamma_n)_n$ une suite réelle à valeurs dans $]0,\gamma]$ où $\gamma \in]0,1[$.

Alors $$\operatorname{Sup}_x |f^*_{n_2}(x) - \bar{f}_{n_2}(x)| \xrightarrow[n\to\infty]{p.s.} 0\ .$$

Preuve : Nous avons :

$$\operatorname{Sup}_x |f^*_{n_2}(x) - \bar{f}_{n_2}(x)| \leq$$

$$\frac{1}{(\sqrt{1-\gamma_n^2})^k} \operatorname{Sup}_x \int_{\mathbb{R}^k} K(y) \left| f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\bar{X}-\gamma_n S_n^{1/2} y}{\sqrt{1-\gamma_n^2}}\right) - f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu-\gamma_n V^{1/2} y}{\sqrt{1-\gamma_n^2}}\right)\right| dy \qquad (2.0)$$

ceci d'après la définition de $f^*_{n_2}$ et de $\bar{f}_{n_2}$, d'autre part on a : pour tout $y \in \mathbb{R}^k$:

$$\left\|\frac{x-(1-\sqrt{1-\gamma_n^2})\bar{X}-\gamma_n S_n^{1/2} y}{\sqrt{1-\gamma_n^2}} - \frac{x-(1-\sqrt{1-\gamma_n^2})\mu-\gamma_n V^{1/2} y}{\sqrt{1-\gamma_n^2}}\right\|$$

$$\leq \frac{1-\sqrt{1-\gamma_n^2}}{\sqrt{1-\gamma_n^2}}\,||\bar{X}-\mu|| + \frac{\gamma_n ||y||}{\sqrt{1-\gamma_n^2}}\,||S_n^{1/2}-V^{1/2}||^* \qquad (2.1)$$

où $||.||^*$ est la norme matricielle subordonnée à la norme vectorielle $||.||$ i.e.

$$||W||^* = \operatorname*{Sup}_{z\in\mathbb{R}^k} \frac{||W.z||}{||z||} \quad \text{où} \quad W \quad \text{est une matrice} \quad (k \times k),$$

la loi forte des grands nombres appliquée à l'échantillon $(X_i)_{i=1,\dots,n}$ nous permet d'affirmer que $||\bar{X}-\mu|| \xrightarrow[n\to\infty]{p.s.} 0$ et $||S_n^{1/2}-V^{1/2}||^* \xrightarrow[n\to\infty]{p.s.} 0$,

ce qui entraîne que (2.1) converge p.s. vers 0 lorsque $n \to +\infty$, ainsi l'uniforme continuité de f sur $\mathbb{R}^k$ assure la convergence p.s., vers 0, de

$$\operatorname*{Sup}_{x} \left| f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\bar{X}-\gamma_n S_n^{1/2} y}{\sqrt{1-\gamma_n^2}}\right) - f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu-\gamma_n V^{1/2} y}{\sqrt{1-\gamma_n^2}}\right)\right|$$

d'autre part, l'hypothèse f bornée sur $\mathbb{R}^k$ implique que :

$$\forall y \in \mathbb{R}^k \; g_n(y) = K(y) \operatorname*{Sup}_{x} \left| f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\bar{X}-\gamma_n S_n^{1/2} y}{\sqrt{1-\gamma_n^2}}\right) - f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu-\gamma_n V^{1/2} y}{\sqrt{1-\gamma_n^2}}\right)\right|$$

$$\leq 2M\, K(y) \quad \text{avec} \quad M = \operatorname*{Sup}_{x} f(x) \; .$$

Nous sommes donc en mesure d'appliquer le théorème de la convergence dominée à la suite $\left(g_n(y)\right)_n$ et on obtient donc

$$\operatorname*{Sup}_{x} \int_{\mathbb{R}^k} K(y) \left| f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\bar{X}-\gamma_n S_n^{1/2} y}{\sqrt{1-\gamma_n^2}}\right) - f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu-\gamma_n V^{1/2} y}{\sqrt{1-\gamma_n^2}}\right)\right| dy \xrightarrow[n]{p.s.} 0$$

on en déduit que (2.0) converge p.s. vers 0 , puisque $\sqrt{1-\gamma_n^2}$ reste bornée. ■

Lemme 2.3.- Si f est uniformément continue, bornée sur $\mathbb{R}^k$, K une densité de probabilité sur $\mathbb{R}^k$ et $\gamma_n \xrightarrow[n\to\infty]{} 0$. Alors :

i) $$\operatorname*{Sup}_x \left| \bar{f}_{n_2}(x) - \frac{1}{(\sqrt{1-\gamma_n^2})^k} f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu}{\sqrt{1-\gamma_n^2}}\right) \right| \xrightarrow[n\to\infty]{} 0$$

ii) $$\operatorname*{Sup}_x \left| f(x) - \frac{1}{(\sqrt{1-\gamma_n^2})^k} f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu}{\sqrt{1-\gamma_n^2}}\right) \right| \xrightarrow[n\to\infty]{} 0$$

Preuve : i) K étant une densité de probabilité sur $\mathbb{R}^k$, nous avons :

$$\operatorname*{Sup}_x \left| \bar{f}_{n_2}(x) - \frac{1}{(\sqrt{1-\gamma_n^2})^k} f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu}{\sqrt{1-\gamma_n^2}}\right) \right| \leqslant$$

$$\frac{1}{(\sqrt{1-\gamma_n^2})^k} \operatorname*{Sup}_x \int_{\mathbb{R}^k} K(y) \left| f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu-\gamma_n V^{1/2} y}{\sqrt{1-\gamma_n^2}}\right) - f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu}{\sqrt{1-\gamma_n^2}}\right) \right| dy \quad .$$

Donc, pour tout $\delta > 0$

$$\operatorname*{Sup}_x \left| \bar{f}_{n_2}(x) - \frac{1}{(\sqrt{1-\gamma_n^2})^k} f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu}{\sqrt{1-\gamma_n^2}}\right) \right| \leqslant$$

$$\begin{aligned} &\frac{1}{(\sqrt{1-\gamma_n^2})^k} \operatorname*{Sup}_x \int_{\Omega_n} K(y) \left| f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu-\gamma_n V^{1/2} y}{\sqrt{1-\gamma_n^2}}\right) - f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu}{\sqrt{1-\gamma_n^2}}\right) \right| dy \\ &+ \frac{1}{(\sqrt{1-\gamma_n^2})^k} \operatorname*{Sup}_x \int_{\Omega_n^c} K(y) \left| f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu-\gamma_n V^{1/2} y}{\sqrt{1-\gamma_n^2}}\right) - f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu}{\sqrt{1-\gamma_n^2}}\right) \right| dy \end{aligned} \qquad (3.2)$$

avec $\quad \Omega_n^c = \left\{ y \in \mathbb{R}^k : \left\| \dfrac{\gamma_n V^{1/2} y}{\sqrt{1-\gamma_n^2}} \right\| > \delta \right\}$

la première quantité à droite de (3.2) est majorée par :

$$\frac{1}{(\sqrt{1-\gamma_n^2})^k} \operatorname*{Sup}_x \operatorname*{Sup}_{\|z\| \leq \delta} |f(x-z)-f(x)| . \int_{\Omega_n} K(y)dy \leq \frac{1}{(\sqrt{1-\gamma_n^2})^k} \operatorname*{Sup}_x \operatorname*{Sup}_{\|z\| \leq \delta} |f(x-z)-f(x)| \quad (3.3)$$

car K est une densité de probabilité.

la deuxième quantité est majorée par :

$$2M \frac{1}{(\sqrt{1-\gamma_n^2})^k} \int_{\Omega_n^c} K(y) \; dy \quad (3.4)$$

Pour établir (i), il suffit d'injecter (3.3) et (3.4) dans (3.2), de remarquer que $\Omega_n^c \subset \left\{ y : \|y\| \geq \dfrac{\sqrt{1-\gamma_n^2}}{\gamma_n} . \dfrac{\delta}{\|V^{1/2}\|^*} \right\}$ et de faire tendre n vers $+\infty$ pour δ fixé puis de faire tendre δ vers 0 .

Pour montrer (ii), on utilise la majoration

$$\operatorname*{Sup}_x \left| f(x) - \frac{1}{(\sqrt{1-\gamma_n^2})^k} f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu}{\sqrt{1-\gamma_n^2}} \right) \right| \leq \left| 1 - \frac{1}{(\sqrt{1-\gamma_n^2})^k} \right| . \operatorname*{Sup}_x |f(x)| +$$

$$\frac{1}{(\sqrt{1-\gamma_n^2})^k} \operatorname*{Sup}_x \left| f(x) - f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu}{\sqrt{1-\gamma_n^2}} \right) \right|$$

et on conclut, en utilisant, la convergence de γ_n vers 0 et le fait que f est uniformément continue bornée sur $\mathbb{R}^k$. ■

Preuve du théorème 2.1. :

(1) est vérifié par construction de f_{n_2} , pour avoir (2), il suffit d'appliquer les lemmes 2.1, 2.2 et 2.3 aux termes de droite de l'inégalité :

$$\sup_x |f_{n_2}(x)-f(x)| \leq \sup_x |f_{n_2}(x)-f^*_{n_2}(x)|$$

$$+ \sup_x |f^*_{n_2}(x)-\bar{f}_{n_2}(x)| + \sup_x \left|\bar{f}_{n_2}(x) - \frac{1}{(\sqrt{1-\gamma_n^2})^k} f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu}{\sqrt{1-\gamma_n^2}}\right)\right|$$

$$+ \sup_x \left|\frac{1}{(\sqrt{1-\gamma_n^2})^k} f\left(\frac{x-(1-\sqrt{1-\gamma_n^2})\mu}{\sqrt{1-\gamma_n^2}}\right) - f(x)\right| \quad \blacksquare$$

Remarques :

1°) Afin de mettre en évidence le fait que f_{n_2} dépend d'un paramètre γ on écrira $f_{n_2}(x,\gamma)$ au lieu de $f_{n_2}(x)$, $f_{n_2}(x,\gamma)$ désignera donc

$$f_{n_2}(x,\gamma) = \frac{1}{n\gamma^k |\text{Det } S_n|^{1/2}} \sum_{j=1}^{n} K\left(S_n^{-1/2}\left(\frac{x-\sqrt{1-\gamma^2}X_j-(1-\sqrt{1-\gamma^2})\bar{X}}{\gamma}\right)\right).$$

Si K est un noyau de Parzen-Rosenblatt sur $\mathbb{R}^k$ (<u>i.e.</u> K une densité telle que $||x||^k K(x) \xrightarrow[||x|| \to +\infty]{} 0$) et $K(0) > 0$ alors $f_{n_2}(x,\gamma) \to \infty$ si $x \in \{X_1,\ldots,X_n\}$ et $\gamma \to 0$, n étant fixé, car si $\exists\, i_o \in \overline{1,n}$: $x = X_{i_o}$ alors

$$f_{n_2}(X_{i_o},\gamma) = \frac{1}{n\gamma^k |\text{Det } S_n|^{1/2}} \sum_{\substack{j=1 \\ j\neq i_o}}^{n} K\left(S_n^{-1/2}\left(\frac{X_{i_o}-\sqrt{1-\gamma^2}X_j-(1-\sqrt{1-\gamma^2})\bar{X}}{\gamma}\right)\right)$$

$$+ \frac{1}{n\gamma^k |\text{Det } S_n|^{1/2}} K\left(S_n^{-1/2}(X_{i_o}-\bar{X}) \frac{(1-\sqrt{1-\gamma^2})}{\gamma}\right) .$$

or $\sqrt{1-\gamma^2} \underset{o}{\sim} 1 - \frac{\gamma^2}{2} \Longrightarrow \frac{1-\sqrt{1-\gamma^2}}{\gamma} \xrightarrow[\gamma\to 0]{} 0$ par conséquent

$$\frac{1}{n\gamma^k |\text{Det } S_n|^{1/2}} \, K\left(\frac{1-\sqrt{1-\gamma^2}}{\gamma} \, S_n^{-1/2} \, (X_{i_o} - \bar{X}) \right) \xrightarrow[\gamma\to 0]{} \infty$$

et pour tout $j \neq i_o$ nous avons :

$$\frac{1}{\gamma^k |\text{Det } S_n|^{1/2}} \, K\left(S_n^{-1/2} \left(\frac{X_{i_o} - \sqrt{1-\gamma^2} X_j - (1-\sqrt{1-\gamma^2})\bar{X}}{\gamma} \right)\right) =$$

$$\frac{1}{||S_n^{-1/2}(X_{i_o} - \sqrt{1-\gamma^2}X_j - (1-\sqrt{1-\gamma^2})\bar{X}||^k} \cdot \frac{||S_n^{-1/2}(X_{i_o} - \sqrt{1-\gamma^2}X_j - (1-\sqrt{1-\gamma^2})\bar{X}||^k}{\gamma^k \, |\text{Det } S_n|^{1/2}}$$
$$K\left(S_n^{-1/2} \left(\frac{X_{i_o} - \sqrt{1-\gamma^2}X_j - (1-\sqrt{1-\gamma^2})\bar{X}}{\gamma} \right)\right)$$

qui converge vers 0 lorsque γ tend vers 0 , car K est un noyau de Parzen-Rosenblatt et $||S_n^{-1/2} \, (X_{i_o} - \sqrt{1-\gamma^2}X_j - (1-\sqrt{1-\gamma^2}) \, \bar{X}||$ converge vers $||S_n^{-1/2}(X_{i_o} - X_j)||$ lorsque $\gamma \to 0$. On obtient donc la conclusion

$$f_{n_2}(x,\gamma) \xrightarrow[\gamma\to 0]{} \begin{cases} \infty & \text{si } \quad x \in \{X_1,\dots,X_n\} \\ \\ 0 & \text{sinon} \end{cases} .$$

Résultat qu'on peut interpréter de la manière suivante :

" l'estimateur $f_{n_2}(.,\gamma)$ a tendance à avoir des pics lorsque l'échantillon $(X_i)_{i=1,\dots,n}$ contient des observations qui sont très voisines du point où l'on veut estimer la densité, la fenêtre γ est proche de 0 et que le noyau K est un noyau de Parzen-Rosenblatt vérifiant $K(0) \neq 0$. (Le noyau normal satisfait à ces conditions) ".

Pour remédier à cela H.J. BIERENS a modifié f_{n_2} comme suit :

$$\tilde{f}_{n_2}(x,\gamma) = \frac{1}{n\gamma^k |\mathrm{Det}\ S_n|^{1/2}} \cdot \sum_{j=1}^{n} 1_{\{X_j \neq x\}}\ K\left(S_n^{-1/2}\left(\frac{x-\sqrt{1-\gamma^2}X_j-(1-\sqrt{1-\gamma^2})\bar{X}}{\gamma}\right)\right)$$

l'estimateur $f_{n_2}(.,\gamma)$ vérifie bien les conditions, a, b') et c') et évite d'avoir des pics lorsque x coïncide avec une observation X_{i_o}, $i_o \in \overline{1,n}$ et le paramètre γ est proche de 0, en fait $\tilde{f}_{n_2}(x,\gamma) \xrightarrow[\gamma\to 0]{} 0$, $\forall x \in \mathbb{R}^k$,

$$f_{n_2}(x,\gamma) - \tilde{f}_{n_2}(x,\gamma) = \begin{cases} 0 & \text{si} \quad x \notin \{X_1,\ldots,X_n\} \\ \dfrac{1}{n\gamma^k}\dfrac{1}{|\mathrm{Det}\ S_n|^{1/2}}\ K\left(S^{-1/2}(X_{i_o}-X).\left(\dfrac{1-\sqrt{1-\gamma^2}}{\gamma}\right)\right) & \text{si} \quad x = X_{i_o} \end{cases}$$

or si K est borné et si $(\xi_n)_{n\geq 1}$ est une suite de réels positifs, indépendante des observations, telle que $\xi_n \xrightarrow[n\to\infty]{} 0$, $n\,\xi_n^k \xrightarrow[n\to\infty]{} +\infty$ alors

$$\sup_{\gamma\in[\xi_n,1]}\ \sup_{x\in\mathbb{R}^k} |f_{n_2}(x,\gamma)-\tilde{f}_{n_2}(x,\gamma)| \xrightarrow[n\to\infty]{p.s.} 0 \qquad (*)$$

car, dans ce cas, on a

$$\sup_{x\in\mathbb{R}^k} |f_{n_2}(x,\gamma)-\tilde{f}_{n_2}(x,\gamma)| \leq \frac{1}{n\gamma^k|\mathrm{Det}\ S_n|^{1/2}} \cdot M_1 \quad , \quad \text{où} \quad M_1 = \sup_{y\in\mathbb{R}^k} K(y)$$

$$\Longrightarrow \sup_{\gamma\in[\xi_n,1]}\ \sup_{x\in\mathbb{R}^k} |f_{n_2}(x,\gamma)-\tilde{f}_{n_2}(x,\gamma)| \leq \frac{1}{n\xi_n^k} \cdot \frac{1}{|\mathrm{Det}\ S_n|^{1/2}} \cdot M_1$$

ensuite la loi des grands nombres permet d'avoir $|\mathrm{Det}\ S_n| \xrightarrow[n\to\infty]{p.s.} |\mathrm{Det}\ V|$, d'où (*).

III - CHOIX DU PAS γ.-

La mise en oeuvre de cette méthode se heurte au problème du choix du paramètre γ, choix plus délicat que celui de la fenêtre dans l'estimateur de Parzen-Rosenblatt, puisque dans ce cas on arrive à déterminer le pas optimal qui minimise le M.I.S.E. asymptotique, tandis que pour l'estimateur de H.J. BIERENS on peut difficilement procéder de la même façon, vu que l'estimateur dépend de la variance et moyenne empiriques. Le critère que nous retiendrons donc est l'écart quadratique intégré i.e. :

$$d(\tilde{f}_{n_2},f) = d(f_{n_2},f) = \int_{\mathbb{R}^k} (f_{n_2}(x,\gamma) - f(x))^2 \, dx \quad .$$

Minimiser, en γ, $d(f_{n_2},f)$ est équivalent à minimiser

$$R(\gamma) = \int_{\mathbb{R}^k} f^2_{n_2}(x,\gamma) \, dx - 2 \int_{\mathbb{R}^k} f_{n_2}(x,\gamma) \, f(x) \, dx$$

quantité qui fait intervenir l'inconnue f, par conséquent inexploitable en pratique, raison pour laquelle H.J. BIERENS propose de remplacer $\int_{\mathbb{R}^k} \tilde{f}_{n_2}(x,\gamma) \, f(x) \, dx$ par $\frac{1}{n} \sum_{j=1}^{n} \tilde{f}_{n_2}(X_j,\gamma)$, technique qui a été utilisée par M. RUDEMO (1982) et P. HALL (1983) (voir chapitre I, validation croisée, page 10). Ce choix se justifie par le fait que sous certaines conditions sur γ_n , K et f la quantité :

$$\left|\frac{1}{n} \sum_{j=1}^{n} \tilde{f}_{n_2}(X_j,\gamma) - \int_{\mathbb{R}^k} \tilde{f}_{n_2}(x,\gamma) \, f(x) \, dx\right| \text{ converge p.s. vers } 0 \text{ lorsque } n \to \infty \, .$$

Proposition 2.1.- Si f est uniformément continue bornée sur $\mathbb{R}^k$, K une densité bornée et à variation bornée sur $\mathbb{R}^k$ et $(\gamma_n)_{n\geq 1}$ est une suite dans $]0,1]$ telle que :

$$\gamma_n \xrightarrow[n\to\infty]{} 0 \quad , \quad n\,\gamma_n^k \xrightarrow[n\to\infty]{} +\infty \quad \text{et} \quad \frac{n\,\gamma_n^{2k}}{\text{Log } n} \xrightarrow[n\to\infty]{} +\infty \quad .$$

Alors :

$$\left|\int_{\mathbb{R}^k} \tilde{f}_{n_2}(x,\gamma_n)\, f(x)\, dx - \frac{1}{n}\sum_{j=1}^{n} \tilde{f}_{n_2}(X_j,\gamma_n)\right| \xrightarrow[n\to\infty]{p.s.} 0 \quad .$$

Preuve : on a

$$\left|\int_{\mathbb{R}^k} \tilde{f}_{n_2}(x,\gamma_n)\, f(x)\, dx - \frac{1}{n}\sum_{j=1}^{n} \tilde{f}_{n_2}(X_j,\gamma_n)\right| \leq$$

$$\left|\int_{\mathbb{R}^k} \tilde{f}_{n_2}(x,\gamma_n) f(x) dx - \int_{\mathbb{R}^k} \bar{f}_{n_2}(x,\gamma_n) f(x) dx\right| + \left|\int_{\mathbb{R}^k} \bar{f}_{n_2}(x,\gamma_n) f(x) dx - \frac{1}{n}\sum_{j=1}^{n} \bar{f}_{n_2}(X_j,\gamma_n)\right|$$
$$+ \left|\frac{1}{n}\sum_{j=1}^{n} (\bar{f}_{n_2}(X_j,\gamma_n) - \tilde{f}_{n_2}(X_j,\gamma_n))\right| \; .$$

Il suffit donc d'établir la convergence p.s. de chacun des trois termes de droite, de cette inégalité, en effet :

$$\left|\int_{\mathbb{R}^k} \tilde{f}_{n_2}(x,\gamma_n) f(x) dx - \int_{\mathbb{R}^k} \bar{f}_{n_2}(x,\gamma_n) f(x) dx\right| = \tag{P1}$$
$$\left|\int_{\mathbb{R}^k} f_{n_2}(x,\gamma_n) f(x) dx - \int_{\mathbb{R}^k} \bar{f}_{n_2}(x,\gamma_n) f(x) dx\right|$$

car $\tilde{f}_{n_2}(.,\gamma_n) = f_{n_2}(.,\gamma_n)$ λ^k-p.p. où λ^k désigne la mesure de Lebesgue sur $\mathbb{R}^k$, ensuite $\left|\int_{\mathbb{R}^k} f_{n_2}(x,\gamma_n) f(x) dx - \int_{\mathbb{R}^k} \bar{f}_{n_2}(x,\gamma_n) f(x) dx\right|$ est majoré par

$$\sup_x |f_{n_2}(x,\gamma_n) - \bar{f}_{n_2}(x,\gamma_n)| \leq \sup_x |f_{n_2}(x,\gamma_n) - f^*_{n_2}(x,\gamma_n)| + \sup_x |f^*_{n_2}(x,\gamma_n) - \bar{f}_{n_2}(x,\gamma_n)|$$

une application des lemmes 2.1 et 2.2 permet d'avoir la convergence p.s. de ces deux derniers termes et on obtient ainsi la convergence p.s. de (P1) vers 0.

Par ailleurs nous avons

$$\left|\int_{\mathbb{R}^k} \bar{f}_{n_2}(x,\gamma_n)\ f(x)\ dx - \frac{1}{n}\sum_{j=1}^{n} \bar{f}_{n_2}(X_j,\gamma_n)\right| = \tag{P2}$$

$$\left|\int_{\mathbb{R}^k} f(y)\left\{\frac{1}{\gamma_n^k}\ \frac{1}{|\mathrm{Det}\ V|^{1/2}} \int_{\mathbb{R}^k} K\left(V^{-1/2}\left(\frac{x-\sqrt{1-\gamma_n^2}\,y-(1-\sqrt{1-\gamma_n^2})\mu}{\gamma_n}\right)\right)\right.\right.$$

$$\left.\left.((dF_n(x)-dF(x))\right\}\ dy\ \right|$$

cette égalité est justifiée par le théorème de Fubini, son terme de droite est borné par :

$$\frac{1}{\gamma_n^k|\mathrm{Det}\ V|^{1/2}}\ \sup_x\ |F_n(x)-F(x)|\int_{\mathbb{R}^k}\int_{\mathbb{R}^k} f(y)\left|dK\left(V^{-1/2}\left(\frac{x-\sqrt{1-\gamma_n^2}\,y-(1-\sqrt{1-\gamma_n^2})\mu}{\gamma_n}\right)\right)\right| dx\ dy$$

$$\leq \frac{1}{\gamma_n^k}\cdot\frac{1}{|\mathrm{Det}\ V|^{1/2}}\ V_K\ \sup_x\ |F_n(x)-F(x)|\quad ,$$

et on conclut de la même façon que le lemme 2.1.

Enfin le terme $\left|\frac{1}{n}\sum_{j=1}^{n}\left(\bar{f}_{n_2}(X_j,\gamma_n) - \tilde{f}_{n_2}(X_j,\gamma_n)\right)\right|$ est majoré par

$$\sup_x |\bar{f}_{n_2}(x,\gamma_n) - \tilde{f}_{n_2}(x,\gamma_n)| \leq \sup_x |\bar{f}_{n_2}(x,\gamma_n) - f_{n_2}(x,\gamma_n)|$$

$$+ \sup_x |f_{n_2}(x,\gamma_n) - \tilde{f}_{n_2}(x,\gamma_n)|$$

quantités qui convergent p.s. vers 0 car d'une part les lemmes 2.1 et 2.2 entraînent que $\sup_x |\bar{f}_{n_2}(x,\gamma_n) - f_{n_2}(x,\gamma_n)|$ converge p.s. vers 0 quand $n \to \infty$ et d'autre part les conditions K bornée et $n\,\gamma_n^k \xrightarrow[n\to\infty]{} \infty$ impliquent que $\sup_x |f_{n_2}(x,\gamma_n) - \tilde{f}_{n_2}(x,\gamma_n)| \xrightarrow[n\to\infty]{p.s.} 0$. ■

Remarques : la proposition précédente entraîne que :

$$\left|\int_{\mathbb{R}^k}\left(\tilde f_{n_2}(x,\gamma_n)-f(x)\right)^2 dx - \int_{\mathbb{R}^k} f^2(x)dx - \left(\int_{\mathbb{R}^k} \tilde f^2_{n_2}(x,\gamma_n)dx - \frac{2}{n}\sum_{j=1}^{n} \tilde f_{n_2}(X_j,\gamma_n)\right)\right| \xrightarrow[n\to\infty]{p.s.} 0$$

i.e. $|R(\gamma_n)-R_n(\gamma_n)| \xrightarrow[n\to\infty]{p.s.} 0$ où $R(\gamma) = \int_{\mathbb{R}^k} \tilde f^2_{n_2}(x,\gamma)\, dx - 2\int_{\mathbb{R}^k} \tilde f_{n_2}(x,\gamma) f(x)dx$

et
$$R_n(\gamma) = \int_{\mathbb{R}^k} \tilde f^2_{n_2}(x,\gamma)\, dx - \frac{2}{n}\sum_{j=1}^{n} \tilde f_{n_2}(X_j,\gamma) .$$

Soit $(\xi_n)_{n\geq 1}$ une suite de réels positifs, telle que :

$$\forall n \geq 1 \qquad \xi_n \in \left]0,1\right] , \quad \xi_n \xrightarrow[n\to\infty]{} 0 \quad \text{et} \quad \frac{n\,\xi_n^{2k}}{\text{Log } n} \xrightarrow[n\to+\infty]{} +\infty \qquad (R_o)$$

cette suite va nous permettre de montrer que $R_n(.)$ est minoré sur $[\xi_n,1]$ et que $d(\tilde f_n(.,\gamma_n^*),f) \xrightarrow[n\to\infty]{p.s.} 0$ où γ_n^* est défini par

$$R_n(\gamma_n^*) = \underset{\gamma\in[\xi_n,1]}{\text{Min}}\, R_n(\gamma) \quad . \qquad (R_1)$$

En effet si K est borné alors, par définition de $\tilde f_{n_2}(x,\gamma)$, nous aurons

$$\tilde f_{n_2}(x,\gamma) \leq \frac{M_1}{\gamma^k |\text{Det } S_n|^{1/2}} \cdot \frac{n-1}{n} \quad , \quad \forall x \in \mathbb{R}^k$$

d'où

$$|R_n(\gamma)| \leq \int_{\mathbb{R}^k} \tilde f^2_{n_2}(x,\gamma)\, dx + \frac{2}{n}\sum_{j=1}^{n} \tilde f_{n_2}(X_j,\gamma)$$

$$\leq \frac{n-1}{n} \cdot \frac{M_1}{\gamma^k |\text{Det } S_n|^{1/2}} \left(\int_{\mathbb{R}^k} \tilde f_{n_2}(x,\gamma)\, dx + 2\right)$$

$$\leq \frac{3\, M_1}{\gamma^k |\text{Det } S_n|^{1/2}} \quad , \text{ car } \tilde f_{n_2}(.,\gamma) \text{ est une densité.}$$

on en déduit donc que $\forall\gamma \in [\xi_n, 1] \quad R_n(\gamma) \geq \dfrac{-3M_1}{\xi_n^k \, |\mathrm{Det}\ S_n|^{1/2}}$ ce qui assure l'existence d'un $\gamma_n^* \in [\xi_n, 1]$ tel que

$$R_n(\gamma_n^*) = \min_{\gamma\in[\xi_n,1]} R_n(\gamma) \quad ,$$

nous allons maintenant procéder à la démonstration de la convergence p.s. de $f_n(.,\gamma_n^*)$ vers $f(.)$, au sens du risque quadratique.

Proposition 2.2.- Si f est uniformément continue, bornée sur $\mathbb{R}^k$, K une densité bornée et à variation bornée sur $\mathbb{R}^k$, γ_n^* est défini par (R_1) , ξ_n vérifie (R_o) et $n\,\xi_n^k \xrightarrow[n\to\infty]{} +\infty$.

Alors $$\int_{\mathbb{R}^k} (\tilde{f}_{n_2}(x,\gamma_n^*) - f(x))^2\, dx \xrightarrow[n\to\infty]{p.s.} 0 \qquad (R_2)$$

Preuve : Nous avons

$$\int_{\mathbb{R}^k} (\tilde{f}_{n_2}(x,\gamma_n^*)-f(x))^2\, dx = \int_{\mathbb{R}^k} \tilde{f}^2_{n_2}(x,\gamma_n^*)\, dx - \frac{2}{n}\sum_{j=1}^{n} \tilde{f}_{n_2}(X_j,\gamma_n^*)$$

$$+ \int_{\mathbb{R}^k} f^2(x)\, dx + 2\left[\frac{1}{n}\sum_{j=1}^{n} \tilde{f}_{n_2}(X_j,\gamma_n^*) - \int_{\mathbb{R}^k} \tilde{f}_{n_2}(x,\gamma_n^*)\, f(x)\, dx\right]$$

$$= R_n(\gamma_n^*) + \int_{\mathbb{R}^k} f^2(x)\, dx + 2\left[\frac{1}{n}\sum_{j=1}^{n} \tilde{f}_{n_2}(X_j,\gamma_n^*) - \int_{\mathbb{R}^k} \tilde{f}_{n_2}(x,\gamma_n^*)\, f(x)\, dx\right]$$

or par définition de γ_n^* nous avons $R_n(\gamma_n^*) \leq R_n(\xi_n)$, pour prouver (R_2) il suffira donc de montrer que $R_n(\xi_n) + \int_{\mathbb{R}^k} f^2(x)\, dx \xrightarrow[n\to\infty]{p.s.} 0$ et que

$$\left[\frac{1}{n}\sum_{j=1}^{n} \tilde{f}_{n_2}(X_j,\gamma_n^*) - \int_{\mathbb{R}^k} \tilde{f}_{n_2}(x,\gamma_n^*)\, f(x)\, dx\right] \xrightarrow[n\to\infty]{p.s.} 0 \ .$$

a) $\underline{R_n(\xi_n) + \int_{\mathbb{R}^k} f^2(x)\,dx \xrightarrow[n\to\infty]{p.s.} 0}$

en effet on a :

$$\left|R_n(\xi_n)+\int_{\mathbb{R}^k} f^2(x)dx\right| = \left|d\left(\tilde{f}_{n_2}(.,\xi_n),f\right)+2\left[\int_{\mathbb{R}^k}\tilde{f}_{n_2}(x,\xi_n)f(x)dx - \frac{1}{n}\sum_{j=1}^{n}\tilde{f}_{n_2}(X_j,\xi_n)\right]\right|$$

$$\leq \operatorname*{Sup}_x |\tilde{f}_{n_2}(x,\xi_n)-f(x)|.\int_{\mathbb{R}^k}|\tilde{f}_{n_2}(x,\xi_n)-f(x)|\,dx$$

$$+ 2\left|\int_{\mathbb{R}^k}\tilde{f}_{n_2}(x,\xi_n)f(x)dx - \frac{1}{n}\sum_{j=1}^{n}\tilde{f}_{n_2}(X_j,\xi_n)\right|$$

$$\leq 2\operatorname*{Sup}_x |\tilde{f}_{n_2}(x,\xi_n)-f(x)| + 2\left|\int_{\mathbb{R}^k}\tilde{f}_{n_2}(x,\xi_n)f(x)dx - \frac{1}{n}\sum_{j=1}^{n}\tilde{f}_{n_2}(X_j,\xi_n)\right|$$

ce qui permet de conclure, puisque les hypothèses de la proposition 2.1. sont vérifiées, d'autre part on a :

b) $\underline{\left|\frac{1}{n}\sum_{j=1}^{n}\tilde{f}_{n_2}(X_j,\gamma_n^*) - \int_{\mathbb{R}^k}\tilde{f}_{n_2}(x,\gamma_n^*)\,f(x)\,dx\right| \xrightarrow[n]{p.s.} 0}$

car $\left|\frac{1}{n}\sum_{j=1}^{n}\tilde{f}_{n_2}(X_j,\gamma_n^*)-\int_{\mathbb{R}^k}\tilde{f}_{n_2}(x,\gamma_n^*)f(x)dx\right| \leq \left|\frac{1}{n}\sum_{j=1}^{n}(\tilde{f}_{n_2}(X_j,\gamma_n^*)-f^*_{n_2}(X_j,\gamma_n^*))\right|$

$$+\left|\frac{1}{n}\sum_{j=1}^{n}f^*_{n_2}(X_j,\gamma_n^*)-\int_{\mathbb{R}^k}f^*_{n_2}(x,\gamma_n^*)f(x)dx\right| + \left|\int_{\mathbb{R}^k}f(x)(f^*_{n_2}(x,\gamma_n^*)-\tilde{f}_{n_2}(x,\gamma_n^*))dx\right|$$

$$\leq 2\operatorname*{Sup}_x|\tilde{f}_{n_2}(x,\gamma_n^*)-f^*_{n_2}(x,\gamma_n^*)| + \left|\frac{1}{n}\sum_{j=1}^{n}f^*_{n_2}(X_j,\gamma_n^*)-\int_{\mathbb{R}^k}f^*_{n_2}(x,\gamma_n^*)f(x)dx\right|$$

par un raisonnement analogue à celui utilisé pour l'expression (P2) dans la proposition 2.1. nous obtenons :

$$\left|\frac{1}{n}\sum_{j=1}^{n} f^*_{n_2}(X_j,\gamma^*_n) - \int_{\mathbb{R}^k} f^*_{n_2}(x,\gamma^*_n)\ f(x)\ dx\right| \leq \frac{1}{(\gamma^*_n)^k} \cdot \frac{V_K}{|\text{Det } S_n|^{1/2}} \sup_x |F_n(x)-F(x)|$$

par ailleurs le fait que $\tilde{f}_{n_2}(x,\gamma) = f_{n_2}(x,\gamma)$ p.s. , $\forall\gamma \in (0,1]$, implique que $\sup_x |\tilde{f}_{n_2}(x,\gamma^*_n) - f^*_{n_2}(x,\gamma^*_n)| \leq \sup_x |f_{n_2}(x,\gamma^*_n) - f^*_{n_2}(x,\gamma^*_n)|$ et d'après (1.3) (page 25), on a $\forall\gamma_n \in]0,1]$ $\sup_x |f_{n_2}(x,\gamma_n) - f^*_{n_2}(x,\gamma_n)|$ est inférieure à $\frac{V_K}{\gamma_n^k} \frac{1}{|\text{Det } S_n|^{1/2}} \sup_x |F_n(x)-F(x)|$ donc, en utilisant $\gamma^*_n \geq \xi_n, \forall n$, on obtient :

$$\left|\frac{1}{n}\sum_{j=1}^{n} \tilde{f}_{n_2}(X_j,\gamma^*_n) - \int_{\mathbb{R}^k} \tilde{f}_{n_2}(x,\gamma^*_n)\ f(x)\ dx\right| \leq \frac{3\ V_K}{\xi_n^k} \frac{1}{|\text{Det } S_n|^{1/2}} \sup_x |F_n(x)-F(x)|$$

ainsi les hypothèses faites sur ξ_n permettent d'avoir b).

<u>*Remarque*</u> : Dans le cas du noyau normal, H.J. BIERENS a démontré que $\sup_x |\tilde{f}_{n_2}(x,\gamma^*_n) - f(x)|$ converge en probabilité, vers 0 , lorsque n tend vers + ∞. Résultat obtenu à partir de l'uniforme continuité de $\bar{f}_{n_2}(x,\gamma)$ sur $\mathbb{R}^k \times]0,1]$, on peut obtenir la convergence p.s., dans notre cas, en imposant à K des conditions suffisantes de régularité par exemple : l'uniforme continuité.

CHAPITRE III

CONVERGENCE UNIFORME PRESQUE SURE D'UNE CLASSE D'ESTIMATEURS DE LA DENSITE D'UNE LOI DE PROBABILITE.

I - *INTRODUCTION*.-

L'estimation de la densité d'une loi de probabilité dans $\mathbb{R}^k$, $k \geq 1$, par la méthode du noyau a fait l'objet de nombreuses études. La majorité de ces travaux, fait intervenir une fenêtre ne dépendant que du nombre d'observations, ce qui n'est pas toujours réaliste et raisonnable (voir chapitre I).

L. DEVROYE et T.J. WAGNER (1980) ont proposé un estimateur à noyau dont le paramètre de lissage dépend des observations et pour lequel, ils ont établi les conditions de convergence uniforme presque sûre, leur résultat a été élargi, au cas où la fenêtre dépend des observations et du point où l'on veut estimer la densité, par CHAI GENXIANG (1984). Dans ce chapitre on se propose de généraliser ces résultats à une classe d'estimateurs beaucoup plus vaste que les deux précédentes et comme application on donnera des résultats de convergence, concernant des estimateurs connus.

Soit $(X_i)_{i=1,\dots,n}$ une suite de variables aléatoires définies sur un espace probabilisé (Ω,A,P) à valeurs dans $\mathbb{R}^k$ $(k \geq 1)$, indépendantes et de même loi à densité admettant une version continue, bornée f. On définit l'estimateur de f en $x \in \mathbb{R}^k$ par

$$f_n(x,\omega) = \frac{1}{n} \sum_{i=1}^{n} K(\theta_n(\omega,x,X_i(\omega))) \; |J_n(\omega,x,X_i(\omega))| \qquad (1)$$

$\omega \in \Omega$, où

- θ_n est une application définie sur $\Omega \times \mathbb{R}^k \times \mathbb{R}^k$ à valeurs réelles vérifiant les conditions suivantes :

(c_1) Pour P-presque tout ω, $\theta_n(\omega,x,.)$ est, quel que soit x, un homéomorphisme de $\mathbb{R}^k$ à dérivées partielles continues dont le jacobien est noté $J_n(\omega,x,.)$.

(c_2) $\forall(x,i)$, $\theta_n(.,x,X_i(.))$ et $J_n(.,x,X_i(.))$ sont des applications mesurables.

- K est une application de $\mathbb{R}^k$ dans $\mathbb{R}$, intégrable par rapport à la mesure de Lebesgue λ^k sur $\mathbb{R}^k$ et $\int K \, d\lambda^k = 1$.

Dans toute la suite on supposera que ces conditions sont réalisées.

II - CONVERGENCE UNIFORME PRESQUE SURE DE f_n .-

On suppose que la fonction K vérifie les conditions suivantes :

i) K est une densité de probabilité sur $\mathbb{R}^k$, $k \geq 1$, bornée, on notera

$$M_2 = \sup_y K(y)$$

ii) K est Riemann intégrable.

On posera

$$L(t) = \inf \{u \,/\, 1_{[t,+\infty[}(||s||)\, K(s) \leq u \ , \ \lambda^k \text{ . p.p.}\} \quad \text{où} \quad ||s|| = \sup_{j \in \overline{1,k}} |s_j|$$

iii) $\displaystyle\int_0^{+\infty} t^{k-1}\, L(t)\, dt < \infty$.

En plus des conditions (c_1) et (c_2) on imposera à θ_n et à J_n de vérifier certaines des conditions suivantes :

iv) P - presque sûrement, pour tout y : $\lim_{n\to\infty} \theta_n^{-1}(\omega,x,y) = x$, $\forall x \in \mathbb{R}^k$.

v) P - presque sûrement, on a :

$$\forall\delta > 0 \quad \operatorname{Sup} \{R \ / \ \lambda^k(B(0,R) \cap E_{n,\delta}(\omega)) = 0\} \xrightarrow[n\to\infty]{} +\infty$$

où :

$$E_{n,\delta}(\omega) = \bigcup_{x\in A} E_{n,\delta}(x,\omega) \ , \quad E_{n,\delta}(x,\omega) = \{y \ / \ ||\theta_n^{-1}(\omega,x,y) - x|| \geq \delta\}$$

et A est une partie non vide de $\mathbb{R}^k$,

vi) P - presque sûrement on a :

$$\forall R > 0 \qquad \operatorname*{Sup}_{\substack{x\in A \\ ||y||\leq R}} ||\theta_n^{-1}(\omega,x,y) - x|| \longrightarrow 0$$

vii) il existe une suite de réels $(\beta_n)_{n\geq 1}$ telle que : $\beta_n \xrightarrow[n\to+\infty]{} +\infty$ et

P - presque sûrement on a : $\displaystyle\operatorname*{Sup}_{x,y\in\mathbb{R}^k} (|J_n(\omega,x,y)|) \leq \left(\frac{\beta_n}{n} \operatorname{Log} n\right)^{-1/2}$

Remarques :

La condition (iii) est réalisée dès que $||x||^{k+\delta} K(x) \longrightarrow 0$ lorsque $||x|| \to \infty$, pour $\delta > 0$.

Les conditions imposées à K ne sont pas trop fortes puisqu'on ne lui impose pas d'être à variation bornée ou d'avoir une fonction caractéristique intégrable, conditions souvent rencontrées (voir E. NADARAJA (1965) et J. VAN RYZIN (1969).

La condition (v) est entraînée par la condition (vi), souvent plus simple à utiliser.

Dans toute la suite A désignera une partie non vide de $\mathbb{R}^k$.

Théorème 1.- Si la fonction K vérifie (i), (ii) et (iii), θ_n vérifie (v) ou (vi) et J_n satisfait (vii). f est uniformément continue, bornée sur $\mathbb{R}^k$.

Alors $$\sup_{x\in A} |f_n(x,\omega) - f(x)| \xrightarrow[n\to\infty]{P\text{-p.s.}} 0 \quad . \qquad (T1)$$

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

$$|f_n(x,\omega) - f(x)| \leq |f_n(x,\omega) - g_n(x,\omega)| + |g_n(x,\omega) - f(x)|$$

où

$$g_n(x,\omega) = \int_{\mathbb{R}^k} K(\theta_n(\omega,x,y)) \ |J_n(\omega,x,y)| \ f(y) \ d\lambda^k(y) \quad .$$

Sous les hypothèses faites l'existence de g_n est garantie par le théorème de changement de variable (L. SCHWARTZ (1965) dont on déduit également que :

$$g_n(x,\omega) = \int_{\mathbb{R}^k} K(y) \ f(\theta_n^{-1}(\omega,x,y)) \ d\lambda^k(y) \ .$$

Pour démontrer le théorème 1, nous aurons besoin des lemmes suivants :

Lemme 1.-

i) Si (i) est vérifiée alors $\forall x \in \mathbb{R}^k$ $\quad g_n(x,\omega) \xrightarrow[n\to\infty]{} f(x)$ P-p.s. .

ii) Si f est uniformément continue et (v) est satisfaite alors

$$\sup_{x\in A} |g_n(x,\omega) - f(x)| \xrightarrow[n\to\infty]{} 0 \quad \text{P-p.s.} \quad .$$

Démonstration du lemme 1 :

On note $M_1 = \sup_x f(x)$.

i) On a pour tout $y \in \mathbb{R}^k$ et tout $x \in \mathbb{R}^k$

$$|K(y)\ f(\theta_n^{-1}(\omega,x,y))| \leq M_1 \ .\ K(y)$$

P - p.s. pour tout ω, ensuite la continuité de f et la condition (iv) donnent : $K(y)\ f(\theta_n^{-1}(\omega,x,y)) \xrightarrow{p.s.} K(y)\ f(x)$, le théorème de convergence dominée permet donc de conclure.

ii) Fixons $\varepsilon > 0$, soit $\delta(\varepsilon)$ tel que

$$||z-z'|| < \delta(\varepsilon) \Longrightarrow |f(z) - f(z')| < \frac{\varepsilon}{2} \ .\left(\int_{\mathbb{R}^k} |K|\ d\lambda^k\right)^{-1}$$

et $R(\varepsilon) > 0$ tel que : $\displaystyle\int_{||y||>R(\varepsilon)} |K(y)|\ d\lambda^k(y) < \frac{1}{4M_1} \ .\ \varepsilon$.

Soit Ω_o de probabilité 1, sur lequel (v) est vérifiée et $\omega \in \Omega_o$, il existe $N(\varepsilon)$ tel que :

$$n \geq N(\varepsilon) \Longrightarrow \lambda^k(B(0,R(\varepsilon)) \cap E_{n,\delta(\varepsilon)}(\omega)) = 0 \ .$$

Soit $x \in \mathbb{R}^k$, on a :

$$|g_n(x,\omega) - f(x)| = \left|\int_{\mathbb{R}^k} K(y)\ [f(\theta_n^{-1}(\omega,x,y)) - f(x)]\ d\lambda^k(y)\right|$$

$$\leq \frac{\varepsilon}{2} + \int_{E_{n,\delta(\varepsilon)}(x,\omega)} |K(y)|\ |f(x) - f(\theta_n^{-1}(\omega,x,y))|\ d\lambda^k(y)$$

d'où $$\sup_{x \in A} |g_n(x,\omega) - f(x)| \leq \frac{\varepsilon}{2} + 2\ M_1 \ .\int_{E_{n,\delta(\varepsilon)}(\omega)} |K(y)|\ d\lambda^k(y)$$

$$\leq \frac{\varepsilon}{2} + 2\ M_1 \ .\int_{||y||\geq R(\varepsilon)} |K(y)|\ d\lambda^k(y) \leq \varepsilon \ .\ \blacksquare$$

Lemme 2.- Soit K une fonction de $\mathbb{R}^k$, positive, bornée et Riemann intégrable. Pour tout $\eta, \delta, \rho > 0$ il existe une fonction :

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

où

1) $\alpha_1, \dots, \alpha_N$ sont des réels positifs

2) $E_1, \dots, E_N$ sont des pavés ouverts, disjoints, de $[-\rho, \rho]^k$

3) $K^*(x) \leq \operatorname*{Sup}_y K(y) = M_2 \quad , \quad \forall x \in \mathbb{R}^k$

4) $|K^*(x) - K(x)| \leq \eta$ sur $[-\rho,\rho]^k \cap D^c$

5) $D \subset B \subset \bigcup_{i=1}^{M} B_i$ où les B_i, $i \in \overline{1,M}$ sont des pavés de $[-\rho, \rho]^k$

et $\lambda^k(B) < \delta$.

Pour la preuve de ce lemme nous renvoyons à L. DEVROYE et T.J. WAGNER (1980) (lemme 3 pages 67-68).

Lemme 3.- *(de SCHEFFÉ : cf. BILLINGSLEY p.* 223 *).*

Pour toute mesure signée μ de masse totale nulle sur un espace mesurable $(X,\mathcal{M})$ on a :

$$|\mu| \; (X) = 2 \; . \operatorname*{Sup}_{E \in \mathcal{M}} |\mu(E)|$$

où $|\mu|$ est la mesure de variation totale de μ .

Notations et Remarques :

On note : $\mathcal{A}_1$ la σ-algèbre engendrée par les ouverts de $\mathbb{R}^k$.

F la fonction de répartition associée à f et μ la mesure correspondante .

F_n la fonction de répartition empirique , μ_n la mesure associée.

$$E^*_{j,n}(x,\omega) = \{y \mid \theta_n(\omega,x,y) \in E_j\},\ j \in \overline{1,N}\ ,$$

(les E_j sont définis au lemme 2)

$$C_{1,n}(x,\omega) = \{y \mid \theta_n(\omega,x,y) \in [-\rho,\rho]^k\}^C$$

$$C_{2,n}(x,\omega) = \{y \mid \theta_n(\omega,x,y) \in [-\rho,\rho]^k \cap D^c\}$$

$$C_{3,n}(x,\omega) = \{y \mid \theta_n(\omega,x,y) \in D\}$$

$$B_n(x,\omega) = \{y \mid \theta_n(\omega,x,y) \in B\}\ .$$

<u>N.B.</u> : L'application $\theta_n(\omega,x,.)$ étant, pour P - presque tout ω et tout x un homéomorphisme de $\mathbb{R}^k$ les ensembles $E^*_{j,n}(x,\omega)$ sont des ouverts de $\mathbb{R}^k$.

Preuve du théorème 1 : On a

$$|f_n(x,\omega) - f(x)| \leq |f_n(x,\omega) - g_n(x,\omega)| + |g_n(x,\omega) - f(x)|$$

or d'après le lemme 1 $\sup\limits_{x\in A} |g_n(x,\omega) - f(x)| \xrightarrow[n\to\infty]{} 0$ P - p.s. , il suffit

donc de montrer que $\sup\limits_{x\in A} |f_n(x,\omega) - g_n(x,\omega)| \xrightarrow[n\to\infty]{} 0$ P - p.s. (T2) la démonstration de ce résultat va se faire en trois étapes et est essentiellement basée sur le lemme 2. On note

$$I_{1,n}(\omega,x) = \int_{\mathbb{R}^k} |J_n(\omega,x,y)| \; |K(\theta_n(\omega,x,y)) - K^*(\theta_n(\omega,x,y))| \; dF(y)$$

$$I_{2,n}(\omega,x) = \Big| \int_{\mathbb{R}^k} |J_n(\omega,x,y)| \; K^*(\theta_n(\omega,x,y)) \; dF(y) - \int_{\mathbb{R}^k} |J_n(\omega,x,y)| \; K^*(\theta_n(\omega,x,y)) \, dF_n(y) \Big|$$

$$I_{3,n}(\omega,x) = \int_{\mathbb{R}^k} |J_n(\omega,x,y)| \; |K^*(\theta_n(\omega,x,y)) - K(\theta_n(\omega,x,y))| \; dF_n(y)$$

pour tout x dans A et tout ω dans Ω_o on a :

$$|f_n(x,\omega) - g_n(x,\omega)| \leqslant \sum_{i=1}^{3} I_{i,n}(\omega,x) \quad .$$

Majorons successivement $I_{1,n}(\omega,x)$, $I_{2,n}(\omega,x)$ et $I_{3,n}(\omega,x)$.

a $\quad \sup_{x\in A} I_{1,n}(\omega,x) \leqslant \sum_{i=1}^{3} \sup_{x\in A} I^i_{1,n}(\omega,x) \quad$ où :

$$I^i_{1,n}(\omega,x) = \int_{C_{i,n}(x,\omega)} |J_n(\omega,x,y)| \; |K(\theta_n(\omega,x,y)) - K^*(\theta_n(\omega,x,y))| \; dF(y)$$

a1 $\quad \sup_{x\in A} I^1_{1,n}(\omega,x) \leqslant \sup_{x\in A} \int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)| \; K(\theta_n(\omega,x,y)) \; dF(y)$

car si $y \in C_{1,n}(x,\omega)$ alors $\theta_n(\omega,x,y) \in ([-\rho,\rho]^k)^c$ et donc $K^*(\theta_n(\omega,x,y)) = 0$.

a2 $\quad \sup_{x\in A} I^2_{1,n}(\omega,x) \leqslant \eta \, M_1 \, (2\rho)^k$.

En effet d'après le lemme 2 -4) on a :

$$\theta_n(\omega,x,y) \in [-\rho,\rho]^k \cap D^c \Longrightarrow |K(\theta_n(\omega,x,y)) - K^*(\theta_n(\omega,x,y))| \leqslant \eta$$

par conséquent : $\operatorname{Sup}_{x\in A} I^2_{1,n}(\omega,x) \leqslant \eta \,.\, \operatorname{Sup}_{x\in A} \int_{C_{2,n}(x,\omega)} |J_n(\omega,x,y)|\ dF(y)$

or $\int_{C_{2,n}(x,\omega)} |J_n(\omega,x,y)|\ dF(y) \leqslant M_1 \,.\, \int_{C_{2,n}(x,\omega)} |J_n(\omega,x,y)|\ dy$

$$\leqslant M_1 \,.\, \int_{\{y:\theta_n(\omega,x,y)\in[-\rho,\rho]^k\}} |J_n(\omega,x,y)|\ dy \leqslant M_1 \,.\, \lambda^k([-\rho,\rho]^k) \leqslant M_1(2\rho)^k$$

car d'une part pour tout ω et pour tout x dans $\mathbb{R}^k$ on a :

$$\{y : \theta_n(\omega,x,y) \in [-\rho,\rho]^k \cap D^c\} \subset \{y : \theta_n(\omega,x,y) \in [-\rho,\rho]^k\}$$

et d'autre part le théorème de changement de variable nous permet d'écrire

$$\lambda^k\left([-\rho,\rho]^k\right) = \int_{\{y:\theta_n(\omega,x,y)\in[-\rho,\rho]^k\}} |J_n(\omega,x,y)|\ dy \,.$$

(a3) $\operatorname{Sup}_{x\in A} I^3_{1,n}(\omega,x) \leqslant 2\delta\, M_1\, M_2$.

En effet d'après le lemme 2-3) $\forall z \in \mathbb{R}^k \quad K^*(z) \leqslant M_2$ ce qui implique que pour P - p.s. tout ω et pour tout x dans A

$$I^3_{1,n}(\omega,x) \leqslant 2 \,.\, M_2 \,.\, \int_{C_{3,n}(x,\omega)} |J_n(\omega,x,y)|\ dF(y)$$

$$\leqslant 2\, M_2\, M_1 \,.\, \int_{C_{3,n}(x,\omega)} |J_n(\omega,x,y)|\ dy$$

$$\leqslant 2\, M_2\, M_1\, \lambda^k(D) \leqslant 2\delta\, M_1\, M_2 \quad \text{(Lemme 2 - 5))} \,.$$

(b) $\operatorname{Sup}_{x\in A} I_{2,n}(\omega,x) \leqslant \operatorname{Sup}_{x\in A} \left| \sum_{j=1}^{N} \alpha_j \int_{E^*_{j,n}(x,\omega)} |J_n(\omega,x,y)| \left(dF(y) - dF_n(y)\right)\right|$

ceci par définition de K^* , (lemme 2 - 1)).

$$\Longrightarrow \operatorname{Sup}_{x\in A} I_{2,n}(\omega,x) \leqslant N \,.\, M_2 \,.\, \operatorname{Sup}_{x\in A,\, j\in\overline{1,N}} \left|\int_{E^*_{j,n}(x,\omega)} |J_n(\omega,x,y)|\ (dF(y) - dF_n(y))\right|$$

or

$$\left|\int_{E^*_{j,n}(x,\omega)} |J_n(\omega,x,y)|\ (dF(y)-dF_n(y))\right| \leq \left(\beta_n\ .\ \frac{\text{Log } n}{n}\right)^{-1/2} .\ \int_{E^*_{j,n}(x,\omega)} |dF_n(y)-dF(y)|$$

$$\leq \left(\beta_n \frac{\text{Log } n}{n}\right)^{-1/2} .\ \int_{\mathbb{R}^k} |dF_n(y)-dF(y)|$$

$$= 2\left(\beta_n \frac{\text{Log } n}{n}\right)^{-1/2} .\ \operatorname*{Sup}_{E\in\mathcal{A}_1} \left|\int_E (dF_n(y)-dF(y))\right| \qquad \text{(lemme de SCHEFFÉ)} ,$$

on obtient donc :

$$\operatorname*{Sup}_{x\in A} I_{2,n}(\omega,x) \leq 2\ .\ N\ .\ M_2\ .\ \left(\beta_n \frac{\text{Log } n}{n}\right)^{-1/2} .\ \operatorname*{Sup}_{E\in\mathcal{A}_1} |\mu_n(E)-\mu(E)|$$

c) $\operatorname*{Sup}_{x\in A} I_{3,n}(\omega,x) \leq \sum_{i=1}^{3} \operatorname*{Sup}_{x\in A} I^i_{3,n}(\omega,x)$ avec :

$$I^i_{3,n}(\omega,x) = \int_{C_{i,n}(x,\omega)} |K(\theta_n(\omega,x,y)) - K^*(\theta_n(\omega,x,y))|\ |J_n(\omega,x,y)|\ dF_n(y)$$

c1) en procédant de la même façon qu'en a1) on obtient :

$$\operatorname*{Sup}_{x\in A} I^1_{3,n}(\omega,x) \leq \operatorname*{Sup}_{x\in A} \int_{C_{1,n}(x,\omega)} K(\theta_n(\omega,x,y))\ |J_n(\omega,x,y)|\ dF_n(y)$$

c2) $$\operatorname*{Sup}_{x\in A} I^2_{3,n}(\omega,x) \leq \eta\ M_1(2\rho)^k + 2\eta\left(\beta_n \frac{\text{Log } n}{n}\right)^{-1/2} \operatorname*{Sup}_{E\in\mathcal{A}_1} |\mu_n(E)-\mu(E)|$$

car d'après le lemme 2 - 4) si $\theta_n(\omega,x,y) \in [-\rho\ ,\rho]^k \cap D^c$ alors

$$|K(\theta_n(\omega,x,y)) - K^*(\theta_n(\omega,x,y))| \leq \eta \quad ,$$

d'où :

$$\operatorname*{Sup}_{x\in A} I^2_{3,n}(\omega,x) \leq \eta\ .\ \operatorname*{Sup}_{x\in A} \int_{C_{2,n}(x,\omega)} |J_n(\omega,x,y)|\ dF_n(y) \quad ,$$

d'autre part, si $C'_{1,n}(x,\omega)$ désigne le complémentaire de $C_{1,n}(x,\omega)$ alors on aura

$C_{2,n}(x,\omega) \subset C'_{1,n}(x,\omega)$, pour tout ω et pour tout x dans $\mathbb{R}^k$, par conséquent :

$$\int_{C_{2,n}(x,\omega)} |J_n(\omega,x,y)| \ dF_n(y) \leq \int_{C'_{1,n}(x,\omega)} |J_n(\omega,x,y)| \ dF_n(y)$$

$$\leq \int_{C'_{1,n}(x,\omega)} |J_n(\omega,x,y)| \ dF(y) + \left|\int_{C'_{1,n}(x,\omega)} |J_n(\omega,x,y)| \ (dF_n(y)-dF(y))\right|$$

ensuite $\displaystyle\int_{C'_{1,n}(x,\omega)} |J_n(\omega,x,y)| \ dF(y) \leq M_1 \ . \int_{\{y:\theta_n(\omega,x,y)\in[-\rho,\rho]^k\}} |J_n(\omega,x,y)| \ dy$

$$\leq M_1 \ \lambda^k\left([-\rho,\rho]^k\right) \leq M_1(2\rho)^k$$

et

$$\left|\int_{C'_{1,n}(x,\omega)} |J_n(\omega,x,y)| \ (dF_n(y)-dF(y))\right| \leq$$

$$\left(\beta_n \ . \ \frac{\text{Log } n}{n}\right)^{-1/2} . \int_{C'_{1,n}(x,\omega)} |dF_n(y)-dF(y)|$$

$$\leq 2 \ . \left(\beta_n \ \frac{\text{Log } n}{n}\right)^{-1/2} . \ \sup_{E\in A_1} |\mu_n(E)-\mu(E)| \quad , \quad \text{d'où (c2)} .$$

(c3) $\displaystyle\sup_{x\in A} I^3_{3,n}(\omega,x) \leq 2 \ M_1 \ M_2 \ \delta + 4 \ M \ M_2 \left(\beta_n \ \frac{\text{Log } n}{n}\right)^{-1/2} \sup_{E\in A_1} |\mu_n(E)-\mu(E)|$

en effet d'après le lemme 2 : $|K(\theta_n(\omega,x,y) - K^*(\theta_n(\omega,x,y))| \leq 2 \ M_2$.

Donc :

$$I^3_{3,n}(\omega,x) \leq 2 \ M_2 \ . \int_{B_n(x,\omega)} |J_n(\omega,x,y)| \ dF_n(y)$$

$$\leq 2 \ M_2\left\{\int_{B_n(x,\omega)} |J_n(\omega,x,y)| dF(y) + \int_{B_n(x,\omega)} |J_n(\omega,x,y)| (dF_n(y)-dF(y))|\right\}$$

Or $B_n(x,\omega) = \bigcup_{i=1}^{M} B_{n,i}(x,\omega)$ où $B_{n,i}(x,\omega) = \{y \mid \theta_n(\omega,x,y) \in B_i\}$

d'où

$$\left|\int_{B_n(x,\omega)} |J_n(\omega,x,y)|(dF_n(y)-dF(y))\right| \leq M \cdot \underset{j\in\overline{1,M}}{\mathrm{Sup}} \left|\int_{B_{n,j}(x,\omega)} |J_n(\omega,x,y)|(dF_n(y)-dF(y))\right|$$

$$\leq M \cdot \left(\beta_n \frac{\mathrm{Log}\ n}{n}\right)^{-1/2} \cdot \underset{j\in\overline{1,M}}{\mathrm{Sup}} \int_{B_{n,j}(x,\omega)} |dF_n(y)-dF(y)| \qquad \text{(condition (vii))}$$

$$\leq 2 \cdot M \cdot \left(\beta_n \frac{\mathrm{Log}\ n}{n}\right)^{-1/2} \cdot \underset{E\in A_1}{\mathrm{Sup}} \left|\int_E (dF_n(y)-dF(y))\right| \qquad \text{(lemme de SCHEFFE)}$$

par ailleurs nous avons :

$$\int_{B_n(x,\omega)} |J_n(\omega,x,y)|dF(y) \leq M_1 \cdot \int_{\{y:\theta_n(\omega,x,y)\in B\}} |J_n(\omega,x,y)|\, d\lambda^k(y)$$

$$\leq M_1 \cdot \lambda^k(B) \leq \delta \cdot M_1$$

car $\lambda^k(B) < \delta$ (lemme 2-5).

En regroupant les différentes majorations obtenues précédemment on obtient : pour P-presque tout ω :

$$\begin{aligned}
\underset{x\in A}{\mathrm{Sup}} |f_n(x,\omega)-g_n(x,\omega)| \leq \Bigg\{ & \underset{x\in A}{\mathrm{Sup}} \int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|K(\theta_n(\omega,x,y))\, dF(y) \\
& + 2\, M_1\, \eta(2\rho)^k + 4\, M_1\, M_2\, \delta + 2N \cdot M_2 \left(\frac{\beta_n}{n} \cdot \mathrm{Log}\ n\right)^{-1/2} \underset{E\in A_1}{\mathrm{Sup}} |\mu_n(E)-\mu(E)| \\
& + \underset{x\in A}{\mathrm{Sup}} \int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\ K(\theta_n(\omega,x,y))\, dF_n(y) \\
& + [2\eta + 4\, M\, M_2] \left(\frac{\beta_n\ \mathrm{Log}\ n}{n}\right)^{-1/2} \underset{E\in A_1}{\mathrm{Sup}} |\mu_n(E)-\mu(E)| \Bigg\}
\end{aligned} \qquad \text{(T3)}$$

On s'intéresse à la quantité $Q_n(\omega)$ définie par :

$$Q_n(\omega) = \sup_{x\in A} \int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\ K(\theta_n(\omega,x,y))\ dF(y)$$

$$+ \sup_{x\in A} \int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\ K(\theta_n(\omega,x,y))\ dF_n(y) \quad .$$

Il est clair que lorsque K est à support compact, on peut choisir ρ de façon à ce que $Q_n(\omega) = 0 \quad \forall\omega \in \Omega$, sinon par définition de L on a :

$$\lambda^k\text{-p.p.} \quad K(z) \leq L(||z||) \ .$$

Donc $\qquad Q_n(\omega) \leq Q_{n_1}(\omega) + Q_{n_2}(\omega) \quad$ avec

$$Q_{n_1}(\omega) = 2 \sup_{x\in A} \int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\ L(||\theta_n(\omega,x,y)||)\ dF(y)$$

et $$Q_{n_2}(\omega) = \sup_{x\in A} \left| \int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\ L(||\theta_n(\omega,x,y)||)\ (dF_n(y)-dF(y)) \right| \quad .$$

<u>Majorons</u> $Q_{n_1}(\omega)$

Nous avons $Q_{n_1}(\omega) \leq 2\ M_1 \sup_{x\in A} \int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\ L(||\theta_n(\omega,x,y)||)\ dy$

or $C_{1,n}(x,\omega) = \{y/\theta_n(\omega,x,y) \in [-\rho,\rho]^k\}^c$ donc en posant $z = \theta_n(\omega,x,y)$ on obtient d'après le théorème de changement de variable,

$$\int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\ L(||\theta_n(\omega,x,y)||)\ dy = \int_{([-\rho,\rho]^k)^c} L(||z||)\ dz$$

Par ailleurs d'après un résultat de L. DEVROYE et T.J. WAGNER (1980)

$$\int_{([-\rho,\rho]^k)^c} L(||z||)\ dz \leq 2^{2k-1} \int_\rho^{+\infty} t^{k-1}\ L(t)\ dt \quad ,$$

d'où $$Q_{n_1}(\omega) \leq 2^{2k}\ M_1 \int_\rho^{+\infty} t^{k-1}\ L(t)\ dt$$

<u>Majorons</u> $Q_{n_2}(\omega)$

Pour cela on introduit un entier ℓ arbitraire et on pose

$$S_j = \{x : \frac{j-1}{\ell} L(\rho) < L(||x||) \leq \frac{j}{\ell} L(\rho)\} \qquad j \in \overline{1,\ell}$$

$$T_j = \{x : \frac{j-1}{\ell} L(\rho) < L(||x||) \leq L(\rho)\} = \bigcup_{i=j}^{\ell} S_i \quad , \qquad j \in \overline{1,\ell}$$

et $\qquad L'(x) = \sum_{j=1}^{\ell} \frac{j-1}{\ell} L(\rho)\, 1_{S_j}(x)$.

On a : $\forall x \in ([-\rho,\rho]^k)^c \quad |L(||x||) - L'(x)| \leq \frac{1}{\ell} L(\rho)$ (*)

car L étant décroissante sur $\mathbb{R}_+$ on a $||x|| > \rho \Longrightarrow L(||x||) \leq L(\rho)$ ce qui entraîne qu'il existe $j_o \in \overline{1,\ell}$ tel que $x \in S_{j_o}$ ou que $L(||x||) = 0$; si $x \in S_{j_o}$ alors $L'(x) = \frac{j_o-1}{\ell} L(\rho)$ et $\frac{j_o-1}{\ell} L(\rho) < L(||x||) \leq \frac{j_o}{\ell} L(\rho)$ et on obtient : $0 < L(||x||) - L'(x) \leq \frac{L(\rho)}{\ell}$, d'autre part si $L(||x||) = 0$ alors $L'(x) = 0$ et (*) est vérifiée.

Considérons la quantité :

$$Q_{n_2}(\omega) = \operatorname*{Sup}_{x\in A} \left|\int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\; L'(\theta_n(\omega,x,y))\; (dF_n(y)-dF(y))\right|$$

on a

$$Q_{n_2}(\omega) \leq \sum_{j=1}^{3} Q_{n_2}^j(\omega) \quad ,$$

où :

$$Q_{n_2}^1(\omega) = \operatorname*{Sup}_{x\in A} \left|\int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\; L'(\theta_n(\omega,x,y))\; (dF_n(y)-dF(y))\right|$$

$$Q_{n_2}^2(\omega) = \operatorname*{Sup}_{x\in A} \int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\; |L(\theta_n(\omega,x,y)) - L'(\theta_n(\omega,x,y))|\; dF_n(y)$$

$$Q_{n_2}^3(\omega) = \operatorname*{Sup}_{x\in A} \int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\; |L'(\theta_n(\omega,x,y)) - L(\theta_n(\omega,x,y))|\; dF(y)$$

Une application de (*) nous donne :

$$Q^2_{n_2}(\omega) + Q^3_{n_2}(\omega) \leq \frac{L(\rho)}{\ell} \sup_{x \in A} \int_{C_{1,n}(x,\omega)} |J_n(\omega,x,y)| \; [dF_n(y)-dF(y)]$$

$$\leq \frac{2L(\rho)}{\ell} \cdot \left(\frac{\beta_n \text{ Log } n}{n}\right)^{-1/2} \qquad \text{(hypothèse (vii))} .$$

D'autre part nous avons :

$$Q^1_{n_2}(\omega) = \sup_{x \in A} \left| \frac{L(\rho)}{\ell} \cdot \sum_{j=1}^{\ell} (j-1) \left[\int_{S_j(x,\omega)} |J_n(\omega,x,y)| \; (dF_n(y)-dF(y) \right] \right|$$

où :

$$S_j(x,\omega) = \{y : \theta_n(\omega,x,y) \in S_j\} \qquad j \in \overline{1,\ell} \; .$$

Notons $T_j(x,\omega) = \{y : \theta_n(\omega,x,y) \in T_j\}$, $j \in \overline{1,\ell}$, donc $T_j(x,\omega) = \bigcup_{i=j}^{\ell} S_i(x,\omega)$

le fait que les $S_i(x,\omega)$ $i \in \overline{j,\ell}$ soient disjoints entraîne que :

$$\sum_{j=2}^{\ell} \int_{T_j(x,\omega)} |J_n(\omega,x,y)|(dF_n(y)-dF(y)) = \sum_{j=2}^{\ell} \left[\sum_{i=j}^{\ell} \int_{S_i(x,\omega)} |J_n(\omega,x,y)|(dF_n(y)-dF(y)) \right]$$

$$= \sum_{j=2}^{\ell} (j-1) \int_{S_j(x,\omega)} |J_n(\omega,x,y)|(dF_n(y)-dF(y))$$

$$= \sum_{j=1}^{\ell} (j-1) \int_{S_j(x,\omega)} |J_n(\omega,x,y)|(dF_n(y)-dF(y))$$

par conséquent :

$$Q^1_{n_2}(\omega) \leq \frac{L(\rho)}{\ell} \sup_{x \in A} \left| \sum_{j=2}^{\ell} \int_{T_j(x,\omega) \cap C_{1,n}(x,\omega)} |J_n(\omega,x,y)|(dF_n(y)-dF(y)) \right|$$

$$\leq \frac{L(\rho)}{\ell} \sup_{\substack{x \in A \\ j \in \overline{1,\ell}}} \left| \int_{T_j(x,\omega) \cap C_{1,n}(x,\omega)} |J_n(\omega,x,y)|(dF_n(y)-dF(y) \right| \; .$$

Enfin l'hypothèse (vii) permet d'avoir :

$$\left|\int_{T_j(x,\omega)\,\cap\, C_{1,n}(x,\omega)} |J_n(\omega,x,y)|\ [dF_n(y) - dF(y)]\right|$$
$$\leqslant \left(\frac{\beta_n \operatorname{Log} n}{n}\right)^{-1/2} \cdot \int_{T_j(x,\omega)} |dF_n(y)-dF(y)| \ .$$

Le lemme de SCHEFFÉ nous donne :

$$Q^1_{n_2}(\omega) \leqslant 2\, L(\rho)\ .\left(\frac{\beta_n \operatorname{Log} n}{n}\right)^{-1/2} \sup_{E\in A_1} |\mu_n(E)-\mu(E)|\ .$$

En injectant ces nouvelles majorations dans (T3) on obtient :

$$\sup_{x\in A} |f_n(x,\omega) - g_n(x,\omega)| \leqslant$$
$$\left[2\eta\ (2\rho)^k M_1 + 4\ \delta M_1 M_2 + 2^{2k} M_1 \int_\rho^{+\infty} t^{k-1} L(t)\, dt + 2L(\rho) \left(\frac{\beta_n \operatorname{Log} n}{n}\right)^{-1/2}\right]$$
$$+ \left[2\ N\ M_2 + 2\eta + 4\ M\ M_2 + 2L(\rho)\right] \left(\frac{\beta_n \operatorname{Log} n}{n}\right)^{-1/2} \sup_{E\in A_1} |\mu_n(E)-\mu(E)|$$

ℓ étant un entier arbitraire, prenons-le égal à $(1+\ell^*)$ où ℓ^* est égal à la partie entière de $\left(\frac{\beta_n \operatorname{Log} n}{n}\right)^{-1/2}$, ainsi on aura :

$$\frac{1}{\ell} \cdot \left(\frac{\beta_n \operatorname{Log} n}{n}\right)^{-1/2} < 1\ .$$

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

$$\text{(C1)}\quad \rho > \gamma(\varepsilon) \Longrightarrow \begin{cases} \displaystyle\int_\rho^{+\infty} t^{k-1} L(t)\, dt < \frac{\varepsilon}{4} \cdot (2^{2k} M_1)^{-1} \\ \text{et} \\ L(\rho) < \dfrac{\varepsilon}{8} \end{cases}$$

(ceci est possible grâce aux conditions imposées à K).

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

(C2) $$\delta \leq \frac{\varepsilon}{16\, M_1\, M_2} \quad \text{et} \quad \eta \leq \frac{\varepsilon}{8\, M_1 (2\rho)^k}$$

sous ces conditions (T4) devient :

(T5) $$\sup_{x\in A} |f_n(x,\omega) - g_n(x,\omega)| \leq \varepsilon + \alpha \cdot \left(\frac{\beta_n \operatorname{Log} n}{n}\right)^{-1/2} \sup_{E\in A_1} |\mu_n(E)-\mu(E)|$$

où $\alpha = 2\, N\, M_2 + 2\eta + 4\, M\, M_2 + 2L(\rho)$.

ρ, δ, η étant fixés de façon à ce que (C1) et (C2) soient vérifiées, le lemme 2 nous permet de fixer N et M donc aussi α. Une application de l'inégalité de A. DVORETZY, J. KIEFER et J. WOLFOWITZ (1956) à :

$$\alpha\left(\frac{\beta_n \operatorname{Log} n}{n}\right)^{-1/2} \sup_{E\in A_1} |\mu_n(E)-\mu(E)| \quad ,$$

nous donnera :

$$\forall \varepsilon' > 0 \, , \Pr\{\alpha \left(\frac{\beta_n \operatorname{Log} n}{n}\right)^{-1/2} \sup_{E\in A_1} |\mu_n(E)-\mu(E)| \geq \varepsilon'\} \leq c_o \operatorname{Exp}(- c_1 \frac{\varepsilon'^2}{\alpha^2} \beta_n \operatorname{Log} n)$$

où c_o et c_1 sont deux constantes, comme $\beta_n \xrightarrow[n\to\infty]{} \infty$ la série de terme général $\exp(- c_1 \frac{\varepsilon'^2}{\alpha^2} \beta_n \operatorname{Log} n)$, est une série convergente, par conséquent le lemme de Borel-Cantelli entraîne que :

$$\forall \varepsilon > 0 \qquad \limsup_{n\to\infty} \left(\sup_{x\in A} |f_n(x,\omega) - g_n(x,\omega)|\right) \leq \varepsilon \qquad P - \text{p.s.} \quad \blacksquare$$

III - APPLICATIONS.-

On s'intéresse dans cette partie, aux applications du résultat précédent à des estimateurs connus.

1°) Estimateur de P. DEHEUVELS (1977-b).

P. DEHEUVELS (1977-b) a défini un estimateur à noyau qui généralise l'estimateur classique de Parzen-Rosenblatt ; comme étant égal à

$$f_{n_1}(x) = \frac{|\mathrm{Det}\ A_n|}{n} \sum_{i=1}^{n} K(A_n(x-X_i))$$

où A_n est une matrice $(k \times k)$ inversible qui ne dépend que de n.

K une densité de probabilité sur $\mathbb{R}^k$.

Théorème 2.- Si f est uniformément continue, bornée sur $\mathbb{R}^k$

K est une densité de probabilité bornée, Riemann intégrable et telle que $\int_0^{+\infty} t^{k-1}\ L(t)\ dt < \infty$.

Si A_n vérifie $\dfrac{n\ |\mathrm{Det}\ A_n|^{-2}}{\mathrm{Log}\ n} \xrightarrow[n\to+\infty]{} +\infty$ et $||A_n^{-1}||^* \xrightarrow[n\to\infty]{} 0$.

Alors $$\sup_{x\in\mathbb{R}^k} |f_{n_1}(x) - f(x)| \xrightarrow[n\to\infty]{p.s.} 0 .$$

$||.||^*$ désigne la norme matricielle subordonnée à la norme vectorielle $||.||$ i.e.

$$||A_n||^* = \sup_{x\neq 0} \frac{||A_n\ x||}{||x||} .$$

Pour prouver ce résultat nous allons montrer que f_{n_1} fait partie de la classe des estimateurs définis par (1) et que les conditions du théorème 1 sont satisfaites.

Si on pose $\theta_n(\omega,x,X_i(\omega)) = A_n(x-X_i(\omega))$ $\forall i \in \overline{1,n}$, $\forall x \in \mathbb{R}^k$, $\forall \omega \in \Omega_o$, alors f_{n_1} s'écrira bien sous la forme (1), reste à montrer que les conditions du théorème 2 entraînent celles du théorème 1 ; en effet :

La condition $||A_n^{-1}||^* \xrightarrow[n\to\infty]{} 0$ entraîne (vi) car $\forall R > 0$

$$\sup_{\substack{x\in\mathbb{R}^k \\ ||y||\leq R}} ||\theta_n^{-1}(\omega,x,y) - x|| = \sup_{\substack{x\in\mathbb{R}^k \\ ||y||\leq R}} ||A_n^{-1}\, y|| \leq ||A_n^{-1}||^* R$$, pour

obtenir (vii) il suffira de choisir $\beta_n = \dfrac{n}{\text{Log } n} |\text{Det } A_n|^{-2}$, $\forall n > 1$, puisque dans ce cas

$$\forall \omega \in \Omega_o \quad \sup_{x,y} |J_n(\omega,x,y)| = |\text{Det } A_n| = \left(\frac{\beta_n}{n} \text{ Log } n\right)^{-1/2} .$$

Remarque : En imposant à K des conditions plus fortes que les nôtres, P. DEHEUVELS (1974) a établi le résultat suivant :

" Une condition nécessaire et suffisante pour que $\sup_x |f_{n_1}(x) - f(x)|$ converge p.s. vers 0 est que :

- K soit une fonction de $\mathbb{R}^k$ dans $\mathbb{R}$, bornée et telle que $\int K \, d\lambda^k = 1$ et $||x|| \; K(x) \xrightarrow[||x|| \to +\infty]{} 0$, $\forall \ell \geq 0$

- A_n soit un automorphisme de $\mathbb{R}^k$ tel que : $||A_n^{-1}||^* \xrightarrow[n\to\infty]{} 0$ et

$$\frac{n\, |\text{Det } A_n|^{-1}}{\text{Log } n} \xrightarrow[n\to\infty]{} +\infty$$ "

2°) Estimateur de L. DEVROYE et T.J. WAGNER (1980).

L'estimateur de L. DEVROYE et T.J. WAGNER est défini par :

$$f_{n_2}(x,\omega) = \frac{1}{n\, h_n^k(\omega)} \sum_{i=1}^{n} K\left(\frac{x - X_i(\omega)}{h_n(\omega)}\right)$$

où $h_n(\omega)$ est une suite de réels positifs, qui dépend des observations X_i , $i \in \overline{1,n}$ et K est une densité de probabilité sur $\mathbb{R}^k$.

f_{n_2} est un cas particulier de (1), il suffit de poser :

$$\theta_n(\omega,x,y) = \frac{x - y}{h_n(\omega)} \quad , \quad \forall x , y \in \mathbb{R}^k \quad , \quad \forall \omega \in \Omega_o \; , \qquad \forall\, n \geq 1$$

une version du théorème 1 pour f_{n_2} est la suivante :

Théorème 3.- Si f est uniformément continue bornée sur $\mathbb{R}^k$ si K vérifie (i), (ii) et (iii).

Si $h_n(\omega) \xrightarrow[n\to\infty]{} 0$ P - p.s. et s'il existe une suite de réels positifs $(\beta_n)_{n\geq 1}$ telle que $\beta_n \xrightarrow[n\to\infty]{} \infty$

et $\forall n > 1$, $\dfrac{n\, h_n^{2k}(\omega)}{\text{Log } n} \geq \beta_n$ P - p.s.

Alors $\operatorname{Sup}_{x\in\mathbb{R}^k} |f_{n_2}(x,\omega) - f(x)| \xrightarrow[n\to\infty]{} 0$ P - p.s. .

Preuve : Il est aisé de vérifier que $h_n(\omega) \xrightarrow[n\to\infty]{P-p.s.} 0$ entraîne (vi) et que (vii) est assurée par l'existence de $(\beta_n)_n$.

3°) *Estimateur de CHAI GENXIANG (1984).*

CHAI GENXIANG a repris l'estimateur précédent et a considéré une fenêtre qui dépend à la fois des observations et du point où l'on estime la densité,son estimateur s'écrit :

$$f_{n_3}(x) = \frac{1}{n\, h_n^k(x,\omega)} (\alpha\rho)^k \, . \sum_{i=1}^{n} K\left(\frac{\alpha\rho}{h_n(x,\omega)} \, . \, (x - X_i(\omega))\right)$$

où α est une constante supérieure ou égale à 1

$h_n(x,\omega)$ est une suite de réels positifs, qui dépend des observations et de x

K est une densité de probabilité à support compact $[-\rho,\rho]^k$.

f_{n_3} s'obtient à partir de (1) en posant :

$$\theta_n(\omega,x,y) = \frac{\alpha\rho}{h_n(x,\omega)} \, . \, (x-y) \quad , \quad \forall x,\ y \in \mathbb{R}^k \quad , \ \forall \omega \in \Omega_o$$

à partir du théorème 1 on obtient :

Théorème 4.- Si f est uniformément continue bornée sur $\mathbb{R}^k$ K vérifie les conditions (i) à (iii).

Soit $A \subseteq \mathbb{R}^k$ tel que : $\sup_{x\in A} h_n(x,\omega) \xrightarrow[n\to\infty]{} 0$ P - p.s.

et il existe une suite $(\beta_n)_{n\geq 1} : \beta_n \geq 0$, $\beta_n \xrightarrow[n\to\infty]{} +\infty$

$$\inf_{x\in A} (h_n^{2k}(x,\omega)) \, . \, \frac{n}{\mathrm{Log}\ n} \geq \beta_n \ , \ \text{P - p.s.}$$

Alors : $\sup_{x\in A} |f_{n_3}(x,\omega) - f(x)| \xrightarrow[n\to\infty]{} 0$ P - p.s. .

Preuve : Il suffit de montrer que les conditions (vi) et (vii) sont réalisées, en effet on a : $\theta_n^{-1}(\omega,x,y) = \left(x - \frac{h_n(x,\omega)}{\alpha\rho} \cdot y\right)$, $\forall y \in \mathbb{R}^k$,

ce qui entraîne que

$$\forall R > 0 \qquad \sup_{\substack{x\in A \\ ||y||\leq R}} ||\theta_n^{-1}(\omega,x,y) - x|| \leq \frac{1}{\alpha\rho} R \cdot \sup_{x\in A} h_n(x,\omega) \quad ,$$

d'où (vi) ensuite par définition de θ_n on a :

$$|J_n(\omega,x,y)| = (\alpha\rho)^k \cdot \frac{1}{h_n^k(x,\omega)}$$

la condition $\inf_{x\in A} (h_n^{2k}(x,\omega)) \geq \frac{\beta_n \cdot \operatorname{Log} n}{n}$ permet donc d'avoir (vii).

Remarque : Ce résultat entraîne celui de CHAI GENXIANG, qui a imposé à $h_n(x,\omega)$ la condition de Lipschitz suivante :

$$\forall x \text{ , } y \in \mathbb{R}^k \text{ ,} \qquad |h_n(x,\omega) - h_n(y,\omega)| \leq ||x-y|| \text{ ,} \quad P - p.s.$$

condition qu'il fait intervenir lors de la démonstration de :

$$\sup_{x\in A} |g_n(x,\omega) - f(x)| \xrightarrow[n\to\infty]{P-p.s.} 0 \quad .$$

Nous signalons que, comme application de ce résultat, CHAI GENXIANG montre la convergence uniforme p.s. de l'estimateur du plus proche voisin.

4°) Estimateur de H.J. BIERENS (1983) (voir chapitre II).

L'estimateur de H.J. Bierens s'écrit :

$$f_{n_5}(x,\omega) = \frac{1}{n\,\gamma_n^k} \frac{1}{|\operatorname{Det} S_n|^{1/2}} \sum_{i=1}^{n} K\left(S_n^{-1/2}\left(\frac{x-\sqrt{1-\gamma_n^2}\,X_i(\omega)-(1-\sqrt{1-\gamma_n^2})\bar{X}_n}{\gamma_n}\right)\right)$$

où γ_n est une suite de réels appartenant à $]0,1]$, $\bar{X}_n$ est la moyenne empirique ,

S est la matrice variance - covariance empirique de l'échantillon $(X_i)_{i\in\overline{1,n}}$,

K est une densité de probabilité, de moyenne nulle, de matrice variance - covariance égale à l'identité de $\mathbb{R}^k$.

Théorème 5.- Si K vérifie (i), (ii) et (iii) et

$\int_{\mathbb{R}^k} x\, K(x)\, dx = 0$; $\int_{\mathbb{R}^k} x\, x^t\, K(x)\, dx = I$.

Si f est uniformément continue bornée sur $\mathbb{R}^k$ et admet un moment d'ordre 1 noté μ et une matrice de variance - covariance V inversible.

Si $(\gamma_n)_{n\geq 1}$ est une suite de réels dans $]0,1]$, telle que :

$$\gamma_n \xrightarrow[n\to\infty]{} 0 \qquad \text{et} \qquad \frac{n\,\gamma_n^{2k}}{\text{Log } n} \xrightarrow[n\to\infty]{} +\infty \ .$$

Alors, en supposant que l'on peut prendre pour A tout sous-ensemble compact du support de la mesure de densité f, on a :

$$\sup_{x\in A} |f_{n_5}(x,\omega) - f(x)| \xrightarrow[n\to\infty]{} 0 \qquad P - \text{p.s.} \ .$$

Preuve : Posons :

$$\theta_n(\omega,x,y) = S_n^{-1/2}\left(\frac{x - \sqrt{1-\gamma_n^2}\; y - (1-\sqrt{1-\gamma_n^2})\;\bar{X}_n}{\gamma_n}\right)$$

d'où $|J_n(\omega,x,y)| = \dfrac{1}{\gamma_n^k\;|\text{Det } S_n|^{1/2}}\;(\sqrt{1-\gamma_n^2})^k$, on constate que l'estimateur $\tilde{f}_{n_5}(x,\omega)$ défini par $\tilde{f}_{n_5}(x,\omega) = (\sqrt{1-\gamma_n^2})^k\; f_{n_5}(x,\omega)$, s'écrit sous la forme (1). On a, par ailleurs :

$$\sup_{x\in A} |f_{n_5}(x,\omega) - f(x)| \leq \frac{1}{(\sqrt{1-\gamma_n^2})^k} \cdot \sup_{x\in A} |\tilde{f}_{n_5}(x,\omega) - f(x)| + \frac{1-(\sqrt{1-\gamma_n^2})^k}{(\sqrt{1-\gamma_n^2})^k} \sup_{x\in A} f(x)$$

or la condition $\gamma_n \xrightarrow[n\to\infty]{} 0$ entraîne que $\dfrac{1-(\sqrt{1-\gamma_n^2})^k}{(\sqrt{1-\gamma_n^2})^k} \xrightarrow[n\to\infty]{} 0$, donc pour établir le théorème 5 , il suffit de montrer que $\tilde{f}_{n_5}(x,\omega)$ vérifie bien les conditions du théorème 1 ; en effet θ_n^{-1} est définie par :

$$\theta_n^{-1}(\omega,x,y) = \frac{1}{\sqrt{1-\gamma_n^2}} (x - (1 - \sqrt{1-\gamma_n^2})\, \bar{X}(\omega) - \gamma_n\, S_n^{1/2}(\omega) \,.\, y)$$

d'où $\forall R > 0$

$$\sup_{\substack{x\in A\\ ||y||\leq R}} ||\theta_n^{-1}(\omega,x,y)-x|| \leq \frac{1-\sqrt{1-\gamma_n^2}}{\sqrt{1-\gamma_n^2}} (\rho_A + ||\bar{x}(\omega)||) + \frac{\gamma_n}{\sqrt{1-\gamma_n^2}} R.||S_n^{1/2}||^*$$

où $\rho_A = \sup_{x\in A} ||x||$.

Or par hypothèse $\mu = \int_{\mathbb{R}^k} x\, f(x)\, dx$ $\quad V = \int_{\mathbb{R}^k} (x-\mu)\,(x-\mu)^t f(x)\, dx$ et les $(X_i)_{i=1,\dots,n}$ sont des v.a.i.i.d., la loi forte des grands nombres entraîne donc que :

$$\bar{X}(\omega) \xrightarrow{p.s.} \mu \quad \text{et} \quad S_n \xrightarrow{p.s.} V \quad , \text{ ce qui permet d'avoir (vi).}$$

Par ailleurs soit $\sigma_o > 0$ tel que $|\mathrm{Det}\ V|^{1/2} > \sigma_o$; la convergence p.s. de S_n vers V assure l'existence de N_o : $\forall n \geq N_o$ $|\mathrm{Det}\ S_n|^{1/2} > \sigma_o$ p.s.

donc $\forall n \geq N_o$ $\sup_{x,y\in A} |J_n(\omega,x,y)| < \dfrac{1}{\gamma_n^k} \cdot \dfrac{1}{|\mathrm{Det}\ S_n|^{1/2}} < \dfrac{1}{\sigma_o\, \gamma_n^k}$ $\quad P$ - p.s. .

Enfin en posant $\beta_n = \sigma_o^2 \dfrac{n\, \gamma_n^{2k}}{\mathrm{Log}\ n}$ on aboutit à (vii) .

CHAPITRE IV

CONDITIONNEMENT SEQUENTIEL ET APPLICATION A l'ESTIMATION DE LA DENSITE PAR LA METHODE DU NOYAU.

I - *INTRODUCTION*.-

A. BERLINET (1984) a étudié des propriétés des paramètres statistiques qu'il a qualifiées de locales dans l'espace des observations, et a proposé une technique qui améliore des estimateurs connus ; cette technique consiste à conditionner par rapport à des statistiques qui ont de bonnes propriétés locales. Nous avons repris ce principe pour définir le conditionnement séquentiel, que nous avons particularisé à l'estimation de la densité univariée, par la méthode du noyau.

II - *CONDITIONNEMENT SEQUENTIEL*.-

Soit $(E,\mathcal{B},\mathcal{P})$ un modèle statistique, où E est un espace topologique et $\mathcal{B}$ sa tribu borélienne. On considère une suite de boréliens propres $(B_i)_{i\in\mathbb{N}}$ appartenant à $\mathcal{B}$, pour $P \in \mathcal{P}$ et $B \in \mathcal{B}$ on notera $\mathcal{P}(P,B) = \{Q : Q \in \mathcal{P}$ et $Q|_{B^c} = P|_{B^c}\}$ où $P|_B$ désigne la restriction de P à B.

Une statistique S est dite B-exhaustive si et seulement si

$\forall P \in \mathcal{P}$, S est exhaustive dans le modèle $(E,\mathcal{B},\mathcal{P}(P,B))$. Intuitivement on peut dire qu'une fois que l'on s'est fixé $P|_{B^c}$, une statistique B-exhaustive résume l'information que l'on peut avoir sur $P|_B$ i.e. ce qui manque pour connaître entièrement la loi P. On se propose, dans ce qui suit, de définir une statistique $\prod_{i=1}^{n} B_i$ - exhaustive dans le modèle $(E^n, \mathcal{B}^{\otimes n}, \mathcal{P}^n)$, qui nous servira à définir le conditionnement séquentiel avec $\mathcal{P}^n$ = l'ensemble des probabilités de la forme $\bigotimes_{i=1}^{n} P_i$ où $P_i \in \mathcal{P}_i$ et $\mathcal{P}_i$ un ensemble de probabilités sur $(E,\mathcal{B})$.

Remarque 1 :

Si $\left(\bigotimes_{i=1}^{n} Q_i\right)\Big|_{\left(\prod_{i=1}^{n} B_i\right)^c} = \left(\bigotimes_{i=1}^{n} P_i\right)\Big|_{\left(\prod_{i=1}^{n} B_i\right)^c}$ alors

$\forall i \in \overline{1,n}$ $\quad Q_i\big|_{B_i^c} = P_i\big|_{B_i^c}$

en effet pour tout $A_i \in \mathcal{B}$, $i \in \overline{1,n}$, tel que $A_i \subset B_i^c$ on a

$$E \times E \times \dots \times E \times A_i \times \dots \times E \subset \left(\prod_{i=1}^{n} B_i\right)^c \quad \text{ce qui entraîne}$$

que :

$$\bigotimes_{j=1}^{n} Q_j(E\times E\times\dots\times E\times A_i\times\dots\times E) = \bigotimes_{j=1}^{n} P_j(E\times E\times\dots\times E\times A_i\times E\times\dots\times E) \Longrightarrow$$

$$Q_i(A_i) = P_i(A_i) \Longrightarrow Q_i\big|_{B_i^c} = P_i\big|_{B_i^c} \quad .$$

Lorsqu'on considère le modèle $(E^n, \mathcal{B}^{\otimes n}, \mathcal{P}^n)$ on a donc

$$\mathcal{P}\left(\bigotimes_{i=1}^{n} P_i , \prod_{i=1}^{n} B_i\right) \subset \bigotimes_{i=1}^{n} \mathcal{P}(P_i,B_i) \subset \mathcal{P}^n \quad \blacksquare$$

Soit $(X_i)_{i=1,\dots,n}$ une suite de v.a.i. à valeurs dans $(E,\mathcal{B})$. On suppose que les X_i ne sont pas forcément de même loi, le modèle est choisi égal à $(E^n,\mathcal{B}^{\otimes n}, \mathcal{P}^n)$ où $\mathcal{P}^n = \bigotimes_{i=1}^n \mathcal{P}_i = \{ \bigotimes_{i=1}^n P_i \text{ , } P_i \in \mathcal{P}_i\}$ et $\mathcal{P}_i$ est une famille de lois de probabilités contenant la loi de X_i , $i \in \overline{1,n}$, $\mathcal{P}_i$ est dominé par une mesure positive μ σ-finie.

Proposition 4.1.- La statistique $S = (S_1,\dots,S_n)$ à valeurs dans $(E,\mathcal{B})^{\otimes n}$ définie par :

$$\begin{cases} S_i = X_i & \text{si} \quad X_i \in B_i \\ \quad = \tilde{x}_i & \text{sinon, avec} \quad \tilde{x}_i \notin B_i \end{cases}$$

est $\prod_{i=1}^n B_i$ - exhaustive dans le modèle $(E^n, \mathcal{B}^{\otimes n},\mathcal{P}^n)$.

Démonstration : On doit montrer que $\forall \bigotimes_{i=1}^n P_i \in \mathcal{P}^n$, S est exhaustive dans le modèle dominé $(E^n, \mathcal{B}^{\otimes n} , \mathcal{P}(\bigotimes_{i=1}^n P_i , \prod_{i=1}^n B_i)$. Supposons que le loi de $(X_1,\dots,X_n)$ soit $\bigotimes_{i=1}^n Q_i \in \mathcal{P}(\bigotimes_{i=1}^n P_i , \prod_{i=1}^n B_i)$.
On a alors, d'après la remarque 1, $\forall i \in \overline{1,n}$ $Q_i\big|_{B^c} = P_i\big|_{B_i^c}$. On note f_{Q_i} une version de $\frac{dQ_i}{d\mu}$ et f_{P_i} une version de $\frac{dP_i}{d\mu}$, coïncidant avec f_{Q_i} sur B_i^c la vraisemblance au point $(X_1,\dots,X_n)$ s'écrit :

$$L(X_1,\dots,X_n) = \prod_{i=1}^n f_{Q_i}(X_i)$$

or $f_{Q_i}(X_i) = (f_{Q_i}(X_i)\ 1_{B_i}(X_i) + 1_{B_i^c}(X_i))\ (f_{Q_i}(X_i)\ 1_{B_i^c(X_i)} + 1_{B_i}(X_i))$

en utilisant la définition de S_i et le fait que $f_{Q_i} 1_{B_i^c} = f_{P_i} 1_{B_i^c}$,

on obtient

$$L(X_1,\dots,X_n) = \prod_{i=1}^{n} (f_{Q_i}(S_i)\, 1_{B_i}(S_i) + 1_{B_i^c}(S_i))$$

$$\times \prod_{i=1}^{n} (f_{P_i}(X_i)\, 1_{B_i^c}(X_i) + 1_{B_i}(X_i)) \quad ,$$

l'exhaustivité de S dans le modèle $(E^n, \mathcal{B}^{\otimes n}, \mathcal{P}(\bigotimes_{i=1}^{n} P_i , \prod_{i=1}^{n} B_i))$ résulte du théorème de factorisation. ■

Remarque 2 :

a) Si les $(X_i)_{i=1,\dots,n}$ sont de même loi alors la statistique S est encore exhaustive dans les modèles $(E^n, \mathcal{B}^{\otimes n}, \mathcal{P}(P^{\otimes n}, \prod_{i=1}^{n} B_i))$ et $(E^n, \mathcal{B}^{\otimes n}, [\mathcal{P}(P, \bigcap_{i=1}^{n} B_i)]^{\otimes n})$.

b) Si les $(X_i)_{i=1,\dots,n}$ sont de même loi et $\forall i \in \overline{1,n}$ $B_i = B$ alors S est exhaustive dans le modèle $(E^n, \mathcal{B}^{\otimes n}, \mathcal{P}(P^{\otimes n}, B^n))$ et on retrouve ainsi le résultat de A. BERLINET (1984-a).

Une façon classique d'améliorer un estimateur est, d'après le théorème de Rao-Blackwell, de le remplacer par son espérance conditionnelle par rapport à une statistique exhaustive. Nous allons nous intéresser, plus particulièrement, à l'application du conditionnement aux estimateurs à noyaux, séquentiels ou non, de la densité sur $\mathbb{R}$, mais auparavant reprenons un résultat tout à fait général dû à A. BERLINET (1984-a) p. VI-06 :

" Soit $(X_i)_{i=1,\dots,n}$ une suite de v.a.i. telle que : pour tout $i \in \overline{1,n}$ X_i est définie sur un espace probabilisé $(\Omega_i, \mathcal{A}_i, Pr_i)$, de loi

Q_i inconnue appartenant à un ensemble $\mathcal{P}_i$ de probabilités sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, qu'on désignera par $\mathcal{P}$ lorsque les X_i, $i \in \overline{1,n}$ sont de même loi Q

et $\psi : \mathcal{P}^n \longrightarrow \mathbb{R}$

$$\bigotimes_{i=1}^{n} Q_i \longmapsto \psi(Q_1 \otimes Q_2 \otimes \ldots\ldots \otimes Q_n)$$

un paramètre réel localisé (cf. A. BERLINET (1984-a) p. X-07) par un fermé F $(F \neq \mathbb{R}^n)$, on se fixe une suite d'ouverts propres U_i, $i \in \overline{1,n}$ tels que F soit inclus dans $\prod_{i=1}^{n} U_i$, et une suite de probabilités $P_i \in \mathcal{P}_i$, $i \in \overline{1,n}$, telles que $P_i(U_i^c) > 0$.

On se donne un estimateur d'ordre n de la forme

$$T(X_1,\ldots,X_n) = \sum_{i=1}^{n} K_i(X_i) \quad ,$$

du paramètre $\psi(\bigotimes_{i=1}^{n} Q_i)$, où les $(K_i)_{i\in\overline{1,n}}$ (qui dépendent ou non de n) sont des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ $\mathcal{P}_i$-intégrables.

On définit S en posant $B_i = U_i$, $i \in \overline{1,n}$

Proposition 4.2.- La tribu $\mathcal{B}_S$, engendrée par S, est $\prod_{i=1}^{n} U_i$ - exhaustive dans le modèle $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \mathcal{P}^n)$ et

$$\sum_{i=1}^{n} \left[K_i(X_i)\, 1_{U_i}(X_i) + m_i(P_i)\, 1_{U_i^c}(X_i)\right]$$

est une version de

$$E^{\mathcal{B}_S}_{\bigotimes_{j=1}^{n} P_j}(T(X_1,\ldots,X_n)) \quad \text{où} \quad m_i(P_i) = \frac{1}{P_i(U_i^c)} \int_{U_i^c} K_i \, dP_i \quad .$$

III - APPLICATION A L'ESTIMATION PAR LA METHODE DU NOYAU.-

III. 1. - Introduction et notations.

Les éléments Q de $\mathcal{P}$ sont supposés admettre une densité par rapport à la mesure de Lebesgue λ, dont une version f_Q est bornée et continue au voisinage d'un réel fixé x. On cherche à estimer $f_Q(x)$, pour $u \in \mathbb{R}$ et $i \in \overline{1,n}$ on pose $K_i(u) = \frac{1}{n\,h_n} K(\frac{x-u}{h_n})$ où K est une fonction λ-intégrable.

Le voisinage U du point x considéré est l'intervalle $]x-H, x+H[$ où H est un réel qui dépend ou non du nombre d'observations n.

Si les observations $(X_i)_{i=1,\dots,n}$ sont indépendantes et de même loi Q alors, avec ces notations, l'estimateur $T(X_1,\dots,X_n)$ de la proposition 4.2., prend la forme de l'estimateur à noyau $f_n(x)$ de $f_Q(x)$ et une version de $E_{P^{\otimes n}}^{\mathcal{B}_S}(f_n(x))$ est

$$\tilde{f}_n(x) = \frac{1}{n\,h_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n}\right) 1_U(X_j) + m(P)\, 1_{U^c}(X_j)$$

où
$$m(P) = \frac{1}{P(U^c)} \int_{U^c} K(\frac{x-y}{h_n})\, f_P(y)\, dy$$

avec P une probabilité de $\mathcal{P}$, telle que $P(U^c) > 0$.

L'apport de cet estimateur est que le voisinage U ne dépend pas forcément de la fenêtre h_n, il généralise donc l'estimateur à noyau conditionné proposé par A. BERLINET. On se propose d'étudier le biais, le risque quadratique et la loi limite de $\tilde{f}_n$, on s'intéressera ensuite à la loi du Log itéré de $\tilde{f}_n$, lorsque $h_n = H$, après cela on examinera le comportement asymptotique de $\tilde{f}_n$ dans le cas où f_P est remplacé par son estimateur à noyau, en fait nous avons obtenu la convergence presque sûre, en moyenne et en moyenne quadratique. Enfin on envisage l'application du conditionnement à l'estimateur à noyau séquentiel, univarié, estimateur défini par P. DEHEUVELS (1974). L'étude du cas multivarié sera faite ultérieurement.

III. 2. - Etude du risque quadratique.-

1°) Comparaison des risques quadratiques de $f_n(x)$ *et* $\tilde{f}_n(x)$.

Dans toute la suite les lois Q et P seront prises dans $\mathcal{P}$, le noyau K sera supposé $\{\lambda\} \cup \mathcal{P}$-intégrable.

Proposition 4.3.-

i) $$(E_Q\, f_n(x) - E_Q\, \tilde{f}_n(x)) = Q(U^c) \left[\frac{1}{Q(U^c)} \int_{]-H/h_n, H/h_n[^c} K(y)\, f_Q(x-h_n y)\, dy - \frac{1}{P(U^c)} \int_{]-H/h_n, H/h_n[^c} K(y)\, f_P(x-h_n y)\, dy \right]$$

ii) si $m(P) = m(Q)$ alors $E_Q\, f_n(x) = E_Q\, \tilde{f}_n(x)$.

Pour la preuve de cette proposition voir A. BERLINET (1984-a) page VI-13 ; des corollaires 1 et 2 page VI-14 on tire la proposition suivante :

Proposition 4.4.- Si K est de carré intégrable alors :

i) $V\, f_n(x) - V\, \tilde{f}_n(x) \geq Q(U^c) \left\{ \frac{1}{Q(U^c)} \int_{]-H/h_n,H/h_n[^c} K(y)\, f_Q(x-h_n\, y)\, dy - \right.$

$$\left. \frac{1}{P(U^c)} \int_{]-H/h_n,H/h_n[^c} K(y) f_P(x-h_n y)\, dy \right\} \cdot \left\{ Q(U) \left[\frac{1}{Q(U^c)} \int_{]-H/h_n,H/h_n[^c} K(y) f_Q(x-h_n y)\, dy + \right.\right.$$

$$\left.\left. \frac{1}{P(U^c)} \int_{]-H/h_n,H/h_n[^c} K(y) f_P(x-h_n y)\, dy \right] + 2 \int_{]-H/h_n,H/h_n[} K(y)\, f_Q(x-h_n y)\, dy \right\}$$

ii) si $m(P) = m(Q)$ alors :

$$R_o - R = Vf_n(x) - V\tilde{f}_n(x) = \frac{1}{n} \left\{ \frac{1}{h_n} \int_{]-H/h_n,H/h_n[^c} K^2(y) f_Q(x-h_n y)\, dy - \right.$$

$$\left. \frac{1}{Q(U^c)} \left(\int_{]-H/h_n,H/h_n[^c} K(y) f_Q(x-h_n y)\, dy \right)^2 \right\} \geq 0$$

où R_o est le risque quadratique associé à $f_n(x)$, R celui de $\tilde{f}_n(x)$.

Remarques :

- La proposition 4.4. i') fournit une minoration de $R_o - R$ tandis que ii') fournit une partie de l'ensemble des éléments de $\mathcal{P}$ pour lesquels on a diminution du risque quadratique, ensemble difficile à déterminer en toute généralité,

- L'hypothèse $m(P) = m(Q)$ entraîne que $R_o \geq R$ et la quantité $R_o - R$ *reste invariante si on remplace* P *par* P' *telle que* $m(P) = m(P')$. Si $K\, 1_{U^c} \neq a\, 1_{U^c} (a \in \mathbb{R})$ Q - p.s. alors $R_o > R$ i.e. l'amélioration est stricte.

2°) *Etude asymptotique du risque quadratique de* $\tilde{f}_n(x)$.

La méthode du conditionnement a pour but une amélioration de l'estimation à n fixé. Il est néanmoins important d'étudier le comportement asymptotique du risque quadratique, il nous fournira des indications sur le choix de H, h_n et éventuellement du noyau K.

a) Etude du biais :

Proposition 4.5.-

a_1) Si $|K|$ est λ-intégrable, $r = Q(U^c)/P(U^c)$, $H \to 0$, $h_n \xrightarrow[n\to\infty]{} 0$, $H/h_n \xrightarrow[n\to\infty]{} \alpha \in [0, +\infty]$.

$$\text{Alors } \lim_{n\to\infty} E_Q \tilde{f}_n(x) = f_Q(x) \int_{]-\alpha,\alpha[} K \, d\lambda + f_P(x) \int_{]-\alpha,\alpha[^c} K \, d\lambda$$

a_2) Si les versions $(f_Q)_{Q\in \mathcal{P}}$ admettent une dérivée d'ordre 4 en x la fonction K vérifie :

$$\begin{cases} K \text{ est paire} \\ y^i K \text{ et } y^i K^2 \text{ sont } \lambda\text{-intégrables, pour } i \in \overline{0,4} \end{cases}$$

$$\text{Alors } [E_Q \tilde{f}_n(x) - f_Q(x)]^2 = [I - f_Q(x)]^2 + h_n^2 [I - f_Q(x)] [f''_Q(x) I_{21} + r f''_P(x) J_{21}]$$

$$+ \frac{h_n^4}{4} [(f''_Q(x) I_{21} + r f''_P(x) J_{21})^2 + \frac{1}{3}(I - f_Q(x)) (f_Q^{(4)}(x) I_{41} + r f_P^{(4)}(x) J_{41})] + h_n^4 o(1).$$

où

$$I = f_Q(x) I_{01} + r f_P(x) J_{01}$$

et

$$I_{k\ell} = \int_{]-H/h_n, H/h_n[} y^k K^\ell(y) \, dy, \quad J_{k\ell} = \int_{]-H/h_n, H/h_n[^c} y^k K^\ell(y) \, dy,$$

$$k \in \overline{0,4} \text{ et } \ell \in \overline{1,2}$$

Preuve de la proposition 4.5. :

$a_1)$ on sait que

$$E_Q \tilde{f}_n(x) = \int_{]-H/h_n,H/h_n[} K(y)\, f_Q(x-h_n y)\, dy + r \int_{]-H/h_n,H/h_n[^c} K(y)\, f_P(x-h_n y)\, dy$$

f_Q étant continue au voisinage de x , $K(y)\, f_Q(x-h_n y)\, 1_{]-H/h_n,H/h_n[}(y)$ converge λ p-p. vers $K(y)\, f_Q(x)\, 1_{]-\alpha,\alpha[}(y)$; par ailleurs

$$|K(y)\, f_Q(-xh_n y)\, 1_{]-H/h_n,H/h_n[}(y)| \leq |K(y)|\, M_1 \quad (\text{où } M_1 = \operatorname*{Sup}_y f_Q(y))$$

le théorème de la convergence dominée, nous donne donc ,

$$\int_{]-H/h_n,H/h_n[} K(y)\, f_Q(x-h_n y)\, dy \xrightarrow[n\to\infty]{} f_Q(x) \int_{]-\alpha,\alpha[} K(y)\, dy$$

de même on obtient : $\int_{]-H/h_n,H/h_n[^c} K(y)\, f_P(x-h_n y)\, dy \xrightarrow[n\to\infty]{} f_P(x) \int_{]-\alpha,\alpha[^c} K(y)\, dy.$

Ensuite $H \xrightarrow[n\to\infty]{} 0 \implies Q(U^c) \xrightarrow[n\to\infty]{} 1$ et $P(U^c) \xrightarrow[n\to\infty]{} 1$ donc $r \xrightarrow[n\to\infty]{} 1$

d'où

$$\lim_{n\to\infty} E_Q \tilde{f}_n(x) = f_Q(x) \int_{]-\alpha,\alpha[} K\, d\lambda + f_P(x) \int_{]-\alpha,\alpha[^c} K\, d\lambda\ .$$

$a_2)$ voir A. BERLINET (1984-a) proposition 8 page VI-18 .

Remarques :

* Si $\int K\, d\lambda = 1$ alors $\lim_{n\to\infty} (E_Q \tilde{f}_n(x) - f_Q(x)) =$ $(f_P(x)-f_Q(x)) \cdot \int_{]-\alpha,\alpha[^c} K\, d\lambda$, par conséquent $\tilde{f}_n(x)$ sera asymptotiquement

sans biais si $\int_{]-\alpha,\alpha[^c} K\,d\lambda = 0$ ou bien si $f_P(x) = f_Q(x)$; dans le cas où $f_P(x) \neq f_Q(x)$ le biais asymptotique de $\tilde{f}_n(x)$ sera strictement plus petit que $|f_P(x) - f_Q(x)|$ si et seulement si $\left|\int_{]-\alpha,\alpha[^c} K\,d\lambda < 0\right.$ entraîne que $(\lim\limits_{n\to\infty} E_Q\,\tilde{f}_n(x) - f_Q(x))$ et $(f_P(x) - f_Q(x))$ sont de signes contraires, ce qui peut présenter un intérêt si on veut construire les intervalles de confiance pour $\tilde{f}_n(x)$.

** Nous avons supposé que H dépend de n et tend vers 0 lorsque $n \to \infty$ car si H ne tend pas vers 0 il existe un réel β et une sous-suite $(H_{n_k})_k$ telle que $H_{n_k} \geq \beta$, $\forall k$, ce qui entraine que $\tilde{f}_n(x)$ et $f_n(x)$ coïncident asymptotiquement puisque dans le cas où $\beta = +\infty$, U contient λ-p.s. toutes les observations ce qui est dépourvu d'intérêt pour nous, et dans le cas où $\beta \in\,]0\,,+\infty[$ la quantité

$\frac{1}{P(U^c)} \int_{]-H/h_{n_k},H/h_{n_k}[^c} K(y)\, f_P(x-h_{n_k}y)\,dy$ tend vers 0 lorsque $n_k \to \infty$,

par conséquent, les observations qui tombent hors de U auront un poids voisin de 0 i.e. $\tilde{f}_n(x)$ reste voisin de $f_n(x)$.

*** La condition $H/h_n \xrightarrow[n\to\infty]{} \alpha \in [0,+\infty]$ est parfaitement réalisable on aura $\alpha = 0$ si $H = O(h_n)$, $\alpha = +\infty$ si $h_n = O(H)$, le cas $\alpha \in\,]0,+\infty[$ englobe le cas $H = \alpha\, h_n$, $\forall\, n$, qui à son tour généralise le choix de A. BERLINET (1984-a) i.e. $(\alpha = 1)$.

b) Etude de la variance :

Proposition 4.6.- Si K^2 est λ-intégrable

$$\lim_{n\to\infty} n\,h_n = +\infty$$

Alors :
$$V(\tilde{f}_n(x)) = \frac{1}{n\,h_n} f_Q(x) \int_{]-H/h_n,H/h_n[} K^2\,d\lambda + O(1/n\,h_n) .$$

Preuve : On a :

$$\frac{1}{1/nh_n}\left(V\,\tilde{f}_n(x) - \frac{1}{n\,h_n}\,f_Q(x)\int_{]-H/h_n,H/h_n[} K^2\,d\lambda\right) = \int_{]-H/h_n,H/h_n[} K^2(y)\ f_Q(x-h_ny)dy$$

$$- f_Q(x)\int_{]-H/h_n,H/h_n[} K^2(y)\ dy - h_n\left[\int_{]-H/h_n,H/h_n[} K(y)\ f_Q(x-h_ny)dy\right]^2 +$$

$$h_n\ r\int_{]-H/h_n,H/h_n]^c} K(y)f_P(x-h_ny)dy\left\{\frac{Q(U)}{P(U^c)}\int_{]-H/h_n,H/h_n[^c} K(y)f_P(x-h_ny)dy - \right.$$

$$\left. 2\int_{]-H/h_n,H/h_n[} K(y)f_Q(x-h_ny)dy\right\} .$$

Les hypothèses faites sur H , h_n , K , f_Q et f_P permettent de conclure. On se propose de déterminer les premiers termes du développement asymptotique de l'erreur quadratique associée à $\tilde{f}_n(x)$ et d'en donner quelques applications pour des noyaux usuels : normal, unité, Epanechnikov ...

Le risque quadratique associé à $\tilde{f}_n(x)$ sera défini par

$$R = V\,\tilde{f}_n(x) + (E_Q\,\tilde{f}_n(x) - f_Q(x))^2$$

<u>*Lemme 4.1.*</u>- Si $H \longrightarrow 0$, si f_P et f_Q admettent une dérivée d'ordre 4 en x . Alors :

$$\begin{aligned} r = 1 &+ 2H(f_P(x) - f_Q(x) \\ &+ 4\,H^2\,f_P(x)\ (f_P(x) - f_Q(x)) \\ &+ H^3\ (8\,f_P^3(x) - 8\,f_Q(x)\ f_P^2(x) + \tfrac{1}{3}\,f_P''(x) - \tfrac{1}{3}\,f_Q''(x)) \\ &+ H^4\ (16\,f_P^{(4)}(x) - 16\,f_Q(x)\ f_P^3(x) + \tfrac{4}{3}\,f_P(x)\ f_P''(x) \\ &\qquad - \tfrac{2}{3}\,f_P(x)\ f_Q''(x) - \tfrac{2}{3}\,f_Q(x)\ f_P''(x)) \\ &+ o(H^4) \quad . \end{aligned}$$

Preuve : Par application de la formule de Taylor - Young à l'ordre 4 on a :

$$\forall Q_1 \in \mathcal{P} \quad F_{Q_1}(x+H) = F_{Q_1}(x) + \sum_{i=1}^{4} \frac{H^i}{i!} f_{Q_1}^{(i-1)}(x) + O(H^4)$$

où F_{Q_1} est la f.r. associée à Q_1

d'où :

$$① \qquad Q_1(U^c) = 1 - Q_1(U) = 1 - 2H\, f_{Q_1}(x) - \frac{H^3}{3} f''_{Q_1}(x) + O(H^3)$$

donc

$$\frac{1}{P(U^c)} = 1 + 2H\, f_P(x) + 4\, H^2\, f_P^2(x) + H^3\, [8\, f_P^3(x) + \frac{1}{3} f''_P(x)]$$

$$② \qquad + H^4\, [\frac{4}{3} f_P(x)\, f''_P(x) + 16\, f_P^4(x)] + O(H^4) \quad .$$

Enfin en effectuant le produit de ① par ② , on obtient le lemme 4.1.

Il est facile de déduire du lemme 4.1. le développement de $[E_Q \tilde{f}_n(x) - f_Q(x)]$ et du risque quadratique en fonction des puissances de H et h. Nous allons voir en quoi cela consiste pour des noyaux usuels.

Applications

a) Noyau normal $\quad K(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$.

Lemme 4.2.- On a :

$$I_{01} = [1 - \exp(-H^2/(2h^2))]^{1/2}$$

$$I_{21} = ([1 - \exp(-H^2/(2h^2))]^{1/2} - \frac{2H}{h\sqrt{2\pi}} \exp(-H^2/(2h^2)))$$

$$J_{21} = 1 - I_{21}$$

$$I_{41} = - \frac{2}{\sqrt{2\pi}} \exp(- H^2/(2h^2)) \left[\frac{H^3}{h^3} - \frac{3H}{h}\right] + 3 \, (1 - \exp(- H^2/(2h^2)))^{1/2}$$

$$J_{41} = 3 - I_{41}$$

$$I_{02} = \frac{1}{2\sqrt{\pi}} \, . \, (1 - \exp(- H^2/h^2))^{1/2} \quad .$$

Preuve : On pose $L = \frac{1}{\sqrt{2\pi}} \int_0^{H/h} e^{-y^2/2} \, dy \Longrightarrow I_{01} = 2L$ et

$$L^2 = \frac{1}{2\pi} \int_0^{H/h} \int_0^{H/h} e^{-(y^2+z^2)/2} \, dy \, dz$$

avec le changement de variable $\begin{cases} y = r \cos \theta \\ z = r \sin \theta \end{cases}$ on obtient :

$$L^2 = \frac{1}{2\pi} \int_0^{H/h} \int_0^{\pi/2} r \, e^{-r^2/2} \, dr \, d\theta = \frac{1}{4} \, (1 - \exp(- H^2/(2h^2)))$$

d'où $\quad I_{01} = \left[1 - \exp(- H^2/(2h^2))\right]^{1/2} \quad$ car $\quad I_{01} \geq 0$

et $\quad \int_{\mathbb{R}} K \, d\lambda = 1 \Longrightarrow J_{01} = 1 - I_{01} \quad .$

Par ailleurs une intégration par parties nous donne

$$\forall p \geq 0 \quad I_{2p+2,1} = (2p+1) \, I_{2p,1} + \frac{1}{\sqrt{2\pi}} \left[- y^{2p+1} \, e^{-y^2/2}\right]_{-H/h}^{H/h}$$

$$= (2p+1) \, I_{2p,1} - \frac{2}{\sqrt{2\pi}} \, (H/h)^{2p+1} \exp(- H^2/(2h^2))$$

par conséquent $\quad I_{21} = I_{01} - \frac{2}{\sqrt{2\pi}} \frac{H}{h} \exp(- H^2/(2 \, h^2))$

et $\quad I_{41} = 3 \, I_{01} - \frac{2}{\sqrt{2\pi}} \frac{H}{h} \exp(- H^2/(2 \, h^2)) \left[3 + \frac{H^2}{h^2}\right] \quad .$

D'autre part $\quad \int_{\mathbb{R}} y^{2p} \, K(y) \, dy = (\frac{1}{2})^p \, \frac{(2p)!}{p!} \quad$ entraîne que

$$J_{21} = 1 - I_{21} \quad \text{et} \quad J_{41} = 3 - I_{41} \, .$$

Pour I_{02} le changement de variable $y = z/\sqrt{2}$ permet de conclure.

Remarques :

1) si $H/h \xrightarrow[n\to\infty]{} 0$ alors $I_{k\ell} \to 0$ et $J_{2p,1} \to (\frac{1}{2})^p \frac{(2p)!}{p!}$

par conséquent $\lim\limits_{n\to\infty} R = \lim\limits_{n\to\infty} (E_Q \tilde{f}_n(x) - f_Q(x))^2 = (f_Q(x) - f_P(x))^2$.

On a donc intérêt à se contenter de la pré-estimation f_P .

2) si $H/h \xrightarrow[n\to\infty]{} \infty$ alors $J_{k\ell} \to 0$ et $I_{2p,1} \to (\frac{1}{2})^p \frac{(2p)!}{p!}$

et $I_{02} \to 1/2\sqrt{\pi}$.

Dans ce cas $\lim\limits_{n\to\infty} R = \lim\limits_{n\to\infty} (E_Q \tilde{f}_n(x) - f_Q(x))^2 = 0$

ceci peut être obtenu lorsqu'on prend $H = h^\varepsilon$ avec $\varepsilon < 1$.

3) si $H/h \xrightarrow[n\to\infty]{} \alpha \in]0, +\infty[$ alors on obtient

$$\lim_{n\to\infty} R = \left[1 - (1 - \exp(-\alpha^2/2))^{1/2}\right] (f_Q(x) - f_P(x))^2$$

ceci incite à suivre la démarche suivante :

- se fixer un α assez grand, afin que le terme $(1 - (1 - \exp(-\alpha^2/2))^{1/2})$ soit très petit.

- remplacer H/h par α dans I_{ij} et J_{ij} , ce qui permettra d'avoir un risque quadratique plus simple à minimiser.

b) <u>Noyau unité</u> $K(y) = \frac{1}{2a} 1_{[-a,a]}(y)$ <u>où</u> $a > 0$.

Dans ce cas on doit imposer à H/h de vérifier $H/h < a$, $\forall n$, sinon $\tilde{f}_n(x) = f(x)$.

Lemme 4.3.- On a :

$$I_{01} = \frac{1}{a} \cdot \frac{H}{h} \qquad J_{01} = 1 - \frac{1}{a} \cdot \frac{H}{h}$$

$$I_{21} = \frac{1}{3a} \frac{H^3}{h^3} \qquad J_{21} = \frac{1}{3a} \left(a^3 - \frac{H^3}{h^3}\right)$$

$$I_{41} = \frac{1}{5a} \frac{H^5}{h^5} \qquad J_{41} = \frac{1}{5a} \left(a^5 - \frac{H^5}{h^5}\right)$$

$$I_{02} = \frac{1}{2a^2} \frac{H}{h} \quad .$$

Démonstration : immédiate.

Remarques : Dans ce cas il n'est plus question de faire tendre H/h vers un $\alpha \in [a, +\infty]$, reste donc deux cas.

a) $H/h \xrightarrow[n\to\infty]{} 0$ voir remarque cas normal .

b) si $H/h \xrightarrow[n\to\infty]{} \alpha < a$ alors $\lim\limits_{n\to\infty} R = (1 - \frac{\alpha}{a})^2 \ (f_Q(x) - f_P(x))^2$.

Donc si on se fixe un α très voisin de a et on remplace H/h par α dans I_{ij} , J_{ij} alors on peut espérer obtenir, asymptotiquement, un risque quadratique voisin de 0 , dont l'expression est plus facile à minimiser.

c) Noyau d'Epanechnikov. $K(y) = \frac{3}{4\beta} (1 - y^2/\beta^2) \ 1_{[-\beta,\beta]}(y)$, $\beta > 0$.

Pour ne pas avoir $f_n(x) \equiv \tilde{f}_n(x)$, H/h doit vérifier : $H/h < \beta$, $\forall n$.

Lemme 4.4.- On a :

$$I_{01} = \frac{3}{2\beta} \frac{H}{h} \left(1 - \frac{1}{3} \left(\frac{H}{\beta h}\right)^2\right) \qquad J_{01} = 1 - I_{01}$$

$$I_{21} = \frac{3}{2\beta} \left(\frac{H}{h}\right)^3 \left(\frac{1}{3} - \frac{1}{5}\left(\frac{H}{\beta h}\right)^2\right) \qquad J_{21} = \frac{\beta^2}{5} - I_{21}$$

$$I_{41} = \frac{3}{2\beta} \left(\frac{H}{h}\right)^5 \left(\frac{1}{5} - \frac{1}{7}\left(\frac{H}{\beta h}\right)^2\right) \qquad J_{41} = \frac{3}{35} \beta^4 - I_{41}$$

$$I_{20} = \frac{9}{8\beta^2} \frac{H}{h} \left[1 - \frac{2}{3} \left(\frac{H}{\beta h}\right)^2 + \frac{1}{5} \left(\frac{H}{\beta h}\right)^4\right]$$

Preuve : immédiate.

Remarque : On constate que les I_{ij} et J_{ij} associés au noyau d'Epanechnikov et ceux associés au noyau unité ont la même forme, par conséquent leurs risques quadratiques ont un même comportement asymptotique. Ce sera d'ailleurs la même chose pour tout autre noyau symétrique à support compact et coïncidant sur ce support avec un polynôme de degré quelconque.

Conclusion : A partir de ces trois applications on constate que le risque quadratique de l'estimateur conditionné est beaucoup plus difficile à étudier que celui de l'estimateur de Parzen - Rosenblatt, estimateur pour lequel on obtient un risque quadratique en $O(\frac{1}{nh} + h^4)$, en imposant à K d'être pair, λ-intégrable, $\int K\, d\lambda = 1$, $\int K^2\, d\lambda < \infty$ et d'avoir des moments d'ordre 4 finis, alors que pour l'estimateur conditionné on doit imposer des conditions du style : $\int_{]-H/h,H/h[} K\, d\lambda = 1$, $\int_{]-H/h,H/h[^c} K\, d\lambda = 0$ afin d'avoir un risque quadratique facile à minimiser, ce qui, par ailleurs, impose au noyau K de dépendre des quantités inconnues h et H.

III.3. - Convergence presque - complète de $\tilde{f}_n(x)$.-

Nous allons établir une condition suffisante pour que $\tilde{f}_n(x)$ converge presque complètement vers $\bar{f}_Q(x) = \lim_{n\to\infty} E_Q\, \tilde{f}_n(x)$, sous l'hypothèse que f_Q et f_P appartiennent à $\mathcal{P}$ et que K est un noyau d'intégrale 1.

Proposition 4.8.- Si f_Q et f_P sont majorées par M_1

K est borné par M_2 et $\int K\, d\lambda = 1$, $\int K^2\, d\lambda < \infty$

$\forall \beta > 0$, $\sum_{n\geq 1} \operatorname{Exp}(-n\beta h_n) < +\infty$

Alors $\tilde{f}_n(x) \xrightarrow[n\to\infty]{p.co.} \bar{f}_Q(x)$.

Preuve : On a $\tilde{f}_n(x) - \bar{f}_Q(x) = (\tilde{f}_n(x) - E_Q\,\tilde{f}_n(x)) + (E_Q\,\tilde{f}_n(x) - \bar{f}_Q(x))$ par définition de $\bar{f}_Q(x)$, le terme $E_Q(\tilde{f}_n(x)) - \bar{f}_Q(x)$ converge vers 0, il suffit donc d'établir la convergence p.co. de $\tilde{f}_n(x) - E_Q\,\tilde{f}_n(x)$ vers 0.

Pour cela posons

$$\overline{K_{h_n}}(x,X_i) = \frac{1}{nh_n}\left\{K\left(\frac{x-X_i}{h_n}\right) 1_U(X_i) - \int_U K\left(\frac{x-y}{h_n}\right) f_Q(y)\,dy \right.$$

$$\left. + \int_{U^c} K\left(\frac{x-y}{h_n}\right) f_P(y)\,dy \left(\frac{1}{P(U^c)}\, 1_{U^c}(X_i) - r\right)\right\}$$

avec $\qquad U = \left]x - H,\ x + H\right[$ et $r = \dfrac{Q(U^c)}{P(U^c)}$

$$\bar{S}_n(x) = \tilde{f}_n(x) - E_Q\,\tilde{f}_n(x) = \sum_{i=1}^{n} \bar{K}_{h_n}(x,X_i)\ .$$

L'inégalité de Hoeffding va nous permettre de majorer $\Pr(\bar{S}_n(x) \geq \varepsilon)$, en effet : si $(Y_n)_{n\geq 1}$ est une suite de v.a. indépendantes, centrées et majorées par un réel $b > 0$, alors pour $\varepsilon \in \left]0,b\right[$

$$(*) \qquad \Pr(S_n \geq n\varepsilon) \leq \mathrm{Exp}\left(-\frac{n\varepsilon}{b}\left(1 + \frac{\sigma^2}{b\varepsilon}\right)\mathrm{Log}\left(1 + \frac{b\varepsilon}{\sigma^2}\right) - 1\right)$$

où $\qquad S_n = \sum_{k=1}^{n} Y_k \quad$ et $\quad \sigma^2 = \frac{1}{n}\sum_{k=1}^{n} E\,Y_k^2\ .$

Donc si on remplace Y_k par $n\,\bar{K}_{h_n}(x,X_k)$ alors on aura bien des v.a. indépendantes centrées telles que :

$$\forall k \in \overline{1,n} \quad \left|n\,\overline{K_{h_n}}(x,X_k)\right| \leq \left|\frac{1}{h_n} K\left(\frac{x-X_k}{h_n}\right) 1_U(X_k) - \frac{1}{h_n}\int_U K\left(\frac{x-y}{h_n}\right) f_Q(y)\,dy\right|$$

$$+ \left(\left|1_{U^c}(X_i) - Q(U^c)\right|\right)\frac{1}{h_n P(U^c)}\left|\int_{U^c} K\left(\frac{x-y}{h_n}\right) f_P(y)\,dy\right|$$

or $\quad \left|\frac{1}{h_n} K(\frac{x-X_k}{h_n}) 1_U(X_k) - \frac{1}{h_n}\int_U K(\frac{x-y}{h_n}) f_Q(y)\, dy\right| \leq \frac{2 M_2}{h_n}$

et $\quad |1_{U^c}(X_i) - Q(U^c)|\left|\frac{1}{h_n P(U^c)}\right| \left|\int_{U^c} K(\frac{x-y}{h_n}) f_P(y)\, dy\right| \leq \frac{2 M_2}{h_n P(U^c)}$

donc $\quad |n\, \overline{K_{h_n}}(x,X_k)| \leq \frac{2 M_2}{h_n} (1 + \frac{1}{P(U^c)})$, par ailleurs on a :

$P(U^c) = 1 - \int_{x-H}^{x+H} f_P(y)\, dy \geq 1 - 2 M_1 H \geq \frac{1}{2}$ pour n assez grand car $H \xrightarrow[n\to\infty]{} 0$,

par conséquent $\quad |n\, \overline{K_{h_n}}(x,X_k)| \leq \frac{6 M_2}{h_n} = b$,

d'autre part nous avons : $\quad \sigma^2 = \frac{1}{n}\sum_{k=1}^{n} E_Q(n\, \overline{K_{h_n}}(x,X_k))^2$

$$= E_Q(n\, \overline{K_{h_n}}(x,X_1))^2 = n\, \mathrm{Var}(f_n(x))$$

$$\Longrightarrow \sigma^2 = \frac{1}{h_n^2}\int_U K^2(\frac{x-y}{h_n})\, f_Q(y)\, dy + Q(U^c)\, (\frac{1}{h_n P(U^c)}\int_{U^c} K(\frac{x-y}{h})\, f_P(y)\, dy)^2$$

$$- (\frac{1}{h_n}\int_U K(\frac{x-y}{h_n})\, f_Q(y)\, dy + r\int_{U^c} K(\frac{x-y}{h_n})\, f_P(y)\, dy)^2$$

d'où

$$\sigma^2 \leq \frac{1}{h_n^2}\int_U K^2(\frac{x-y}{h_n})\, f_Q(y)\, dy + \frac{r}{h_n^2\, P(U^c)}\, (\int_{U^c} K(\frac{x-y}{h_n})\, f_P(y)\, dy)^2$$

en utilisant l'inégalité de Schwarz on obtient

$$(\int_{U^c} K(\frac{x-y}{h_n})\, f_P(y)\, dy)^2 \leq \int_{U^c} K^2(\frac{x-y}{h_n})\, dy \int_{U^c} f_P^2(y)\, dy$$

$$\Longrightarrow \sigma^2 \leq \frac{M_1}{h_n^2}\left[\int_{]x-H,x+H[} K^2(\frac{x-y}{h_n})\, dy + r\int_{]x-H,x+H[^c} K^2(\frac{x-y}{h_n})\, dy\right]$$

$$\leq \frac{M_1}{h_n}\int_{]-H/h_n,H/h_n[} K^2(y)\, dy + \frac{1}{P(U^c)}\int_{]-H/h_n,H/h_n[^c} K^2(y)\, dy$$

$$\leq \frac{M_1}{h_n} \cdot B \cdot (1 + \frac{1}{P(U^c)}) \qquad \text{où} \quad B = \int K^2\, d\lambda$$

ainsi $\sigma^2 \leq \frac{2\, M_1}{h_n} \cdot B$, ce qui implique que le terme $\frac{\sigma^2}{b\varepsilon}$, de l'inégalité de Hoeffding, est majoré par $\frac{1}{\varepsilon} \cdot \frac{M_1\, B}{3\, M_2}$, par ailleurs la fonction

$h(u) = (1+u) \log(1 + \frac{1}{u}) - 1$ est strictement positive et décroissante sur $]0\ ,\ +\infty[$, donc

$$(1 + \frac{\sigma^2}{b\varepsilon}) \log(1 + \frac{b\varepsilon}{\sigma^2}) - 1 \geq (1 + \frac{M_1\, B}{3\, M_2\, \varepsilon}) \operatorname{Log}(1 + \frac{3\, M_2\, \varepsilon}{M_1\, B}) - 1$$

$$= c > 0$$

l'inégalité (*) et les hypothèses faites sur h_n entraînent donc que

$$\Pr(|\bar{S}_n(x)| \geq \varepsilon) = \exp(- n\, \varepsilon .\, c \cdot \frac{h_n}{6\, M_2})$$

et

$$\sum_{n\geq 1} \Pr(|\bar{S}_n(x)| \geq \varepsilon) \leq \infty \quad \text{i.e.} \quad \bar{S}_n(x) \xrightarrow[n\to\infty]{\text{p.co.}} 0 \ . \ \blacksquare$$

III.4. - Etude de la loi limite.-

Dans ce paragraphe, nous commençons par reprendre un résultat de J. AKONOM (1978), concernant la loi limite d'une classe d'estimateurs de la partie continue d'une densité de probabilité ; résultat que nous avons légèrement modifié et auquel nous avons donné une seconde démonstration. Ensuite nous avons établi les conditions sous lesquelles une vaste classe d'estimateurs conditionnés, qui contient notamment les estimateurs à noyaux, les estimateurs splines, les estimateurs par les fonctions orthogonales etc ..., converge en loi vers une v.a. normale. Enfin en application nous avons donné une condition suffisante de convergence en loi de $f_n(x)$ et de $\tilde{f}_n(x)$ (définis page 70).

III. 4. 1. - Loi limite de la classe d'estimateurs de J. BLEUEZ.

J. BLEUEZ et D. BOSQ (1976) ont défini une vaste classe d'estimateurs de la densité d'une loi de probabilité, classe qui contient de nombreux estimateurs classiques : estimateurs à noyaux, estimateurs splines, estimateurs par les fonctions orthogonales etc ... ; J. AKONOM (1978), a repris cette classe, pour estimer la densité de la partie continue d'une probabilité mixte et a établi une condition suffisante pour la convergence en loi, nous allons reprendre ce résultat, avec quelques légères modifications, mais auparavant fixons quelques notations :

Soit (Ω, A, Pr) un espace probabilisé, $(E, \mathcal{B}, \mu)$ un espace mesuré où E est un espace topologique, $\mathcal{B}$ sa tribu borélienne. On définit une v.a. X sur (Ω, A) et à valeurs dans $(E, \mathcal{B})$; dont la loi P_X admet une densité $\frac{dP_X}{d\mu}$ par rapport à μ (dont une version sera notée f).

On considère un échantillon $(X_1, \ldots, X_n)$ de la v.a. X et I une partie non bornée de $\mathbb{R}_+$; on se donne une famille $(K_r(x,t),\ r \in I)$ de fonctions réelles mesurables à deux variables $(x,t) \in E \times E$, l'estimateur de $f(x)$ défini par J. BLEUEZ est donné par : $f_n^1(x) = \frac{1}{n} \sum_{j=1}^{n} K_{r(n)}(x, X_j)$, $x \in E$ où $r(n) \in I$ et $r(n) \xrightarrow[n \to \infty]{} +\infty$.

Dans toute la suite μ sera supposée non atomique et α désignera un nombre strictement positif attaché à la famille $\{K_r(x,t),\ r \in I\}$. Nous utiliserons les hypothèses suivantes :

I) f est bornée, on posera $M_1 = \sup_y f(y)$,

II) $\forall x \in E$, $\int_E K_r(x,y)\, f(y)\, d\mu(y)$ existe ,

III) $\forall x \in E$, $\exists A_2(x) > 0$ tel que pour n assez grand

$$\frac{1}{r^{\alpha}(n)} \operatorname{Var}(K_{r(n)}(x,X_1)) \leq A_2(x) .$$

IV) $\forall x \in E$, $\exists A_3(x) > 0$ tel que pour n assez grand

$$\int_E |K_{r(n)}(x,y)|^3 \, d\mu(y) \leq A_3(x) \; r^{2\alpha}(n)$$

V) $\dfrac{r^{\alpha}(n)}{n} \xrightarrow[n\to\infty]{} 0$

Nous verrons par la suite que ces conditions ne sont pas très restrictives et qu'elles sont vérifiées dans les cas usuels.

Proposition 4.9.- Si les hypothèses (I) à (V) sont vérifiées. Alors :

(1) la v.a. $\dfrac{f_n^1(x) - E\, f_n^1(x)}{\sqrt{\operatorname{Var} f_n^1(x)}}$ converge en loi vers une v.a. normale centrée réduite lorsque $n \to \infty$.

Preuve : Posons $V_{nj} = K_{r(n)}(x,X_j)$, $j \in \overline{1,n}$, $n \geq 1$, ainsi $f_n^1(x) = \frac{1}{n} \sum_{j=1}^{n} V_{nj}$ est une moyenne de v.a. indépendantes de même loi que $V_n = K_{r(n)}(x,X)$.

(1) est équivalent à

$$(1') \quad \forall c \in \mathbb{R} , \quad \lim_{n\to\infty} \Pr\left\{\frac{f_n^1(x) - E\, f_n^1(x)}{\sigma(f_n^1(x))} \leq c\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c} e^{-y^2/2} \, dy$$

or d'après M. LOEVE (1963) p. 316 une C.N.S. pour que (1') soit vérifiée est :

$$(2) \quad \forall \varepsilon > 0 , \quad n \Pr\left\{\frac{|V_n - E\, V_n|}{\sigma(V_n)} \geq \varepsilon \sqrt{n}\right\} \xrightarrow[n\to\infty]{} 0$$

une application de l'inégalité de Markov nous donne

$$(3) \qquad \Pr\left\{\frac{|V_n - E\,V_n|}{\sigma(V_n)} \geq \varepsilon\sqrt{n}\right\} \leq \frac{E|V_n - E\,V_n|^3}{\varepsilon^3\, n^{3/2}\, \sigma^3(V_n)}$$

il suffira de montrer que $\lim\limits_{n\to\infty} \dfrac{E|V_n - E\,V_n|^3}{n^{1/2}\,\sigma^3(V_n)} = 0$, en effet en utilisant l'inégalité

$$(a + b)^3 \leq 4(a^3 + b^3) \qquad \forall a,\ b \geq 0$$

nous obtenons

$$E|V_n - E\,V_n|^3 \leq 4(E|V_n|^3 + |E\,V_n|^3)$$

par ailleurs l'inégalité de Jensen entraîne que $|E\,V_n|^3 \leq E|V_n|^3$ ainsi nous avons $E|V_n - E\,V_n|^3 \leq 8\,E|V_n|^3$

or
$$\begin{aligned} E|V_n|^3 &= \int_E |K_{r(n)}(x,y)|^3\, f(y)\, d\mu(y) \\ &\leq M_1 \int_E |K_{r(n)}(x,y)|^3\, d\mu(y) \\ &\leq M_1\, A_3(x)\, r^{2\alpha}(n) \text{ , ceci d'après l'hypothèse (IV)} \end{aligned}$$

nous obtenons donc
$$\begin{aligned} \frac{E|V_n - E\,V_n|^3}{n^{1/2}\,\sigma^3(V_n)} &\leq \frac{8\,E|V_n|^3}{\sqrt{n}\,\sqrt{\mathrm{Var}(V_n)}^3} \\ &\leq \frac{8\,M_1\,A_3(x)\,r^{2\alpha}(n)}{n^{1/2}\,\left(\frac{1}{r^\alpha(n)}\,\mathrm{Var}(V_n)\right)^{3/2} .\, r^{3\alpha/2}(n)} \end{aligned}$$

par conséquent

$$\lim_{n\to\infty} \frac{E|V_n - E\,V_n|^3}{n^{1/2}\,\sigma^3(V_n)} \leqslant \lim_{n\to\infty} \left\{ \frac{1}{\left(\frac{1}{r^{\alpha}(n)}\ \mathrm{Var}(V_n)\right)^{3/2}} \cdot \left(\frac{r^{\alpha}(n)}{n}\right)^{1/2} \right\} .\ 8\ M_1\ A_3(x)$$

quantité qui converge vers 0 lorsque n tend vers $+\infty$ puisque d'après l'hypothèse (III) $\lim\limits_{n\to\infty} \frac{1}{r^{\alpha}(n)}$ $\mathrm{Var}(V_n) \leqslant A_2(x)$ et d'après (V) $r^{\alpha}(n)/n \xrightarrow[n\to\infty]{} 0$, ce qui permet de conclure. ∎

Nous obtenons le même résultat de convergence en utilisant le théorème central limite de LINDEBERG-FELLER et les hypothèses suivantes :

a) $\exists m > 0$ tel que $\forall y \in E \quad f(y) \geqslant m$

b) $\sup\limits_{n} |\bar{f}_n(x)| < +\infty$ où $\bar{f}_n(x) = \int_E K_{r(n)}(x,y)\ f(y)\ d\mu(y)$

c) $\forall x \in E$, $\exists A_1(x) > 0$ tel que pour $r(n)$ assez grand dans I

$$\sup_{y} |K_{r(n)}(x,y)| \leqslant A_1(x)\ r^{\alpha}(n)$$

d) $\forall x \in E$, $\exists A_2'(x) > 0$ tel que pour r assez grand dans I

$$\int_E K^2_{r(n)}(x,y)\ d\mu(y) \geqslant A_2'(x)\ r^{\alpha}(n) \quad .$$

<u>*Proposition 4.10*</u>.- Si f, $\{K_r(.,.),\ r \in I\}$ vérifient les hypothèses (a) à (d) et $\frac{r^{\alpha}(n)}{n} \xrightarrow[n\to\infty]{} 0$.

Alors la v.a. $\dfrac{f_n^1(x) - \bar{f}_n(x)}{\sqrt{\mathrm{Var}(f_n^1(x))}}$ converge en loi vers une v.a. normale centrée réduite.

<u>*Preuve*</u> : Posons $T_{r(n),j} = K_{r(n)}(x,X_j) - \bar{f}_n(x)$, $n \geqslant 1$, $j \in \overline{1,n}$ et $s_n = \sqrt{n}\ \sigma\ \{K_{r(n)}(x,X_1)\}$.

Nous avons $\left[\sigma(f_n^1(x))\right]^{-1} \{f_n^1(x) - \bar{f}_n(x)\} = \frac{1}{s_n} \sum_{j=1}^{n} T_{r(n),j}$ (4)

or $\forall n \geq 1$ les $T_{r(n),j}$ sont des v.a. i.i.d. de moyenne nulle et de variance $\sigma^2\{K_{r(n)}(x,X_1)\}$ et d'après le théorème central limite de LINDEBERG-FELLER une C.S. pour que (4) soit réalisée est :

(5) $$\forall \varepsilon > 0 \quad , \quad \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|T_{r(n),j}| \geq \varepsilon s_n} T_{r(n),j}^2 \, dP \xrightarrow[n\to\infty]{} 0 \quad .$$

Nous avons

$$\sigma^2(K_{r(n)}(x,X_1)) = \int_E K_{r(n)}^2(x,y)\ f(y)\ d\mu(y) - \bar{f}_n^2(x)$$
$$\geq m\ A_2'(x)\ r^\alpha(n) - \bar{f}_n^2(x)$$

grâce à l'hypothèse (a).

L'hypothèse (b) est le fait que $r^\alpha(n) \xrightarrow[n\to\infty]{} +\infty$ nous permettent d'affirmer que pour n assez grand

$$\sigma^2(K_{r(n)}(x,X_1)) > \frac{1}{2}\ m\ A_2'(x)\ r^\alpha(n) \quad .$$

Par ailleurs l'inégalité de Schwarz dans $L^2(f.d\mu)$ nous donne :

$$\bar{f}_n^2(x) \leq \int_E K_{r(n)}^2(x,y)\ f(y)\ d\mu(y) \ . \int_E f(y)\ d\mu(y)$$
$$\leq (A_1(x)\ r^\alpha(n))^2 \qquad \text{(hypothèse (c))}$$

ainsi $$T_{r(n),j}^2 \leq 4\ (A_1(x)\ r^\alpha(n))^2 \tag{7}$$

par conséquent en utilisant la définition de s_n , (6) et (7) nous obtenons :

$$\frac{\sup T^2_{r(n),j}}{s_n^2} \leqslant 4\ (A_1(x)\ r^\alpha(n))^2 \cdot \frac{2}{n\ m\ A_2'(x)\ r^\alpha(n)}$$

$$= \frac{8\ A_1^2(x)}{m\ A_2'(x)} \cdot \left(\frac{r^\alpha(n)}{n}\right)$$

donc $$\int_{|T_{r(n),j}| \geq \varepsilon s_n} T^2_{r(n),j}\ dP \leqslant \sup T^2_{r(n),j}\ \Pr\{|T_{r(n),j}| \geqslant \varepsilon\ s_n\}$$

$$\leqslant \sup T^2_{r(n),j}\ \frac{\sigma^2(K_{r(n)}(x,X_1)}{\varepsilon^2\ s_n^2}$$

$$\leqslant \frac{8\ A_1^2(x)}{m\ A_2'(x)}\ \frac{\sigma^2(K_{r(n)}(x,X_1))}{\varepsilon^2} \cdot \left(\frac{r^\alpha(n)}{n}\right)$$

par conséquent (5) est vérifiée puisque $\frac{r^\alpha(n)}{n} \xrightarrow[n\to\infty]{} 0$ ■

III.4.2. - Loi limite de la classe d'estimateurs de J. BLEUEZ conditionnés.

Afin d'obtenir une nouvelle classe d'estimateurs, qui localement améliorent l'estimation d'une certaine famille de densités, on se propose d'appliquer la proposition 4.2. à l'estimateur $f_n^1(x)$.

L'échantillon $(X_i)_{i=1,\dots,n}$ est donc, défini sur un espace probabilisé $(\Omega,\mathcal{A},\Pr)$ de loi Q appartenant à un ensemble $\mathcal{P}$, de probabilités sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, la suite d'ouverts $(U_i)_i$ est réduite à $U_i = U_n$, $\forall i \in \overline{1,n}$ et telle que $x \in U_n$. On se fixe également une probabilité $P \in \mathcal{P}$ telle que $P(U_n^c) > 0$. Une version de $f_n^1(x)$ conditionné par rapport à $\mathcal{B}_S$, est définie par :

$$\tilde{f}_n^1(x) = \frac{1}{n} \sum_{j=1}^{n} \{K_{r(n)}(x,X_j)\ 1_U(X_j) + m_{r(n)}(P)\ 1_{U^c}(X_j)\}$$

où $$m_{r(n)}(P) = \frac{1}{P(U^c)} \int_{U^c} K_{r(n)}(x,y)\, f_P(y)\, dy \quad .$$

On notera

$$f_n^*(x) = E_Q(\tilde{f}_n^1(x)) = \{\int_U K_{r(n)}(x,y)\, f_Q(y)\, dy + \frac{Q(U^c)}{P(U^c)} \int_{U^c} K_{r(n)}(x,y)\, f_P(y)\, dy\}.$$

Avec des hypothèses similaires à celles de la proposition 4.10. nous obtenons la convergence en loi de $\dfrac{\tilde{f}_n^1(x) - f_n^*(x)}{\sigma(\tilde{f}_n^1(x))}$ vers une v.a. normale centrée réduite

nous imposerons à f_Q , f_P , $\{K_r(.,.),\ r \in I\}$, $r(n)$, les hypothèses suivantes :

i) $\exists m_o > 0$ tel que $\forall y \in U$ $f_Q(y) > m_o$,

ii) $\sup_n |f_n^*(x)| < +\infty$,

iii) $\exists A_{o1}(x) > 0$ tel que $\sup_{y \in E} |K_{r(n)}(x,y)| \leq A_{o1}(x)\, r^\alpha(n)$ pour n assez grand ,

iv) $\exists A_{o2}(x) > 0$ tel que $\int_U K^2_{r(n)}(x,y)\, dy \geq A_{o2}(x)\, r^\alpha(n)$ pour n assez grand ,

v) $\dfrac{r^\alpha(n)}{n} \xrightarrow[n\to\infty]{} 0$.

Proposition 4.11.- Si f_Q , f_P , $\{K_r(.,.),\ r \in I\}$, $r(n)$, $n \geq 1$ vérifient les hypothèses (i) à (v).

Alors :

$$\frac{\tilde{f}_n^1(x) - f_n^*(x)}{\sigma(\tilde{f}_n^1(x))} \xrightarrow{\mathcal{L}} N(0,1) \quad .$$

Preuve : Posons

$$\tilde{V}_{r(n),j} = K_{r(n)}(x,X_j)\ 1_U(X_j) + m_{r(n)}(P)\ 1_{U^c}(X_j) - f_n^*(x)$$

pour $n \geq 1$ et $j \in \overline{1,n}$

$$\tilde{s}_n = \sqrt{n}\ \sigma\{K_{r(n)}(x,X_1)\ 1_U(X_1) + m_{r(n)}(P)\ 1_{U^c}(X_1)\}$$

d'où

$$\frac{\tilde{f}_n^1(x) - f_n^*(x)}{\sigma(\tilde{f}_n^1(x))} = \frac{1}{\tilde{s}_n} \sum_{j=1}^{n} \tilde{V}_{r(n),j}$$

la condition de LINDEBERG-FELLER s'écrit :

$$\forall \varepsilon > 0 \qquad \frac{1}{\tilde{s}_n^2} \sum_{j=1}^{n} \int_{|\tilde{V}_{r(n),j}| \geq \varepsilon \tilde{s}_n} \tilde{V}^2_{r(n),j}\ dQ \xrightarrow[n\to\infty]{} 0 \tag{8}$$

or $\quad \sigma^2\ \{K_{r(n)}(x,X_1)\ 1_U(X_1) + m_{r(n)}(P)\ 1_{U^c}(X_1)\} =$

$$\int_U K^2_{r(n)}(x,y)\ f_Q(y)\ dy + Q(U^c)\ \left[\frac{1}{P(U^c)} \int_{U^c} K_{r(n)}(x,y) f_P(y) dy\right]^2 - \{f_n^*(x)\}^2$$

et $\displaystyle\int_U K^2_{r(n)}(x,y)\ f_Q(y)\ dy + Q(U^c)\ \left[\frac{1}{P(U^c)} \int_{U^c} K_{r(n)}(x,y)\ f_P(y)\ dy\right]^2 \geq$

$$\int_U K^2_{r(n)}(x,y)\ f_Q(y)\ dy \geq m_o\ A_{o2}(x)\ r^\alpha(n)\ ;$$

par application de (i) et (iv), d'autre part la condition (ii) et le fait que $r^\alpha(n) \xrightarrow[n\to\infty]{} \infty$ impliquent que, pour n assez grand,

$$\sigma^2\{K_{r(n)}(x,X_1)\ 1_U(X_1) + m_{r(n)}(P)\ 1_{U^c}(X_1)\} \geq \frac{m_o}{2}\ A_{o2}(x)\ r^\alpha(n)\ ,$$

ensuite nous avons $f_n^*(x) \leq A_{o1}(x)\ r^\alpha(n)$ pour n assez grand et $\quad |f_n^*(x)\ [K_{r(n)}(x,X_1)\ 1_U(X_1) + m_{r(n)}(P)\ 1_{U^c}(X_1)]| \leq (A_{o1}(x)\ r^\alpha(n))^2$

(il faut distinguer les 2 cas $X_1 \in U$ et $X_1 \in U^c$) , en regroupant ces différentes majorations nous obtenons

$$\tilde{V}^2_{r(n),j} \leq 4(A_{o1}(x)\ r^\alpha(n))^2$$

et on achève la preuve de la même façon que dans la proposition 4.10. ■

En procédant de la même manière qu'à la proposition 1, on pourrait montrer que $\dfrac{\tilde{f}^1_n(x) - f^*_n(x)}{\sigma(\tilde{f}^1_n(x))} \xrightarrow[n\to\infty]{\mathcal{L}} N(0,1)$, à condition d'avoir :

(c1) $\quad f_Q$ est bornée

(c2) $\quad K_r(x,.)$ est Q, P-intégrable

(c3) $\quad \exists\ A_{12}(x) > 0$ tel que :

$$\lim_{n\to\infty} \frac{1}{r^\alpha(n)} \mathrm{Var}\{K_{r(n)}(x,X_1)\ 1_U(X_1) + m_{r(n)}(P)\ 1_{U^c}(X_1)\} \leq A_{12}(x)$$

(c4) $\quad \exists A_{13}(x) > 0$ tel que pour n assez grand

$$\int_E |K_{r(n)}(x,y)|^3\ dy \leq A_{13}(x)\ r^{2\alpha}(n)$$

(c5) $\quad \dfrac{r^\alpha(n)}{n} \xrightarrow[n\to\infty]{} 0 \quad ,$

bien entendu x est un réel fixé, puisque nous conditionnons afin de corriger localement $\tilde{f}^1_n$.

III.4.3. - Application aux estimateurs à noyaux.

1°) Estimateur à noyau de Parzen-Rosenblatt.

Comme cela a été indiqué précédemment, la classe d'estimateurs introduite par J. BLEUEZ contient les estimateurs obtenus par la méthode du noyau, elle correspond au choix d'une fonction K à une variable avec :

$$K_r(x,y) = r\, K(r(x-y)) \quad , \quad r \geqslant 0$$

où l'on suppose que E est un espace vectoriel sur $\mathbb{R}$, il suffit alors de poser $h_n = 1/r(n)$ pour retrouver l'estimateur de Parzen-Rosenblatt.

Nous allons montrer que les conditions imposées à la famille $\{K_r(.,.),\ r \in I\}$ dans les propositions 4.9 et 4.10 sont vérifiées dans ce cas usuel, en effet : en prenant $\alpha = 1$ les conditions que doit vérifier $r(n)$ s'écrivent :

$$h_n \xrightarrow[n\to\infty]{} 0 \qquad \text{et} \qquad n\, h_n \xrightarrow[n\to\infty]{} \infty$$

ensuite pour que les hypothèses (I) à (IV) (p. III.3.1.) soient réalisées il suffit de supposer que f est continue bornée et que $|K|^i$ est λ-intégrable, avec $i \in \overline{1,3}$. Nous retrouvons ainsi le résultat de E. PARZEN (1962). Quant à la proposition 4.10, elle est applicable dès que f ne s'annule pas et est continue bornée et K est bornée, λ-intégrable.

2°) Estimateur à noyau conditionné.

On se place dans le cas particulier où $E = \mathbb{R}$ et $\mu = \lambda$ la mesure de Lebesgue sur $\mathbb{R}$, l'ouvert U est choisi égal à $]x-H, x+H[$ où $H = H(n)$ est une suite de réels strictement positifs, on retrouve ainsi l'estimateur $\tilde{f}_n(x)$, défini page 70 ;

$$\tilde{f}_n(x) = \frac{1}{n} \sum_{j=1}^{n} \left(\frac{1}{h} K\left(\frac{x-X_j}{h}\right) 1_U(X_j) + m_n(P)\ 1_{U^c}(X_j)\right)$$

où $m_n(P) = \dfrac{1}{h\ P(U^c)} \displaystyle\int_{U^c} K\left(\frac{x-y}{h}\right) f_P(y)\, dy$,

afin que $\tilde{f}_n(x)$ vérifie les hypothèses (a), (b), (c), (d) de la proposition 4.10 il suffit que

(*) f_Q et f_P sont bornées continues au voisinage de x

f_Q est non nulle sur un voisinage de x

(**) $|K|$ est bornée par M_2 , λ-intégrable

(***) $h_n \xrightarrow[n\to\infty]{} 0$, $n\ h_n \xrightarrow[n\to\infty]{} \infty$ et $H_n/h_n \xrightarrow[n\to\infty]{} \gamma (\in\ ]0,+\infty[)$.

En effet :

(a) est vérifiée d'après (*)

(b) l'est également car

$$f_n^*(x) = \left\{ E_Q(\tilde{f}_n(x)) = \frac{1}{h} \int_U K\left(\frac{x-y}{h}\right) f_Q(y)\, dy + \frac{Q(U^c)}{P(U^c)} \frac{1}{h} \int_{U^c} K\left(\frac{x-y}{h}\right) f_P(y)\, dy \right\}$$

existe et est borné d'après (*) et (**).

D'autre part $|K|$ étant bornée par M_2 , $\sup\limits_{x\in\mathbb{R}} \left|\frac{1}{h} K\left(\frac{x-y}{h}\right)\right| \leq \frac{1}{h} - M_2$.

<u>i.e.</u> (c) est vérifiée, enfin nous avons

$$\int_U \frac{1}{h^2} K^2\left(\frac{x-y}{h}\right) dy = \frac{1}{h} \int_{]-H/h,H/h[} K^2(y)\, dy$$. L'hypothèse que $\frac{H}{h}$

converge vers un réel γ non nul, positif, comibnée avec $|K|$ λ-intégrable et bornée nous permet d'affirmer qu'il existe une constante $\beta > 0$ telle que pour n assez grand on ait $\frac{1}{h} \int_{]-H/h,H/h[} K^2(y)\, dy \geq \beta/h$ i.e. l'hypothèse (d) est réalisée (K est supposé non nulle au voisinage de 0). Enfin si on suppose que :

(a1) f_Q et f_Q sont bornées et continues au voisinage de x

(a2) K^i est $\{\lambda\} \cup \mathcal{P}$-intégrable $i = 1,2,3$

(a3) $h_n \xrightarrow[n\to\infty]{} 0$, $n\, h_n \xrightarrow[n\to\infty]{} +\infty$, $H_n/h_n \xrightarrow[n\to\infty]{} \gamma \in \left]0,+\infty\right]$

alors les conditions (c1) à (c5) seront satisfaites et par conséquent la convergence en loi de $\tilde{f}_n(x)$ est assurée.

III.5. - Loi du Log itéré.-

P. HALL (1981) a déterminé les conditions sous lesquelles la classe d'estimateurs de J. BLEUEZ vérifie la loi du log itéré, nous allons reprendre ses résultats et les appliquer à l'estimateur à noyau conditionné qui correspond au cas particulier $U = \left]x-h\, ,\, x+h\right[$ où h est la fenêtre. Auparavant nous fixons quelques notations et rappelons le résultat de P. HALL.

On posera $$S_n(r) = \sum_{i=1}^{n} (K_r(x,X_i) - E(K_r(x,X_i)))$$

$$\sigma_{rs} = \mathrm{Cov}(K_r(x,X_1),\, K_s(x;X_1)) \quad , \quad \sigma_r^2 = \sigma_{rr}$$

où $\{K_r(.,.),\ r \in I\}$ est la famille de fonctions définie précédemment.

P. HALL a établi le résultat suivant :

Proposition 4.12.-

i) si $\{K_r(.,.),\ r \in I\}$ est une famille de fonctions à variations bornées

ii) si $$\frac{(\mathrm{Log}\ n)^4 \left(\int |dK_{r(n)}(.,u)|\right)^2}{n\, \sigma_{r(n)}^2\ \mathrm{Log}\ \mathrm{Log}\ n} \xrightarrow[n\to\infty]{} 0$$

iii) si $\lim_{\varepsilon\to 0} \limsup_{n\to\infty} \sup_{m} \left| \frac{\sigma_{r(m)r(n)}}{\sigma^2_{r(n)}} - 1 \right| = 0$

où le sup est pris pour les m tels que : $|m-n| \leq \varepsilon n$. Alors :

$$\limsup_{n\to\infty} \pm \frac{S_n(r(n))}{\sqrt{2n\,\sigma^2_{r(n)}\ \mathrm{Log\,Log}\ n}} = 1 \quad \text{p.s.}$$

la condition (iii) peut se formuler comme suit :

iii') pour toutes suites $m,n \longrightarrow \infty$ avec $\frac{m}{n} \longrightarrow 1$ on a

$$\frac{\sigma_{r(m)r(n)}}{\sigma^2_{r(n)}} \xrightarrow[n\to\infty]{} 1 \ .$$

Ce théorème est applicable à l'estimateur à noyau conditionné $\tilde{f}_{n2}(x)$, obtenu à partir de $\tilde{f}_n(x)$ en prenant $U =]x-h\ ,\ x+h[$, car il fait partie de la classe d'estimateurs de J. BLEUEZ, en effet l'estimateur conditionné correspondant au cas $U =]x-h, x+h[$ s'écrit :

$$\tilde{f}_{n2}(x) = \frac{1}{n} \sum_{i=1}^{n} \tilde{K}_{h_n}(x,X_i) \qquad \text{où}$$

$$\tilde{K}_{h_n}(x,X_i) = \frac{1}{h_n} K\left(\frac{x-X_i}{h_n}\right) 1_{]-1,1[}\left(\frac{x-X_i}{h_n}\right) + 1_{]-1,1[^c}\left(\frac{x-X_i}{h_n}\right) \frac{1}{P(U^c)} \int_{]-1,1[^c} K(y) f_P(x-h_y)\,dy$$

avec K une fonction réelle $\{\lambda\} \cup P$-intégrable, dans la suite nous écrirons indifféremment $\tilde{K}_h(x,X)$ ou $\tilde{K}_h(X)$.

Proposition 4.13.- Sous les hypothèses suivantes :

(i1) f_P continue, f_Q continue bornée sur un voisinage de x et non non nulle en x.

(i2) K est un noyau continu borné sur $\mathbb{R}$ et à variation bornée sur $]-1,1[$ et tel que $\int |K| < \infty$.

(i3) $\forall m$, $n \longrightarrow \infty$ avec $\frac{m}{n} \longrightarrow 1$ on a $\frac{h(m)}{h(n)} \longrightarrow 1$

(i4) $\frac{(\mathrm{Log}\ n)^4}{n\ h\ \mathrm{Log}\ \mathrm{Log}\ n} \underset{n\to\infty}{\longrightarrow} 0$

Nous avons :

(LI)
$$\limsup_{n\to\infty} \pm \frac{\tilde{f}_{n2}(x) - E_Q\, \tilde{f}_{n2}(x)}{\sqrt{2\,\frac{\mathrm{Log}\ \mathrm{Log}\ n}{n\ h}}} \underset{\text{p.s.}}{=} \sqrt{f_Q(x) \int_{-1}^{1} K^2\, d\lambda}$$

Preuve : Posons $\bar{K}_h(X_i) = h\, \tilde{K}_h(X_i)$

et
$$S_n(r) = S_n(h) = \sum_{i=1}^{n} (\bar{K}_h(x,X_i) - E_Q(\bar{K}_h(x,X_i))$$

$$\sigma_{r(m)r(n)} = \sigma_{h(m)h(n)} \quad .$$

Dans ce cas

$$\frac{S_n(r(n))}{(2n\ \sigma^2_{r(n)}\ \mathrm{Log}\ \mathrm{Log}\ n)^{1/2}} = \frac{\sum_{i=1}^{n} (\bar{K}_h(x,X_i) - E_Q(\bar{K}_h(x,X_i)))}{(2n\ \mathrm{Var}(\bar{K}_h(x,X_1))\ \mathrm{Log}\ \mathrm{Log}\ n)^{1/2}}$$

$$= \frac{\tilde{f}_{n2}(x) - E_Q\, \tilde{f}_{n2}(x)}{\left[2\ \frac{\mathrm{Log}\ \mathrm{Log}\ n}{n\ h^2}\ \mathrm{Var}(\bar{K}_h(x,X_1))\right]^{1/2}}$$

or
$$\mathrm{Var}(\bar{K}_h(x,X_1)) = h \int_{-1}^{1} K^2(y)\ f_Q(x-hy)\ dy - h^2 \left[\int_{-1}^{1} K(y)\ f_Q(x-hy)\ dy\right]^2$$

$$+ h^2\, \frac{Q(U^c)}{P(U^c)} \int_{]-1,1[^c} K(y) f_P(x-hy)dy \left\{\frac{Q(U)}{P(U^c)} \int_{]-1,1[^c} K(y) f_P(x-hy)dy - \right.$$

$$\left. - 2\int_{-1}^{1} K(y) f_Q(x-hy)dy\right\}$$

et d'après A. BERLINET (1984-a), proposition 7 p. VI-16, si $f_Q(x) \neq 0$ et $\int_{-1}^{1} K^2(y)\, dy$ existe alors

$$\mathrm{Var}(\bar{K}_h(x,X_1)) \underset{h\to 0}{\sim} h\, f_Q(x) \int_{-1}^{1} K^2(y)\, dy$$

$$\frac{S_n(r(n))}{(2n\, \sigma^2_{r(n)} \mathrm{Log\,Log}\, n)^{1/2}} \underset{n\to\infty}{\sim} \frac{\tilde{f}_{n2}(x) - E_Q\, \tilde{f}_{n2}(x)}{\left[2\, \frac{\mathrm{Log\,Log}\, n}{nh}\, f_Q(x) \int_{-1}^{1} K^2(y)\, dy\right]^{1/2}} .$$

Donc pour avoir (LI) il suffit de montrer que $\bar{K}_h$ et h vérifient les conditions du théorème 1 page 49 de P. HALL (1981).

En effet :

$$\bar{K}_h(u) = K(\tfrac{x-u}{h})\, 1_{]-1,1[}(\tfrac{x-u}{h}) + \frac{h}{P(]x-h,x+h[^c)} \int_{]-1,1[^c} K(y) f_P(x-hy)\, dy\, 1_{]-1,1[^c}(\tfrac{x-u}{h})$$

est à variation bornée sur $\mathbb{R}$ puisque

$$\begin{aligned} d\, \bar{K}_h(u) &= dK(\tfrac{x-u}{h}) && \text{si} \quad u \in]x-h,x+h[\\ &= 0 && \text{sinon} \end{aligned}$$

ce qui entraîne que

$$\int |d\, \bar{K}_h(u)| = \int_U |dK(\tfrac{x-u}{h})| = \int_{-1}^{1} |dK(u)| < \infty$$

car par hypothèse K est à variation bornée sur $]-1,1[$ ainsi la condition (i) du théorème de P. HALL est satisfaite.

Ensuite

$$\frac{(\mathrm{Log}\; n)^4 \left[\int |d\, \bar{K}_h(u)|\right]^2}{n\, \sigma_h^2\, \mathrm{Log\,Log}\, n} = \frac{(\mathrm{Log}\; n)^4 \left[\int_{-1}^{1} |dK(u)|\right]^2}{n\, \mathrm{Var}(\bar{K}_h(x,X_1))\, \mathrm{Log\,Log}\, n}$$

$$\underset{n\to\infty}{\sim} \frac{(\mathrm{Log}\; n)^4 \left[\int_{-1}^{1} |dK(u)|\right]^2}{n\, h\, \mathrm{Log\,Log}\, n\, f_Q(x) \int_{-1}^{1} K^2(u)\, du}$$

quantité qui converge vers 0 lorsque n tend vers $+\infty$, d'après (i4).
La condition (ii) de P. HALL est donc vérifiée, enfin pour avoir la condition (iii) ou (iii') du théorème de P. HALL, il suffit de montrer que : pour toutes suites $h, k \to 0$ telles que $h/k \to 1$ on a

$$\text{(LI2)} \qquad \frac{\mathrm{Cov}(\bar{K}_h(X_1),\ \bar{K}_k(X_1))}{h\ f_Q(x) \int_{-1}^{1} K^2(y)\ dy} \longrightarrow 1$$

or

$$\mathrm{Cov}(\bar{K}_h(X_1),\ \bar{K}_k(X_1)) = E_Q(\bar{K}_h(X_1)\ \bar{K}_k(X_1)) - E_Q(\bar{K}_h(X_1))\ E_Q(K_k(X_1))$$

$$= E_Q\left\{K(\frac{x-X_1}{h})\ K(\frac{x-X_1}{k})\ 1_{]-1,1[}(\frac{x-X_1}{h})\ 1_{]-1,1[}(\frac{x-X_1}{k})\right\}$$

$$+ k\ .\ E_Q\left(K(\frac{x-X_1}{h})\ 1_{]-1,1[}(\frac{x-X_1}{h})\ 1_{]-1,1[^c}(\frac{x-X_1}{k})\right)$$

$$.\ \frac{1}{P(]x-k,x+k[^c)} \int_{]-1,1[^c} K(y) f_P(x-ky)\,dy$$

$$+ h\ E_Q\left(K(\frac{x-X_1}{k})\ 1_{]-1,1[}(\frac{x-X_1}{k})\ 1_{]-1,1[^c}(\frac{x-X_1}{h})\right)$$

$$\frac{1}{P(]x-h,x+h[^c)} \int_{]-1,1[^c} K(y) f_P(x-hy)\,dy$$

$$+ hk\ \frac{Q(]x-h,x+h[^c \cap]x-k,x+k[^c)}{P(]x-h,x+h[^c).P(]x-k,x+k[^c)}\ .\ \int_{]-1,1[^c} K(y) f_P(x-hy)\,dy \int_{]-1,1[^c} K(y) f_P(x-ky)\,dy$$

$$- hk\left\{\int_{-1}^{1} K(y) f_Q(x-hy)\,dy + \frac{Q(]x-h,x+h[^c)}{P(]x-h,x+h[^c)} \int_{]-1,1[^c} K(y) f_P(x-hy)\,dy\right\}$$

$$.\left\{\int_{-1}^{1} K(y) f_Q(x-ky)\,dy + \frac{Q(]x-k,x+k[^c)}{P(]x-k,x+k[^c)} \int_{]-1,1[^c} K(y) f_P(x-ky)\,dy\right\}$$

h et k jouant des rôles symétriques on peut supposer que $h \leq k$

on aura donc $1_{]-1,1[}\left(\frac{x-X_1}{h}\right) \cdot 1_{]-1,1[^c}\left(\frac{x-X_1}{k}\right) = 0$ par conséquent

$$\mathrm{Cov}(\bar{K}_h(X_1), \bar{K}_k(X_1)) = h \int_{-1}^{1} K(y)\, K(\tfrac{h}{k}\, y)\, f_Q(x-hy)\, dy$$

$$+ \frac{kh}{P(]x-h,x+h[^c)} \int_{]-1,1[^c} K(y) f_P(x-hy) dy \int_{]-1,-\frac{h}{k}] \cup [\frac{h}{k},1[} K(y) f_Q(x-ky) dy$$

$$+ \frac{hk\; Q(]x-k,x+k[^c)}{P(]x-h,x+h[^c)\; P(]x-k,x+k[^c)} \cdot \int_{]-1,1[^c} K(y) f_P(x-hy) dy \int_{]-1,1[^c} K(y) f_P(x-ky) dy$$

$$- hk \left\{ \int_{-1}^{1} K(y) f_Q(x-hy) dy + \frac{Q(]x-h,x+h[^c)}{P(]x-h,x+h[^c)} \int_{]-1,1[^c} K(y) f_P(x-hy) dy \right\}$$

$$\cdot \left\{ \int_{-1}^{1} K(y) f_Q(x-ky) dy + \frac{Q(]x-k,x+k[^c)}{P(]x-k,x+k[^c)} \int_{]-1,1[^c} K(y) f_P(x-ky) dy \right\}$$

nous savons que les conditions imposées à f_P , f_Q et à K entraînent que

* $Q(]x-h,x+h[^c)$, $Q(]x-k,x+k[^c)$, $P(]x-h,x+h[^c)$ et $P(]x-k,x+k[^c)$ convergent vers 1 lorsque h et $k \longrightarrow 0$.

* $\int_{-1}^{1} K(y) f_Q(x-hy) dy$ et $\int_{-1}^{1} K(y) f_Q(x-ky) dy \xrightarrow[h,k\to 0]{} f_Q(x) \int_{-1}^{1} K(y) dy$

* $\int_{]-1,1[^c} K(y) f_Q(x-hy) dy$ et $\int_{]-1,1[^c} K(y) f_Q(x-ky) dy \xrightarrow[h,k\to 0]{} f_Q(x) \int_{]-1,1[^c} K(y) dy$.

D'autre part la continuité de K et les conditions $h \to 0$, $k \to 0$, $\frac{h}{k} \to 1$ entraînent que :

$$K(y)\, K(\tfrac{h}{k}\, y)\, f_Q(x-hy) \longrightarrow K^2(y)\, f_Q(x) \quad \text{pour tout } y \in]-1,1[$$

ainsi le théorème de la convergence dominée nous permet d'avoir :

$$\int_{-1}^{1} K(y)\, K(\frac{h}{k}\, y)\, f_Q(x-hy)\, dy \xrightarrow[\substack{h,k\to 0 \\ h/k\to 1}]{} f_Q(x) \int_{-1}^{1} K^2(y)\, dy$$

en regroupant tout cela nous obtenons (LI2). ∎

III.6. - Estimateur conditionné à noyau.-

III.6.1. - Introduction et notations

Nous savons que la technique du conditionnement suppose la connaissance d'une bonne approximation P sur U^c , de la loi inconnue Q. En général, cette approximation provient d'une préestimation pour laquelle on pourra utiliser des méthodes énumérées dans la première partie. On se propose, dans ce qui suit, d'étudier le comportement asymptotique de l'estimateur conditionné $\tilde{f}_n(x)$, dans lequel on a remplacé $f_P(y)$ par son estimateur à noyau $f_n(y) = \frac{1}{nh_n} \sum_{k=1}^{n} K(\frac{y-X_k}{h_n})$, l'estimateur obtenu sera appelé estimateur conditionné à noyau et il est défini par :

$$\tilde{f}_{n1}(x) = g_{U_n}(x) + g_{U_n^c}(x)$$

où

$$g_{U_n}(x) = \frac{1}{nh_n} \sum_{j=1}^{n} K(\frac{x-X_j}{h_n})\, 1_{U_n}(X_j)$$

et

$$g_{U_n^c}(x) = \frac{\psi_n(x)}{\emptyset_n(x)} \qquad \text{avec}$$

$$\psi_n(x) = \frac{1}{n^2h_n^2} \sum_{i,j=1}^{n} 1_{U_n^c}(X_j) \int_{U_n^c} K(\frac{y-X_i}{h_n})\, K(\frac{x-y}{h_n})\, dy$$

$$\emptyset_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} \int_{U_n^c} K(\frac{y-X_i}{h_n})\, dy$$

$\phi_n(x)$ est une approximation de $P(U_n^c)$, pour n assez grand, elle est presque sûrement non nulle.

Pour tout $n \in \mathbb{N}^*$ nous posons $\bar{\psi}_n(x) = E(\psi_n(x))$ et $\bar{\phi}_n(x) = E(\phi_n(x))$, et désignons par $B_n(x)$ et, sous réserve d'existence, par $V_n(x)$ et $R_n(x)$ le biais la variance et l'erreur quadratique de l'estimateur $\tilde{f}_{n1}(x)$.

III.6.2. - Etude du biais.

Proposition 4.14.-

- Si f_Q est continue bornée, de dérivées d'ordre 2 et 3 bornées.

- K est une densité de probabilité sur $\mathbb{R}$, bornée, paire et telle que $\int |z^i K(z)| \, dz < \infty$, $i \in \overline{1,3}$.

- $h_n \xrightarrow[n\to\infty]{} 0$, $n h_n^2 \xrightarrow[n\to\infty]{} + \infty$.

Alors $B_n(x) = (E \tilde{f}_{n1}(x) - f_Q(x)) \xrightarrow[n\to\infty]{} 0$, plus exactement

$$B_n(x) = \frac{h_n^2}{2} \Delta(x) + O(h_n^3 + \frac{1}{h_n \sqrt{n}})$$

avec
$$\Delta(x) = f''_Q(x) \left[1 + \int_{]-1,1[^c} K(t) \, dt\right] \int_{\mathbb{R}} s^2 K(s) \, ds \, .$$

<u>*Preuve :*</u> Nous avons $E(\tilde{f}_{n1}(x)) = E(g_{U_n}(x)) + E(g_{U_n^c}(x))$.

D'après la définition de g_{U_n} on a :

$$E(g_{U_n}(x)) = \frac{1}{h_n} \int_{U_n} K\left(\frac{x-y}{h_n}\right) f_Q(y)\, dy = \int_{-1}^{1} K(y)\, f_Q(x-h_n\, y)\, dy \quad .$$

Une application de la formule de Taylor à $f_Q(x-h_n\, y)$ nous donne :

$$E(g_{U_n}(x)) = f_Q(x) \int_{-1}^{1} K(y)\, dy + \frac{h_n^2}{2} f''_Q(x) \int_{-1}^{1} y^2\, K(y)\, dy + h_n^2\, \varepsilon_1(h_n)$$

avec $\varepsilon_1(h_n) \xrightarrow[n\to\infty]{} 0$

$$\left(\varepsilon_1(h_n) = h_n \int_{-1}^{1} z^3\, K(z)\, f_Q^{(3)}(x - \theta\, h_n\, z)\, dz\right)$$

$$= O(h_n)$$

d'autre part la forme de $g_{U_n^c}(x)$ ne nous permet pas de connaître exactement l'expression de $E\left(g_{U_n^c}(x)\right)$. On se contentera des premiers termes de son développement asymptotique, pour cela l'identité :

$$\forall z \in \mathbb{R}^* \qquad \frac{1}{z} = 1 - \frac{z-1}{z}$$

appliquée à $z = \dfrac{\phi_n(x)}{\bar{\phi}_n(x)}$, pour n assez grand , nous donne :

$$g_{U_n^c}(x) = \frac{\psi_n(x)}{\bar{\phi}_n(x)} \cdot \frac{1}{\dfrac{\phi_n(x)}{\bar{\phi}_n(x)}} = \frac{\psi_n(x)}{\bar{\phi}_n(x)} \cdot \left\{1 - \frac{\phi_n(x) - \bar{\phi}_n(x)}{\phi_n(x)}\right\}$$

$$= \frac{\psi_n(x)}{\bar{\phi}_n(x)} - \frac{\psi_n(x)}{\phi_n(x)} \cdot \frac{\phi_n(x) - \bar{\phi}_n(x)}{\bar{\phi}_n(x)}$$

d'où :

$$E(g_{U_n^C}(x)) = E\left[\frac{\psi_n(x)}{\bar{\varphi}_n(x)}\right] - E\left[\frac{\psi_n(x)}{\varphi_n(x)} \cdot \frac{\varphi_n(x) - \bar{\varphi}_n(x)}{\bar{\varphi}_n(x)}\right] .$$

Pour déterminer le développement en h_n de $E(g_{U_n^c}(x))$. Nous allons procéder en trois étapes : nous commencerons par étudier $\bar{\varphi}_n(x)$ ensuite $\bar{\psi}_n(x)$ et enfin le terme $E\left[\frac{\psi_n(x)}{\varphi_n(x)} \cdot \frac{\varphi_n(x) - \bar{\varphi}_n(x)}{\bar{\varphi}_n(x)}\right]$.

a) $\bar{\varphi}_n(x) = 1 - 2\, h_n\, f_Q(x) + O(h_n^3)$

en effet $\bar{\varphi}_n(x) = \frac{1}{h_n} \int_{\mathbb{R}} \int_{U_n^c} K(\frac{y-z}{h_n})\, f_Q(z)\, dy\, dz = \int_{U_n^c} \int_{\mathbb{R}} K(t)\, f_Q(y-h_n\, t)\, dt\, dy$

$$= \int_{\mathbb{R}} K(t) \int_{]x-h_n(1+t),x+h_n(1-t)[^c} f_Q(z)\, dz\, dt$$

$$= \int_{\mathbb{R}} K(t)\, Q(]x-h_n(1+t),\ x+h_n(1-t)[^c)\, dt$$

quantité qui existe, puisque $K(t)\, Q(]x-h_n(1+t),x+h_n(1-t)[^c) \leq K(t)$, $\forall t \in \mathbb{R}$.
D'après la formule de Taylor, à l'ordre 2, appliquée à $F_Q(x+h_n(1-t))$ et à $F_Q(x-h_n(1+t))$ nous avons :

$$F_Q(x+h_n(1-t)) = F_Q(x) + h_n(1-t)\, f_Q(x) + \frac{h_n^2}{2}\,(1-t)^2\, f'_Q(x) + h_n^2(1-t)\, \varepsilon_{1x}(h_n(1-t))$$

où

$$\varepsilon_{1x}(h_n(1-t)) = \frac{1}{6}\, h_n(1-t)\, f''_Q(x + \theta_1\, h_n(1-t)) \quad \text{et} \quad \theta_1 \in]0,1[$$

et d'autre part

$$F_Q(x-h_n(1+t)) = F_Q(x) - h_n(1+t)\, f_Q(x) + \frac{h_n^2}{2}\,(1+t)^2\, f'_Q(x) + h_n^2\,(1+t)^2\, \varepsilon_{2x}(h_n(1+t))$$

où $\varepsilon_{2x}(h_n(1+t)) = -\frac{1}{6} h_n(1+t) f''_Q(x - \theta_2 h_n(1+t))$ et $\theta_2 \in]0,1[$.

On en déduit que

$$Q(]x-h_n(1+t), x + h_n(1-t)[^c) = 1 - 2 h_n f_Q(x) + 2t h_n^2 f'_Q(x) + h_n^2 \varepsilon_n(x,t)$$

avec

$$\varepsilon_n(x,t) = (1+t)^2 \varepsilon_{2x}(h_n(1+t)) - (1-t)^2 \varepsilon_{1x}(h_n(1-t))$$

ensuite la parité de K et les conditions imposées à f''_Q et à K permettent d'écrire

$$\bar{\varphi}_n(x) = 1 - 2 h_n f_Q(x) + \bigcirc(h_n^3) .$$

En effet $\displaystyle\int_{\mathbb{R}} K(t) \varepsilon_n(x,t) dt = \bigcirc(h_n)$ car

$$|\varepsilon_n(x,t)| \leq \frac{h_n}{6} (|1+t|^3 + |1-t|^3) \sup_y |f''_Q(y)| .$$

b) $$\frac{\bar{\psi}_n(x)}{\bar{\varphi}_n(x)} = f_Q(x) \int_{]-1,1[^c} K(s) ds + \frac{h_n^2}{2} \alpha f''_Q(x) + \bigcirc(\frac{1}{n} + h_n^3)$$

où $$\alpha = \int_{]-1,1[^c} s^2 K(s) ds + \int_{\mathbb{R}} s^2 K(s) ds \int_{]-1,1[^c} K(t) dt$$

les v.a. X_i étant indépendantes nous avons

$$\bar{\psi}_n(x) = \frac{1}{n h_n^2} \int_{U_n^c} \int_{U_n^c} K(\frac{x-y}{h_n}) K(\frac{y-z}{h_n}) f_Q(z) dy dz +$$

$$+ \frac{n-1}{n} \frac{1}{h_n^2} Q(U_n^c) \int_{\mathbb{R}} \int_{U_n^c} K(\frac{x-y}{h_n}) K(\frac{y-z}{h_n}) f_Q(z) dy dz$$

en posant $s = \frac{x-y}{h_n}$ et $t = \frac{x-z}{h_n}$ on obtient :

$$\bar{\psi}_n(x) = Q(U_n^c) \int_{\mathbb{R}} \int_{]-1,1[^c} K(s)\, K(t-s)\, f_Q(x-h_n t)\, ds\, dt$$

$$+ \frac{1}{n} \Big[\int_{]-1,1[^c} \int_{]-1,1[^c} K(s)\, K(t-s)\, f_Q(x-h_n t)\, ds\, dt$$

$$- Q(U_n^c) \int_{\mathbb{R}} \int_{]-1,1[^c} K(s)\, K(t-s)\, f_Q(s-h_n t)\, ds\, dt \Big]$$

Notons M_1 un majorant de f_Q , nous avons :

$$\Big| \int_{]-1,1[^c} \int_{]-1,1[^c} K(s)\, K(t-s)\, f_Q(x-h_n t)\, ds\, dt$$

$$- Q(U_n^c) \int_{\mathbb{R}} \int_{]-1,1[^c} K(s)\, K(t-s)\, f_Q(x-h_n t)\, ds\, dt \Big|$$

$\leq 2\, M_1$, ce qui entraîne que le second terme, dans l'expression de $\bar{\psi}_n(x)$ est un $O(1/n)$. Pour le premier terme, la formule de Taylor appliquée à $f_Q(x-h_n t)$ nous donne :

$$\int_{\mathbb{R}} \int_{]-1,1[^c} K(s)\, K(t-s)\, f_Q(x-h_n t)\, ds\, dt = f_Q(x) \int_{\mathbb{R}} \int_{]-1,1[^c} K(s)\, K(t-s)\, ds\, dt$$

$$- h_n\, f'_Q(x) \int_{\mathbb{R}} \int_{]-1,1[^c} t\, K(s)\, K(t-s)\, ds\, dt$$

$$+ \frac{h_n^2}{2}\, f''_Q(x) \int_{\mathbb{R}} \int_{]-1,1[^c} t^2\, K(s)\, K(t-s)\, ds\, dt$$

$$- h_n^2\, \varepsilon_x(h_n)$$

avec $\varepsilon_x(h_n) = \frac{h_n}{6} \int_{\mathbb{R}} \int_{]-1,1[^c} t^3\, K(s)\, K(t-s)\, f_Q^{(3)}(x-h_n\theta t)\, ds\, dt$, $\theta \in]0,1[$

or par hypothèse $f_Q^{(3)}$ est bornée et $\int_{\mathbb{R}} |t^i K(t)|\, dt < \infty$, $i \in \overline{0,3}$ donc

$$\varepsilon_x(h_n) = O(h_n) \quad ,$$

d'autre part la condition : K paire, entraîne que
$\int_{\mathbb{R}} \int_{]-1,1[^c} t\, K(s)\, K(t-s)\, ds\, dt = 0$ car la fonction
$G(t) = t \int_{]-1,1[^c} K(s)\, K(t-s)\, ds$ existe et est impaire, elle existe car K est bornée et λ-intégrable. Elle est impaire car

$$G(-t) = -t \int_{]-1,1[^c} K(s)\, K(-t-s)\, ds = -t \int_{]-1,1[^c} K(-z)\, K(z-t)\, dz \quad \text{(on a posé } z = -s)$$

$$= -t \int_{]-1,1[^c} K(z)\, K(t-z)\, dz \quad \text{car K est paire}$$

<u>i.e.</u> $G(-t) = -G(t)$.

Par ailleurs $\alpha = \int_{\mathbb{R}} \int_{]-1,1[^c} t^2\, K(s)\, K(t-s)\, ds\, dt =$

$$= \int_{\mathbb{R}} \int_{]-1,1[^c} (s-z)^2\, K(s)\, K(z)\, ds\, dz$$

$$= \int_{]-1,1[^c} s^2\, K(s)\, ds + \int_{]-1,1[^c} K(s)\, ds \int_{\mathbb{R}} z^2\, K(z)\, dz \ ,$$

car K est paire et $\int K\, d\lambda = 1$, de même
$\int_{\mathbb{R}} \int_{]-1,1[^c} K(s)\, K(t-s)\, ds\, dt = \int_{]-1,1[^c} K(s)\, ds$ enfin les conditions imposées à f_Q nous donnent

$$Q(U^c) = 1 - F_Q(x+h_n) + F_Q(x-h_n) = 1 - 2\, h_n\, f_Q(x) + O(h_n^3) \ .$$

En regroupant tous ces résultats on obtient :

$$\bar{\psi}_n(x) = (1 - 2\, h_n\, f_Q(x) + \bigcirc(h_n^3))\, (f_Q(x) \int_{]-1,1[^c} K(s)\, ds + \alpha \, . \, \frac{h_n^2}{2}\, f''_Q(x))$$

$$+ \bigcirc(h_n^3 + 1/n) \, .$$

Ensuite en appliquant le développement $\frac{1}{1-\varepsilon} = 1 + \varepsilon + \varepsilon^2 + o(\varepsilon^2)$ à $\frac{1}{\bar{\emptyset}_n(x)}$ on aboutit à b)

c) $$E\left\{\frac{\psi_n(x)}{\emptyset_n(x)} \, . \, \frac{\emptyset_n(x) - \bar{\emptyset}_n(x)}{\bar{\emptyset}_n(x)}\right\} = \bigcirc\left[\left(\frac{1}{n\, h_n^2}\right)^{1/2}\right]$$

puisque K est une densité de probabilité bornée, notons M_2 sa borne supérieure. Nous avons pour tout échantillon $(X_i)_{i=1,\dots,n}$, tout réel x et tout n

$$\frac{\psi_n(x)}{\emptyset_n(x)} = \frac{1}{n\, h_n} \; \frac{\sum_{i=1}^{n} 1_{U_n^c}(X_i) \sum_{j=1}^{n} \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-X_j}{h_n})\, dy}{\sum_{j=1}^{n} \int_{U_n^c} K(\frac{y-X_j}{h_n})\, dy}$$

$$\leq \frac{M_2}{h_n}$$, nous avons majoré $1_{U_n^c}(X_i)$ par 1 et

$\int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-X_j}{h_n})\, dy$ par $M_2 \int_{U_n^c} K(\frac{y-X_j}{h_n})\, dy$, $\forall j \in \overline{1,n}$

ainsi $$\left|E(\frac{\psi_n(x)}{\emptyset_n(x)} - \frac{\emptyset_n(x) - \bar{\emptyset}_n(x)}{\bar{\emptyset}_n(x)})\right| \leq \frac{M_2}{h_n} \, . \, \frac{1}{\bar{\emptyset}_n(x)}\; E|\emptyset_n(x) - \bar{\emptyset}_n(x)|$$

qui, d'après l'inégalité de Schwarz, est inférieur à

$$\frac{M_2}{h_n} \, . \, \frac{1}{\bar{\emptyset}_n(x)}\; (\mathrm{Var}(\emptyset_n(x)))^{1/2}$$

ensuite par définition de $\emptyset_n(x)$, nous avons

$$V(\vartheta_n(x)) \leqslant \frac{1}{n\, h_n^2} \int_{\mathbb{R}} \left(\int_{U_n^c} K\left(\frac{y-z}{h_n}\right) dy\right)^2 f_Q(z)\, dz \leqslant \frac{1}{n} \;, \quad \text{car } \forall z \in \mathbb{R}$$

on a $\frac{1}{h_n} \int_{U_n^c} K\left(\frac{y-z}{h_n}\right) dy \leqslant 1$, enfin le fait que $\lim\limits_{n\to\infty} \bar{\vartheta}_n(x) = 1$ permet d'assurer l'existence d'une constante $m > 0$, telle que pour n assez grand $\bar{\vartheta}_n(x) \geqslant m \quad \forall x \in \mathbb{R}$.

Finalement nous avons $\left| E\left(\frac{\psi_n(x)}{\bar{\vartheta}_n(x)} \cdot \frac{\vartheta_n(x) - \bar{\vartheta}_n(x)}{\bar{\vartheta}_n(x)}\right)\right| \leqslant \frac{M_2}{m} \cdot \frac{1}{h_n \sqrt{n}}$ <u>i.e.</u> c)

III.6.3. - <u>Etude de la variance</u>.

On se propose dans ce paragraphe, d'étudier le développement asymptotique de la variance de $\tilde{f}_{n1}(x)$.

<u>*Proposition 4.15.*</u>- Si f_Q est continue bornée sur $\mathbb{R}$, et admet des dérivées d'ordre 2 et 3 bornées. Si K est une densité de probabilité sur $\mathbb{R}$, bornée et telle que

$$\int_{\mathbb{R}} t^i K(t)\, dt < \infty \quad i=1,2 \qquad \text{et} \qquad \int_{\mathbb{R}} K^2(t)\, dt < \infty \; .$$

Si $h_n \xrightarrow[n\to\infty]{} 0$ et $n\, h_n^2 \xrightarrow[n\to\infty]{} \infty$.

Alors $$V(\tilde{f}_{n1}(x)) = O\left(\frac{1}{n\, h_n^2}\right) .$$

<u>*Preuve*</u> *:* Nous savons que $\tilde{f}_{n1}(x)$ s'écrit sous la forme :

$$\tilde{f}_{n1}(x) = g_{U_n}(x) + \frac{\psi_n(x)}{\bar{\vartheta}_n(x)} - \frac{\psi_n(x)}{\vartheta_n(x)} \cdot \frac{\vartheta_n(x) - \bar{\vartheta}_n(x)}{\bar{\vartheta}_n(x)}$$

donc

$$V(\tilde{f}_{n1}(x)) = V(g_{U_n}(x)) + \frac{1}{\bar{\vartheta}_n^2(x)} V(\psi_n(x)) + V\left(\frac{\psi_n(x)}{\vartheta_n(x)} \cdot \frac{\vartheta_n(x) - \bar{\vartheta}_n(x)}{\bar{\vartheta}_n(x)}\right)$$

$$+ 2 \operatorname{Cov}\left(g_{U_n}(x), \frac{\psi_n(x)}{\bar{\vartheta}_n(x)}\right) - 2 \operatorname{Cov}\left(g_{U_n}(x), \frac{\psi_n(x)}{\vartheta_n(x)} \frac{\vartheta_n(x) - \bar{\vartheta}_n(x)}{\bar{\vartheta}_n(x)}\right)$$

$$- 2 \operatorname{Cov}\left(\frac{\psi_n(x)}{\bar{\vartheta}_n(x)}, \frac{\psi_n(x)}{\vartheta_n(x)} \cdot \frac{\vartheta_n(x) - \bar{\vartheta}_n(x)}{\bar{\vartheta}_n(x)}\right) .$$

Nous allons déterminer le développement asymptotique de chaque terme du second membre de cette égalité ,

1) $V(g_{U_n}(x)) = \dfrac{1}{n\, h_n} f_Q(x) \displaystyle\int_{-1}^{1} K^2(y)\, dy + o(1/n\, h_n)$

en effet un calcul direct nous donne $V(g_{U_n}(x)) = \dfrac{1}{n\, h_n^2} E(K^2(\frac{x-X_1}{h_n})\, 1_{U_n}(X_1)) - \frac{1}{n} (E\, g_{U_n}(x))^2$, nous savons, d'après l'étude du biais, que $E\, g_{U_n}(x)^n$ converge vers $f_Q(x) \int_{-1}^{1} K(t)\, dt$ lorsque $n \to \infty$, par conséquent $\frac{1}{n} E^2\, g_{U_n}(x)$ est un $O(1/n)$ ainsi $V(g_{U_n}(x)) = \dfrac{1}{n\, h_n} \displaystyle\int_{-1}^{1} K^2(t)\, f_Q(x-h_n\, t)\, dt + O(1/n)$ ou encore, en utilisant la formule de Taylor, $V(g_{U_n}(x)) = \dfrac{1}{n\, h_n} \displaystyle\int_{-1}^{1} K^2(t)\, dt\, f_Q(x) + o(\frac{1}{n\, h_n})$.

2) $V(\vartheta_n(x)) = \dfrac{1}{n\, h_n} f_Q(x) \displaystyle\int_{\mathbb{R}} \left(\int_{]-1,1[^c} K(s)\, K(t-s)\, ds\right)^2 dt + o(\frac{1}{n\, h_n})$

$\psi_n(x)$ étant défini par $\dfrac{1}{n^2\, h_n^2} \displaystyle\sum_{i,j=1}^{n} 1_{U_n^c}(X_i) \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-X_j}{h_n})\, dy$

nous avons

$$E(\psi_n^2(x)) = \frac{1}{n^4\, h_n^4} E\left[\sum_{i,j,k,\ell=1}^{n} 1_{U_n^c}(X_i)\, 1_{U_n^c}(X_j) \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-X_k}{h_n})\, dy \right.$$

$$\left. \cdot \int_{U_n^c} K(\frac{x-z}{h_n})\, K(\frac{z-X_\ell}{h_n})\, dz\right]$$

Afin d'alléger les notations et provisoirement nous poserons :

$$\tilde{S}_i = 1_{U_n^c}(X_i) \quad \text{et} \quad T_i = \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-X_i}{h_n})\, dy$$

avec ces notations nous obtenons :

$$\begin{aligned} n^4 h_n^4 E(\psi_n^2(x)) = {} & n\, E(\tilde{S}_1 \tilde{S}_1 T_1 T_1) + n(n-1)\, E(\tilde{S}_1 \tilde{S}_1 T_1 T_2) + n(n-1)\, E(\tilde{S}_1 \tilde{S}_1 T_2 T_1) \\ & + n(n-1)\, E(\tilde{S}_1 \tilde{S}_1 T_2 T_2) + n(n-1)(n-2)\, E(\tilde{S}_1 \tilde{S}_1 T_2 T_3) + n(n-1)\, E(\tilde{S}_1 \tilde{S}_2 T_1 T_1) + \\ & + n(n-1)\, E(\tilde{S}_1 \tilde{S}_2 T_1 T_2) + n(n-1)(n-2)\, E(\tilde{S}_1 \tilde{S}_2 T_1 T_3) + n(n-1)\, E(\tilde{S}_1 \tilde{S}_2 T_2 T_1) \\ & + n(n-1)\, E(\tilde{S}_1 \tilde{S}_2 T_2 T_2) + n(n-1)(n-2)\, E(\tilde{S}_1 \tilde{S}_2 T_2 T_3) + n(n-1)(n-2)\, E(\tilde{S}_1 \tilde{S}_2 T_3 T_1) \\ & + n(n-1)(n-2)\, E(\tilde{S}_1 \tilde{S}_2 T_3 T_2) + n(n-1)(n-2)\, E(\tilde{S}_1 \tilde{S}_2 T_3 T_3) + \\ & + n(n-1)(n-2)(n-3)\, E(\tilde{S}_1 \tilde{S}_2 T_3 T_4) \end{aligned}$$

avec les mêmes notations :

$$h_n^4 (E\, \psi_n(x))^2 = (\frac{1}{n} E(\tilde{S}_1 T_1) + \frac{n-1}{n} E(\tilde{S}_1 T_2))^2$$

qu'on peut encore écrire sous la forme

$$\frac{1}{n^2} E(\tilde{S}_1 T_1 \tilde{S}_2 T_2) + (\frac{n-1}{n})^2 E(\tilde{S}_1 \tilde{S}_2 T_3 T_4) + 2\, \frac{n-1}{n^2} E(\tilde{S}_1 T_1 \tilde{S}_2 T_3)$$

d'où

$$\begin{aligned} V(\psi_n(x)) & = E(\psi_n^2(x)) - E^2(\psi_n(x)) \\ & = \frac{1}{h_n^4} \left\{ \frac{1}{n^3} E(\tilde{S}_1 T_1^2) + \frac{2n(n-1)}{n^4} E(T_1)\, E(\tilde{S}_1 T_1) + \frac{n(n-1)}{n^4} E(\tilde{S}_1)\, E(T_1^2) \right. \\ & + \frac{2n(n-1)}{n^4} E(\tilde{S}_1 T_1^2)\, E(\tilde{S}_1) + \frac{n^2-2n}{n^4} E^2(\tilde{S}_1 T_1) + \frac{n(n-1)(n-2)}{n^4} E(\tilde{S}_1)\, E^2(T_1) \\ & + \frac{n(n-1)(n-2)}{n^4} (E\, \tilde{S}_1)^2\, E(T_1^2) + \frac{-4n^3+10n^2-6n}{n^4} (E\, \tilde{S}_1)^2\, (E\, T_1)^2 \\ & \left. + \frac{2n^3-10n^2+8n}{n^4}\, E(\tilde{S}_1)\, E(T_1)\, E(\tilde{S}_1 T_1) \right\} \end{aligned}$$

par ailleurs, d'après la définition de $\tilde{S}_i$ et T_i , nous avons

$$E(\tilde{S}_1) = Q(U_n^c) = 1 - 2\, h_n\, f_Q(x) + \bigcirc(h_n^3)$$

$$E(T_1) = \int_{\mathbb{R}} \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-z}{h_n})\, f_Q(z)\, dy\, dz$$

$$= h_n^2 \int_{\mathbb{R}} \int_{]-1,1[^c} K(s)\, K(t-s)\, f_Q(x-h_n t)\, dt\, ds$$

$$E(\tilde{S}_1\, T_1) = \int_{U_n^c} \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-z}{h_n})\, f_Q(z)\, dy\, dz$$

$$= h_n^2 \int_{]-1,1[^c} \int_{]-1,1[^c} K(y)\, K(t-y)\, f_Q(x-h_n\, t)\, dt\, dy$$

$$E(T_1^2) = \int_{\mathbb{R}} (\int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-z}{h_n})\, dy)^2\, f_Q(z)\, dz$$

$$= h_n^3 \int_{\mathbb{R}} (\int_{]-1,1[^c} K(s)\, K(t-s)\, ds)^2\, f_Q(x-h_n t)\, dt$$

$$E(\tilde{S}_1\, T_1^2) = \int_{U_n^c} (\int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-z}{h_n})\, dy)^2\, f_Q(z)\, dz$$

$$= h_n^3 \int_{]-1,1[^c} (\int_{]-1,1[^c} K(s)\, K(t-s)\, ds)^2\, f_Q(x-h_n t)\, dt$$

les conditions imposées à f_Q , K et à h_n permettent d'avoir

(V2) $$\int_A (\int_{]-1,1[^c} K(s)\, K(t-s)\, ds)^i\, f_Q(x-h_n t)\, dt$$

$$= f_Q(x) \int_A (\int_{]-1,1[^c} K(s)\, K(t-s)\, ds)^i\, dt + \bigcirc(h_n)$$

avec $A = \mathbb{R}$ ou $]-1,1[^c$ et $i=1,2$

(il suffit d'appliquer la formule de Taylor à l'ordre 1 à $f_Q(x-h_n t)$ et de remarquer que $H^i(t) = (\int_{]-1,1[^c} K(s)\, K(t-s)\, ds)^i$ est paire d'où $t\, H^i(t)$ est impaire et donc $\int_A t\, H^i(t)\, dt = 0)$.

Pour conclure il suffit de reprendre (VI), de remplacer $E(\tilde{S}_1^i\ T_1^j)$, $i \in \overline{0,1}$, $j \in \overline{1,2}$ par son développement obtenu à partir de (V2), et on obtient

$$V\ \psi_n(x) = \frac{1}{n\ h_n}\ f_Q(x) \int_{\mathbb{R}} \left(\int_{]-1,1[^c} K(s)\ K(t-s)\ ds\right)^2 dt + o(1/(n\ h_n))$$

3) $$\mathrm{Cov}(g_{U_n}(x),\ \psi_n(x)) = \frac{1}{n\ h_n}\ f_Q(x) \int_{-1}^{1} \int_{]-1,1[^c} K(s)\ K(t)\ K(t-s)\ ds\ dt + \bigcirc(1/n)$$

en effet : le fait que les X_i soient i.i.d. nous donne

$$E(g_{U_n}(x)\ \psi_n(x)) = \frac{1}{n^3\ h_n^3} \sum_{i,j,k=1}^{n} E(L_i\ \tilde{S}_j\ T_k)$$

$$= \frac{1}{n^3\ h_n^3}\ ((n-1)\ E(L_1\ T_1)\ E(\tilde{S}_2) + n(n-1)\ E(L_1)\ E(T_2\ \tilde{S}_2) + n(n-1)(n-2)\ E(L_1)\ E(T_2)\ E(\tilde{S}_3))$$

avec $\tilde{S}_i$ et T_i définis de la même façon que précédemment et L_i donné par :

$$L_i = K\left(\frac{x-X_i}{h_n}\right)\ 1_{U_n}(X_i) \quad , \quad \forall i \in \overline{1,n} \ .$$

Ensuite nous savons que :

$$E(L_1) = h_n \int_{-1}^{1} K(s)\ f_Q(x-h_n s)\ ds = h_n\ f_Q(x) \int_{-1}^{1} K(s)\ ds + \bigcirc(h_n^3)$$

$$E(\tilde{S}_1) = Q(U_n^c) = 1 - 2\ h_n\ f_Q(x) + \bigcirc(h_n^3)$$

$$E(\tilde{S}_1\ T_1) = h_n^2\ f_Q(x) \int_{]-1,1[^c} \int_{]-1,1[^c} K(s)\ K(t-s)\ ds\ dt + \bigcirc(h_n^4)$$

$$E(T_1) = h_n^2\ f_Q(x) \int_{\mathbb{R}} \int_{]-1,1[^c} K(s)\ K(t-s)\ ds\ dt + \bigcirc(h_n^4)$$

par ailleurs avec le changement de variables $\frac{x-y}{h_n} = s$ et $\frac{x-z}{h_n} = t$ dans

$$E(T_1\ L_1) = \int_{U_n} \int_{U_n^c} K\left(\frac{x-z}{h_n}\right) K\left(\frac{x-y}{h_n}\right) K\left(\frac{y-z}{h_n}\right) f_Q(z)\ dy\ dz$$

on a

$$E(T_1\ L_1) = h_n^2 \int_{]-1,1[} \int_{]-1,1[^c} K(t)\ K(s)\ K(t-s)\ f_Q(x-h_n t)\ ds\ dt$$

or d'après la formule de Taylor

$$\int_{]-1,1[} \int_{]-1,1[^c} K(t)\ K(s)\ K(t-s)\ f_Q(x-h_n t)\ ds\ dt$$

$$= f_Q(x) \int_{-1}^{1} \int_{]-1,1[^c} K(t)\ K(s)\ K(t-s)\ ds\ dt$$

$$+ \frac{h_n^2}{2} \int_{-1}^{1} \int_{]-1,1[^c} t^2\ K(t)\ K(s)\ K(t-s)\ f''_Q(x-\theta h_n t)\ ds\ dt$$

où $\theta \in]0,1[$

(on a utilisé la propriété $\int_{]-1,1[^c} K(s)\ K(t-s)\ ds$ est paire en t, ce qui entraîne que $t\ K(t) \int_{]-1,1[^c} K(s)\ K(t-s)\ ds$ est impaire en t).

Donc $E(T_1\ L_1) = h_n^2\ f_Q(x) \int_{-1}^{1} \int_{]-1,1[^c} K(s)\ K(t)\ K(t-s)\ ds\ dt + \bigcirc(h_n^4)$

enfin à partir de $Cov(g_{U_n}(x),\ \psi_n(x)) = E(g_{U_n}(x).\psi_n(x)) - E(g_{U_n}(x))\ E(\psi_n(x))$

et de $E(g_{U_n}(x))\ .\ E(\psi_n(x)) = \frac{1}{n\ h_n^3}\ \left[E(L_1)\ E(\tilde{S}_1\ T_1) + (n-1)\ E(L_1)\ E(\tilde{S}_1\ T_2)\right]$

on obtient 3)

4) $$V\left(\frac{\psi_n(x)}{\phi_n(x)}\ \frac{\phi_n(x) - \bar{\phi}_n(x)}{\bar{\phi}_n(x)}\right) = \bigcirc\left(\frac{1}{n\ h_n^2}\right)$$

En effet $$V\left(\frac{\psi_n(x)}{\phi_n(x)}\ .\ \frac{\phi_n(x) - \bar{\phi}_n(x)}{\bar{\phi}_n(x)}\right) \leq E\left\{\frac{\psi_n(x)}{\phi_n(x)}\ .\ \frac{\phi_n(x) - \bar{\phi}_n(x)}{\bar{\phi}_n(x)}\right\}^2$$

$$\leq \frac{M_2^2}{h_n^2}\ \frac{1}{\bar{\phi}_n^2(x)}\ .\ V(\phi_n(x))$$

car $\dfrac{\psi_n(x)}{\varphi_n(x)} \leq \dfrac{M_2}{h_n}$, d'autre part $V(\varphi_n(x)) \leq \frac{1}{n}$ et $\bar{\varphi}_n(x) \geq m > 0$

impliquent que :

$$V\left(\frac{\psi_n(x)}{\varphi_n(x)} \cdot \frac{\varphi_n(x) - \bar{\varphi}_n(x)}{\bar{\varphi}_n(x)}\right) \leq \frac{1}{n\ h_n^2}\ (\frac{M_2}{m})^2$$

$$5) \quad \mathrm{Cov}\left(g_{U_n}(x)\ ,\ \frac{\psi_n(x)}{\varphi_n(x)}\ \frac{\varphi_n(x) - \bar{\varphi}_n(x)}{\bar{\varphi}_n(x)}\right) = O\left(\frac{1}{n\ h_n^2}\right)$$

D'après l'inégalité de Schwarz, nous avons

$$\left|\mathrm{Cov}\left(g_{U_n}(x)\ ,\ \frac{\psi_n(x)}{\varphi_n(x)}\ \frac{\varphi_n(x) - \bar{\varphi}_n(x)}{\bar{\varphi}_n(x)}\right)\right| \leq$$

$$\left\{V(g_{U_n}(x))\ .\ V\left(\frac{\psi_n(x)}{\varphi_n(x)} \cdot \frac{\varphi_n(x) - \bar{\varphi}_n(x)}{\bar{\varphi}_n(x)}\right)\right\}^{1/2}$$

or $$V(g_{U_n}(x)) \leq \frac{1}{n\ h_n^2} \int_{U_n} K^2(\frac{x-y}{h_n})\ f_Q(y)\ dy \leq \frac{M_2^2}{n\ h_n^2}$$

donc en utilisant 4) on obtient

$$\mathrm{Cov}\left(g_{U_n}(x)\ ,\ \frac{\psi_n(x)}{\varphi_n(x)} \cdot \frac{\varphi_n(x) - \bar{\varphi}_n(x)}{\bar{\varphi}_n(x)}\right) \leq \frac{\tau}{n\ h_n^2} \quad \text{où} \quad \tau = \frac{M_2^2}{m}$$

d'où 5).

$$6) \quad \mathrm{Cov}\left(\frac{\psi_n(x)}{\bar{\varphi}_n(x)}\ ,\ \frac{\psi_n(x)}{\varphi_n(x)} \cdot \frac{\varphi_n(x) - \bar{\varphi}_n(x)}{\bar{\varphi}_n(x)}\right) = O\left(\frac{1}{n\ h_n^{3/2}}\right)$$

on sait que $$V\left(\frac{\psi_n(x)}{\varphi_n(x)} \cdot \frac{\varphi_n(x) - \bar{\varphi}_n(x)}{\bar{\varphi}_n(x)}\right) \leq \frac{1}{n\ h_n^2}\left(\frac{M_2}{m}\right)^2$$

et que $$V\left(\frac{\psi_n(x)}{\bar{\varphi}_n(x)}\right) = O(\frac{1}{n\ h_n})$$ d'où 6).

III.6.4. - Consistance de $\widetilde{f_{n1}}(x)$

Nous allons montrer que sous certaines conditions sur h_n, K et f_Q le nouvel estimateur $\widetilde{f_{n1}}(x)$ converge en probabilité ou presque complètement vers $f_Q(x)$.

a) Convergence en probabilité.

On note $V_{ij} = \frac{1}{\bar{\varphi}_n(x)} \frac{1}{h_n^2} 1_{U_n^c}(X_i) \int_{U_n^c} K(\frac{x-y}{h_n}) K(\frac{y-X_j}{h_n})\, dy$, $i,j = 1,\ldots,n$

ce qui permet d'écrire

$$\frac{\psi_n(x)}{\bar{\varphi}_n(x)} = \frac{1}{n^2} \sum_{i,j=1}^{n} V_{ij} \ .$$

Lemme 4.5.- Si $h_n \to 0$ et $n\, h_n \to \infty$ lorsque $n \to \infty$, f_Q est continue bornée au voisinage de x.

$$\int K \, d\lambda = 1 \quad , \quad K \text{ est positif borné et pair}$$

Alors

a) $\dfrac{\psi_n(x) - \bar{\psi}_n(x)}{\bar{\varphi}_n(x)} \xrightarrow[n\to\infty]{P} 0$

b) $\dfrac{\varphi_n(x) - \bar{\varphi}_n(x)}{\bar{\varphi}_n(x)} \xrightarrow[n\to\infty]{P} 0$

c) $g_{U_n}(x) \xrightarrow[n\to\infty]{P} f_Q(x) \displaystyle\int_{-1}^{1} K(y)\, dy$.

Preuve : a) et b) résultent immédiatement des majorations de $V(\psi_n(x))$ et $V(\varphi_n(x))$ obtenues précédemment et de la minoration de $\bar{\psi}_n(x)$ pour n assez grand.

c) $g_{U_n}(x)$ peut s'écrire sous la forme :

$$g_{U_n}(x) = \frac{1}{n\,h_n} \sum_{j=1}^{n} \left\{ K\left(\frac{x-X_j}{h_n}\right) 1_{U_n}(X_j) - \int_{U_n} K\left(\frac{x-y}{h_n}\right) f_Q(y)\, dy \right\} + \frac{1}{h_n} \int_{U_n} K\left(\frac{x-y}{h_n}\right) f_Q(y)\, dy \,.$$

On montre facilement que $\frac{1}{h_n} \int_{U_n} K\left(\frac{x-y}{h_n}\right) f_Q(y)\, dy \xrightarrow[n\to\infty]{} f_Q(x) \int_{]-1,1[} K\, d\lambda$

et que $\left(\frac{1}{n\,h_n} \sum_{j=1}^{n} K\left(\frac{x-X_j}{h_n}\right) 1_{U_n}(X_j)\right) - \frac{1}{h_n} \int_{U_n} K\left(\frac{x-y}{h_n}\right) f_Q(y)\, dy \xrightarrow[n\to\infty]{P} 0$

finalement en écrivant

$$\tilde{f}_{n1}(x) = g_{U_n}(x) + \frac{\dfrac{\psi_n(x) - \bar{\psi}_n(x)}{\bar{\psi}_n(x)} + \dfrac{\bar{\psi}_n(x)}{\bar{\psi}_n(x)}}{1 + \dfrac{\psi_n(x) - \bar{\psi}_n(x)}{\bar{\psi}_n(x)}}$$

et en utilisant le lemme 4.5. on obtient la proposition suivante :

<u>*Proposition 4.16.*</u>- Si $h_n \xrightarrow[n\to\infty]{} 0$, $n\,h_n \xrightarrow[n\to\infty]{} \infty$

K est une densité de probabilité bornée paire .

f_Q continue bornée au voisinage de x.

Alors $\tilde{f}_{n1}(x) \xrightarrow[n\to\infty]{P} f_Q(x)$.

b) <u>Convergence presque complète</u>.

Nous allons procéder de la même façon que précédemment, pour montrer que $\tilde{f}_{n1}(x)$ converge presque complètement vers $f_Q(x)$.

<u>*Proposition 4.17.*</u>- Si f_Q est continue bornée au voisinage de x.

K est positif, borné et $\int_{\mathbb{R}} K\, d\lambda = 1$

$h_n \xrightarrow[n\to\infty]{} 0$ et $\forall \alpha > 0 \;\; \sum_{n\geq 1} e^{-\alpha n h_n} < \infty$.

Alors $\tilde{f}_{n1}(x) \xrightarrow{p.co.} f_Q(x)$.

Preuve : La démonstration de ce résultat va se faire en quatre étapes, il suffit de décomposer $\tilde{f}_{n1}(x)$ comme dans la preuve du lemme 4.5.

et de montrer que $g_{U_n}(x) \xrightarrow[n\to\infty]{p.co.} f_Q(x) \int_{-1}^{1} K\, d\lambda$; $\dfrac{\psi_n(x) - \bar{\psi}_n(x)}{\bar{\emptyset}_n(x)} \xrightarrow[n\to\infty]{p.co.} 0$

$\dfrac{\emptyset_n(x) - \bar{\emptyset}_n(x)}{\bar{\emptyset}_n(x)} \xrightarrow[n\to\infty]{p.co.} 0$; $\dfrac{\bar{\psi}_n(x)}{\bar{\emptyset}_n(x)} \xrightarrow[n\to\infty]{} f_Q(x) \int_{]-1,1[^c} K\, d\lambda$.

Nous utiliserons l'inégalité de Hoeffding, pour établir la convergence p.co. de ces différentes quantités :

" Si $(T_i)_{i=1,\dots,n}$ est une suite de v.a. indépendantes, centrées, et majorées par $b > 0$, si $S_n = \sum_{i=1}^{n} T_i$, $\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} E(T_i^2)$ alors

$$\forall \varepsilon \in]0,b[\quad \Pr\{S_n \geq n_\varepsilon\} \leq \exp\left(- \frac{n\varepsilon}{b} \left(\left(1 + \frac{\sigma^2}{b\varepsilon}\right) \operatorname{Log}\left(1 + \frac{b\varepsilon}{\sigma^2}\right) - 1\right)\right)$$

i) $\dfrac{\psi_n(x) - \bar{\psi}_n(x)}{\bar{\emptyset}_n(x)} \xrightarrow[n\to\infty]{p.co.} 0$

Afin de faire apparaître la somme de v.a. indépendantes, on décompose

$\dfrac{\psi_n(x)-\bar{\psi}_n(x)}{\bar{\emptyset}_n(x)}$ en 2 parties :

$$\frac{\psi_n(x) - \bar{\psi}_n(x)}{\bar{\emptyset}_n(x)} = \frac{1}{n^2} \sum_{i=1}^{n} (V_{ii} - E\, V_{ii}) + \frac{1}{n^2} \cdot \sum_{i \neq j} (V_{ij} - E\, V_{ij})$$

avec $V_{ij} = \dfrac{1}{h_n^2\, \bar{\emptyset}_n(x)}\, 1_{U_n^c}(X_i) \displaystyle\int_{U_n^c} K\left(\frac{x-y}{h_n}\right) K\left(\frac{y-X_j}{h_n}\right) dy$, $i\ ,\ j = 1,\dots,n$.

Les v.a. $\frac{1}{n}(V_{ii} - E(V_{ii})$ $i \in \overline{1,n}$ sont indépendantes, centrées de valeurs absolues majorées par : $\dfrac{2\, M_2}{n\, h_n\, \bar{\emptyset}_n(x)} = b$, car :

$$\left|\frac{1}{n}(V_{ii}-E(V_{ii}))\right| \leqslant$$

$$\frac{1}{n}\frac{1}{\bar{\psi}_n(x)h_n^2}\int_{U_n^c} K(\frac{x-y}{h_n})\,K(\frac{y-X_i}{h_n})\,dy + \frac{1}{n\,h_n^2}\frac{1}{\bar{\psi}_n(x)}\int_{U_n^c}\int_{U_n^c} K(\frac{x-y}{h_n})\,K(\frac{y-z}{h_n})\,f_Q(z)dy\,dz$$

d'où

$$\left|\frac{1}{n}(V_{ii}-E(V_{ii}))\right| \leqslant \frac{M_2}{n\,h_n\,\bar{\psi}_n(x)}\left\{\frac{1}{h_n}\int_{U_n^c} K(\frac{x-y}{h_n})\,dy + \frac{1}{h_n}Q(U_n^c)\int_{U_n^c} K(\frac{x-y}{h_n})\,dy\right\}$$

$$\leqslant \frac{2.M_2}{n\,h_n\,\bar{\psi}_n(x)}\ .$$

Notons $\sigma_1^2 = \frac{1}{n}\sum_{i=1}^{n} E\{\frac{1}{n}(V_{ii} - E(V_{ii}))\}^2 = \frac{1}{n^2}\mathrm{Var}(V_{11})$

$$\leqslant \frac{1}{n^2\,h_n^4}\frac{1}{\bar{\psi}_n^2(x)}\int_{U_n^c}\left\{\int_{U_n^c} K(\frac{x-y}{h_n})\,K(\frac{y-z}{h_n})\,dy\right\}^2 f_Q(z)\,dz$$

$$\leqslant \frac{M_2}{n^2\,\bar{\psi}_n^2(x)}\frac{1}{h_n^3}\int_{U_n^c}\int_{U_n^c} K(\frac{x-y}{h_n})\,K(\frac{y-z}{h_n})\,f_Q(z)\,dy\,dz$$

$$= \frac{M_2}{n^2\,\bar{\psi}_n(x)}\frac{1}{h_n}\frac{\int_{]-1,1[^c}\int_{]-1,1[^c} K(y)\,K(t-z)\,f_Q(x-h_nt)\,dy\,dt}{\int_{\mathbb{R}} K(t)\,Q(]x-h_n(1+t),\ x+h_n(1-t)[^c)\,dt}$$

$$\leqslant \frac{M_1\,M_2}{n^2\,\bar{\psi}_n(x)}\frac{1}{h_n}\cdot\frac{1}{\int_{\mathbb{R}} K(t)\,Q(]x-h_n(1+t),x+h_n(1-t)[^c)\,dt}$$

or $\int_{\mathbb{R}} K(t)\ Q(]x-h_n(1+t),x+h_n(1-t)[^c)\ dt \xrightarrow[n\to\infty]{} 1$, donc on peut toujours minorer, à partir d'un certain rang N_o , $\int_{\mathbb{R}} K(t)\ Q(]x-h_n(1+t),\ x+h_n(1-t)[^c)\ dt$ par un réel η strictement positif. D'où

$$\sigma_1^2 \leqslant \frac{M_1\cdot M_2}{n^2\,\bar{\psi}_n(x)}\cdot\frac{1}{h_n\eta}$$

L'inégalité de Hoeffding nous donne donc :

$$\Pr\{|\frac{1}{n}\sum_{i=1}^{n}(V_{ii}-E\,V_{ii})| \geq n\varepsilon\} \leq 2\exp\left[-\frac{n\varepsilon}{b}\left((1+\frac{\sigma_1^2}{b\varepsilon})\,\mathrm{Log}(1+\frac{b\varepsilon}{\sigma_1^2})-1\right)\right]$$

pour $\varepsilon \in]0,b[$.

Le fait que la fonction $\ell(u) = (1+u)\log(1+\frac{1}{u}) - 1$ soit strictement positive et décroissante sur $]0,+\infty[$ et que

$$\frac{\sigma_1^2}{b\varepsilon} \leq \left(\frac{M_1\,M_2}{n^2\,\bar{\psi}_n(x)\,b_n\,\eta}\right)\frac{1}{\varepsilon}\left(\frac{n\,h_n\,\bar{\psi}_n(x)}{2\,M_2}\right) = \frac{M_1}{2n\varepsilon\eta} \leq \frac{M_1}{2\varepsilon\eta}$$

entraînent que :

$$(1+\frac{\sigma_1^2}{b\varepsilon})\,\mathrm{Log}(1+\frac{b\varepsilon}{\sigma_1^2}) - 1 \geq (1+\frac{M_1}{2\varepsilon\eta})\,\mathrm{Log}(1+\frac{2\varepsilon\eta}{M_1}) - 1$$

$$= \beta > 0\ .$$

D'où $\forall n > N_o$

$$\Pr\left\{|\frac{1}{n}\sum_{i=1}^{n}(V_{ii}-EV_{ii})| \geq n\varepsilon\right\} \leq \exp(-\frac{n^2\varepsilon}{2M_2}\,.\,\beta\,.\,h_n\,\bar{\psi}_n(x)) \quad ,$$

or par définition de N_o , $\forall n \geq N_o$ $\quad \bar{\psi}_n(x) \geq \eta$; donc

$$\forall\varepsilon \in]0,b[\ ,\quad \sum_{n\geq N_o}\Pr\left\{\frac{1}{n^2}\sum_{i=1}^{n}(V_{ii}-EV_{ii}) \geq \varepsilon\right\} \leq$$

$$\sum_{n\geq N_o}\exp(-n^2\,h_n\,\alpha) < +\infty \quad \text{où} \quad \alpha = \frac{\beta\eta\varepsilon}{2M_2}\ .$$

On en déduit que :

$$\frac{1}{n^2}\sum_{i=1}^{n}(V_{ii}-EV_{ii}) \xrightarrow[n\to\infty]{p.co.} 0\ .$$

D'autre part

$$\sum_{i\neq j} \frac{1}{n^2} (V_{ij} - E(V_{ij})) = \frac{1}{n^2\, h_n^2\, \bar{\psi}_n(x)} \sum_{i\neq j} \left\{ 1_{U_n^c}(X_i) \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-X_j}{h_n})\, dy \right.$$

$$\left. - Q(U_n^c) \int_{\mathbb{R}} \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-z}{h_n})\, f_Q(z)\, dy\, dz \right\}$$

est majoré par : $S_{2n} + S_{3n}$ où

$$S_{2n} = \frac{n-1}{n^2} \frac{1}{h_n^2\, \bar{\psi}_n(x)} \sum_{j=1}^{n} \left(\int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-X_j}{h_n})\, dy - \right.$$

$$\left. - \int_{\mathbb{R}} \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-z}{h_n})\, f_Q(z)\, dy\, dz \right)$$

et

$$S_{3n} = \frac{n-1}{n} \cdot \frac{1}{h_n^2} \cdot \frac{1}{\bar{\psi}_n(x)} (1 - Q(U_n^c)) \int_{\mathbb{R}} \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-z}{h_n})\, f_Q(z)\, dy\, dz$$

il suffit donc de montrer que $S_{2n} \xrightarrow[n\to\infty]{p.co.} 0$ et $S_{3n} \xrightarrow[n\to\infty]{} 0$.

Les conditions imposées à h_n et à f_Q permettent d'établir, sans peine, que $S_{3n} \xrightarrow[n\to\infty]{} 0$. Pour S_{2n} nous allons procéder de la même façon que précédemment ; posons pour $j \in \overline{1,n}$

$$T_j = \frac{n-1}{n} \frac{1}{h_n^2\, \bar{\psi}_n(x)} \left\{ \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-X_j}{h_n})\, dy - \right.$$

$$\left. - \int_{\mathbb{R}} \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-z}{h_n})\, f_Q(z)\, dy\, dz \right\}$$

ces v.a. sont indépendantes, centrées, majorées par $\frac{n-1}{n} \frac{1}{\bar{\psi}_n(x)} \frac{2\, M_2}{h_n}$

et $\sigma_2^2 = \frac{1}{n} \sum_{j=1}^{n} E\, T_j^2 = E\, T_1^2$

$$\leq (\frac{n-1}{n})^2 \frac{1}{h_n^4} \frac{1}{\bar{\psi}_n^2(x)} \int_{\mathbb{R}} \left\{ \int_{U_n^c} K(\frac{x-y}{h_n})\, K(\frac{y-z}{h_n})\, dy \right\}^2 f_Q(z)\, dz$$

$$\leq (\frac{n-1}{n})^2 \frac{1}{h_n} \cdot \frac{1}{\bar{\psi}_n^2(x)} \cdot M_1 \cdot M_2^2 \ .$$

Pour $n \geq N_o$ $\quad \sigma_2^2 \leq (\frac{n-1}{n})^2 \frac{1}{h_n} \frac{1}{\bar{\psi}_n(x)} \frac{M_1 M_2}{\eta}$; par un raisonnement analogue au précédent on conclut que $S_{3n} \xrightarrow[n\to\infty]{p.co.} 0$.

ii) $\dfrac{\psi_n(x) - \bar{\psi}_n(x)}{\bar{\psi}_n(x)} \xrightarrow[n\to\infty]{p.co.} 0$.

En effet, posons $\dfrac{\psi_n(x) - \bar{\psi}_n(x)}{\bar{\psi}_n(x)} = \dfrac{1}{n} \sum_{j=1}^{n} (L_j - E\,L_j) = \dfrac{1}{n} S_{n4}$

où $\quad L_j = \dfrac{1}{\bar{\psi}_n(x)} \dfrac{1}{h_n} \displaystyle\int_{U_n^c} K(\frac{y-X_j}{h_n})\, dy \quad , \quad j \in \overline{1,n}$

les v.a. $(L_j - E\,L_j)_{j\in\overline{1,n}}$ sont indépendantes, centrées et majorées par $\dfrac{2}{\bar{\psi}_n(x)}$

notons $\sigma_3^2 = \dfrac{1}{n} \sum_{i=1}^{n} E(L_j - E\,L_j)^2$.

On a $\quad \sigma_3^2 \leq \dfrac{1}{h_n^2\, \bar{\psi}_n^2(x)} \displaystyle\int_{\mathbb{R}} \left\{\int_{U_n^c} K(\frac{y-z}{h_n})\, dy\right\}^2 f_Q(z)\, dz \leq \dfrac{1}{\bar{\psi}_n^2(x)}$.

L'inégalité de Hoeffding appliquée à S_{n4} nous donne, pour $n \geq N_o$;

$$\forall \varepsilon \in]0,b'[\qquad Pr\{|S_{n4}| \geq n\varepsilon\} \leq 2\, e^{-n\theta\varepsilon} \quad \text{où}$$

$$b' = \frac{2}{\bar{\psi}_n(x)} \quad , \quad \theta = \frac{1}{2}\, \eta\, \{(1 + \frac{1}{2\varepsilon\eta}) \log(1+2\varepsilon\eta) - 1\}$$

ainsi $\dfrac{\psi_n(x) - \bar{\psi}_n(x)}{\bar{\psi}_n(x)} \xrightarrow[n\to\infty]{p.co.} 0$ ■

Enfin nous savons, d'après III.3., que $g_{U_n}(x) \xrightarrow{p.co.} f_Q(x) \displaystyle\int_{]-1,1[} K\, d\lambda$, ce qui achève la preuve de $\tilde{f}_{n1}(x) \xrightarrow{p.co.} f_Q(x)$.

III.7. - Estimation séquentielle par la méthode du noyau.-

On se propose, dans cette dernière partie, d'appliquer le conditionnement, par rapport à la statistique S (définie par la proposition 4.1), à une classe d'estimateurs à noyaux séquentiels ; classe dûe à P. DEHEUVELS (1974) et qui généralise les estimateurs de Parzen-Rosenblatt, YAMATO (1971) etc ...

Soit $(h_n)_n$ une suite de réels strictement positifs tels que : $h_n \xrightarrow[n\to\infty]{} 0$

H une fonction arbitraire définie sur $\mathbb{R}_+^*$ dans $\mathbb{R}_+^*$

K une fonction réelle λ-intégrable, bornée ,

l'estimateur de P. DEHEUVELS (1974) est défini par :

$$\bar{f}_n(x) = \frac{1}{\sum_{i=1}^{n} h_i \, H(h_i)} \sum_{k=1}^{n} H(h_k) \, K\left(\frac{x-X_k}{h_k}\right) .$$

Si les v.a. $(X_i)_{i\in\mathbb{N}^*}$, sont indépendantes de même loi, et si la suite d'ouverts (U_i) , $i \in \mathbb{N}^*$ est définie par : $U_i =]x-h_i , x+h_i[$ alors, d'après la proposition 4.2. une version de $\bar{f}_n(x)$ conditionné par rapport à la statistique S , est donnée par

$$\bar{\bar{f}}_n(x) = \frac{1}{\sum_{i=1}^{n} h_i \, H(h_i)} \sum_{k=1}^{n} H(h_k) \left\{ K\left(\frac{x-X_k}{h_k}\right) 1_{U_k}(X_k) + m_k(P) \; 1_{U_k^c}(X_k) \right\}$$

où

$$m_k(P) = \frac{1}{P(U_k)} h_k \int_{]-1,1[^c} K(y) \, f_P(x-h_k y) \, dy \qquad \forall k \geq 1 .$$

On obtient ainsi un estimateur récursif, puisque :

$$\bar{\bar{f}}_{n+1}(x) = \frac{\sum_{i=1}^{n} h_i \, H(h_i)}{\sum_{i=1}^{n+1} h_i \, H(h_i)} \bar{\bar{f}}_n(x)$$

$$\frac{H(h_{n+1})}{\sum_{i=1}^{n+1} h_i \, H(h_i)} \left\{ K\left(\frac{x-X_{n+1}}{h_{n+1}}\right) 1_{U_{n+1}}(X_{n+1}) + m_{n+1}(P) \; 1_{U_{n+1}^c}(X_{n+1}) \right\}$$

On se propose dans ce qui suit d'étudier le risque quadratique asymptotique de $\bar{\bar{f}}_n(x)$ et de le comparer avec celui de $\bar{f}_n(x)$.

III.7.1. - Biais asymptotique.-

Proposition 4.18.- Si K est λ-intégrable

$$h_n \xrightarrow[n\to+\infty]{} 0 \quad \text{et} \quad \sum_{i=1}^{n} h_i\, H(h_i) \xrightarrow[n\to\infty]{} \infty$$

Alors $E_Q\, \bar{\bar{f}}_n(x) \xrightarrow[n\to\infty]{} f_Q(x) \int_{]-1,1[} K\, d\lambda + f_P(x) \int_{]-1,1[^c} K\, d\lambda$.

Preuve : Pour démontrer cette proposition et pour la suite, nous aurons besoin du lemme de Toeplitz ; dont un énoncé partiel est le suivant :

Lemme de Toeplitz.- Soit (a_i) , $i \in \mathbb{N}^*$ une suite de réels positifs ou nuls.

Si $b_n = \sum_{i=1}^{n} a_i \xrightarrow[n\to\infty]{} \infty$ alors on a pour toute suite $(u_n)_{n\geq 1}$ convergeant vers u $\frac{1}{b_n} \sum_{i=1}^{n} a_i\, u_i \xrightarrow[n\to\infty]{} u$.

Pour un énoncé complet cf. M. LOEVE (1963) page 238.

Par définition de $\bar{\bar{f}}_n(x)$ et des X_i $i \in \overline{1,n}$, nous avons :

$$E_Q\, \bar{\bar{f}}_n(x) = \frac{1}{\sum_{i=1}^{n} h_i\, H(h_i)} \sum_{j=1}^{n} h_j\, H(h_j) \left\{ \int_{]-1,1[} K(y)\, f_Q(x-h_j y)\, dy + \frac{Q(U_j^c)}{P(U_j^c)} \int_{]-1,1[^c} K(y)\, f_P(x-h_j y)\, dy \right\}.$$

Du lemme de Toeplitz il résulte qu'une condition suffisante, pour avoir la convergence du biais, est que :

$$\int_{]-1,1[} K(y)\, f_Q(x-h_j y)\, dy + \frac{Q(U_j^c)}{P(U_j^c)} \int_{]-1,1[^c} K(y)\, f_P(x-h_j y)\, dy$$

converge vers

$$f_Q(x) \int_{]-1,1[} K(y)\, dy + f_P(x) \int_{]-1,1[^c} K(y)\, dy \quad \text{lorsque} \quad j \to +\infty \;;$$

or cette condition est vérifiée grâce aux hypothèses imposées à $(h_n)_n$, f_Q , f_P et K .

III. 7. 2. - Variance asymptotique.-

On suppose que les conditions de la proposition 4.18 sont vérifiées

Proposition 4.19.- Si K est bornée

$$\frac{\sum_{i=1}^{n} h_i\, H^2(h_i)}{\left(\sum_{i=1}^{n} h_i\, H(h_i)\right)^2} \xrightarrow[n\to\infty]{} 0$$

et

$$\sum_{i=1}^{n} h_i\, H^2(h_i) \xrightarrow[n\to+\infty]{} +\infty \quad . \quad \text{Alors}$$

$$V\, \overset{=}{f}_n(x) \underset{n\to\infty}{\sim} f_Q(x) \int_{]-1,1[} K^2\, d\lambda \left\{ \frac{\sum_{i=1}^{n} h_i\, H^2(h_i)}{\left(\sum_{i=1}^{n} h_i\, H(h_i)\right)^2} \right\}$$

Preuve : Nous avons, d'après l'indépendance des X_i et l'expression de $\overset{=}{f}_n(x)$

$$V\, \overset{=}{f}_n(x) = \frac{1}{\left(\sum_{i=1}^{n} h_i\, H(h_i)\right)^2} \sum_{k=1}^{n} H^2(h_k)\; V\left(K\left(\frac{x-X_k}{h_k}\right) 1_{U_k}(X_k) + m_k(P)\; 1_{U_k^c}(X_k) \right)$$

d'où $\dfrac{\left(\sum\limits_{i=1}^{n} h_i\, H(h_i)\right)^2}{\sum\limits_{i=1}^{n} h_i\, H^2(h_i)} \cdot V(\overset{=}{f}_n(x)) =$

$$\frac{1}{\sum\limits_{i=1}^{n} h_i\, H^2(h_i)} \sum_{k=1}^{n} h_k^2\, H^2(h_k)\, Q(U_k^c) \left\{ \frac{1}{P(U_k^c)} \int_{]-1,1[^c} K(y)\, f_P(x-h_k y)\, dy \right\}^2$$

$$+ \frac{1}{\sum\limits_{i=1}^{n} h_i\, H^2(h_i)} \sum_{k=1}^{n} h_k\, H^2(h_k) \int_{]-1,1[} K^2(y)\, f_Q(x-h_k y)\, dy$$

$$- \frac{1}{\sum\limits_{i=1}^{n} h_i\, H^2(h_i)} \sum_{k=1}^{n} h_k^2\, H^2(h_k) \left\{ \int_{]-1,1[} K(y)\, f_Q(x-h_k y)\, dy + \frac{Q(U_k^c)}{P(U_k^c)} \cdot \int_{]-1,1[^c} K(y)\, f_P(x-h_k y)\, dy \right\}^2$$

posons $a_i = h_i\, H^2(h_i)$

et $$u_i = h_i\, Q(U_i^c) \left\{ \frac{1}{P(U_i^c)} \int_{]-1,1[^c} K(y)\, f_P(x-h_i y)\, dy \right\}^2$$

les conditions imposées à h_i , Q , P et à K entraînent que $u_i \xrightarrow[i\to\infty]{} 0$, par conséquent le lemme de Toeplitz permet d'obtenir la convergence vers 0 , du premier terme du membre de droite, de même en posant $a_i = h_i\, H^2(h_i)$ et

$$u_i = h_i \left\{ \int_{]-1,1[} K(y)\, f_Q(x-h_i y)\, dy + \frac{Q(U_i^c)}{P(U_i^c)} \int_{]-1,1[^c} K(y)\, f_P(x-h_i y)\, dy \right\}^2$$

on obtient la convergence vers 0 du troisième terme.

Reste donc à montrer que

$$\frac{1}{\sum_{i=1}^{n} h_i\, H^2(h_i)} \sum_{k=1}^{n} h_k\, H^2(h_k) \int_{]-1,1[} K^2(y)\, f_Q(x-h_k y)\, dy \xrightarrow[n\to\infty]{} f_Q(x) \int_{]-1,1[} K^2\, d\lambda .$$

En effet on peut écrire $\displaystyle \frac{1}{\sum_{i=1}^{n} h_i\, H^2(h_i)} \sum_{k=1}^{n} h_k\, H^2(h_k) \int_{]-1,1[} K^2(y)\, f_Q(x-h_k y)\, dy$

$$= f_Q(x) \int_{]-1,1[} K^2(y)\, dy + \frac{1}{\sum_{i=1}^{n} h_i\, H^2(h_i)} \sum_{k=1}^{n} h_k\, H^2(h_k) \int_{]-1,1[} K^2(y)\, (f_Q(x-h_k y)-f_Q(x))\, dy$$

or f_Q étant continue bornée et K^2 λ-intégrable, une application du théorème de la convergence dominée donne : $\displaystyle \int_{]-1,1[} K^2(y)\, (f_Q(x-h_k y)-f_Q(x))\, dy \xrightarrow[h_k\to 0]{} 0$

une application du lemme de Toeplitz aux suites $a_i = h_i\, H^2(h_i)$ et $u_i = \displaystyle\int_{]-1,1[} K^2(y)\, (f_Q(x-h_k y) - f_Q(x))\, dy$, permet de conclure.

Remarques

* Les conditions $\displaystyle\int_{]-1,1[^c} K\, d\lambda = 0$ et $\displaystyle\int K\, d\lambda = 1$ entraînent que $\bar{\bar{f}}_n(x)$ est asymptotiquement sans biais.

** Si on définit les variances asymptotiques de $\bar{f}_n$ (resp.) $\bar{\bar{f}}_n$ comme étant égales à

$$\frac{1}{(\sum_{i=1}^{n} h_i\, H(h_i))^2} \sum_{i=1}^{n} h_i\, H^2(h_i) \;.\; f_Q(x) \int_{\mathbb{R}} K^2\, d\lambda$$

resp. $$\frac{1}{(\sum_{i=1}^{n} h_i\, H(h_i))^2} \sum_{i=1}^{n} h_i\, H^2(h_i) \quad f_Q(x) \int_{]-1,1[} K^2\, d\lambda$$

Alors le choix de la variance asymptotique comme critère de comparaison des deux estimateurs $\bar{f}_n$ et $\bar{\bar{f}}_n$ donnera $\bar{\bar{f}}_n$ préférable à $\bar{f}_n$. Par ailleurs nous savons que pour une certaine classe de lois $\bar{\bar{f}}_n$ est préférable à $\bar{f}_n$ au sens du risque quadratique.

REFERENCES

J. AKONOM (1978) - *Estimation de la densité de la partie continue d'une probabilité mixte.*
Thèse de 3ème cycle. Université de Lille I.

A. BERLINET (1984-a) - *Sur quelques problèmes d'estimation fonctionnelle et de statistique des processus.*
Thèse de Doctorat ès Sciences Mathématiques. Université de Lille I.

A. BERLINET (1984-b) - *Propriétés locales d'un paramètre. Application à l'estimation.*
C.R.A.S. *t. 298, série I, p. 345-348.*

P.J. BICKEL, M. ROSENBLATT (1973) - *On some global measures of the deviations of density function estimates.*
Ann. Statist. 1, p. 1071-1095.

H.J. BIERENS (1983) - *Sample moments integrating normal kernel estimators of multivariate density and regression functions.*
Sankhya - The Indian Journal of Statistics. Vol. 45. Series B pt 2, pp. 160-192.

P. BILLINGSLEY (1968) - *Convergence of probability measures.*
John Wiley and Sons.

J. BLEUEZ, D. BOSQ (1976) - *Conditions nécessaires et suffisantes de convergence pour une classe d'estimateurs de la densité.*
C.R.A.S. Paris 282 Série A, 63-66.

D. BOSQ, J.P. LECOUTRE (1985) - *Théorie de l'estimation fonctionnelle*
Livre à paraître.

A.W. BOWMAN (1980) - *A note on consistency of the kernel method for the analysis of categorical data.*
Biometrika 67, p. 682-684.

A.W. BOWMAN (1982) - *A comparative study of some kernel - based non-parametric density estimators.*
Manchester - Sheffield School of Probability and Statistics Research Report n° 84/AWB/1.

A.W. BOWMAN (1984) - *An alternative method of cross - validation for the smoothing of density estimates.*
Biometrika 71 p. 353-360.

L. BREIMAN, W. MEISEL, E. PURCELL (1977) - *Variable kernel estimates of multivariate densities and their calibration.*
Technometrics 19, p. 443-507.

CHAI GENXIANG (1984) - *Consistent estimation of random window-width kernel of density function.*
Scientia Sinica (Series A) vol. XXVII n° 11 p. 1155-1163.

Y.S. CHOW, S. GEMAN, L.D. WU (1983) - *Consistent cross - validated density estimation.*
Ann. *Statist.* 11 n° 1, p. 25-38.

G. COLLOMB (1976) - *Estimation non paramétrique de la régression par la méthode du noyau.*
Thèse Docteur Ingénieur, Université Paul Sabatier de Toulouse.

P. DEHEUVELS (1974) - *Estimation séquentielle de la densité.*
Thèse présentée à l'Université de Paris VI.

P. DEHEUVELS (1977-a) - *Estimation non paramétrique de la densité par histogrammes généralisés.*
Revue de *Statistique* Appliquée, *Vol.* 25, p. 5-42.

P. DEHEUVELS (1977-b) - *Estimation non paramétrique de la densité par histogrammes généralisés (II).*
Pub. *Inst. Stat. Univ. Paris,* XXII, 1-23.

P. DEHEUVELS *et* P. HOMINAL (1980) - *Estimation automatique de la densité.*
Revue de *Statistique* Appliquée, *Vol.* 28, p. 25-52.

L.P. DEVROYE *et* T.J. WAGNER (1980) - *The strong uniform convergence of kernel density estimates.*
Multivariate Analysis, V, *North-Holland.*
Publ. Amsterdam p. 59-77.

L. DEVROYE, L. GYÖRFI (1984) - *Nonparametric density estimation : the* L_1.
View - Wiley, New York.

L. DEVROYE, C.S. PENROD (1984) - *Distribution Free lower bounds in density estimation.*
Ann. *Statist* 12, p. 1250-1262.

R.P.W. DUIN (1976) - *On the choice of smoothing parameters for Parzen estimators of probability density functions.*
I.E.E.E. *Trans. Computers* C. 25 p. 1175-1179.

A. DVORETZKY, J. KIEFER *et* T. WOLFOWITZ (1956) - *Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator.*
Ann. Math. Statist. 27, 642-669.

P. HALL (1981) - *Laws of the iterated logarithm for nonparametric density estimators.*
Z. f. W. 56, 47-61.

P. HALL (1983) - *Large sample optimality of least squares cross-validation in density estimation.*
Ann. Statist. 11 p. 1156-1174.

J.D.F. HABBEMA, J. HERMANS, K. VANDENBROEK (1974) - *A stepwise discriminaiton analysis program using density estimation.*
In Compstat 1974. Ed. G. Bruckman p. 109-110.
Vienna : Physica Verlag.

P.J. HUBER (1981) - *Robust statistics.*
Wiley.

D.O. LOFTSGAARDEN et C.P. QUESENBERRY (1965) - *A nonparametric estimate of a multivariate density function.* Ann. Math. Statist. 36 p. 1049-1051.

M. LOEVE (1963) - *Probability theory.* D. Van Nostrand, Princeton N.J.

E.A. NADARAYA (1965) - *On non parametric estimates of density functions and regression curves.* Theory of probability and its applications, 10, p. 186-90.

E. PARZEN (1962) - *On estimation of probability density function and mode.* Ann. Math. Stat. t. 33 p. 1065-1076.

B.L.S. PRAKASA RAO (1983) - *Nonparametric functional estimation.* Academic Press. New York.

M. ROSENBLATT (1956) - *Remarks on some non-parametric estimates of a density function.* Ann. Math. Statist. 27 p. 832-837.

M. RUDEMO (1982) - *Empirical choice of histograms and kernel density estimators.* Scand. J. Statist. 9 p. 65-78.

E.F. SCHUSTER, G.C. GREGORY (1981) - *On the nonconsistency of maximum likelihood nonparametric density estimators.* In : Computer Science and Statistics : Proceeding of the 13th Symposium on the interface, Ed. W.F. Eddy p. 295-298. Springer-Verlag, New York.

L. SCHWARTZ (1965) - *Méthodes mathématiques pour les sciences physiques.* Hermann.

B.W. SILVERMAN (1978) - *Choosing the window width when estimating a density.* Biometrika 65, p. 1-11.

M. STONE (1974) - *Cross - validatory choice and assessment of statistical predictions.* J.R. Statist. Soc. B. 36 p. 111-47.

M. STONE (1974) - *Cross - validation and multinomial prediction.* Biometrika 61 p. 509-15.

M. STONE (1977) - *An asymptotic equivalence of choice of model by cross - validation and Akaike's criterion.* J.R. Statist. Soc. B. 39, p. 44-47.

M. STONE (1977) - *Asymptotics for and against cross - validation.* Biometrika 64, p. 29-35.

C.J. STONE (1984) - *An asymptotically optimal window selection rule for kernel density estimates.* Ann. Statist. V. 12 n° 4, p. 1285-1297.

R. TAPIA, J. THOMPSON (1978) - *Non parametric probability density estimation. The John University Press. Baltimore and London.*

J. VAN RYZIN (1969) - *On strong consistency of density estimates, Ann. Math. Statist. 40, 1765-1772.*

G.S. WATSON, M.R. LEADBETTER (1963) - *On the estimation of the probability density. Ann. Math. Statist., 34 p. 480-491.*

YAMATO (1971) - *Sequential estimation of a continuous probability density function and mode. Bull. Math. Statist. (14) p. 1-12.*

R E S U M E

Ce travail se compose de quatre parties.

Dans un premier temps nous dressons le bilan 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 et nous proposons une procédure d'estimation en trois étapes.

Dans la seconde partie nous montrons la convergence uniforme presque sûre d'un estimateur à noyau, construit de telle façon que ses moments d'ordre 1 et 2 coïncident avec les moments empiriques de la loi à estimer.

La troisième partie est consacrée à l'étude de la convergence uniforme presque sûre d'une classe d'estimateurs, qui recouvre la plupart des estimateurs à noyaux connus.

La dernière partie porte sur la correction locale de l'estimateur à noyau au moyen du conditionnement par rapport à une statistique localement exhaustive ; le risque quadratique ainsi que la convergence suivant différents modes y sont étudiés. On envisage pour terminer le cas séquentiel.

M O T S   C L E S

ESTIMATEURS A NOYAUX - FENETRE OPTIMALE - CONVERGENCE UNIFORME PRESQUE COMPLETE - EXHAUSTIVITE LOCALE - ESTIMATEURS SEQUENTIELS.