N° d'Ordre 672

UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE

THESE

Présentée à

l'UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE

pour obtenir

LE GRADE DE DOCTEUR DE 3ème CYCLE

par

Jacques AKONOM

Estimation de la densité de la partie continue d'une probabilité mixte

Membres du Jury : MM. POUZET, Président
BOSQ, Rapporteur
Mle MARQUET,
M. DELPORTE.

Soutenue le 23 Février 1978

Nous remercions vivement Monsieur le Professeur POUZET qui nous a fait l'honneur d'accepter de présider le jury de cette thèse.

Nous remercions également Mademoiselle le Professeur MARQUET et Monsieur le Professeur DELPORTE qui ont bien voulu faire partie de la commission d'examen.

Nous tenons à exprimer notre reconnaissance à Monsieur le Professeur BOSQ qui nous a proposé le sujet de cette thèse, qui nous a permis d'utiliser certains de ses résultats, et surtout qui a su nous initier à la recherche, en nous prodiguant conseils et encouragements.

Nous remercions aussi l'équipe des Probabilistes et Statisticiens de l'U.E.R. de Mathématiques de LILLE I qui nous a aidé de ses remarques et entouré de sa sympathie.

Nous remercions enfin Madame BERAT et Madame LENGAIGNE pour tout le soin et la diligence dont elles ont fait preuve quant à la réalisation de cet ouvrage.

P L A N.

INTRODUCTION.

L'estimation de la densité s'est faite par diverses méthodes, ROSENBLATT a défini la méthode du noyau en 1956 ([1]), ÇENCOV ([2]) a introduit celle des fonctions orthogonales en 1962, depuis WAHBA ([3]) a construit des estimateurs utilisant les fonctions splines.

Des conditions de convergence de l'estimateur obtenu par la méthode des fonctions orthogonales ont été données par BOSQ ([4]) et cette étude a été poursuivie par BLEUEZ ([5]). BLEUEZ définit une classe d'estimateurs de la densité qui contient ceux obtenus par les deux premières méthodes et précise des conditions nécessaires et suffisantes de convergence de ces estimateurs.

Dans ces travaux, la famille des probabilités est dominée par une mesure μ, et les densités de probabilité sont prises par rapport à μ. Nous étudierons ici le cas de mesures définies sur la droite réelle, somme d'une mesure absolument continue par rapport à la mesure de LEBESGUE sur $\mathbb{R}$ et d'une mesure dont le support est fini ou dénombrable (combinaison de mesures de DIRAC). Elles seront appelées respectivement partie continue et partie discrète.

Dans une première partie, on étudiera le cas où le support de la partie discrète est connu, l'ensemble des mesures est alors dominé. Dans ce cas, les résultats établis par BOSQ et BLEUEZ ([5]) s'appliquent sans modification et on rappellera brièvement ceux qui concernent la convergence de l'estimateur de la densité de la partie continue par rapport à la mesure de LEBESGUE et les conditions d'existence d'un estimateur sans biais.

Dans la seconde partie, le support de la partie discrète est inconnu et la famille des probabilités non dominée. On construira alors un estimateur de la densité de la partie continue et on donnera des conditions nécessaires et suffisantes de convergence suivant divers modes.

Dans la troisième partie, plus brève, on appliquera les résultats précédents aux mesures définies sur $[-\pi, \pi]$, dont la densité de la partie continue appartient à $\mathbb{L}^2(-\pi, \pi)$. La densité étant donnée par son développement en série de Fourier.

La quatrième étendra les résultats de la seconde lorsque les variables aléatoires ne sont plus indépendantes, mais forment un processus mélangeant. La définition adoptée pour la mélangeance est due à BILLINGSLEY ([11]). Nous utiliserons également des publications internes de l'U.E.R. de Math. de Lille ([13]), ([14]).

NOTATIONS.-

Soit $(\Omega, \mathcal{A}, \Pi)$ un espace probabilisable, E un borélien de $\mathbb{R}$ de mesure strictement positive - E sera le plus souvent un intervalle de $\mathbb{R}$ borné ou infini - $\mathcal{B}$ la tribu borélienne de E, λ la restriction à E de la mesure de Lebesgue et X une variable aléatoire définie sur $(\Omega, \mathcal{A}, \Pi)$ à valeurs dans $(E, \mathcal{B})$.

On suppose que la loi de X, P_X peut se décomposer comme somme d'une mesure absolument continue par rapport à la mesure de Lebesgue et d'une mesure dont le support est l'ensemble fini ou dénombrable $\Delta = \{d_1, d_2, \ldots, d_m, \ldots\}$ on les appellera partie continue et partie discrète de P_X ; l'ensemble Δ sera encore appelé ensemble des masses.

On sait que si une telle décomposition existe, elle est unique. D'autre part, d'après le théorème de Radon-Nikodym - par exemple ([6]), p. 128 - il existe une fonction f à valeurs p.s. finies et positives et une suite a_m de nombres réels positifs telles que :

(1) $$dP_X = f\, d\lambda + \sum_{d_m \in \Delta} a_m\, \delta_{d_m}$$

où δ_{d_m} désigne la mesure de Dirac au point d_m.

En outre la classe d'équivalence de f dans $L^1(E, \lambda)$ est unique, et la relation

(2) $$\int_E f\, d\lambda + \sum_{d_m \in \Delta} a_m = 1$$

est vérifiée.

Dans toute la suite, on écrira $\sum_m a_m\, \delta_{d_m}$ sans préciser l'ensemble d'indices décrit par m.

Soit $\mathcal{D}_\lambda$ l'ensemble des densités (par rapport à la mesure de Lebesgue) des mesures de masse totale inférieure à 1 et $\mathcal{D}$ un sous-ensemble de $\mathcal{D}_\lambda$. Ces mesures peuvent être considérées comme la partie continue des probabilités P_X.

On notera $\mathfrak{S}_n$ - ou $\mathfrak{S}$ s'il n'y a pas d'ambiguité - un échantillon $(X_1,X_2,\ldots,X_n)$ de taille n de la variable aléatoire X. Sachant que la densité que l'on veut estimer appartient à $\mathcal{D}$, on se propose de trouver des estimateurs de cette densité qui convergent, suivant des modes que l'on précisera plus loin, lorsque la taille de l'échantillon augmente indéfiniment.

CHAPITRE I

ETUDE DE LA DENSITE DE LA PARTIE CONTINUE DANS LE CAS OU LE SUPPORT DES MASSES EST CONNU.

1.1. - CONSTRUCTION D'UN ESTIMATEUR DE LA DENSITE.-

Si $\Delta = \{d_1, d_2, \ldots, d_m, \ldots\}$ support de la partie discrète est connu, la mesure $\mu = \lambda + \sum_m \delta_{d_m}$ est une mesure σ-finie qui domine l'ensemble des mesures P_X.

En effet, E est la réunion des ensembles de mesure finie $(E-\Delta) \cap [-n, n]$ et $\{d_m\}$ lorsque m et n décrivent $\mathbb{N}$.
La méthode des fonctions orthogonales s'applique sans modification. $L^2(E,\lambda)$ est un espace de Hilbert séparable.

Soit $(e_i)_i$ une base orthonormale de $L^2(E,\lambda)$ et $(e_i^*)_i$ une famille de versions des e_i, alors la réunion des familles $(e_i^* \, 1_{E-\Delta})_i$ et $(1_{d_m})_m$ sont des versions d'une base orthonormale de $L^2(E,\mu)$.

(Toute fonction $f \in L^2(\mu)$ s'écrit $f = f.1_{E-\Delta} + \sum_m f(d_m).1_{d_m}$ où $f.1_{E-\Delta} \in L^2(\lambda)$ et $\sum_m f^2(d_m)$ converge).

Une version $\hat{f}_n^*$ de l'estimateur $\hat{f}_n$ de $f = \frac{dP}{d\mu}$ sera définie par la relation :

$$\hat{f}_n^*(t;X_1,X_2,\ldots,X_n) = \sum_{i=1}^{r(n)} \frac{1}{n} \left[e_i^*(X_1)\ 1_{E-\Delta}(X_1) + \ldots + e_i^*(X_n)\ 1_{E-\Delta}(X_n)\right] e_i^*(t)\ 1_{E-\Delta}(t)$$

$$+ \sum_m \frac{1}{n} \left[1_{d_m}(X_1) + \ldots + 1_{d_m}(X_n)\right] 1_{d_m}(t)$$

Note : CAMBANIS (7) a étudié des bases orthonormales dans un espace $L^2(E,\mathcal{B}_E,\mu)$ où E est un intervalle de $\mathbb{R}$, $\mathcal{B}_E$ la tribu borélienne et μ une mesure finie.

Et une version $\hat{f}_n^*$ de l'estimateur de la densité de la partie continue sera définie par :

$$\hat{f}_n^*(t;X_1,\ldots,X_n) = \sum_{i=1}^{r(n)} \frac{1}{n} \left[e_i^*(X_1)\ 1_{E-\Delta}(X_1) + \ldots + e_i^*(X_n)\ 1_{E-\Delta}(X_n)\right]\ e_i^*(t)\ 1_{E-\Delta}(t)$$

Comme les fonctions $e_i^*.1_{E-\Delta}$ et e_i^* sont égales λ-p.p. , on prendra plus facilement comme version

$$(3) \qquad \hat{f}_n(t) = \sum_{i=1}^{r(n)} \frac{1}{n} \left[e_i(X_1)\ 1_{E-\Delta}(X_1) + \ldots + e_i(X_n)\ 1_{E-\Delta}(X_n)\right]\ e_i(t)$$

En convenant de confondre les fonctions et leurs représentants dans $L^2(E,\lambda)$ chaque fois qu'il n'y a pas ambiguïté.

BLEUEZ ([5]) a montré que les classes des estimateurs obtenus par la méthode du noyau et par la méthode des fonctions orthogonales étaient contenues dans celle obtenues par la méthode suivante :

Soit I une partie infinie, non bornée de $\mathbb{R}$; $K_r(x,t)$, $r \in I$, une famille de fonctions réelles mesurables définies sur $E \times E$, indexée par I. On lui associe une famille d'estimateurs de la densité par :

$$(4) \qquad \hat{f}_n(t) = \frac{1}{n} \sum_{j=1}^{n} K_{r(n)}\ (X_j,t)$$

où $r(n) \in I$ et $\lim_{n\to\infty} r(n) = \infty$.

Remarque : L'estimateur obtenu par la méthode des fonctions orthogonales correspond au noyau

$$(4') \qquad K_{r(n)}(x,t) = \sum_{i=1}^{r(n)} e_i(x)\ e_i(t)$$

Cette méthode s'applique à l'estimation de la densité de la partie continue des mesures P_X.

Proposition 1.- Soit $\mathcal{D}_1$ un ensemble de densités de mesures de probabilité absolument continues par rapport à la mesure de Lebesgue et

$\mathcal{D} = \{k\, f \mid f \in \mathcal{D}_1 \;,\; 0 < k < 1\}$ l'ensemble des densités de la partie continue des probabilités mixtes (on dira que $\mathcal{D}$ est associé à $\mathcal{D}_1$).

Soit $K_r(x,t)$ une famille de fonctions réelles mesurables-Lebesgue.

Si pour toute fonction f de $\mathcal{D}_1$, l'estimateur de la densité $\hat{f}$ associé converge suivant un certain mode lorsque $n \to \infty$, $r(n)$ vérifiant une condition (C).

Si pour tout $k > 0$, $r([kn])$ vérifie la même condition (C) alors

$$\hat{g}_n(t;X_1,\dots,X_n) = \sum_{j=1}^{n} \frac{1}{n} K_{r(n)}(X_j,t)\; 1_{E-\Delta}(X_j) \tag{5}$$

est un estimateur convergent suivant le même mode, pour toute densité g de $\mathcal{D}$.

Démonstration : Si, pour tout $k > 0$, $r([kn])$ vérifie la condition (C) alors $\sum_{j=1}^{n} \frac{1}{n} K_{r([kn])}(X_j,t)\; 1_{E-\Delta}(X_j)$ converge vers $f(t)$, $\forall f \in \mathcal{D}_1$.

Soit P une loi de probabilité dont la densité g de la partie continue appartient à $\mathcal{D}$ et $\vartheta(n)$ le nombre d'éléments de l'échantillon qui tombent dans $E-\Delta$. On a $\lim \frac{\vartheta(n)}{n} = \int_E g\, d\lambda$ p.s. et

$$\lim_{n\to\infty} \sum_{j=1}^{n} \frac{1}{\vartheta(n)} K_{r(n)}(X_j,t)\; 1_{E-\Delta}(X_j) = \lim_{n} \sum_{j=1}^{n} \frac{1}{\vartheta(n)} K_{r(\vartheta(n))}(X_j,t)\; 1_{E-\Delta}(X_j)$$

converge vers la densité de la loi conditionnée par $X \in E-\Delta$, c'est-à-dire vers $\dfrac{g}{\int_E g\, d\lambda}$; d'où le résultat.

1.2. - CONDITIONS NECESSAIRES ET SUFFISANTES DE CONVERGENCE DE L'ESTIMATEUR DE LA DENSITE.-

Tous les résultats établis pour une famille de probabilités dont les densités (par rapport à la mesure de Lebesgue) appartiennent à $\mathcal{D}_1$ sont donc valables pour la famille de probabilités mixtes associée. On se contentera de les énoncer pour les familles de probabilités absolument continues par rapport à la mesure de Lebesgue. Les théorèmes qui suivent, donnés sans démonstration, sont dus à BLEUEZ (5), p. 41-71.

a - E est l'intervalle fini $[a, b]$.

Soit (e_i) la base orthonormale de $L^2(E,\lambda)$ formée des fonctions trigonométriques :

$$e_0(t) = \frac{1}{\sqrt{b-a}} \quad ;$$

$$e_{2i-1}(t) = \sqrt{\frac{2}{b-a}} \sin 2i \frac{t-a}{b-a} \quad ; \quad e_{2i}(t) = \sqrt{\frac{2}{b-a}} \cos 2i \frac{t-a}{b-a} \quad ; \quad i \geq 1$$

et $\hat{f}$ l'estimateur associé au noyau $\sum_{i=1}^{r(n)} e_i(x)\, e_i(t)$.

Proposition 2.- Soit $\mathcal{D}$ l'ensemble des densités définies sur $[a,b]$ et possédant une série de Fourier convergeant en tout point t de $[a,b]$. Alors les propriétés suivantes sont équivalentes :

a) $\lim\limits_{n\to\infty} \frac{r(n)}{n} = 0$

b) $\forall f \in \mathcal{D}, \ \forall t \in [a,b] \quad \hat{f}_n(t) \xrightarrow{P} f(t)$

c) $\forall f \in \mathcal{D}, \ \forall t \in [a,b] \quad E(|\hat{f}_n(t) - f(t)|^2) \to 0$

d) $\forall f \in \mathcal{D}, \ E\left[\int_a^b (\hat{f}_n(t) - f_n(t))^2\, dt\right] \to 0$

Proposition 3.- Soit $\mathcal{D}$ l'ensemble des densités définies sur $[a,b]$ bornées et dont la série de Fourier converge uniformément sur $[a,b]$.
Alors les propriétés suivantes sont équivalentes :

a) $\lim\limits_{n\to\infty} \frac{r(n)}{n} \operatorname{Log} n = 0$

b) $\forall f \in \mathcal{D}, \quad \sup\limits_{t\in[a,b]} |\hat{f}_n(t) - f(t)| \xrightarrow{p.c_o.} 0$ (*)

c) $\forall f \in \mathcal{D}, \quad \sup\limits_{t\in[a,b]} |\hat{f}_n(t) - f(t)| \xrightarrow{P} 0$

d) $\forall f \in \mathcal{D} \quad E[\sup\limits_{t\in[a,b]} |\hat{f}_n(t) - f(t)|] \longrightarrow 0$

b - *E est la droite réelle.*

Les fonctions d'Hermite forment une base orthonormale de $L^2(\mathbb{R},\lambda)$.
Elles sont données par :

$$e_i(t) = (2^i\ i!\ i)^{-1/2}\ e^{t^2/2}\ \frac{d^i}{dt^i}\ (e^{-t^2}) \ ; \quad t \in \mathbb{R} \ ; \quad i \in \mathbb{N} .$$

L'estimateur $\hat{f}$ est obtenu par la méthode des fonctions orthogonales.

Proposition 4.- Soit $[a,b]$ un intervalle compact de R et $\mathcal{D}_{[a,b]}$ l'ensemble des densités définies sur $\mathbb{R}$, bornées et dont le développement en série de fonctions d'Hermite converge vers f uniformément sur $[a,b]$.
Alors une condition nécessaire et suffisante pour que :

$$\forall f \in \mathcal{D}_{[a,b]} \quad E[\sup_{t\in[a,b]} |\hat{f}_n(t) - f(t)|] \to 0 \qquad \text{est que l'on ait :}$$

$$\lim_{n\to\infty} \frac{\sqrt{r(n)}}{n} \operatorname{Log} n = 0$$

Proposition 5.- Soit $\mathcal{D}$ l'ensemble des densités définies et bornées sur $\mathbb{R}$ dont le développement en série d'Hermite converge uniformément vers

---

(*) On dit qu'une suite de variables aléatoires Y_n converge presque complètement vers une variable Y si
$\forall \varepsilon > 0 \ , \sum_n P(|Y_n - Y| > \varepsilon) < \infty$ et on le notera $Y_n \xrightarrow{p.c_o.} Y$

$f(t)$. Alors une condition suffisante pour que l'on ait :

$$\sup_{t\in\mathbb{R}} |\hat{f}_n(t) - f(t)| \xrightarrow{p.c_o.} 0$$

est

$$\lim_{n\to\infty} \frac{r(n)}{n} \operatorname{Log} n = 0 \quad .$$

1.3. - ESTIMATION SANS BIAIS.-

On sait que $\sum_m \frac{1}{n} [1_{d_m}(X_1) + \ldots + 1_{d_m}(X_n)] \, 1_{d_m}(t)$ est un estimateur sans biais de la partie discrète et qu'il est optimal dans la famille des estimateurs sans biais.

Il n'existe pas en général d'estimateurs sans biais de la partie continue. Cependant, on a le résultat suivant :

Proposition 6.- Si $\mathcal{D}$ ensemble des densités de la partie continue des probabilités définies sur E est un sous-ensemble convexe de $L^2(E,\lambda)$, s'il existe une densité f_o telle que $\inf_t f_o(t) > 0$ et si les densités de $\mathcal{D}$ sont bornées, une condition nécessaire et suffisante pour qu'il existe un estimateur sans biais, de carré intégrable, de la densité est que l'espace vectoriel engendré par $\mathcal{D}$ soit de dimension finie.

La démonstration reprend celle de ([8]) cor. 4, p. 15 et utilise le fait que la partie discrète est estimable sans biais.

CHAPITRE II

ETUDE DE LA DENSITE DE LA PARTIE CONTINUE DANS LE CAS OU LE SUPPORT DES MASSES EST INCONNU.

2.1. - CONSTRUCTION D'UN ESTIMATEUR DE LA DENSITE DE LA PARTIE CONTINUE.-

a - *Choix d'un estimateur de* Δ.

Dans l'échantillon, les éléments peuvent appartenir à Δ ou avoir été tirés suivant une loi "continue". Il faut un critère permettant de répartir les points de l'échantillon en deux groupes.

On a les résultats suivants :

Lemme 1.- Dans l'échantillon, si deux variables aléatoires différentes prennent la même valeur d, le point d appartient presque sûrement à l'ensemble des masses.

Démonstration : Si les variables X_i et X_j sont indépendantes, $P(X_i = X_j , X_i \in E - \Delta) = 0$. En effet la diagonale de $E-\Delta$ est négligeable pour la mesure de Lebesgue.

Lemme 2.- Si la distribution ne comporte pas que des masses, la proportion de masses parmi l'ensemble des points de l'échantillon sortis une seule fois tend vers 0 presque sûrement.

Démonstration : Soit V_n le nombre de points sortis une fois et W_n le nombre de masses (éléments de Δ) ne figurant qu'une fois dans l'échantillon.

On va établir que $\dfrac{W_n}{V_n}$ converge presque complètement vers 0.

Une loi P étant donnée, on peut ranger les masses $\{d_1, d_2, \ldots, d_m, \ldots\}$ et supposer que $\forall m$, $P(\{d_m\}) = a_m \geq a_{m+1}$.

Soit $N_{\varepsilon'}$, le premier entier tel que

$$\sum_{N_{\varepsilon'}+1} a_m < \varepsilon' \qquad 0 < \varepsilon' < 1 .$$

Les masses sorties peuvent être l'une des $N_{\varepsilon'}$ premières ou appartenir à l'ensemble $\Delta - \{d_1, \ldots, d_{N_{\varepsilon'}}\}$. Or la probabilité pour qu'il y ait plus de $2n\,\varepsilon'$ éléments dans l'ensemble $\Delta - \{d_1, \ldots, d_{N_{\varepsilon'}}\}$ est majoré par $e^{-n\varepsilon'}$ (inégalité du type de celle de Hoeffding ([9]) appliquée à une loi binomiale $\mathcal{B}(n, \varepsilon')$) d'où :

(6) $$P(W_n > N_{\varepsilon'} + 2\,n\varepsilon') < e^{-n\varepsilon'} \qquad \text{pour } n \text{ assez grand}$$

D'autre part, $P(E-\Delta) = 1 - \sum_m a_m$ et $V_n - W_n$ est le nombre de points de $E-\Delta$ dans l'échantillon. $V_n - W_n$ suit une loi binomiale $\mathcal{B}(n, 1 - \sum_m a_m)$ et

(6') $$\frac{V_n - W_n}{n} \xrightarrow{p.c_o.} (1 - \sum_m a_m)$$

$P(V_n - W_n < \frac{n}{2}(1 - \Sigma\, a_m))$ est le terme général d'une série convergente.

Or $$P\left(\frac{W_n}{V_n} > 2\,\frac{N_{\varepsilon'} + 2n\,\varepsilon'}{n(1 - \Sigma\, a_m)}\right) \leq P(W_n > N_{\varepsilon'} + 2n\,\varepsilon') + P(V_n < \frac{n}{2}(1 - \Sigma\, a_m))$$

En choisissant n suffisamment grand pour que $\frac{N_{\varepsilon'}}{n} < \varepsilon'$

et en posant $$\varepsilon = \frac{6}{1 - \sum_m a_m}\,\varepsilon'$$

on établit que

$$P\left(\frac{W_n}{V_n} > \varepsilon\right) \quad \text{est le terme général d'une série convergente.}$$

Les lemmes 1 et 2 fournissent un estimateur de Δ

$$\hat{\Delta}_n = \{X_j \mid \exists k \neq j \ ; \ X_k = X_j\}$$

Comparaison des critères.

Un critère est un estimateur de Δ, soit $\hat{\Delta}'_n$. On supposera toujours que $\hat{\Delta}'_n$ est symétrique par rapport aux observations et que $\hat{\Delta}'_n \subset \{X_1,\dots,X_n\}$. $\hat{\Delta}_n$ désignera le critère précédent $\hat{\Delta}_n = \{X_i \mid \exists\ k \neq i\ ,\ X_i = X_k\}$ qui sera utilisé dans la suite.

α) On compare les critères avec $E|1_\Delta(X_1) - 1_{\hat{\Delta}'_n}(X_1)|$ puisque $E(\sum_i |1_\Delta(X_i) - 1_{\hat{\Delta}'_n}(X_i)|)$ représente l'espérance du nombre d'erreurs commises. Un critère n'acceptant pas les "points doubles" comme masses ne peut être préférable à $\hat{\Delta}_n$.

En effet, s'il existe $(x_1,x_2,\dots,x_n)$ avec $x_i = x_j$ pour $i \neq j$ et $x_i \notin \hat{\Delta}'_n$ alors, en prenant une probabilité P telle que $\Delta = \{x_i\}$ $(P(x_i) = p)$

$$\begin{aligned} E|1_\Delta(X_1) - 1_{\hat{\Delta}_n}(X_1)| &= P(X_1 \in \Delta,\ X_1 \notin \hat{\Delta}_n) + P(X_1 \notin \Delta,\ X_1 \in \hat{\Delta}_n) \\ &= p(1-p)^{n-1} \end{aligned}$$

$$\begin{aligned} E|1_\Delta(X_1) - 1_{\hat{\Delta}'_n}(X_1)| &= P(X_1 \in \Delta\ ,\ X_1 \notin \hat{\Delta}'_n) + P(X_1 \notin \Delta,\ X_1 \in \hat{\Delta}'_n) \\ &\geq P(X_1 = x_i) = p \end{aligned}$$

On en déduit que tout critère préférable à $\hat{\Delta}_n$ doit vérifier $\hat{\Delta}'_n \supset \hat{\Delta}_n$.

β) Il existe des critères préférables à $\hat{\Delta}_n$.

Si D est un ensemble de mesure de Lebesgue nulle, le critère défini par

$$\hat{\Delta}'_n = \hat{\Delta}_n \cup \{X_i \mid X_i \in D\}$$

est préférable.

Si $\Delta \cap D = \emptyset$, $\hat{\Delta}'_n = \hat{\Delta}_n$ p.s.

sinon $P(X_1 \notin \Delta, X_1 \in \hat{\Delta}'_n) = 0$ et $P(X_1 \in \Delta, X_1 \in \hat{\Delta}'_n) \leqslant P(X_1 \in \Delta, X_1 \in \hat{\Delta}_n)$

Si D est de mesure strictement positive, le critère cesse d'être préférable. Il suffit de prendre $P = \frac{1}{2\lambda(D)} \cdot 1_D \cdot \lambda + \frac{1}{2}\delta_d$.

γ) Si l'on adopte la fonction de risque $\mathcal{R}_1(P, \hat{\Delta}'_n) = P(X_1 \notin \Delta, X_1 \in \hat{\Delta}'_n)$ alors $\hat{\Delta}_n$ est admissible $(\mathcal{R}_1(P, \hat{\Delta}_n) = 0)$.

L'autre erreur définie par $\mathcal{R}_2(P, \hat{\Delta}'_n) = P(X_1 \in \Delta, X_1 \notin \hat{\Delta}'_n)$ n'est pas minimisée par $\hat{\Delta}_n$. Cependant le lemme 2 a montré que

$$\mathcal{R}_2(P, \hat{\Delta}_n) \longrightarrow 0 \text{ p.s.}$$

En fait le choix de l'estimateur de Δ dépend du problème étudié et dans l'étude de la densité de la partie continue, il est préférable de conserver quelques points parasites (des masses), que d'enlever des points appartenant à la partie continue.

b - *Estimation sans biais.*

Proposition 7.- Si le support des masses est inconnu, si l'ensemble $\mathcal{P}$ des mesures est convexe, si $\mathcal{P}$ contient les mesures de Dirac et une mesure absolument continue par rapport à la mesure de Lebesgue, alors il n'existe aucun estimateur sans biais de la densité de la partie continue.

Démonstration : Soit f la densité d'une mesure de $\mathcal{P}$ absolument continue par rapport à la mesure de Lebesgue.

Soit $\mathcal{P}'$ l'ensemble des probabilités

$$P = a\, f\, \lambda + (1-a) \sum_m a_m \delta_{d_m} \quad ; \quad 0 \leqslant a \leqslant 1 \;;\; \sum_m a_m = 1$$

$\mathcal{P}'$ est un sous-ensemble convexe de $\mathcal{P}$.

Si la densité de la partie continue des mesures de $\mathcal{P}$ est estimable sans biais, il en est de même de celle de $\mathcal{P}'$.

Estimer la densité de la partie continue d'une mesure de $\mathcal{P}'$ revient à estimer a. Or, d'après le théorème de Bickel-Lehmann, si la densité est estimable sans biais, il existe un estimateur d'ordre 1.

Il existe donc une fonction $T : E \times E \to \mathbb{R}$ telle que

$$\forall a > 0 , \forall t \in E , \quad a f(t) = \int_E T(x,t) \, dP(x)$$

$$a f(t) = a \int_E T(x,t) \, f(x) \, dx + \sum_m a_m T(d_m,t)$$

il faudrait avoir

$$\forall t \in E \qquad f(t) = \int_E T(x,t) \, f(x) \, dx$$

et $\quad \forall d_m \qquad T(d_m,t) = 0$

ce qui est contradictoire.

c - *Choix d'un estimateur de la densité.*

Après avoir construit un estimateur $\hat{\Delta}_n$ de Δ, on reprend la méthode utilisée par Bosq et Bleuez ([5]) - estimateur donné par la formule (3). Pour cela étant donné I, partie infinie, non bornée, de $\mathbb{R}^+$; $K_r(x,t)$ une famille de fonctions réelles, mesurables, de deux variables $(x,t) \in E \times E$, on associe à cette famille une suite d'estimateurs de la densité de la partie continue définis par :

$$\hat{f}_n(t) = \frac{1}{n} \sum_{j=1}^{n} K_{r(n)}(X_j,t) \, 1_{E-\hat{\Delta}_n}(X_j)$$

ou encore en posant

(7) $$S_j = \{(x_1,\ldots,x_j,\ldots,x_n) \mid \forall h \ , \ h \neq j \ , \ x_h \neq x_j\}$$

(8) $$\hat{f}_n(t) = \frac{1}{n} \sum_{j=1}^{n} K_{r(n)}(X_j,t) \ 1_{S_j}(X_1,\ldots,X_n)$$

avec $r(n) \in I$ et $\lim\limits_{n\to\infty} r(n) = \infty$.

2.2. - CONVERGENCE SIMPLE ET EN MOYENNE QUADRATIQUE.-

Nous supposerons le noyau K_r quelconque. On notera

(9) $$\bar{f}_n(t) = \int_E K_{r(n)}(x,t)\, f(x)\, dx$$

a - Espérance et écart-type de $\hat{f}_n(t)$.

Lemme 3.- L'espérance de l'estimateur $\hat{f}_n(t)$ est donnée par :

(10) $$E_p(\hat{f}_n(t)) = \bar{f}_n(t) + \sum_m a_m(1-a_m)^{n-1} K_{r(n)}(d_m,t)$$

Démonstration : Les variables étant indépendantes et de même loi, il vient :

$$E_p(\hat{f}_n(t)) = E_P(K_{r(n)}(X_1,t)\, 1_{S_1}(X_1,\ldots,X_n))$$

$$E_p(\hat{f}_n(t)) = E_P(K_{r(n)}(X_1,t)\, 1_{S_1}(X_1,\ldots,X_n)\big[1_{E-\Delta}(X_1) + 1_\Delta(X_1)\big])$$

Or $$1_{S_1}(X_1,\ldots,X_n)\, 1_{E-\Delta}(X_1) = 1_{E-\Delta}(X_1) \text{ p.s.}$$

(un hyperplan de $\mathbb{R}^n$ est négligeable pour la mesure de Lebesgue) et

$$E_P(K_{r(n)}(X_1,t)\, 1_{E-\Delta}(X_1)) = \int_{E-\Delta} K_{r(n)}(x,t)\, f(x)\, dx = \bar{f}_n(t)$$

D'autre part,

$$E_P(K_{r(n)}(X_1,t)\, 1_{S_1}(X_1,\ldots,X_n)\, 1_\Delta(X_1) =$$

$$= \sum_m a_m K_{r(n)}(d_m,t) \int_{(E-\{d_m\})^{n-1}} dP(x_1)\ \ldots\ dP(x_n)$$

$$= \sum_m a_m K_{r(n)}(d_m,t)\ (1-a_m)^{n-1}$$

d'où l'expression de $E_p(\hat{f}_n(t))$.

Lemme 4.- La variance de $\hat{f}_n(t)$ est donnée par

(11) $$\sigma^2(\hat{f}_n(t)) = \frac{1}{n}\left[\int_E K^2_{r(n)}(x,t)\, f(x)\, dx - (\bar{f}_n(t))^2\right]$$
$$+ \frac{1}{n}\sum_m a_m(1-a_m)^{n-1} K^2_{r(n)}(d_m,t)$$
$$+ 2\,\bar{f}_n(t) \sum_m a_m \left(a_m - \frac{1}{n}\right)(1-a_m)^{n-2} K_{r(n)}(d_m,t)$$
$$+ \frac{n-1}{n} \sum_{m\neq m'} a_m a_{m'} (1-a_m-a_{m'})^{n-2} K_{r(n)}(d_m,t)\, K_{r(n)}(d_{m'},t)$$
$$- \left[\sum_m a_m (1-a_m)^{n-1} K_{r(n)}(d_m,t)\right]^2$$

Démonstration : Calculons $E_P(\hat{f}^2_n(t))$:

$$E_P(\hat{f}^2_n(t)) = E_P\left[\frac{1}{n}\sum_{j=1}^{n} K_{r(n)}(X_j,t)\,(1_{E-\Delta}(X_j) + 1_\Delta(X_j)\, 1_{S_j}(X_1,\ldots,X_n))^2\right]$$

Les variables $K_{r(n)}(X_j,t)\, 1_{E-\Delta}(X_j)$ sont de même loi, ainsi que les variables $K_{r(n)}(X_j,t)\, 1_\Delta(X_j)\, 1_{S_j}(X_1,\ldots,X_n)$.

En développant, il vient après regroupement des termes :

$$E_P(\hat{f}^2_n(t)) = \frac{1}{n} E_P\left[K^2_{r(n)}(X_1,t)\, 1_{E-\Delta}(X_1)\right] + \frac{1}{n} E_P\left[K^2_{r(n)}(X_1,t)\, 1_\Delta(X_1)\, 1_{S_1}(X_1,\ldots,X_n)\right]$$
$$+ \frac{n-1}{n} E_P\left[K_{r(n)}(X_1,t)\, 1_{E-\Delta}(X_1)\, K_{r(n)}(X_2,t)\, 1_{E-\Delta}(X_2)\right]$$
$$+ 2\,\frac{n-1}{n} E_P\left[K_{r(n)}(X_1,t)\, 1_{E-\Delta}(X_1)\, K_{r(n)}(X_2,t)\, 1_\Delta(X_2)\, 1_{S_2}(X_1,\ldots,X_n)\right]$$
$$+ \frac{n-1}{n} E_P\left[K_{r(n)}(X_1,t)\, 1_\Delta(X_1)\, 1_{S_1}(X_1,\ldots,X_n)\, K_r(X_2,t)\, 1_\Delta(X_2)\, 1_{S_2}(X_1,\ldots,X_n)\right]$$

Des calculs analogues à ceux du lemme 3 conduisent à :

$$E_P[K^2_{r(n)}(X_1,t)\ 1_{E-\Delta}(X_1)] = \int_E K^2_{r(n)}(x,t)\ f(x)\ dx$$

$$E_P[K^2_{r(n)}(X_1,t)\ 1_\Delta(X_1)\ 1_{S_1}(X_1,\dots,X_n)] = \sum_m a_m\ (1-a_m)^{n-1}\ K^2_{r(n)}(d_m,t)$$

$$E_P[K_{r(n)}(X_1,t)\ K_{r(n)}(X_2,t)\ 1_{E-\Delta}(X_1)\ 1_{E-\Delta}(X_2)] =$$

$$(E_P[K_{r(n)}(X_1,t)\ 1_{E-\Delta}(X_1)])^2 = \bar{f}^2_n(t)$$

les variables étant indépendantes.

$$E_P[K_{r(n)}(X_1,t)\ K_{r(n)}(X_2,t)\ 1_{E-\Delta}(X_1)\ 1_\Delta(X_2)\ 1_{S_2}(X_1,\dots,X_n)] =$$

$$\int_{E-\Delta} K_{r(n)}(x_1,t)\ [\sum_m a_m\ K_{r(n)}(d_m,t) \int_{(E-d_m)^{n-2}} dP(X_3)\dots dP(X_n)]\ f(x_1)dx_1$$

$$= \bar{f}_n(t) \sum_m a_m\ (1-a_m)^{n-2}\ K_{r(n)}(d_m,t)$$

$$E_P[K_{r(n)}(X_1,t)\ K_{r(n)}(X_2,t)\ 1_\Delta(X_1)\ 1_\Delta(X_2)\ 1_{S_1}(X_1,\dots,X_n)\ 1_{S_2}(X_1,\dots,X_n)] =$$

$$= \sum_{\substack{m,m' \\ m\neq m'}} a_m\ a_{m'}\ K_r(d_m,t)\ K_r(d_{m'},t) \int_{(E-\{d_m,d_{m'}\})^{n-2}} dP(X_3)\ \dots\ dP(X_n)$$

$$= \sum_{m\neq m'} a_m\ a_{m'}\ (1-a_m-a_{m'})^{n-2}\ K_{r(n)}(d_m,t)\ K_{r(n)}(d_{m'},t)$$

D'où
$$E_P(\hat{f}^2_n(t)) = \frac{1}{n}\int_E K^2_{r(n)}(x,t)\ f(x)\ dx + \frac{n-1}{n}\ \bar{f}^2_n(t)$$
$$+ \frac{1}{n}\sum_m a_m\ (1-a_m)^{n-1}\ K^2_{r(n)}(d_m,t)$$
$$+ 2\ \frac{n-1}{n}\ \bar{f}_n(t) \sum_m a_m(1-a_m)^{n-2}\ K_{r(n)}(d_m,t)$$
$$+ \frac{n-1}{n} \sum_{m\neq m'} a_m\ a_{m'}(1-a_m-a_{m'})^{n-2}\ K_{r(n)}(d_m,t)\ K_{r(n)}(d_{m'},t)$$

On en déduit l'expression de $\sigma^2(\hat{f}_n(t))$ en utilisant (10)

$$\sigma^2(\hat{f}_n(t)) = E_P(\hat{f}^2_n(t)) - \bar{f}^2_n(t) - 2\ \bar{f}_n(t) \sum_m a_m\ (1-a_m)^{n-1}\ K_{r(n)}(d_m,t)$$
$$- [\sum_m a_m\ (1-a_m)^{n-1}\ K_{r(n)}(d_m,t)]^2$$

c'est-à-dire la relation (11).

b - *Lemmes relatifs aux séries.*

Les calculs précédents ont fait apparaître des quantités du type

$\sum_m a_m (1-a_m)^{n-1}$, $\sum_m a_m^2 (1-a_m)^{n-2}$,...

Etant donnée une série à termes positifs, de somme $L_0 = \sum_m a_m$ majorée par 1, on définit

$$L_1(n) = \sum_m a_m (1-a_m)^{n-1} \tag{13}$$

$$L_2(n) = \sum_m a_m^2 (1-a_m)^{n-2} \tag{13'}$$

Donnons des majorations de $L_1(n)$ et $L_2(n)$ en fonction de la vitesse de convergence de la série $\sum_m a_m$.

Lemme 5.- Pour toute série convergente à termes positifs, de somme inférieure à 1 , $\lim_{n\to\infty} L_1(n) = 0$ et $\lim_{n\to\infty} n\, L_2(n) = 0$.

Démonstration : En écrivant

$$L_1(n) \leq \sum_{m=1}^{p} a_m(1-a_m)^{n-1} + \sum_{m>p} a_m$$

le premier résultat est immédiat.

Comme $L_2(n) = L_1(n-1) - L_1(n)$

on peut écrire $\qquad L_1(1) = L_0 = L_1(n) + \sum_{k=2}^{n} L_2(k)$

La série de terme général $L_2(n)$ est donc convergente, ce qui implique

$$\lim n\, L_2(n) = 0$$

Lemme 6.- S'il existe une infinité dénombrable d'indices m tels que $a_m \neq 0$, alors $\limsup_{n\to\infty} n\, L_1(n) \neq 0$.

Démonstration : Pour tout $\varepsilon > 0$, il existe m_o tel que

$$\frac{1}{p+1} < a_{m_o} \leqslant \frac{1}{p} < \varepsilon$$

et
$$L_1(p) \geqslant a_{m_o} (1-a_{m_o})^{p-1} \geqslant \frac{1}{p+1} (1-\frac{1}{p})^{p-1}$$

d'où $\limsup p\, L_1(p) \geqslant \frac{1}{e}$.

Lemme 7.- S'il existe un nombre $s > 1$ tel que $\limsup a_m m^s \leqslant \ell$, alors
$$L_1(n) \leqslant C_1\, n^{\frac{1}{s} - 1}$$
$$L_2(n) \leqslant C_2\, n^{\frac{1}{s} - 2}$$

où C_1 et C_2 sont des constantes ne dépendant que de la série $\sum_m a_m$.

Démonstration : Pour simplifier l'écriture, on supposera que $\forall\, m, a_m \leqslant \ell\, m^{-s}$. La fonction définie sur $[0,1]$ par $x \to x\,(1-x)^{n-1}$ admet en $\frac{1}{n}$ un maximum inférieur ou égal à $\frac{1}{2n}$ $(n \geqslant 2)$.

Soit p_o l'entier vérifiant $p_o - 1 < (n\ell)^{1/s} \leqslant p_o$

$$L_1(n) \leqslant \sum_{m=1}^{p_o} a_m (1-a_m)^{n-1} + \sum_{m>p_o} a_m$$

$$\leqslant p_o \frac{1}{2n} + \int_{(n\ell)^{1/s}}^{\infty} \frac{\ell}{x^s}\, dx$$

$$\leqslant \frac{1 + (n\ell)^{1/s}}{2n} + \frac{(n\ell)^{1/s}}{n\,(s-1)} \leqslant C_1\, n^{\frac{1}{s} - 1}$$

La démonstration est analogue pour $L_2(n)$.
La fonction $x \mapsto x^2\,(1-x)^{n-2}$ admet en $\frac{2}{n}$ un maximum inférieur ou égal à $\frac{1}{n^2}$ $(n \geqslant 4)$. On prend p_o tel que $p_o - 1 < (\frac{n\ell}{2})^{1/s} \leqslant p_o$

Lemme 8.- Si $\liminf a_m m^s = \ell$, alors
$$L_1(n) \geqslant C'_1\, n^{\frac{1}{s} - 1}$$
$$L_2(n) \geqslant C'_2\, n^{\frac{1}{s} - 2}$$

On prend n assez grand et p_o un entier tel que les relations suivantes soient vérifiées

$$p_o \leqslant (n\ell)^{1/s} < p_o + 1 \ ; \ \forall\, m \geqslant p_o, \frac{1}{2}\, \ell\, m^{-s} < a_m \leqslant \frac{2}{n}$$

alors $L_1(n) \geqslant \sum_{m \geqslant p_o} a_m (1-a_m)^{n-1}$

$$\geqslant \frac{1}{e^2} \sum_{m \geqslant p_o} a_m \geqslant \frac{1}{e^2} \int_{(n\ell)^{1/s}}^{\infty} \frac{\ell}{x^s}\, dx \geqslant C'_1\, n^{1/s-1}$$

La seconde inégalité se démontre de la même manière.

Lemme 9.- S'il existe un nombre $\gamma > 1$, tel que $\limsup \gamma^m a_m \leqslant \ell$ alors $L_1(n) \leqslant \frac{C}{n}$ et $L_2(n) \leqslant \frac{C'}{n^2}$ où C et C' sont deux constantes.

Pour simplifier les calculs, on supposera que $\forall m \geqslant 1 \quad a_m = \gamma^{-m}$

Soit p_o un entier tel que

$$p_o \leqslant \frac{\text{Log } n}{\text{Log } \gamma} < p_o + 1 \qquad\qquad a_{p_o+1} \leqslant \frac{1}{n} \leqslant a_{p_o}$$

$$L_1(n) = \sum_{m=1}^{p_o-1} \gamma^{-m} (1-\gamma^{-m})^{n-1} + \gamma^{-p_o} (1-\gamma^{-p_o})^{n-1} + \sum_{m=p_o+1}^{\infty} \gamma^{-m}(1-\gamma^{-m})^{n-1}$$

$$\leqslant \sum_{m=1}^{p_o-1} \gamma^{-m} (1-\gamma^{-m})^{n-1} + \frac{1}{2n} + \gamma^{-(p_o+1)} \frac{1}{1-\gamma^{-1}}$$

Pour majorer la première somme, on va calculer le rapport de deux termes consécutifs. En posant $v_k = \gamma^{-k}(1-\gamma^{-k})^{n-1}$

$$\frac{v_k}{v_{k+1}} = \gamma \left[\frac{1 - \gamma^{-k}}{1 - \gamma^{-(k+1)}}\right]^{n-1} < \gamma(1 - (\gamma-1)\, \gamma^{-(k+1)})^{n-1}$$

Or, si $k < p_o - 1$, $(\gamma-1)\, \gamma^{-k-1} > (\gamma-1)\, \gamma^{-p_o} \geqslant (\gamma-1)\, \frac{1}{n}$

d'où $$\frac{v_k}{v_{k+1}} < \gamma\ (1 - \frac{\gamma-1}{n})^{n-1}$$

Pour tout $\varepsilon \in]0,\ 1 - \gamma\ e^{1-\gamma}[$, on a pour n assez grand

$$\frac{v_k}{v_{k+1}} < (1+\varepsilon)\ \gamma\ e^{1-\gamma} < 1 - \varepsilon^2$$

on peut donc majorer

$$\sum_{m=1}^{p_o-1} \gamma^{-m}\ (1-\gamma^{-m})^{n-1}$$

par $$a_{p_o-1}\ (1 - a_{p_o-1})^{n-1} \sum_{m=1}^{p_o-1} (1-\varepsilon^2)^k \leqslant \frac{1}{2\ n\ \varepsilon^2}$$

On en déduit le résultat annoncé sur $L_1(n)$.

On démontre de même que le rapport des deux termes consécutifs $a_k^2\ (1-a_k)^{n-2} : \gamma^2\ (1 - (\gamma-1)\ \gamma^{-k-1})^{n-2}$ peut être majoré par un nombre strictement inférieur à 1 dès que $k < p_o - \frac{2}{\text{Log}\ \gamma}$ pour n assez grand et que $L_2(n) \leqslant \frac{C'}{n^2}$.

c - *Convergence en moyenne d'ordre 1 et 2.*

Dans la suite, on supposera que le noyau K_r vérifie les deux conditions :

(14) $$\forall\ t \in E \quad \exists\ A(t) \quad |K_r(x,t)| \leqslant A(t)\ r^\alpha$$

(15) $$\forall\ t \in E \quad \exists\ B(t) \int_E K_r^2(x,t)\ dx \leqslant B(t)\ r^\alpha$$

où α est un nombre réel strictement positif, ne dépendant que de la famille K_r.

Les conditions de convergence en moyenne d'ordre 1 et 2 de l'estimateur $\hat{f}_n(t)$ sont données par la

Proposition 8.- Soit $\mathcal{D}$ l'ensemble des densités de la partie continue des probabilités définies sur E ; si les hypothèses suivantes sont vérifiées :

α) les densités de $\mathcal{D}$ sont bornées,

β) l'intégrale $\int_E K_r(x,t)\, f(x)\, dx$ converge vers $f(t)$ lorsque $r \to \infty$,

γ) $\lim\limits_{n\to\infty} r(n) = \infty$.

Si le support des masses est infini, la condition $(r(n))^\alpha \; L_1(n) \to 0$ implique la convergence de $\hat{f}_n(t)$ vers $f(t)$ en moyenne quadratique. Si le support des masses est fini, ou s'il existe un nombre $\gamma > 1$, tel que $\lim\limits_{m\to\infty} a_m \, \gamma^m = 0$.

Alors les conditions :

i) $\dfrac{(r(n))^\alpha}{n} \to 0$

ii) $\forall f \in \mathcal{D} \quad E(|\hat{f}_n(t) - f(t)|) \longrightarrow 0$

iii) $\forall f \in \mathcal{D} \quad E(|\hat{f}_n(t) - f(t)|^2) \longrightarrow 0$

sont équivalentes.

Démonstration : D'après le lemme 6, la condition $(r(n))^\alpha \; L_1(n) \to 0$ implique $\dfrac{r(n)^\alpha}{n} \to 0$.

En notant $\bar{f}_n(t) = \int K_r(x,t)\, f(x)\, dx$

$$E_P(|\hat{f}_n(t) - f(t)|^2) = \sigma^2(\hat{f}_n(t)) + (E_P(\hat{f}_n(t)) - f(t))^2$$

$$= \sigma^2(\hat{f}_n(t)) + \left[\bar{f}_n(t) - f(t) + \frac{1}{n}\sum_m a_m \; (1-a_m)^{n-1} \; K_{r(n)}(d_m,t)\right]^2$$

Les conditions β) et $(r(n))^\alpha \; L_1(n) \to 0$ impliquent que le terme entre crochets a une limite nulle.

Nous allons chercher une majoration de $\sigma^2(\hat{f}_n(t))$.

La fonction f est bornée sur E ; et la fonction $\bar{f}_n$ converge simplement vers f ; on peut donc écrire pour n assez grand $\bar{f}_n(t) \leqslant M$.
En utilisant $K_{r(n)}(x,t) \leqslant A(t)\,(r(n))^\alpha$, il vient :

$$\sigma^2(\hat{f}_n(t)) \leqslant \frac{1}{n} M\, A(t)(r(n))^\alpha + \frac{r^{2\alpha}}{n} L_1(n)\, A^2(t) + 2\, M\, (L_2(n) + \frac{L_1(n)}{n})\, A(t)(r(n))^\alpha$$

$$+ \sum_{m\neq m'} a_m\, a_{m'}\, (1-a_m-a_{m'})^{n-1}\, A^2(t)(r(n))^{2\alpha}$$

Or $$\sum_{m\neq m'} a_m\, a_{m'}\, (1-a_m-a_{m'})^{n-1} \leqslant \sum_{m,m'} a_m\, a_{m'}\, (1-a_m-a_{m'}+a_m a_{m'})^{n-1} \leqslant L_1(n)^2$$

Les conditions

$$\frac{(r(n))^\alpha}{n} \longrightarrow 0 \quad \text{et} \quad L_1(n)\,(r(n))^\alpha \longrightarrow 0 \quad \text{impliquent} \quad \sigma^2(\hat{f}_n(t)) \longrightarrow 0$$

ce qui établit la convergence de l'estimateur en moyenne quadratique.

De plus, si le support des masses est fini,
$L_1(n)$ tend vers 0 plus vite que $\frac{1}{n}$;

s'il existe un nombre $\gamma > 1$, tel que $\lim a_m\, \gamma^m = 0$,
$L_1(n)$ est de l'ordre de $\frac{1}{n}$;

La majoration de $\sigma^2(\hat{f}_n(t))$ nous donne i) $\Longrightarrow$ iii) et iii) $\Longrightarrow$ ii) est la conséquence de l'inégalité de Schwarz.

La condition nécessaire ii) $\Longrightarrow$ i) a été établie par Bleuez - ([5]), prop. I.1 - dans le cas où la famille de probabilités est dominée par une mesure μ.

En considérant des probabilités, sommes d'une mesure absolument continue - par rapport à la mesure de Lebesgue - et d'une mesure de Dirac, on peut se ramener au cas où la famille est dominée.

En effet, l'estimateur de l'ensemble des masses converge presque complètement et la famille est dominée par $\lambda + \delta_d$.

2.3. - CONVERGENCE SIMPLE PRESQUE COMPLETE.-

Proposition 9.- Sous les hypothèses de la proposition 8, on peut pour toute répartition des masses (d_m, a_m) construire une suite $r(n) \to \infty$ telle que $\hat{f}_n(t) \xrightarrow{p.c_o.} f(t)$ quand $n \to \infty$, pour tout $f \in \mathcal{D}$.

Si la vitesse de convergence de la suite (a_m) est connue, on a les résultats suivants :

i) s'il existe un nombre $s > 1$ tel que $a_m m^s \to 0$ quand $m \to \infty$

$$\frac{(r(n))^\alpha}{n^{1-1/s}} \longrightarrow 0 \Longrightarrow \forall f \in \mathcal{D}\ ,\ \hat{f}_n(t) \xrightarrow{p.c_o.} f(t)$$

ii) s'il existe un nombre $\gamma > 1$ tel que $a_m \gamma^m \to 0$ quand $m \to \infty$

$$\frac{(r(n))^\alpha \operatorname{Log} n}{n} \longrightarrow 0 \Longrightarrow \forall f \in \mathcal{D},\ \hat{f}_n(t) \xrightarrow{p.c_o.} f(t)$$

iii) si le support des masses est fini

s'il existe une densité $f_o \in \mathcal{D}$ et une fonction $A_1(t)$ telle que $\forall x \in E \quad \int K_r^2(x,t)\ f_o(x)\ dx \geq A_1(t)\ r^\alpha$

les conditions suivantes sont équivalentes :

a) $\forall \beta > 0 \quad \sum_n \exp\left(-\frac{\beta n}{(r(n))^\alpha}\right) < \infty$

b) $\forall f \in \mathcal{D}\ ,\ \forall t \in E \quad \hat{f}_n(t) \xrightarrow{p.c_o.} f(t)$ et $E(|\hat{f}_n(t) - f(t)|^2) \to 0$

Démonstration :

$$\hat{f}_n(t) - f(t) = \sum_{j=1}^{n} \frac{1}{n} K_{r(n)}(X_j,t)\ 1_{S_j}(X_1,\ldots,X_n) - f(t)$$

(16) $$= I_1 + I_2 + I_3$$

avec $I_1 = \sum_{j=1}^{n} \frac{1}{n} K_{r(n)}(X_j,t)\ 1_{E-\Delta}(X_j)\ 1_{S_j}(X_1,\ldots,X_n) - \int_E K_r(x,t)\ f(x)\ dx$

$$= \sum_{j=1}^{n} \frac{1}{n} K_{r(n)}(X_j,t)\ 1_{E-}\ (X_j) - \int_E K_{r(n)}(x,t)\ f(x)\ dx \quad \text{p.s.}$$

$$I_2 = \sum_{j=1}^{n} \frac{1}{n} K_{r(n)}(X_j,t)\ 1_{\Delta}(X_j)\ 1_{S_j}(X_1,\ldots,X_n)$$

$$I_3 = \int_E K_{r(n)}(x,t)\ f(x)\ dx - f(t)$$

Comme $\int_E K_{r(n)}(x,t)\ f(x)\ dx$ converge simplement vers $f(t)$, lorsque $r \to \infty$, il suffit d'établir que I_1 et I_2 converge vers 0 p.c_o.

a - Convergence p.c_o. *de* I_1.

Nous reprenons rapidement la démonstration de Bleuez ([5]), lemme 5. Celle-ci utilise l'inégalité de Hoeffding ([9]) p. 58

$$P(S_n > n\varepsilon) \leq \exp\left(-\frac{n\varepsilon}{b}\left[\left(1 + \frac{\sigma^2}{b\varepsilon}\right) \text{Log}\left(1 + \frac{b\varepsilon}{\sigma^2}\right) - 1\right]\right)$$

$S_n = \sum_{k=1}^{n} X_k$, où les X_k sont des variables aléatoires, indépendantes, de même loi, centrées, majorées par b.

$$0 < \varepsilon < b \quad \text{et} \quad \sigma^2 = \frac{1}{n}\sum_{k=1}^{n} E(X_k^2)$$

Les variables $K_{r(n)}(X_j,t)\ 1_{E-\Delta}(X_j) - \int_E K_{r(n)}(x,t)\ f(x)\ dx$ sont indépendantes, de même loi, centrées, majorées par $b = 2\ A(t)\ (r(n))^{\alpha}$ de variance $\int_E K^2_{r(n)}(x,t)\ f(x)\ dx - \bar{f}_n^2(t)$.

La quantité $\frac{\sigma^2}{b\varepsilon}$ est majorée par $\frac{M.B(t)}{2\ \varepsilon\ A(t)}$ où $M = \sup_{t\in E} f(t)$.

La fonction $u \mapsto (1+u)\ \text{Log}(1+\frac{1}{u}) - 1$ est strictement positive et décroissante sur $]0\ ,\ \infty[$. En posant

$$\beta(t) = \frac{1}{2\ A(t)}\left[\left(1 + \frac{M\ B(t)}{2\ \varepsilon\ A(t)}\right) \text{Log}\left(1 + \frac{2\ \varepsilon\ A(t)}{M\ B(t)}\right) - 1\right]$$

il vient :

(17) $$P(|I_1| \geq \varepsilon) \leq 2 \exp(- \varepsilon\, \beta(t) \frac{n}{(r(n))^{\alpha}})$$

b - Convergence p.c$_o$. *de* I_2.

On étudiera tout d'abord le cas particulier où Δ n'a qu'un nombre fini N_o de masses.

$$\sum 1_{\Delta}(X_j)\, 1_{S_j}(X_1,\ldots,X_n) \quad \text{est majoré par } N_o$$

et $$|I_2| \leq \frac{N_o}{n} A(t)\, (r(n))^{\alpha}$$

la condition $\frac{(r(n))^{\alpha}}{n} \to 0$, implique la convergence de I_2 vers 0.

Si le nombre des masses est infini, pour tout entier k , on définit N_k par :

(18) N_k est le plus petit entier tel que $\sum_{m>N_k} a_m \leq \frac{1}{k}$

On pose

$$\Delta_k = \Delta - \{d_1 , d_2 ,\ldots, d_{N_k}\} , \quad \text{alors} \quad P(\Delta_k) \leq \frac{1}{k}$$

$$\sum_{j=1}^{n} 1_{\Delta}(X_j)\, 1_{S_j}(X_1,\ldots,X_n) = \sum_{j=1}^{n} (1_{\Delta_k}(X_j) + 1_{\Delta-\Delta_k}(X_j))\, 1_{S_j}(X_1,\ldots,X_n)$$

$$\leq \sum_{j=1}^{n} 1_{\Delta-\Delta_k}(X_j)\, 1_{S_j}(X_1,\ldots,X_n) + \sum_{j=1}^{n} 1_{\Delta_k}(X_j)$$

$$\leq N_k + \sum_{j=1}^{n} 1_{\Delta_k}(X_j)$$

d'où (19) $$|I_2| \leq \frac{(r(n))^{\alpha}}{n} A(t)\, (N_k + \sum_{j=1}^{n} 1_{\Delta_k}(X_j))$$

Les variables $1_{\Delta_k}(X_j)$ sont indépendantes, de même loi binomiale. d'espérance $\leq \frac{1}{k}$

En utilisant l'inégalité de Hoeffding, avec $b = 1$, $\varepsilon = \frac{1}{k}$, $\sigma^2 \leq \frac{1}{k}$;

$$P\left(\frac{1}{n}\sum_{j=1}^{n} (1_{\Delta_k}(X_j) - E(1_{\Delta_k}(X_j))) > \frac{1}{k}\right) \leq \exp\left(-\frac{n}{k}(2 \text{ Log } 2 - 1)\right)$$

d'où
$$P\left(\frac{1}{n}\sum_{j=1}^{n} 1_{\Delta_k}(X_j) > \frac{2}{k}\right) \leq \exp\left(-\frac{n}{k}(2 \text{ Log } 2 - 1)\right)$$

On en déduit

$$P\left(|I_2| > \frac{3}{k} A(t)\ (r(n))^{\alpha}\right) \leq \exp\left(-\frac{n}{k}(2 \text{ Log } 2 - 1)\right)$$

dès que $\frac{N_k}{n} \leq \frac{1}{k}$.

La démonstration sera achevée si on peut construire deux suites $k(n)$ et $r(n)$ vérifiant les conditions

$$k(n) \to \infty \quad , \quad r(n) \to \infty$$

(20) $$k(n)\, N_{k(n)} \leq n \quad (c) \quad , \quad \frac{(r(n))^{\alpha}}{k(n)} \to 0 \quad (d)$$

$$\frac{k(n) \text{ Log } n}{n} \longrightarrow 0 \quad (e)$$

La condition (d) implique

$$\forall \varepsilon > 0 \quad , \frac{3}{k(n)} A(t)\ (r(n))^{\alpha} < \varepsilon \quad \text{pour } n \text{ assez grand}$$

et
$$P(|I_2| > \varepsilon) \leq \exp\left(-\frac{n}{k(n)}(2 \log 2 - 1)\right) .$$

Les relations $N_k \leq N_{k+1}$ et $k\, N_k < (k+1)\, N_{k+1}$ sont toujours vérifiées.

On peut donc définir une suite $k(n)$ par :

(21) $$k(n) = \inf\left(k, \left[\frac{n}{(\text{Log } n)^{\eta}}\right]\right) \quad \text{si} \quad k\, N_k \leq n < (k+1)\, N_{k+1}$$

où $[x]$ désigne la partie entière de x , et $\eta > 1$.

On en déduit l'existence d'une suite $r(n)$ vérifiant $r(n) \to \infty$ et (d).

La convergence simple p.c$_o$. de I_2 est donc établie et celle de I_1 résulte de $\frac{r(n)^\alpha \text{ Log } n}{n} \to 0$

c - Lorsque l'ordre de grandeur des masses a_m est connu, on peut préciser l'ordre de $k(n)$ et de $r(n)$. Nous le ferons dans deux cas.

* $a_m = 0 \left(\frac{1}{m^s}\right)$, $s > 1$

$$\sum_{m=N_k}^{\infty} \frac{1}{m^s} = 0 \left(N_k^{\ 1-s}\right)$$

d'après (17), $\frac{1}{k} = 0 \left(N_k^{\ 1-s}\right)$ ou $N_k = 0 \left(k^{1/s-1}\right)$ et en utilisant (19)

$$k\ N_k = 0(n)$$

on peut choisir une suite $k(n)$, $k(n) = 0 \left(n^{1-1/s}\right)$.

La convergence presque complète sera assurée si $r(n)$ vérifie :

$$\lim_{n\to\infty} \frac{(r(n))^\alpha}{n^{1-1/s}} = 0$$

* $a_m = 0 \left(\gamma^{-m}\right)$, $\gamma > 1$

$$\sum_{m>N_k} \gamma^{-m} = \frac{1}{\gamma-1} \gamma^{-N_k}$$

d'où $N_k = 0$ (Log k).

Les conditions (20) et (21) seront vérifiées avec $k(n) = o \left(\frac{n}{\text{Log } n}\right)$

$\frac{r(n)^\alpha \text{ Log } n}{n} \to 0$ implique donc la convergence p.c$_o$. de $\hat{f}_n(t)$.

d - Pour établir la partie iii) de la proposition, on remarque que, si l'ensemble des masses est fini, la convergence presque complète de $\hat{f}_n(t)$ est équivalente à celle de I_1.

D'après la proposition 1, on est donc ramené au cas où l'ensemble des masses est connu, et les théorèmes démontrés lorsque l'ensemble des mesures est dominé s'appliquent.

La partie iii) résulte alors de la proposition 3 de Bleuez ([5]).

2.4. - CONVERGENCE UNIFORME PRESQUE COMPLETE.-

Pour étudier la convergence uniforme presque complète, des hypothèses supplémentaires sur le noyau K_r sont nécessaires. Dans le cas où l'ensemble E est un intervalle compact de $\mathbb{R}$, ou si l'on se restreint à l'étude de la convergence uniforme sur un intervalle compact de $\mathbb{R}$, on supposera que K_r vérifie une condition de Lipschitz sur l'intervalle E_1 de longueur strictement positive, plus précisément que

il existe des constantes réelles, positives C, η, μ telles que

(22) $$\forall r \in I \ , \forall x \in E_1 \ , \forall t \ , t_o \in E_1$$

$$|K_r(x,t) - K_r(x,t_o)| \leqslant C\, r^\eta \,|t - t_o|^\mu$$

On a alors la

Proposition 10.- Soit $\mathcal{D}$ l'ensemble des densités de la partie continue des probabilités définies sur E.

Si les hypothèses suivantes sont vérifiées

a) les densités de $\mathcal{D}$ sont bornées,

b) $\exists\, A \ , \forall x \in E \ \ \forall t \in E \ \ |K_r(x,t)| \leqslant A\, r^\alpha$

b') $\exists\, B \ \int_E K_r^2(x,t)\, dx \leqslant B\ r^\alpha$ A et B indépendants de t.

c) l'intégrale $\int_E K_r(x,t)\, f(x)\, dx$ converge uniformément vers $f(t)$ lorsque $r \to \infty$.

d) la famille K_r vérifie la condition (22).

Pour toute suite (d_m, a_m) de masses, on peut construire une suite $r(n)$ telle que $\lim r(n) = \infty$

et $$\sup_{t \in E_1} |\hat{f}_n(t) - f(t)| \xrightarrow{p.c_o.} 0$$

De plus, si l'ordre de grandeur de la suite a_m est connu, l'une des deux conditions implique la convergence uniforme presque complète :

i) $\exists\ s > 1\ ,\quad a_m = 0\,(m^{-s})\quad$ et $\quad (r(n))^{\alpha}\, n^{\frac{1}{s}-1} \to 0$

ii) $\exists\ \gamma > 1\ ,\quad a_m = 0\,(\gamma^{-m})\quad$ et $\quad \dfrac{(r(n))^{\alpha}\ \mathrm{Log}\ n}{n} \to 0$.

Démonstration : La relation (16) nous donne une majoration de $\sup\limits_{t\in E_1} |\hat{f}_n(t) - f(t)|$; $\sup|\hat{f}_n(t) - f(t)| \leqslant |I_1|+|I_2|+|I_3|$ avec

$$I_1 = \sup_{t\in E_1} \sum_{j=1}^{n} \frac{1}{n} K_{r(n)}(X_j,t)\ 1_{E-\Delta}(X_j) - \int_E K_{r(n)}(x,t)\ f(x)\ dx \qquad \text{p.s.}$$

$$I_2 = \sup_{t\in E_1} \sum_{j=1}^{n} \frac{1}{n} K_{r(n)}(X_j,t)\ 1_{\Delta}(X_j)\ 1_{S_j}(X_1,\dots,X_n)$$

$$I_3 = \sup_{t\in E_1} \int_E K_{r(n)}(x,t)\ f(x)\ dx - f(t)$$

Il suffit d'établir que I_1 converge presque complètement vers 0 , puisque la condition c) exprime la convergence de I_3 , et que la démonstration de la convergence de I_2 reprend celle de la proposition 9 , partie b) sans autre modification que : l'inégalité (19) est remplacée par :

$$I_2 \leqslant \frac{(r(n))^{\alpha}}{n}\ A\ (N_k + \sum_{j=1}^{n} 1_{\Delta_k}(X_j))$$

Soit (J_k) un recouvrement de E_1 par $\left[\lambda(E_1)\ \ r(n)^{\frac{\eta+1}{\mu}} + 1\right]$ intervalles de centre t_k , de rayon $\dfrac{1}{r(n)^{\eta+1/\mu}}$.

En posant $\hat{H}_n(t) = \sum\limits_{j=1}^{n} \frac{1}{n} K_{r(n)}(X_j,t)\ 1_{E-\Delta}(X_j) - \int_E K_{r(n)}(x,t)\ f(x)\ dx$ il vient, en utilisant la condition de Lipschitz (22)

$$\forall\ t \in J_k\ \ |K_{r(n)}(x,t_k) - K_{r(n)}(x,t)| \leqslant C\ (r(n))^{\eta}\ |t_k - t|^{\mu}$$

$$\leqslant C\ (r(n))^{\eta} \left(\frac{1}{r(n)^{\eta+1/\mu}}\right)^{\mu} = \frac{C}{r(n)}$$

d'où : $\forall\, t \in J_k \,,\quad |\hat{H}_n(t) - \hat{H}_n(t_k)| \leq 2\,\dfrac{C}{r(n)}$.

Alors
$$\sup_{t\in E_1} |\hat{H}_n(t)| \leq \sup_k |\hat{H}_n(t_k)| + 2\,\frac{C}{r(n)}$$

et, pour n assez grand

$$\begin{aligned} P(\sup_{t\in E_1} |\hat{H}_n(t)| \geq \varepsilon) &\leq P(\sup_k |\hat{H}_n(t_k)| \geq \frac{\varepsilon}{2}) \\ &\leq \sum_k P(|\hat{H}_n(t_k)| \geq \frac{\varepsilon}{2}) \\ &\leq \left[\lambda(E_1)\quad (r(n))^{\eta+1/\mu} + 1\right] P(|\hat{H}_n(t_k)| \geq \frac{\varepsilon}{2}) \end{aligned}$$

Si l'on pose $\qquad \beta = \dfrac{1}{2A}\left[(1 + \dfrac{M\,B}{\varepsilon\,A})\ \mathrm{Log}(1 + \dfrac{\varepsilon\,A}{B\,M}) - 1\right]$

d'après la relation (17) de la prop. 9 , il vient :

$$P(|I_1| \geq \varepsilon) \leq \left[\lambda(E_1)\quad r(n)^{\eta+1/\mu} + 1\right] \exp(-\frac{\beta\, n}{(r(n))^{\alpha}})$$

La convergence presque complète de I_1 est assurée dès que

$$\sum_n r(n)^{\eta+1/\mu} \exp(-\frac{\beta\, n}{(r(n))^{\alpha}}) < \infty$$

en particulier, si la condition

$$\text{ii)} \qquad \frac{(r(n))^{\alpha}\ \mathrm{Log}\ n}{n} \longrightarrow 0$$

est vérifiée.

2.5. - LOI LIMITE DE L'ESTIMATEUR.

Un noyau K_r étant donné, nous lui avons associé un estimateur $\hat{f}_n(t)$ de la densité de la partie continue. Nous cherchons, pour t_o fixé dans E, la loi limite de $\hat{f}_n(t_o)$.

Proposition 11. - Si les conditions suivantes sont remplies :

a) la densité f est bornée par M ;

a') l'intégrale $\bar{f}_n(t_o) = \int_E K_{r(n)}(x,t_o)f(x)dx$ converge simplement vers $f(t_o)$;

b) $\exists A(t_o) = A$ tel que $\forall x \in E$, $|K_{r(n)}(x,t_o)| \leq A(r(n))^{\alpha}$;

c) $\exists B_1$, $\exists B_2$, tels que $0 < B_1 r^{\alpha} \leq \int_E K^2_{r(n)}(x,t)f(x)dx$;

et $\int_E K^2_{r(n)}(x,t)dx \leq B_2(r(n))^{\alpha}$;

d) $\exists s > 2$ tel que $\limsup a_m m^s = \ell < \infty$ (éventuellement $\ell = 0$)

e) $\lim\limits_{n\to\infty} r(n) = \infty$ et $\lim\limits_{n\to\infty} \frac{(r(n))^{\alpha}}{n^{1-2/s}} = 0$

alors

$$\hat{S}_n(t_o) = \frac{\hat{f}_n(t_o) - \bar{f}_n(t_o)}{\sigma(\hat{f}_n(t_o))}$$

converge en loi vers une variable aléatoire dont la loi et normale de moyenne nulle et d'écart-type 1.

Démonstration : On peut écrire $\hat{S}_n(t_o)$ comme somme de variables aléatoires :

$$\hat{S}_n(t_o) = \sum_{j=1}^{n} Y_{j,n} + \sum_{j=1}^{n} Z_{j,n}$$

avec

$$Y_{j,n} = \frac{1}{n\,\sigma(\hat{f}_n(t_o))} (K_{r(n)}(X_j,t_o)1_{E-\Delta}(X_j) - \bar{f}_n(t_o))$$

$$Z_{j,n} = \frac{1}{n\,\sigma(\hat{f}_n(t_o))} K_{r(n)}(X_j,t_o)1_{\Delta}(X_j)1_{S_j}(X_1,\dots,X_n)$$

Dans la suite de la démonstration, on les notera simplement Y_j et Z_j sans mentionner l'indice n ; on prendra également $t_o = t$.

Pour établir la convergence en loi de $\hat{S}_n(t)$, il suffit de montrer que :

i) $\sum_{j=1}^{n} Z_j$ converge en probabilité vers 0, et

ii) $\sum_{j=1}^{n} Y_j$ converge en loi vers $\mathcal{N}(0,1)$.

En effet, si V_n est une suite de variables aléatoires convergeant en loi vers V et W_n une suite convergeant en probabilité vers 0, la somme converge en loi vers V.

- dém. : PETROV (9) p. 16, th. 12.

Nous commencerons par rechercher un encadrement de $\sigma(\hat{f}_n(t))$.

Lemme 10.- Sous les hypothèses de la proposition 11

$$\sigma^2(\hat{f}_n(t)) = \frac{1}{n}\left[\int_E K^2_{r(n)}(x,t)f(x)dx - \bar{f}^2_n(t)\right](1+\eta_n(t)) \quad \text{où} \quad \eta_n(t) \to 0.$$

Démonstration :

La condition a') entraîne la majoration de $\bar{f}_n(t)$ par $1+M$ dès que n est assez grand.

et $\frac{1}{n}\left[\int_E K_{r(n)}(x,t)f(x)dx - \bar{f}^2_n(t)\right]$ est de l'ordre de $\frac{(r(n))^\alpha}{n}$.

Nous reprenons l'expression de $\sigma^2(\hat{f}_n(t))$ donnée par le lemme 4.

$$(11) \quad \sigma^2(\hat{f}_n(t)) = \frac{1}{n}\left(\int_E K^2_{r(n)}(x,t)f(x)dx - \bar{f}^2_n(t)\right) + G_n(t)$$
$$+ 2\,\bar{f}_n(t) \sum_m a_m\left(a_m - \frac{1}{n}\right)(1 - a_m)^{n-2} K_{r(n)}(d_m,t)$$

où $G_n(t) = \frac{n-1}{n} \sum_{m\neq m'} a_m a_{m'} (1 - a_m - a_{m'})^{n-2} K_{r(n)}(d_m,t) K_{r(n)}(d_{m'},t)$

(23) $$+ \frac{1}{n} \sum_m a_m (1-a_m)^{n-1} K^2_{r(n)}(d_m,t) - (\sum_m a_m (1-a_m)^{n-1} K_r(d_m,t))^2 .$$

En reprenant les notations du lemme 7,

$$\left| \frac{2}{A(M+1)(r(n))^{\alpha}} \bar{f}_n(t) \sum_m a_m (a_m - \frac{1}{n})(1 - a_m)^{n-2} K_r(d_m,t) \right| \leq L_2(n) + \frac{1}{n} L_1(n-1)$$

or, sous la condition d) $\frac{1}{n} L_1(n-1)$ et $L_2(n)$ sont majorés par $C\, n^{\frac{1}{s} - 2}$.

D'autre part, $G_n(t) =$

(24) $$- \frac{1}{n} \sum_{m\neq m'} a_m a_{m'} (1-a_m-a_{m'})^{n-2} K_r(d_m,t) K_r(d_{m'},t)$$

$$+ \sum_{m,m'} a_m a_{m'} (1-a_m-a_{m'})^{n-2} K_r(d_m,t) K_r(d_{m'},t) - \sum_m a_m^2 (1-2a_m)^{n-2} K_r^2(d_m,t)$$

$$+ \frac{1}{n} \sum_m a_m (1-a_m)^{n-1} K^2_{r(n)}(d_m,t) - (\sum_m a_m (1-a_m)^{n-1} K_{r(n)}(d_m,t))^2.$$

En regroupant les 2° et 5° termes de la somme

$$\left| \sum_{m,m'} a_m a_{m'} \left[(1-a_m-a_{m'})^{n-2} - (1-a_m)^{n-1}(1-a_{m'})^{n-1}\right] K_r(d_m,t) K_r(d_{m'},t) \right|$$

est majoré par : (*)

$$\sum_{m,m'} (a_m a_{m'})^2 (n-2)(1-a_m)^{n-3}(1-a_{m'})^{n-3} \left| K_r(d_m,t) K_r(d_{m'},t) \right|$$

$$+ \sum_{m,m'} (a_m^2 a_{m'}^2 + 2a_m^2 a_{m'})(1-a_m)^{n-2}(1-a_{m'})^{n-2} | K_r(d_m,t) K_r(d_m,t) K_r(d_{m'},t) |$$

d'où

$$G_n(t) \leq \frac{1}{n} (L_1(n))^2 A^2 (r(n))^{2\alpha} + L_2(n) A^2 (r(n))^{2\alpha} + \frac{1}{n} L_1(n)(r(n))^{2\alpha} A^2 +$$

$$+ \left[(n-2)(L_2(n-1))^2 + (L_2(n))^2 + 2L_1(n-1)L_2(n)\right] A^2 (r(n))^{2\alpha}$$

et sous la condition d)

(*) il suffit d'utiliser les relations :
si $0 < x < y$, $y^n - x^n = (y-x)(y^{n-1} + xy^{n-2} + \ldots + x^{n-1}) < n(y-x)y^{n-1}$.

$$G_n(t) \leq C'(r(n))^{2\alpha} n^{\frac{1}{s}-2} .$$

Les majorations précédentes donnent un ordre de grandeur du terme $\eta_n(t)$

$$\eta_n(t) = \frac{G_n(t) + O((r(n))^{\alpha} n^{\frac{1}{s}-2})}{O((r(n))^{\alpha} n^{-1})} = O((r(n))^{\alpha} n^{\frac{1}{s}-1})$$

Démonstration de i).

Pour n assez grand, la condition b) et le lemme 11 entraînent

$$|Z_j| \leq \frac{1}{n}\left(\frac{B_1(t)(r(n))^{\alpha}}{2n}\right)^{-1/2} A(t)(r(n))^{\alpha} 1_{\Delta}(X_j) 1_{S_j}(X_1,\dots,X_n)$$

et

$$|\sum Z_j| \leq \frac{\sqrt{2}\,A}{\sqrt{B_1}} \sqrt{\frac{(r(n))^{\alpha}}{n}} \sum_j 1_{\Delta}(X_j) 1_{S_j}(X_1,\dots,X_n)$$

il suffit d'établir que

$$\forall \varepsilon > 0,\ P\left(\sqrt{\frac{(r(n))^{\alpha}}{n}} \sum_j 1_{\Delta}(X_j) 1_{S_j}(X_1,\dots,X_n) > \varepsilon\right) \to 0$$

La méthode sera très proche de celle utilisée dans la démonstration de la Proposition 9, partie b).

On supposera que pour m assez grand $a_m \leq \frac{1}{m^s}$ et on note $\Delta_{n,s} = \Delta - \{d_1, d_2, \dots, d_{[n^{1/s}]}\}$ où $[x]$ représente la partie entière de x.

$$\sum_j 1_{\Delta}(X_j) 1_{S_j}(X_1,\dots,X_n) \leq n^{1/s} + \sum_{j=1}^{n} 1_{\Delta_{n,s}}(X_j)$$

or

$$P(\Delta_{n,s}) \leq \sum_{[n^{1/s}]}^{\infty} \frac{1}{m^s} \leq \frac{C}{n^{1-1/s}}$$

$$|\sum_j Z_j| \leq \frac{\sqrt{2}A}{\sqrt{B_1}} \left(\sqrt{(r(n))^\alpha\, n^{\frac{2}{s}-1}} + \sqrt{\frac{(r(n))^\alpha}{n}} \sum_j 1_{\Delta_{n,s}}(X_j)\right).$$

D'après la condition e), $\lim \frac{(r(n))^\alpha}{n^{1-2/s}} = 0$

d'où

$$P(|\sum_j Z_j| \geq \frac{\sqrt{2}A}{B_1}\ \varepsilon) \leq P(\sqrt{\frac{(r(n))^\alpha}{n}} \sum_j 1_{\Delta_{n,s}}(X_j) > \frac{\varepsilon}{2}) = P_n(\varepsilon)$$

$$\sqrt{\frac{(r(n))^\alpha}{n}} \sum_j 1_{\Delta_{n,s}}(X_j)$$

est une variable aléatoire dont l'espérance et la variance sont majorées par :

$$C\sqrt{(r(n))^\alpha\, n^{\frac{2}{s}-1}}\ .$$

La convergence en probabilité de $\sum_j Z_j$ vers 0 résulte alors d'une inégalité de Bienaymé pour les lois binomiales.

$$P_n(\varepsilon) \leq \frac{16\ C^2}{\varepsilon^2} \frac{(r(n))^\alpha}{n^{1-2/s}}$$

et $P_n(\varepsilon)$ tend vers 0 d'après la condition e)

Démonstration de ii).

Les Y_j sont des variables aléatoires indépendantes, centrées, de même loi.

Nous allons montrer que le théorème central limite - Loève ([10]) p. 316 - s'applique aux variables Y_j.

Pour cela, il suffit de montrer que

α) $\forall\ \varepsilon > 0,\ \sum_{j=1}^{n} P(|Y_j| > \varepsilon)$ tend vers 0 quand n tend vers ∞.

β) $\exists\ \tau > 0$, tel que $\lim_{n\to\infty} \sum_j \sigma_j^2(Y_j(\tau)) = 1$ et $\lim_{n\to\infty} E_P(\sum_{j=1}^{n} Y_j(\tau)) = 0$

où $Y_j(\tau)$ représente la variable Y_j tronquée à τ.

La condition α) est vérifiée, en effet :

$$|Y_j| \leqslant \frac{\sqrt{2}\ A(r(n))^{\alpha}}{n\ (\hat{f}_n(t_o))} \leqslant \frac{2A\ (r(n))^{\alpha}}{\sqrt{B_1}\ \ \sqrt{n}} \to 0$$

et pour n assez grand $Y_j(\tau) = Y_j$

$$\sum_{j=1}^{n} \sigma_j^2(Y_j) = \frac{\int K_{r(n)}^2(x,t) f(x) dx - \bar{f}_n^2(t)}{n\sigma^2(\hat{f}_n(t))} = \frac{1}{1 + \eta_n(t)}$$

La condition β) résulte du lemme 10 et du fait que les variables sont centrées.

Ceci achève la démonstration de la proposition 11 et

$$\frac{\hat{f}_n(t_o) - \bar{f}_o(t_o)}{(\hat{f}_n(t_o))} \xrightarrow{L} \mathcal{N}(0,1)$$

<u>*Remarque*</u> : La condition imposée sur les masses

(25) $$\lim_{m\to\infty} a_m\ m^2 = 0$$

est nécessaire pour assurer la convergence en loi vers $\mathcal{N}(0,1)$.

On va construire un estimateur et une loi de probabilité ne vérifiant pas (25) et pour lesquels, il n'y aura pas convergence de

$$\frac{\hat{f}_n(t_o) - \bar{f}_n(t_o)}{\sigma(\hat{f}_n(t_o))}\ .$$

<u>*Contre-exemple.*</u>

On considère l'ensemble des probabilités définies sur $[-\pi,\pi]$ dont la densité de la partie continue admet un développement en série de Fourier qui converge uniformément sur $[-\pi,\pi]$. Soit $\mathcal{D}$ l'ensemble de ces densités.

On prendra le système des fonctions trigonométriques

$$e_o(t) = \frac{1}{\sqrt{2\pi}}$$

$$e_{2i-1}(t) = \frac{\cos it}{\sqrt{\pi}} \quad ; \quad e_{2i}(t) = \frac{\sin it}{\sqrt{\pi}} \qquad i = 1,2,\ldots$$

et le noyau K_r défini par

$$K_r(x,t) = \sum_{i=0}^{2r} e_i(x)e_i(t) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{i=1}^{r} \cos i(x-t)$$

(26) $$K_r(x,t) = \frac{1}{2\pi} \frac{\sin(r+1/2)(x-t)}{\sin \frac{x-t}{2}}$$

le noyau K_r vérifie les deux relations

(14') $$|K_r(x,t)| \leqslant \frac{r+1/2}{\pi}$$

(15') $$\int_{-\pi}^{\pi} K_r^2(x,t)dx = \frac{2r+1}{2\pi}$$

l'estimateur est alors donné par

(27) $$\hat{f}_n(t) = \frac{1}{2\pi n} \sum_{j=1}^{n} \frac{\sin(r+1/2)(X_j-t)}{\sin \frac{X_j-t}{2}} 1_{S_j}(X_1,\ldots,X_n).$$

Soit P la probabilité définie par

(28) $$dP = \frac{1}{4\pi} d\lambda + \sum_{m=1}^{\infty} \frac{3}{\pi^2 m^2} \delta_{2^{-m}}$$

la densité de la partie continue est uniforme

les masses sont situées aux points d'abscisse 2^{-m}.

On vérifie que

$$\int_{[-\pi,\pi]} dP = \int_{-\pi}^{\pi} \frac{1}{4\pi}\, dx + \sum_{m=1}^{\infty} \frac{3}{\pi^2 m^2} = 1 .$$

On va établir que $\dfrac{\hat{f}_n(0) - \bar{f}_n(0)}{\sigma(\hat{f}_n(0))}$ ne converge pas en loi vers $\mathcal{N}(0,1)$.

$$f(0) = \bar{f}_n(0) = \frac{1}{4\pi}$$

$$\hat{f}_n(0) = \sum_{j=1}^{n} Y_j + \sum_{j=1}^{n} Z_j \qquad \text{p.s.}$$

avec

(29) $$Y_j = \frac{1}{2\pi n \sigma(\hat{f}_n(0))} \frac{\sin(r+\frac{1}{2})X_j}{\sin \frac{X_j}{2}} 1_{E-\Delta}(X_j)$$

(30) $$Z_j = \frac{1}{2\pi n \sigma(\hat{f}_n(0))} \frac{\sin(r+\frac{1}{2})X_j}{\sin \frac{X_j}{2}} 1_{\Delta}(X_j) 1_{S_j}(X_1,\dots,X_n)$$

Comme $\displaystyle\int_{-\pi}^{\pi} K_r^2(x,t) f(x)\,dx = \int_{-\pi}^{\pi} \frac{1}{4\pi^2} \Big(\frac{1}{2} + \sum_{i=1}^{r} \cos ix\Big)^2 dx = \frac{r+1/2}{4\pi^2}$

il vient

$$\sigma^2(\hat{f}_n(0)) = \frac{2r(n)+1}{8n\pi^2}(1+\eta_n) \qquad \eta_n = O\Big(\frac{r(n)}{\sqrt{n}}\Big)$$

La démonstration se fera en deux étapes :

i) $$\sum_{j=1}^{n} Y_j \xrightarrow{L} \mathcal{N}(0,1)$$

ii) $$\lim\Big(P\Big\{\sum_{j=1}^{n} Z_j > 1\Big\}\Big) = 1.$$

Pour la première partie de la démonstration, il suffit de reprendre la proposition 11.

La majoration de $\sigma(\hat{f}_n(0))$ calculée dans le lemme 10 - et, en particulier, la majoration de $G_n(t)$ - est encore valable si la relation $(r(n))\, n^{-1/2} \to 0$ est vérifiée.

Dans la seconde partie, on notera

$$\Delta' = \{d_m | m > [\log_2 n]\} = \{2^{-m} | m > [\log_2 n]\}$$

si $X_j \in \Delta'$ $\quad (r(n))X_j \leqslant \frac{r(n)}{n}$

$$\text{et} \quad Z_j = \frac{1}{2\pi n \sqrt{\frac{r(n)+1/2}{4\pi^2 n}}\,(1+\eta_n)} \quad \frac{\frac{\sin(r(n)+1/2)X_j}{\sin \frac{X_j}{2}}} {} 1_{S_j}(X_1,\ldots,X_n)$$

$$\geqslant \sqrt{\frac{r(n)}{n}}\; 1_{S_j}(X_1,\ldots,X_n) \; . \tag{30'}$$

On va montrer que

$$V_n = \sum_{j=1}^{n} 1_{\Delta-\Delta'}(X_j)\, 1_{S_j}(X_1,\ldots,X_n) \xrightarrow{\text{p.s.}} 0$$

$$E(V_n) = n\, E(1_{\Delta-\Delta'}(X_1)\, 1_{S_j}(X_1,\ldots,X_n))$$

$$= n \sum_{d_m \in \Delta-\Delta'} a_m (1-a_m)^{n-1} \; .$$

$$E(V_n) = n \sum_{m<\log_2 n} \frac{3}{\pi^2 m^2} \left(1 - \frac{3}{\pi^2 m^2}\right)^{n-1}$$

$$\leqslant n \log_2 n \frac{3}{\pi^2 \log_2^2 n} \left(1 - \frac{3}{\pi^2 \log_2^2 n}\right)^{n-1}$$

$$\leqslant \frac{3n}{\pi^2 \log_2 n} \left(1 - \frac{3}{\pi^2 \log_2^2 n}\right)^{n-1}$$

qui est le terme général d'une série convergente.

Or, V_n est à valeurs entières et $P(V_n \geq 1) \leq E(V_n)$ ce qui établit $\lim\limits_{n\to\infty} \text{p.s. } V_n = 0$.

Il suffit donc de montrer que :

$$\lim_n P(W_n \geq \sqrt{r(n)}) = 1$$

où

(31) $$W_n = \sum_{j=1}^{n} 1_{\Delta'}(X_j) 1_{S_j}(X_1,\ldots,X_n)$$

$$E(W_n) = n \sum_{d_m \in \Delta'} a_m(1-a_m)^{n-1} \geq n \sum_{m>\sqrt{\alpha n}} a_m(1-a_m)^{n-1}$$

$$\geq n \sum_{m>\sqrt{\alpha n}} \frac{3}{\pi^2 m^2} \left(1 - \frac{3}{\pi^2 \alpha n}\right)^{n-1}$$

(31') $$\geq \sqrt{n} \frac{3}{\pi^2 \sqrt{\alpha}} \exp\left(-\frac{3}{\pi^2 \alpha}\right) \geq \frac{3}{2} \sqrt{n}$$

la variance de W_n peut être calculée de la même manière

$$W_n^2 = \sum_{j=n}^{n} 1_{\Delta'}(X_j) 1_{S_j}(X_1,\ldots,X_n) + 2 \sum_{1<j<k<n} 1_{\Delta'}(X_j) 1_{\Delta'}(X_k) 1_{S_j \cap S_k}(X_1,\ldots,X_n)$$

$$E(W_n^2) = n \sum_{d_m \in \Delta'} a_m(1-a_m)^{n-1} + n(n-1) \sum_{\substack{m \neq m' \\ d_m \in \Delta' ; d_{m'} \in \Delta'}} a_m a_{m'} (1-a_m-a_{m'})^{n-2}$$

$$\sigma^2(W_n) = E(W_n^2) - (EW_n)^2$$

une transformation analogue à celle de $G_n(t)$ (23) et (24) dans la démonstration du lemme 10 montre que

$$\sigma^2(W_n) = O(n\, L_1(n)^2 + n^2 L_2(n) + n\, L_1(n) + (n-2)n^2 L_2(n)^2$$

avec $L_1(n) = O(\frac{1}{\sqrt{n}})$ et $L_2(n) = O(\frac{1}{n^{3/2}})$ - lemmes 7 et 8 -

d'où $\sigma^2(W_n) = O(\sqrt{n})$; en utilisant une inégalité de Markov, il vient :

$$P(W_n < \sqrt{n}) \to 0 \qquad \text{(d'après (31'))}$$

comme

$$\sum_{j=1}^{n} Z_j - \sqrt{\frac{r(n)}{n}}\, W_n \geq 0 \quad \text{p.s.} \qquad (30')$$

$$\lim_n P\left(\sum_{j=1}^{n} Z_j > \sqrt{r(n)}\right) = 1$$

ce qui achève la partie ii) de la démonstration.

Conclusion.

$$\left\{\sum_{j=1}^{n} Y_j + \sum_{j=1}^{n} Z_j > 0\right\} \supset \left\{\sum_{j=1}^{n} Y_j > -1\right\} \cap \left\{\sum_{j=1}^{n} Z_j > 1\right\}$$

on en déduit, en utilisant

$$\lim_n P\left(\sum_j Z_j > 1\right) = 1$$

que

$$\lim P\left(\sum_{j=1}^{n} (Y_j + Z_j) > 0\right) \geq \lim_n P\left(\sum_{j=1}^{n} Y_j > -1\right)$$

$$\geq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+1} e^{-\frac{u^2}{2}}\, du \qquad \text{- d'après i) -}$$

et $\dfrac{\hat{f}_n(0) - \bar{f}_n(0)}{\sigma(\hat{f}_n(0))}$ ne converge pas en loi vers $\mathcal{N}(0,1)$.

CHAPITRE III

ETUDE DE QUELQUES VITESSES DE CONVERGENCE

Dans cette partie, nous nous limiterons au cas où les probabilités sont définies sur l'intervalle $[-\pi,\pi]$. La densité de la partie continue sera donnée par son développement en série de Fourier. On recherchera un ordre de grandeur de $r(n)$ qui minimise l'erreur faite en estimant $f(t)$ par $\hat{f}_n(t)$.

L'estimateur est défini par

$$\hat{f}_n(t) = \sum_{i=0}^{r(n)} \left(\sum_{j=1}^{n} \frac{1}{n} e_i(X_j) 1_{S_j}(X_1,\ldots,X_n) e_i(t) \right.$$

avec
$$e_o(t) = \frac{1}{\sqrt{2\pi}}$$

$$e_{2i-1}(t) = \frac{\cos it}{\sqrt{\pi}} \qquad e_{2i}(t) = \frac{\sin it}{\sqrt{\pi}} \qquad i = 1,2,\ldots$$

On peut supposer que le coefficient $r(n)$ est pair $\hat{f}_n(t)$ s'écrit encore

$$\hat{f}_n(t) = \frac{1}{2\pi n} \sum_{j=1}^{n} \frac{\sin \frac{(r(n)+1)}{2}(X_j-t)}{\sin \frac{X_j-t}{2}} 1_{S_j}(X_1,\ldots,X_n).$$

On définit le M.I.S.E. par

$$J_n = E_P\left(\int_E (\hat{f}_n(t) - f(t))^2 dt\right).$$

En notant b_i les coefficients de Fourier de f

$$b_i = \int_E f(t) e_i(t) dt$$

et $$\hat{b}_i = \frac{1}{n} \sum_{j=1}^{n} e_i(X_j) 1_{S_j}(X_1,\dots,X_n)$$

les e_i sont des fonctions orthogonales, il vient donc

$$J_n = \sum_{i=0}^{r(n)} E_p((\hat{b}_i - b_i)^2) + \sum_{i>r(n)} b_i^2$$

$$J_n = \sum_{i=0}^{r(n)} E_P(\frac{1}{n^2} [\sum_{j=1}^{n} (e_i(X_j) 1_{E-\Delta}(X_j) - b_i) + \sum_{j=1}^{n} e_i(X_j) 1_\Delta(X_j) 1_{S_j}(X_1,\dots,X_n)]^2)$$
$$+ \sum_{i>r(n)} b_i^2 .$$

Un calcul analogue à celui du lemme 4 - relation (12) - conduit à

$$J_n = \frac{1}{n} \int_E f(t) \sum_{i=0}^{r(n)} e_i^2(t)dt - \frac{1}{n} \sum_{i=0}^{r(n)} b_i^2 + \sum_{i>r(n)} b_i^2$$
$$+ \sum_{i=0}^{r(n)} \frac{1}{n} \sum_m a_m(1-a_m)^{n-1} e_i^2(d_m)$$
$$+ \sum_{i=0}^{r(n)} \frac{n-1}{n} \sum_{m\neq m'} a_m a_{m'}(1-a_m-a_{m'})^{n-2} e_i(d_m) e_i(d_{m'})$$
$$+ \sum_{i=0}^{r(n)} 2\,\frac{n-1}{n} \sum_m a_m^2(1-a_m)^{n-2} e_i(d_m) b_i$$

et en remplaçant, dans le cas particulier des fonctions trigonométriques

$$J_n = \frac{1+r(n)}{2\pi n} \int_{-\pi}^{\pi} f(t)dt - \frac{1}{n} \sum_{i=0}^{r(n)} b_i^2 + \sum_{i>r(n)} b_i^2$$
$$+ \frac{1+r(n)}{2\pi n} L_1(n) + 2\,\frac{n-1}{n} \sum_m a_m^2(1-a_m)^{n-2} \bar{f}_n(d_m)$$
$$+ \frac{n-1}{2\pi n} \sum_{m\neq m'} a_m a_{m'}(1-a_m-a_{m'})^{n-2} \frac{\sin \frac{r(n)+1}{2}(d_m-d_{m'})}{\sin \frac{d_m-d_{m'}}{2}} .$$

En regardant l'ordre de grandeur de chacun des termes qui figurent dans la somme précédente

$$\frac{1+r(n)}{n}\int_{-\pi}^{\pi} f(t)dt = 0(\frac{r(n)}{n})$$

l'ordre de $\sum\limits_{i>r(n)} b_i^2$ dépend de la vitesse de convergence de la série de Fourier $\sum\limits_i b_i e_i(t)$ vers $f(t)$.

Le dernier terme est de l'ordre de $r(n)(L_1(n))^2$.

Tous les autres termes sont d'ordre inférieur à $\frac{r(n)}{n}$.

On en déduit les résultats suivants :

S'il existe un nombre $s > 2$, tel que $\lim m^2 a_m = 0$

$$J_n = 0(\frac{r(n)}{n}) + \sum_{i>r(n)} b_i^2 .$$

Si $b_k = 0(\frac{1}{k^p})$, $p > \frac{1}{2}$,

et si l'on prend $r(n) = 0(n^{1/2p})$, $J_n = 0(\frac{1}{n^{1-\frac{1}{2p}}})$.

En effet, si $b_k = 0(\frac{1}{k^p})$, $\sum\limits_{i>r(n)} b_k^2 = 0(\frac{1}{(r(n))^{2p-1}})$.

Sinon, on supposera qu'il existe un nombre s, $1 < s < 2$ tel que $a_m = 0\ (m^{-s})$

$$J_n = 0(r(n)(L_1(n))^2) + \sum_{i>r(n)} b_i^2$$

si $b_k = 0(\frac{1}{k^p})$, $p > \frac{1}{2}$,

$$J_n = 0(r(n)n^{\frac{2}{s}-2}) + 0(\frac{1}{(r(n))^{2p-1}})$$

et si l'on prend

$$r(n) = 0(n^{\frac{1}{p}(1-\frac{1}{s})}),\quad J_n = 0(\frac{1}{n^{\frac{2p-1}{p}(1-\frac{1}{s})}}) .$$

Une autre manière de mesurer "l'erreur" faite en estimant $f(t)$ par $\hat{f}_n(t)$ est de définir la quantité

$$Q_n = \sup_{t\in E}(E_p[\hat{f}_n(t) - f(t)]^2)$$

$$Q_n = \sup_{t\in E}(E_p[\hat{f}_n(t) - E_p(\hat{f}_n(t))]^2 + E_p(\hat{f}_n(t) - f(t))^2)$$

$$Q_n = \sup_{t\in E}(\sigma^2(\hat{f}_n(t)) + \left[\sum_{i=0}^{r(n)} (\sum_m a_m(1-a_m)^{n-1} e_i(d_m))e_i(t) - \sum_{i>r(n)} b_i e_i(t)\right]^2).$$

Nous nous limiterons au cas des séries de Fourier, et nous supposerons que la densité f est bornée.

Le lemme 10 nous donne l'ordre de grandeur de $\sigma^2(\hat{f}_n(t))$ lorsque la condition $r(n)n^{1/s-1} \to 0$ est remplie

$$\sigma^2(\hat{f}_n(t)) = O(\frac{r(n)}{n})$$

d'où

L'erreur en moyenne quadratique uniforme

$$Q_n = O(\frac{r(n)}{n}) + (r(n)L_1(n) + \sum_{i>r(n)} |b_i|)^2 .$$

On peut préciser l'ordre de grandeur de Q_n si $b_k = O(\frac{1}{k^p})$, $p > 1$, si les masses a_m décroissent suffisamment vite, c'est-à-dire si

$$\lim a_m m^s = 0 \quad \text{avec} \quad s > \frac{2p-1}{p-1}$$

alors pour

$$r(n) = O(n^{\frac{1}{2p-1}}) \qquad Q_n = O(\frac{1}{n^{1 - \frac{1}{2p-1}}})$$

CHAPITRE IV

ETUDE DE LA CONVERGENCE DE L'ESTIMATEUR DANS LE CAS DE VARIABLES DEFINISSANT UN PROCESSUS Φ-MELANGEANT.

4.1. - Définitions.

Dans cette partie, on abandonne l'hypothèse : les variables X_j sont indépendantes et on suppose qu'elles définissent un processus stationnaire et ϕ-mélangeant.

Les définitions sont celles de Billingsley [11].

La suite est stationnaire : $\forall\, k \in N$, $\forall (t_1,t_2,\ldots,t_k) \in N^k$, $\forall\, \ell \in N$ la loi de $(X_{t_1},X_{t_2},\ldots,X_{t_k})$ est la même que la loi de $(X_{t_1+\ell},\ldots,X_{t_k+\ell})$.

Pour $k \in N$, on note M_1^k et M_k^∞ les tribus engendrées respectivement par les variables $(X_1,X_2,\ldots,X_k)$ et $(X_k,X_{k+1},\ldots)$.

Soit ϕ une fonction positive, définie sur N, vérifiant $\phi(0) = 1$ la suite est ϕ-mélangeante si

$$\forall\, k \in N,\ \forall\, n \geq 1,\ \forall E_1 \in M_1^k,\ \forall\, E_2 \in M_{k+n}^\infty\ |P(E_1 \cap E_2)-P(E_1)P(E_2)| \leq \phi(n)P(E_1)$$

on ne considérera que des fonctions ϕ décroissantes, telles que

$$\lim_{n\to\infty} \phi(n) = 0.$$

Deux lemmes classiques seront utilisés.

Lemme 11.- Si Y_1 est M_1^k mesurable et Y_2 M_{k+n}^∞ mesurable $(n \geq 1)$ alors $E(|Y_1|^r)<\infty$; $E(|Y_2|^s)<\infty$; $r > 1$, $s > 1$, $\frac{1}{r} + \frac{1}{s} = 1$ implique

$$|E(Y_1Y_2) - E(Y_1)E(Y_2)| \leq 2(\phi(n))^{1/r}(E(|Y_1|^r))^{1/r}(E(|Y_2|^s))^{1/s}$$

Lemme 12. - On suppose que la suite des variables X_j est stationnaire et ψ-mélangeante, que

$$E(X_1) = 0 \quad \text{et} \quad \sum_{k=1}^{\infty} |E(X_1 X_k)| < \infty$$

alors $\frac{1}{n} E(S_n^2) \leqslant 2 \sum_{k=1}^{\infty} E(X_1 X_k)$ et $\frac{1}{n} E(S_n^2) \to \sigma^2$

où $\sigma^2 = E(X_1^2) + 2 \sum_{k=2}^{\infty} E(X_1 X_k)$ et $S_n = \sum_{k=1}^{n} X_k$.

Enoncé et démonstration : Billingsley ([11]) p. 170-173.

Nous avons utilisé dans la démonstration de la proposition 9, une inégalité du type de Bernstein pour des variables indépendantes. M. Bosq en a donné un énoncé pour les processus stationnaires et mélangeants.

Lemme 13. - Soit $(X_j)_{j \geqslant 1}$, une suite de variables réelles qui forme un processus stationnaire et ψ-mélangeant.

On suppose que

$$|X_1| \leqslant d, \quad E(X_1) = 0, \quad E(X_1^2) > 0$$

$\sigma^2 = E\, X_1^2 + 2 \sum_{\nu > 1} E\, X_1 X_\nu$ est fini et strictement positif.

$k = k(n)$ est une suite d'entiers vérifiant $\lim k(n) = \lim \frac{n}{k(n)} = \infty$.

On pose $S_n = \sum_{j=1}^{n} X_j$ alors

$$P(|S_n| > \varepsilon\sigma(S_n)) \leqslant 2e\, e^{-\frac{\tau\varepsilon}{2}} (e^{\frac{3}{2}\tau^2} + \psi(k) e^{n/k})$$

$$\varepsilon > 0, \quad 0 < \tau \leqslant \frac{\sigma(S_n)}{4kd}, \; n \text{ assez grand.}$$

Lemme 14. - Sous les conditions du lemme précédent et si en outre $\psi(n) = a\rho^n$, $a > 0$, $0 < \rho < 1$, $n \geqslant 1$, on a pour n assez grand

$$\forall \eta \geq 0, \quad P\left(\left|\frac{S_n}{n}\right| \geq \eta\right) \leq 4e \exp\left[-\frac{(\mathrm{Log}\ 1/\rho)^{1/2}\eta}{8d(1+\sqrt{2})}\sqrt{n}\right] .$$

Ces deux lemmes sont dus à M. Bosq ([12]).

4.2. - *Choix d'un estimateur* - *Etude de la convergence presque complète.*

L'estimateur $\hat{f}_n(t)$ défini dans le chapitre II ne converge pas nécessairement vers $f(t)$ lorsque les variables X_j sont dépendantes, même si elles vérifient des conditions de mélangeance.

En effet, si les Z_j sont des variables aléatoires indépendantes, de même loi, la suite de variables définie par :

$Z_1, Z_1, Z_2, Z_2, \ldots, Z_n, Z_n, Z_{n+1}, \ldots$ est stationnaire et φ-mélangeante (on peut prendre $\varphi(0) = \varphi(1) = 1$, $\varphi(n) = 0$ pour $n \geq 2$).

L'estimateur $\hat{f}_n$ est tel que $\hat{f}_{2i}(t) = 0$.

On est donc amené à modifier les critères du chapitre II et à ne garder comme masses que les points sortis deux fois au moins, à des moments suffisamment éloignés.

On peut cependant exclure les cas "pathologiques" et supposer que la probabilité d'obtenir deux fois un point qui n'appartient pas à l'ensemble des masses est nulle.

Dans ce cas, on va montrer que l'estimateur défini dans la partie II par (8)

$$\hat{f}_n(t) = \frac{1}{n} \sum_{j=1}^{n} K_{r(n)}(X_j, t) 1_{S_j}(X_1, X_2, \ldots, X_n)$$

converge presque complètement vers $f(t)$ si les hypothèses suivantes sont vérifiées.

Proposition 12. - Soit (X_j) une suite de variables aléatoires stationnaire et φ-mélangeante, la fonction φ vérifiant

$$\psi(n) = a\rho^n, \quad a > 0, \quad 0 < \rho < 1, \quad \text{pour} \quad n \quad \text{assez grand.}$$

Si les hypothèses suivantes sont vérifiées

(14) $\forall\, t \in E, \ \exists A(t) \qquad |K_r(x,t)| \leqslant A(t) r^\alpha$

(15) $\forall\, t \in E, \ \exists B(t) \quad \int_E K_r^2(x,t)dx \leqslant B(t) r^\alpha$

α) les densités de $\mathcal{D}$ sont bornées

β) l'intégrale $\int_E K_r(x,t)f(x)dx$ converge vers $f(t)$ lorsque $r \to \infty$.

Si en outre

γ) $P(X_i = X_j, \ X_i \in E-\Delta, \ i \neq j) = 0$

alors

Pour toute suite (d_m, a_m) on peut construire une suite $r(n)$ telle que l'estimateur $\hat{f}_n(t)$ converge p.c_o vers $f(t)$.

Si la suite a_m vérifie $\exists \gamma > 1, \ \lim \gamma^m a_m = 0$ il suffit que $\dfrac{(r(n))^\alpha \text{Log } n}{\sqrt{n}} \to 0$.

Si la suite a_m vérifie $\exists s > 1, \ \lim m^s a_m = 0$ il suffit que $\dfrac{(r(n))^\alpha \text{Log } n}{n^{\frac{s-1}{2s-1}}} \to 0$.

Démonstration :

La démonstration reprend en grande partie celle de la proposition 9.

$f_n(t) - f(t) = I_1 + I_2 + I_3$

avec

$$I_1 = \sum_{j=1}^{n} \frac{1}{n} K_{r(n)}(X_j,t) 1_{E-\Delta}(X_j) 1_{S_j}(X_1,\dots,X_n) - \int_E K_{r(n)}(x,t)f(x)dx$$

(16) $$I_2 = \sum_{j=1}^{n} \frac{1}{n} K_{r(n)}(X_j,t) 1\ (X_j) 1_{S_j}(X_1,\dots,X_n)$$

$$I_3 = \int_E K_{r(n)}(x,t)f(x)dx - f(t)$$

Convergence $p.c_o$ de I_1.

On sait que $I_1 = \frac{1}{n} \sum_{j=1}^{n} Y_j$ p.s.

(32) $$Y_j = K_{r(n)}(X_j,t)\,1_{E-\Delta}(X_j) - \int_E K_{r(n)}(x,t)f(x)dx.$$

La suite des variables Y_j est stationnaire et $\emptyset$-mélangeante. La stationnarité résulte de la même propriété pour les X_j et la fonction K_r étant mesurable, la tribu engendrée par les Y_j est contenue dans celle engendrée par les X_j, d'où la $\emptyset$-mélangeance des Y_j.

Les conditions du lemme 14 sont vérifiées

$$|Y_1| \leq d = 2\,A(t)\,(r(n))^{\alpha} \qquad E(Y_1) = 0$$

$$E(Y_1^2) = \int_E K^2_{r(n)}(x,t)f(x)dx - \bar{f}_n^2(t) \leq B(t)M(r(n))^{\alpha}$$

$$\sigma^2 = E(Y_1^2) + 2 \sum_{\nu>1} E(Y_1 Y_\nu)$$

et en utilisant le lemme 11 - avec $r = s = 2$ -

$$\sigma^2 \leq E(Y_1^2)(1 + 4 \sum_{\nu>1} \sqrt{\emptyset(\nu)})$$

$$\sigma^2 \leq (r(n))^{\alpha}\beta^2 \qquad \text{où } \beta^2 = MB(t)(1 + 4 \sum_{\nu} \sqrt{\emptyset(\nu)})$$

il vient :

$$P(I_1 > \varepsilon) \leq 4e \exp\left[- \frac{\text{Log } 1/\rho)^{1/2}\ \varepsilon\sqrt{n}}{16(1+\sqrt{2})A(t)(r(n))^{\alpha}}\right].$$

Il suffit pour assurer la convergence $p.c_o$ de I_1 vers 0 que

$$\lim_{n\to\infty} \frac{(r(n))^{\alpha}}{\sqrt{n}} \text{ Log } n = 0.$$

Convergence de I_2.

Les notations reprennent celles de la proposition 9.

k étant un entier positif, on désigne par N_k le plus petit entier tel que

$$\sum_{m \geq N_k+1} a_m \leq \frac{1}{k} \quad \text{et} \quad \Delta_k = \{d_{N_k+1}, d_{N_k+2}, \ldots\}$$

I_2 convergera vers 0 presque complètement si on peut construire une suite $k(n) \to \infty$ telle que

$$\frac{(r(n))^\alpha}{k(n)} \to 0$$

et

$$\sum_n P\Big(\sum_{j=1}^{n} \frac{1}{n} K_{r(n)}(X_j,t)\,1_\Delta(X_j)\,1_{S_j}(X_1,\ldots,X_n) \geq \frac{3\,A(t)(r(n))^\alpha}{k(n)}\Big) < \infty$$

Or $$\sum_{j=1}^{n} 1_\Delta(X_j)\,1_{S_j}(X_1,\ldots,X_n) \leq N_k + \sum_{j=1}^{n} 1_{\Delta_k}(X_j)$$

et

(19) $$I_2 \leq \frac{A(t)(r(n))^\alpha}{n}\Big(N_k + \sum_{j=1}^{n} 1_{\Delta_k}(X_j)\Big).$$

Les variables $1_{\Delta_k}(X_j)$ forment une suite stationnaire, φ-mélangeante, d'espérance $\leq \frac{1}{k}$.

Le lemme 14 peut s'appliquer avec $d = 1$

(33) $$P\Big(\sum_{j=1}^{n} 1_{\Delta_k}(X_j) \geq \frac{2}{k(n)}\Big) \leq P\Big(\sum_{j=1}^{n} (1_{\Delta_k}(X_j) - E(1_{\Delta_k}(X_j))) \geq \frac{1}{k(n)}\Big)$$

$$\leq 4\,e\,\exp\Big(-\frac{(\mathrm{Log}\ 1/\rho)^{1/2}\sqrt{n}}{8(1+\sqrt{2})\,k(n)}\Big)$$

et dès que $k(n)N_{k(n)} \leq n$

(33') $$P\Big(I_2 \geq \frac{3\,A(t)(r(n))^\alpha}{k(n)}\Big) \leq 4\,e\,\exp\Big(-\frac{(\mathrm{Log}\ 1/\rho)^{1/2}\sqrt{n}}{8(1+\sqrt{2})\,k(n)}\Big).$$

La convergence presque complète résulte de l'existence de deux suites $k(n)$ et $r(n)$ vérifiant :

$$(34)\qquad k(n) \to \infty \quad ; \quad r(n) \to \infty \quad ; \quad k(n)N_{k(n)} \leq n$$

$$\frac{(r(n))^{\alpha}}{k(n)} \to \infty \quad ; \quad \frac{k(n)\,\mathrm{Log}\, n}{\sqrt{n}} \to 0.$$

Ces conditions sont analogues aux conditions (20) de la proposition 9 au remplacement près de n par $\sqrt{n}$. La démonstration s'achève de la même manière.

Remarque.

La condition $\gamma)$ exprimait que la probabilité de sortir deux fois un point qui n'appartient pas à l'ensemble des masses est nulle.
Cette condition peut être abandonnée mais il faut modifier l'estimateur.

On choisira comme estimateur $\hat{\Delta}_n$ de Δ :

$$\hat{\Delta}_{n,\nu(n)} = \{X_j \mid \exists k < j-\nu(n) \quad X_k = X_j\}$$

où $\nu(n)$ est un entier positif et l'on peut supposer $\lim\limits_{n\to\infty} \nu(n) = \infty$
c'est-à-dire on ne considère comme masses que les points sortis deux fois à des moments suffisamment éloignés.

Si on pose :

$$R_{j,\nu(n)} = \{(x_1,x_2,\ldots,x_j) \in \mathbb{R}^j \mid x_j \neq x_1 \quad \text{et} \quad x_j \neq x_2 \ldots \text{ et } x_j \neq x_{j-\nu(n)}\}$$

on prendra comme estimateur de la densité

$$\hat{f}_n(t) = \frac{1}{n} \sum_{j=1}^{n} K_{r(n)}(X_j,t)\, 1_{R_{j,\nu}}(X_1,\ldots,X_j).$$

On peut alors construire des suites $r(n)$ et $\nu(n)$ telles que l'estimateur $\hat{f}_n(t)$ converge presque complètement vers $f(t)$ sous les hypothèses de la proposition 12, la condition $\gamma)$ étant abandonnée.

Démonstration :

On décompose $\hat{f}_n(t) - f(t)$ sous la forme

$$\hat{f}_n(t) - f(t) = I_1 + I_2' + I_2'' + I_3.$$

Les notations I_1 et I_3 désignent les mêmes quantités que précédemment

$$I_2' = \frac{1}{n} \sum_{j=1}^{n} K_{r(n)}(X_j,t) 1_{E-\Delta}(X_j)(1 - 1_{R_{j,\nu}}(X_1,\dots,X_j))$$

$$I_2'' = \frac{1}{n} \sum_{j=1}^{n} K_{r(n)}(X_j,t) 1_{\Delta}(X_j) 1_{R_{j,\nu}}(X_1,\dots,X_j).$$

Convergence de I_2'.

$1_{E-\Delta}(X_j)(1 - 1_{R_{j,\nu}}(X_1,\dots,X_j))$ ne prend que les valeurs 0 et 1, donc

$$P(I_2' \neq 0) \leqslant \sum_{j=1}^{n} P(X_j \in E-\Delta, X_j = X_1 \text{ ou } \dots \text{ ou } X_j = X_{j-\nu})$$

$$\leqslant \sum_{j=1}^{n} \sum_{h=1}^{j-\nu} P(X_j \in E-\Delta, X_j = X_h).$$

La restriction de P à $E-\Delta$ est non atomique, il existe donc une partition de $E-\Delta$ en ensembles A_ε de probabilité $\leqslant \varepsilon$.

Il vient :

$$P(I_2' \neq 0) \leqslant \sum_{j=1}^{n} \sum_{h=1}^{j-\nu} \sum_{A_\varepsilon} P(X_j \in A_\varepsilon \text{ et } X_h \in A_\varepsilon).$$

En utilisant la définition de la ψ-mélangeance,

$$P(I_2' \neq 0) \leqslant \sum_{j=1}^{n} \sum_{h=1}^{j-\nu} P(E-\Delta)(\varepsilon + \psi(j-h)).$$

Comme ε est arbitraire, et en utilisant le fait que la fonction ψ est décroissante, il vient :

(35) $$P(I_2' \neq 0) \leqslant \frac{n^2}{2} \psi(\nu(n)).$$

Convergence de I_2''.

$$I_2'' \leqslant \frac{A(t)(r(n))^{\alpha}}{n} \sum_{j=1}^{n} 1_{\Delta}(X_j) 1_{R_{j,\nu}}(X_1,\dots,X_j)$$

$$\sum_{j=1}^{n} 1_{\Delta}(X_j) 1_{R_{j,\nu}}(X_1,\dots,X_j) = \sum_{j=1}^{n} 1_{\Delta_k}(X_j) 1_{R_{j,\nu}}(X_1,\dots,X_n) + \sum_{j=1}^{n} 1_{\Delta-\Delta_k}(X_j) 1_{R_{j,\nu}}(X_1,\dots,X_n)$$

La première somme est majorée par $\sum_{j=1}^{n} 1_{\Delta_k}(X_j)$.

La seconde par $\nu(n)N_k$ (*).

La relation (19) est remplacée par

(19') $$I_2'' \leq \frac{A(t)(r(n))^{\alpha}}{n}\left(\nu(n)N_{k(n)} + \sum_{j=1}^{n} 1_{\Delta_k}(X_j)\right).$$

La relation (33') est vérifiée dès que

$$\nu(n)k(n)N_{k(n)} \leq n$$

(33') $$P\left(I_2 \geq \frac{3A(t)(r(n))^{\alpha}}{k(n)}\right) \leq 4\ e \exp\left(- \frac{(\mathrm{Log}\ 1/\rho)^{1/2}\sqrt{n}}{8(1+\sqrt{2})k(n)}\right).$$

La convergence sera établie s'il existe des suites $\nu(n)$, $k(n)$, $r(n)$ vérifiant :

(36)
(a) $k(n) \to \infty$; $r(n) \to \infty$; $\nu(n) \to \infty$;

(b) $\sum n^2 \vartheta(\nu(n)) < \infty$

(c) $\nu(n)k(n)N_{k(n)} \leq n$

(d) $\dfrac{k(n)\mathrm{Log}\ n}{\sqrt{n}} \to 0$ (e) $\dfrac{(r(n))^{\alpha}}{k(n)} \to 0.$

On peut choisir $\nu(n) = 4\,\dfrac{\mathrm{Log}\ n}{\mathrm{Log}\ 1/\rho}$.

Sachant que $\vartheta(n) = a\rho^n$; (36 b) est vérifié.

Les relations (c) et (d) permettent de construire $k(n)$.

En particulier, s'il existe $s > 1$, tel que $\lim_m a_m m^s = 0$, on peut choisir $k(n) = 0\left(\dfrac{n^{\beta}}{\mathrm{Log}\ n}\right)$ $\beta = \inf\left(\dfrac{s-1}{s}, \dfrac{1}{2}\right)$.

On achève la démonstration en prenant une suite $r(n)$ qui vérifie (a) et (e).

(*) Si $d \in \Delta-\Delta_k$, si $\forall\, h \leq j-\nu$ $X_h \neq d$ et $X_{j-\nu+1} = \dots = X_j = d$

alors $1_{R_j}(X_1,\dots,X_j) = \dots = 1_{R_{j-\nu+1}}(X_1,\dots,X_{j-\nu+1}) = 1$

et $\forall\, p > 0$, si $X_{j+p} = d$, $1_{R_{j+p}}(X\ ,\dots,X_{j+p}) = 0$

QUELQUES REMARQUES SUR LES METHODES UTILISEES.

a) Sur le choix de l'estimateur.

Nous avons choisi un estimateur de la densité de la partie continue $\hat{f}_n$ symétrique par rapport aux variables. Les calculs de l'espérance, de la variance, et plus généralement l'étude de la convergence de l'estimateur en ont été facilités. Comme dans toute méthode d'estimation, l'adjonction d'un point supplémentaire à l'échantillon modifie le coefficient $\frac{1}{n}$ en $\frac{1}{n+1}$, et on supposera le plus souvent que $r(n+1) = r(n)$, c'est-à-dire que le même noyau est utilisé, qu'il y ait n points ou $n+1$ points dans l'échantillon. Mais ici, le choix de l'estimateur présente un inconvénient : si le $n+1$-ème point est une masse qui n'était apparue qu'une seule fois dans l'échantillon, le terme correspondant $K_{r(n)}(X_{n+1},t)$ doit être enlevé.

Il est commode de ne comparer un élément X_j de l'échantillon qu'à ceux qui le précèdent, le calcul pratique de l'estimateur serait alors du type suivant :

- si X_j apparaît pour la 1ère fois, on ajoute le terme correspondant $K_{r(n)}(X_j,t)$ à la somme.
- si X_j apparaît pour la 2ème fois, on retranche ce terme
- si X_j apparaît pour la ième fois, $i \geqslant 2$, le terme $K_{r(n)}(X_j,t)$ ne figure pas dans la somme.

Le procédé pourrait être simplifié de la manière suivante :
Ne figurent dans la somme que les termes correspondants aux X_j qui apparaissent pour la première fois.

Ce qui revient à choisir l'estimateur défini par :

$$\hat{f}_n(t) = \frac{1}{n} \sum_{j=1}^{n} K_{r(n)}(X_j,t)\, 1_{R_j}(X_1,\ldots,X_j) \qquad (37)$$

où R_j est l'ensemble $\{(x_1,x_2,\ldots,x_j) \in \mathbb{R}^j \mid x_j \neq x_1 \text{ et } \ldots\ x_j \neq x_{j-1}\}$

L'étude de la convergence suivant divers modes de ce dernier estimateur n'a pas été faite, cependant la démonstration de la proposition 9 montre également qu'il converge presque complètement vers la densité.

Il pourrait être intéressant, dans les calculs pratiques d'introduire des estimateurs "séquentiels" du type

$$\hat{f}_n(t) = \frac{1}{n} \sum_{j=1}^{n} K_{r(j)}(X_j,t)\, 1_{R_j}(X_1,\ldots,X_j). \tag{37'}$$

b) Précision des mesures.

Dans la pratique, les observations sont faites avec une certaine précision. Pour que les méthodes précédentes s'appliquent, il faut un critère permettant de dire qu'un point est une masse.

Deux éléments sont connus : l'intervalle E décrit par les observations et la précision des mesures.

On peut donc se donner une partition de E en intervalles de diamètre η et considérer que deux points coïncident si leur distance est inférieure à η. Soit B_i cette partition.

Des résultats intéressants ne semblent pouvoir être obtenus que si les conditions suivantes sont vérifiées :

- les masses décroissent suffisamment vite, ou encore, la proportion d'intervalles B_i qui contiennent des masses non négligeables est faible.

- on suppose que la probabilité pour que deux points qui ne sont pas des masses tombent dans le même intervalle B_i est faible.

En particulier, si le nombre d'observations était grand (de l'ordre du nombre d'intervalles B_i), il faudrait modifier le critère et ne comparer une observation qu'aux k précédentes (k petit par rapport au nombre d'intervalles B_i).

BIBLIOGRAPHIE

(11) BILLINGSLEY - *Convergence of probability measures*, Wiley, 1968.

(5) BLEUEZ - *Conditions nécessaires et suffisantes de convergence pour une classe d'estimateurs de la densité*, (Thèse 3ème cycle) - Lille, 1976.

(4) BOSQ - *Contribution à la théorie de l'estimation fonctionnelle*, Publ. Inst. Stat. de l'Université de Paris, 1970.

(8) BOSQ - *Sur l'estimation sans biais d'un paramètre fonctionnel*, Publ. Int. USTL Lille I, 1977 n° 104.

(12) BOSQ - *Inégalité de Bernstein pour les processus stationnaires et mélangeants. Applications*. CRASP tome 281 (1975) série A p. 1095-1098.

(7) CAMBANIS - *Bases in* L^2 *spaces with applications to stochastic processes with orthogonal increments*, Proc. of the Amer. Math. Soc. Vol. 29 n° 2 (1971).

(2) CENCOV - *Evaluation of an unknown distribution density from observations*, Soviet Math. 3 p. 1559-1562 (1962).

(13) DELECROIX - *Applications d'une inégalité de Stein à des lois limites d'estimateurs de la densité*, Publ. Int. USTL - Lille I, 1976 n° 93.

(14) DEVOLDER - *Inégalité de type exponentiel*, Publ. Int. de l'U.E.R de Maths pures et appliquées, USTL, Lille I, 1976 n° 98.

(6) HALMOS - *Measure theory*, Van Nostrand Princeton, 1950.

(10) LOEVE - *Probability theory*, Van Nostrand Princeton.

(9) PETROV - *Sums of independent random variables*, Springer Verlag, 1975.

(1) ROSENBLATT - *Remark on some non parametric estimates of a density function*, AMS 27 p. 832-837 (1956).

(3) WAHBA - *Interpolating splines methods for density estimation (I-Equispaced knots)*, Annals of Statistics p. 30-48 (1975).