STATISTIQUE ET ANALYSE
DES DONNÉES

M. DELECROIX
M. E. NOGUEIRA
A. C. ROSA

Sur l'estimation de la densité d'observations ergodiques

Statistique et analyse des données, tome 16, nº 3 (1991), p. 25-38
<http://www.numdam.org/item?id=SAD_1991__16_3_25_0>

Statistique et Analyse des Données
1992 - Vol. 16 n° 3 p. 25 - 38.

SUR L'ESTIMATION DE LA DENSITE D'OBSERVATIONS ERGODIQUES

M. DELECROIX

G.R.E.M.A.Q. (U.R.A. CNRS 947)
Université des Sciences Sociales de Toulouse
Place Anatole France - 31042 Toulouse cedex

Laboratoire de Statistique et Probabilités
(U.R.A. C.N.R.S. D0745)
Université Paul Sabatier, 118 route de Narbonne
31062 Toulouse cedex

M.E. NOGUEIRA et A.C. ROSA

Departamento de Matematica
F.C.T.U.C.
3049 Coimbra Codex

Résumé.

On démontre dans cet article la convergence uniforme, presque sûre et en moyenne, d'une classe d'estimateurs récursifs de la densité commune d'observations X_t, quand celles-ci forment un processus ergodique.

Abstract.

In this paper we prove the mean uniform convergence and the a.e. uniform convergence of a family of recursive estimates of the common density of a set of observations X_t, when X_t is a realization of an ergodic process.

Classification A.M.S. - 62G05, 62M09.

Mots clés.

estimateur récursif de la densité, méthode du noyau, processus ergodiques.

Key words.

Density's recursive estimators, Kernel's method, ergodic processes.

(Ce travail a bénéficié de l'appui du C.M.U.C. (I.N.I.C.) et du projet CEN-MAT n° 395/90).

Manuscrit reçu le 10 novembre 1991, révisé le 20 juillet 1992.

I. INTRODUCTION ET HYPOTHESES GENERALES.

Dans cet article nous nous proposons d'étudier les propriétés de convergence uniforme, presque sûre (p.s.) et en moyenne, d'estimateurs de la densité f de variables X_t, $t \in \mathbb{Z}$, formant un processus stationnaire à valeurs dans $\mathbb{R}^d$ ($\mathbb{Z}$ représente ici l'ensemble des entiers). Ces estimateurs sont construits par la méthode du noyau de Parzen-Rosenblatt (Parzen (1962), Rosenblatt (1956)), et s'écrivent

$$\hat{f}_n(x) = \sum_{t=1}^{n} \beta_{t,n}\, h_t^{-d}\, K((x-X_t)/h_t), \quad x \in \mathbb{R}^d \tag{1}$$

(ici $(\beta_{t,n})$, $1 \le t \le n$, $n \ge 1$, représente une suite de séquences de réels positifs, (h_n), $n \ge 1$, une suite de réels positifs décroissant vers zéro, et la fonction K une densité de probabilité sur $\mathbb{R}^d$, bornée).

Les estimateurs définis en (1) ont été introduits par Deheuvels (1974), ils comprennent les estimateurs récursifs "standard" ($\beta_{t,n} = n^{-1}$, $1 \le t \le n$, $n \ge 1$) étudiés, entre autres, par Devroye et Györfi (1985) quand les X_t, $t \in \mathbb{Z}$, forment une suite d'observations indépendantes.

De très nombreux résultats de convergence des estimateurs à noyau, ont, en effet, été démontrés dans ce cas (cf. Bosq et Lecoutre (1987), par exemple, pour une introduction), puis dans le cadre des processus (cf. Györfi et al. (1989), pour une synthèse récente). En ce domaine, les études réalisées jusqu'alors concernent essentiellement les processus strictement stationnaires et mélangeants en un des sens usuels (cf. Györfi et al. (1989), pour une définition).

Plus récemment, des auteurs comme Delecroix (1991), Györfi (1981), Györfi et Masry (1988) et Yakowitz (1989) ont obtenu les convergences ponctuelles, L^1 et L^2 d'estimateurs du noyau lorsque les X_t constituent un processus ergodique, condition impliquée par toutes les hypothèses de mélangeance mais strictement plus faible qu'elles (cf. Ash et Gardner (1975) ou Rosenblatt (1978)). C'est pourquoi nous démontrerons la convergence uniforme de $\hat{f}_n$ vers f, sous ce type d'hypothèse. Le résultat est, à notre connaissance, original.

Cependant, les contre-exemples de P. Shields (cf. Györfi et al, p. 60) et de Györfi et Lugosi (1991) montrent que l'ergodicité de l'échantillon ne suffit pas à assurer la convergence L_1 des estimateurs construits par les méthodes de l'histogramme et du noyau, respectivement.

Le problème de la détermination des conditions supplémentaires minimales qu'on doit imposer à la structure de dépendance de $(X_t, t \in \mathbb{Z})$ pour obtenir la convergence d'estimateurs de paramètres fonctionnels reste ouvert (cf. Györfi et Lugosi (1991)).

Dans la lignée des travaux de Györfi (1981), Györfi et al. (1989) et Delecroix (1991), on impose une condition sur les densités de transition du processus. Ainsi, supposant l'existence des densités conditionnelles, $f_{X_t}^{\mathcal{F}_{t-1}}$, des variables de X_t par rapport à la tribu engendrée par le passé du processus $(\mathcal{F}_{t-1} = \sigma(X_{t-i}, i \geq 1))$ et posant $U_n = n^{-1} \sum_{t=1}^{n} f_{X_t}^{\mathcal{F}_{t-1}}$, nous utiliserons exactement l'hypothèse suivante pour établir les convergences souhaitées :

(P) $f \in C_0(\mathbb{R}^d)$, $f_{X_t}^{\mathcal{F}_{t-1}} \in C_0(\mathbb{R}^d)$ et on a

(a) $\sup_{x \in \mathbb{R}^d} |U_n(x) - f(x)| \xrightarrow[n \to \infty]{p.s.} 0$,

(b) $\lim_{n \to \infty} E\left(\sup_{x \in \mathbb{R}^d} |U_n(x) - f(x)|\right) = 0$,

$C_0(\mathbb{R}^d)$ désignant l'espace des fonctions continues tendant vers zéro à l'infini.

La condition (P), qui traduit la convergence de U_n vers la densité stationnaire f du processus, est évidemment vérifiée par les processus ergodiques à densités suffisamment régulières (Cf. Delecroix (1991), Györfi (1981), Györfi et al (1989)).

La convergence uniforme d'estimateurs à noyau a déjà été étudiée, par exemple, par Bosq (1988), Collomb (1985), Roussas (1988) dans le cas des processus mélangeants. Les résultats qui suivent généralisent donc leurs travaux. Ils constituent une amélioration d'un résultat antérieur de Delecroix (1987), obtenue grâce à l'utilisation d'une majoration exponentielle (cf. Lemme 1) dérivée de techniques type "différence de martingales" (cf. Azuma (1967)).

Après cette introduction, le deuxième paragraphe sera consacré à l'exposition de nos résultats, le troisième à la démonstration du théorème principal et le quatrième à quelques commentaires.

II. RESULTATS PRINCIPAUX.

Pour obtenir la convergence à zéro de $\|\hat{f}_n - f\|_0$ ($\|.\|_0$ représente la norme de la convergence uniforme sur $C_0(\mathbb{R}^d)$), nous majorerons cette quantité par la somme de

$$A_n = \|\sum_{t=1}^{n} Y_{t,n}\|_0 \quad \text{et}$$

$$B_n = \|\sum_{t=1}^{n} \beta_{t,n}\, h_t^{-d}\, E^{\mathcal{F}_{t-1}} (K((.-X_t)/h_t)) - f\|_0 \ , \tag{2}$$

où $f(.)$ représente la fonction qui à y associe $f(y)$ et où l'on a posé, pour x élément de $\mathbb{R}^d$,

$$Y_{t,n}(x) = \beta_{t,n}\, h_t^{-d} \left\{ K((x-X_t)/h_t) - E^{\mathcal{F}_{t-1}} (K((x-X_t)/h_t)) \right\} . \tag{3}$$

Les variables $Y_{t,n}(x)$ formant, à x et n fixés, une différence de martingales (Azuma (1967)), la convergence à zéro du terme A_n résultera du lemme suivant :

Lemme 1.

Soit $(Z_t,\ t \in \mathbb{N})$ une différence de martingales par rapport à une suite croissante de tribus, telle que, pour tout entier positif t, on ait

$$|Z_t| \leq \alpha_t < +\infty \quad \text{p.s.} \quad .$$

Alors, pour tout ε, $\varepsilon > 0$, et tout entier positif n, on a

$$P\left(|\sum_{t=1}^{n} b_{t,n} Z_t| > \varepsilon \right) \leq 2 \exp\left(-\frac{\varepsilon^2}{2} \left(\sum_{t=1}^{n} b_{t,n}^2\, \alpha_t^2\right)^{-1} \right) ,$$

quelle que soit la suite de séquences de réels $(b_{t,n})$, $1 \leq t \leq n$, $n \geq 1$. ∎

Nous pouvons, d'autre part, écrire B_n sous la forme

$$\|\sum_{t=1}^{n} \beta_{t,n} [H_{h_t}] (f_{X_t}^{\mathcal{F}_{t-1}}) - f\|_0 \ ,$$

où $[H_{h_t}]$ est l'opérateur qui, à une fonction g de $C_0(\mathbb{R}^d)$ associe la fonction $[H_{h_t}](g)$ définie ponctuellement par

$$([H_{h_t}](g))\,(x) = \int_{\mathbb{R}^d} h_t^{-d}\, K((x-u)/h_t)\; g(u)\; du.$$

Cet opérateur jouit des propriétés suivantes :

Lemme 2. (Delecroix (1987))

Soit, pour $\beta \in \mathbb{R}^+$, $[H_\beta]$ l'opérateur défini ponctuellement par

$$([H_\beta](g))\,(x) = \int_{\mathbb{R}^d} \beta^{-d}\, K((x-u)/\beta)\; g(u)du\ , \quad x \in \mathbb{R}^d$$

K étant une densité bornée sur $\mathbb{R}^d$.

Alors $[H_\beta]$ est un opérateur de $C_0(\mathbb{R}^d)$ dans $C_0(\mathbb{R}^d)$, de norme inférieure à 1. De plus, $[H_\beta]$ converge fortement vers l'identité de $C_0(\mathbb{R}^d)$, si β tend vers zéro. ■

La convergence à zéro du terme B_n se déduira alors de ces propriétés et d'un lemme d'analyse fonctionnelle dû à Fritz (1974) :

Lemme 3.

Soit $([K_{t,n}])$, $1 \le t \le n$, $n \ge 1$, une suite de séquences d'opérateurs définis sur un Banach B à valeurs dans B.

Supposons que :

(c1) La suite des opérateurs $\left(\sum_{t=1}^{n} [K_{t,n}]\right)$, $n \ge 1$, converge fortement vers un opérateur C de B dans B,

(c2) t étant fixé, $\lim_{n\to\infty} \|K_{t,n}\| = 0$,

(c3) $\sup_n \left(\sum_{t=1}^{n} \|[K_{t,n}]\|\right) < \infty$,

où $\|.\|$ désigne la norme usuelle des opérateurs de B dans B.

Alors, pour toute suite (y_t), $t \ge 1$, de B convergeant vers y, on aura

$$\lim_{n\to\infty}\left(\sum_{t=1}^{n}[K_{t,n}](y_t)\right) = Cy\,.$$

Si les $[K_{t,n}]$ vérifient de plus

$$\text{(c4)}\qquad \sup_n \left(\sum_{t=1}^{n-1} t\,\|[K_{t,n}] - [K_{t+1,n}]\| + n\|[K_{n,n}]\|\right) < +\infty\,,$$

la conclusion reste valable lorsque la suite (y_t) converge vers y au sens de Césaro. ∎

Nous obtenons précisément le théorème qui suit.

Théorème 1.

Supposons que, outre les hypothèses générales du paragraphe I, le noyau K est à support compact, Lipschitzien, et qu'il existe C, C > 0, tel que

$$\text{K)}\quad \forall\, a, b \in \mathbb{R}^+,\ \int_{\mathbb{R}^d} |a^{-d} K(x/a) - b^{-d} K(x/b)|\, dx \;\le\; C\, \left|1 - \left(\frac{b}{a}\right)^d\right|.$$

Admettons de plus que les coefficients h_n et $\beta_{t,n}$, $1 \le t \le n$, $n \ge 1$, vérifient :

$$\text{H1)}\quad \exists\, \alpha > 0 \ \text{ tel que } \sum_{n=1}^{+\infty} (n^{-\alpha} h_n^{-2d-1}) < +\infty \ ;$$

$$\beta1)\quad \sum_{t=1}^{n} \beta_{t,n}^2 h_t^{-2d} = o\left(\frac{1}{\text{Log } n}\right);$$

$$\beta2)\quad \lim_{n\to\infty} \sum_{t=1}^{n} \beta_{t,n} = 1 \ ;$$

$$\beta3)\quad t \text{ étant fixé, } \lim_{n\to\infty} \beta_{t,n} = 0\,;$$

$$\beta4)\quad \sup_n \left\{\sum_{t=1}^{n-1} t\, \beta_{t+1,n}\, (1 - (h_{t+1}/h_t)^d)\right\} < +\infty \ ;$$

$$\beta5)\quad \sup_n \left\{\sum_{t=1}^{n-1} t\, |\beta_{t,n} - \beta_{t+1,n}|\right\} < +\infty \ ;$$

$$\beta6)\quad \sup_n \{n|\beta_{n,n}|\} < +\infty\,.$$

Alors, si $E(\|X_1\|_e^2) < \infty$ ($\|.\|_e$ désigne la norme euclidienne sur $\mathbb{R}^d$), $\hat{f}_n$ converge vers f, dans $C_0(\mathbb{R}^d)$, presque sûrement et en moyenne. ■

Deux choix particuliers de suites $\beta_{t,n}$ ont été effectués jusqu'alors : $\beta_{t,n} = n^{-1}$, $\beta_{t,n} = h_t^d \Big/ \left(\sum_{t=1}^{n} h_t^d\right)$, $1 \le t \le n$, $n \ge 1$ (Deheuvels (1974), Györfi (1981)). Les deux corollaires qui suivent précisent sous quelles conditions les estimateurs correspondants convergent vers f. On les obtient par un calcul direct, en vérifiant successivement les conditions βi) du Théorème 1.

Corollaire 1.

Supposons que, outre les hypothèses générales du paragraphe I, le noyau K soit à support compact, lipschitzien et vérifie la condition K). Si de plus le processus $(X_t, t \in \mathbb{Z})$ vérifie $E(\|X_1\|_e^2) < \infty$ et la suite (h_n), $n \ge 1$, est telle que

H2) $(h_n/h_{n+1})^d < 1 + \frac{1}{n}$, pour n suffisamment grand,

H3) $\lim_{n\to\infty} nh_n^{2d}/\mathrm{Log}n = +\infty$,

alors $\|f_n^1 - f\|_0 \underset{n\to\infty}{\longrightarrow} 0$, p.s. et en moyenne,

où
$$f_n^1(x) = n^{-1} \sum_{t=1}^{n} h_t^{-d} K((x-X_t)/h_t),\ x \in \mathbb{R}^d . \quad ■$$

Corollaire 2.

Sous les hypothèses du corollaire 1, on a
$$\|f_n^2 - f\| \underset{n\to\infty}{\longrightarrow} 0, \text{ p.s. et en moyenne,}$$
où
$$f_n^2(x) = \left(\sum_{t=1}^{n} h_t^d\right)^{-1} \sum_{t=1}^{n} K((x-X_t)/h_t),\ x \in \mathbb{R}^d . \quad ■$$

La condition K) est vérifiée par de nombreux noyaux. Devroye et Györfi (1985, p. 186) montrent que si K est "radialement décroissant" (i.e., à x fixé, $K(\lambda x)$ est monotone décroissante en λ, $0<\lambda<\infty$), alors on a

$$\int_{\mathbb{R}^d} |a^{-d}K(ax) - K(x)|\, dx \leq 2\,(1-a^d),\ 0 < a < 1,$$

et dès lors K vérifie la condition K). Delecroix (1987, p. 61) montre qu'il en est de même pour les noyaux continuement différentiables réguliers. En particulier, les produits de densités type uniforme ou triangulaire en fournissent de bons exemples.

Notons enfin que si h_n est de la forme n^{-b} $(0 < b < \frac{1}{2d})$ la suite des h_n vérifie les conditions des corollaires 1 et 2.

III. DEMONSTRATION DU THEOREME 1.

Soit α un réel vérifiant H1) et n un entier positif.

Posant

$$A_n^1 = \sup_{x \in [-n^\alpha, n^\alpha]^d} \left| \sum_{t=1}^{n} Y_{t,n}(x) \right|$$

et

$$A_n^2 = \sup_{x \notin [-n^\alpha, n^\alpha]^d} \left| \sum_{t=1}^{n} Y_{t,n}(x) \right| \quad (Y_{t,n}(x)) \text{ défini en } (3)),$$

nous allons montrer successivement la convergence à zéro de A_n^1 . A_n^2 et B_n (défini en (2)).

A) Convergence de A_n^1 .

Soit $\delta_n = h_n^{d+2}$. On peut diviser l'intervalle $[-n^\alpha, n^\alpha]$ en $\ell_n = \dfrac{2n^\alpha}{\delta_n}$ sous-intervalles d'amplitude δ_n, si $\dfrac{2n^\alpha}{\delta_n}$ est entier (sinon, $\ell_n = \left[\dfrac{2n^\alpha}{\delta_n}\right] + 1$, où [a] désigne la partie entière du réel a, et l'amplitude du dernier intervalle est inférieure à δ_n). De même le compact $[-n^\alpha, n^\alpha]^d$ est divisé en $m_n = \ell_n^d$ pavés de côtés inférieurs ou égaux à δ_n, appelés $\mathcal{J}_{n,1}, \ldots, \mathcal{J}_{n,m_n}$.

On peut écrire

$$\sup_{x\in[-n^{\alpha},n^{\alpha}]^d}\left|\sum_{t=1}^{n} Y_{t,n}(x)\right| = \max_{1\le i\le m_n}\ \sup_{x\in \mathcal{I}_{n,i}}\left|\sum_{t=1}^{n} Y_{t,n}(x)\right|$$

$$\le A_n^3 + A_n^4 ,$$

où

$$A_n^3 = \max_{1\le i\le m_n}\ \sup_{x\in \mathcal{I}_{n,i}}\left|\sum_{t=1}^{n} \{Y_{t,n}(x) - Y_{t,n}(x_{n,i})\}\right| ,$$

$$A_n^4 = \max_{1\le i\le m_n}\left|\sum_{t=1}^{n} Y_{t,n}(x_{n,i})\right|$$

et $x_{n,i}$ est un point arbitraire de $\mathcal{I}_{n,i}$, $i = 1,\dots,m_n$.

Le caractère lipschitzien de K permet d'écrire, pour tous i, $1 \le i \le m_n$, et x, $x \in \mathcal{I}_{n,i}$,

$$\left|\sum_{t=1}^{n} \{Y_{t,n}(x) - Y_{t,n}(x_{n,i})\}\right| = 0\left(\|x-x_{n,i}\|_e \sum_{t=1}^{n} \beta_{t,n}\, h_t^{-d-1}\right) .$$

D'après la définition des $\mathcal{I}_{n,i}$, nous obtenons

$$A_n^3 = 0\left(\delta_n \sum_{t=1}^{n} \beta_{t,n}\, h_t^{-d-1}\right) .$$

Finalement, l'hypothèse $\beta2$), le choix de δ_n et la décroissance de (h_n), $n\ge1$, donnent

$$A_n^3 = 0(h_n).$$

D'autre part, pour tout ε, $\varepsilon > 0$, on peut écrire

$$P(A_n^4 > \varepsilon) \le m_n \sup_{x\in\mathbb{R}^d} P\left(\left|\sum_{t=1}^{n} Y_{t,n}(x)\right| > \varepsilon\right).$$

Or, pour tout $x, x \in \mathbb{R}^d$, les $Y_{t,n}(x)$, $t \geq 1$, forment une différence de martingales par rapport aux tribus $(\mathcal{F}_t, t \geq 0)$, et $|Y_{t,n}(x)| \leq 2K_1 \beta_{t,n} h_t^{-d}$ $(K_1 = \sup_{x \in \mathbb{R}^d} |K(x)|)$.

Le lemme 1 permet alors de conclure que

$$P(A_n^4 > \varepsilon) \leq 2m_n \exp\left(-\frac{\varepsilon^2}{8K_1^2}\left(\sum_{t=1}^n \beta_{t,n}^2 h_t^{-2d}\right)^{-1}\right).$$

La définition de m_n donne enfin

$$P(A_n^4 > \varepsilon) = 0\left(n^{\alpha d} h_n^{-d(d+2)} \exp\left(-\frac{\varepsilon^2}{8K_1^2}\left(\sum_{t=1}^n \beta_{t,n}^2 h_t^{-2d}\right)^{-1}\right)\right),$$

série convergente sous les hypothèses H1) et β1), ce qui assure la convergence, presque complète, à zéro de A_n^4.

B) Convergence de A_n^2.

Posons $$\bar{X}_t^n = X_t \mathbb{1}_{\{\|X_t\|_e \leq n^\alpha - K_2 h_t\}},$$

la constante K_2 étant déterminée par l'implication

$$\|z\|_e > K_2 \Rightarrow K(z) = 0 \text{ (K est à support compact).}$$

Soit

$$\bar{Y}_{t,n}(x) = \beta_{t,n} h_t^{-d} \{K((x - \bar{X}_t^n)/h_t) - E^{\mathcal{F}_{t-1}} (K((x - \bar{X}_t^n)/h_t))\}.$$

Alors

$$A_n^2 = \sup_{x \notin [-n^\alpha, n^\alpha]^d} \left|\sum_{t=1}^n \{Y_{t,n}(x) - \bar{Y}_{t,n}(x)\}\right|,$$

car

$$\bar{Y}_{t,n}(x) = 0, \text{ pour tout } x, x \notin [-n^\alpha, n^\alpha]^d.$$

L'hypothèse lipschitzienne faite sur K permet d'affirmer l'existence d'une constante K_3 telle que, pour tout $\varepsilon, \varepsilon > 0$,

$$P(A_n^2 > \varepsilon) \le P\left(\left\{\sum_{t=1}^{n} \beta_{t,n}\, h_t^{-d-1}\, (\|X_t - \tilde{X}_t^n\|_e + E^{\mathcal{F}_{t-1}}(\|X_t - \tilde{X}_t^n\|_e))\right\} > \frac{\varepsilon}{K_3}\right).$$

Il vient, alors, d'après les inégalités de Markov et de Schwarz,

$$P(A_n^2 > \varepsilon) = 0\left(\sum_{t=1}^{n} \beta_{t,n}\, h_t^{-d-1}\,(n^{\alpha} - K_2 h_t)^{-1}\right).$$

Finalement, la décroissance de la suite (h_n) et les hypothèses $\beta2)$ et H1) assurent la convergence de la série de terme général $P(A_n^2 > \varepsilon)$.

Des points précédents découle que $\|\sum_{t=1}^{n} Y_{t,n}\|_0 \underset{n\to\infty}{\overset{p.s.}{\to}} 0$.

On déduit la convergence, en moyenne, à zéro de A_n de l'inégalité

$$\forall\, a > 0,\ E(A_n) \le a + \{\text{p.s. sup } A_n\}\, P(A_n > a).$$

C) Convergence de B_n.

On peut écrire, pour tout g, $g \in C_0(\mathbb{R}^d)$,

$$\|\sum_{t=1}^{n} \beta_{t,n}[H_{h_t}](g) - g\|_0 \le$$
$$\sum_{t=1}^{n} \beta_{t,n}\|[H_{h_t}](g) - g\|_0 + \left|\left(\sum_{t=1}^{n} \beta_{t,n}\right) - 1\right| \|g\|_0 .$$

Le lemme 2 et les hypothèses $\beta2)$, $\beta3)$ permettent alors de conclure quc $\sum_{t=1}^{n} \beta_{t,n}[H_{h_t}]$ converge vers $Id_{C_0(\mathbb{R}^d)}$.

L'hypothèse C1) du lemme 3 est donc vérifiée.

On vérifie de même C2) et C3) sous $\beta3)$ et $\beta2)$, puisque $\|\beta_{t,n}[H_{h_t}]\| = 0(\beta_{t,n})$.

Reste à vérifier C4). Or on a

$$\|\beta_{t,n}[H_{h_t}] - \beta_{t+1,n}\,[H_{h_{t+1}}]\| \le$$

$$|\beta_{t,n} - \beta_{t+1,n}| \ \|[H_{h_t}]\| + \beta_{t+1,n} \ \|[H_{h_t}] - [H_{h_{t+1}}]\|.$$

Maintenant, pour tout g, $g \in C_0(\mathbb{R}^d)$,

$$\|[H_{h_t}](g) - [H_{h_{t+1}}](g)\|_0 \leq$$
$$\leq \|g\|_0 \int_{\mathbb{R}^d} |h_t^{-d} K(z/h_t) - h_{t+1}^{-d} K(z/h_{t+1})| \, dz.$$

Il vient, alors, d'après l'hypothèse K) et la décroissance de la suite (h_n),

$$\|\beta_{t,n}[H_{h_t}] - \beta_{t+1,n}[H_{h_{t+1}}]\| =$$
$$= 0 \left\{ |\beta_{t,n} - \beta_{t+1,n}| + \beta_{t+1,n}(1 - (h_{t+1}/h_t)^d) \right\}.$$

Dès lors C4) est vérifiée pour les $(\beta_{t,n}[H_{h_t}])$, si $\beta 4)$, $\beta 5)$ et $\beta 6)$ le sont.

D'après ce qui précède, la suite d'opérateurs $\left(\sum_{t=1}^{n} \beta_{t,n} [H_{h_t}] \right)$ vérifie les hypothèses du lemme 3.

La condition générale (P), imposée au processus, permet de conclure la convergence, p.s. et en moyenne, à zéro de B_n.

IV. REMARQUES FINALES.

Les techniques développées ci-dessus s'adaptent évidemment à une étude de convergence de l'estimateur usuel du noyau

$$f_n(x) = n^{-1} h_n^{-d} \sum_{t=1}^{n} K((x-X_t)/h_n), \quad x \in \mathbb{R}^d.$$

La démonstration du théorème 1 s'adapte point par point : on majore $\|f_n-f\|_0$ par une somme de deux termes équivalents à A_n et B_n. La démonstration de la convergence de A_n à zéro se calque sur celle du paragraphe III, celle de B_n (nettement simplifiée) s'obtient aussi par des techniques "banachiques", et se trouve in extenso dans Delecroix (1987).

On obtient précisément :

Théorème 2.

Supposons que, outre les hypothèses générales introduites au paragraphe I, le noyau K soit lipschitzien et à support compact.
Admettons que le processus $(X_t, t \in \mathbb{Z})$ est tel que $E(\|X_t\|_e^2) < +\infty$.

Si la suite $(h_n, n \in \mathbb{N})$ vérifie

$$\lim_{n \to \infty} nh_n^{2d} / \operatorname{Log} n = +\infty,$$

alors $\|f_n - f\|_0 \underset{n \to \infty}{\longrightarrow} 0$, p.s. et en moyenne. ■

En se basant sur des résultats de convergence du même type (dans le cas de processus mélangeants), des auteurs comme Collomb (1985), Bosq (1988), Sarda et Vieu (1985) ont démontré la convergence du prédicteur non paramétrique d'une observation future X_{T+h} à partir de $X_1, \ldots, X_T$. Nos résultats permettront sans doute d'obtenir cette convergence dans le cas d'un processus ergodique, et de même d'obtenir la convergence des estimateurs de divers paramètres fonctionnels dépendant de f (taux de hasard, régression...). Ceci fera l'objet de travaux ultérieurs.

BIBLIOGRAPHIE

Ash, R. et Gardner, M., (1975), Topics in stochastic processes, Academic Press.

Azuma, K., (1967), Weighted sums of certain dependent random variables, Tohoku Math. J., 19, 357-367.

Bosq, D., (1988), Prévision non paramétrique d'un processus stationnaire non borné, Pub. LSTA, 81, Université Pierre et Marie Curie, Paris VI.

Bosq, D. et Lecoutre, J.P. (1987), Théorie de l'estimation fonctionnelle, Economica.

Collomb, G., (1985), Nonparametric time series analysis and prediction : uniform almost sure convergence, Statistics, 2, 297-307.

Deheuvels, P., (1974), Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre et uniforme presque sûre des estimateurs de la densité, C.R.A.S., A, 278, 1217-1220.

Delecroix, M., (1987), Sur l'estimation et la prévision non-paramétrique des processus ergodiques, Thèse de Doctorat d'Etat présentée le 2 Juillet 1987 à l'U.S.T.L.F.A.

Delecroix, M., (1991), Convergence d'estimateurs de paramètres fonctionnels sous une hypothèse ergodique, soumis pour publication aux Annales de l'I.H.P.

Devroye, L. et Gyorfi, L., (1985), Nonparametric density estimation : the L_1 view, John Wiley & Sons.

Gyorfi, L., (1981), Strong consistent density estimate from ergodic sample, J. Mult. Anal., 11, 81-84.

Gyorfi, L., Hardle, W., Sarda, P. et Vieu, P., (1989), Nonparametric curve estimation from time series, Lecture Notes in Statistics, 60, Springer-Verlag.

Gyorfi, L. et Lugosi, G., (1991), Kernel density estimation from ergodic sample is not universally consistent, soumis pour publication à Computational Statistics and Data Analysis.

Gyorfi, L. et Masry, E., (1988), The L_1 and L_2 strong consistency of recursive kernel density estimation from dependent samples, Report, TU Budapest.

Fritz, J., (1974), Learning from an ergodic training sequence, Limit Theorems of Probability Theory (P. Révész, Ed), 79-91, North-Holland, Amsterdam.

Parzen, E., (1962), On estimation of a probability density function and the mode, Ann. Mat. Stat., 33, 1065-1076.

Rosenblatt, M., (1956), Remarks on some nonparametric estimates of a density function, Ann. Mat. Stat., 27, 832-835.

Rosenblatt, M., (1978), Dependence and asymptotic independence for random processes, M.A.A. Studies in Mathematics, vol. 18 (Studies in Proba Theory), 24-45.

Roussas, G., (1988), Nonparametric estimation in mixing sequences of random variables, J. of Stat. Plan. and Inf., 18, 135-149.

Sarda, P. et Vieu, P., (1985), Nonparametric regression estimation, application to prediction, Proceedings of the 4th E.Y.S.M., Pliska, Studia Mathematica Bulgarica.

Yakowitz, S. (1989), Nonparametric density and regression estimation for Markov sequences without mixing assumptions, J. Mult. Anal., 30, 1,124-136.