



HAL
open science

Estimation Non Paramétrique De La Médiane Conditionnelle: Medianogramme Et Methode Du Noyau. Application A La Prévision Des Processus.

Ali Gannoun

► **To cite this version:**

Ali Gannoun. Estimation Non Paramétrique De La Médiane Conditionnelle: Medianogramme Et Methode Du Noyau. Application A La Prévision Des Processus.. Annales de l'ISUP, 1990, XXXV (1), pp.11-22. hal-03666444

HAL Id: hal-03666444

<https://hal.science/hal-03666444v1>

Submitted on 12 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ESTIMATION NON PARAMETRIQUE DE LA MEDIANE CONDITIONNELLE :
MEDIANOGRAMME ET METHODE DU NOYAU.
APPLICATION A LA PREVISION DES PROCESSUS.

par Ali GANNOUN

Laboratoire de Statistique Théorique et Appliquée.

Tour 45-55- 3ème E. 4 Place Jussieu - 75252 PARIS CEDEX 05.

Résumé- Nous proposons deux méthodes non paramétriques pour l'estimation de la médiane conditionnelle à partir de processus α -mélangeant. La première est empirique. L'estimateur, appelé médianogramme, est à rapprocher du régressogramme étudié notamment par BOSQ (1970) [1] et GEFFROY (1975) [2]. La deuxième est la méthode du noyau abondamment utilisée dans le cas de l'espérance conditionnelle. Nous démontrons la convergence uniforme presque complète de ces deux estimateurs vers la médiane conditionnelle théorique et nous les utilisons pour la construction de prédicteurs non paramétriques.

Abstract- We suggest two non parametric estimation methods for conditional median from α -mixing processes. the first is empirical, the estimator, called medianogram, is compared to regressogram studied especially by BOSQ (1970) and GEFFROY (1975). The second is the kernel method used for conditional regression estimation. We prove the uniform almost completely convergence for this two estimators to the theoretical conditional median and we use them for construction of non parametric predictors.

I- INTRODUCTION :

Soit $(Z_n)_{n \in \mathbb{N}}$ un processus mélangeant, on considère le problème de la prédiction de Z_{N+1} à partir de la suite $\{Z_i, i=1, \dots, N\}$.

Un cadre général d'étude est l'observation d'un processus $(X_t, Y_t)_{t \in \mathbb{Z}}$ à valeur dans $\mathbb{R}^d \times \mathbb{R}$ où Y_t est intégrable.

La plupart des prédicteurs non paramétriques sont basés sur des estimateurs non paramétriques de l'espérance conditionnelle $E(Y_t | X_t = x)$. Citons par exemple les travaux de COLLOMB (1984) [3], COLLOMB et HARDLE (1986) [4], VIEU (1987) [5], SARDA (1988) [6], ROUSSAS (1988) [7], BOSQ (1988) [8] et (1989) [9].

Dans ce travail, nous proposons une nouvelle approche de ces prédicteurs par l'estimation de la médiane conditionnelle $\mu(x) = \mu(Y_t | X_t = x)$.

Dans le paragraphe (II), nous détaillons la définition de la médiane conditionnelle et nous présentons les principales hypothèses faites sur le processus (X_t, Y_t) ainsi qu'un lemme servant pour la démonstration des théorèmes de convergence (1) et (2).

Dans le paragraphe (III) (resp. (IV)), nous définissons le médianogramme (resp. la méthode du noyau) et nous démontrons la convergence uniforme presque complète de ces estimateurs vers la médiane conditionnelle théorique. Un cas particulier est traité par la

suite.

En posant

$$n = N-d, X_t = (Z_t, \dots, Z_{t+d-1}) \text{ et } Y_t = (Z_{t+d}),$$

on construit, dans le paragraphe (V), deux prédicteurs non paramétriques de Z_{n+1} à partir des estimateurs définis dans les paragraphes (III) et (IV).

II- DEFINITION, HYPOTHESES GENERALES ET RESULTATS PRELIMINAIRES :

Soit X un vecteur aléatoire (vec. a.) de \mathbb{R}^d et Y une variable aléatoire réelle (v. a. r.) intégrable, on désigne par $P_{X,Y}$ la probabilité du couple (X,Y) dans $(\mathbb{R}^{d+1}, \mathcal{B}(\mathbb{R}^{d+1}))$.

Soit P_Y^X une version régulière de la probabilité conditionnelle de Y par rapport à X et $F(\cdot|x)$ une version de la fonction de répartition conditionnelle (F. r. c.) de Y sachant $X=x$, une médiane conditionnelle (m. c.) de Y sachant $X=x$, $\mu(x)$, est telle que

$$\mu(x) = F^{-1}(1/2|x) = \inf \{ y \mid F(y|x) \geq 1/2 \}.$$

C'est aussi la meilleure approximation de Y par une fonction de X au sens de la norme L^1 (cf. [10]), soit

$$E(|Y-\mu(x)| \mid X=x) = \inf_{a \in \mathbb{R}} E(|Y-a| \mid X=x).$$

Pour l'estimation de μ , nous considérons une suite d'observations $(X_t, Y_t)_{t \leq n}$ d'un processus stationnaire et α -mélangeant (cf. [11]) à valeurs dans $\mathbb{R}^d \times \mathbb{R}$. Soit $h(n)$ une suite positive de limite nulle qui vérifie conjointement avec $\alpha(n)$:

$$(H1) \quad \lim_{n \rightarrow \infty} nh^d(n)/k(n) \log n = \infty.$$

$k(n)$ étant une suite entière croissante de limite infini vérifiant :

$$(H2) \quad \lim_{n \rightarrow \infty} \alpha(k(n)) \frac{2k(n)}{3^n} = 0.$$

\mathcal{C} est un compact de \mathbb{R}^d , on suppose :

- (H3) μ est unique et vérifie l'hypothèse suivante dite d'unicité uniforme (cf. [12]) : $\forall \epsilon > 0, \exists \lambda > 0, \forall t : \mathcal{C} \rightarrow \mathbb{R}$,

$$\sup_{x \in \mathcal{C}} |\mu(x) - t(x)| \geq \epsilon \Rightarrow \sup_{x \in \mathcal{C}} |F(\mu(x)|x) - F(t(x)|x)| \geq \lambda.$$

- (H4) $F(y|\cdot)$ est lipschitzienne sur \mathcal{C} en x uniformément par rapport à y .

Si $\hat{F}_n(\cdot|x)$ (resp. $\hat{\mu}_n(x)$) désigne un estimateur non paramétrique de $F(\cdot|x)$ (resp. de $\mu(x)$), on a :

Lemma 1-Sous (H3), si

$$\sup_{x \in \mathcal{C}} \sup_{y \in \mathbb{R}} |\hat{F}_n(y|x) - F(y|x)| \xrightarrow{p. co.} 0,$$

alors

$$\boxed{\sup_{x \in \mathcal{C}} |\hat{\mu}_n(x) - \mu(x)| \xrightarrow{p. co.} 0.}$$

Démonstration-

Dans (H3), on remplace t par $\hat{\mu}_n$ et on obtient :

$$(a) \forall \epsilon > 0, \exists \lambda > 0, \forall t : \mathcal{C} \rightarrow \mathbb{R}, \sup_{x \in \mathcal{C}} |\hat{\mu}_n(x) - \mu(x)| \geq \epsilon \rightarrow$$

$$\sup_{x \in \mathcal{C}} |F(\hat{\mu}_n(x)|x) - F(\mu(x)|x)| \geq \lambda.$$

$$|F(\hat{\mu}_n(x)|x) - F(\mu(x)|x)| \leq$$

$$|F(\hat{\mu}_n(x)|x) - \hat{F}_n(\hat{\mu}_n(x)|x)| + \underbrace{|\hat{F}_n(\hat{\mu}_n(x)|x) - \hat{F}_n(\mu(x)|x)|}_{1/2} + \underbrace{|\hat{F}_n(\mu(x)|x) - F(\mu(x)|x)|}_{1/2}.$$

Ceci donne :

$$|F(\hat{\mu}_n(x)|x) - F(\mu(x)|x)| \leq 2 \sup_{x \in \mathcal{C}} \sup_{y \in \mathbb{R}} |F(y|x) - \hat{F}_n(y|x)|.$$

En utilisant (a), on obtient :

$$\forall \epsilon > 0, \exists \lambda > 0 \text{ telle que } \sup_{x \in \mathcal{C}} |\hat{\mu}_n(x) - \mu(x)| \geq \epsilon \rightarrow$$

$$\sup_{x \in \mathcal{C}} \sup_{y \in \mathbb{R}} |F(y|x) - \hat{F}_n(y|x)| \geq \lambda/2.$$

D'où :

$$\forall \epsilon > 0, \exists \lambda > 0 \text{ telle que } P(\sup_{x \in \mathcal{C}} |\hat{\mu}_n(x) - \mu(x)| \geq \epsilon) \leq$$

$$P(\sup_{x \in \mathcal{C}} \sup_{y \in \mathbb{R}} |F(y|x) - \hat{F}_n(y|x)| \geq \lambda/2).$$

Si la série dont le terme général est la partie droite de l'inégalité de dessus est convergente, alors il sera de même pour la série dont le terme général est la partie gauche. C. Q. F. D.

III-ESTIMATION EMPIRIQUE DE LA MEDIANE CONDITIONNELLE :

III.1. Construction du médianogramme :

Soit $J_n = \{ I_{n,m}, m \in \mathbb{Z} \}$ le pavage de \mathbb{R}^d défini par :

$$m = (m_1, \dots, m_d) \in \mathbb{Z}^d \text{ et } I_{n,m} = \prod_{k=1}^d [m_k h(n), (m_k + 1)h(n)].$$

Pour tout $x \in \mathbb{R}^d$, on désigne par $I(x, h(n))$ le pavé de J_n contenant x .

$F_{h(n)}(\cdot|x)$ représente la F. r. c. de Y sachant que $X \in I(x, h(n))$, on suppose qu'elle est absolument continue.

$\mu_{h(n)}(x) = F_{h(n)}^{-1}(1/2|x)$ est une m. c. de Y sachant que $X \in I(x, h(n))$ supposée unique.

On note $q(x, h(n))$ le nombre de X_1 appartenant à $I(x, h(n))$: c'est la réalisation d'une v. a. $Q(x, h(n))$.

$Y_{11}, Y_{12}, \dots, Y_{1q(x, h(n))}$ sont les Y_1 telles que $X_1 \in I(x, h(n))$.

On définit l'estimateur empirique de $F_{h(n)}(y|x)$ qu'on note $\hat{F}_{h(n)}(y|x)$ de la façon suivante :

$$\hat{F}_{h(n)}(y|x) = F_{q(x, h(n))}(Y_{11}, Y_{12}, \dots, Y_{1q(x, h(n))}, y)$$

$$= \begin{cases} \frac{1}{q(x, h(n))} \sum_{j=1}^{q(x, h(n))} 1_{\{Y_{1j} \leq y\}} & \text{si } q(x, h(n)) \neq 0, \\ 0 & \text{si } q(x, h(n)) = 0. \end{cases}$$

L'estimateur empirique $\hat{\mu}_{h(n)}(x)$ de $\mu_{h(n)}(x)$ est défini comme suit :

$$\hat{\mu}_{h(n)}(x) = \hat{F}_{h(n)}^{-1}(1/2|x).$$

$$= \begin{cases} Y_{([1q(x, h(n))]/2)} & \text{si } q(x, h(n)) \text{ est impair,} \\ (Y_{([1q(x, h(n))]/2)} + Y_{([1q(x, h(n))]/2+1)}) & \text{si } q(x, h(n)) \text{ est pair,} \\ \mu_0 & \text{si } q(x, h(n)) = 0. \end{cases}$$

(μ_0 , étant une valeur arbitraire).

III.2. - Convergence uniforme (p. co.) de $\hat{\mu}_{h(n)}$ vers μ :

Soit les hypothèses suivantes :

(M1) La f. r. $F(\dots)$ de (X, Y) est absolument continue et sa densité est bornée et lipschitzienne sur $I(x, h(n)) \times \mathbb{R}$.

(M2) La f. r. c. $F(\cdot|x)$ est absolument continue.

(M3) La loi marginale de X a une densité strictement minorée par $m > 0$ et est lipschitzienne sur $I(x, h(n))$.

Lemme 2- Sous (M1), (M2) et (M3) on a :

$$\sup_{x \in \mathbb{R}} |F_{h(n)}(y|x) - F(y|x)| \rightarrow 0, n \rightarrow \infty.$$

Démonstration- (cf. [13]).

Soit $G(\cdot)$ la f. r. d'une v. a. Z et $\hat{G}_n(\cdot)$ l'estimateur empirique de G défini, à partir d'observations $(Z_i)_{i \in \mathbb{N}}$ α -mélangeantes, alors :

Lemme 3- Sous (M2), pour n assez grand et $\varepsilon > 0$, on a :

$$\sup_{z \in \mathbb{R}} P(|\hat{G}_n(z) - G(z)| > \varepsilon) \leq 2 \exp\left(-\frac{n\varepsilon}{32ek(n)}\right)$$

Démonstration- (cf. [13]).

THEOREME 1- Sous (H1), (H2), (H3), (H4), (M1), (M2) et (M3), on a :

$$\sup_{x \in \mathcal{E}} (\hat{\mu}_{h(n)}(x) - \mu(x)) \xrightarrow{p. co.} 0$$

Démonstration- Il suffit de démontrer que sous ces hypothèses, on a :

(*) $\sup_{x \in \mathcal{E}} \sup_{y \in \mathbb{R}} |\hat{F}_{h(n)}(y|x) - F(y|x)| \xrightarrow{p. co.} 0$ et d'appliquer le lemme 1.

En utilisant la formule de probabilité totale (cf. [13]) et le lemme 3, on obtient pour n assez grand,

$$P(\sup_{y \in \mathbb{R}} |\hat{F}_{h(n)}(y|x) - F_{h(n)}(y|x)| \geq \varepsilon) \leq 2 \exp^{-\frac{\varepsilon m h^d(n)}{32ek(n)}}.$$

\mathcal{E} étant un compact de \mathbb{R}^d , on peut lui construire un recouvrement par un nombre ℓ_n fini de cubes I de \mathbb{R}^d de côté $h(n)$ en se sens qu'il existe $x_1, x_2, \dots, x_{\ell_n}$ dans \mathcal{E} tels que :

$$\mathcal{E} \subset \bigcup_{i=1}^{\ell_n} I(x_i, h(n))$$

On choisit $h(n)$ de sorte que ℓ_n vérifie :

$$\ell_n \leq \frac{B}{h^d(n)} \quad (B > 0, \text{ indépendant de } n)$$

D'après (H4), $|F(y|x) - F(y|x_1)| \leq \delta \|x - x_1\| \quad \forall x \in \mathcal{E} \text{ et } \forall i \in \{1, \dots, \ell_n\}$.

Ainsi pour tout $x \in I(x_i, h(n))$, on a $|F(y|x) - F(y|x_1)| \leq \delta h(n)$
 $\forall i \in \{1, \dots, \ell_n\}$.

Compte tenu du fait que $\lim_{n \rightarrow \infty} h(n) = 0$ et de l'inégalité

$$|\hat{F}_{h(n)}(y|x) - F(y|x)| \leq$$

$$|\hat{F}_{h(n)}(y|x) - F_{h(n)}(y|x_1)| + |F_{h(n)}(y|x_1) - F(y|x_1)| + |F(y|x_1) - F(y|x)|$$

$\forall x \in I(x_i, h(n)), \forall i \in \{1, \dots, \ell_n\}$, on a :

$$P(\sup_{x \in \mathcal{E}} \sup_{y \in \mathbb{R}} |\hat{F}_{h(n)}(y|x) - F(y|x)| > \varepsilon) \leq$$

$$P(\sup_{1 \leq i \leq \ell_n} \sup_{y \in \mathbb{R}} |\hat{F}_{h(n)}(y|x_i) - F(y|x_i)| > \varepsilon/3) \leq 2\ell_n \exp^{-\frac{\varepsilon m h^d(n)}{96ek(n)}}.$$

Comme $\lim_{n \rightarrow \infty} n h^d(n) = \infty$, il existe η tel que $(n h^d(n))^{-1} < \eta$ et donc

$\ell_n < B n \eta$, par conséquent :

$$\sum 2L_n \exp - \frac{cmnh^d(n)}{96ek(n)} \leq \sum 2B\eta n \cdot 1 - \frac{cmnh^d(n)}{96ek(n)\text{Log}n} < \infty \text{ à cause de (H1)}.$$

Ainsi, nous avons démontré (*).

III.3- Cas particulier: Processus géométriquement mélangé.

Les hypothèses (H1) et (H2) sont vérifiées, en effet :

$$\alpha(k) = a\rho^k, \quad a > 0, \quad 0 < \rho < 1.$$

Posons :

$$k(n) = \beta_n n^{(1+\beta)/2}, \quad 1 < \beta_n < 2, \text{ alors}$$

$$\begin{aligned} \alpha(k(n)) &= \exp \left(\frac{2k(n)}{3n} \text{Log} \alpha(k(n)) \right) \\ &= \exp \left(\frac{2k(n)}{3n} \text{Log} a \right) \exp \left(\text{Log} \rho \frac{2k(n)^2}{3n} \right) \\ &= \exp \left(\frac{2n^{(\beta-1)/2}}{3} \text{Log} a \right) \exp \left(\text{Log} \rho \frac{2\beta_n n^\beta}{3} \right) \rightarrow 0. \end{aligned}$$

(H2) est donc vérifiée.

Pour l'hypothèse (H1), un simple calcul permet d'avoir :

$$\lim_{n \rightarrow \infty} nh^d(n)/k(n)\text{Log}n = \infty \Leftrightarrow \lim_{n \rightarrow \infty} n^{(1-\beta)/2} h^d(n)/\text{Log}n = \infty.$$

Il suffit de choisir $h(n) = O\left(\left(\frac{n^\gamma \text{Log}n}{n^{(1-\beta)/2}}\right)^{1/d}\right)$, $\gamma \in]0, (1-\beta)/2[$.

IV. ESTIMATION DE LA MEDIANE CONDITIONNELLE PAR LA METHODE DU NOYAU :

Soit K un noyau de \mathbb{R} (cf. [14]), on supposera dans tout ce qui suit que ce noyau vérifie :

$$(N1) \quad 0 < K(\cdot) < K_1 \text{ et } \int K(x)dx = 1.$$

K_1 est une constante positive.

IV.1- Construction de l'estimateur :

Notons $K_n(x) = \frac{1}{h^d(n)} K\left(\frac{x}{h(n)}\right) \forall x \in \mathbb{R}^d$, pour tout $x \in \mathbb{R}^d$, on considère

l'estimateur $(P_Y^{X=x})_n$ de $(P_Y^{X=x})$ définie par :

$$(P_Y^{X=x})_n = \frac{\sum_{i=1}^n \delta_{Y_i} K_n(x-X_i)}{\sum_{i=1}^n K_n(x-X_i)}$$

$$F_n(y|x) = \int_{\{Y \leq y\}} (P_Y^{X=x})_n(dy) = \frac{\sum_{i=1}^n 1_{\{Y_i \leq y\}} K_n(x-X_i)}{\sum_{i=1}^n K_n(x-X_i)} \text{ est un estimateur}$$

non paramétrique de la f. r. c. de Y sachant $X=x$.

Un estimateur non paramétrique de la médiane conditionnelle de Y sachant $X=x$ est défini par :

$$\mu_n(x) = F_n^{-1}(1/2|x) = \text{Inf} \left\{ y \text{ telle que } F_n(y|x) \geq 1/2 \right\}$$

IV.2- Convergence uniforme (p.co.) de μ_n vers μ :

On suppose l'existence ($\forall y$) d'une application définie sur \mathbb{R}^d telle que :

$$(N2) \quad E(1_{\{Y_1 \leq y\}} | \cdot) = F(y | \cdot)$$

La loi marginale de X_1 vérifie avec la mesure (λ) de Lebesgue :

$$(N3) \quad \begin{aligned} \exists \Gamma < \infty : P(X_1 \in B) \leq \Gamma \lambda(B), \quad \forall 1 \in \mathbb{N}, \quad \forall B \in \mathcal{B}_{\mathbb{R}^d} \\ \exists \gamma < \infty : P(X_1 \in B) \geq \gamma \lambda(B), \quad \forall 1 \in \mathbb{N}, \quad \forall B \in \mathcal{B}_{\mathbb{R}^d} \end{aligned}$$

Posons :

$$F^{1,n}(x, y) = \frac{1}{n} \sum_{i=1}^n 1_{\{Y_i \leq y\}} K_n(x - X_i), \quad f_n(x) = \frac{1}{n} \sum_{i=1}^n K_n(x - X_i),$$

alors :

$$F_n(y|x) = \frac{F^{1,n}(x, y)}{f_n(x)}$$

Lemme 3- Si $k(n)$ est une suite vérifiant (H2), alors :

Pour tout $\epsilon > 0$ et pour n suffisamment grand,

$$\sup_x \sup_y P(|F^{1,n}(x, y) - E(F^{1,n}(x, y))| > \epsilon) \leq 2 \cdot \exp\left(-\frac{bnh^d(n)}{k(n)}\right)$$

Démonstration- C'est une application d'un lemme de CARBON (1983) [15],

$$\text{posons} \quad \Delta_1 = 1_{\{Y_1 \leq y\}} K\left(\frac{x - X_1}{h(n)}\right) - E(1_{\{Y_1 \leq y\}} K\left(\frac{x - X_1}{h(n)}\right)),$$

D'après (N1) et (N3), on a $|\Delta_1| \leq 2K_1 = d$ et $E(\Delta_1^2) \leq h^d(n)K_1\Gamma = D_n$.

On choisit $\alpha = \frac{1}{4kde}$, ($k = k(n)$, $h = h(n)$), on pose $\xi(k) = \sum_{i=1}^n \alpha(1)$.

$$P\left(\frac{1}{n} |\Delta_1| > \epsilon\right) \leq 2 \exp\left(\frac{n}{k} \left[\frac{\epsilon}{8K_1e} - \frac{1}{16K_1e} \left(\frac{\Gamma h^d}{k} + 32K_1 \frac{\xi(k)}{k}\right) - 2\sqrt{e}(\alpha k)^{2k/n} \right]\right).$$

Comme $\lim_{n \rightarrow \infty} k(n) = \infty \Rightarrow \lim_{n \rightarrow \infty} \alpha(k(n)) = 0$, par conséquent $\lim_{n \rightarrow \infty} \frac{\xi(k(n))}{k(n)} = 0$.

$$\text{De même } \lim_{n \rightarrow \infty} \frac{\Gamma h^d(n)}{k(n)} = 0.$$

L'hypothèse (H2) implique $\lim_{n \rightarrow \infty} 2\sqrt{e}(\alpha(k))^{2k/n} = 0$.

Ainsi, pour n assez grand, on obtient :

$$P\left(\frac{1}{n} |\Delta_1| > \epsilon\right) \leq 2 \cdot \exp\left(-\frac{n\epsilon}{32K_1ek(n)}\right) \text{ et}$$



$$P(|F^{1,n}(x,y) - E(F^{1,n}(x,y))| > c) \leq P\left(\frac{1}{n} |\Delta_1| > cnh^d(n)\right) \\ \leq 2 \exp\left(-\frac{cnh^d(n)}{32K_1ek(n)}\right).$$

Pour finir la démonstration, il suffit de poser $b = \frac{c}{32K_1ek}$ et de remarquer que le dernier majorant ne dépend pas ni de x ni de y .

Lemme 4 - Sous (H1) et (H2), si K est lipschitzien on a :

$$\sup_{x \in \mathcal{E}} \sup_{y \in \mathbb{R}} |F^{1,n}(x,y) - E(F^{1,n}(x,y))| \xrightarrow{P. CO.} 0$$

Démonstration - K est lipschitzien, on note L son rapport, donc :

$$\left| K\left(\frac{x-X_1}{h(n)}\right) - K\left(\frac{x'-X_1}{h(n)}\right) \right| \leq Lh(n)^{-\rho} \|x-x'\|^\rho, \quad \forall (x,x') \in \mathbb{R}^{2d}.$$

On se fixe τ tel que :

$$\tau > d/\rho + 1 \quad (I)$$

Pour tout $n \in \mathbb{N}$, on désigne par $\{I_{tk}, k = 1, \dots, \ell_n\}$ un recouvrement de \mathcal{E} par des boules de centre tk et de rayon $\leq h(n)^\tau$.

Puisque \mathcal{E} est compact, donc \mathcal{E} est borné et on peut choisir un recouvrement de manière à ce que, $\forall n \in \mathbb{N}$, on ait :

$$\ell_n \leq Bh(n)^{-d\tau} \quad (B > 0, \text{ indépendant de } n). \quad (**)$$

On pose :

$$\Delta_n(x,y) = \frac{1}{nh^d(n)} \left[\sum_{i=1}^n 1_{\{Y_i < y\}} K\left(\frac{x-X_1}{h(n)}\right) - E\left(\sum_{i=1}^n 1_{\{Y_i < y\}} K\left(\frac{x-X_1}{h(n)}\right)\right) \right]$$

$$\Delta_n(x,y) = \Delta_n(tk,y) + \hat{\Delta}_n(x,y) - \Delta_n(tk,y) \quad (***)$$

$$|\Delta_n(x,y)| \leq |\Delta_n(tk,y) - \Delta_n(x,y)| + |\Delta_n(tk,y)|$$

$$|\Delta_n(x,y)| \leq |\hat{\Delta}_n(x,y)| + |\Delta_n(tk,y)|,$$

$$\hat{\Delta}_n(x,y) = \Delta_n(x,y) - \Delta_n(tk,y)$$

$$= \frac{1}{nh^d(n)} \left(\sum_{i=1}^n 1_{\{Y_i < y\}} \left[K\left(\frac{x-X_1}{h(n)}\right) - K\left(\frac{tk-X_1}{h(n)}\right) \right] \right)$$

$$- E \sum_{i=1}^n 1_{\{Y_i < y\}} \left[K\left(\frac{x-X_1}{h(n)}\right) - K\left(\frac{tk-X_1}{h(n)}\right) \right].$$

Le fait que $1_{\{Y_i < y\}} \leq 1$ et que K est lipschitzien permet d'avoir :

$$|\hat{\Delta}_n(x,y)| \leq \frac{1}{nh^d(n)} \sum_{i=1}^n (|K\left(\frac{x-X_1}{h(n)}\right) - K\left(\frac{tk-X_1}{h(n)}\right)| + E|K\left(\frac{x-X_1}{h(n)}\right) - K\left(\frac{tk-X_1}{h(n)}\right)|)$$

$$\leq 2Lh(n)^{-(d+\rho)} \|x - tk\|^\rho.$$

Puisque tk a été choisi de manière à ce que la boule I_{tk} contienne x ,

alors :

$$\|x - tk\| \leq h(n)^T \Rightarrow |\hat{\Delta}_n(x, y)| \leq 2Lh(n)^{T\rho - (d+\rho)}$$

Cette majoration est indépendante de (x, y) . En utilisant l'inégalité (I) et $\lim_{n \rightarrow \infty} h(n) = 0$, on peut conclure que :

$$\sup_{x \in \mathcal{E}} \sup_{y \in \mathbb{R}} |\hat{\Delta}_n(x, y)| \longrightarrow 0, \quad n \longrightarrow \infty. \quad (\text{convergence certaine}).$$

Pour finir la démonstration du lemme, il faut prouver :

$$\text{Max}_{1, \ell_n} |\Delta_n(tk, y)| \xrightarrow{\text{p. co.}} 0. \quad (***)$$

Pour tout $\varepsilon > 0$, on pose :

$$\beta(n) = P(\text{Max}_{1, \ell_n} |\Delta_n(tk, y)| > \varepsilon),$$

en utilisant le lemme 3 et les inégalités (I) et (*), on a :

$$\beta(n) \leq 2Bh(n)^{-d(1+d/\rho)} \exp\left(-\frac{bnh^d(n)}{k(n)}\right).$$

(H1) implique

$$\exp\left(-\frac{bnh^d(n)}{k(n)}\right) = n^{-\frac{bnh^d(n)}{k(n)\text{Log}n}} = n^{-b\gamma(n)} \quad \text{où } \gamma(n) = \frac{nh^d(n)}{k(n)\text{Log}n} \longrightarrow 0.$$

Ceci implique qu'il existe $\delta > 0$ pour laquelle on a : Pour n assez grand

$$(nh^d(n))^{-\eta} < \delta,$$

et il s'en suit :

$$\beta(n) \leq 2B\delta n^{\eta - b\gamma(n)},$$

et donc :

$$\sum_{i=1}^n \beta(n) \leq \infty \text{ et donc } (***).$$

Lemme 5- Sous les hypothèses du lemme 4, on a :

$$\sup_{x \in \mathcal{E}} |f_n(x) - E(f_n(x))| \xrightarrow{\text{p. co.}} 0.$$

Démonstration- Il suffit de remplacer dans la démonstration du lemme précédent $1_{\{Y_1 \leq y\}}$ par 1.

Lemme 6- Sous les hypothèses du lemme 4, on a :

$$\exists \delta > 0 : \sum_{i=1}^n P(\inf_{x \in \mathcal{E}} f_n(x) \leq \delta) < \infty$$

Démonstration- (*) $\inf_{x \in \mathcal{E}} f_n(x) \geq \inf_{x \in \mathcal{E}} E(f_n(x)) - \sup_{x \in \mathcal{E}} |f_n(x) - E(f_n(x))|.$

$\forall x \in \mathcal{E}$, on a :

$$E(f_n(x)) = (nh^d(n))^{-1} \sum_{i=1}^n E\left(K\left(\frac{x - X_i}{h(n)}\right)\right) = n^{-1} \sum_{i=1}^n \int h(n)^{-d} K\left(\frac{x - z}{h(n)}\right) P_x(dz)$$

Les hypothèses (N1) et (N3) permettent d'affirmer :

$\exists \tau > 0, \exists n_0 > 0 : E(f_n(x)) \geq \tau \int K(z) dz, \forall x \in \mathcal{C}, \forall n \in \mathbb{N} \text{ et } n \geq n_0.$

On pose $\delta = \tau/2$; (*) et le lemme 5 permettent de finir la démonstration.

Lemme 7- Sous (H4), si $\int \|u\| K(u) du < \infty$, on a :

$$\sup_{x \in \mathcal{C}} \sup_{y \in \mathbb{R}} |E(F^{1,n}(x,y) - F(y|x)E(f_n(x)))| \rightarrow 0, n \rightarrow \infty.$$

Démonstration- Pour tout $x \in \mathcal{C}, y \in \mathbb{R}$, on a :

$$\begin{aligned} & E(F^{1,n}(x,y) - F(y|x)E(f_n(x))) \\ &= E\left(\frac{1}{nh^d(n)} \sum_{i=1}^n 1_{\{Y_i < y\}} K\left(\frac{x-X_i}{h(n)}\right) - F(y|x) \frac{1}{nh^d(n)} E \sum_{i=1}^n K\left(\frac{x-X_i}{h(n)}\right)\right) \\ &= \frac{1}{n} \sum_{i=1}^n (E(h^{-d}(n) 1_{\{Y_i < y\}} K\left(\frac{x-X_i}{h(n)}\right)) - F(y|x)E(h^{-d}(n)K\left(\frac{x-X_i}{h(n)}\right))) \\ &= \frac{1}{n} \sum_{i=1}^n E\left[1_{\{Y_i < y\}} - F(y|x)\right] h^{-d}(n)K\left(\frac{x-X_i}{h(n)}\right). \quad (*) \end{aligned}$$

Un résultat bien connu sur l'espérance conditionnelle permet d'écrire :

$$(*) = \frac{1}{n} \sum_{i=1}^n E\left[E(1_{\{Y_i < y\}}|X_i) - F(y|x)\right] h^{-d}(n)K\left(\frac{x-X_i}{h(n)}\right).$$

Or d'après (N2), on a :

$$E(1_{\{Y_i < y\}}|X_i) = F(y|X_i),$$

donc

$$\begin{aligned} (*) &= \frac{1}{n} \sum_{i=1}^n E\left[F(y|X_i) - F(y|x)\right] h^{-d}(n)K\left(\frac{x-X_i}{h(n)}\right) \\ &\leq \frac{1}{n} \sum_{i=1}^n E\left[L|x-X_i\right] h^{-d}(n)K\left(\frac{x-X_i}{h(n)}\right). \end{aligned}$$

On pose :

$$t = \frac{x-X_i}{h(n)}.$$

L'hypothèse (H4) et l'hypothèse (N3) sur la loi marginale de X_1 , permettent d'avoir :

$(*) \leq h(n) L \int \|t\| K(t) dt \rightarrow 0$ car $h(n) \rightarrow 0, n \rightarrow \infty$. C. Q. F. D.

THEOREME 2: Sous (H1), (H2), (H3), (H4), (N1), (N2) et (N3), si le noyau K est lipschitzien et vérifie $\int \|u\| K(u) du < \infty$, alors

$$\sup_{x \in \mathcal{C}} |\mu_n(x) - \mu(x)| \xrightarrow{p. co.} 0.$$

Démonstration- Il suffit de démontrer que

$$(*) \sup_{x \in \mathcal{C}} \sup_{y \in \mathbb{R}} |F_n(y|x) - F(y|x)| \rightarrow 0$$

et d'utiliser le lemme 1 pour conclure.

Remarquons :

$$\sup_{x \in \mathcal{C}} \sup_{y \in \mathbb{R}} |F_n(y|x) - F(y|x)| \leq \frac{\sup_{x \in \mathcal{C}} \sup_{y \in \mathbb{R}} |F^{1,n}(x,y) - E(F^{1,n}(x,y))|}{\inf_{x \in \mathcal{C}} f_n(x)} + \frac{\sup_{x \in \mathcal{C}} |f_n(x) - E(f_n(x))| + \sup_{x \in \mathcal{C}} \sup_{y \in \mathbb{R}} |E(F^{1,n}(x,y)) - F(y|x)E(f_n(x))|}{\inf_{x \in \mathcal{C}} f_n(x)}$$

Les lemmes 3, 4, 5, 6 et 7 constituent les différentes étapes pour montrer (*) et le théorème en découle.

IV.3- Cas particulier :

Soit le processus α -mélangeant de coefficient $\alpha(k) = \rho^k$, $0 < \rho < 0$.

On choisit $k(n) = \beta n^\beta$, ($1 < \beta < 2$);

$$\alpha(k(n)) \frac{2k(n)}{3n} = \exp\left(\frac{2k(n)}{3n}\right) e^{-k(n)} \text{Log } \rho = \exp\left(\frac{2}{3} \beta n^{\beta-1} e^{-\beta n^\beta} \text{Log } \rho\right) \leq \exp\left(\text{Log } \rho \left[\frac{2}{3} n^{\beta-1} \exp n^\beta\right]\right).$$

Le majorant tend vers 0 quand n tend vers ∞ et (H2) est donc vérifiée.

Un simple calcul montre que $nh(n)^d/k(n)\text{Log}n \approx \frac{n^{1-\beta}h(n)^d}{\text{Log}n}$, il suffit de

choisir $h(n) = O\left[\left(\frac{n^\gamma \text{Log}n}{(1-\beta)/2}\right)^{1/d}\right]$, $\gamma \in]0, (1-\beta)/2[$ pour que (H1)

soit vérifiée.

V. APPLICATION A LA PREVISION :

Soit (Z_n) un processus réel strictement stationnaire et α -mélangeant.

On veut prédire Z_{N+1} à partir des observations Z_1, \dots, Z_N .

Pour cela, on cherche à évaluer $Z_{N+1}^* = \mu(Z_{N+1} | Z_1, \dots, Z_N)$ où d est choisi de façon convenable (cf. [12]).

Il suffit de poser pour $N > d$

$X_t = (Z_t, \dots, Z_{t+d-1})$, $Y_t = (Z_{t+d})$, $n = N-d$ pour retrouver le schéma précédent.

Le prédicteur non paramétrique est alors défini par :

$$\hat{Z}_{N+1} = \mu_n(X_n) \text{ (resp } \hat{Z}_{N+1} = \hat{\mu}_n(X_n)) \text{ et on a :}$$

Corollaire- Sous les conditions du théorème 2 (resp. du théorème 1),

$$\left| \hat{Z}_{N+1} - Z_{N+1}^* \right|_1 \Big|_{(Z_{N-d+1}, \dots, Z_N) \in \mathcal{C}} \xrightarrow{\text{P.S.}} 0$$

Démonstration- Remarquons que la condition d' α -mélangeance sur $(Z_n)_N$ implique que le processus $X_t = (Z_t, \dots, Z_{t+d-1})$ est aussi α -mélangeant.

Ensuite $|\hat{Z}_{N+1} - Z_{N+1}^*|_1 |_{(Z_{N-d+1}, \dots, Z_N) \in \mathcal{G}} \leq \sup_{x \in \mathcal{G}} |\mu_n(x) - \mu(x)| \xrightarrow{P. co.} 0$

d'après le théorème 2 (resp. le théorème 1).

VI. CONCLUSION :

Les simulations et les comparaisons avec d'autres méthodes de prévisions paramétrique (BOX et JENKINS) et non paramétrique (Régression) montrent l'efficacité de la méthode exposée en raison de sa robustesse (cf. [13]), elles feront l'objet d'une prochaine publication.

REFERENCES BIBLIOGRAPHIQUES :

- [1] BOSQ, D.: Contribution à l'estimation fonctionnelle. Pub. ISUP., (19), fasc. (II) et (III) (1970).
- [2] GEFFROY, J.: Etude de la convergence du regressogramme. Séminaire de ISUP. (1975).
- [3] COLLOMB, G. : Propriétés de convergence presque complète du prédicteur à noyau. Z. Wahr. verw. geb., 66, 441-460 (1984).
- [4] COLLOMB, G. et HARDLE, W. : Stochastic Process. App., 23, 77-89 (1986).
- [5] VIEU, P. : Thèse d'Université, Université P. SABATIER, Toulouse (1987).
- [6] SARDA, P. : Thèse d'Université, Université P. SABATIER, Toulouse (1987).
- [7] ROUSSAS, G. : Journal of Statistical Planning and Inference, 18, 135-149 (1988).
- [8] BOSQ, D. : Prévision non paramétrique d'un processus stationnaire non borné. Pub. LSTA (81) (1988).
- [9] BOSQ, D. : Estimation et prévision non paramétrique d'un processus stationnaire. C. R. Acad. Sci. Paris, t. 308, Série I, 453-456 (1989).
- [10] STONE, C. : Consistant nonparametric regression. Ann. Stat., Vol. 10, N°4, 1040-1053 (1976).
- [11] BRADLEY, R. : Dependence in probability and statistics. Eberlin Taqqu, 165-192 (1986).
- [12] COLLOMB, G., HARDLE, W. et HASSANI, S. : Journal of Statistical Planning and Inference, Vol. 15, 227-236 (1987).
- [13] GANNOUN, A. : Thèse d'Université, Université PARIS VI (1989).
- [14] ROSEMBLATT, M. : Remarks on some non parametric estimates. A.M.S., 27, 832-837 (1956).
- [15] CARBON, M. : Inégalité de BERNSTEIN pour les processus. C. R. Acad. Sci. Paris, t. 297, Série I (1983).
- [16] CARBON, M. et DELECROIX, M. : Nonparametric forecasting in time series. A paraître (1989).