STATISTIQUE ET ANALYSE DES DONNÉES

GÉRARD COLLOMB

Prédiction non paramétrique : étude de l'erreur quadratique du prédictogramme

Statistique et analyse des données, tome 9, nº 1 (1984), p. 1-34

<http://www.numdam.org/item?id=SAD_1984__9_1_1_0>

© Association pour la statistique et ses utilisations, 1984, tous droits réservés.

L'accès aux archives de la revue « Statistique et analyse des données » implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Statistiques et Analyse de données .1.
mai 1984 - Vol. 9 n° 1 pp. 1-34

PREDICTION NON PARAMETRIQUE : ETUDE DE L'ERREUR QUADRATIQUE DU PREDICTOGRAMME.

Gérard COLLOMB

Laboratoire de Statistique et Probabilités
E.R.A.-C.N.R.S. n°591
Université Paul Sabatier - 31062 TOULOUSE CEDEX

Résumé : *Soit $(X_n)_{\mathbb{N}}$ un processus strictement stationnaire à valeurs dans $[0,1]^p$, g une fonction réelle définie sur $[0,1]^p$, s et k deux entiers et $(\mathcal{P}_n)_{\mathbb{N}}$ une suite de partition de $[0,1]^{kp}$. Le prédictogramme associé à X_i, $i=1,\ldots,n$ et $\mathcal{P}_n$ est la moyenne des $g(X_{i+s})$, $i=k,\ldots,n-s$ tels que les vecteurs $(X_{i-k+1},\ldots,X_i)$ et $(X_{n-k+1},\ldots,X_n)$ appartiennent au même élément de $\mathcal{P}_n$. La v.a.r. ainsi définie est étudiée en tant qu'estimateur de la v.a.r. $E(g(X_{n+s})/[X_{n-k},\ldots,X_n])$, qui est le meilleur prédicteur de $g(X_{n+s})$ à partir de $(X_i,\ i=1,\ldots,n)$ lorsque $(X_n)_{\mathbb{N}}$ est markovien d'ordre k.*

Abstract :
Let $(X_n)_{\mathbb{N}}$ be a strictly stationary process which is valued in $[0,1]^p$, g a real measurable function defined on $[0,1]^p$, s and k two positive integers and $(\mathcal{P}_n)_{\mathbb{N}}$ a sequence of partitions of $[0,1]^{kp}$. The predictogram associated with $(X_i,\ i=1,\ldots,n)$ and $\mathcal{P}_n$ is the average of the $g(X_{i+s})$, $i=k,\ldots,n-s$, such that the vectors $(X_{i-k+1},\ldots,X_i)$ and $(X_{n-k+1},\ldots,X_n)$ belong to the same element of $\mathcal{P}_n$. This predictogram is considered as an estimator of the r.r.v. $E(g(X_{n+s})/[X_{n-k+1},\ldots,X_n])$ which is the best predictor of $g(X_{n+s})$ from $(X_i,\ i=1,\ldots,n)$ when $(X_n)_{\mathbb{N}}$ is markovian of order k. In the more general case of an uniformly strong mixing process and under unrestrictive non parametric assumptions on the distribution of $(X_1,\ldots,X_{k+s})$ we give for this estimate a necessary and sufficient condition (on $\mathcal{P}_n$) of L^2-consistency and a majorization of its mean square error leading to a result on its rate of convergence.

Mots clés : *analyse d'une série temporelle ; autoregression ; erreur quadratique moyenne ; non paramétrique ; prédicteur ; prédictogramme ; processus markovien, ϕ-mélangeant, stationnaire; vitesse de convergence.*

1 - INTRODUCTION

Soit $(X_n)_{\mathbb{N}}$ un processus strictement stationnaire dont l'espace d'états est $\mathbb{R}^p$, $p \in \mathbb{N}_*$, ou une partie non vide C de $\mathbb{R}^p$. On désigne par g une fonction réelle mesurable définie sur $\mathbb{R}^p$ et par s et k deux entiers positifs.

On considère le problème de la prediction de $g(X_{n+s})$ à l'aide de la suite $\{X_i, i=1,\dots,n\}$ et on s'intéresse au problème de "l'estimation", à partir de cette suite, de la v.a.r.

(1-1) $\quad R(X_{n-k+1},\dots,X_n) = E[g(X_{n+s})/(X_{n-k+1},\dots,X_n)]$

Nous étudions l'*erreur quadratique* d'un "estimateur" non paramétrique (e.n.p.) de cette v.a.r., qui est défini au début du paragraphe 2 et dénommé *prédictogramme* en raison de l'analogie de sa définition avec celle de l'histogramme (e.n.p de la densité) du periodogramme (e.n.p. de la densité spectrale) ou du regressogramme (e.n.p. de la régression).

Lorsque $(X_n)_{\mathbb{N}}$ est un *processus markovien d'ordre k*, on a

(1-2) $\quad R(X_{n-k+1},\dots,X_n) = E[g(X_{n+s}) \; / (X_1,\dots,X_n)]$.

Cette v.a.r. est alors le meilleur (pour une fonction de perte quadratique) *prédicteur probabiliste* (Bosq, 1979) de $g(X_{n+s})$. Le *predictogramme*, en tant qu'estimateur du predicteur probabiliste, est un *prédicteur statistique* (Bosq, 1979) ou *prédicteur*.

Ce prédicteur est qualifié de *non paramétrique* parce que le problème de la prédiction ne se réduit pas à celui de l'estimation d'un nombre fini de paramètres réels définissant R : on supposera uniquement que R est bornée, ou, parfois, Lipschitzienne, et la loi de $[X_1,\dots,X_k]$ sera seulement assujettie à de simples hypothèses de régularité, $(X_n)_{\mathbb{N}}$ étant uniformément fortement mélangeant.

La définition du prédictogramme et les résultats obtenus sur son erreur quadratique sont énoncés dans le paragraphe 2 qui donne également diverses remarques concernant les hypothèses introduites et la nature de ces résultats. Leurs démonstrations sont regroupées dans le paragraphe 3. Dans le paragraphe 4 nous considérons leur application à l'estimation de la probabilité de transition d'un processus de Markov. Le paragraphe 5 est constitué de remarques sur la définition et l'utilisation du prédictogramme.

Travaux portant sur des problèmes analogues.

Un premier problème très proche de celui abordé ici est le problème de l'estimation de la fonction R ou de la fonction d'autoregression d'ordre k, à valeurs dans $\mathbb{R}^p$, notée $\bar{R}$ et définie par

$$\bar{R}(u) = E(X_{k+s} \,/\, [X_1,\ldots,X_k] = u) \ , \quad \forall\, u \in \mathbb{R}^{kp} .$$

Pour un problème de statistique appliquée (prévision de température en météorologie), Watson (1964, p.369-370) propose la méthode du noyau consistant à estimer la fonction $\bar{R}$ par la fonction

$$(1\text{-}3) \quad \rho_n(.) = \sum_{i=k}^{n-s} X_{i+s}\, K((.-[X_{i-k+1},\ldots,X_i])/h_n) \Big/ \sum_{i=k}^{n-s} K((.-[X_{i-k+1},\ldots,X_i])\,/\,h_n)$$

où K est un noyau de $\mathbb{R}^p$ et $(h_n)_{\mathbb{N}}$ une suite réelle strictement positive de limite nulle. Divers auteurs ont ensuite donné des résultats concernant les propriétés asymptotiques de cet e.n.p. fonctionnel, sous diverses conditions portant sur $(X_n)_{\mathbb{N}}$ et $(h_n)_{\mathbb{N}}$: Roussas (1969) obtient d'abord une propriété de convergence ponctuelle en probabilité, Robinson (1982) donne la loi limite de $\rho_n(u) - \bar{R}(u)$ convenablement normalisé pour u fixé, Collomb et Doukhan (1983) fournissent une majoration de l'erreur quadratique moyenne ponctuelle ou intégrée et Collomb (1982 a) obtient une propriété de convergence uniforme (sur un compact) presque sûre.

Cette même propriété est établie par Collomb (1982 b) pour la méthode des k points les plus proches, Banon (1978) Nguyen et Pham (1981) et Pham (1981) s'intéressent aux propriétés de convergence ponctuelle d'estimateurs récursifs définis également à l'aide d'un noyau. Par ailleurs, Kalaidjian (1981), Doukhan (1983) et Bosq (1983) donnent une étude par simulation de l'estimateur fonctionnel (1-3).

Il est clair cependant que ce problème d'estimation fonctionnelle ne s'identifie pas au problème abordé dans le présent article. En effet le problème de la prédiction se pose de manière naturelle (cf. le début de la présente introduction) non pas en un point x donné a priori ou pour un ensemble de points x, mais, si k=1 par exemple, au point X_n qui, de par la nature même du problème, n'est pas indépendant de $(X_1,\dots,X_{n-1})$. L'étude d'un e.n.p. ρ_n de $\bar{R}$ doit donc porter alors sur la v.a. $\rho_n(X_n) - \bar{R}(X_n)$. Cette approche est celle de Bosq (1979,1983) qui étudie la quantité $E|\rho_n(X_n) - \bar{R}(X_n)|^2$ pour une classe d'e.n.p. $\rho_n(X_n)$ contenant à la fois l'estimateur à noyau et le prédictogramme. Cependant, ce dernier auteur suppose que la loi de X_1 est connue, ce qui constitue une hypothèse assez restrictive en raison de la nature non paramétrique du problème. La convergence presque sûre vers 0 de la v.a.r. $|\rho_n(X_n) - \bar{R}(X_n)|$ est obtenue par Collomb (1982 a) [resp. 1982 b] pour le prédicteur non paramétrique $\rho_n(X_n)$ défini par la méthode du noyau [resp. par la méthode des k points les plus proches].

Les résultats que nous avons obtenus sur le prédictogramme, notamment les propositions 1 et 4, sont à rapprocher de ceux de Stone (1977), Devroye et Wagner (1980) et Spiegelman et Sacks (1980) relatifs à la propriété de "convergence universelle" de certains e.n.p., de la régression. Nous renvoyons à Collomb (1981) pour une revue bibliographique sur les e.n.p. de la régression, le paragraphe 2 de ce dernier papier concernant uniquement le régressogramme.

2 - DEFINITIONS ET PROPOSITIONS.

Nous considérons le cas où le processus $(X_n)_{\mathbb{N}}$ prend ses valeurs dans $C = [0,1]^p$. Le *prédictogramme* est défini comme suit (on suppose $n-s > k$)

> pour tout n fixé, soit $\mathcal{S}_n$ une partition de $[0,1]^{kp}$: le prédictogramme $R_n(X_{n-k+1},\dots,X_n)$ associé à $\mathcal{S}_n$ est égal à la moyenne des $g(X_{i+s})$, $i = k,\dots,n-s$, tels que le vecteur $[X_{i-k+1},\dots,X_i]$ appartienne à l'élément de $\mathcal{S}_n$ contenant $[X_{n-k+1},\dots,X_n]$, à 0 si aucun vecteur $[X_{i-k+1},\dots,X_i]$ n'appartient à cet élément de $\mathcal{S}_n$,

ce qui peut s'écrire

$$
\begin{aligned}
R_n(X_{n-k+1},\dots,X_n) &= \frac{1}{|J|}\sum_{i=k}^{n-s} g(X_{i+s})1_{\{(X_{i-k+1},\dots,X_i)\in J\}} \quad \text{si } |J| \neq 0, \\
&= 0 \text{ sinon}, \qquad (2\text{-}1)
\end{aligned}
$$

$$\text{avec } J \in \mathcal{S}_n : (X_{n-k+1},\dots,X_n)\in J \text{ et } |J| = \sum_{i=k}^{n-s} 1_{\{(X_{i-k+1},\dots,X_i)\in J\}}.$$

On pose, pour toute partition $\mathcal{S}_n$,

(2-2) $$h_n = \inf\{h \in \mathbb{R}^+ : \forall B \in \mathcal{S}_n,\ \exists\, u \in [0,1]^{kp} : B \subset [u-h,\ u+h]^{kp}\}$$

et on suppose que $(h_n)_{\mathbb{N}}$ est de limite nulle, ainsi que

(2-3) $$\exists d > 0 : \forall n \in \mathbb{N},\ \mu(A) \geq d h_n^{kp},\ \forall A \in \mathcal{S}_n,$$

où μ est la mesure de Lebesgue sur $[0,1]^{kp}$ (si chaque élément de $\mathcal{S}_n$ est un pavé de côté h_n, on a $d = 1$).

On associe au processus $(X_n)_{\mathbb{N}}$ la condition de mélange suivante (Billingsley, 1968, p.166)

(I) *le processus $(X_n)_{\mathbb{N}}$ est ϕ-mélangeant, en ce sens que il existe une suite réelle positive $(\phi_n)_{\mathbb{N}}$ de limite nulle satisfaisant pour tous entiers $n \geq 0$ et $i > 0$*

(2-4) $$|P(A \cap B) - P(A)\,P(B)| \leq \phi_n\, P(A)$$

dès que A [resp. B] est un ensemble $\sigma(X_1,\dots,X_i)$ mesurable [resp. $\sigma(X_{i+n}, X_{i+n+1},\dots)$ mesurable],

en notant que (I) est satisfaite dès que $(X_n)_{\mathbb{N}}$ est un processus melangeant au sens de la théorie ergodique et pour lequel $\exists\, \nu \in \mathbb{N} : \phi_\nu < 1$ (Bradley, 1980); on définit également pour le processus $(X_n)_{\mathbb{N}}$ l'hypothèse

(II) | *la loi du vecteur aléatoire* $(X_1,\dots,X_k)$ *dans* $[0,1]^{kp}$ *vérifie*

(2-5) | $\exists m, M, m > 0, 0 < M < \infty : m\mu(B) \leq P[(X_1,\dots,X_k) \in B] \leq M\mu(B),\ \forall B \in \mathcal{B}_{[0,1]^{kp}}$.

On suppose en outre que la fonction g vérifie

(2-6) $\quad |g(z)| < G \quad \forall z \in [0,1]^p$

où G est une constante de $\mathbb{R}$.

Le premier résultat que nous donnons est un résultat de convergence en moyenne quadratique pour le prédictogramme.

Proposition 1 .

Si le processus $(X_n)_{\mathbb{N}}$ *satisfait les hypothèses* (I) *et* (II) *et si* $(h_n)_{\mathbb{N}}$ *vérifie*

(2-7) $$nh_n^{kp} / \sum_{i=0}^{n} \phi_i \xrightarrow[n \to \infty]{} \infty$$

on a

(2-8) $$R_n(X_{n-k+1},\dots,X_n)-R(X_{n-k+1},\dots,X_n) \xrightarrow[n \to \infty]{L_2} 0 .$$

Remarque 1

Ce résultat sur l'estimation de l'*espérance conditionnelle* $R(X_{n-k+1},\dots,X_n)$ est à rapprocher du résultat relatif à l'estimation de l'espérance d'une v.a.r. bornée X_1 à partir d'une suite $\{X_1,\dots,X_m\}$ stationnaire et φ-mélangeante (on a $\frac{1}{m}\sum_{i=1}^{m} X_i \xrightarrow[m \to \infty]{L_2} E X_1$ dès que $m/(\sum_{\ell=0}^{m} \phi_\ell) \xrightarrow[m \to \infty]{L_2} \infty$, voir Billingsley, 1968, formules (20.32) et (20.29), p. 172) : nh_n^{kp} joue ici le rôle de m.

Nous précisons maintenant le résultat (2-8) dans le cas où le processus $(X_n)_{\mathbb{N}}$ vérifie une hypothèse supplémentaire concernant la fonction R :

(III) | *la fonction de prédiction R est Lipschitzienne, en ce sens que*

(2-9) | $\exists c > 0, c < \infty$ et $\gamma > 0 : |R(z)-R(z')| \leq c\,||z-z'||^{\gamma},\ \forall z,z' \in [0,1]^{kp}$

où $||\cdot||$ désigne la norme de la convergence uniforme sur $[0,1]^{kp}$.

Proposition 2.

On pose

(2-10) $\quad t = k + s - 1\ ,\ N = n - t\ $ et $\ \phi_n^* = t + \sum_{i=1}^{N} \phi_i\ .$

Si le processus $(X_n)_N$ vérifie (I), (II) et (III), la v.a.r.

$$Q_n = R_n(X_{n-k+1},\dots,X_n) - R(X_{n-k+1},\dots,X_n)$$

vérifie

(2-11) $$EQ_n^2 \leq M\left(\frac{4G}{md}\right)^2 \frac{1+8\phi_n^*}{N\, h_n^{kp}} + 2c^2\, h_n^{2\gamma} + O\left(\frac{\phi_n^*}{nh^{kp}}\right)^2$$

Remarques 2

Si nous examinons l'inégalité (2-11) nous voyons que ce résultat mathématique sur le prédictogramme confirme les affirmations intuitives suivantes : pour toute partition $\mathcal{S}_n$ (et donc h_n) fixée, la prédiction sera d'autant meilleure que

. la dimension p du vecteur X_1 ou le nombre k est petit

. la fonction R est "lisse" : la constante c est petite ou l'ordre γ est grand

. la loi de $(X_1,\dots,X_k)$ est "plutôt" uniforme : les constantes m et M ne sont pas trop différentes

. la "dépendance" entre X_i et X_j , pour tous entiers i et j, est petite en moyenne lorsque $|i-i'|$ est grand : même si $\sum_{i=1}^{n} \phi_i$ tend vers l'infini lorsque n croît, cette quantité doit rester petite devant n.

Nous introduisons enfin sur le processus $(X_n)_{\mathbb{N}}$ la condition

(IV) *la suite $(\phi_n)_{\mathbb{N}}$ vérifie*

$$\sum_{n=1}^{\infty} \phi_n \leq L < \infty$$

en notant que cette condition est satisfaite dans le cas particulier important (en raison de l'application privilégiée au problème de la prédiction - formule 1-2 - évoqué en introduction) où $(X_n)_{\mathbb{N}}$ est *markovien* et vérifie la *condition de Doeblin* (Doob, 1953, p.192) ou la *condition de norme* L_p, p=1 ou ∞ , de Rosenblatt (1971, p.206-211), ou lorsque ce processus s'écrit sous la forme plus classique "$X_{n+1} = f(X_n) + \varepsilon_n$" dès que la fonction f et le bruit $(\varepsilon_n)_{\mathbb{N}}$ sont assujettis à des conditions de régularité données par Doukhan et Ghindès (1980).

La proposition 2 donne immédiatement le résultat suivant concernant un "cas usuel" de processus.

<u>Corollaire</u>

Si le processus $(X_n)_{\mathbb{N}}$ vérifie (II), (III) avec $\gamma=1$ et (IV), on a

(2-12) $$EQ_n^2 \leq \frac{a^2}{Nh_n^{kp}} + 2\, c^2\, h_n^2 + O\left(\frac{1}{nh_n^{kp}}\right)^2 \text{ avec } a^2 = \left(\frac{4G}{md}\right)^2 \|\Gamma\|1+8(\ell+L)| .$$

<u>Note</u> : pour k=1, le résultat (2-12) est à rapprocher des resultats relatifs à l'erreur quadratique de l'histogramme ou du régressogramme (majoration de la forme $a^2/(nh_n^p) + b^2\, h_n^2$, voir Collomb, 1978, p. 24) en un point x fixé.

La formule (2-11) permet d'obtenir aisément une information sur la vitesse optimale de L_2-convergence du prédictogramme. Cette information est donnee par une majoration du minimum de la fonction $\mathcal{S}_n \longrightarrow E\, Q_n^2(\mathcal{S}_n)$, où $E\,Q_n^2(\mathcal{S}_n)$ est l'erreur quadratique du prédictogramme associé à une partition $\mathcal{S}_n$.

<u>Proposition 3</u>

Si le processus $(X_n)_{\mathbb{N}}$ vérifie les hypothèses de la proposition 2, on a

(2-13) $$\inf_{\mathcal{S}_n \in \pi} E\,Q_n^2(\mathcal{S}_n) \leq A\left(\frac{\phi_n^*}{n}\right)^{\frac{2\gamma}{2\gamma+kp}}$$

où π est l'ensemble des partitions de $[0,1]^{kp}$ vérifiant les hypothèses (2-2) et (2-3) *et A est une constante (indépendante de n) positive.*

<u>Remarques 3</u>

Si de plus $\gamma=1$ et (IV) est vérifiée, on a pour k=1

$$\inf_{\mathcal{S}_n \in \pi} E\,Q_n^2(\mathcal{S}_n) \leq A\, n^{-2/(kp+2)} , \quad 0 < A < \infty .$$

On remarque alors que la vitesse de convergence définie par la partie droite de l'inégalité ci-dessus est identique à la vitesse optimale de convergence de l'histogramme ou du régressogramme, lorsqu'on considère l'erreur quadratique intégrée (voir Collomb, 1978). Par ailleurs, ce résultat sur un e.n.p. de la fonction R, définie sur $\mathbb{R}^p$ et Lipschitzienne d'ordre 1, est à rapprocher des résultats de Parzen (1962), Collomb (1977) et Collomb et Doukhan (1983) relatifs à la vitesse de convergence d'e.n.p. d'une densité, d'une régression ou d'une fonction de prédiction R, définie sur $\mathbb{R}^p$ et dérivable à l'ordre q, $q \in \mathbb{N}$: les divers e.n.p. considérés ici atteignent alors la vitesse optimale de convergence donnée par Bretagnolle et Huber (1979) à propos des e.n.p. de la densité et par

Stone (1982) à propos des e.n.p. de la régression.

Enfin, nous complétons la proposition 1 en donnant une condition sur h_n qui est, en un certain sens, nécessaire et suffisante pour (2-8).
On désigne par $\mathcal{D}_0$ l'ensemble des processus $\xi = (X_n)_{\mathbb{N}}$ dont l'espace d'états est $[0,1]^p$ et qui vérifient les hypothèses (II) et (IV). [On suppose en outre ici que la fonction g est telle qu'il existe dans $\mathcal{D}_0$ un processus $(X_n)_{\mathbb{N}}$ pour lequel on a $[Eg(X_1)]\, E[g(X_1) - Eg(X_1)]^2 \neq 0$].

Proposition 4

Pour que

(2-14) $\forall\, \xi \in \mathcal{D}_0\ ,\quad R_n(X_{n-k+1},\dots,X_n) - R(X_{n-k+1},\dots,X_n) \xrightarrow[n\to\infty]{L_2} 0$

il faut et il suffit que $(h_n)_{\mathbb{N}}$ *vérifie*

(2-15) $n\, h_n^{kp} \xrightarrow[n\to\infty]{} \infty$.

Note : on retrouve là sur la suite $(h_n)_{\mathbb{N}}$ une condition nécessaire de convergence L_2 analogue aux conditions nécessaires de convergence L_2 d'e.n.p. classiques de la régression, considérées par Devroye (1982).

Nous formulons enfin une remarque concernant l'ensemble des résultats que nous venons d'énoncer.

Remarque 4.

Les propriétés asymptotiques obtenues ici pour le prédictogramme $R_n(X_{n-k+1},\dots,X_n)$ s'étendent immédiatement à l'estimateur

$$\bar{R}_n(X_{n-k+1},\dots,X_n) = |J|^{-1} \sum_{i=k}^{n-s} X_{i+s}\, 1_{\{(X_{i-k+1},\dots,X_i)\in J\}} \text{ si } |J| \neq 0, \text{ 0 sinon,}$$

de $E(X_{n+s} / [X_{n-k+1},\dots,X_n])$, prédicteur (probabiliste) de X_{n+s} lorsque le processus $(X_n)_{\mathbb{N}}$ est markovien d'ordre k : il suffit d'appliquer les propositions ci-dessus, successivement pour tout $\lambda=1,\dots,p$, à la fonction g qui associe à tout vecteur de $[0,1]^p$ sa λième coordonnée.

3 - DEMONSTRATIONS.

On donne d'abord un ensemble de notations simplificatrices et quelques lemmes préliminaires.

3.1. notations.

On simplifie les notations en posant, outre (2-10),

$$\tilde{X}_i = [X_{i-k+\ell}, \ \ell=1,\dots,t+1] \quad \text{pour} \quad i=k,\dots,n ,$$

$$\text{"}\sum_j\text{"} = \text{"}\sum_{j\in\mathcal{S}_n}\text{"} \quad \text{et} \quad \text{"}\sum_i\text{"} = \sum_{i=k}^{\text{"}n-s\text{"}} , \text{ avec éventuellement } i\neq i' ,$$

et, pour toute fonction ρ définie sur $[0,1]^{kp}$, en désignant également par ρ la la fonction $\tilde{\rho}$ définie sur $[0,1]^{p(t+1)}$ par

$$\tilde{\rho}(z_1,\dots,z_{p(t+1)}) = \rho(z_1,\dots,z_{kp}) .$$

On note, pour tout j de $\mathcal{S}_n$,

$$I_j(z_1,\dots,z_{p(t+1)}) = 1_{\{(z_1,\dots,z_{kp})\in j\}}$$

et on pose 0/0=0, en remarquant tout de suite que (2-1) s'écrit alors

$$(3\text{-}1) \qquad R_n(\tilde{X}_n) = N^{-1}\, D_n^{-1} \sum_i \sum_j g(X_{i+s})\, I_j(\tilde{X}_i)\, I_j(\tilde{X}_n)$$

où

$$D_n = N^{-1} \sum_i \sum_j I_j(\tilde{X}_i)\, I_j(\tilde{X}_n) .$$

Par ailleurs, on pose $\tilde{X} = \tilde{X}_k$,

$$p_j = E\, I_j(\tilde{X}) = P(j) \ , \ \forall\, j\in\mathcal{S}_n ,$$

et

$$H = h^{kp}$$

en remarquant tout de suite que les hypothèses (2-2), (2-3) et (2-5) impliquent

$$(3\text{-}2) \qquad m\,d\,H \le p_j \le M\,H \quad , \quad \forall\, j\in\mathcal{S}_n .$$

3.2. lemmes préliminaires.

Dans le lemme 1 et à l'intérieur de la démonstration des lemmes qui le suivent, on étend la définition (2-4, pour $\ell \in \mathbb{N}$) au cas $\ell \in \mathbb{Z}$, en posant

$$\phi_\ell = 1 \quad , \quad \forall\, \ell \le 0 \; .$$

lemme 1

Si a et b sont deux fonctions réelles définies sur $[0,1]^{p(t+1)}$ avec

$$E|a(\tilde{X})| \le \bar{a} \quad \text{et} \quad |b(\tilde{X})| \le b \; ,$$

on a, pour tout $\nu \in \mathbb{N}_*$ et $\nu' \in \mathbb{N}_*$, $\nu' > \nu$,

$$Ea(\tilde{X}_\nu)\, b(\tilde{X}_{\nu'}) = Ea(\tilde{X})Eb(\tilde{X}) + O_1(2\bar{a}b\, \phi_{\nu'-\nu-t})$$

où $O_1(.)$ vérifie $|O_1(z)| \le |z|$.

démonstration

Pour $\nu'-\nu>t$, on utilise le résultat (20-28) de Billingsley (1968;p.171, avec ξ et ζ définis dans l'énoncé du lemme 1, p.170) où ξ est remplacé par $a(\tilde{X}_\nu)$, k par $\nu+s$, ζ par $b(\tilde{X}_{\nu'})$, k+n par $\nu'-k+1$ et C par b.

Pour $\nu'-\nu \le t$, on utilise l'inégalité

$$|E\,a(\tilde{X}_\nu)\, b(\tilde{X}_\nu) - E\,a(\tilde{X}_\nu)\, E\,b(\tilde{X}_\nu)| \le E|a(\tilde{X}_\nu)||b(\tilde{X}_{\nu'})| + E|a(\tilde{X})|\, E|b(\tilde{X})|$$

en tenant compte des hypothèses sur $|a(\tilde{X})|$ et $|b(\tilde{X})|$ et de la notation $\phi_\ell = 1$ pour $\ell \le 0$.

lemme 2

Si ζ est une fonction réelle positive bornée vérifiant

$$E\,\zeta(X) < \varepsilon$$

alors la v.a.r. $T = U$ ou V, avec

$$U = N^{-1} \sum_i \sum_j \zeta(\tilde{X}_i)\, I_j(\tilde{X}_i)\, I_j(\tilde{X}_n)$$

et

$$V = N^{-1} \sum_i \sum_j \zeta(\tilde{X}_n)\, I_j(\tilde{X}_i)\, I_j(\tilde{X}_n)$$

vérifie

$$FH^{-1}T \le M\varepsilon + M' \frac{\psi_n^*}{HN} \; , \quad M' < \infty .$$

démonstration

Le lemme 1 permet d'écrire

$$E([\zeta(\tilde{X}_i)\ I_j(\tilde{X}_i)]\ I_j(\tilde{X}_n)) = E\zeta(\tilde{X})\ I_j(\tilde{X})p_j + O_1(2Cp_j\ \phi_{n-i-t})$$

et

$$E(I_j(\tilde{X}_i)[\zeta(\tilde{X}_n)I_j(\tilde{X}_n)]) = E\zeta(X)\ I_j(X)p_j + O_1(2p_j C\phi_{n-i-t})$$

où C est une constante bornant ζ .

Ces deux inégalités et le résultat (3-2) donnent, en tenant compte du fait que $\sum_j P_j = 1$ et $\sum_j \zeta(\tilde{X})\ I_j(\tilde{X}) = \zeta(\tilde{X})$

$$EH^{-1}\,T \le M\varepsilon + \frac{M'}{HN} a_n \quad , \quad M' < \infty,$$

où

$$a_n = \sum_{i=k}^{n-s} \phi_{n-i-t} .$$

On a

$$a_n = \sum_{\ell=1-k}^{n-2k-s} \phi_\ell$$

qui implique

$$a_n \le \phi_n^*$$

ce qui achève la démonstration du lemme 2.

lemme 3

Si $\{\gamma_j,\ j \in \mathcal{S}_n\}$ est une suite de fonctions réelles mesurables définies sur $[0,1]^{p(t+1)}$ et telles que pour tout j de $\mathcal{S}_n$

$$E\gamma_j(\tilde{X})=0\ ,\ E\gamma_j^2(\tilde{X}) \le C^2 p_j,\ E|\gamma_j(\tilde{X})| \le Cp_j \quad \text{et} \quad |\gamma_j(\tilde{X}_i)| \le C\ ,$$

avec $C < \infty$, alors la v.a.r.

$$\Gamma = (HN)^{-1} \sum_i \sum_j \gamma_j(\tilde{X}_i)\ I_j(\tilde{X}_n)$$

vérifie

$$E\Gamma^2 \le MC^2 \frac{1+8\phi_n^*}{HN} + C'\left(\frac{\phi_n^*}{HN}\right)^2 , \quad C' < \infty \ .$$

démonstration

On a

(*1) $$E\Gamma^2 = H^{-2}(EA + 2EB)$$

avec, puisque $I_j(z)\, I_{j'}(z) = 1_{\{j=j'\}}\, I_j(z)$,

$$A = N^{-2} \sum_i \sum_j \gamma_j^2(\tilde{X}_i)\; I_j(\tilde{X}_n)$$

et

(*2) $$B = N^{-2} \sum_i \sum_{i'>i} \sum_j B_{ii'j}$$

où

$$B_{ii'j} = \gamma_j(\tilde{X}_i)\; \gamma_j(\tilde{X}_{i'})\; \gamma_j(\tilde{X}_n).$$

Le lemme 1 permet d'écrire

$$E\gamma_j^2(\tilde{X}_i)\; I_j(\tilde{X}_n) = E\gamma_j^2(\tilde{X})\, E\, I_j(\tilde{X}) + 0_1(2E\gamma_j^2(\tilde{X})\phi_{n-i-t})$$

qui donne

$$E\gamma_j^2(\tilde{X}_i)\; I_j(\tilde{X}_n) \le C^2 p_j^2 + 2C^2\; p_j\; \phi_{n-i-t}$$

d'où il vient, puisque $\sum_j p_j = 1$ et par (3-2),

(*3) $$EA \le MC^2\; HN^{-1} + 2C^2\; a_n\; N^{-2}$$

où

$$a_n = \sum_{i=k}^{n-s} \phi_{n-i-t} \quad .$$

En utilisant à nouveau le lemme 1, on obtient d'abord

$$EB_{ii'j} = E\gamma_j(X_i)\gamma_j(\tilde{X}_{i'})EI_j(\tilde{X}) + 0_1(2E|\gamma_j(\tilde{X}_i)\gamma_j(\tilde{X}_{i'})|\phi_{n-i'-t}),$$

ensuite

$$E\gamma_j(\tilde{X}_i)\gamma_j(\tilde{X}_{i'}) = 0_1(2E|\gamma_j(\tilde{X})|C\phi_{i'-i-t})$$

et enfin

$$E|\gamma_j(\tilde{X}_i)\gamma_j(\tilde{X}_{i'})| = E^2|\gamma_j(\tilde{X})| + 0_1(2E|\gamma_j(\tilde{X})|C\phi_{i'-i-t})$$

qui impliquent

$$EB_{ii'j} \le 2C^2 \left(p_j^2\; \phi_{i'-i-t} + p_j^2\phi_{n-i'-t} + 2p_j\; \phi_{n-i'-t}\; \phi_{i'-i-t} \right)$$

d'où il vient par (*2)

(*4) $|EB| \le 2MC^2[(b_n+c_n)HN^{-2} + 2d_n N^{-2}]$

où

$$b_n = \sum_{i=k}^{n-s} \sum_{i'=i+1}^{n-s} \phi_{i'-i-t}$$

$$c_n = \sum_{i=k}^{n-s} \sum_{i'=i+1}^{n-s} \phi_{n-i'-t}$$

et

$$d_n = \sum_{i=k}^{n-s} \sum_{i'=i+1}^{n-s} \phi_{i'-i-t}\, \phi_{n-i'-t} \,.$$

L'égalité (*1) et les inégalités (*3) et (*4) amènent

(*5) $E\Gamma^2 \le MC^2[1+4N^{-1}(b_n+c_n)] \,/\, (HN) + De_n \,, \quad D < \infty \,,$

où

$$e_n = (a_n + d_n) \,/\, (HN)^2 .$$

La quantité b_n peut s'écrire

$$b_n = \sum_{i=k}^{n-s} \sum_{\ell=1-t}^{n-s-t-i} \phi_\ell$$

et vérifie donc

$$b_n \le \sum_{i=k}^{n-s} \sum_{\ell=1-t}^{n-t} \phi_\ell$$

d'où il vient

(*6) $b_n \le N\phi_n^*$

On a

$$c_n = \sum_{i=k}^{n-s} \sum_{\ell=s-t}^{n-i-1-t} \phi_\ell$$

d'où il vient

$$c_n \le \sum_{i=k}^{n-s} \sum_{\ell=1-t}^{n-t} \phi_\ell$$

et donc également

(*7) $c_n \le N\phi_n^*$.

Nous avons vu, lors de la fin de la démonstration du lemme 2, qu'on a $a_n \leq \phi_n^*$.
Puisque $\phi_n \geq 1$, on a aussi

(*8) $\quad a_n \leq \phi_n^{*2}$.

Par ailleurs

$$d_n = \sum_{i=k}^{n-s} \sum_{\ell=1-t}^{n-s-t-i} \phi_\ell \, \phi_{n-i-\ell-2t}$$

peut s'écrire

$$d_n = \sum_{i=k}^{n-s} \sum_{\ell=1-t}^{n-s-t} \phi_\ell \, \phi_{n-i-\ell-2t} \; 1_{\{\ell \leq n-s-t-i\}}$$

d'où il vient

$$d_n = \sum_{\ell=1-t}^{n-s-t} \phi_\ell \, \delta_\ell$$

où

$$\delta_\ell = \sum_{i=k}^{n-s-t-\ell} \phi_{n-i-\ell-2t}$$

$$= \sum_{u=s-t}^{n-k-\ell-2t} \phi_u$$

vérifie

$$\delta_\ell \leq \phi_n^* - 1, \quad \ell = 1-t,\dots,n-s-t ,$$

ce qui donne

$$d_n \leq \phi_n^{*2}$$

Ce dernier résultat et l'inégalité (*8) amènent

$$e_n \leq 2(\phi_n^* \,/\, HN)^2$$

qui avec (*5), (*6) et (*7) amène l'inégalité recherchée.

3.3. <u>démonstration de la proposition 1</u>.

Il s'agit de montrer que la v.a.r.

$$Q_n = R_n(\tilde{X}_n) - R(\tilde{X}_n)$$

vérifie

(3-3) $\quad EQ_n^2 \xrightarrow[n \to \infty]{} 0$.

On a, d'après (3-1),

(3-4) $\quad Q_n = A'_n + B'_n$

où, D_n étant défini dans (3-1),

$$A'_n = A_n \,/\, D_n \quad \text{avec} \quad A_n = \frac{1}{N} \sum_i \sum_j [g(X_{i+s}) - R(\tilde{X}_i)] I_j(\tilde{X}_i)\; I_j(\tilde{X}_n)$$

et

(3-5) $$B'_n = B_n \,/\, D_n \quad \text{avec} \quad B_n = \frac{1}{N} \sum_i \sum_j [R(\tilde{X}_i) - R(\tilde{X}_n)]\; I_j(\tilde{X}_i)\; I_j(\tilde{X}_n)\ .$$

La démonstration de (3-3) comporte deux parties:

partie A. - On montre $E\, A_n^2 \xrightarrow[n\to\infty]{} 0$ en donnant une majoration de $E\, A_n^2$ qui nous sera utile lors de la démonstration de la proposition 2.

partie B. - On montre que $E\, B_n^2 \xrightarrow[n\to\infty]{} 0$.

On posera

$$\Delta_n = \sum_j p_j\; I_j(\tilde{X}_n)$$

en remarquant tout de suite que le résultat (3-2) implique

(3-6) $\quad \Delta_n \geq md\, H$

A. L'identité

$$A'_n = [A_n - (A_n/D_n)(D_n - \Delta_n)] \,/\, \Delta_n$$

l'inégalité

$$|A_n / D_n| \leq 2G \quad ,$$

qui découle de (2-6), et le résultat (3-6) permettent d'écrire

$$|A'_n| \leq (|A_n| + 2G|D_n - \Delta_n|) \,/\, (mdH)$$

d'où il vient

(*1) $$EA'^2_n \leq 2(md)^{-2} [\, EH^{-2} A_n^2 + 4G^2\, EH^{-2}(D_n - \Delta_n)^2 \,] .$$

Si on applique le lemme 3 aux v.a.r.

$$\gamma_j(\tilde{X}_i) = [g(X_{i+s}) - R(\tilde{X}_i)]\, I_j(\tilde{X}_i) \quad [\text{resp.}\ \gamma_j(\tilde{X}_i) = I_j(\tilde{X}_i) - p_j]$$

qui vérifient

$$E\gamma_j(\tilde{X}_i) = 0$$

puisque

$$E\gamma_j(X_i) = E\left\{ E(g(\tilde{X}_{i+s}) \,/\, [X_{i-k+1},\dots,X_i]) - R(\tilde{X}_i)\right\} I_j(\tilde{X}_i)$$

et que le terme entre accolades dans cette formule est nul de par la définition (1-1) de R [resp. parce que $p_j = E\, I_j(\tilde{X}_i)$] et pour lesquelles on a

$$C = 2G \qquad [\text{resp.}\ C=1]$$

ainsi que

$$\Gamma = H^{-1} A_n \quad [\text{resp.}\ \Gamma = H^{-1}(D_n - \Delta_n)]$$

on obtient

$$EH^{-2} A_n^2 \leq 4MG^2(1+8\phi_n^*) / (HN) + C'[\phi_n^* / (HN)]^2 \;, \quad C' < \infty \;,$$

et

(*2) $$EH^{-2}(D_n - \Delta_n)^2 \leq M(1+8\phi_n^*) \,/\, (HN) + C''[\phi_n^* \,/\, (HN)]^2 \;, \quad C'' < \infty \;,$$

qui avec (*1) donnent

(3-7) $$EA_n'^2 \leq \frac{8MG^2}{(md)^2}\; \frac{1+8\phi_n^*}{HN} + C\left(\frac{\phi_n}{HN}\right)^2 \;, \quad C < \infty .$$

L'hypothèse (2-7) implique

$$\phi_n^* / (HN) \xrightarrow[n\to\infty]{} 0$$

d'où il vient par (*1)

(*3) $$E\, A_n'^2 \xrightarrow[n\to\infty]{} 0$$

et également par (*2)

(*4) $$E\, H^{-2}\, (D_n - \Delta_n)^2 \xrightarrow[n\to\infty]{} 0$$

B. Nous utilisons maintenant le fait que l'ensemble des fonctions continues (à support compact) est dense dans $L_2(R^q,P), q<\infty$ (Dunford et Schwartz 1957, p.298). Pour tout $\varepsilon > 0$, il existe donc une fonction réelle $\rho = \rho_\varepsilon$, continue sur $[0,1]^{kp}$ et telle que la fonction

$$\lambda(.) = \lambda_\varepsilon(.) = R(.) - \rho(.)$$

vérifie

$$E\ \lambda^2(\tilde{X}) \leq \varepsilon\ .$$

La v.a.r. B_n' peut s'écrire

(*5) $$B_n' = L_n + M_n / D_n$$

avec

$$L_n = \frac{1}{N} \sum_i \sum_j [\lambda(\tilde{X}_i) - \lambda(\tilde{X}_n)]\ I_j(\tilde{X}_i)\ I_j(\tilde{X}_n) / D_n$$

(*6) $$M_n = \frac{1}{N} \sum_i \sum_j [\rho(\tilde{X}_i) - \rho(\tilde{X}_n)]\ I_j(\tilde{X}_i)\ I_j(\tilde{X}_n)\ .$$

L'inégalité de Schwartz donne

$$L_n^2 \leq \frac{1}{N} \sum_i \sum_j [\lambda(\tilde{X}_i) - \lambda(\tilde{X}_n)]^2\ I_j(\tilde{X}_i)\ I_j(\tilde{X}_n) / D_n$$

d'où il vient

(*7) $$E\ L_n^2 \leq 2[E(U / D_n) + E(V / D_n)]$$

où U et V sont les v.a.r. définies dans le lemme 2 avec $\zeta = \lambda^2$. L'identité

$$T / D_n = [T - (T / D_n)\ (D_n - \Delta_n)] / \Delta_n$$

avec

(*8) $$T = U \text{ ou } V\ ,$$

le résultat (3-6) et l'inégalité de Schwartz permettent d'écrire

$$E(T/D_n) \leq (md)^{-1}\ (EH^{-1}\ T + G' \sqrt{EH^{-2}(D_n - \Delta_n)^2}\ )$$

où G'est une constante bornant λ^2. Le lemme 2 et le résultat (*4) amènent alors

$$E(T/D_n) \leq M(md)^{-1}\varepsilon + \alpha_n \ , \quad \text{avec} \quad \alpha_n \xrightarrow[n\to\infty]{} 0 \ ,$$

qui par (*7) et (*8) donne

(*9) $$E\,L_n^2 \leq 4M(md)^{-1}\ \varepsilon + \beta_n \ , \quad \text{avec} \quad \beta_n \xrightarrow[n\to\infty]{} 0 \ .$$

La continuité de la fonction ρ implique l'existence d'un réel positif $\alpha = \alpha_\varepsilon$ pour lequel on a

(*10) $$|\rho(\tilde{X}_i) - \rho(\tilde{X}_n)| \leq \varepsilon + 2G''\ 1_{\{\|\tilde{X}_i - \tilde{X}_n\| \geq \alpha\}}$$

où G" est une constante positive bornant ρ et où $\|z_1,\ldots,z_{p(t+1)}\| = \|z_1,\ldots,z_{kp}\|$ d'après nos notations 3-1.

Si nous remarquons que l'hypothèse (2-2) amène

(3-8) $$\sum_j I_j(\tilde{X}_i)\ I_j(\tilde{X}_n) \leq 1_{\{\|\tilde{X}_i - \tilde{X}_n\| \leq h_n\}} \ ,$$

nous voyons que l'hypothèse $h_n \xrightarrow[n\to\infty]{} 0$ implique

$$\exists\ n_0 \in N_* : \forall\ n > n_0 \ , \ 1_{\{\|\tilde{X}_i - \tilde{X}_n\| \geq \alpha\}} \sum_j I_j(\tilde{X}_i)\ I_j(\tilde{X}_n) = 0 \ ,$$

d'où il vient, en considérant l'inégalité (*10) et la formule (*6)

$$\exists\ n_0 \in N_* : \forall\ n > n_0 \ : \ |M_n / D_n| \leq \varepsilon \ .$$

Ce dernier résultat, l'inégalité (*9) et l'égalité (*5) qui implique

$$E\ B_n'^2 \leq 2(E\,L_n^2 + E(M_n / D_n)^2)$$

amènent

$$E\,B_n'^2 \leq 2(4M(md)^{-1}\varepsilon + \varepsilon^2) + \gamma_n \ , \quad \text{avec } \gamma_n \xrightarrow[n\to\infty]{} 0 \ ,$$

d'où il vient, puisque ε peut être choisi quelconque

$$E\,B_n'^2 \xrightarrow[n\to\infty]{} 0$$

ce qui avec (*3) achève la démonstration de la proposition 1.

3.4. démonstration des propositions 2 et 3.

Pour tout $i=1,\dots,n$ on a, d'après (3-8)

$$|R(\tilde{X}_i)-R(\tilde{X}_n)|\sum_j I_j(\tilde{X}_i)\; I_j(\tilde{X}_n) \le |R(X_i)-R(X_n)|\; 1_{\{\|\tilde{X}_i-\tilde{X}_n\|\le h_n\}} \sum_j I_j(\tilde{X}_i) I_j(\tilde{X}_n)$$

d'où il vient, en tenant compte de (2-9)

$$|R(\tilde{X}_i)-R(\tilde{X}_n)|\sum_j I_j(\tilde{X}_i)\; I_j(\tilde{X}_n) \le c\, h_n^{\gamma} \sum_j I_j(\tilde{X}_i)\; I_j(\tilde{X}_n)$$

Il suffit alors de considérer la définition (3-5) de B'_n pour obtenir

$$|B'_n| \le c\; h_n^{\gamma}$$

Ce resultat, l'inégalité (3-7) et la formule (3-4) qui donne

$$E\, Q_n^2 \le 2(E\, A_n'^2 + E\, B_n'^2)$$

amènent immédiatement l'inégalité de la proposition 2.

Pour prouver la proposition 3, on utilise d'abord le lemme de Parzen (1962, p.1074) pour voir que la fonction

$$h \longrightarrow \psi(h) = M\left(\frac{4G}{md}\right)^2 \frac{1+8\phi_n^*}{Nh^{kp}} + 2c^2\; h^{2\gamma}$$

est minimisée par $h^*=B\left(\frac{\phi_n^*}{N}\right)^{\gamma/(kp+2\gamma)}$ pour lequel on a $\psi(h^*) = C\left(\frac{\phi_n^*}{N}\right)^{2\gamma/(kp+2\gamma)}$ et ensuite (2-12) qui donne (2-13) pour $A > C$, où A, B, $C \in]0,\infty[$ ne dépendent pas de n.

3.5. démonstration de la proposition 4.

Pour prouver que (2-15) est une conséquence de (2-14), on choisit dans $\mathcal{D}_0$ un processus constitué de v.a. X_n, $n\in\mathbb{N}$, satisfaisant

(*1) $\quad r = E\, g(X_1) \neq 0$ et $v = E(g(X_1)-r)^2 > 0$

et indépendantes, ce qui permet d'écrire pour $Q_n = R_n(\tilde{X}_n) - R(\tilde{X}_n)$,

(*2) $\quad Q_n = R_n(\tilde{X}_n) - r$,

la condition (2-14) amenant

(*3) $\quad EQ_n^2 \xrightarrow[n\to\infty]{} 0$.

On remarque d'abord que la définition (2-1) de $R_n(X_n)$ peut s'écrire, puisque $|J| = ND_n$,

$$R_n(\tilde{X}_n) = \begin{cases} D_n^{-1} \sum_i \sum_j g(X_{i+s}) I_j(\tilde{X}_i) I_j(\tilde{X}_n) \text{ si } ND_n \geq 1 , \\ 0 \text{ si } ND_n = 0 , \end{cases}$$

qui donne par (*2)

$$(*4) \quad Q_n = -r \, \mathbb{1}_{\{ND_n=0\}} + [A_n/D_n], \text{ avec } [A_n/D_n] = \begin{cases} A_n/D_n \text{ si } D_n \neq 0, \\ 0 \text{ si } D_n = 0 , \end{cases}$$

où on a posé

$$(*5) \quad A_n = N^{-1} \sum_i \sum_j (g(X_{i+s})-r) \, I_j(\tilde{X}_i) I_j(\tilde{X}_n) ,$$

et d'où il vient

$$Q_n^2 = ([A_n / D_n])^2 + r^2 \, \mathbb{1}_{\{nD_n=0\}}$$

et donc

$$(*6) \quad Q_n^2 \geq ([A_n/D_n])^2$$

ainsi que

$$E \, Q_n^2 \geq r^2 \, E \, \mathbb{1}_{\{ND_n=0\}} .$$

Cette dernière inégalité amène

$$r^2 - EQ_n^2 \leq r^2 \, E \, \mathbb{1}_{\{ND_n \geq 1\}}$$

qui donne par (*1) et puisque $\mathbb{1}_{\{ND_n \geq 1\}} \leq ND_n$,

$$1 - EQ_n^2 / r^2 \leq E \, 1_{\{ND_n \geq 1\}} \leq E \, ND_n ,$$

d'où il vient par (*3)

$$(*7) \quad \exists \, \alpha_0 > 0 , \exists \, n_0 \in \mathbb{N} \; : \; \forall \, n > n_0 , \; NH \leq \alpha_0 ,$$

après avoir utilisé, en tenant compte du fait que $(h_n)_{\mathbb{N}}$ est de limite nulle, l'inégalité

(*8) $\quad E N D_n \leq a_0 \, NH + a_1 \, h^{ps}$, $\forall \, n \in \mathbb{N}$, $0 < a_0 < \infty$, $0 < a_1 < \infty$, qui sera démontrée dans le point a ci-dessous.

La définition de $[A_n \,/\, D_n]$ dans (*4) amène, puisque $A_n = 0$ dès que $D_n = 0$,

$$A_n = [A_n \,/\, D_n] \, D_n$$

d'où il vient par (*6)

$$A_n^2 \leq Q_n^2 \; D_n^2$$

et donc en utilisant l'inégalité de Schwartz

$$(E A_n^2)^2 \leq E Q_n^4 \; E D_n^4$$

qui donne, en tenant compte de l'hypothèse (2-6),

$$(E A_n^2)^2 \leq 4 G^2 E Q_n^2 \; E D_n^4 .$$

Cette dernière majoration et le résultat

(*9) $\quad \exists \, b_0, \; 0 < b_0 < \infty$, $\exists \, n_1 \in \mathbb{N} : \forall \, n > n_1$, $E D_n^4 \leq b_0 \, H^4 (1 + \sum_{u=2}^{4} (NH)^{-u})$,

qui sera démontré dans le point b ci-dessous, amenent alors, en tenant compte de (*7),

$$E(H^{-2} \, A_n^2) \leq \beta_0 (EQ_n^2)^{1/2} \; , \; \forall \, n \in \mathbb{N} , \; n > \max(n_0, n_1) \; , \; 0 < \beta_0 < \infty.$$

Cette dernière inégalité et la minoration

(*10) $\quad E A_n^2 \geq c_0 \, v \, H/N - c_1 \, h^{ps} / N^2$, $\forall \, n \in \mathbb{N}$, $0 < c_0 < \infty$, $0 < c_1 < \infty$,

qui sera démontrée dans le point c ci-dessous donnent, en utilisant la seconde partie de (*1) et le résultat (*7),

$$(NH)^{-1} \leq \gamma_0 (E Q_n^2)^{1/2} + \gamma_1 \, h^{ps} \; , \quad 0 < \gamma_0 < \infty \; , \quad 0 < \gamma_1 < \infty,$$

pour n suffisamment grand, ce qui, puisque $(h_n)_{\mathbb{N}}$ est de limite nulle et que les comportements asymptotiques de NH et nh_n^{kp} sont identiques, prouve que (*3) implique (2-15).

Point a : démonstration de (*8).

On a

$$ND_n = U_n + V_n$$

avec

$$U_n = \sum_{i=k}^{n-k} \sum_j I_j(\tilde{X}_i)\, I_j(\tilde{X}_n) \quad \text{et} \quad V_n = \sum_{i=n-k+1}^{n-s} \sum_j I_j(\tilde{X}_i)\, I_j(\tilde{X}_n) \, .$$

Le lemme ci-dessous donne immédiatement

$$EU_n \le (n-2k+1)a_o\, H \le a_o\, nH$$

et

$$EV_n \le \mathbb{1}_{\{k-s>o\}}\, a_o \sum_{i=n-k+1}^{n-s} h^{p(n-i)} = \mathbb{1}_{\{k-s>o\}} a_o\, h^{ps} \sum_{i=n-k+1}^{n-s} h^{p(n-i-s)}$$

ce qui prouve (*8) puisque $(h_n)_{\mathbb{N}}$ est de limite nulle (dans le cas $s \ge k$ l'inégalité de (*8) est précisée par $a_1 = 0$).

Lemme 4.

Il existe une constante a_o, $0 < a_o < \infty$, telle que pour tous entiers n, i et $i' \ge i$ on a

$$E \sum_j I_j(\tilde{X}_i)\, I_j(\tilde{X}_{i'}) \le \left| \begin{array}{l} a_o\, H \text{ si } i'-i \ge k \, , \\ a_o\, h^{p(i'-i)} \text{ si } i'-i \le k . \end{array} \right.$$

Démonstration.

L'hypothèse (2.2) montre qu'il est possible d'associer à tout élément j de $\mathcal{P}_n$ un cube C_j de $\mathbb{R}^p$ ayant h pour côté et satisfaisant pour tout $[u_1,\dots,u_k]$ de $\mathbb{R}^{kp}$

$$I_j(u_1,\dots,u_k) \le \prod_{m=1}^{k} \mathbb{1}_{\{u_m \in C_j\}} \, , \; u_m \in \mathbb{R}^p \, , \quad m=1,\dots,k \, ,$$

ce qui amène, pour tous entiers i et $i' \ge i$,

$$I_j(\tilde{X}_i)\, I_j(\tilde{X}_{i'}) \le F = \prod_{a=i-k+1}^{i} \mathbb{1}_{\{X_a \in C_j\}} \prod_{b=i'-k+1}^{i'} \mathbb{1}_{\{X_b \in C_j\}}$$

où F peut s'écrire pour $b_o = \max(i'-k+1,\, i+1)$

$$F = F_1 F_2 \quad \text{avec} \quad F_1 = \prod_{a=i-k+1}^{i} \mathbb{1}_{\{X_a \in C_j\}} \quad \text{et} \quad F_2 = \prod_{b=b_0}^{i'} \mathbb{1}_{\{X_b \in C_j\}} .$$

L'indépendance des v.a. X_i , $i \in \mathbb{N}$, implique (outre l'indépendance de F_1 et F_2) l'indépendance des facteurs constituant F_1 et F_2, ce qui donne

$$EF = (E\, \mathbb{1}_{\{X_1 \in C_j\}})^c \quad \text{avec} \quad c = k+i' - b_0 ,$$

d'où il vient, de par la dernière inégalité obtenue et en tenant compte de la partie droite de (2.5), pour tout j de $\mathcal{P}_n$

$$E\, I_j(\tilde{X}_i)\, I_j(\tilde{X}_{i'}) \le (\tilde{M}\, h^p)^c \ , \ 0 < \tilde{M} < \infty, \ (\tilde{M} = M^{1/k}).$$

Cette majoration et un calcul simple montrant que

$$c = 2k \ \text{ si } \ i'-i \ge k, \ c = k+(i'-i) \ \text{ si } \ i'-i \le k, \ h^{pc} = H\, h^{p(c-k)} ,$$

amènent immédiatement le résultat énoncé dans le lemme ci-dessus, en tenant compte de l'inégalité $\sum_j H \le (md)^{-1} < \infty$ découlant de la relation $\sum_j p_j = 1$ et de la partie gauche de la double inégalité (3-2).

Point b : demonstration de (*9)

En écrivant

$$D_n = B_n + \tilde{B}_n + \Delta_n$$

où

$$B_n = \sum_j N^{-1} \sum_{i=k}^{n-k} (I_j(\tilde{X}_i) - p_j)\, I_j(\tilde{X}_n),\ \tilde{B}_n = \sum_j N^{-1} \sum_{i=n-k+1}^{n-s} (I_j(\tilde{X}_i) - p_j)\, I_j(\tilde{X}_n) ,$$

$$\Delta_n = \sum_j p_j\, I_j(\tilde{X}_n) ,$$

et en remarquant que les v.a.r. $\tilde{B}_n$ et Δ_n vérifient

$$|\tilde{B}_n| \le \sum_j N^{-1}\ 2 \max(k-s,0)\ I_j(\tilde{X}_n) \le b_1 N^{-1} \ , \ 0 < b_1 < \infty ,$$

et

$$0 < \Delta_n \le MH$$

en raison de (3-2), on obtient la majoration

$$E D_n^4 \le b_2(H^4(1+(NH)^{-4}) + E B_n^4), \quad 0 < b_2 < \infty .$$

On a

$$B_n^4 = \sum_j V_j^4 \; I_j(\tilde{X}_n) \quad \text{où} \quad V_j = N^{-1} \sum_{i=k}^{n-k} (I_j(\tilde{X}_i) - p_j)$$

est pour tout j de $\mathcal{P}_n$ une v.a;r. indépendante de $\tilde{X}_n$, ce qui donne

$$E B_n^4 = \sum_j E V_j^4 p_j .$$

En écrivant, pour tout j de $\mathcal{P}_n$, $N V_j = \sum_{u=1}^{k} V_{j,u}$ avec

$$V_{j,u} = \sum_{v=1}^{N_u} (I_j(\tilde{X}_{u+kv}) - p_j) \text{ où } N_u = \mathrm{Pe}[(n-k)/k]$$

vérifie $N_u \le N/k$ pour tout u=1,...,k , on obtient

$$E V_j^4 \le b_3 \sum_{u=1}^{k} (N_u^{-1} V_{j,u})^4 , \quad 0 < b_3 < \infty .$$

En remarquant alors que $V_{j,u}$ est une somme de N_u v.a.r. indépendantes centrées et en appliquant un résultat général relatif au moment d'ordre 4 d'une somme de telles v.a.r., on obtient pour u = 1,...,k,

$$E(N_u^{-1} V_{j,u})^4 = N_u^{-4}(N_u m_{j,4} + 3N_u(N_u - 1)(m_{j,2})^2)$$

où $m_{j,4} = E(I_j(\tilde{X}_1)-p_j)^4 = p_j^4 - 3p_j^4 + 6p_j^3 - 3p_j^2 + p_j$ et $m_{j,2} = E(I_j(\tilde{X}_1)-p_j)^2 = p_j(1-p_j)$

vérifient, en raison de (3-2) et puisque $(h_n)_{\mathbb{N}}$ est de limite nulle, les inégalités $m_{j,4} \le b_4 H$ et $m_{j,2} \le b_5 H$ pour n suffisamment grand, b_4 et b_5 étant des constantes positives. On a donc le résultat

$$\exists n_1 \in \mathbb{N} , \exists b_6, 0 < b_6 < \infty : E(N_u^{-1} V_{j,u})^4 \le b_6 H^4(1+(NH)^{-2} + (NH)^{-3}), \forall n \in \mathbb{N} , n > n_1 ,$$

d'où il vient, en considérant les dernières formules dont les membres de gauche sont respectivement EV_j^4 et EB_n^4 ,

$$E B_n^4 \le b_7 H^4(1+(NH)^{-2} + (NH)^{-3}), \; 0 < b_7 < \infty, \; \forall n > n_1 .$$

La majoration de ED_n^4 donnée plus haut amene alors immédiatement le résultat (*9).

Point c : démonstration de (*10).

On remarque d'abord que (*5) peut s'ecrire

$$A_n = B_n + C_n$$

avec

$$B_n = \sum_j N^{-1} \sum_{i=k}^{n-k-s} A_{i,j}\ I_j(\tilde{X}_n) \quad \text{et} \quad C_n = \sum_j N^{-1} \sum_{i=n-k-s+1}^{n-s} A_{i,j}\ I_j(\tilde{X}_n)$$

où

$$A_{i,j} = (g(X_{i+s})-r)\ I_j(\tilde{X}_i)\ .$$

L'expression $B_n = A_n - C_n$ donne

$$E\,B_n^2 \leq 2(E\,A_n^2 + E\,C_n^2)$$

d'où il vient

$$E\,A_n^2 \geq 2^{-1}\,E\,B_n^2 - E\,C_n^2\ .$$

On a

$$B_n^2 = \sum_j (N^{-1} \sum_{i=k}^{n-k-s} A_{i,j})^2\ I_j(\tilde{X}_n)$$

d'où il vient, de par l'indépendance, pour tout j de $\mathcal{P}_n$, de $I_j(\tilde{X}_n)$ et $A_{i,j}$ pour tout $i \leq n-k-s$,

$$E\,B^2 = \sum_j p_j\ E(n^{-1} \sum_{i=k}^{n-k-s} A_{i,j})^2\ .$$

En écrivant

$$E(N^{-1} \sum_{i=k}^{n-k-s} A_{i,j})^2 = N^{-2} \sum_{a=k}^{n-k-s} \sum_{b=k}^{n-k-s} E\,F_{a,b}$$

où

$$F_{a,b} = (g(X_{a+s})-r)(g(X_{b+s}) - r)\ I_j(\tilde{X}_a)\ I_j(\tilde{X}_b)$$

vérifie d'abord

$$E\,F_{a,b} = 0 \quad \text{si} \quad a \neq b$$

en raison de l'indépendance de la v.a.r. centree $g(X_{\max(a,b)+s})-r$ et de la v.a.r. égale au produit des autres facteurs constituant F , et ensuite

$$E\,F_{a,a} = v\,p_j$$

en raison de l'indépendance de $g(X_{a+s})-r$ et $I_j(\tilde{X}_a)$, on obtient

$$E\,B_n^2 = N^{-2}(n-2k-s+1)\ v \sum_j p_j^2$$

d'où il vient par (2-10), (2-3), (2-5) et (3-2) où $\sum_j p_j = 1$,

$$E\,B_n^2 \le c_2\ v\ H\ /\ N\ ,\quad c_2 = md > 0.$$

On a

$$C_n^2 = \sum_j N^{-2}\ T_j^2$$

où $T_j = \sum_{i=n-k-s+1}^{n-s} A_{i,j}\ I_j(\tilde{X}_n)$ verifie

$$T_j^2 \le k \sum_{i=n-k-s+1}^{n-s} A_{i,j}^2\ I_j(\tilde{X}_n)\ .$$

Par passage à l'espérance et en tenant compte de (2-6) on obtient

$$E\,T_j^2 \le 4G^2\,k \sum_{i=n-k-s+1}^{n-s} E\ I_j(\tilde{X}_i)\ I_j(\tilde{X}_n)$$

d'où il vient, en utilisant le lemme 4,

$$E\,C_n^2 \le c_3\ N^{-2} \sum_{i=n-k-s+1}^{n-s} h^{p\psi(i)} \quad \text{où} \quad \psi(i) = \max(k,n-i)\ ,\quad 0 < c_3 < \infty\ ,$$

qui donne, puisque $(h_n)_{\mathbb{N}}$ est de limite nulle,

$$E\,C_n^2 \le c_4\ N^{-2}\ h^{ps}\ ,\qquad 0 < c_4 < \infty\ .$$

Cette majoration et les deux dernières inégalités concernant respectivement $E\,A_n^2$ et $E\,B_n^2$ amènent immédiatement le résultat (*10).

4 - APPLICATIONS : ESTIMATION DE LA PROBABILITE DE TRANSITION.

On pose, pour tout u de $[0,1]^p$

$$F_n(u;X_{n-k+1},\dots,X_n) = \frac{1}{|J|}\sum_{i=k}^{n-s} 1_{\{X_{i+s}\le u\}}\, 1_{\{(X_{i-k+1},\dots,X_i)\in J\}}$$

si $|J| \neq 0$, 0 sinon, avec J et $|J|$ definis en (2-1),

et

$$F(u;X_{n-k+1},\dots,X_n) = P(X_{n+s}\le u \,/\, (X_{n-k+1},\dots,X_n))$$

qui est *la fonction de répartition de la probabilité de transition* du processus $(X_n)_{\mathbb{N}}$ lorsque ce dernier est *markovien d'ordre k*.

Les resultats concernant le predictogramme se transposent aisément à cet estimateur non paramétrique de la probabilite de transition, ce qui permet d'énoncer les propositions suivantes.

Proposition 1'.

Sous les hypothèses de la proposition 1 on a

$$\forall u\in[0,1]^p \quad , \quad F_n(u;X_{n-k+1},\dots,X_n) - F(u;X_{n-k+1},\dots,X_n) \xrightarrow[n\to\infty]{L_2} 0$$

Proposition 2' et 3'.

Enoncé identique à ceux des propositions 2 et 3, avec G=1, c *défini par* (2-9) *où* $R(.)=F(u;.)$ *et* $Q_n = F_n(u;X_{n-k+1},\dots,X_n) - F(u;X_{n-k+1},\dots,X_n)$.

Proposition 4'.

Pour que

$$\forall\ \xi\in\mathcal{D}_0\ ,\ \forall\ u\in[0,1]^p \quad F_n(u;X_{n-k+1},\dots,X_n) - F(u;X_{n-k+1},\dots,X_n) \xrightarrow[n\to\infty]{L_2} 0$$

il faut et il suffit que $(h_n)_{\mathbb{N}}$ *vérifie*

$$n\,h_n^{kp} \xrightarrow[n\to\infty]{} \infty .$$

démonstration des propositions 1', 2', 3' et 4'. On applique les propositions 1, 2, 3 et 4, pour tout u de $[0,1]^p$, à la fonction

$$g(z) = 1_{\{z \le u\}} \quad , \quad \forall\, z \in [0,1]^p \ .$$

Autres applications.

L'estimateur de la probabilité de transition que nous venons de considérer permet de définir

- un estimateur de la densité

$$f(.;X_{n-k+1},\dots,X_n) = dF(.;X_{n-k+1},\dots,X_n)/d\mu$$

de cette probabilité de transition (ici μ est la mesure de Lebesgue sur $[0,1]^p$) : il suffit d'associer à $F_n(.;X_{n-k+1},\dots,X_n)$ un histogramme.

- pour $p=1$ et tout α de $[0,1[$, un estimateur du *α-quantile*

$$A^{(\alpha)}(X_{n-k+1},\dots,X_n) = \sup\{\, a : F(a\,;X_{n-k+1},\dots,X_n) \le \alpha \,\}$$

de la probabilité de transition : il suffit d'estimer $A^{(\alpha)}$ à l'aide de $A_n^{(\alpha)}$ défini par la partie droite de la formule ci-dessus dans laquelle F est remplacée par F_n. On remarquera en particulier que $A_n^{0.5}(X_{n-k+1},\dots,X_n)$ est un estimateur de la *médiane conditionnelle* $A^{0.5}(X_{n-k+1},\dots,X_n)$ qui est le meilleur prédicteur probabiliste pour la fonction de perte $\lambda(.) = |\cdot|$.

Par ailleurs, il est évident que

- dans le cas $p=1$, on obtient immédiatement la définition et les propriétés de convergence dans L_2 d'estimateurs de *moments* de tout ordre $\beta \in \mathbb{N}$

$$R^{(\beta)}(X_{n-k+1},\dots,X_n) = E(X_{n+s}^{\beta} \,/\, X_{n-k+1},\dots,X_n)$$

de la probabilité de transition : il suffit de choisir $g(.) = (\cdot)^{\beta}$.

- un estimateur de la *densité stationnaire*

$$\pi = dQ\,/\,d\mu \quad , \quad \text{où } Q \text{ est la loi de } (X_1,\dots,X_k),$$

est défini par l'*histogramme*

$$\pi_n(z) = \frac{1}{nh_n^{kp}} \sum_{i=1}^{n} 1_{\{(X_{i-k+1},\dots,X_i)\in J\}} , \quad J \in \mathcal{S}_n : z \in J \quad , \forall z \in [0,1]^{kp}$$

qui est lui-même une fonction de densité (puisque $\int_C \pi_n(z)dz = 1$) et pour lequel les méthodes de démonstration utilisées dans le paragraphe 3 permettent de prouver qu'il jouit de propriétés analogues à celles de l'histogramme.

Enfin, nous remarquerons, à propos de ces divers e.n.p. définis à partir du prédictogramme, que les techniques de démonstrations employées ici permettent d'obtenir *a fortiori* des résultats sur l'estimation, pour tous z fixés (certains) dans $[0,1]^{kp}$, de $R(z)$, $f(.;z)$, $\pi(z)$ [resp. et $F(.;z)$, $A^{(\alpha)}(z)$, $\alpha \in [0,1[$, $R^{(\beta)}(z)$, $\beta \in \mathbb{N}$] analogues à ceux obtenus par Collomb et Doukhan (1983) [resp. Roussas, 1969] à propos de la méthode du noyau, et ce avec des hypothèses moins restrictives sur les fonctions à estimer : la continuité n'est plus utile et la dérivabilité est remplacée par une condition de Lipschitz.

5. REMARQUES SUR LA DEFINITION ET L'UTILISATION DU PREDICTOGRAMME.

La définition et l'utilisation du prédictogramme sont analogues à celles du régressogramme ou de l'histogramme, avec les mêmes

- simplicité d'utilisation, notamment par rapport à certaines *méthodes paramétriques* classiques (du type Box et Jenkins) plus souvent employées dans des problèmes de prédiction et, également, par rapport aux méthodes de prédiction qui utilisent des *e.n.p.* de la *densité spectrale* (voir par exemple Parzen, 1957). Toutefois, dans le cas $kp > 1$, le prédictogramme présente des *difficultés de visualisation*, identiques à celles du régressogramme ou de l'histogramme.
- problème du choix de la partition $\mathcal{S}_n$: les conditions (2-2) et (2-3) laissent une grande liberté dans le choix de la partition $\mathcal{S}_n$ et il est alors possible de tenir compte dans le choix de $\mathcal{S}_n$ d'une certaine connaissance *a priori* du phénomène observé. Lorsque $\mathcal{S}_n$ est constitué de pavés égaux de côté h_n, le problème du choix de h_n peut être abordé à l'aide d'une *méthode heuristique* exposée par Collomb (1983).

• faible coût informatique (inférieur notamment à celui provenant de l'emploi des méthodes du noyau ou des k points les plus proches), en particulier lorsque la partition $\mathcal{P}_n$ est constituée de pavés égaux, grâce à l'utilisation de la fonction partie entière (voir Collomb, 1978, p.14).

• manque de précision, inhérent aux méthodes non paramétriques : la proposition 3 montre que la *vitesse de convergence* du prédictogramme, identique à celle de l'histogramme ou du régressogramme est tres inférieure à la vitesse de convergence des méthodes paramétriques classiques et un simple calcul effectué à partir de la formule (2-13) montre qu'une bonne précision ne peut pas être assurée lorsque n n'est pas "très" grand : les études par simulation de Kalaidjian (1981) sont à cet égard tres significatives.

• usage statistique specifique, lorsque le statisticien est en présence d'observations et ne dispose d'aucune connaissance *a priori* sur le phénomène probabiliste dont elles sont issues. De même que l'histogramme est alors utilisé pour "reconnaître" une distribution de probabilité, le prédictogramme permet de déceler une tendance dans la série temporelle $(X_n)_{\mathbb{N}}$, sans autre hypothèse que celle relative à la valeur de k. Cette méthode de prédiction peut aider le statisticien dans le choix d'un modèle paramétrique pour le processus observé.

Références. (Les références precédées du signe "*" concernent plus particulièrement les problèmes d'analyse d'une série temporelle et de prédiction non paramétriques.)

* Banon, G. (1978). Nonparametric identification for diffusion processes. *Siam Journal on Control and Optimization*, 16, 3, 380-395.

Billingsley, P. (1968). Convergence of Probability Measures. *Wiley*, New-York.

* Bosq, D. (1979). Sur la prediction non parametrique des variables aleatoires et de mesures aléatoires. *Zeitschrift für W.u.v.G.*, *64*, 541-553.

* Bosq, D. (1983). Non Parametric Prediction in Stationary Processes. *Lecture Notes in Statistics*, 16, 69-84.

Bradley, R.C. (1980). On the ϕ-mixing condition for stationary random sequences. *Duke Mathematical Journal*, 47, 2, 421-433.

Bretagnolle, J. et Huber, C. (1979). Estimation des densités : risques minimaux. *Zeitschrift für Wahrscheinlichkeitstheorie verw. Gebiete*, 47, 119-137

Collomb, G. (1977). Quelques propriétés de la methode du noyau pour l'estimation non paramétrique de la régression en un point fixe. *Comptes Rendus Acad. Sci. Paris*, 285, série A, 289-292.

Collomb, G. (1978). Estimation non paramétrique de la régression : régressogramme et methode du noyau. *Publications* du Laboratoire de Statistique et Probabilites de l'Université de Toulouse, n° 07-78, 1-59.

Collomb, G. (1981). Estimation non paramétrique de la régression : revue bibliographique. *International Statistical Review*, 49, 75-93.

* Collomb, G. (1982a). Propriétés de convergence presque complète du prédicteur à noyau. *Zeitschrift für W.u.v.G.*, 66, 441-460.

Collomb, G. (1982b). Analyse d'une série temporelle et prédiction non paramétrique : méthodes de la fenêtre mobile et des k points les plus proches en estimation de régression pour des observations dépendantes. *Prépublication.*

* Collomb, G. (1983). From non parametric regression to non parametric prediction : survey of the mean square error and original results on the predictogram. *Lecture Notes in Statistics*, 16, 182-204.

* Collomb, G. et Doukhan, P. (1983). Estimation non paramétrique de la fonction d'autorégression d'un processus stationnaire et mélangeant : risques quadratiques pour la méthode du noyau. *Comptes Rendus Acad. Sci. Paris*, à paraître.

Devroye, L. et Wagner, T.S. (1980). Distribution Free Consistency Results in Nonparametric Discrimination and Regression Function Estimation. *Annals of Statistics*, 8, 231-239.

Devroye, L. (1982), Necessary and sufficient conditions for the pointwise convergence of nearest neighbor regression function estimates. *Z. Wahr. verw. Geb.*, 61, 467-481.

* Doob, J. (1953). Syochastic Processes, *Wiley*, New-York.

* Doukhan, P. et Ghindès, M. (1980). Estimation dans le processus "$X_{n+1}=f(X_n)+\varepsilon_n$". *Comptes-Rendus. Acad. Sci. Paris*, 297, serie A, 61-64.

* Doukhan, P. (1983). Simulations in the General First Order Autoregressive Process (Unidimensional Normal Case). *Lecture Notes in Statistics*, 16, 50-68.

Dunford, N., Schwartz, J.T. (1957). Linear Operators. Part I. General Theory. *New-York, Interscience.*

Kalaidjian, R. (1981). Efficacités comparées de certaines méthodes de prédiction pour un arma perturbé. *Statistique et Analyse des Données*, 6,3,47-64.

* Nguyen, H.T. et Pham, D.T. (1981). Nonparametric estimation in diffusion model by discrete sampling. *Publications de L'Institut de Statistique de l'Université de Paris*, XX VI, 2, 89-109.

Parzen, E. (1957). On consistent estimates of the spectrum of a stationary time series. *Ann. Math. Statist.*, 28, 329-348.

Parzen, E. (1962). On estimation of a probability density and mode. *Ann. Math. Statist.* 35, 1065-1075.

* Pham, D.T. (1981). Nonparametric Estimation of the Drift Coefficient in the Diffusion Equation. *Math. Operationsforsch. Statist.*, 12, 1, 61-74.

* Robinson, P.M. (1982). Non parametric estimators for time series. *Journal of Time Series Analysis*, 4,3, 185-207.

* Rosenblatt, M. (1971). Markov Processes. Structure and Asymptotic Behavior, Springer, Berlin

* Roussas. G. (1969) Nonparametric Estimation of the Transition Distribution Function of a Markov Process. *Annals of Mathematical Statistics*,40, 1386-1400.

Spiegelman, G. et Sacks, J. (1980). Consistent Window Estimation in Nonparametric Regression Estimation, *Annals of Statistics*,8, 240-246.

Stone, C.J. (1977). Consistent nonparametric regression. *Annals of Statistics*, 5, 595-620.

Stone, C.J. (1982). Optimal global rates of convergence for nonparametric regresregression. *Annals of Statistics*, 10, 1040-1053.

* Watson, G.S. (1964). Smooth regression analysis. *Sankhya*, 26. A, 359-372.