STATISTICA, anno LXX, n. 1, 2010

NOTE ON CONDITIONAL MODE ESTIMATION FOR FUNCTIONAL DEPENDENT DATA

S. Dabo-Niang, A. Laksaci

1. INTRODUCTION

Let us introduce n pairs of random variables $(X_i, Y)_{i=1,\dots,n}$ that we suppose drawn from the pair (X, Y) valued in $F \times R$ where F is a semi-metric space. Let d denotes the semi-metric. Assume that there exists a regular version of the conditional probability distribution of Y given X which is absolutely continuous with respect to the Lebesgue measure on R and has bounded density. Assume that for a given x the conditional density f^x of Y given $X = x$ is unimodal and the conditional mode, denote by $\theta(x)$ is defined by

$$\theta(x) = argmax_{y \in R} f^x(y)$$

In the remainder of the paper, x is fixed in F and N_x denotes a neighborhood of x. We define the kernel estimator $\hat{f}^x$ of f^x as follows:

$$\hat{f}^x(y) = \frac{h_H^{-1} \sum_{i=1}^n K(h_K^{-1}(d(x, X_i)) H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^n K(h_K^{-1}(d(x, X_i))}, \qquad \forall\, y \in R$$

here the numerator is equal to zero when the denominator approaches zero. K and H are kernels functions and $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers. Note that a similar estimate was already introduced in the special case where X is a real random variable by many authors, Rosenblatt (1969) and Youndjé (1993) among others. For the functional case, see Ferraty *et al.* (2006). A natural and usual estimator of $\theta(x)$, denoted $\hat{\theta}(x)$ is given by:

$$\hat{\theta}(x) = argmax_{y \in R} \hat{f}^x(y) \qquad (1)$$

Note that this estimate $\hat{\theta}(x)$ is not necessarily unique, so the remainder of the paper concerns any value $\hat{\theta}(x)$ satisfying (1).

The main goal of this paper is to study the nonparametric estimate $\hat{\theta}(x)$ of $\theta(x)$, when the explanatory variable X is valued in the space F of eventually infinite dimension and when the observations $(X_i, Y_i)_{i=1,\ldots,n}$ are strongly mixing.

The interest of the nonparametric estimation of the conditional mode for strong mixing processes comes mainly from the fact that the mode regression provides better estimations than the classical mean regression in some situations (see for instance Collomb *et al.* (1987), Quintela and Vieu (1997), Berlinet *et al.* (1998) or Louani and Ould-Saïd (1999) for the multivariate case).

Currently, the progress of informatics tools permits the recovery of increasingly bulky data. These large data sets are available essentially by real time monitoring, and computers can manage such databases. The object of statistical study can then be curves (consecutive discrete recordings are aggregated and viewed as sampled values of a random curve) not numbers or vectors. Functional data analysis (FDA) (see Bosq (2000), Ferraty and Vieu, (2006), Ramsay and Silverman, (2002)) can help to analyze such high-dimensional data sets. The statistical problems involved in the modelization of functional random variables have received increasing interests in the recent literature (see for example Dabo-Niang (2002), Dabo-Niang and Rhomari (2003, 2009), Masry (2005) for the nonparametric context). In this functional area, the first results concerning the conditional mode estimation were obtained by Ferraty *et al.* (2006). They established the almost complete convergence of the kernel estimator in the i.i.d. case. This last result has been extended to dependent case by Ferraty *et al.* (2005). Ezzahrioui and Ould-said (2006, 2008) have studied the asymptotic normality of the kernel estimator of the conditional mode for both i.i.d. and strong mixing cases. The monograph of Ferraty and Vieu (2006) presents an important collection of statistical tools for nonparametric prediction of functional variables. Recently, Dabo-Niang and Laksaci (2007) stated the convergence in L^p norm of the conditional mode function in the independent case.

In this paper, we consider the case where the data are both dependent and of functional nature. We prove the *p*-integrated consistency by giving the upper bounds for the estimation error. Theses results can be applied to predict time series, by cutting the past of the series in continuous paths.

The paper is organized as follows: the following Section is devoted to fixe the notations and hypotheses. We state our results on Section 3. All proofs are given in the Appendix.

2. NOTATION AND ASSUMPTIONS

We begin by recalling the definition of the strong mixing property. For this we introduce the following notations. Let $F_i^k(Z)$ denote the σ-algebra generated by $\{Z_j,\ i \le j \le k\}$.

Definition: *Let* $\{Z_i,\ i = 1, 2, ...\}$ *denote a sequence of random variables. Given a positive integer* n *, set*

$$\alpha(n) = sup\{\left|P(A \cap B) - P(A)P(B)\right|: \quad A \in F_1^k(Z) \text{ and } \quad B \in F_{k+n}^{\infty}(Z), \quad k \in N^*\}$$

The sequence is said to be α*-mixing (strong mixing) if the mixing coefficient* $\alpha(n) \to 0$ *as* $n \to \infty$.

There exist many processes fulfilling the strong mixing property. We quote here, the usual ARMA processes which are geometrically strongly mixing, *i.e.,* there exist $\rho \in (0,1)$ and $a > 0$ such that, for any $n \ge 1$, $\alpha(n) \le a\rho^n$ (see, *e.g.*, Jones (1978)). The threshold models, the EXPAR models (see, Ozaki (1979)), the simple ARCH models (see Engle (1982)), their GARCH extension (see Bollerslev (1986)) and the bilinear Markovian models are geometrically strongly mixing under some general ergodicity conditions.

Throughout the paper, when no confusion is possible, we will denote by C or C' some strictly positive generic constants, $g^{(j)}$ the j^{th} order derivative of the function g, and we suppose that the response variable Y is p-integrable. Our nonparametric model will be quite general in the sense that we will just need the following assumptions:

- (H1) $P(X \in B(x,r)) = \phi_x(r) > 0$ where $B(x,r) = \{x' \in F : d(x',x) < r\}$
- (H2) For all $k \ge 2$, we suppose that:

$$\begin{cases} (i) \ \textit{For all } 1 \le i_1 < \ldots i_k \le n, \textit{ the conditional density of } (Y_{i_1},...Y_{i_k}) \textit{ given } (X_{i_1},\ldots X_{i_k}) \\ \quad \textit{exists and is uniformely boundes} \\ (ii) \ \textit{There exists } v_k > 0, \textit{ such that} \\ \quad \max(max_{1\le i_1<\ldots i_k\le n} P(d(X_{i_j},x) \le r, 1 \le j \le k), \phi_x^k(r)) = O(\phi_x^{1+v_k}(r)). \end{cases}$$

- (H3) $(X_i, Y_i)_{i \in N}$ is an α-mixing sequence of mixing coefficient $\alpha(n)$ satisfying

$$\exists a > max\left(max_{2\le k\le p}(k-1)\frac{1+v_k}{v_k}, p \right), \text{ such that } \forall n \in N\ \alpha(n) \le Cn^{-a}$$

- (H4) f^x is $2-$times continuously differentiable with respect y on R such that, $\forall(y_{1,}y_2)\in \mathrm{R}^2$;

$$\forall(x_{1,}x_2)\in N_x^2 \quad \left|f^{x_1(j)}(y_1)-f^{x_2(j)}(y_2)\right|\leq (d^{b_1}(x_1,x_2),+\left|y_1-y_2\right|^{b_2}),\ \ b_1>0,b_2>0$$

for $j=0,1,2$ with the convention $f^{x(0)}=f^x$

- (H5) K is a function with support $(0,1)$ such that $0<C'<K(t)<C<\infty.$
- (H6) H is of classe C^2, of compact support and satisfies

$$\int H(t)dt=1, \quad and$$

$$\forall(y_{1,}y_2)\in \mathrm{R}^2 \quad \left|H^{(j)}(y_1)-H^{(j)}(y_2)\right|\leq C\left|y_1-y_2\right|,\ \ j=0,1,2 \ \ with\ H^{(0)}=H.$$

- (H7) There exists $\gamma_{1,},\gamma_2>0$, such that

$$Cn^{\frac{3-a}{a+1}+\gamma_1}\leq h_H\phi_x(h_K)<n^{\frac{1}{1-a}} \ \ and\ \ lim_{n\to\infty}n^{\gamma}h_H=\infty \ \text{ with } \ a>(5+\sqrt{17})/2\,.$$

The concentration propriety (H1) is less restrictive that the fractal condition introduced by Gasser *et al.* (1998) and is known to hold for several continuous time processes (see for instance Bogachev (1999) for a gaussian measure, Li and Shao (2001) for a general gaussian process and Ferraty *et al.* (2006) for more discussion). In order to establish the same convergence rate as in the i.i.d. case (see Dabo-Niang and Laksaci (2007)), we reinforce the mixing by introducing (H2) and (H3). Note that we can establish the convergence results without these mixing assumptions, however, the convergence rate expression will be perturbed, it will contain the covariance term of the observations. Assumptions (H4) is the regularity condition which characterizes the functional space of our model and is needed to evaluate the bias term in our asymptotic developments. Assumptions (H6) and (H7) are standard technical conditions in nonparametric estimation. They are imposed for sake of simplicity and brevity of the proofs.

3. MAIN RESULTS

We establish the p-mean rate of convergence of the estimate $\hat{\theta}(x)$ to $\theta(x)$.

Theorem 1: *Under hypotheses (H1)-(H7), and if* $nh_H^3\phi_x(h_K)\to\infty$, *we have for all* $p\in[1,\infty[$

$$\hat{\theta}(x)-\theta(x)_p = O(h_K^{b_1}+h_H^{b_2})+O\left(\left(\frac{1}{nh_H^3\phi_x(h_K)}\right)^{\frac{1}{2}}\right).$$

where $.=(E^{1/p}|.|^p)$.

Proof of Theorem 1: For all $i=1,\ldots,n$, let $K_i=K(h_K^{-1}d(x,X_i))$, $H_i(y)=H(h_H^{-1}(y-Y_i))$ and

$$\hat{f}_N^x(y)=\frac{1}{nh_H EK_1}\sum_{i=1}^n K_iH_i(y), \qquad \hat{f}_D^x=\frac{1}{nEK_1}\sum_{i=1}^n K_i.$$

We consider a Taylor development of the function $\hat{f}_N^{x(1)}$ at the vicinity of $\hat{\theta}(x)$, in particular for the point $\theta(x)$,

$$\hat{f}_N^{x(1)}(\hat{\theta}(x))=\hat{f}_N^{x(1)}(\theta(x))+(\hat{\theta}(x)-\theta(x))\hat{f}_N^{x(2)}(\theta^*(x))$$

where $\theta^*(x)$ is between $\hat{\theta}(x)$ and $\theta(x)$. By the unimodality of f^x and assumption (H6), we have

$$f^{x(1)}(\theta(x))=\hat{f}_N^{x(1)}(\hat{\theta}(x))=0, \text{ and } f^{x(2)}(\theta(x))<0$$

Thus, if $\hat{f}_D^x\neq 0$, we have

$$\hat{\theta}(x)-\theta(x)=\frac{1}{\hat{f}_N^{x(2)}(\theta^*(x))}(\hat{f}_N^{x(1)}(\theta(x))-f^{x(1)}(\theta(x))) \tag{2}$$

It is shown in Theorem 11.15 of Ferraty and Vieu (2006, P.179) that, under (H1)-H(7),

$$\hat{\theta}(x)-\theta(x)\to 0 \qquad \textit{almost completely } (a.co.)$$

So, by combining this consistency and the result of Lemma 11.17 in Ferraty and Vieu (2006, P.181) together with the fact that $\theta^*(x)$ is lying between $\hat{\theta}(x)$ and $\theta(x)$, it follows that

$$\hat{f}_N^{x(2)}(\theta^*(x))-f^{x(2)}(\theta(x))\to 0. \qquad a.co.$$

Since $\left|f^{x(2)}(\theta(x))\right|>0$, we can write

$$\exists C>0, \quad \text{such that} \quad \left|\frac{1}{\hat{f}_N^{x(2)}(\theta^*(x))}\right| \le C \ \ a.s.$$

It follows that

$$\hat{\theta}(x)-\theta(x) \le C\|\hat{f}_N^{x(1)}(\theta(x)) - f^{x(1)}(\theta(x))\|_p + (P(\hat{f}_D^x = 0))^{1/p}.$$

Theorem 1 is then a consequence of the following lemmas.

Lemma 1: *Under hypotheses (H1)-(H3), (H5)-(H6), we have,*

$$\left\|\hat{f}_N^{x(1)}(\theta(x)) - E[\hat{f}_N^{x(1)}(\theta(x))]\right\|_p = O\left(\left(\frac{1}{nh_H^3\phi_x(h_K)}\right)^{\frac{1}{2}}\right) \tag{3}$$

Lemma 2: *If the hypotheses (H1), (H4)-(H6) are satisfied, we get*

$$E[\hat{f}_N^{x(1)}(\theta(x))] - f^{x(1)}(\theta(x)) = O(h_K^{b_1} + h_H^{b_2})$$

Lemma 3: *Under the conditions of Lemma 1, we get*

$$(P(\hat{f}_D^x = 0))^{1/p} = O\left(\left(\frac{1}{n\phi_x(h_K)}\right)^{\frac{1}{2}}\right).$$

Labo. EQUIPPE, Maison de Recherche, SOPHIE DABO-NIANG
Univ. Lille3,
BP60149, 59653 Villeneuve d'Ascq cedex Lille, France

Laboratoire de Mathématiques ALI LAKSACI
Univ. Djillali Liabès,
BP 89, 22000 Sidi Bel Abbès, Algérie

ACKNOWLEDGEMENTS

The authors would like to thank the anonym referee, whose remarks permit us to improve substantially the quality of the paper and clarify some points.

APPENDIX

Proof of Lemma 1: Let

$$\Delta_i = K_i H^{(1)}(h_H^{-1}(\theta(x)-Y_i)) - E[K_i H^{(1)}(h_H^{-1}(\theta(x)-Y_i))].$$

Then

$$\left\|\hat{f}_N^{x(1)}(\theta(x)) - E[\hat{f}_N^{x(1)}(\theta(x))]\right\|_p = \frac{1}{nh_H^2 EK_1}\left\|\sum_{i=1}^{n}\Delta_i\right\|_p$$

Because of (H1) and (H5) we can write $EK_1 = O(\phi_x(h_K))$. So, it remains to show that

$$\sum_{i=1}^{n}\Delta_{i\,p} = O(\sqrt{nh_H\phi_x(h_K)})$$

The evaluation of this quantity is based on ideas similar to that used by Yokoyama (1980). More preciously, we prove the case where $p = 2m$ (for all $m \in N^*$) and we use the Holder inequality for lower values of p. Indeed, if $p = 2m$, we have

$$E\left[\left|\sum_{i=1}^{n}\Delta_i\right|^{2m}\right] = \sum_{k=1}^{2m}\sum_{\begin{cases} p_j,\, p_j \geq 1 \\ \sum_{j=1}^{k} p_j = 2m\end{cases}} C_{2m}^{p_1 p_2,\ldots,p_k}\sum_{1\leq i_1<\ldots<i_k\leq n} E\left[\prod_{j=1}^{k}\Delta_{i_j}^{p_j}\right] \tag{4}$$

where $C_{2m}^{p_1 p_2,\ldots,p_k} = \dfrac{(2m)!}{p_1!\,p_2!\ldots p_k!}$. The stationarity of the couples $(X_i, Y_i)_{i=1,\ldots,n}$ allows us to write

$$\sum_{1\leq i_1<\ldots<i_k\leq n} E\left[\prod_{j=1}^{k}\Delta_{i_j}^{p_j}\right] \leq \sum_{1\leq i_1<\ldots<i_k\leq n}\left|E\left[\prod_{j=1}^{k}\Delta_{i_j}^{p_j}\right]\right| \leq n\sum_{1\leq e_1,\ldots,e_{k-1}\leq n}\left|E\left[\prod_{j=1}^{k}\Delta_{1+\sum_{b=1}^{j-1}e_b}^{p_j}\right]\right|$$

Now, we show by induction on m that the last term above is bounded for all k and p_j, $1 \leq j \leq k$ as follows

$$A(m): \quad \begin{aligned} &\forall k \leq 2m, \forall p_j \in N^* \ \sum_{i=1}^{k} p_j = q \leq 2m,\ \exists C_1, \forall n \in N \\ &n\sum_{1\leq e_1,\ldots,e_{k-1}\leq n}\left|E\left[\prod_{j=1}^{k}\Delta_{1+\sum_{b=1}^{j-1}e_b}^{p_j}\right]\right| \leq C_1\ (nh_H\phi_x(h_K))^{q/2}\end{aligned}$$

Firstly, in the case where $m=1$, we have $q=1$ or $q=2$. If $q=1$, then $A(1)$ is necessarily satisfied since the Δ_i are centered. Indeed, $q=1$ implies that $k=1$ and then $nE[\Delta_1]\leq C_1\,(nh_H\phi_x(h_K))^{\frac{1}{2}}$. However, if $q=2$, we distinguish two cases: $k=2$ or $k=1$. For the last case, we have under (H1), that $\sum_{i=1}^{n}E[\Delta_i^2]\leq Cnh_H\phi_x(h_K)$ while for the first case we use the same ideas as those used in the proof of Lemma 11.13 of Ferraty and Vieu (2006, P. 175) and obtain

$$\left|\sum_{1\leq i\neq j\leq n}E[\Delta_i\Delta_i]\right|\leq nh_H\phi_x(h_K) \tag{5}$$

Thus, $A(m)$ is true for $m=1$. The next step is to show that $A(m)$ is hereditary. In other words, we suppose that $A(m)$ is true and we prove $A(m+1)$. Obviously, we can suppose that $q\geq 2$ (because the upper bound to prove is always true if $q=1$). Thus, for all $k\leq 2m+2$ and $\forall p_j\in N^*\ \sum_{i=1}^{k}p_j=q\leq 2m+2$, we have

$$n\sum_{1\leq e_1,\ldots,e_{k-1}\leq n}\left|E\left[\prod_{j=1}^{k}\Delta^{p_j}_{1+\sum_{b=1}^{j-1}e_b}\right]\right|\leq n\sum_{s=1}^{k-1}\sum_{e_s=1}^{n}\sum_{1\leq e_j\leq e_s,\,j\neq s}\left|E\left[\prod_{j=1}^{k}\Delta^{p_j}_{1+\sum_{b=1}^{j-1}e_b}\right]\right| \tag{6}$$

We split the above sum as follows

$$\begin{aligned}\sum_{e_s=1}^{n}\sum_{1\leq e_j\leq e_s,\,j\neq s}\left|E\left[\prod_{j=1}^{k}\Delta^{p_j}_{1+\sum_{b=1}^{j-1}e_b}\right]\right|\leq&\sum_{e_s=1}^{u_n}\sum_{1\leq e_j\leq e_s,\,j\neq s}\left|E\left[\prod_{j=1}^{k}\Delta^{p_j}_{1+\sum_{b=1}^{j-1}e_b}\right]\right|\\&+\sum_{e_s=u_n+1}^{n}\sum_{1\leq e_j\leq e_s,\,j\neq s}\left|E\left[\prod_{j=1}^{k}\Delta^{p_j}_{1+\sum_{b=1}^{j-1}e_b}\right]\right|\end{aligned} \tag{7}$$

The first term of (7) is such that

$$A:=\sum_{e_s=1}^{u_n}\sum_{1\leq e_j\leq e_s,\,j\neq s}\left|E\left[\prod_{j=1}^{k}\Delta^{p_j}_{1+\sum_{b=1}^{j-1}e_b}\right]\right|\leq Cu_n^{k-1}h_H^{k}\phi_x^{1+v_k}(h_K) \tag{8}$$

For the second term, we have

$$B \le \sum_{e_s=u_n+1}^{n} \sum_{1\le e_j\le e_s,\, j<s} \left| E\left[\prod_{j=1}^{k} \Delta^{p_j}_{1+\sum_{b=1}^{j-1} e_b}\right] \right| \sum_{1\le e_j\le e_s,\, j>s} \left| E\left[\prod_{j=1}^{k} \Delta^{p_j}_{1+\sum_{b=1}^{j-1} e_b}\right] \right|$$
$$+ \sum_{e_s=u_n+1}^{n} \sum_{1\le e_j\le e_s,\, j\ne s} \alpha(e_s) \tag{9}$$
$$\le B_1 + B_2$$

Let us first consider the term B_1 and observe that if $k = 2m+2$ and $s = k-1$ then $p_j = 1$, for all $1 \le j \le 2m+2$ and the second expectation is null (because the Δ_i are centered). So, in the case where none of the expectations is null, we have $k < 2m+2$ or $s < k-1$ which imply that $s \le 2m$. Moreover, we can prove in the same way that $k - s \le 2m$. We have that, when $B_1 \ne 0$, then

$\sum_{j=1}^{s} p_j \le 2m$ ($\sum_{j=1}^{k} p_j \le 2m+2$ and $\sum_{j=s+1}^{k} p_j \ge 2$ since either $s < k-1$ or $p_k \ge 2$ and $\sum_{j=s+1}^{s} p_j \le 2m$. From what precedes, notice that either $B_1 = 0$ or $q \ge 4$ (since $\sum_{j=s+1}^{k} p_j \ge 2$ and $\sum_{j=1}^{s} p_j \ge 2$). So, we apply $A(m)$ and get

Hence,

$$B_1 \le \frac{1}{n} C \sqrt{n h_H \phi_x(h_K)}^{\,q}. \tag{10}$$

For the B_2 term, we use the fact that

$$B_2 \le \sum_{e_s=u_n}^{n} e_s^{k-2} \alpha(e_s) \le C \sum_{i=u_n}^{n} i^{k-2-a+1+\eta-1-\eta}$$
$$\le C u_n^{k-2-a+1+\eta} \sum_{i=u_n}^{n} i^{-1-\eta} \le C u_n^{k-1-a+\eta}, \quad \eta > 0. \tag{11}$$

It is easy to see that the upper bounds (8) and (11) are respectively increasing and decreasing with respect to u_n. This gives the idea to choose u_n in order to balance these two terms. We choose $u_n = \left(h_H^{k} (\phi_x(h_K))^{1+v_k}\right)^{\frac{1}{\eta-a}}$. Then

$$A + B_2 \le C h_H \phi_x(h_K) h_H^{\frac{k(1-k)}{a-\eta}+k-1} (\phi_x(h_K))^{\frac{(1-k)(1+v_k)}{a-\eta}+v_k}$$

Furthermore, since $a > \max\left(max_{2\le k\le p}(k-1)\frac{1+v_k}{v_k}, p \right)$, then, we can get some small $\eta > 0$ such that $(1-k)(1+v_k)+v_k(a-\eta) > 0$ and $(1-k)k+(k-1)(a-\eta) > 0$. Consequently

$$A + B_2 \le Ch_H \phi_x(h_K). \tag{12}$$

We easily deduce from (6), (7), (10) and (12), the following upper bound

$$\begin{aligned} n\sum_{1\le e_1, e_2,\ldots,e_{k-1}\le n} \left| E\left[\prod_{j=1}^{k} \Delta^{p_j}_{1+\sum_{v=1}^{j-1} e_v} \right] \right| &\le C\sqrt{nh_H\phi_x(h_K)}^{q} + C' nh_H\phi_x(h_K) \\ &\le C\sqrt{nh_H\phi_x(h_K)}^{q}. \end{aligned} \tag{13}$$

The last line comes from the fact that $q \ge 2$ and $nh_H\phi_x(h_K) \to \infty$. So, we conclude that $A(m)$ is hereditary, so it is true for all $m \in N^*$. We directly conclude from (13) and (4) that

$$E\left[\left| \sum_{i=1}^{n} \Delta_i \right|^{2m} \right] \le C(nh_H\phi_x(h_K))^m.$$

Finally, it suffices to use the Holder inequality to show that for all $p < 2m$

$$\left\| \sum_{i=1}^{n} \Delta_i \right\|_p \le \left\| \sum_{i=1}^{n} \Delta_i \right\|_{2m} \le C\sqrt{nh_H\phi_x(h_K)}.$$

This yields the proof of this lemma.

■

Proof of Lemma 2: It is easy to see that

$$E[\hat{f}_N^{x(1)}(\theta(x))] - f^{x(1)}(\theta(x)) = \frac{1}{h_H^2 EK_1} E[K_1 E[H_1^{(1)}(\theta(x)) \mid X_1]] - f^{x(1)}(\theta(x)).$$

By an integration by part and the change of variables $= \frac{y-z}{h_H}$, we have

$$E[\hat{f}_N^{x(1)}(\theta(x))] - f^{x(1)}(\theta(x)) \le \frac{1}{EK_1}(EK_1 \int H(t) \left| f^{X_1(1)}(\theta(x) - h_H t) - f^{x(1)}(\theta(x)) \right| dt).$$

Hypotheses (H4) and (H6) allow to get the desired result.

■

Proof of Lemma 3: It is clear that, for all $\varepsilon < 1$, we have

$$P(\hat{f}_D^x = 0) \leq P(\hat{f}_D^x \leq 1-\varepsilon) \leq P(\left|\hat{f}_D^x - E[\hat{f}_D^x]\right| \geq \varepsilon).$$

The Markov's inequality allows to get, for any $p > 0$,

$$P(\left|\hat{f}_D^x - E[\hat{f}_D^x]\right| \geq \varepsilon) \leq \frac{E[\left|\hat{f}_D^x - E[\hat{f}_D^x]\right|^p]}{\varepsilon^p}.$$

So

$$(P(\hat{f}_D^x = 0))^{1/p} = O(\left\|\hat{f}_D^x - E[\hat{f}_D^x]\right\|_p).$$

The computation of $\left\|\hat{f}_D^x - E[\hat{f}_D^x]\right\|_p$ can be done by following the same arguments as those invoked to get (3). This yields the proof
■

REFERENCES

A. BERLINET, A. GANNOUN and E. MATZNER-LOBER, (1998), *Normalité asymptotique d'estimateurs convergents du mode conditionnel.* "Canad. J. Statist", 26, pp. 365-380.
V. I. BOGACHEV, (1999), *Gaussian measures.* Math surveys and monographs, 62, Amer. Math. Soc.
T. BOLLERSLEV, (1986), *General autoregressive conditional heteroskedasticity.* "J. Econom", 31, pp. 307-327.
D. BOSQ, (2000), *Linear Processes in Function Spaces: Theory and applications.* Lecture Notes in Statistics, 149, Springer.
G. COLLOMB, W. HARDLE, and S. HASSANI, S. (1987), *A note on prediction via conditional mode estimation.* "J. Statist.Plann. Inference", 15, pp. 227-236.
S. DABO-NIANG, (2002), *Estimation de la densité dans un espace de dimension infinie: application aux diffusions.* "C. R., Math., Acad. Sci. Paris", 334, pp. 213-216.
S. DABO-NIANG and N. RHOMARI, (2003), *Estimation non paramétrique de la régression avec variable explicative dans un espace métrique.* "C. R., Math., Acad. Sci. Paris", 336, pp. 75-80.
S. DABO-NIANG, and N. RHOMARI, (2009), *Kernel regression estimate in a Banach space.* "J. Statist. Plann. Inference", 139, 1421-1434.
S. DABO-NIANG, and A. LAKSACI, (2007), *Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle.* "C. R., Math., Acad. Sci. Paris", 344, pp. 49-52.
R. F. ENGLE, (1982), *Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation.* "Econometrica", 50, pp. 987-1007.
M. EZZAHRIOUI, and E. OULD-SAID, (2008). *Asymptotic normality of nonparametric estimators of the conditional mode function for functional data.* "J. Nonparametr. Stat", 20, 3-18.
M. EZZAHRIOUI and E. OULD-SAID, (2006). *On the asymptotic properties of a nonparametric estimator of the conditional mode function for functional dependent data.* LMPA, janvier 2006 (Preprint).
F. FERRATY, A. LAKSACI, and PH. VIEU, (2005), *Functional times series prediction via conditional mode.* "C. R., Math., Acad. Sci. Paris", 340, pp. 389-392.

F. FERRATY, A. LAKSACI, and PH. VIEU, (2006), *Estimating some characteristics of the conditional distribution in nonparametric functional models*. "Stat. Inference Stoch. Process", 9, pp. 47-76.

F. FERRATY, and PH. VIEU, (2006), *Nonparametric functional data analysis*. Springer-Verlag, New-York.

T. GASSER, P. HALL, and B. PRESNELL, (1998). *Nonparametric estimation of the mode of a distribution of random curves*. "J. R.Statist. Soc, B", 60, pp. 681-691.

D. A. JONES, (1978), *Nonlinear autoregressive processes*. "Proc. Roy. Soc. London A", 360, pp. 71-95.

W. V. LI, and Q.M. SHAO, (2001), *Gaussian processes: inequalities, small ball probabilities and applications*. In Stochastic processes: Theory and Methods, Ed. C.R. Rao and D. Shanbhag. Hanbook of Statistics, 19, North-Holland, Amsterdam.

D. LOUANI, and E. OULD-SAID, (1999), *Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis*. "J. Nonparam. Stat", 11, pp. 413-442.

E. MASRY, (2005), *Nonparametric regression estimation for dependent functional data: asymptotic normality*. "Stochastic Process. Appl", 115, pp. 155-177.

T. OZAKI, (1979), *Nonlinear time series models for nonlinear random vibrations*. Technical report. Univ. of Manchester.

J. O. RAMSAY and B.W. SILVERMAN, (2002), *Applied functional data analysis. Methods and case studies*. Springer-Verlag, New York.

A. QUINTELA DEL RIO, and PH. VIEU, (1997), *A nonparametric conditional mode estimate*. "J. Nonparam. Stat", 8, pp. 253-266.

M. ROSENBLATT, (1969), *Conditional probability density and regression estimators*. Multivariate Analysis II, Ed. P.R. Krishnaiah. Academic Press, New York and London.

E. YOUNDJÉ, (1993), *Estimation non paramétrique de la densité conditionnelle par la méthode du noyau*. Thèse de Doctorat, Université de Rouen.

R. YOKOYAMA, (1980), *Moments bounds for stationary mixing sequences*. "Z. Wahrs. verw. Gebiete", 52, pp. 45-57.

SUMMARY

Note on conditional mode estimation for functional dependent data

We consider α-mixing observations and deal with the estimation of the conditional mode of a scalar response variable Y given a random variable X taking values in a semi-metric space. We provide a convergence rate in L^p norm of the estimator.