

SOME ASYMPTOTIC RESULTS OF A NONPARAMETRIC CONDITIONAL HAZARD ESTIMATOR WITH FUNCTIONAL EXPLANATORY VARIABLE UNDER STRONG MIXING CONDITIONS

ABBES RABHI, AICHA HAMZA and AMINA ANGELIKA BOUCHENTOUF

Communicated by Lucian Beznea

We consider the problem of nonparametric estimation of the conditional hazard function for functional mixing data. More precisely, given strictly stationary random variables $Z_i = (X_i, Y_i)_{i \in \mathbb{N}}$, we investigate a kernel estimate of the conditional hazard function of a scalar response variable Y_i given the functional variable X_i . The principal aim of this paper is to give the mean squared convergence rate and to prove the asymptotic normality of the proposed estimator in the context of strong mixing dependence.

AMS 2010 Subject Classification: Primary 62H12; Secondary 62N02, 62J05.

Key words: functional data, kernel conditional hazard function, kernel estimation, strong mixing process.

1. INTRODUCTION

The statistical problems involved in the modelization of functional data have received an increasing interest in the literature. The infatuation for this topic is linked with many fields of applications in which the data are collected in the functional order. Under this supposition, the statistical analysis focuses on a framework of infinite dimension for the data under study. This field of modern statistics has received much attention in the last 20 years, and it has been popularized in the book of Ramsay and Silverman [21]. This type of data appears in many fields of applied statistics: environmetrics [8], meteorological sciences [3], etc.

In this paper, we are interested in the nonparametric estimation of the conditional hazard function when the co-variate are of functional nature.

The nonparametric estimation of the hazard and/or the conditional hazard function is quite important in a variety of fields including medicine, reliability, survival analysis or in seismology. The literature on this model in multivariate statistics is abundant. Historically, the hazard estimate was intro-

duced by Watson and Leadbetter [28], since, several results have been added, see for example, Roussas [24] (for previous works), Li and Tran (2007) (for recent advances and references).

From a theoretical point of view, a sample of functional data can be involved in many different statistical problems, such as for example: classification and principal components analysis (PCA) [4,22] or longitudinal studies, regression and prediction [2, 7].

The recent monograph by Ferraty and Vieu [12] summarizes many of their contributions to the non-parametric estimation with functional data; among other properties, consistency of the conditional density, conditional distribution and regression estimates are established in the i.i.d. case as well as under dependence conditions (strong mixing). Almost complete rates of convergence are also obtained, and the different techniques are applied to several examples of functional data samples. Related work can be seen in the paper of Masry [17], where the asymptotic normality of the functional non-parametric regression estimate is proven, considering strong mixing dependence conditions for the sample data. For automatic smoothing parameter selection in the regression setting, see Rachdi and Vieu [20].

The literature is strictly not limited in the case where the data is of functional nature (a curve). The first result in this context, was given by Ferraty *et al.* [11]. They established the almost complete convergence of the kernel estimate of the conditional hazard function in the i.i.d. case and under α -mixing condition. Recently, Rabhi *et al.* [19] studied the mean quadratic convergence in the i.i.d. case of this estimate. More recently Mahiddine *et al.* [16] give the uniform version of the almost complete convergence rate in the i.i.d. case.

The estimation of the hazard function is a problem of considerable interest, especially to inventory theorists, medical researchers, logistics planners, reliability engineers and seismologists. The non-parametric estimation of the hazard function has been extensively discussed in the literature. Beginning with Watson and Leadbetter [28], there are many papers on these topics: Ahmad [1], Singpurwalla and Wong [25], etc. We can cite Quintela [18] for a survey.

When hazard rate estimation is performed with multiple variables, the result is an estimate of the conditional hazard rate for the first variable, given the levels of the remaining variables. Many references, practical examples and simulations in the case of non-parametric estimation using local linear approximations can be found in Spierdijk [26].

The main aim of this paper, is to study, under general conditions, the asymptotic proprieties of the functional data kernel estimate of the conditional hazard function introduced by Ferraty *et al.* [11]. More precisely, we treat the L^2 -convergence rate by giving the exact expression involved in the leading terms

of the quadratic error. In addition, we establish the asymptotic normality of the construct estimator. We point out that our asymptotic results are useful in some statistical problems such as the choice of the smoothing parameters, the determination of confidence intervals and in risk analysis. The present work extended to dependent case the result of Rabhi *et al.* [19] given in i.i.d. case functional. We note that, one of the main difficulty, when dealing with functional variables, relies on the difficulty for choosing some appropriate measure of reference in infinite dimensional spaces. The main feature of our approach is to build estimates and to derive their asymptotic properties without any notion of density for the functional variable X . This approach allows us to avoid the use of a reference measure in such functional spaces. In each of the above described sections, we will give general asymptotic results without assuming existence of such a density, and each of these results will be discussed in relation with earlier literature existing in the usual finite dimensional case.

Our paper presents some asymptotic properties related with the non-parametric estimation of the conditional hazard function. In a functional data setting, the conditioning variable is allowed to take its values in some abstract semi-metric space. In this case, Ferraty *et al.* [27] define non-parametric estimators of the conditional density and the conditional distribution. They give the rates of convergence (in an almost complete sense) to the corresponding functions, in an a dependence (α -mixing) context. In Rabhi *et al.* [19], the same properties are shown in an i.i.d. context in the data sample. We extend their results to dependent case by calculating the bias and variance of these estimates, and establishing their asymptotic normality, considering a particular type of kernel for the functional part of the estimates. Because the hazard function estimator is naturally constructed using these two last estimators, the same type of properties is easily derived for it. Our results are valid in a real (one- and multi-dimensional) context.

The paper is organized as follows: in the next section we present our model. Section 3 is dedicated to fixing notations and hypotheses. We state our main results in Section 4. Section 5 is devoted to some discussions on the applicability of our asymptotic result in some statistical problems such as the choice of the smoothing parameters, the determination of confidence intervals and in seismology analysis.

2. THE MODEL

Consider $Z_i = (X_i, Y_i)$, $i \in \mathbb{N}$ be a $\mathcal{F} \times \mathbb{R}$ -valued measurable strictly stationary process, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, where (\mathcal{F}, d) is a semi-metric space. In most practical applications, \mathcal{F} is a normed space which

can be of infinite dimension (e.g., Hilbert or Banach space) with norm $\| \cdot \|$ so that $d(x, x') = \|x - x'\|$.

In the following, x will be a fixed point in \mathcal{F} and N_x will denote a fixed neighborhood of x . We assume that the regular version of the conditional probability of Y given X exists. Moreover, we suppose that, for all $x \in N_x$ the conditional distribution function of Y given $X = x$, $F^x(\cdot)$, is 3-times continuously differentiable and we denote by f^x its conditional density with respect to (w.r.t.) Lebesgue's measure over \mathbb{R} . In this paper, we consider the problem of the nonparametric estimation of the conditional hazard function defined, for all $t \in \mathbb{R}$ such that $F^x(t) < 1$, by

$$\lambda^x(t) = \frac{f^x(t)}{1 - F^x(t)}.$$

In this context, we estimate this function by

$$\widehat{\lambda}^x(t) = \frac{\widehat{f}^x(t)}{1 - \widehat{F}^x(t)}$$

where

$$\widehat{F}^x(t) = \sum_{i=1}^n W_{ni}(x) \mathbf{1}_{\{Y_i \leq t\}}, \quad \forall t \in \mathbb{R}$$

and

$$\widehat{f}^x(t) = \sum_{i=1}^n W_{ni}(x) H(h^{-1}(t - Y_i)), \quad \forall t \in \mathbb{R}$$

with $\mathbf{1}_{\{\cdot\}}$ being the indicator function and where $W_{ni}(x) = \frac{h^{-1}K(h^{-1}\|x - X_i\|)}{\sum_{j=1}^n K(h^{-1}\|x - X_j\|)}$,

K is a kernel function and $h := h_n$ is a sequence of positive real numbers which goes to zero as n goes to infinity. Furthermore, the estimator $\widehat{h}^x(t)$ can be written as

$$(2.1) \quad \widehat{\lambda}^x(t) = \frac{\widehat{f}_N^x(t)}{\widehat{F}_D^x - \widehat{F}_N^x(t)}$$

where

$$(2.2) \quad \begin{aligned} \widehat{F}_D^x &:= \frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K(h^{-1}\|x - X_i\|), \quad K_1(x) = K(h^{-1}\|x - X_1\|) \\ \widehat{F}_N^x(t) &:= \frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K(h^{-1}\|x - X_i\|) \mathbf{1}_{\{Y_i \leq t\}} \\ \widehat{f}_N^x(t) &:= \frac{1}{nh\mathbb{E}[K_1(x)]} \sum_{i=1}^n K(h^{-1}\|x - X_i\|) H(h^{-1}(t - Y_i)). \end{aligned}$$

Our main purpose is to study the L^2 - consistency and the asymptotic normality of the nonparametric estimate \widehat{h}^x of h^x when the random filed $(Z_i, i \in \mathbb{N})$ satisfies the following mixing condition.

3. NOTATIONS AND HYPOTHESES

All along the paper, when no confusion is possible, we will denote by C, C', C_1, C_2, \dots some strictly positive generic constants. In order to establish our asymptotic results we need the following hypotheses:

(H0) $\forall r > 0, \mathbb{P}(X \in B(x, r)) =: \phi_x(r) > 0$, where $B(x, r) = \{x' \in \mathcal{F} / \|x - x'\| < r\}$.

(H1) $(X_i, Y_i)_{i \in \mathbb{N}}$ is an α -mixing sequence whose coefficients of mixture verify:

$$\exists a > (5 + \sqrt{17})/2, \exists c > 0 : \forall n \in \mathbb{N}, \alpha(n) \leq cn^{-a}.$$

(H2) $0 < \sup_{i \neq j} \mathbb{P}((X_i, X_j) \in B(x, h) \times B(x, h)) = \mathcal{O}\left(\frac{(\phi_x(h))^{(a+1)/a}}{n^{1/a}}\right)$.

Note that (H0) can be interpreted as a concentration hypothesis acting on the distribution of the *f.r.v.* X , whereas (H2) concerns the behavior of the joint distribution of the pairs (X_i, X_j) . In fact, this hypothesis is equivalent to assume that, for n large enough

$$\sup_{i \neq j} \frac{\mathbb{P}((X_i, X_j) \in B(x, h) \times B(x, h))}{\mathbb{P}(X \in B(x, h))} \leq C \left(\frac{\phi_x(h)}{n}\right)^{1/a}.$$

(H3) For $l \in \{0, 2\}$, the functions $\Psi_l(s) = \mathbb{E}\left[\frac{\partial^l F^X(t)}{\partial t^l} - \frac{\partial^l F^x(t)}{\partial t^l} \middle| \|x - X\| = s\right]$ and

$$\Phi_l(s) = \mathbb{E}\left[\frac{\partial^l f^X(t)}{\partial t^l} - \frac{\partial^l f^x(t)}{\partial t^l} \middle| \|x - X\| = s\right] \text{ are derivable at } s = 0.$$

(H4) The bandwidth h satisfies:

$$\begin{cases} \text{(H4a)} & \lim_{n \rightarrow \infty} h_n = 0, \lim_{n \rightarrow \infty} \frac{\log n}{nh^j \phi_x(h)} = 0, j = 0, 1 \\ \text{(H4b)} & \exists \eta > 0 \text{ such that } C_1 n^{\frac{3-a}{a+1} + \eta} \leq h \phi_x(h) \text{ and } \phi_x(h) \leq C_2 n^{\frac{1}{1-a}}. \end{cases}$$

(H5) The kernel K is measurable function with support $[0, 1]$ and satisfies $0 < C\mathbf{1}_{[0,1]}(\cdot) < K(\cdot) < C'\mathbf{1}_{[0,1]}(\cdot) < \infty$.

(H6) H is a bounded function, such that for $b_1 > 0$

$$\int_{\mathbb{R}} zH(z)dz = 0, \int_{\mathbb{R}} z^2H(z)dz < \infty, \text{ and } \int_{\mathbb{R}} |z|^{b_1} H^2(z)dz < \infty.$$

3.1. Remarks on the assumptions

Remark 3.1. Assumption (H0) plays an important role in our methodology. It is known as (for small h) the “concentration hypothesis acting on the distribution of X ” in infinite-dimensional spaces. This assumption is not at all restrictive and overcomes the problem of the non-existence of the probability density function. In many examples, around zero the small ball probability $\phi_x(h)$ can be written approximately as the product of two independent functions $\psi(z)$ and $\varphi(h)$ as $\phi_x(h) = \psi(z)\varphi(h) + o(\varphi(h))$. This idea was adopted by Masry [17] who reformulated the Gasser *et al.* [13] one.

Without the differentiability of $\phi_x(\cdot)$, this assumption has been used by many authors where $\psi(\cdot)$ is interpreted as a probability density, while $\varphi(\cdot)$ may be interpreted as a volume parameter. In the case of finite-dimensional spaces, that is $\mathcal{S} = \mathbb{R}^d$, it can be seen that $\phi_x(h) = C(d)h^d\psi(x) + o(h^d)$, where $C(d)$ is the volume of the unit ball in \mathbb{R}^d . Furthermore, in infinite dimensions, there exist many examples fulfilling the decomposition mentioned above. We quote the following (which can be found in Ferraty *et al.* [9]):

- (1) $\phi_x(h) \approx \psi(h)h^\gamma$ for some $\gamma > 0$.
- (2) $\phi_x(h) \approx \psi(h)h^\gamma \exp\{C/h^p\}$ for some $\gamma > 0$ and $p > 0$.
- (3) $\phi_x(h) \approx \psi(h)/|\ln h|$.

Assumption (H2) is classical and permits to make the variance term negligible.

Remark 3.2. Assumption (H3) is a regularity condition which characterizes the functional space of our model and is needed to evaluate the bias.

Remark 3.3. Assumptions (H5) and (H6) are classical in functional estimation for finite or infinite dimension spaces.

3.2. Main results

3.3. Mean squared convergence

The first result concerns the L^2 -consistency of $\widehat{\lambda}^x(y)$.

THEOREM 3.1. *Under assumptions (H0)–(H6), we have*

$$\mathbb{E} \left[\widehat{\lambda}^x(t) - \lambda^x(t) \right]^2 = B_n^2(x, t) + \frac{\sigma_\lambda^2(x, t)}{nh\phi_x(h)} + o(h^4) + o\left(\frac{1}{nh\phi_x(h)}\right),$$

where

$$B_n(x, t) = \frac{h^2 \left(B_H^f - \lambda^x(t)B^F \right) + h \left(B_K^f - \lambda^x(t)B_K^F \right)}{1 - F^x(t)}$$

with

$$\begin{aligned}
 B_H^f(x, t) &= \frac{1}{2} \frac{\partial^2 f^x(t)}{\partial t^2} \int_{\mathbb{R}} z^2 H(z) dz \\
 B_K^f(x, t) &= h \Phi'_0(0) \frac{\left(K(1) - \int_0^1 (sK(s))' \beta_x(s) ds \right)}{\left(K(1) - \int_0^1 K'(s) \beta_x(s) ds \right)} \\
 B^F(x, t) &= \frac{1}{2} \frac{\partial^2 F^x(t)}{\partial t^2} \int_{\mathbb{R}} z^2 H(z) dz \\
 B_K^F(x, t) &= h \Psi'_0(0) \frac{\left(K(1) - \int_0^1 (sK(s))' \beta_x(s) ds \right)}{\left(K(1) - \int_0^1 K'(s) \beta_x(s) ds \right)}.
 \end{aligned}$$

and

$$\sigma_\lambda^2(x, t) = \frac{\beta_2 \lambda^x(t)}{(\beta_1^2 (1 - F^x(t)))} \quad (\text{with } \beta_j = K^j(1) - \int_0^1 (K^j)'(s) \beta_x(s) ds, \text{ for } j=1, 2).$$

Proof. By using the same decomposition used in Theorem 3.1 (Rabhi *et al.* [19], p. 408), we show that the proof of Theorem 3.1 can be deduced from the following intermediate results:

LEMMA 3.1. *Under the hypotheses of Theorem (3.1), we have*

$$\mathbb{E} \left[\widehat{f}_N^x(t) \right] - f^x(t) = h^2 B_H^f(x, t) + h B_K^f(x, t) + o(h^2)$$

and

$$\mathbb{E} \left[\widehat{F}_N^x(t) \right] - F^x(t) = h^2 B^F(x, t) + h B_K^F(x, t) + o(h^2).$$

Remark 3.4. Observe that, the result of this lemma permits to write

$$\left[\mathbb{E} \widehat{F}_N^x(t) - F^x(t) \right] = \mathcal{O}(h^2),$$

and

$$\left[\mathbb{E} \widehat{f}_N^x(t) - f^x(t) \right] = \mathcal{O}(h^2).$$

LEMMA 3.2. *Under the hypotheses of Theorem (3.1), we have*

$$\text{Var} \left[\widehat{f}_N^x(t) \right] = \frac{\sigma_f^2(x, t)}{nh\phi_x(h)} + o \left(\frac{1}{nh\phi_x(h)} \right),$$

$$\text{Var} \left[\widehat{F}_N^x(t) \right] = o \left(\frac{1}{n\phi_x(h)} \right)$$

and

$$\text{Var} \left[\widehat{F}_D^x \right] = o \left(\frac{1}{n\phi_x(h)} \right).$$

where $\sigma_f^2(x, t) := f^x(t) \int_{\mathbb{R}} H^2(z) dz$.

LEMMA 3.3. *Under the hypotheses of Theorem (3.1), we have*

$$\text{Cov}(\widehat{f}_N^x(t), \widehat{F}_D^x) = o\left(\frac{1}{nh\phi_x(h)}\right),$$

$$\text{Cov}(\widehat{f}_N^x(t), \widehat{F}_N^x(t)) = o\left(\frac{1}{nh\phi_x(h)}\right)$$

and

$$\text{Cov}(\widehat{f}_D^x, \widehat{F}_N^x(t)) = o\left(\frac{1}{nh\phi_x(h)}\right).$$

Remark 3.5. It is clear that, the results of Lemmas (3.2 and 3.3) allows us to write

$$\text{Var}\left(\widehat{F}_D^x - \widehat{F}_N^x(t)\right) = o\left(\frac{1}{n\phi_x(h)}\right)$$

3.4. Asymptotic normality

Now, we study the asymptotic normality of $\widehat{\lambda}^x(t)$. We replace (H0), (H4) and (H5) by the following hypotheses, respectively.

(H0') The concentration property (H0) holds. Moreover, there exists a function $\beta_x(\cdot)$ such that

$$\forall s \in [0, 1] \quad \lim_{h \rightarrow 0} \frac{\phi_x(sh)}{\phi_x(h)} = \beta_x(s).$$

(H4') There exist sequences of integers (u_n) and (v_n) increasing to infinity such that $(u_n + v_n) \leq n$, satisfying

(i) $v_n = o((nh\phi_x(h))^{1/2})$ and $\left(\frac{n}{h\phi_x(h)}\right)^{1/2} \alpha(v_n) \rightarrow 0$ as $n \rightarrow 0$,

(ii) $q_n v_n = o((nh\phi_x(h))^{1/2})$ and $q_n \left(\frac{n}{h\phi_x(h)}\right)^{1/2} \alpha(v_n) \rightarrow 0$ as $n \rightarrow 0$

where q_n is the largest integer such that $q_n(u_n + v_n) \leq n$.

(H5') The kernel K satisfies (H5) and is a differentiable function supported on $]0, 1[$ with derivative K' such that $-\infty < C < K'(\cdot) < C' < 0$.

Remark 3.6. The increasing property of $\phi_x(\cdot)$ implies that $\beta_h^x(\cdot)$ is bounded and then integrable (all the more so $\beta_0^x(\cdot)$ is integrable).

The function $\beta_h^x(\cdot)$ is increasing for all fixed h . Its pointwise limit $\beta_0^x(\cdot)$ also plays a fundamental role in the asymptotic normality result. It permits giving the variance term explicitly. With simple algebra, it is possible to specify this function (with $\beta_0(u) := \beta_0^x(u)$) in the above examples by:

- (1) $\beta_0(u) = u^\gamma,$
- (2) $\beta_0(u) = \delta_1(u)$ where $\delta_1(\cdot)$ is Dirac function,
- (3) $\beta_0(u) = \mathbf{1}_{]0,1[}(u).$

THEOREM 3.2. Assume that (H0)–(H6) hold, and under hypotheses (H0'), (H4') and (H5'), then we have for any $x \in \mathcal{A}$,

$$\left(\frac{nh\phi_x(h)}{\sigma_\lambda^2(x,t)}\right)^{1/2} \left(\widehat{\lambda}^x(t) - \lambda^x(t) - B_n(x,t)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad \text{as } n \rightarrow \infty.$$

where

$$\mathcal{A} = \{x \in \mathcal{F}, f^x(t)(1 - F^x(t)) \neq 0\}$$

and $\xrightarrow{\mathcal{D}}$ means the convergence in distribution.

Obviously, if one imposes some additional assumptions on the function $\phi_x(\cdot)$ and the bandwidth parameter h we can improve our asymptotic normality by removing the bias term $B_n(x,t)$.

COROLLARY 3.1. Under the hypotheses of Theorem 3.2 and if the bandwidth parameter h and if the function $\phi_x(h)$ satisfies:

$$\lim_{n \rightarrow \infty} \sqrt{nh\phi_x(h)} = 0$$

we have

$$\left(\frac{nh\phi_x(h)}{\sigma_\lambda^2(x,t)}\right)^{1/2} \left(\widehat{\lambda}^x(t) - \lambda^x(t)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad \text{as } n \rightarrow \infty.$$

Proofs of Theorem and Corollary. We consider the decomposition

$$\begin{aligned} (3.1) \quad \widehat{\lambda}^x(t) - \lambda^x(t) &= \frac{1}{\widehat{F}_D^x - \widehat{F}_N^x(t)} \left(\widehat{f}_N^x(t) - \mathbb{E}\widehat{f}_N^x(t)\right) \\ &+ \frac{1}{\widehat{F}_D^x - \widehat{F}_N^x(t)} \left\{ \lambda^x(t) \left(\mathbb{E}\widehat{F}_N^x(t) - F^x(t)\right) + \left(\mathbb{E}\widehat{f}_N^x(t) - f^x(t)\right) \right\} \\ &\quad + \frac{\lambda^x(t)}{\widehat{F}_D^x - \widehat{F}_N^x(t)} \left\{ 1 - \mathbb{E}\widehat{F}_N^x(t) - \left(\widehat{F}_D^x - \widehat{F}_N^x(t)\right) \right\} \end{aligned}$$

Therefore, Theorem 3.2 and Corollary 3.1 are a consequence of Lemma 3.1, remark 3.4 and the following results.

LEMMA 3.4. Under the hypotheses of Theorem 3.2

$$\left(\frac{nh\phi_x(h)}{\sigma_f^2(x,t)}\right)^{1/2} \left(\widehat{f}_N^x(t) - \mathbb{E}\left[\widehat{f}_N^x(t)\right]\right) \rightarrow N(0,1).$$

LEMMA 3.5. Under the hypotheses of Theorem 3.2

$$\widehat{F}_D^x - \widehat{F}_N^x(t) \rightarrow 1 - F^x(t) \quad \text{in probability}$$

and

$$\left(\frac{nh\phi_x(h)}{\sigma_\lambda^2(x,t)}\right)^{1/2} \left(\widehat{F}_D^x - \widehat{F}_N^x(t) - 1 + \mathbb{E}[\widehat{F}_N^x(t)]\right) = o_{\mathbb{P}}(1).$$

4. DISCUSSIONS AND APPLICATIONS

In this section we emphasize the potential impact of our work by studying its practical interest in some important statistical problems. Moreover, in order to show the easy implementation of our approach on a concrete case, we discuss in the second part of this section the practical utilization of our model in risk analysis.

4.1. Some derivatives

- *On the choices of the bandwidths parameters.* As all smoothing by a kernel method, the choice of bandwidths parameters has crucial role in determining the performance of the estimators. The mean quadratic error given in Theorem (3.1) is a basic ingredient to solve this problem. Usually, the ideal theoretical choices are obtained by minimizing this error. Here, we have explained its leading term which is

$$B_n^2(x, t) + \frac{\sigma_\lambda^2(x, t)}{nh\phi_x(h)}.$$

Then, the smoothing parameters minimizing this leading term is asymptotically optimal with respect to the L^2 -error. However, the practical utilization of this criterium requires some additional computational efforts. More precisely, it requires the estimation of the unknown quantities Ψ'_0 , Φ'_0 , $f^x(t)$ and $F^x(t)$. Clearly, all these estimations can be obtained by using pilot estimators of the conditional distribution function $F^x(t)$ and of the conditional density $f^x(t)$. Such estimations are possible by using the kernel methods, with a separate choice of the bandwidths parameters between both models. More preciously, for the conditional density, we propose to adopt, to the functional case, the bandwidths selectors studied by Bouraine *et al.* [6] by considering the following criterion

$$(4.1) \quad \text{CVPDF} = \frac{1}{n} \sum_{i=1}^n W_1(X_i) \int_{\mathbb{R}} \widehat{f}^{X_i^{-i}}(y) W_2(y) dy - \frac{2}{n} \sum_{i=1}^n \widehat{f}^{X_i^{-i}}(Y_i) W_1(X_i) W_2(Y_i)$$

while, for the conditional distribution function we can use the cross-validation rule proposed by De Gooijer and Gannoun (2000) (in vectorial case)

$$\text{CVCDF} = \frac{1}{n} \sum_{k,l \in I_n} \left[\mathbf{1}_{\{Y_k \leq Y_l\}} - \widehat{F}^{X_k^{-k}}(Y_l) \right]^2 W(X_k),$$

where W_1 , W_2 and W are some suitable trimming functions and

$$\widehat{F}^{X_k^{-k}}(Y_i) = \frac{\sum_{i \in I_{n, \varsigma_n}^{k, l}} K(h^{-1} \|X_k - X_i\|) \mathbf{1}_{\{Y_i \leq Y_i\}}}{\sum_{i \in I_{n, \varsigma_n}^{k, l}} K(h^{-1} \|X_k - X_i\|)}$$

and

$$\widehat{f}^{X_i^{-i}}(t) = \frac{h^{-1} \sum_{j \in I_{n, \varsigma_n}^i} K(h^{-1} \|X_i - X_j\|) H(h^{-1}(t - Y_j))}{\sum_{j \in I_{n, \varsigma_n}^i} K(h^{-1} \|X_i - X_j\|)}$$

with

$$I_{n, \varsigma_n}^{k, l} = \{i \text{ such that } |i - k| \geq \varsigma_n \text{ and } |i - l| \geq \varsigma_n\} \text{ and}$$

$$I_{n, \varsigma_n}^i = \{j \text{ such that } |j - i| \geq \varsigma_n\}.$$

Of course, we can also adopt another selection methods, such that, the parametric bootstrap method, proposed by Hall *et al.* (1999) and Hyndman *et al.* (1996) for, respectively, the conditional cumulative distribution function and the conditional density in the finite dimensional case. Nevertheless, a data-driven method that allows to overcome this additional computation is very important in practice and is one of the natural prospects of the present work.

• *Confidence intervals.* The main application of Theorem 3.2 is to build confidence band for the true value of $\lambda^x(t)$. Similarly to the previous application, the practical utilization of our result in this topic requires the estimation of the quantity $\sigma_\lambda^2(x, t)$. A plug-in estimate for the asymptotic standard deviation $\sigma_\lambda^2(x, t)$ can be obtained by using the estimators $\widehat{f}^x(t)$ and $\widehat{F}^x(t)$ of $f^x(t)$ and $F^x(t)$. Then we get

$$\widehat{\sigma}_\lambda^2(x, t) := \frac{\widehat{\zeta}_2 \widehat{f}^x(t)}{\left(\widehat{\zeta}_1^2 (1 - \widehat{F}^x(t))^2\right)}$$

where

$$\widehat{\zeta}_1 = \frac{1}{n\phi_x(h)} \sum_{i=1}^n K(h^{-1} \|x - X_i\|) \text{ and } \widehat{\zeta}_2 = \frac{1}{n\phi_x(h)} \sum_{i=1}^n K^2(h^{-1} \|x - X_i\|).$$

Clearly, the function $\phi_x(\cdot)$ does not appear in the calculation of the confidence interval by simplification. More precisely, we obtain the following approximate $(1 - \zeta)$ confidence band for $\lambda^x(t)$

$$\widehat{\lambda}^x(t) \pm t_{1-\zeta/2} \times \left(\frac{\widehat{\sigma}_\lambda^2(x, t)}{nh\phi_x(h)} \right)^{1/2},$$

where $t_{1-\zeta/2}$ denotes the $1 - \zeta/2$ quantile of the standard normal distribution.

5. APPENDIX

In the following, we will denote $\forall i$

$$K_i(x) = K(h_H^{-1}\|x - X_i\|) \quad \text{and} \quad H_i(t) = H(h^{-1}(t - Y_i)).$$

Proof of Lemma 3.1. Firstly, for $\mathbb{E}[\widehat{f}_N^x(t)]$, we start by writing

$$\mathbb{E}[\widehat{f}_N^x(t)] = \frac{1}{\mathbb{E}[K_1(x)]} \mathbb{E} [K_1(x)\mathbb{E}[h^{-1}H_1(t)|X]]$$

with $h^{-1}\mathbb{E} [H_1(t)|X] = \int_{\mathbb{R}} H(z)f^X(t - hz)dt$.

The latter can be re-written, by using a Taylor expansion under (H3), as follows

$$h^{-1}\mathbb{E}[H_1(t)|X] = f^X(t) + \frac{h^2}{2} \left(\int_{\mathbb{R}} z^2 H(z)dz \right) \frac{\partial^2 f^X(t)}{\partial t^2} + o(h^2).$$

Thus, we get

$$\mathbb{E} \left[\widehat{f}_N^x(t) \right] = \frac{1}{\mathbb{E}[K_1(x)]} \left(\mathbb{E} [K_1(x)f^X(t)] + \left(\int_{\mathbb{R}} z^2 H(z)dz \right) \mathbb{E} \left[K_1(x) \frac{\partial^2 f^X(t)}{\partial t^2} \right] + o(h^2) \right).$$

Let $\psi_l(\cdot, t) := \frac{\partial^l f(\cdot, t)}{\partial t^l}$: for $l \in \{0, 2\}$, since $\Phi_l(0) = 0$, we have

$$\begin{aligned} \mathbb{E} [K_1(x)\psi_l(X, t)] &= \psi_l(x, t)\mathbb{E}[K_1(x)] + \mathbb{E} [K_1(x) (\psi_l(X, t) - \psi_l(x, t))] \\ &= \psi_l(x, t)\mathbb{E}[K_1(x)] + \mathbb{E} [K_1(x) (\Phi_l(\|x - X\|))] \\ &= \psi_l(x, t)\mathbb{E}[K_1(x)] + \Phi'_l(0)\mathbb{E} [\|x - X\|K_1(x)] \\ &\quad + o(\mathbb{E} [\|x - X\|K_1(x)]). \end{aligned}$$

So,

$$\begin{aligned} \mathbb{E} \left[\widehat{f}_N^x(t) \right] &= f^x(t) + \frac{h^2}{2} \frac{\partial^2 f^x(t)}{\partial t^2} \int_{\mathbb{R}} z^2 H(z)dz + o \left(h^2 \frac{\mathbb{E} [\|x - X\|K_1(x)]}{\mathbb{E}[K_1(x)]} \right) \\ &\quad + \Phi'_0(0) \frac{\mathbb{E} [\|x - X\|K_1(x)]}{\mathbb{E}[K_1(x)]} + o \left(\frac{\mathbb{E} [\|x - X\|K_1(x)]}{\mathbb{E}[K_1(x)]} \right). \end{aligned}$$

Similarly to Ferraty *et al.* (2007), we show that

$$\frac{1}{\phi_x(h)} \mathbb{E} [\|x - X\|K_1(x)] = h \left(K(1) - \int_0^1 (sK(s))' \beta_x(s)ds + o(1) \right)$$

and

$$\frac{1}{\phi_x(h)} \mathbb{E} [K_1(x)] = K(1) - \int_0^1 K'(s)\beta_x(s)ds + o(1).$$

Hence,

$$\begin{aligned} \mathbb{E} \left[\widehat{f}_N^x(t) \right] &= f^x(t) + \frac{h^2}{2} \frac{\partial^2 f^x(t)}{\partial t^2} \int_{\mathbb{R}} z^2 H(z) dz \\ &\quad + h \Phi'_0(0) \frac{\left(K(1) - \int_0^1 (sK(s))' \beta_x(s) ds \right)}{\left(K(1) - \int_0^1 K'(s) \beta_x(s) ds \right)} + o(h^2). \end{aligned}$$

Secondly, concerning $\mathbb{E}[\widehat{F}_N^x(t)]$. First note that $K_1(x) = K_1(x)\mathbf{1}_{B(x,h)}(X_1)$, under (H5). We deduce from (2.2) and (H0)

$$\mathbb{E}[\widehat{F}_N^x(t)] = \frac{1}{\mathbb{E}[K_1(x)]} \mathbb{E} \left(K_1(x) \mathbf{1}_{B(x,h)}(X_1) F^{X_1}(t) \right)$$

we write by an integration by part

$$\mathbb{E}[\widehat{F}_N^x(t)] = \frac{1}{\mathbb{E}[K_1(x)]} \mathbb{E} \left[K_1(x) \mathbb{E}[\mathbf{1}_{B(x,h)}(X_1) F^{X_1}(t) | X_1] \right]$$

with

$$\mathbb{E} \left[\mathbf{1}_{B(x,h)}(X_1) F^{X_1}(t) | X_1 \right] = \int_{\mathbb{R}} H(z) F^{X_1}(t - hz) dz.$$

The same steps used to studying $\mathbb{E}[\widehat{f}_N^x(t)]$ can be followed to prove that

$$\begin{aligned} \mathbb{E} \left[\widehat{f}_N^x(t) \right] &= F^x(t) + \frac{h^2}{2} \frac{\partial^2 F^x(t)}{\partial t^2} \int_{\mathbb{R}} z^2 H(z) dz \\ &\quad + h \Psi'_0(0) \frac{\left(K(1) - \int_0^1 (sK(s))' \beta_x(s) ds \right)}{\left(K(1) - \int_0^1 K'(s) \beta_x(s) ds \right)} + o(h^2) \quad \square. \end{aligned}$$

Proof of Lemma 3.2. For the first quantity $\text{Var}[\widehat{f}_N^x(t)]$, we have

$$s_n^2 = \text{Var}[\widehat{f}_N^x(t)] = \frac{1}{(nh \mathbb{E}[K_1(x)])^2} \text{Var} \left[\sum_{i=1}^n \Gamma_i(x) \right]$$

where

$$\Gamma_i(x) = K_i(x) H_i(t) - \mathbb{E}[K_i(x) H_i(t)].$$

Thus

$$\begin{aligned} \text{Var}[\widehat{f}_N^x(t)] &= \frac{1}{(nh \mathbb{E}[K_1(x)])^2} \underbrace{\sum_{i \neq j} \text{Cov}(\Gamma_i(x), \Gamma_j(x))}_{s_n^{\text{Cov}}} + \underbrace{\sum_{i=1}^n \text{Var}(\Gamma_i(x))}_{s_n^{\text{var}}} \\ &= \frac{1}{n(h \mathbb{E}[K_1(x)])^2} \text{Var}[\Gamma_1(x)] + \frac{1}{(nh \mathbb{E}[K_1(x)])^2} \sum_{i \neq j} \text{Cov}(\Gamma_i(x), \Gamma_j(x)). \end{aligned}$$

Let us calculate the quantity $\text{Var} [\Gamma_1(x)]$. We have:

$$\begin{aligned} \text{Var} [\Gamma_1(x)] &= \mathbb{E} [K_1^2(x)H_1^2(t)] - (\mathbb{E} [K_1(x)H_1(t)])^2 \\ &= \mathbb{E} [K_1^2(x)] \frac{\mathbb{E} [K_1^2(x)H_1^2(t)]}{\mathbb{E} [K_1^2(x)]} - (\mathbb{E} [K_1(x)])^2 \left(\frac{\mathbb{E} [K_1(x)H_1(t)]}{\mathbb{E} [K_1(x)]} \right)^2. \end{aligned}$$

So, by using the same arguments as those used in previous lemma we get

$$\begin{aligned} \frac{1}{\phi_x(h_K)} \mathbb{E} [K_1^2(x)] &= K^2(1) - \int_0^1 (K^2(s))' \beta_x(s) ds + o(1) \\ \frac{\mathbb{E} [K_1^2(x)H_1^2(t)]}{\mathbb{E} [K_1^2(x)]} &= h_H f^x(t) \int_{\mathbb{R}} H^2(z) dz + o(h) \\ \frac{\mathbb{E} [K_1(x)H_1(t)]}{\mathbb{E} [K_1(x)]} &= h f^x(t) + o(h) \end{aligned}$$

which implies that

$$\begin{aligned} (5.1) \quad \text{Var} [\Gamma_i(x)] &= h\phi_x(h) f^x(t) \int_{\mathbb{R}} H^2(z) dz \left(K^2(1) - \int_0^1 (K^2(s))' \beta_x(s) ds \right) + o(h\phi_x(h)). \end{aligned}$$

Now, let us focus on the covariance term. To do that, we need to calculate the asymptotic behavior of quantity defined as

$$\sum_{i \neq j} \left| \text{Cov}(\Gamma_i(x), \Gamma_j(x)) \right| = \sum_{1 \leq |i-j| \leq c_n} \left| \text{Cov}(\Gamma_i(x), \Gamma_j(x)) \right| = J_{1,n} + J_{2,n}$$

with $c_n \rightarrow \infty$, as $n \rightarrow \infty$.

For all (i, j) we write

$$\text{Cov} (\Gamma_i(x), \Gamma_j(x)) = \mathbb{E} [K_i(x)K_j(x)H_i(t)H_j(t)] - (\mathbb{E} [K_i(x)H_i(t)])^2$$

and we use the fact that

$$\begin{cases} \mathbb{E} [H_i(t)H_j(t)|(X_i, X_j)] = \mathcal{O}(h^2); & \forall i \neq j \\ \mathbb{E} [H_i(t)|X_i] = \mathcal{O}(h) & \forall i. \end{cases}$$

For $J_{1,n}$: by means of the integral realized above and under (H2) and (H5), we get

$$\mathbb{E} [K_i(x)K_j(x)H_i(t)H_j(t)] \leq Ch^2 \mathbb{P} [(X_i, X_j) \in B(x, h) \times B(x, h)]$$

and

$$\mathbb{E} [K_i(x)H_i(t)] \leq Ch \mathbb{P} (X_i \in B(x, h)).$$

It follows that, the hypotheses (H0), (H2) and (H5), imply that

$$\text{Cov} (\Gamma_i(x), \Gamma_j(x)) \leq C'h^2 \phi_x(h) \left(\phi_x(h) + \left(\frac{\phi_x(h)}{n} \right)^{1/a} \right)$$

So

$$J_{1,n} \leq C' \left(nc_n h^2 \left(\frac{\phi_x(h)}{n} \right)^{1/a} \phi_x(h) \right).$$

Hence

$$J_{1,n} = \mathcal{O} \left(nc_n h^2 \left(\frac{\phi_x(h)}{n} \right)^{1/a} \phi_x(h) \right).$$

On the other hand, these covariances can be controlled by mean of the usual Davydov-Rios's covariance inequality for mixing processes (see Rio 2000, formula 1.12a). Together with (H1), this inequality leads to:

$$\forall i \neq j, |\text{Cov}(D_i(x), D_j(x))| \leq c |i - j|^{-a}.$$

By the fact, $\sum_{k \geq c_n+1} k^{-a} \leq \int_{c_n}^{\infty} t^{-a} dt = \frac{c_n^{-a+1}}{a-1}$, we get by applying (H1),

$$J_{2,n} \leq \sum_{|i-j| \geq c_n+1} |i-j|^{-a} \leq \frac{nc_n^{-a+1}}{a-1}.$$

Thus, by using the following classical technique (see Bosq, 1998 [5]), we can write

$$s_n^{cov} = \sum_{0 < |i-j| \leq c_n} |\text{Cov}(\Gamma_i(x), \Gamma_j(x))| + \sum_{|i-j| > c_n} |\text{Cov}(\Gamma_i(x), \Gamma_j(x))|.$$

Thus

$$s_n^{cov} \leq C'n \left(c_n h^2 \left(\frac{\phi_x(h)}{n} \right)^{1/a} \phi_x(h) + \frac{c_n^{-a+1}}{a-1} \right).$$

Choosing $c_n = h^{-2} \left(\frac{\phi_x(h)}{n} \right)^{-1/a}$, and owing to the right inequality in (H4b), we can deduce

$$(5.2) \quad s_n^{cov} = o(nh\phi_x(h)).$$

Finally,

$$\begin{aligned} s_n^2 &= o(nh\phi_x(h)) + \mathcal{O}(nh\phi_x(h)) \\ &= \mathcal{O}(nh\phi_x(h)). \end{aligned}$$

In conclusion, we have

$$(5.3) \quad \text{Var}[\widehat{f}_N^x(t)] = \frac{f^x(t)}{nh\phi_x(h)} \left(\int_{\mathbb{R}} H^2(z) dz \right) \left(\frac{\left(K^2(1) - \int_0^1 (K^2(s))' \beta_x(s) ds \right)}{\left(K(1) - \int_0^1 K'(s) \beta_x(s) ds \right)^2} \right) + o\left(\frac{1}{nh\phi_x(h)} \right)$$

Now, for $\widehat{F}_N^x(t)$, (resp. \widehat{F}_D^x) we replace $H_i(t)$ by $\mathbf{1}_{\{Y_i \leq t\}}$ (resp. by 1) and we follow the same ideas,

$$\text{Var}[\widehat{F}_N^x(t)] = \frac{F^x(t)}{n\phi_x(h)} \left(\int_{\mathbb{R}} H^2(z) dz \right) \left(\frac{\left(K^2(1) - \int_0^1 (K^2(s))' \beta_x(s) ds \right)}{\left(K(1) - \int_0^1 K'(s) \beta_x(s) ds \right)^2} \right) + o\left(\frac{1}{n\phi_x(h)}\right).$$

and

$$\text{Var}[\widehat{F}_D^x] = \frac{1}{n\phi_x(h)} \left(\frac{\left(K^2(1) - \int_0^1 (K^2(s))' \beta_x(s) ds \right)}{\left(K(1) - \int_0^1 K'(s) \beta_x(s) ds \right)^2} \right) + o\left(\frac{1}{n\phi_x(h)}\right).$$

This yields the proof. \square

Proof of Lemma 3.3. The proof of this lemma follows the same steps as the previous Lemma. For this, we keep the same notation and we write

$$\begin{aligned} \text{Cov}(\widehat{f}_N^x(t), \widehat{F}_N^x(t)) &= \frac{1}{nh(\mathbb{E}[K_1(x)])^2} \text{Cov}(\Gamma_1(x), \Delta_1(x)) \\ &\quad + \frac{1}{n^2 h(\mathbb{E}[K_1(x)])^2} \sum_{i \neq j} \text{Cov}(\Gamma_i(x), \Delta_j(x)) \end{aligned}$$

where

$$\Delta_i(x) = K_i(x) \mathbf{1}_{\{Y_i \leq t\}} - \mathbb{E}[K_i(x) \mathbf{1}_{\{Y_i \leq t\}}].$$

For the first term, we have under (H0) and (H5)

$$\begin{aligned} \text{Cov}(\Gamma_1(x), \Delta_1(x)) &= \mathbb{E}[K_1^2(x) \mathbf{1}_{\{Y_1 \leq t\}} H_1(t)] - \mathbb{E}[K_1(x) \mathbf{1}_{\{Y_1 \leq t\}}] \mathbb{E}[K_1(x) H_1(t)] \\ &= \mathcal{O}(h\phi_x(h)) + \mathcal{O}(h\phi_x^2(h)) \\ &= \mathcal{O}(h\phi_x(h)). \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{1}{nh(\mathbb{E}[K_1(x)])^2} \text{Cov}(\Gamma_1(x), \Delta_1(x)) &= \mathcal{O}\left(\frac{1}{n\phi_x(h)}\right) \\ (5.4) \qquad \qquad \qquad &= o\left(\frac{1}{nh\phi_x(h)}\right). \end{aligned}$$

So, by using similar arguments as those invoked in the proof of Lemma 3.2, we use once again the boundedness of K and H , and the fact that (H1) and (H6) imply that

$$\mathbb{E}(H_i(t)|X_i) = \mathcal{O}(h).$$

Moreover, the right part of (H4b) implies that

$$\text{Cov}(\Gamma_i(x), \Delta_j(x)) = \mathcal{O}\left(h\phi_x(h) \left(\frac{\phi_x(h)}{n}\right)^{1/a} + \phi_x(h)\right),$$

Meanwhile, using the Davydov-Rio's inequality in Rio (2000) for mixing processes leads to

$$|\text{Cov}(\Gamma_i(x), \Delta_j(x))| \leq C\alpha(|i-j|) \leq C|i-j|^{-a},$$

we deduce easily that for any $c_n > 0$:

$$\begin{aligned} \sum_{i \neq j} \text{Cov}(\Gamma_i(x), \Delta_j(x)) &= \mathcal{O}\left(n c_n h \phi_x(h) \left(\frac{\phi_x(h)}{n}\right)^{1/a} + \phi_x(h)\right) \\ &+ \mathcal{O}(n h c_n^{-a}). \end{aligned}$$

It suffices now to take $c_n = h^{-1} \left(\frac{\phi_x(h)}{n}\right)^{-1/a}$ to get the following expression for the sum of the covariances:

$$(5.5) \quad \sum_{i \neq j} \text{Cov}(\Gamma_i(x), \Delta_j(x)) = o(n\phi_x(h)).$$

From (5.4) and (5.5) we deduce that

$$\text{Cov}(\widehat{f}_N^x(t), \widehat{F}_N^x(t)) = o\left(\frac{1}{nh\phi_x(h)}\right)$$

The same arguments can be used to show that

$$\text{Cov}(\widehat{f}_N^x(t), \widehat{F}_D^x) = o\left(\frac{1}{nh\phi_x(h)}\right) \quad \text{and} \quad \text{Cov}(\widehat{F}_N^x(t), \widehat{F}_D^x) = o\left(\frac{1}{n\phi_x(h)}\right). \quad \square$$

Proof of Lemma 3.4. Let

$$S_n = \sum_{i=1}^n \Lambda_i(x)$$

where

$$(5.6) \quad \Lambda_i(x) := \frac{\sqrt{h\phi_x(h)}}{h\mathbb{E}[K_1(x)]} \Gamma_i(x).$$

Obviously, we have

$$\sqrt{nh\phi_x(h)} [\sigma_f(x, t)]^{-1} \left(\widehat{f}_N^x(t) - \mathbb{E}\widehat{f}_N^x(t)\right) = (n(\sigma_f(x, t))^2)^{-1/2} S_n.$$

Thus, the asymptotic normality of $(n(\sigma_f(x, t))^2)^{-1/2} S_n$, is sufficient to show the proof of this Lemma. This last is shown by the blocking method, where the random variables Λ_i are grouped into blocks of defined different sizes.

We consider the classical big- and small-block decomposition. We split the set $\{1, 2, \dots, n\}$ into $2k_n + 1$ subsets with large blocks of size u_n and small blocks of size v_n and put

$$k_n := \left[\frac{n}{u_n + v_n} \right],$$

where $[y]$ stands for the integer part of the real number y .

Assumption (H4')(ii) allows us to define the large block size by

$$u_n =: \left[\left(\frac{nh\phi_x(h)}{q_n} \right)^{1/2} \right].$$

Using Assumption (H4') and simple algebra allows us to prove that

$$(5.7) \quad \frac{v_n}{u_n} \rightarrow 0, \quad \frac{u_n}{n} \rightarrow 0, \quad \frac{u_n}{\sqrt{nh\phi_x(h)}} \rightarrow 0, \quad \text{and} \quad \frac{n}{u_n} \alpha(v_n) \rightarrow 0.$$

Now, let Υ_j , Υ'_j and Υ''_j be defined as follows:

$$\begin{aligned} \Upsilon_j &= \sum_{i=j(u_n+v_n)+1}^{j(u_n+v_n)+u_n} \Lambda_i(x), \quad 0 \leq j \leq k_n + 1 \\ \Upsilon'_j &= \sum_{i=j(u_n+v_n)+u_n+1}^{(j+1)(u_n+v_n)+u_n} \Lambda_i(x), \quad 0 \leq j \leq k_n + 1 \\ \Upsilon''_j &= \sum_{i=k_n(u_n+v_n)+1}^n \Lambda_i(x), \quad 0 \leq j \leq k_n + 1. \end{aligned}$$

Clearly, we can write

$$S_n = \sum_{j=0}^{k_n-1} \Upsilon_j + \sum_{j=0}^{k_n-1} \Upsilon'_j + \Upsilon''_{k_n} r =: S'_n + S''_n + S'''_n.$$

We prove that

$$(5.8) \quad (i) \frac{1}{n} \mathbb{E}(S''_n)^2 \longrightarrow 0, \quad (ii) \frac{1}{n} \mathbb{E}(S'''_n)^2 \longrightarrow 0,$$

$$(5.9) \quad \left| \mathbb{E} \left\{ \exp \left(iwn^{-1/2} S'_n \right) \right\} - \prod_{j=0}^{k_n-1} \mathbb{E} \left\{ \exp \left(iwn^{-1/2} \Upsilon_j \right) \right\} \right| \longrightarrow 0,$$

$$(5.10) \quad \frac{1}{n} \sum_{j=0}^{k_n-1} \mathbb{E}(\Upsilon_j^2) \longrightarrow \sigma_f^2(x, t),$$

$$(5.11) \quad \frac{1}{n} \sum_{j=0}^{k_n-1} \mathbb{E} \left(\Upsilon_j^2 \mathbf{1}_{\{|\Upsilon_j| > \varepsilon \sqrt{n\sigma_j^2(x,t)}\}} \right) \rightarrow 0$$

for every $\varepsilon > 0$.

Expression (5.8) shows that the terms S_n'' and S_n''' are negligible, while Equations (5.9) and (5.10) show that the Υ_j are asymptotically independent, verifying that the sum of their variances tends to $\sigma_j^2(x, t)$. Expression (5.11) is the Lindeberg-Feller's condition for a sum of independent terms. Asymptotic normality of S_n is a consequence of Equations (5.8)–(5.11).

- *Proof of (5.8).* Because $\mathbb{E}(\Lambda_j) = 0$, $\forall j$, we have that

$$\begin{aligned} \mathbb{E}(S_n'')^2 &= \text{Var} \left(\sum_{j=0}^{k_n-1} \Upsilon_j' \right) = \sum_{j=0}^{k_n-1} \text{Var} (\Upsilon_j') \\ &\quad + \sum_{0 \leq i < j \leq k_n-1} \text{Cov} (\Upsilon_i', \Upsilon_j') := \Pi_1 + \Pi_2. \end{aligned}$$

By the second-order stationarity we get

$$\begin{aligned} \text{Var} (\Upsilon_j') &= \text{Var} \left(\sum_{i=j(u_n+v_n)+u_n+1}^{(j+1)(u_n+v_n)} \Lambda_i(x) \right) \\ &= v_n \text{Var}(\Lambda_1(x)) + \sum_{i \neq j}^{v_n} \text{Cov} (\Lambda_i(x), \Lambda_j(x)). \end{aligned}$$

Then

$$\begin{aligned} \frac{\Pi_1}{n} &= \frac{k_n v_n}{n} \text{Var}(\Lambda_1(x)) + \frac{1}{n} \sum_{j=0}^{k_n-1} \sum_{i \neq j}^{v_n} \text{Cov} (\Lambda_i(x), \Lambda_j(x)) \\ &\leq \frac{k_n v_n}{n} \left\{ \frac{\phi_x(h_K)}{h_H \mathbb{E}^2 K_1(x)} \text{Var} (\Gamma_1(x)) \right\} + \frac{1}{n} \sum_{i \neq j}^n \left| \text{Cov} (\Lambda_i(x), \Lambda_j(x)) \right| \\ &\leq \frac{k_n v_n}{n} \left\{ \frac{1}{h_H \phi_x(h_K)} \text{Var} (\Lambda_1(x)) \right\} + \frac{1}{n} \sum_{i \neq j}^n \left| \text{Cov} (\Lambda_i(x), \Lambda_j(x)) \right| \end{aligned}$$

Simple algebra gives us

$$\frac{k_n v_n}{n} \cong \left(\frac{n}{u_n + v_n} \right) \frac{v_n}{n} \cong \frac{v_n}{u_n + v_n} \cong \frac{v_n}{u_n} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Using Equation (5.2) we have

$$(5.12) \quad \lim_{n \rightarrow \infty} \frac{\Pi_1}{n} = 0.$$

Now, let us turn to Π_2/n . We have

$$\begin{aligned} \frac{\Pi_2}{n} &= \frac{1}{n} \sum_{i=0}^{k_n-1} \sum_{i \neq j}^{k_n-1} \text{Cov}(\Upsilon_i(x), \Upsilon_j(x)) \\ &= \frac{1}{n} \sum_{i=0}^{k_n-1} \sum_{j=0}^{k_n-1} \sum_{l_1=1}^{v_n} \sum_{l_2=1}^{v_n} \text{Cov}(\Lambda_{m_j+l_1}, \Lambda_{m_j+l_2}), \end{aligned}$$

with $m_i = i(u_n + v_n) + v_n$. As $i \neq j$, we have $|m_i - m_j + l_1 - l_2| \geq u_n$. It follows that

$$\frac{\Pi_2}{n} \leq \frac{1}{n} \sum_{i=1}^n \sum_{|i-j| \geq u_n}^n \text{Cov}(\Lambda_i(x), \Lambda_j(x)),$$

then

$$(5.13) \quad \lim_{n \rightarrow \infty} \frac{\Pi_2}{n} = 0.$$

By Equations (5.12) and (5.13) we get Part(i) of the Equation(5.8).

We turn to (ii), we have

$$\begin{aligned} \frac{1}{n} \mathbb{E} (S_n''')^2 &= \frac{1}{n} \text{Var} (\Upsilon_k'') \\ &= \frac{\vartheta_n}{n} \text{Var} (\Lambda_1(x)) + \frac{1}{n} \sum_{i=1}^{\vartheta_n} \sum_{i \neq j}^{\vartheta_n} \text{Cov} (\Lambda_i(x), \Lambda_j(x)) \end{aligned}$$

where $\vartheta_n = n - k_n(u_n + v_n)$; by the definition of k_n , we have $\vartheta_n \leq u_n + v_n$.

Then

$$\frac{1}{n} \mathbb{E} (S_n''')^2 \leq \frac{u_n + v_n}{n} \text{Var} (\Lambda_1(x)) + \frac{1}{n} \sum_{i=1}^{\vartheta_n} \sum_{i \neq j}^{\vartheta_n} \text{Cov} (\Lambda_i(x), \Lambda_j(x))$$

and by the definition of u_n and v_n we achieve the proof of (ii) of Equation (5.8).

• *Proof of (5.9).* We make use of Volkonskii and Rozanov’s lemma (see the appendix in Masry, 2005) and the fact that the process (X_i, X_j) is strong mixing.

Note that Υ_a is $\mathcal{F}_{i_a}^{j_a}$ -measurable with $i_a = a(u_n + v_n) + 1$ and $j_a = a(u_n + v_n) + u_n$; hence, with $V_j = \exp(i\omega n^{-1/2} \Upsilon_j)$ we have

$$\begin{aligned} \left| \mathbb{E} \left\{ \exp \left(i\omega n^{-1/2} S_n' \right) \right\} - \prod_{j=0}^{k_n-1} \mathbb{E} \left\{ \exp \left(i\omega n^{-1/2} \Upsilon_j \right) \right\} \right| \\ \leq 16k_n \alpha(v_n + 1) \cong \frac{n}{v_n} \alpha(v_n + 1) \end{aligned}$$

which goes to zero by the last part of Equation (5.7). Now we establish Equation 5.10).

• *Proof of (5.10).* Note that $\text{Var}(S'_n) \rightarrow \sigma_f^2(x, t)$ by Equation (5.8)-(i) and (5.8)-(ii) and since $\text{Var}(S_n) \rightarrow \sigma_f^2(x, t)$ (by the definition of the Λ_i and Equation (5.3)). Then because

$$\mathbb{E}(S'_n)^2 = \text{Var}(S'_n) = \sum_{j=0}^{k_n-1} \text{Var}(\Upsilon_j) + \sum_{i=0}^{k_n-1} \sum_{i \neq j}^{k_n-1} \text{Cov}(\Upsilon_i, \Upsilon_j),$$

all we have to prove is that the double sum of covariances in the last equation tends to zero. Using the same arguments as those previously used for Π_2 in the proof of the first term of Equation (5.8) we obtain by replacing v_n by u_n we get

$$\frac{1}{n} \sum_{j=0}^{k_n-1} \mathbb{E}(\Upsilon_j^2) = \frac{k_n u_n}{n} \text{Var}(\Lambda_1) + o(1).$$

As $\text{Var}(\Lambda_1) \rightarrow \sigma_f^2(x, t)$ and $k_n u_n / n \rightarrow 1$, we get the result.

Finally, we prove Equation (5.11).

• *Proof of (5.11).* Recall that

$$\Upsilon_j = \sum_{i=j(u_n+v_n)+1}^{j(u_n+v_n)+u_n} \Lambda_i.$$

Making use of Assumptions (H5), (H6) and (H5'), we have

$$|\Lambda_i| \leq C(h\phi_x(h))^{-1/2}$$

thus

$$|\Upsilon_j| \leq C u_n (h\phi_x(h))^{-1/2},$$

which goes to zero as n goes to infinity by Equation (5.7). Then for n large enough, the set $\{|\Upsilon_j| > \varepsilon (n\sigma_f^2(x, t))^{-1/2}\}$ becomes empty, this completes the proof and therefore that of the asymptotic normality of $(n(\sigma_f(x, t))^2)^{-1/2} S_n$. \square

Proof of Lemma 3.5. It is clear that, the result of Lemma (3.1) and Lemma (3.2) permits us

$$\mathbb{E} \left(\widehat{F}_D^x - \widehat{F}_N^x(t) - 1 + F^x(t) \right) \rightarrow 0$$

and

$$\text{Var} \left(\widehat{F}_D^x - \widehat{F}_N^x(t) - 1 + F^x(t) \right) \rightarrow 0$$

then

$$\widehat{F}_D^x - \widehat{F}_N^x(t) - 1 + F^x(t) \xrightarrow{\mathbb{P}} 0.$$

Moreover, the asymptotic variance of $\widehat{F}_D^x - \widehat{F}_N^x(t)$ given in Remark (3.5) allows us to obtain

$$\frac{nh\phi_x(h)}{\sigma_h(x,t)^2} \text{Var} \left(\widehat{F}_D^x - \widehat{F}_N^x(t) - 1 + \mathbb{E} \left(\widehat{F}_N^x(t) \right) \right) \longrightarrow 0.$$

By combining this result with the fact that

$$\mathbb{E} \left(\widehat{F}_D^x - \widehat{F}_N^x(t) - 1 + \mathbb{E} \left(\widehat{F}_N^x(t) \right) \right) = 0$$

we obtain the claimed result. \square

REFERENCES

- [1] I.A. Ahmad, *Uniform strong convergence of the generalized failure rate estimate*. Bull. Math. Statist. **17** (1976), 77–84.
- [2] K. Benhenni, F. Ferraty, M. Rachdi and P. Vieu, *Local smoothing regression with functional data*. Comput. Statist. **22** (2007), 353–369.
- [3] P. Besse, H. Cardot and D. Stephenson, *Autoregressive forecasting of some functional climatic variations*. Scand. J. Stat. **27** (2000), 673–687.
- [4] P. Besse and J.O. Ramsay, *Principal component analysis of sampled curves*. Psychometrika **51** (1986), 285–311.
- [5] D. Bosq, *Nonparametric Statistics for Stochastic Processes. Estimation and Prediction. (Second Edition)*. Lect. Notes Stat. **110**, Springer-Verlag, 1998.
- [6] M. Bouraine, A. Ait Saidi, F. Ferraty and P. Vieu, *Choix optimal de l'indice multifonctionnel: méthode de validation croisée*. Rev. Roumaine Math. Pures Appl. **50** (2010), 5, 355–367.
- [7] H. Cardot, F. Ferraty and P. Sarda, *Functional linear model*. Statist. Probab. Lett. **45** (1999), 11–22.
- [8] J. Damon and S. Guillas, *The inclusion of exogenous variables in functional autoregressive ozone forecasting*. Environmetrics **13** (2002), 759–774.
- [9] F. Ferraty, A. Mas and P. Vieu, *Advances in nonparametric regression for functional variables*. Aust. N. Z. J. Stat. **49** (2007), 1–20.
- [10] F. Ferraty, A. Rabhi and P. Vieu, *Conditional quantiles for functional dependent data with application to the climatic El Nino phenomenon*. Sankhya **67** (2005), 378–399.
- [11] F. Ferraty, A. Rabhi and P. Vieu, *Estimation non paramétrique de la fonction de hasard avec variable explicative fonctionnelle*. Rev. Roumaine Math. Pures Appl. **52** (2008), 1–18.
- [12] F. Ferraty and P. Vieu, *Non-Parametric Functional Data Analysis*. Springer-Verlag, NewYork, 2006.
- [13] T. Gasser, P. Hall and B. Presnell, *Nonparametric estimation of the mode of a distribution of random curves*. J. R. Stat. Soc. Ser. B. Stat. Methodol. **60** (1998), 681–691.
- [14] P. Hall, R.C.L. Wolef and Q. Yao, *Methods for estimating a conditional distribution function*. An. Statist. Assoc. **94** (1999), 154–163.
- [15] R.J. Hyndman, D.M. Bashtannyk and G.K. Grunwald, *Estimating and visualizing conditional densities*. J. Comput. Graph. Statist. **5** (1996), 315–336.
- [16] A. Mahiddine, A.A. Bouchentouf and A. Rabhi, *Nonparametric estimation of some characteristics of the conditional distribution in single functional index model*. Malaya J. Matematik (MJM) **2** (2014), 4, 392–410.

- [17] E. Masry, *Non-parametric regression estimation for dependent functional data: Asymptotic normality*. Stoch. Process. Appl. **115** (2005), 155–177.
- [18] A. Quintela, *Plug-in bandwidth selection in kernel hazard estimation from dependent data*. Comput. Statist. Data Anal. **51** (2007), 5800–5812.
- [19] A. Rabhi, S. Benaissa, E.H. Hamel and B. Mechab, *Mean square error of the estimator of the conditional hazard function*. Appl. Math. (Warsaw) **40** (2013), 405–420.
- [20] M. Rachdi and P. Vieu, *Non-parametric regression for functional data: Automatic smoothing parameter selection*. J. Statist. Plann. Inference **137** (2007), 2784–2801.
- [21] J.O. Ramsay and B.W. Silverman, *Functional Data Analysis*. Second Edition. Springer-Verlag, New York, 2005.
- [22] J. Rice and B.W. Silverman, *Estimating the mean and covariance structure non-parametrically when the data are curves*. J.R. Stat. Soc. Ser. B. Stat. Methodol. **53** (1991), 233–243.
- [23] E. Rio, *Théorie asymptotique des processus aléatoires dépendants*. Math. Appl. (Berlin) **31**, Springer, (2000).
- [24] G.G. Roussas, *Hazard rate estimation under dependence conditions*. J. Statist. Plann. Inference. **22** (1989), 81–93.
- [25] N.D. Singpurwalla and M.Y. Wong, *Estimation of the failure rate. A survey of non-parametric methods. Part I: Non-Bayesian methods*. Comm. Statist. Theory Methods **12** (1983), 559–588.
- [26] L. Spierdijk, *Non-parametric conditional hazard rate estimation: A local linear approach*. Comput. Statist. Data Anal. **52** (2008), 2419–2434.
- [27] L.T. Tran and S. Yakowitz, *Nearest neighbor estimators for rom fields*. J. Multivariate Anal. **44** (1993), 23–46.
- [28] G.S. Watson and M.R. Leadbetter, *Hazard analysis*. Sankhya **26** (1964), 101–116.

Received 3 December 2014

*University of Sidi Bel Abbes,
Laboratory of Mathematics,
Algeria
rabhi_abbes@yahoo.fr*

*University of Science and technology of
Oran,
Departement of Mathematics,
Algeria
hamza_naziha@yahoo.fr*

*University of Sidi Bel Abbes,
Laboratory of Mathematics,
Algeria
amina_bouchentouf@yahoo.com*