Asymptotic Normality of the Kernel Estimate of the Conditional Distribution Function for Quasi-associated data

H.DAOUDI
Department of Mathematics.
Ibn Khladoun University, Tiaret, Algeria.
Laboratory of Stochastic Process and Statistics.
daoudiham63@gmail.com

B.MECHAB
Department of Probability and Statistics.
Djillali Liabs University, Algeria.
Laboratory of Stochastic Process and Statistics.
mechaboub@yahoo.fr

Abstract

The main goal of this paper is to study the asymptotic normality of the estimate of the conditional distribution function of a scalar response variable Y given a hilbertian random variable X when the observations are quasi-associated. Our approach is based on the Doob's technique. It is shown that, under the concentration property on small balls of the probability measure of the functional estimator and some regularity conditions, the kernel estimate of the conditional distribution function is asymptotically normally distributed. We performed out simulation experiments to examine the behavior of this asymptotic property over finite sample data.

Keywords: Conditional distribution function; probabilities of small balls; asymptotic normality; nonparametric kernel estimation; quasi-associated data.

1. Introduction

In recent years, the functional estimate has attracted a lot of attention in the statistical literature. Functional data arise in a variety of fields including econometrics, epidemiology, environmental science and many others.

For nonparametric functional estimation, the book of Ferraty and Vieu (2006) gives an excellent synthesis of kernel method for conditional models, where many asymptotic properties of regression, conditional quantile and conditional density estimator have been obtained. The estimation of the conditional distribution function in a functional framework was introduced by Ferraty et al. (2006). They constructed a dual kernel estimator for the conditional distribution function and specified the almost complete convergence rate of this estimator when the observations are independent and identically distributed. The case of α-mixing observations was studied by Ferraty et al. (2007). An example of an application on conditional median prediction, as well as the determination of prediction intervals, was considered in this article. Several authors have treated the estimation of the conditional distribution function as a preliminary study of the estimation of conditional quantile. For example, Ezzahrioui and Ould-Saïd (2008) have studied the asymptotic normality of this estimator in both cases (i.i.d. and mixing).

The associated random variables play an important role in a wide variety of areas, including reliability theory, mathematical physics, multivariate statistical analysis, life sciences and in percolation theory. Many works were treated data under positive and negative dependant random variables, one can quote, Newman (1984) and Matula (1992). The concept of quasi-association is a special case of weak dependence introduced by Doukhan and Louhichi (1999) for real-valued stochastic processes. To the best of our knowledge, there is few papers dealing with the nonparametric estimation for quasi-associated random variables. We quote, Douge (2010) studied a limit theorem for quasi-associated hilbertian random variables, Attaoui and Ling (2016) studied asymptotic results of a nonparametric conditional cumulative distribution estimator in the single functional index modeling of time series data, Tabti and Ait Saidi (2018) studied the estimation and simulation of the conditional hazard function in the quasi-associated framework when the observations are linked via a functional single index structure, the asymptotic normality of this last estimator was studied by Daoudi et al. (2018). Mechab and Laksaci (2016) studied Nonparametric relative regression for associated random variables. Daoudi et al. (2019) studied the asymptotic normality of the nonparametric conditional density function estimate with functional variables for quasi-associated data

The main contribution of this work is the study of the asymptotic normality of the estimator of the conditional distribution function of Ferraty et al. (2006) in case of quasi-associated data. Note that, like all asymptotic statistics nonfunctional parametric, our result is related to the phenomenon of concentration of the probability measure of the explanatory variable and regularity of the functional space of the model. we recall the definition of quasi-association:

Definition 1.1. *A sequence $(X_n)_{n \in \mathbb{N}}$ of real random vectors variables is said to be Quasi-Association (QA), if for any disjoint subsets I and J of $\mathbb{N}$ and all bounded Lipschitz functions $f : \mathbb{R}^{|I|d} \to \mathbb{R}$ and $g : \mathbb{R}^{|J|d} \to \mathbb{R}$ satisfying*

$$Cov(f(X_i, i \in I), g(X_j, j \in J)) \leq Lip(f)Lip(g) \sum_{i \in I} \sum_{j \in J} \sum_{k=1}^{d} \sum_{l=1}^{d} \left| Cov(X_i^k, X_j^l) \right|$$

where X_i^k denotes the kth component of X_i,

$$Lip(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{||x - y||_1} \ \ with \ \ ||(x_1, ..., x_k)||_1 = |x_1| + \cdots + |x_k|.$$

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. An application on simulated data is given to validate our theoretical result in Section 5. The auxiliary results and proofs are given in Section 6. We finalize the paper with a conclusion in Section 7.

2. The model

Consider $Z_i = (X_i, Y_i)_{1 \leq i \leq n}$ be a n quasi-associated random identically distributed as the random $Z = (X, Y)$, with values in $\mathcal{H} \times \mathbb{R}$, where $\mathcal{H}$ is a separable real Hilbert

space with the norm $\| \, . \, \|$ generated by an inner product $< .,. >$.

We consider the semi-metric d defined by $\forall x, x' \in \mathcal{H}$, $d(x, x') = |< x - x', \alpha >|$ where $\alpha \in \mathcal{H}$. In the following x will be a fixed point in $\mathcal{H}$.

We intend to estimate the conditional distribution function $F^x(y)$ using n dependent observations $(X_i, Y_i)_{i \in \mathbb{N}}$ draw from a random variables with the same distribution with $Z := (X, Y)$. To this aim, we introduce the kernel type estimator $\widehat{F}^x$ of F^x defined by:

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

where K is the kernel, H is a given distribution function and $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers.

3. Notations and hypotheses

All along the paper, when no confusion will be possible, we will denote by C or/and C' some strictly positive generic constants whose values are allowed to change. We assume that the random pair $Z_i = \{(X_i, Y_i), i \in \mathbb{N}\}$ is stationary quasi-associated processes.

Let λ_k the covariance coefficient defined as:

$$\lambda_k = sup_{s \geq k} \sum_{|i-j| \geq s} \lambda_{i,j}$$

where:

$$\lambda_{i,j} = \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} | \, cov(X_i^k, X_j^l) \, | + \sum_{k=1}^{\infty} | \, cov(X_i^k, Y_j) \, | + \sum_{l=1}^{\infty} | \, cov(Y_i, X_j^l) \, | + cov(Y_i, Y_j) \, | \, .$$

X_i^k denotes the kth component of X_i defined as $X_i^k := < X_i, e^k >$.

For $h > 0$, let $B(x, h) := \{x' \in \mathcal{H}/d(x', x) < h\}$ be the ball of center x and radius h. To establish the asymptotic normality of the estimator $\widehat{F}^x$, we need to include the following assumptions:

(H1) $\mathbb{P}(d(x', x) < h_K) = \phi_x(h_K) > 0$ and Moreover, there exists a function $\beta(x, .)$ such that:

$$\forall s \in [0, 1], \lim_{h_K \to 0} \frac{\phi(x, sh_K)}{\phi(x, h_K)} = \beta(x, s).$$

.

(H2) For $l \in \{0, 2\}$, the functions $\Phi_l(s) = \mathbb{E}[\frac{\partial^l F^X(y)}{\partial y^l} - \frac{\partial^l F^x(y)}{\partial y^l} \mid d(x, X) = s]$ are differentiable at $s = 0$.

(H3) H is a cumulative distribution has derivative H' such that: $\int H'^{2/r}(t) dt < \infty$ and $\int | \, t \, |^{b_2} H'^2 dt < \infty$.

(**H4**) K is a kernel function and bounded continuous Lipschitz function such that:

$$C\mathbf{1}_{[0,1]}(.) < K(.) < C'\mathbf{1}_{[0,1]}(.)$$

where $\mathbf{1}_{[0,1]}$ is the indicator function on [0,1], and its derivative K' is such that:

$$-\infty < C < K'(t) < C' < 0 \text{ for } 0 \le t \le 1.$$

(**H5**) The bandwidths (h_K, h_H) satisfied:

(i) $\lim\limits_{n\to\infty} h_K = 0$, $\lim\limits_{n\to\infty} h_H = 0$ and (ii) $\lim\limits_{n\to\infty} (h_H^{b_2} + h_K^{b_1})\sqrt{n\phi_{(x,h_K)}} = 0.$

(**H6**) The sequence of random pairs $(X_i, X_j), i \in \mathbb{N}$ is quasi-associated with covariance coefficient λ_k, $k \in \mathbb{N}$ satisfying:

$$\exists \alpha > 0, \exists C > 0, \text{such that } \lambda_k \le Ce^{-\alpha k}.$$

(**H7**)

$$\sup_{i\neq j} \mathbb{P}\left[(X_i, X_j) \in B(x, h_K) \times B(x, h_K)\right] = \Psi_{i,j} > 0$$

$$\Psi_{i,j} = max_{i\neq j}\{\mathbb{P}(d(x, X_i)) < h_K, \mathbb{P}(d(x, X_j)) < h_K)\} = O(\phi_x^2(h_K)).$$

Comments on the hypotheses

Assumption (**H1**) is the concentration property of the explanatory variable in small balls. The function $\beta(x,.)$ plays a fundamental role in all asymptotic, in particular for the variance term. The condition (**H2**) is used to control the regularity of the functional space of our model and these are needed to evaluate the bias term of the convergence rates. The hypotheses (**H3**) and (**H4**) are technical conditions on the cumulative function H and the kernels K, H' and K'. Assumption (**H5**) is also classical in the functional estimation in finite or infinite dimension spaces, in particular, is used to eliminate the term bias in the result of asymptotic normality. The hypothesis (**H6**) is a structural condition used for the quasi-associated data. To establish the asymptotic normality of our model under quasi-association, we need the assumption (**H7**), which describes the asymptotic behavior of the joint distribution of the couple (X_i, X_j).

4. Main Results: Asymptotic Normality

Theorem 4.1. *Under hypotheses (H1)-(H7), as n goes to infinity, we have:*

$$\sqrt{\frac{n\phi_{(x,h_K)}}{\sigma^2(x)}} \left(\widehat{F}^x(y) - F^x(y)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad n \to \infty$$

where

$$\sigma^2(x) = \frac{C_2 F^x(y)(1 - F^x(y))}{C_1^2} \int H'^2(t)dt$$

with

$$C_j = K(1) - \int_0^1 (K^j)'(s)\beta(x,s)ds, \quad \text{for } j = 1, 2.$$

Proof of Theorem 4.1. The proof of this theorem is based on the following decomposition and the lemmas bellow:

$$
\begin{aligned}
\widehat{F}^x(y) - F^x(y) &= \frac{\widehat{F}_N^x(y) - F^x(y)\widehat{F}_D^x(y)}{\widehat{F}_D^x(y)} \\
&= \frac{1}{\widehat{F}_D^x(y)}\{\widehat{F}_N^x(y) - \mathbb{E}(\widehat{F}_N^x(y)) \\
&+ \mathbb{E}(\widehat{F}_N^x(y)) - F_N^x(y))\} \\
&- \frac{1}{\widehat{F}_D^x(y)}\left\{F^x(y)(\widehat{F}_D^x(y) - 1)\right\}
\end{aligned}
$$

where

$$\widehat{F}_N^x(y) = \frac{1}{n\mathbb{E}(K_1(x))}\sum_{i=1}^n K_i(x)H_i(y)$$

and

$$\widehat{F}_D^x = \frac{1}{n\mathbb{E}(K_1(x))}\sum_{i=1}^n K_i(x),$$

with

$$K_i(x) = K\left(h_K^{-1}d(x, X_i)\right) \quad \text{and} \quad H_i(y) = H\left(h_H^{-1}(y - Y_i)\right).$$

Finally, to state the asymptotic normality of $F^x(y)$, we show that the numerator suitably normalized is asymptotically normally distributed (with law $\mathcal{N}(0, \sigma^2(x))$) and that the denominator converges in probability to 1.

Then, the proof of Theorem 4.1 can be deduced from the following lemmas:

Lemma 4.1. *Under the hypotheses of Theorem 4.1, as n goes to infinity, we have:*

$$\sqrt{n\phi(x, h_K)}\left(\widehat{F}_N^x(y) - \mathbb{E}(\widehat{F}_N^x(y))\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x)).$$

Lemma 4.2. *(See, (Laksaci and Mechab, 2014)). Under the hypotheses (H1)-(H5), we have:*

$$\mathbb{E}(\widehat{F}_N^x(y)) - F_N^x(y) = B_H^F(x, y)h_H^2 + B_K^F(x, y)h_K + o(h_H^2) + o(h_K)$$

where

$$B_H^F(x, y) = \frac{1}{2}\frac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H'(t)dt$$

and

$$B_K^F(x, y) = h_K\Phi_0'(0)\frac{C_0}{C_1}.$$

Pak.j.stat.oper.res. Vol.XV No.IV 2019 pp999-1015

1003

Lemma 4.3. *Under the hypothesis (H1)-(H7), as n goes to infinity, we have:*

$$\sqrt{n\phi(x, h_K)} \left(F^x(y)(\widehat{F}_D^x(y) - 1) \right) \to 0 \quad in \quad probability.$$

Lemma 4.4. *Under hypotheses of Theorem 4.1, we have*

$$\sum_{n \in \mathbb{N}} \mathbb{P} \left(\widehat{F}_D^x(y) < 1/2 \right) < \infty.$$

5. Application on simulated data

We aim to evaluate, on a finite sample, performances of the asymptotic normality of the conditional distribution on simulated data. In particular, our main purpose is to show how we can implement easily and quickly this estimator in practice. Of course, the applicability of our asymptotic normality result requires a practical estimation of the asymptotic bias and variance.

The main purpose of this section is to test the effectiveness of the two asymptotic normality results. For this purpose, we consider the functional nonparametric model as follows:

$$Y = r(x) + \epsilon \qquad where \qquad \epsilon \sim \mathcal{N}(0, 1)$$

We are interested in functional data derived from a mixture of two Gaussian stochastic processes $Z_1(t)$ and $Z_2(t)$ over an interval [-1,1] defined by:

$$Z_1(t) = \sqrt{-2log(U)}cos(2\pi(1 - W)t), \quad Z_2(t) = \sqrt{-2log(1 - U)}sin(2\pi W t)$$

where U and W are random variables distributed uniformly over the interval [0,1]. The explanatory functional variables are quasi-associated are constructed by:

$$X(t) = Z_1(t) + Z_2(t)$$

We generate a sample of size 200 $\{X_i(t)\}_{i=1,...,200}$ of $X(t)$, and we observe each variable X_i on $(t_j)_{j=1,...,100} \in [-1, 1])$.

The curves obtained are plotted in Figure 1: On the other hand, for $i = 1, ..., n = 200$,

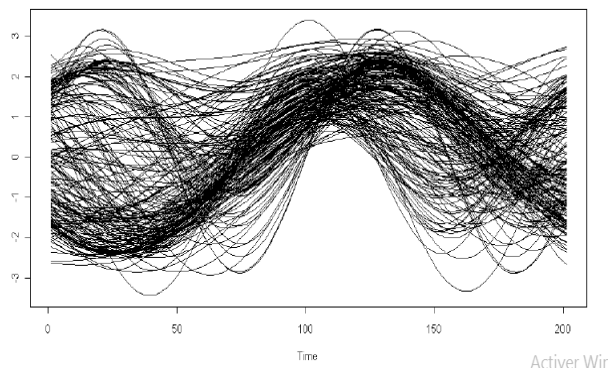Figure 1: A sample of 200 curves

1004

Pak.j.stat.oper.res. Vol.XV No.IV 2019 pp999-1015

the scalar response Y_i is computed by considering the following operator:

$$r(x) = \int_0^1 \frac{dt}{1+ \mid x(t) \mid}$$

Recall that, the conditional distribution of Y given $X = x$ corresponding to this model is explicitly given by the law of ϵ_i shifted by $r(x)$. Then, the corresponding conditional density $f^x(y)$ is:

$$f^x(y) = \frac{1}{2\pi} \exp(-\frac{1}{2}(y - r(x))^2$$

Elsewhere, as it is well-known in FDA, the choice of the metric and the smoothing parameters have crucial roles in the computational issues. To optimize these choices in this illustration, we use firstly the cross-validation procedure method for choosing smoothing parameters, secondly regarding the shape of the curves X_i , it is clear that the PCA-type semi-metric (see Benhenni et al. 2007), is well-adapted to this kind of data. Then, we point out that, we opted for a quadratic kernel which is supported within (0,1) and taken $K = H'$.
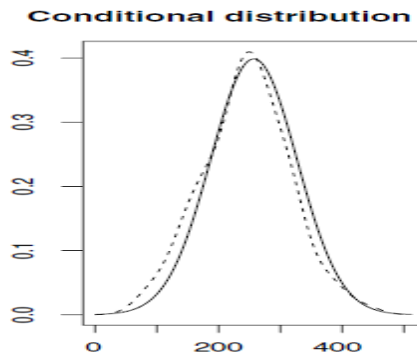

Figure 2: The asymptotic distribution of the conditional distribution function

The obtained results are shown in figure 2. It appears clearly that, asymptotic distributions have good behaviors with respect to the standard normal distribution. This conclusion is confirmed by the Kolmogorov-Smirnov test which, for $n = m = 200$, gives 0.80 as a P value for the first model and 0.69 for the second one.

6.   Auxiliary results and proofs

First of all, we state the following lemmas.

Lemma 6.1. *(See, Douge (2010)). Let $(X_n)_{n \in \mathbb{N}}$ be a quasi-associated sequence of random variables with values in $\mathcal{H}$. Let $f \in BL(\mathcal{H}^{|I|}) \cap \mathbb{L}^\infty$ and $g \in BL(\mathcal{H}^{|J|}) \cap \mathbb{L}^\infty$, for some finite disjoint subsets $I, J \subset \mathbb{N}$. Then*

$$Cov(f(X_i, i \in I), g(X_j, j \in J)) \leq Lip(f)Lip(g) \sum_{i \in I} \sum_{j \in J} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \left| Cov(X_i^k, X_j^l) \right|$$

where $(BL(\mathcal{H}^u; u > 0)$ is the set of bounded Lipschitz functions $f : \mathcal{H}^u \to \mathbb{R}$ and $\mathbb{L}^\infty$

is the set of bounded functions.

Lemma 6.2. *(See, Kallabis and Nemann (2006)). Let $X_1, ..., X_n$ the real random variables such that $\mathbb{E}(X_j) = 0$ and $\mathbb{P}(| X_j | \leq M) = 1$ for all $j = 1, ..., n$ and some $M < \infty$, Let $\sigma_n^2 = Var(\sum_{i=1}^{n} \Delta_i)$.*
Assume, furthermore, that there exist $K < \infty$ and $\beta > 0$ such that, for all u-uplets $(s_1, ..., s_u) \in \mathbb{N}^u$, $(t_1, ..., t_v) \in \mathbb{N}^v$ with $1 \leq s_1 \leq \cdots \leq s_u \leq t_1 \leq \cdots \leq t_v \leq n$.
The following inequality is fulfilled:

$$| cov(X_{s_1}...X_{s_u}, X_{t_1}...X_{t_v}) | \leq K^2 M^{u+v-2} v e^{-\beta(t_1 - s_u)}.$$

Then,

$$\mathbb{P}(| \sum_{j=1}^{n} X_j | > t) \leq \exp\{-\frac{t^2/2}{A_n + B_n^{1/3} t^{5/2}}\}$$

for some

$$A_n \leq \sigma_n^2$$

and

$$B_n = (\frac{16nK^2}{9An(1 - e^{-\beta}) \vee 1}) \frac{2(K \vee M)}{1 - e^{-\beta}}.$$

6.1 Proof of lemma 4.1

We denote

$$Z_{ni}(x, y) = \frac{\sqrt{\phi(x, h_K)}}{\sqrt{n}E(K_1)}(\Gamma_i(x, y) - \mathbb{E}\Gamma_i(x, y))$$

where

$$\Gamma_i(x, y) = K(h_K^{-1}d(x, X_i))H_i(y) - \mathbb{E}[K_1 H_1], 1 \leq i \leq n$$

and

$$S_n := \sum_{i=1}^{n} Z_{ni}(x, y) \tag{2}$$

Therefore,

$$S_n = \sqrt{n\phi(x, h_K)}(\widehat{F}_N(x, y) - \mathbb{E}(\widehat{F}_N(x, y)).$$

Thus, our claimed result is, now:

$$S_n \to \mathcal{N}(0, \sigma^2(x)).$$

To do that, we use the basic technique of Doob (1959). Indeed, we consider $p = p_n$ and $q = q_n$ two sequences of natural numbers tending to infinity, such that

$$p = O(\sqrt{n\phi(x, h_K)}), \qquad q = o(p)$$

and we split S_n into

$$S_n = T_n + T_n' + \xi_k \text{ with } T_n = \sum_{j=1}^{k} \eta_j \text{ and } T_n' = \sum_{j=1}^{k} \eta_j$$

1006

Pak.j.stat.oper.res. Vol.XV No.IV 2019 pp999-1015

where

$$\eta_j = \sum_{i \in I_J} Z_{ni}(x, y), \quad \xi_j = \sum_{i \in I_J} Z_{ni}(x, y), \quad \zeta_k = \sum_{i=k(p+q)+1} Z_{ni}(x, y)$$

with

$$I_j = (j-1)(p+q) + 1, ..., (j-1)(p+q) + p,$$
$$J_j = (j-1)(p+q) + p + 1, ..., j(p+q).$$

Observe that, for $k = \frac{n}{p+q}$, (where [.] stands for the integral part), we have $\frac{kq}{n} \to 0$ and $\frac{kp}{n} \to 1, \frac{q}{n} \to 0$, which imply that $\frac{p}{n} \to 0$ as $n \to \infty$.

Now, our asymptotic result is based on:

$$\mathbb{E}(T'_n)^2 + \mathbb{E}(\zeta_n)^2 \to 0 \tag{3}$$

and

$$T_n \to \mathcal{N}(0, 1). \tag{4}$$

Proof of (3).

By stationarity, we get

$$\mathbb{E}(T'_n)^2 = kVar(\zeta_1) + 2 \sum_{1 \le i < j \le k} | Cov(\zeta_i, \zeta_j) | \tag{5}$$

and

$$kVar(\zeta_1) \le qkVar(Z_{n1}(x, y)) + 2k \sum_{1 \le i < j \le k} Cov(Z_{ni}(x, y), Z_{nj}(x, y)) \tag{6}$$

by the fact that $\frac{kq}{n} \to 0$. We obtain

$$
\begin{aligned}
qkVar(Z_{n1}(x, y)) &= \phi(x, h_K)qk \frac{1}{n(\mathbb{E}(K_1))^2} Var(\Gamma_1(x, y)) \\
&= O(\frac{kq}{n}) \to 0, \text{ as } n \to \infty.
\end{aligned}
$$

On the other hand, we have

$$k \sum_{1 \le i < j \le k} | Cov(Z_{ni}(x, y), Z_{nj}(x, y)) | = \frac{k\phi(x, h_K)}{n(\mathbb{E}(K_1))^2} \sum_{1 \le i < j \le k} Cov(Z_{ni}(x, y), Z_{nj}(x, y))$$

we obtain,

$$\sum_{1 \le i < j \le k} | Cov(Z_{ni}(x, y), Z_{nj}(x, y)) | = o(q\phi(x, h_K)).$$

Then

$$k \sum_{1 \le i < j \le k} | Cov(Z_{ni}(x, y), Z_{nj}(x, y)) | = O(\frac{kq}{n}) \to 0, \text{ as } n \to \infty. \tag{7}$$

From (5)-(7), we obtain

$$kVar(\xi_i) \to 0, \text{ as } n \to \infty. \tag{8}$$

We use the stationarity, to evaluate the second term in the right-hand side of (4)

$$
\begin{aligned}
\sum_{1 \le i < j \le k} | Cov(\xi_{ni}, \xi_{nj}) | &= \sum_{1 \le i < j \le k} (k-l) | Cov(\xi_{ni}, \xi_{nj}) | \\
&\le k \sum_{1 \le i < j \le k} | Cov(\xi_{ni}, \xi_{nj}) | \\
&\le \sum_{l=1}^{k-1} \sum_{(i,j) \in J \times J_{l+1}} Cov(Z_{ni}(x,y), Z_{nj}(x,y))
\end{aligned}
$$

It is clear that, for all $(i,j) \in J_i \times J_j$, we have $|i - j| \ge p + 1 > p$, then

$$
\begin{aligned}
\sum_{1 \le i < j \le k} | Cov(\xi_{ni}, \xi_{nj}) | &\le k \frac{C\phi(x,h_K)(h_K^{-1}Lip(K) + h_H^{-1}Lip(H))^2}{n(\mathbb{E}[K_1)^2} \sum_{i=1}^{p} \sum_{j=2p+q+1, |i-j|>p}^{v} \lambda_{i,j} \\
&\le \frac{Ckp\phi(x,h_K)(h_K^{-1}Lip(K) + h_H^{-1}Lip(H))^2}{n(\mathbb{E}[K_1)^2} \lambda_p \\
&\le \frac{Ckp(h_K^{-1}Lip(K) + h_H^{-1}Lip(H))^2}{n\phi(x,h_K)} e^{-\alpha p} \\
&\le \frac{Ckp}{nh_H^2\phi^3(x,h_K)} e^{-\alpha p} \to 0.
\end{aligned}
$$

Finally, by combining this last result and (7). We can write

$$\mathbb{E}(T_1')^2 \to 0 \text{ as } n \to \infty. \tag{9}$$

Moreover

$$
\begin{aligned}
\mathbb{E}(\zeta_k)^2 &\le (n - k(p+q))Var(Z_{n1}(x,y)) + 2 \sum_{1 \le i < j \le k} | Cov(Z_{ni}(x,y), Z_{nj}(x,y)) | \\
&\le pVar(Z_{n1}(x,y)) + 2 \sum_{1 \le i < j \le k} | Cov(Z_{ni}(x,y), Z_{nj}(x,y)) | \\
&\le \frac{p\phi(x,h_K)}{n\mathbb{E}(K_1)^2} Var(Z_{n1}(x,y)) + \underbrace{\frac{C\phi(x,h_K)}{n\mathbb{E}(K_1)^2} \sum_{1 \le i < j \le k} | Cov(Z_{ni}(x,y), Z_{nj}(x,y)) |}_{o(1)} \\
&\le \frac{Cp}{n} + o(1).
\end{aligned}
$$

Then,

$$\mathbb{E}(\zeta_k)^2 \to 0 \text{ as } n \to \infty.$$

Which combining with (3) completes the proof of (4).
Proof of (6).

The proof of convergence in (6) is based in the following two results

$$| \mathbb{E}(e^{it \sum_{j=1}^{k} \eta_j}) - \prod_{j=1}^{k} \mathbb{E}(e^{it\eta_j}) | \to 0 \tag{10}$$

and

$$kVar(\eta_1) \to \sigma^2(x), \quad k\mathbb{E}(\eta_1^2 \mathbf{1}_{\eta_1 > \epsilon\sigma(x)}) \to 0. \tag{11}$$

Proof of (10).

$$
\begin{aligned}
| \mathbb{E}(e^{it \sum_{j=1}^{k} \eta_j}) - \prod_{j=1}^{k} \mathbb{E}(e^{it\eta_j}) | \quad &\leq \quad | \mathbb{E}(e^{it \sum_{j=1}^{k} \eta_j}) - \mathbb{E}(e^{it \sum_{j=1}^{k-1} \eta_j})\mathbb{E}(e^{it\eta_j}) | \\
&+ \quad | \mathbb{E}(e^{it \sum_{j=1}^{k-1} \eta_j}) - \prod_{j=1}^{k-1} \mathbb{E}(e^{it\eta_j}) | \\
&= \quad | Cov(e^{it \sum_{j=1}^{k-1} \eta_j}, e^{it\eta_k}) | + | \mathbb{E}(e^{it \sum_{j=1}^{k-1} \eta_j}) \\
&- \prod_{j=1}^{k-1} \mathbb{E}(e^{it\eta_j}) |
\end{aligned}
\tag{12}
$$

and successively, we have

$$| \mathbb{E}(e^{it \sum_{j=1}^{k} \eta_j}) - \prod_{j=1}^{k} \mathbb{E}(e^{it\eta_j}) | \leq | Cov(e^{it \sum_{j=1}^{k-1} \eta_j}, e^{it\eta_k}) | + | Cov(e^{it \sum_{j=1}^{k-2} \eta_k - 1}, e^{it\eta_j}) |$$

$$+ \cdots + | Cov(e^{it\eta_2}, e^{it\eta_1}) | . \tag{13}$$

Once again we apply Lemma 6.2, to write

$$| Cov(e^{it\eta_2}, e^{it\eta_1}) | \leq C(h_K^{-1} Lip(K) + h_H^{-1} Lip(H))^2 \frac{\phi(x, h_K)}{n(\mathbb{E}[K_1])^2} \sum_{i \in I_1} \sum_{j \in I_2} \lambda_{i,j}.$$

Applying this inequality to each term on the right-hand side of (13). We obtain

$$| \mathbb{E}(e^{it \sum_{j=1}^{k} \eta_j}) - \prod_{j=1}^{k} \mathbb{E}(e^{it\eta_j}) | \leq C(h_K^{-1} Lip(K) + h_H^{-1} Lip(H))^2 \frac{\phi(x, h_K)}{n(\mathbb{E}K_1)^2}$$

$$\times \left(\sum_{i \in I_1} \sum_{j \in I_2} \lambda_{i,j} + \sum_{i \in I_1 \cup I_2} \sum_{j \in I_3} \lambda_{i,j} + \cdots + \sum_{i \in I_1 \cup \cdots \cup I_{k-1}} \sum_{j \in I_k} \lambda_{i,j} \right).$$

Observe that for every $2 \leq l \leq k-1, (i,j) \in I_l * I_{l+1}$, we have $|i - j| \geq q + 1 > q$, then

$$\sum_{i \in I_1 \cup \cdots \cup I_{k-1}} \sum_{j \in I_k} \lambda_{i,j} \leq p\lambda_q.$$

Therefore, inequality (12) becomes

$$
\begin{aligned}
\mid \mathbb{E}(e^{it\sum_{j=1}^{k}\eta_j}) - \prod_{j=1}^{k}\mathbb{E}(e^{it\eta_j}) \mid &= Ct^2(h_K^{-1}Lip(K) + h_H^{-1}Lip(H))^2 \frac{\phi(x,h_K)}{n(\mathbb{E}[K_1])^2}kp\lambda_q \\
&= Ct^2(h_K^{-1}Lip(K) + h_H^{-1}Lip(H))^2 \frac{\phi(x,h_K)}{(\mathbb{E}[K_1])^2}kpe^{-\alpha q} \\
&= Ct^2(h_K^{-1}Lip(K) + h_H^{-1}Lip(H))^2 \frac{1}{n\phi(x,h_K)}kp\lambda_q \\
&= Ct^2 \frac{kp}{nh_H^2\phi^3(x,h_K)}\lambda_q \to 0.
\end{aligned}
$$

Proof of (11).
By the same arguments used in 5, we have

$$
\begin{aligned}
\lim_{n\to\infty} kVar(\eta_1) &= \lim_{n\to\infty} kpVar(Z_{n1}(x,y)) \\
&= \lim_{n\to\infty} \frac{\phi(x,h_K)}{n(\mathbb{E}K_1)^2}Var(\Gamma_1(x,y))
\end{aligned}
$$

So, by using the same arguments as those used by (Ferraty and Vieu, 2007), we get

$$
\begin{aligned}
\frac{1}{\phi(x,h_K)}\mathbb{E}(K_1^2) &\to K_1^2 - \int_0^1 (K^2)'(s)\beta(x,s)ds + o(1) \\
\frac{\mathbb{E}(K_1^2 H_1^2)}{\mathbb{E}(K_1^2)} &\to F^x(y)(1 - F^x(y))\int H'^2(t)dt \\
\frac{\mathbb{E}(K_1^2 H_1^2)}{\mathbb{E}(K_1^2)} &\to F^x(y)(1 - F^x(y))
\end{aligned}
$$

which imply that

$$
\frac{\phi(x,h_K)}{n(\mathbb{E}K_1)^2}Var(\Gamma_1(x,y)) \to \sigma^2(x).
$$

Hence

$$
kVar(\eta_1) \to \sigma^2(x).
$$

For the second part of (11), we use the fact that:

$$
\mid \eta_1 \mid \le Cp \mid Z_{n1}(x,y) \mid \le \frac{Cp}{\sqrt{n\phi(x,h_K)}}
$$

and by Tchebychev's inequality we get:

$$
\begin{aligned}
k\mathbb{E}(\eta_1^2 \mathbf{1}_{\eta_1 > \epsilon\sigma(x)}) &\le \frac{Cp^2 k}{n\phi(x,h_K)}\mathbb{P}(\eta_1 > \epsilon\sigma(x)) \\
&\le \frac{Cp^2 k}{n\phi(x,h_K)}\frac{Var(\eta_1)}{\epsilon^2\sigma^2(x)} \\
&= O\left(\frac{p^2}{n\phi(x,h_K)}\right).
\end{aligned}
$$

6.2 Proof of lemma 4.3

We have

$$|\widehat{F}_D^x(y) - \mathbb{E}\widehat{F}_D^x(y)| = \frac{1}{n\mathbb{E}[K_1]}\sum_{i=1}^{n}\Delta_i$$

where

$$\Delta_i = K(h_K^{-1}d(x, X_i)) - \mathbb{E}[K_1], 1 \leq i \leq n$$

clearly we have $\mathbb{E}(\Delta_i) = 0$ and Moreover, we can write:

$$\| \Delta_i \|_{\infty}{}^1 \leq 2C \parallel K \parallel_{\infty}$$

and

$$Lip(\Delta_i) \leq Ch_K^{-1}Lip(K).$$

Now, to apply lemma 6.2, we have to evaluate the variance term $Var(\sum_{i=1}^{n}\Delta_i)$ and the covariance term $cov(\Delta_{s_1}...\Delta_{s_u}, \Delta_{t_1}...\Delta_{t_v})$, for all $(s_1, ..., s_u) \in \mathbb{N}^u, (t_1, ..., t_v) \in \mathbb{N}^v$ with $1 \leq s_1 \leq \cdots \leq s_u \leq t_1 \leq \cdots \leq t_v \leq n$.

Firstly, for the covariance term, we consider the following cases: If $t_1 = s_u$. By using the fact that $\mathbb{E}[\parallel K_1 \parallel] = O(\phi(x, h_K))$ we have:

$$
\begin{aligned}
\mid Cov(\Delta_{s_1}...\Delta_{s_u}, \Delta_{t_1}...\Delta_{t_v}) \mid &\leq \left(\frac{C}{n\mathbb{E}[K_1]}\right)^{u+v} \mathbb{E} \mid \Delta_i \mid^{u+v} \\
&\leq \left(\frac{C \parallel K \parallel_{\infty}}{n\mathbb{E}[K_1]}\right)^{u+v} \mathbb{E}[\parallel K_1 \parallel] \\
&\leq \phi(x, h_K)\left(\frac{C}{n\phi(x, h_K)}\right)^{u+v}
\end{aligned}
$$

If $t_1 > s_u$, we use the quasi-association, under (H7), we get :

$$
\begin{aligned}
\mid Cov(\Delta_{s_1}...\Delta_{s_u}, \Delta_{t_1}...\Delta_{t_v}) \mid &\leq \left(\frac{Lip(K)}{nh_K\mathbb{E}[K_1]}\right)^2 \\
&\times \left(\frac{C}{n\mathbb{E}[K_1]}\right)^{u+v-2} \sum_{i=1}^{u}\sum_{j=1}^{v}\lambda_{s_i,t_j} \\
&\leq \left(h_K^{-1}Lip(K)\right)^2 \left(\frac{C}{n\mathbb{E}[K_1]}\right)^{u+v} v\lambda_{t_1-s_u} \\
&\leq \left(h_K^{-1}Lip(K)\right)^2 \left(\frac{C}{\phi(x, h_K)}\right)^{u+v} ve^{-\alpha(t_1-s_u)}.
\end{aligned}
$$

(14)

On the other hand, by (H7) we have:

---

[1]For any function f we denote by $\parallel f \parallel_{\infty}$ the supremun norm.

$| Cov(\Delta_{s_1}...\Delta_{s_u}, \Delta_{t_1}...\Delta_{t_v}) | \leq$

$$
\left(\frac{C \parallel K \parallel_\infty}{n[K_1]}\right)^{u+v-2} (| \, \mathbb{E}[\Delta_{s_u}, \Delta_{t_1}] \, | + \mathbb{E} \mid \Delta_{s_u} \mid \mathbb{E} \mid \Delta_{t_1} \mid
$$

$$
\leq \left(\frac{C \parallel K \parallel_\infty}{n\mathbb{E}[K_1]}\right)^{u+v-2} \left(\frac{C}{n\mathbb{E}[K_1]}\right)^2
$$

$$
\times \quad (sup_{i\neq j}\mathbb{P}\left((X_i, X_j) \in B(x, h_K) \times B(x, h_K) + \mathbb{P}\left(X_1 \in B(x, h_K)\right)^2\right)
$$

$$
\leq \left(\frac{C}{\phi(x, h_K)}\right)^{u+v} (\phi(x, h_K))^2.
$$

$$(15)$$

Furthermore, taking a $\gamma - power$ of (14), $(1 - \gamma) - power$ of (15), with $1 = 4 < \gamma < 1 = 2$, we obtain an upper-bound of the tree terms as follows: for $1 \leq s_1 \leq \cdots \leq s_u \leq t_1 \leq \cdots \leq t_v \leq n$

$$
| \, cov(\Delta_{s_1}...\Delta_{s_u}, \Delta_{t_1}...\Delta_{t_v}) \, | \leq \phi(x, h_K) \left(\frac{C}{n\phi(x, h_K)}\right)^{u+v}.
$$

Secondly, for the variance term $Var(\sum_{i=1}^n \Delta_i)$,
we put, for all $1 \leq i \leq n$,

$$
\begin{aligned}
| \, Var(\Delta_{s_1}...\Delta_{s_u}, \Delta_{t_1}...\Delta_{t_v}) \, | &= \left(\frac{1}{n\mathbb{E}[K_1]}\right)^2 \sum_{i=1}^n \sum_{j=1}^n Cov(K_i, K_j) \\
&= \left(\frac{1}{n\mathbb{E}[K_1]}\right)^2 Var(K_1) \\
&+ \left(\frac{1}{n\mathbb{E}[K_1]}\right)^2 \sum_{i=1}^n \sum_{j=1,i\neq j}^n Cov(K_i, K_j)
\end{aligned}
$$

$$(16)$$

for the first term,
$$
Var(K_1) = \mathbb{E}\left(K_1^2\right) - (\mathbb{E}(K_1))^2
$$

then,
$$
\mathbb{E}[K_1^2] = O\left(\phi(x, h_K)\right).
$$

It follows that:
$$
\left(\frac{1}{n\mathbb{E}[K_1]}\right)^2 Var(K_1) = O(n\phi(x, h_K)). \quad (17)
$$

Now, let us evaluate the asymptotic behavior of the sum in the second term of (16).

For this, we need the following decomposition:

$$\sum_{i=1}^{n}\sum_{j=1,i\neq j}^{n} Cov(K_i, K_j) = \underbrace{\sum_{i=1}^{n}\sum_{j=1,0<|i-j|\leq m_n}^{n} Cov(K_i, K_j)}_{I}$$

$$+ \underbrace{\sum_{i=1}^{n}\sum_{j=1,|i-j|>m_n}^{n} Cov(K_i, K_j)}_{II}$$

where (m_n) is a sequence of positive integer which goes to infinity as $n \to \infty$. From Assumptions (H1), (H4) and (H8), we have, for $i \neq j$

$$\begin{aligned}
I &\leq nm_n \left(max_{i\neq j} \mid \mathbb{E}(K_iK_j) \mid +(\mathbb{E}(K_1))^2\right) \\
&\leq Cnm_n \left(\phi^2(x, h_K) + \phi^2(x, h_K)\right) \\
&\leq Cnm_n \left(\phi^2(x, h_K)\right).
\end{aligned}$$

$$(18)$$

Since the kernels K is bounded and Lipschitzian, we get

$$\begin{aligned}
II &\leq (h_K^{-1}Lip(K))^2 \sum_{i=1}^{u}\sum_{j=1_{|i-j|>m_n}}^{v} \lambda_{i,j} \\
&\leq C(h_K^{-1}Lip(K))^2 \sum_{i=1}^{u}\sum_{j=1_{|i-j|>m_n}}^{v} \lambda_{i,j} \\
&\leq Cn(h_K^{-1}Lip(K))^2 \lambda_{m_n} \\
&\leq Cn(h_K^{-1}Lip(K))^2 e^{-\alpha m_n}.
\end{aligned}$$

$$(19)$$

Then, by (18) and (19), we get

$$\sum_{j=1,i\neq j}^{n} Cov(K_i, K_j) \leq C \left(nm_n(\phi^2(x, h_K)) + n\left(h_K^{-1}Lip(K)\right)^2 e^{-\alpha m_n}\right)$$

by choosing

$$m_n = \log\left(\frac{(h_K^{-1}Lip(K))^2}{\alpha\phi^2(x, h_K)}\right)$$

we get

$$\frac{1}{\phi(x, h_K)} \sum_{j=1,i\neq j}^{n} Cov(K_i, K_j) \to 0, \text{ as } n \to \infty. \qquad (20)$$

Finally, by combining results (16), (17) and (20), we get:

$$Var\left(\sum_{i=1}^{n}\Delta_i\right) = O\left(\frac{1}{n\phi(x,h_K)}\right).$$

6.3 Proof of lemma 4.4

We have

$$
\begin{aligned}
\mathbb{P}\{|\widehat{F}_D^x(y)| \le 1/2\} &\le \mathbb{P}\{|\widehat{F}_D^x(y) - 1| > 1/2\} \\
&\le \mathbb{P}\{|\widehat{F}_D^x(y) - \mathbb{E}\widehat{F}_D^x(y)| > 1/2\}
\end{aligned}
$$

we deduce that

$$\sum_{n\in\mathbb{N}}\mathbb{P}\left(\widehat{F}_D^x(y) < 1/2\right) < \infty.$$

7. Conclusion

In this paper, we established the asymptotic normality property of the kernel estimate of the conditional distribution function in quasi-associated data framework. Our theoretical and practical studies confirm that our kernel estimator has good asymptotic properties.

References

1. Attaoui, S., Laksaci, A. and Ould-Saïd, E. (2015). Asymptotic Results for an estimator of the Regression Function for Quasi-Associated Processes. Functional Statistics and Applications, Contributions to Statistics 10.1007/978-3-319-22476-3-1.

2. Daoudi, H. and B. Mechab. (2019). Rate of pointwise consistency for nonparametric of the conditional distribution function estimate with functional variables for quasi-associated data, Statistics, optimization and information computing, Accepted:27/07/2019.

3. Daoudi, H., Mechab, B. and Elmezouar, C.Z. (2018). Asymptotic Normality of a Conditional Hazard Function Estimate in the Single Index for Quasi-Associated Data, Communication in Statistics Theory and Methods. https://www.tandfonline.com/doi/full/10.1080/03610926.2018.1549248.

4. Daoudi, H., Mechab, B., Benaissa, S., and Rabhi., A. (2019). Asymptotic normality of the nonparametric conditional density function estimate with functional variables for the quasi-associated data, International Journal of Statistics and Economics 20(3), 94-106.

5. Douge, L. (2010). Théorèmes limites pour des variables quasi-associes hilbertiennes, Ann. I.S.U.P 54, 51-60.

6.   Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities, Stoch. Proc. Appl 84, 313-342.

7.   Ezzahrioui, M. and Ould-Saïd, E. (2008). Asymptotic Results of a Nonparametric Conditional Quantile Estimator for Functional Time Series, Communication in Statistics Theory and Methods 37(17), 2735-2759.

8.   Ferraty, F. Mas, A. and Vieu, P., (2007). Advances in nonparametric regression for functional variables. Aust. and New Zeal. J. of Statist 49, 1-20.

9.   Kallabis, R. S. and Neumann, M. H. (2006). An exponential inequality under weak dependence, Bernoulli 12, 333-350.

10.   Laksaci, A. and B. Mechab. (2014). Conditional hazard estimate for functional random fields. Journal of Statistical Theory and Practice 8, 192-200.

11.   Mechab, W. and Laksaci, A. (2016). Nonparametric relative regression for associated random variables, Metron 74, 75-97.

12.   Tabti, H. and Ait Saidi, A. (2018). Estimation and simulation of conditional hazard function in the quasi-associated framework when the observations are linked via a functional single-index structure, Journal Communications in Statistics-Theory and Methods 47, 816-838.