STATISTICS IN TRANSITION new series, June 2022
Vol. 23, No. 2, pp. 89–105, DOI 10.2478/stattrans-2022-0018
Received – 30.01.2021; accepted – 18.01.2022

On the nonparametric estimation of the conditional hazard estimator in a single functional index

Abdelmalek Gagui[1], **Abdelhak Chouaf**[2]

ABSTRACT

This paper deals with the conditional hazard estimator of a real response where the variable is given a functional random variable (i.e it takes values in an infinite-dimensional space). Specifically, we focus on the functional index model. This approach offers a good compromise between nonparametric and parametric models. The principle aim is to prove the asymptotic normality of the proposed estimator under general conditions and in cases where the variables satisfy the strong mixing dependency. This was achieved by means of the kernel estimator method, based on a single-index structure. Finally, a simulation of our methodology shows that it is efficient for large sample sizes.

Key words: single functional index, conditional hazard function, nonparametric estimation, α-mixing dependency, asymptotic normality, functional data.

1. Introduction

The nonparametric estimation of the hazard function plays a crucial role in statistical analyses. This subject can be approached from multiple perspectives depending on the complexity of the problem. Many techniques have been studied in the literature to treat these various situations but all treat only real or multidimensional explanatory random variables. We refer to Watson and Leadbetter (1964), who were the first to study the nonparametric estimation of the hazard function. In the sequel, many authors have been interested in the study of such a function (see, for example Tanner and Wong (1983), Delecroix and Yazourh (1992), Collomb et al. (1985) and Youndjé et al. (1996)).

Focusing on functional data, the first results on the nonparametric estimate of this model, were achieved by Ferraty et al. (2000). They have studied the almost complete convergence of an estimator with kernel for the function of a chance of a real random variable conditioned by a functional explanatory variable. For instance, Masry (2005) showed the asymptotic normality of the estimator for the function of regression, Ferraty et al. (2007) studied the mean squared convergence, Burba et al. (2008) are interested in the estimate of the function of regression by using the method of k-nearest neighbours, Quintela-del-Rio (2008) obtained the asymptotic normality of the non-parametric estimation of the conditional hazard function. Ferraty et al. (2010) they etablished the almost complete convergence uniform on the functional component of this nonparametric model.

[1]Djillali Liabes University, Algeria. E-mail: gagui.abdelmalek@gmail.com. ORCID: https://orcid.org/0000-0002-2715-4304.
[2]Djillali Liabes University, Algeria. E-mail: abdo_stat@yahoo.fr.

The modelling of the spatial data was also considered in nonparametric estimation for functional data. On this subject, Dabo-Niang et al. (2012) studied the almost convergence of an estimator with kernel for the function of regression. Laksaci et al. (2009) treated the almost complete convergence of the estimator with a kernel of the function of conditional distribution and the conditional quantiles. Li and Tran (2007) obtained the asymptotic normality of a kernel estimator of the hazard function. The study of the kernel estimator of the conditional hazard function when the covariates take values in functional statistic was treated by Lakssaci et al. (2010).

Our goal in this work is devoted to the study of the single functional index model. This approach consists of making a projection between the explanatory variable Y on the functional response variable X to the non-parametric context on a function directly θ. In the finite-dimensional, random variables have been widely studied, see for example Hardle et al. (1993), Hristache et al. (2001). Furthermore, when the case is infinite dimensions or when the explanatory variable is functional, the first work which was interested in the single-index model for the nonparametric estimation is Ferraty et al. (2003). They stated for i.i.d. variables and obtained the almost complete convergence under some conditions. In the same context Ait Saidi et al. (2005) studied the dependent case of these estimators, Ait Saidi et al. (2008) proposed cross-validated estimation where the functional index is an unknown, Attaoui et al. (2011) obtained the uniform almost complete convergence of conditional density in the functional single index. More recently Tabti et al. (2017) obtained the pointwise almost complete convergence and the uniform almost complete convergence of a kernel estimator of the hazard function with the quasi-association condition in a single-index approach.

In the present paper, we obtain, under some conditions, the asymptotic normality of the conditional hazard function estimator. This result enables us to obtain the confidence intervals of this estimator. In practice, this study has great importance because it permits us to construct a prediction method based on the maximum risk estimation with a single functional index.

In Section 2, we introduce the estimator of our model in the single-functional index. Section 3 we introduce assumptions and asymptotic properties are given. Practical aspects are discussed in Section 4. Simulations are given in Section 5. Finally, Section 6 is devoted to the proofs of the results.

2. The model

Let $\{(X_i,Y_i),\ 1 \leq i \leq n\}$ be n random variables, identically distributed as the random pair (X,Y) with values in $\mathbb{H}\times\mathbb{R}$, where $\mathbb{H}$ is a separable real Hilbert space with the norm $\|\,.\,\|$ generated by an inner product $<.,.>$. We consider the semi-metric d_θ associated with the single index $\theta \in \mathbb{H}$ defined by $\forall x_1,x_2 \in \mathbb{H} : d_\theta(x_1,x_2) :=|< x_1 - x_2, \theta >|$. Assume that the explanation of Y given X is done through a fixed functional index θ in $\mathbb{H}$. In the sense that there exists a θ in $\mathbb{H}$ (unique up to a scale normalization factor) such that: $\mathbb{E}[Y|X] = \mathbb{E}[Y| < \theta, X >]$. The conditional cumulative distribution function of Y given $< X, \theta >$ is denoted by

$$F^x(\theta,y) := F(y| < \theta,x >) = \mathbb{P}(Y \le y \,|< X,\theta >=< \theta,x >),\ \forall y \in \mathbb{R}$$

Clearly we have for all $x \in \mathbb{H}$,

$$F_1(.| < x,\theta_1 >) = F_2(.| < x,\theta_2 >) \Rightarrow F_1 \equiv F_2 \text{ and } \theta_1 = \theta_2.$$

The natural kernel estimator of $F(\theta,y,x)$ is defined as

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

We suppose that the conditional density of Y given $X = x$ denoted by $f(.|x)$ exists and is given by $\forall y \in \mathbb{R},\ f_\theta(y|x) := f(y| < x,\theta >)$. In the following, we denote by $f(\theta,.,x)$, the conditional density of Y given $< x,\theta >$ and we define the kernel estimator $\widehat{f}(\theta,.,x)$ of $f(\theta,.,x)$ by:

$$\widehat{f}(\theta,y,x) = \frac{h_H^{-1}\sum_{i=1}^{n} K(h_K^{-1} d_\theta(x,X_i))H'(h_H^{-1}(y-Y_i))}{\sum_{i=1}^{n} K(h_K^{-1} d_\theta(x,X_i))},\ \forall y \in \mathbb{R} \tag{2}$$

with the convention $0/0 = 0$, where K and H are kernels function (H' is the derivate of H) and $h_K := h_{n,K}$ (resp $h_H := h_{n,H}$) is a sequence of bandwidths that decrease to zero as n goes to infinity.

We are interested in estimating non parametrically the conditional hazard function λ defined by:

$$\hat{\lambda}(\theta,y,x) = \frac{\widehat{f}(\theta,y,x)}{1-\widehat{F}(\theta,y,x)},\ \forall y \in \mathbb{R}.$$

3. Main results

We begin with introducing some notations. Let $(X_i,Y_i)_{i=1}^{\infty}$ be a sequence of random variables and $\alpha(n)$ be a sequence of real numbers. A stationary process $(X_i,Y_i)_{i=1}^{\infty}$ is called α-mixing or strongly mixing, if
$\alpha(n) = \sup_{A\in\mathscr{A}_1^k} \sup_{B\in\mathscr{A}_{n+k}^{\infty}} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| \to 0$, as $n\to\infty$, where $\mathscr{F}_a^b$ is the σ-algebra generated by $(X_j,Y_j)_{j=a}^{b}$.

In this section, we give some obtained results on the asymptotic normality of the estimator $\widehat{\lambda}(\theta,y,x)$, which require the following additional hypotheses. All along the paper, when no confusion is possible, we will denote by C and C' some strictly positive generic constants. We put, for any $x \in \mathbb{H}$, and $i = 1,...,n$ $K_i(\theta,x) := K(h_K^{-1} d_\theta(x,X_i))$ and, for all $y \in \mathbb{R}$, $H_i^j := H^j(h_H^{-1}(y-Y_i))$ for $j = 0,1$ In the following, for any $x \in \mathbb{H}$ and $y \in \mathbb{R}$, let $\mathscr{N}_x$ be a fixed neighbourhood of x in $\mathbb{H}$, $S_{\mathbb{R}}$ will be a fixed compact subset of $\mathbb{R}$,and we will use the notation $B_\theta(x,h) := \{x_1 \in \mathbb{H} :\ 0 < | < x-x_1,\theta > | < h\}$, the ball centered at x, with radius h. All along the paper, when no confusion will be possible, we will denote by C, C'

and $C_{\theta,x}$ some generic constant in $\mathbb{R}_+^*$

(H1) $\mathbb{P}(X \in B_{\theta,x}(h)) = \phi_{\theta,x}(h) > 0$. Moreover, there exists a function $\beta_{\theta,x}(.)$ such that: $\forall s \in [-1,1],\ \lim_{n\longrightarrow\infty} \frac{\beta_{\theta,x}(sh_K)}{\beta_{\theta,x}(h_K)} = \beta_{\theta,x}(s)$.

(H2) For $l \in \{0,2\}$, the functions $\psi_l(s) = \mathbb{E}\left[\frac{\partial^l f(\theta,y,X)}{\partial y^l} - \frac{\partial^l f(\theta,y,x)}{\partial y^l} | d_\theta(x,X=s)\right]$ and $\Psi_l(s) = \mathbb{E}\left[\frac{\partial^l F(\theta,y,X)}{\partial y^l} - \frac{\partial^l F(\theta,y,x)}{\partial y^l} | d_\theta(x,X=s)\right]$ are differentiable at $s=0$.

(H3) The kernel K is a differentiable function and its derivative K' exists and is such that there exist two constants C and C' with $-\infty < C < K'(t) < C' < 0$, for $t \in [0,1]$.

(H4) The kernels K and H are an even bounded function .

(H5) The bandwidths h_K and h_H satisfy
(1*) $\lim_{n\longrightarrow\infty} \frac{1}{nh_H\phi_{\theta,x}(h_K)} = 0$,
(2*) $\lim_{n\longrightarrow\infty} nh_H^5\phi_{\theta,x}(h_K) = 0$ and $\lim_{n\longrightarrow\infty} nh_H h_k^2\phi_{\theta,x}(h_K) = 0$,
(3*) $\lim_{n\longrightarrow\infty} h_H = 0,\ \lim_{n\longrightarrow\infty} h_K = 0,\ \textit{and}\ \lim_{n\longrightarrow\infty} \frac{\log n}{n\phi_{\theta,x}(h_K)} = 0$,
(4*) $\lim_{n\longrightarrow\infty} h_K^{2b_1}\phi_{\theta,x}(h_K) = 0,\ \textit{and}\ \lim_{n\longrightarrow\infty} h_H^{2b_1}\phi_{\theta,x}(h_K) = 0$.

(H6) $(X_i,Y_i)_{i\in\mathbb{N}}$ is a strongly mixing sequence, whose mixing coefficient $\alpha(n)$ satisfies $\exists a > (5+\sqrt{17})/2,\ \exists C > 0 : \forall n \in \mathbb{N},\ \alpha(n) \le Cn^{-a}$.

(H7) $0 < \sup_{i\neq j} \mathbb{P}((X_i,X_j) \in B_\theta(x,h_K) \times B_\theta(x,h_K)) = O\left(\frac{\phi_{\theta,x}(h_K)^{(a+1)/a}}{n^{1/a}}\right)$.

(H8) $\exists \beta_0 > 0,\ C_1,\ C_2 > 0$, such that: $C_1 n^{\frac{3-a}{a+1}+\beta_0} \le \phi_{\theta,x}(h_K) \le C_2 n^{\frac{1}{1-a}}$.

Comments on the assumptions

Assumptions (H1)-(H4) are technicals and permit to give an explicit asymptotic variance. The function $\beta_{\theta,x}(.)$ will play a major role in our results, it intervenes to compute the exact constant terms involved in our asymptotic expansions (for more of this assumptions, see Ferraty et al. 2007). Finally (H5)-(H8) permits to remove the bias term in the asymptotic normality result.
Now, we give our main result.

Theorem 3.1. *Assume that (H1)-(H5) hold, and (H6)-(H8) hold, as n goes to infinity, we have*

$$(nh_H\phi_{\theta,x}(h_K))^{1/2}(\widehat{\lambda}(\theta,y,x) - \lambda(\theta,y,x) - B_n(\theta,y,x)) \xrightarrow{\mathscr{D}} \mathscr{N}(0,\sigma_h^2(\theta,y,x)),$$

where

$$B_n(\theta,y,x) = \frac{1}{1-F(\theta,y,x)}\left((B_H^f - \lambda(\theta,y,x)B_H^F)h_H^2 + ((B_K^f - \lambda(\theta,y,x)B_K^F)h_K)\right)$$

with

$$\sigma_h^2(\theta,y,x) = \frac{M_2\lambda(\theta,y,x)}{M_1^2(1-F(\theta,y,x))}$$

$M_0 = K(1) - \int_0^1 sK'(s)\beta_{\theta,x}(s)ds, \quad M_j = K^j(1) - \int_0^1 (K^j)'(s)\beta_{\theta,x}(s)ds$
for $j = 1,2$

and

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

$$B_K^f(\theta,y,x) = h_k\psi_0'(0)\frac{M_0}{M_1}h_K.$$

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

$$B_K^F(\theta,y,x) = h_k\Psi_0'(0)\frac{M_0}{M_1}h_K.$$

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

Corollary 3.1. *Under the hypotheses of Theorem 3.1,and if the bandwidth parameters (h_K and h_H) satisfies (H5) and if the function $\phi_{\theta,x}(h_K)$ satisfies :*

$$\lim_{n\longrightarrow\infty}(h_H^2 + h_K)(n\phi_{\theta,x}(h_K))^{1/2} = 0,$$

we have

$$(nh_H\phi_{\theta,x}(h_K))^{1/2}(\widehat{\lambda}(\theta,y,x) - \lambda(\theta,y,x)) \xrightarrow{\mathscr{D}} \mathscr{N}(0,\sigma_h^2(\theta,y,x)),$$

The proof of Theorem 3.1 is based on the following decomposition:

$$\begin{aligned}
\widehat{\lambda}(\theta,y,x) - \lambda(\theta,y,x) &= \frac{1}{\widehat{F}_D(\theta,x) - \widehat{F}_N(\theta,y,x)}\left(\widehat{f}_N(\theta,y,x) - \mathbb{E}[\widehat{f}_N(\theta,y,x)]\right) \\
&+ \frac{1}{\widehat{F}_D(\theta,x) - \widehat{F}_N(\theta,y,x)}\Big\{\lambda(\theta,y,x)\left(\mathbb{E}[\widehat{F}_N(\theta,y,x)] - F(\theta,y,x)\right) \\
&+ \left(\mathbb{E}[\widehat{f}_N(\theta,y,x)] - f(\theta,y,x)\right)\Big\} \\
&+ \frac{\widehat{h}(\theta,y,x)}{\widehat{F}_D(\theta,x) - \widehat{F}_N(\theta,y,x)}\Big\{1 - \mathbb{E}[\widehat{F}_N(\theta,y,x)] \\
&- \left(\widehat{F}_D(\theta,x) - \widehat{F}_N(\theta,y,x)\right)\Big\}
\end{aligned}$$

Lemma 3.1. *Under the Assumptions of Theorem 3.1, as n goes to infinity, we have*

$$(nh_H\phi_{\theta,x}(h_K))^{1/2}(\widehat{f}_N(\theta,y,x)-\mathbb{E}[\widehat{f}_N(\theta,y,x)]) \xrightarrow{\mathscr{D}} \mathscr{N}(0,\sigma_f^2(\theta,y,x)).$$

Proof of lemma 3.1 First, we define that

$$Z_i(\theta,y,x)=\frac{\sqrt{\phi_{\theta,x}(h_K)}}{\sqrt{nh_H}\mathbb{E}[K_1(\theta,x)]}(\zeta_i(\theta,y,x)-\mathbb{E}[\zeta_i(\theta,y,x)]),$$

and

$$T_n:=\sum_{i=1}^{n} Z_i(\theta,y,x).$$

where $\zeta_i(\theta,y,x)=H_i'(\theta,x)K_i(\theta,x)$,
Thus,

$$T_n=\sqrt{nh_H\phi_{\theta,x}(h_K)}(\widehat{f}_N(\theta,y,x)-\mathbb{E}[\widehat{f}_N(\theta,y,x)]).$$

So, our claimed result is now

$$T_n \longrightarrow \mathscr{N}(0,\sigma_f^2(\theta,x)). \tag{3}$$

Therefore, we have

$$\begin{aligned} Var(T_n) &= nh_H\phi_{\theta,x}(h_K)Var(\widehat{f}_N(\theta,y,x)-\mathbb{E}[\widehat{f}_N(\theta,y,x)]) \\ &= nh_H\phi_{\theta,x}(h_K)Var(\widehat{f}_N(\theta,y,x)) \end{aligned} \tag{4}$$

Now, we need to evaluate the variance of $\widehat{f}_N(\theta,y,x)$. For this we have for all $1\le i\le n$, :

$$\begin{aligned} Var(\widehat{f}_N(\theta,y,x)) &= \frac{1}{(nh_H\mathbb{E}[K_1(\theta,x)])^2}\sum_{i=1}^{n}\sum_{j=1}^{n} Cov(\zeta_i(\theta,y,x),\zeta_j(\theta,y,x)) \\ &= I_{1,n}+I_{2,n}. \end{aligned}$$

where

$$\begin{aligned} I_{1,n} &= \frac{1}{n(h_H\mathbb{E}[K_1(\theta,x)])^2}Var(\zeta_1(\theta,y,x)) \\ I_{2,n} &= \frac{1}{(nh_H\mathbb{E}[K_1(\theta,x)])^2}\sum_{i=1}^{n}\sum_{j=1_{i\neq j}}^{n} Cov(\zeta_i(\theta,y,x),\zeta_j(\theta,y,x)). \end{aligned}$$

First, for the quantity $I_{1,n}$, we have

$$\begin{aligned} Var(\zeta_1(\theta,y,x)) &\leq \mathbb{E}\left[H'^2_1(y)K_1^2(\theta,x)\right] \\ &\leq \mathbb{E}\left[K_1^2(\theta,x)\mathbb{E}\left[H'^2_1(y)| < \theta, X_1 >\right]\right]. \end{aligned}$$

$$\begin{aligned} |\mathbb{E}\left[H'^2_1(y)| < \theta, X_1 >\right]| &= \left|\int_{\mathbb{R}} H'^2(h_H^{-1}(y-z))f(\theta,z,x)dz\right| \\ &\leq h_H \int_{\mathbb{R}} H'^2|f(\theta,y-h_Ht,x)f(\theta,y,x)|dt \\ &+ h_H f(\theta,y,x)\int_{\mathbb{R}} H'^2 dt \\ &\leq h_H^{1+b_2}\int_{\mathbb{R}} |t|^{b_2}H'^2dt + h_H f(\theta,y,x)\int_{\mathbb{R}} H'^2 dt \\ &= h_H\left(o(1) + f(\theta,y,x)\left(\int_{\mathbb{R}} H'^2 dt\right)\right). \end{aligned}$$

As $n \longrightarrow \infty$, $\mathbb{E}[K_1^2(\theta,x)] \longrightarrow M_2\phi_{\theta,x}(h_K)$, one gets

$$Var(\zeta_1(\theta,y,x)) = M_2\phi_{\theta,x}(h_K)h_H\left(o(1) + f(\theta,y,x)\left(\int_{\mathbb{R}} H'^2 dt\right)\right).$$

So, using (H5-1*), we get

$$\begin{aligned} I_{1,n} &= \frac{M_2\phi_{\theta,x}(h_K)}{n(M_1h_H\phi_{\theta,x}(h_K))^2}h_H\left(o(1) + f(\theta,y,x)\left(\int_{\mathbb{R}} H'^2 dt\right)\right) \\ &= o\left(\frac{1}{nh_H\phi_{\theta,x}(h_K)}\right) + \frac{M_2 f(\theta,y,x)}{M_1^2 nh_H\phi_{\theta,x}(h_K)}\left(\int_{\mathbb{R}} H'^2 dt\right) \\ &\longrightarrow \frac{M_2 f(\theta,y,x)(\int_{\mathbb{R}} H'^2 dt)}{M_1^2 nh_H\phi_{\theta,x}(h_K)}, \ as\ n \longrightarrow \infty. \end{aligned} \tag{5}$$

Second, for the quantity $I_{2,n}$, we will use the following decomposition:

$$I_{2,n} = \sum_{i=1}^{n}\sum_{\substack{j=1 \\ 0<|i-j|\leq m_n}}^{n} Cov(\zeta_i(\theta,y,x),\zeta_j(\theta,y,x)) + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ |i-j|>m_n}}^{n} Cov(\zeta_i(\theta,y,x),\zeta_j(\theta,y,x)).$$

Similarly to Attaoui said (2014), we can easily write

$$I_{2,n} = O(nh_H^2\phi_{\theta,x}(h_K)).$$

It yields,

$$\frac{1}{nh_H\phi_{\theta,x}(h_K)}I_{2,n} \longrightarrow 0, \ as\ n \longrightarrow \infty. \tag{6}$$

Finally, the proof of the Lemma is completed, to get

$$Var(T_n) \longrightarrow \frac{M_2 f(\theta,y,x)}{M_1^2}\left(\int_{\mathbb{R}} H'^2 dt\right) =: \sigma_f^2(\theta,x).$$

Lemma 3.2. *Under the Assumptions (H1)-(H4), and (H7), as n goes to infinity, we have*

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

Proof of lemma 3.2 First, for $\mathbb{E}[\widehat{F}(\theta,y,x)]$, we start by writing

$$\mathbb{E}[\widehat{F}_N(\theta,y,x)] = \frac{1}{\mathbb{E}[K_1(\theta,x)]}\mathbb{E}\left[K_1(\theta,x)\mathbb{E}[h_H^{-1}H_1'(y)|X]\right]$$

with

$$h_H^{-1}\mathbb{E}[H_1'(y)|X] = \int_{\mathbb{R}} H'(t)F(\theta,y-h_H t,X)dt$$

The latter can be re-written, using a Taylor expansion under (H4), as follows:

$$h_H^{-1}\mathbb{E}[H_1'(y)|X] = F(\theta,y,X) + \frac{h_H^2}{2}\left(\int t^2 H'(t)dt\right)\frac{\partial^2 F(\theta,y,X)}{\partial^2 y} + o(h_H^2).$$

Thus, we get

$$\begin{aligned}\mathbb{E}[\widehat{F}_N(\theta,y,x)] &= \frac{1}{\mathbb{E}[K_1(\theta,x)]}\Bigg(\mathbb{E}[K_1(\theta,x)F(\theta,y,X)] + \left(\int t^2 H'(t)dt\right)\\ &\times \mathbb{E}\left[K_1(\theta,x)\frac{\partial^2 F(\theta,y,X)}{\partial^2 y}\right] + o(h_H^2)\Bigg).\end{aligned}$$

Let $\Psi_l(.,y) := \frac{\partial^l F(.,y,.)}{\partial^l y}$: for $l \in \{0,2\}$, since $\Psi_l(0) = 0$, we have

$$\begin{aligned}\mathbb{E}[K_1(\theta,x)\psi(X,y)] &= \Psi_l(x,y)\mathbb{E}[K_1(\theta,x)] + \mathbb{E}[K_1(\theta,x)(\Psi_l(X,y) - \Psi(x,y))]\\ &= \Psi(x,y)\mathbb{E}[K_1(\theta,x)] + \mathbb{E}[K_1(\theta,x)(\Psi_l(d_\theta(x,X)))]\\ &= \Psi_l(x,y)\mathbb{E}[K_1(\theta,x)] + \Psi_l'(0)\mathbb{E}[d_\theta(x,X)K_1(\theta,x)]\\ &+ o(\mathbb{E}[d_\theta(x,X)K_1(\theta,x)]).\end{aligned}$$

So

$$\begin{aligned}\mathbb{E}[\widehat{F}_N(\theta,y,x)] &= F(\theta,y,x)+\frac{h_H^2}{2}\frac{\partial^2 F(\theta,y,X)}{\partial^2 y}\int t^2H'(t)dt+o\left(h_H^2\frac{\mathbb{E}[d_\theta(x,X)K_1(\theta,x)]}{\mathbb{E}[K_1(\theta,x)]}\right)\\ &+ \Psi_0'(0)\frac{\mathbb{E}[d_\theta(x,X)K_1(\theta,x)]}{\mathbb{E}[K_1(\theta,x)]}+o\left(\frac{\mathbb{E}[d_\theta(x,X)K_1(\theta,x)]}{\mathbb{E}[K_1(\theta,x)]}\right).\end{aligned}$$

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

$$\frac{1}{\phi_{\theta,x}(h_K)}\mathbb{E}[d_\theta(x,X)K_1(\theta,x)]=M_0h_K+o(h_K)$$

and

$$\frac{1}{\phi_{\theta,x}(h_K)}\mathbb{E}[K_1(\theta,x)]\longrightarrow M_1.$$

Hence,

$$\mathbb{E}[\widehat{F}_N(\theta,y,x)]=F(\theta,y,x)+\frac{h_H^2}{2}\frac{\partial^2 F(\theta,y,X)}{\partial^2 y}\int t^2H'(t)dt+\Psi_o'(0)\frac{M_0}{M_1}h_K+o(h_H^2)+o(h_K)$$

Lemma 3.3. *Under the Assumptions (H1)-(H4), and (H7), as n goes to infinity, we have*

$$\mathbb{E}[\widehat{f}_N(\theta,y,x)]-f(\theta,y,x)=B_H^f(\theta,y,x)h_H^2+B_K^f(\theta,y,x)h_K+o(h_H^2)+o(h_K)$$

Proof of lemma 3.3. The proof of this lemma follows the steps as for proving lemma 3.2, to study $\mathbb{E}[\widehat{f}_N(\theta,y,x)]$ it suffices to write by an integration by part

$$\mathbb{E}[\widehat{f}_N(\theta,y,x)]=\frac{1}{\mathbb{E}[K_1]}\mathbb{E}[K_1\mathbb{E}[H_1\mid X]] \text{ with } \mathbb{E}[K_1\mathbb{E}[H_1\mid X]]=\int_{\mathbb{R}}H'(t)f^X(y-h_Ht)dt$$

Then we can follow to prove that

$$\mathbb{E}[\widehat{f}_N(\theta,y,x)]=f(\theta,y,x)+\frac{h_H^2}{2}\frac{\partial^2 f(\theta,y,X)}{\partial^2 y}\int t^2H'(t)dt+\psi_0'(0)\frac{M_0}{M_1}h_K+o(h_H^2)+o(h_K)$$

Lemma 3.4. *Under the hypotheses of Theorem 3.1*

$$\widehat{F}_D(\theta,x)-\widehat{F}_N(\theta,y,x)\longrightarrow 1-F(\theta,y,x),\ \textit{in probability}.$$

And

$$\left(\frac{nh_H\phi_{\theta,x}(h_K)}{\sigma_h^2(\theta,y,x)}\right)^{1/2}\left(\widehat{F}_D(\theta,x)-\widehat{F}_N(\theta,y,x)-1+\mathbb{E}[\widehat{F}_N(\theta,y,x)\right)=o_p(1).$$

Proof of lemma 3.4. It is clear that

$$\mathbb{E}\left[\widehat{F}_D(\theta,x)-\widehat{f}(\theta,y,x)-1+F(\theta,y,x)\right]\longrightarrow 0,$$

and

$$Var\left[\widehat{F}_D(\theta,x)-\widehat{f}(\theta,y,x)-1+F(\theta,y,x)\right]\longrightarrow 0,$$

then

$$\widehat{F}_D(\theta,x)-\widehat{f}(\theta,y,x)-1+F(\theta,y,x)\xrightarrow{\mathbb{P}} 0.$$

Moreover, the asymptotic variance of $\widehat{F}_D(\theta,x)-\widehat{f}_N(\theta,y,x)$ (see Djebbouri et al.(2015)), allows to obtain

$$\frac{nh_H\phi_{\theta,x}(h_K)}{\sigma_h^2(\theta,y,x)}Var\left[\widehat{F}_D(\theta,x)-\widehat{f}_N(\theta,y,x)-1+\mathbb{E}[\widehat{f}_N(\theta,y,x)]\right]\longrightarrow 0.$$

By combining the result with the fact that

$$\mathbb{E}\left[\widehat{F}_D(\theta,x)-\widehat{f}_N(\theta,y,x)-1+\mathbb{E}[\widehat{f}_N(\theta,y,x)]\right]\longrightarrow 0,$$

we obtain the claimed result.

4. Simulation study

We first construct the simulation of the explanatory functional variables. In the second part, we focus on the ability of the nonparametric functional regression to predict responses variable from functional predictors. Finally we illustrated the Monte Carlo methodology and we will test the efficiency of the asymptotic normality results in parallel with the practical experiment.

For this purpose, we consider the following process explanatory functional variables for $n=350$:

$$X_i(t)=1-\sin(2\Omega_i t)\alpha_i+\Omega_i t\,,\ \ \forall t\in[0,\pi]$$

where α_i and Ω_i are n independent real random variables (r.r.v.) uniformly distributed over $[0.3;2]$ ($resp.[1;3]$), t is assumed that these curves are observed on a discretization grid of 100 points in the interval. These functional variables are represented in the Figure 1

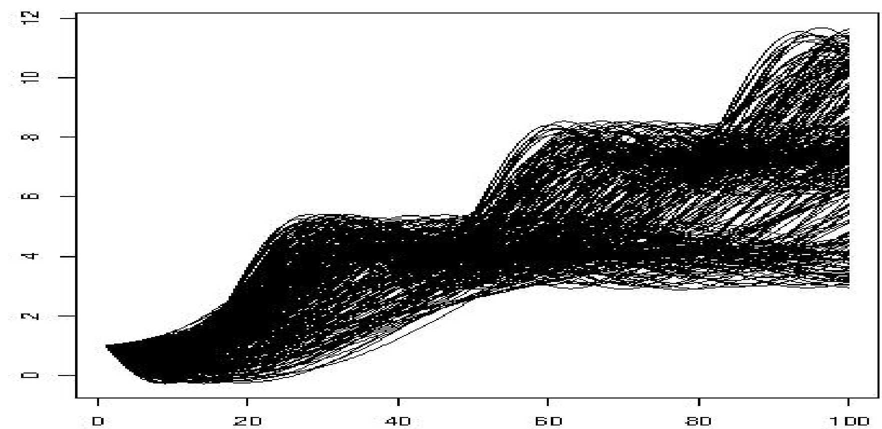

Figure 1. The curves $X_{i=1,\ldots,200}$

For response variables Y_i, we consider the following model for all $i = 1, \ldots n$ and $l = 1, \ldots n$:

$$Y = \lambda(< X_i, \theta_l >) + \varepsilon$$

where $\lambda(\mathscr{X}) = \int_0^{t_j} \frac{1}{1 - \mathscr{X}_i(v)^2} dv$ and ε is a centred normal variable and it is assumed to be independent of $(X_i)_i$. Our goal in this illustration is to show the usefulness of conditional density in the context of forecasting.

Now, we precise the different parameters of our estimators. Indeed, first of all, it is clear that the shape of the curves allows us to use

$$d(x_1, x_2) = \sqrt{\int_0^1 (x_1(t) - x_2(t))^2} \text{ ; } \forall x_1, x_2 \in \mathscr{H} \text{ where H is semi-metric}$$

We choose particularly the quadratic kernels defined by

$$K(x) = \frac{3}{2}(1-x^2)\; x \in [0,1]\; ;\; K_0(x) = \frac{3}{4}(1-x^2)\; x \in [-1,1] \text{ and } H(x) = \int_{-\infty}^{x} K_0(u)du.$$

In this illustration, we select the functional index θ on the set of eigenvectors of the empirical covariance operator.

$$\frac{1}{200}\sum_{i=1}^{200}(X_i - \bar{X})^t((X_i - \bar{X})).$$

Indeed, we recall that the ideas of Aitsaidi (2007) can be adapted to find a method of practical selection for θ. However, this adaptation in the case of the conditional density requires tools and additional preliminary results (see the discussion Attaoui *et al.* (2010) and Attaoui (2014)).

For this purpose, we divide our observations into two packets: learning sample $(X_i,Y_i)_{i=1,\ldots 200}$ and test sample and $(X_i,Y_i)_{i=201,\ldots 250}$(see, Ferraty et al. (2006)). For the choice of smoothing parameters h_K and h_H, we will adopt the selection criterion used by Ferraty and Vieu (2006) in the case of the kernel method for which h_K and h_H are obtained. by minimizing the next criterion

$$\text{for each} \quad X_i \quad \text{in the sample of the test} \qquad err(h_K,h_H) = |Y_{i^*} - \theta(X_{i^*})| \tag{7}$$

where i^* denotes the index of the nearest curve X_i from all the curves of the learning sample.

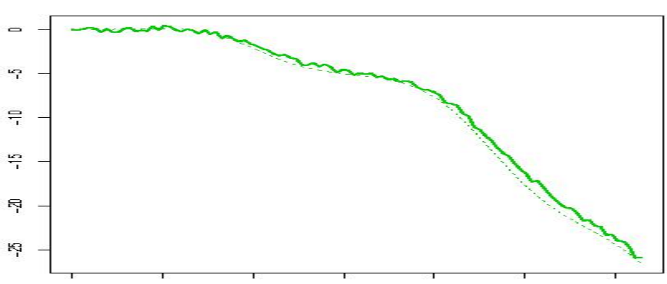

Figure 2. Predicted functional responses (solid lines); observed functional responses (dashed lines).

In this simulation study, we assume the quality of prediction by comparing the predicted functional responses (i.e. $\widehat{\lambda}(\theta,y,x)$ for any X in the testing sample) and the true functional operator (i.e. $\lambda(\theta,y,x)$) as in Figure. 2. However, if one wishes to assess the quality of prediction for the whole testing sample, it is much better to see what happens direction by direction. In other words, displaying the predictions onto the direction θ_l amounts to plotting the 50 points $(\lambda(< X_i,\theta_l >),\widehat{\lambda}(< X_i,\theta_l >))_{i=201,\ldots,250}$. Figure. 3 proposes a componentwise prediction graph for the two first components $(i.e. l = 1,2)$. The quality of componentwise predictions is quite good for each component.

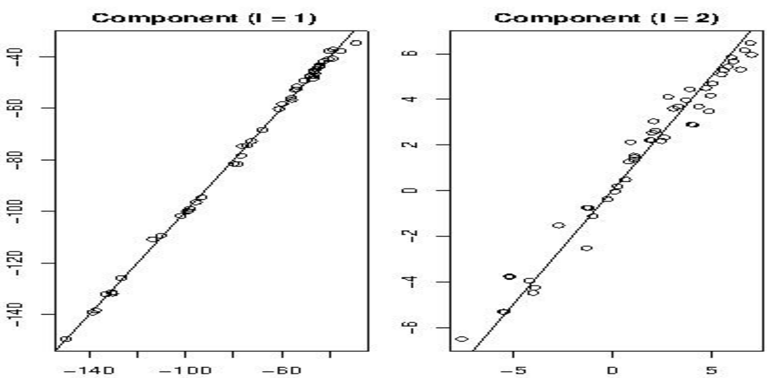

Figure 3. Representation of the prediction quality for each component.

For the next simulation algorithm we used:

- Simulate a sample of size n
- Calculate the smoothing parameters h_K and h_H that are varied over an interval [0,1] and which minimizes in 7
- We compute the quantities

$$(nh_H\phi_{\theta,x})^{1/2}(\widehat{\lambda}(\theta,y,x)-\lambda(\theta,y,x))$$

where $\widehat{\lambda}(\theta,y,x)$ is the functional hazard kernel estimator from the sample $(X_i,Y_i)_{i=1,\dots,200}$,

- compute a standard hazard function estimator by the kernel method .
- compare the estimated $\widehat{\lambda}(\theta,y,x)$ with the corresponding estimated $\lambda(\theta,y,x)$.

The obtained results are shown in Figure. 4.

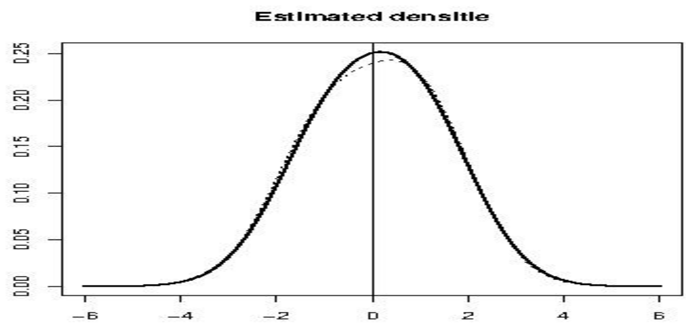

Figure 4. Representation of the asymptotic distribution of the hazard function estimator.

It can be seen that, both are very close and have good behaviours with respect to the standard normal distribution.

5. Conclusions

In this paper, we are mainly interested in the nonparametric estimation of the conditional hazard function estimator for a variable explanatory functionally conditioned to an actual response variable via a functional single index model. We show that the estimator provides good predictions under this model. One of the main contributions of this work is the choice of a semi-metric. Indeed, it is well known that, in non-parametric functional statistics, the semi-metric of the projection type is very important for increasing the concentration property. The functional index model is a special case of this family of semi-metrics because it is based on the projection on a functional direction which is important for the implementation of our method in practice.

Acknowledgements

The authors thank the associate editor and the anonymous reviewers for their valuable comments and suggestions, which substantially improved the quality of an earlier version of this article.

References

Ait Saidi, A., Ferraty, F., Kassa, P. and Vieu, P., (2005). Single functional index model for a time series. *Rev. Roumaine Math. Pures Appl.* 50, pp. 321–330.

Ait Saidi, A., Ferraty, F., Kassa, P. and Vieu, P., (2008). Cross-validated estimations in the single functional index model. *Statistics.* Vol. 42, No. 6, pp. 475–494.

Ait Saidi, A. and Mecheri, K., (2016). The conditional cumulative distribution function in single functional index model. *Comm. Statist. Theory Methods.* 45, pp. 4896–4911.

Arfi, M., (2013). Nonparametric Estimation for the Hazard Function. *Communications in Statistics - Theory and Methods* 42, pp. 2543–2550.

Attaoui, S., (2014). Strong uniform consistency rates and asymptotic normality of conditional density estimator in the single functional index modeling for time series data. *AStA - Advances in Statistical Analysis* 98, pp. 257–286.

Attaoui, S., laksaci, A. and Ould Said, F., (2011). A note on the conditional density estimate in the single functional index model. *Statist. Probab. Lett.,* 81, pp. 45–53.

Bagkavos, D., (2011). Local linear hazard rate estimation and bandwidth selection . *Annals of the Institute of Statistical Mathematics* 63, pp. 1019–1046.

Bouraine, M., Ait Saidi, A., Ferraty, F. and Vieu, P., (2010). Choix optimal de l'indice Multi-fonctionnel: Methode de validation croisée. *Rev. Roumaine Math.Pures Appl.*. 55, pp. 355–367.

Burba, F., Ferraty, F. and Vieu, P., (1996). Convergence of k nearest neighbor kernel estimator in nonparametric functional regression. *Comptes Rendus Mathematique*. 346, pp. 339–342.

Collomb, G., Hassani, S., Sarda, P. and Vieu, P., (1985). Estimation non parametrique de la fonction de hasard pour des observations dependentes. *Statistique et Analyse des Données* 10, pp. 42–49.

Dabo-Niang, S., Kaid, Z. and Laksaci, A., (2012). On spatial conditional mode estimation for a functional regressor. *Statistics and probability letters,*. 82, pp. 1413–1421.

Delecroix, M. , Yazourh , O., (1992). Estimation de la fonction de hazard en présence de censure droite. Méthode des fonctions orthogonales. *Statistique et Analyse des Données* 16, pp. 39–62 .

Estevèz-Pérez, G., Quintela-del-R, A. and Vieu, P., (2002). Convergence rate for cross-validatory bandwith in kernel hazard estimation from dependent samples. *J. Statist. Plann. Inference* 104, pp. 1–30.

Ferraty, F. and Vieu, P., (2006). *Nonparametric functional data analysis. Theory and Practice*. Springer Series in Statistics. New York.

Ferraty, F., Laksaci , A. and Vieu, P., (2005). Functional times series prediction via conditional mode, *C. R., Math. Acad. Sci. Paris* 340(5), pp. 389–392.

Ferraty, F., Laksaci , A. and Vieu, P., (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models, *Stat. Inf. Stoch. Proc.* 9, pp. 47–76.

Ferraty, F., Laksaci, A., Tadj, A., and Vieu, P.,(2010). Rate of uniform consistency for nonparametric estimates with functional variables, *Journal of Statistical Planning and Inference*. 140, pp. 335–352.

Härdle, W., Hall, P. and Ichumira, H., (1993). Optimal smoothing in single-index models *Ann. Statist.*. 27, pp. 157–178.

Hristache, M., Juditsky, A., and Spokoiny, V., (2001). Direct estimation of the index coefficient in the single-index model. *Ann. Statist.*, 29 , pp. 595–623.

Laksaci, A. and Mechab, B., (2010). Estimation nonparamétrique de la fonction de hasard avec variable explicative fonctionnelle : cas des données spatiales. *Rev. Roumaine Math. Pures Appl.* 55, pp. 35–51.

Laksaci, A. and Maref, F., (2009). Nonparametric estimation of conditional quantiles for functional and spatial dependent variables. *Comptes Rendus Mathematique*. 347, pp. 1075–1080.

Lecoutre, J. P. and Ould-Said, E.,(1992). Estimation de la densité et de la fonction de hasard conditionnelle pour un processus fortement mélangeant avec censure. *C.R. Math. Acad. Sci.Paris*. 314, pp. 295–300.

Li, J. and Tran, L. T., (2007). Hazard rate estimation on random fields, *J. Multivariate Anal.* 98 , pp. 1337–1355.

Masry, E., (2005). Nonparametric regression estimation for dependent functional data: asymptotic normality, *Stoch. Process. Appl.* 115, pp. 155–177.

Patil, P. N. , Wells M. T. and Marron J. S., (1993). Some heuristics of kernel based estimators of ratio functions. *Journal of Nonparametric Statistics* 4, pp. 203–209.

Quintela-Del-Río A., (2008). Hazard function given a functional variable: Non-parametric estimation under strong mixing conditions. *Journal of Nonparametric Statistics* 5, pp. 413–430.

Roussas, G. G., (1968). On some properties of nonparametric estimates of probability density functions *Bull. Soc. Math. Greece (N. S.)* 9, pp. 29–43 .

Tabti, H. and Ait Saidi, A., (2017). Estimation and simulation of conditional hazard function in the quasi-associated framework when the observations are linked via a functional single-index structure, *Communications in Statistics - Theory and Methods* 47, pp. 816–838 .

Tanner, M. A. and Wong, W. H., (1983). The Estimation of the Hazard Function from Randomly Censored Data by the Kernel Method *The Annals of Statistics* 11, pp. 989–993 .

Watson G. S. and Leadbetter M. R., (1964). Hazard Analysis. I *Biometrika* 51, pp. 175–184 .

Youndjé, É., Sarda, P. and Vieu, P., (1996). Optimal smooth hazard estimates. *Test.* 5, pp. 379–394.