Filomat 38:19 (2024), 6911–6935
https://doi.org/10.2298/FIL2419911M

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Strong uniform consistency of the conditional hazard function with functional explanatory variable in single functional index model under randomly truncated data

Souad Mekkaoui[a], Amina Angelika Bouchentouf[a], Abbes Rabhi[a]

[a]*Laboratory of Mathematics, University Djillali LIABES of Sidi Bel Abbes, Algeria*

Abstract. The aim of this paper is to estimate non-parametrically the conditional hazard function of a scalar response variable taking values in separable Hilbert space. The response variable is assumed to be left truncated data. We introduce kernel-type estimators for the conditional distribution function and conditional density. Then, we establish the pointwise almost complete convergence and the uniform almost complete convergence (with rate) of the kernel estimators, based on the single index structure. Additionally, the asymptotic properties of the conditional hazard function are provided. Finally, a simulation study is carried to illustrate the performance of our estimator.

1. Introduction

Functional Data Analysis (FDA) is a statistical branch that deals with the analysis of infinite-dimensional variables, such as images, sets, and curves. FDA has experienced remarkable growth in the past 20 years, partly driven by major advances in data collection technology that have ushered in the Big Data era. Literature related to Functional Data Analysis is extensive, the reader can refer to textbooks such as [15, 22, 28, 39] and broad surveys such as [8, 18, 28]. Current developments on various aspects of FDA are also reported in numerous review papers, notably regression analysis [9], functional principal component analysis [11], depth analysis [35], clustering [23], dependent functional data [26], spatial functional data [33], nonparametric modeling [30], semiparametric modeling [43], and testing [19].

Nonparametric estimation of the hazard function has been a topic of considerable interest in the literature. The early works in this area were done by Ferraty *et al.* [14], who proposed a kernel estimator for the conditional hazard function and derived its asymptotic properties and rates under different scenarios, such as censored and/or dependent variables. Then, Rabhi *et al.* [37] obtained the asymptotic mean square error of the conditional hazard function estimator. Merouan *et al.* [34] examined the asymptotic mean square error of the conditional hazard function using the linear method. Recently, Bellatrach *et al.* [7] presented the convergence rate of the hazard function with functional explanatory variable in the case of spatial data with k Nearest Neighbor method. Kebir *et al.* [25] dealt with estimation of the conditional hazard function with

2020 *Mathematics Subject Classification*. Primary 62G05; 62G08; 62G20; Secondary 62E20

Keywords. Conditional density, Conditional hazard function, Functional single index model, Lynden-bell estimator, Nonparametric estimators, Truncated data.

Received: 15 October 2023; Revised: 30 November 2023; Accepted: 24 January 2024

Communicated by Biljana Popović

Email addresses: souad.mekaoui75@gmail.com (Souad Mekkaoui), amina_bouchentouf@yahoo.com (Amina Angelika Bouchentouf), rabhi_abbes@yahoo.fr (Abbes Rabhi)

a recursive kernel from censored functional ergodic data. Omari *et al.* [36] studied the uniform convergence of nonparametric conditional hazard function in the single functional index model for dependent data. For more details on this subject, readers can refer to Ferraty *et al.* [13], Belabbaci *et al.* [6], Rabhi *et al.* [38], and Massim and Mechab [32].

In this paper, we consider the problem of nonparametric estimation of the conditional hazard function in a single index model, where the covariate is functional and the response variable is subject to left truncation. The hazard function measures the risk of an event over time, and it can vary depending on covariates such as age and gender. Truncation happens when we only observe time-to-event data within a certain interval, which can cause bias and complexity in the analysis. Truncated data are common in various domains, such as astronomy, economics, AIDS research, and reliability engineering (e.g., Woodroofe [45], Wang [44], Lawless [24] and Gardes and Stupfler [17]). Therefore, several methods have been proposed for nonparametric estimation and regression with truncated data [3, 10, 20].

Moreover, we note that the single-index models are a flexible and parsimonious way to handle high dimensional data by reducing the dimensionality of the covariate space. These models have attracted significant attention due to their relevance in various scientific fields such as biostatistics, economics, and medicine. Some of the notable works on the single-index models can be found in Ferraty *et al.* [12], Aït Saidi *et al.* [1, 2], Attaoui [4], Tabti and Ait Saidi [42], and Gagui and Chouaf [16].

The paper is organized as follows. In Section 2, we present our model and estimators. In Section 3, we introduce assumptions and state the main results, we establish the almost complete convergence with rates. In Section 4 the consistent uniform of the proposed estimators is studied. A simulation study is carried out to show the good behavior of our estimator in Section 5. Further, Section 6 is dedicated to the technical proofs. Finally, we conclude the paper in Section 6.

2. Model, notations and estimators

Throughout this paper, we shall denote by C, C', or $C_{\theta,x}$ certain constants generated in $\mathbb{R}_+^*$. Let (X_i, Y_i) for $i = 1, \ldots, N$ represent N random variables, independent and identically distributed as (X, Y) with values in $\mathcal{H} \times \mathbb{R}$, where the sample size N is deterministic but unknown, and $\mathcal{H}$ is a real Hilbert space with the norm $\|\cdot\|$ derived from an inner product $\langle\cdot,\cdot\rangle$. We consider the semi-metric d_θ associated with the single index $\theta \in \mathcal{H}$ defined by $\forall x_1, x_2 \in \mathcal{H}$, $d_\theta(x_1, x_2) = |\langle x_1 - x_2, \theta\rangle|$. The sample is not entirely observed; only n variables are observed, where $n \leq N$. Specifically, we assume that the lifetimes y_i, for $i = 1, \ldots, N$, are left-truncated by T_i, for $i = 1, \ldots, N$. In this model, (Y_i, T_i) is observed if $Y_i \geq T_i$, signifying that the random variable of interest Y is inferred from the random variable T. T is assumed to be independent of (X, Y).

We assume that Y and T have unknown distribution functions F and G, respectively. The observed sample size n is a known random variable, where $n \leq N$. It is noteworthy that if the samples (Y_i, T_i) for $i = 1, 2, \ldots, N$ are i.i.d., then the observed data (Y_i, T_i), for $i = 1, 2, \ldots, n$, are also i.i.d. (see Lemdani and Ould-Saïd [29]). As a consequence of the truncation sequence, the size of the observed sample, n, follows a binomial distribution with parameters N and $\mu := \mathbb{P}(Y \geq T)$.

It is evident that if $\mu = 0$, no data can be observed. Hence, we assume throughout this paper that $\mu > 0$.

It is worth noting that, by the strong law of large numbers (SLLN), as $N \to \infty$,

$$\mu_n = \frac{n}{N} \to \alpha, \ \mathbb{P} \text{ a.s.}$$

For any real distribution function L, the left and right endpoints of its support are denoted by

$$a_L = \inf\{t,\ L(t) > 0\} \text{ and } b_L = \sup\{t,\ L(t) < 1\}. \quad (1)$$

The kernel estimator of the conditional function is constructed from the observed variables (X_i, Y_i, T_i), $i = 1, 2, \ldots, n$, based on the Lynden-Bell estimator, G_n, of the cumulative distribution G of the random variable T. We assume that the variables of interest, Y and T, are independent. Following Stute and Wang

[41], the distribution functions of Y and T under the left-truncated condition are expressed respectively as follows:

$$F^{*}(y)=\mu^{-1}\int_{-\infty}^{y}G(u)dF(u) \text{ and } G^{*}(t)=\mu^{-1}\int_{-\infty}^{t}G(t\wedge u)dF(u),$$

where $t\wedge u=\min(t,u)$. These are estimated by

$$F_n^{*}(y)=\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{\{Y_i\le y\}} \text{ and } G_n^{*}(t)=\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{\{T_i\le t\}}.$$

Then, we define

$$\begin{aligned} C(y) &= \mathbb{P}(T\le y\le Y|Y\ge T)\\ &= G^{*}(y)-F^{*}(y)\\ &= \mu^{-1}G(y)\left(1-F(y)\right), \end{aligned}$$

which is empirically estimated by

$$C_n(y)=G_n^{*}(y)-F_n^{*}(y^{-})=\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{\{T_i\le y\le Y_i\}}.$$

The nonparametric maximum likelihood estimators of F and G are well-known as Lynden-Bell estimators (see Lynden-Bell [31]):

$$F_n(y)=1-\prod_{i:Y_i\le y}\left(\frac{nC_n(Y_i)-1}{nC_n(Y_i)}\right) \text{ and } G_n(y)=1-\prod_{i:T_i>y}\left(\frac{nC_n(T_i)-1}{nC_n(T_i)}\right).$$

He and Yang [21] demonstrated that μ can be estimated by $\mu_n=\frac{G_n(y)(1-F_n(y))}{C_n(y)}$, which is independent of y. And the asymptotic properties of F_n and G_n were examined by Woodroofe [45], showing that

$$\sup_{y>a_F}\left|\widetilde{F}_n(y)-F(y)\right|\longrightarrow 0 \text{ and } \sup_{t>a_G}\left|\widetilde{G}_n(t)-G(t)\right|\longrightarrow 0,$$

provided the identifiability conditions, $a_G\le a_F$, $b_G\le b_F$, and $\int_{a_F}^{\infty}\frac{1}{G}dF<\infty$, where a_L, b_L are the endpoints of the support of the distribution function L defined by (1).

Assume that the conditional distribution function of Y given X has a single structure and is given by

$$\forall\ y\in\mathbb{R},\quad F_\theta^x(y)=F\left(y|<X,\theta>=<x,\theta>\right)=F(\theta,y,x).$$

By saying this, we implicitly assume the existence of a regular version for the conditional distribution of Y given $\langle X,\theta\rangle=\langle x,\theta\rangle$.

If this distribution is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$, we will denote by $F^{(1)}(\theta,\cdot,x)=f(\theta,\cdot,x)$ the conditional density of Y given $\langle X,\theta\rangle=\langle x,\theta\rangle$, with the conditional density $f(\theta,y,x)$ given by $\forall y\in\mathbb{R}$, $f_\theta^x(y)=f\left(y|\langle X,\theta\rangle=\langle x,\theta\rangle\right)=f(\theta,y,x)=F^{(1)}(\theta,\cdot,x)$.

We also assume that the conditional hazard function of Y given $\langle X,\theta\rangle=\langle x,\theta\rangle$ denoted by $h_\theta^x(\cdot)$ exists and is given by

$$\forall\ y\in\mathbb{R},\ h_\theta^x(y):=h\left(y|\langle X,\theta\rangle=\langle x,\theta\rangle\right)=\frac{f(\theta,y,x))}{1-F(\theta,y,x)}.$$

Our main objective is to estimate the conditional hazard function $h_\theta^x(y)$ for a fixed θ, in the form

$$\widehat{h}_\theta^x(y)=\frac{\widehat{f}(\theta,y,x)}{1-\widehat{F}(\theta,y,x)} \text{ with } \widehat{F}(\theta,y,x)<1,$$

where $\widehat{F}(\theta, y, x)$ is the conditional cumulative distribution function estimator of $F(\theta, y, x)$, given by

$$\widehat{F}(\theta, y, x) = \frac{\sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(h_K^{-1}\langle x - X_i, \theta\rangle\right) H\left(h_H^{-1}(y - Y_i)\right)}{\sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(h_K^{-1}\langle x - X_i, \theta\rangle\right)}, \tag{2}$$

and $\widehat{f}(\theta, y, x) = \widehat{F}^{(1)}(\theta, y, x)$ is the conditional density estimator for $f(\theta, y, x)$, defined as

$$\widehat{f}(\theta, y, x) = \frac{\sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(h_K^{-1}\langle x - X_i, \theta\rangle\right) H'\left(h_H^{-1}(y - Y_i)\right)}{h_H \sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(h_K^{-1}\langle x - X_i, \theta\rangle\right)}, \tag{3}$$

where K is a Kernel function, H is a distribution function, H' is the derivative of H and $h_K := h_{K,n}$ (resp. $h_H := h_{H,n}$) is a sequence of positive real numbers that decreases to zero as n tends to infinity, G_n is the Lynden-Bell estimator. Note that all sums containing $G_n^{-1}(Y_i)$ are taken for i such that $G_n^{-1}(Y_i) \neq 0$. For the identifiability of the model, the readers may refer to Ferraty *et al.* [12].

Let N_x be a fixed neighborhood of x, where $x \in \mathcal{H}$ and $y \in \mathbb{R}$. Additionally, let $S_{\mathbb{R}}$ be a fixed compact set in $\mathbb{R}$, and define $B_\theta(x, h) = \mathbb{P}\{X \in \mathcal{H} : 0 < |\langle x - X, \theta\rangle| < h\}$.

We make the assumption that $0 = a_G < a_F$, $b_G < b_F$, and that T_i and (X_i, Y_i) for $1 \leq i \leq n$ are independent.

3. Main results and assumptions

3.1. Pointwise almost complete convergence

In this section, we present a pointwise almost complete estimation, along with the corresponding rate, of the conditional cumulative distribution, conditional density, and conditional hazard function. We introduce the following conditions, which ensure the good behaviours of the estimators $\widehat{F}(\theta, y, x)$ and $\widehat{f}(\theta, y, x)$.

(A1) For all $h_K > 0$, $\mathbb{P}(X \in B_\theta(x, h_K)) := \phi_{\theta,x}(h_K) > 0$ and $\phi_{\theta,x}(h_K) \to 0$ as $h_K \to 0$.

(A2) The conditional cumulative distribution $F(\theta, y, x)$ (resp. the conditional density $f(\theta, y, x)$) satisfies the Hölder condition:

For all $(y_1, y_2) \in S_{\mathbb{R}} \times S_{\mathbb{R}}$, for all $(x_1, x_2) \in N_x \times N_x$, and for all $\theta \in \Theta_{\mathcal{H}}$, there exist $a > 0$ and $b > 0$:

- (i) $\left|F(\theta, y_1, x_1) - F(\theta, y_2, x_2)\right| \leq C_{\theta,x}\left(\|x_1 - x_2\|^a + \left|y_1 - y_2\right|^b\right)$.
- (ii) $\left|f(\theta, y_1, x_1) - f(\theta, y_2, x_2)\right| \leq C_{\theta,x}\left(\|x_1 - x_2\|^a + \left|y_1 - y_2\right|^b\right)$.

(A3) (i) The kernel K is positive, with compact support $[0, 1]$, of class C^1 on $[0, 1]$, $K(0) > 0$, $K(1) > 0$, and its derivative K' is such that $-\infty < C_1 < K'(t) < C_2$ on $[0, 1]$.

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

(A4) The kernel H is a positive bounded Lipschitz-continuous function, and H' is a positive bounded Lipschitz-continuous function such that: $\int |z|^b H'(z)dz < \infty$ and $\int H'(z)dz = 1$.

(A5) The bandwidths h_H and h_K satisfy

- (i) $\lim_{n\to\infty} h_K = \lim_{n\to\infty} h_H = 0$.
- (ii) $\lim_{n\to\infty} \frac{\log(n)}{n h_H^j \phi_{\theta,x}(h_K)} = 0$ for $j = 0, 1$.
- (iii) $\lim_{n\to\infty} n^\gamma h_H = +\infty$ for some $\gamma > 0$.

(A6) $\exists \alpha < \infty, \forall (y, x) \in S_{\mathbb{R}} \times N_x$, $f(\theta, y, x) \leq \alpha$.

(A7) $\exists\beta > 0, \forall(y, x) \in S_{\mathbb{R}} \times N_x, F(\theta, y, x) \leq 1 - \beta.$

Remark 3.1. *In our methodology, assumption (A1) assumes a crucial role, often referred to as the 'concentration property' in finite-dimensional spaces. The purpose of (A2) is to regulate the smoothness of the functional space within our model. Conditions (A3) and (A4) are standard in functional estimation, applicable to both finite and infinite-dimensional spaces.*

Proposition 3.2. *Under assumptions (A1)-(A7), we have*

$$\left|\widehat{h}(\theta, y, x) - h(\theta, y, x)\right| = O\left(h_K^a + h_H^b\right) + O_{a.co}\left(\sqrt{\frac{\log n}{nh_H\phi_{\theta,x}(h_K)}}\right).$$

Proof. The proof relies on the following decomposition, applicable for any $x \in S_{\mathbb{R}}$, where C is a strictly positive real constant:

$$\widehat{h}(\theta, y, x) - h(\theta, y, x) = \frac{\widehat{f}(\theta, y, x)}{1 - \widehat{F}(\theta, y, x)} - \frac{f(\theta, y, x)}{1 - F(\theta, y, x)},$$

which can be expressed as:

$$\begin{aligned}
&\left|\widehat{h}(\theta, y, x) - h(\theta, y, x)\right| \\
&\leq \frac{1}{\left|1 - \widehat{F}(\theta, y, x)\right|}\left|\widehat{f}(\theta, y, x) - f(\theta, y, x)\right| \\
&\quad + \frac{\left|h(\theta, y, x)\right|}{\left|1 - \widehat{F}(\theta, y, x)\right|}\left|\widehat{F}(\theta, y, x) - F(\theta, y, x)\right| \\
&\leq C\frac{\left|\widehat{f}(\theta, y, x) - f(\theta, y, x)\right| + \left|\widehat{F}(\theta, y, x) - F(\theta, y, x)\right|}{\left|1 - \widehat{F}(\theta, y, x)\right|}.
\end{aligned} \tag{4}$$

The result of Proposition 3.2 follows from the subsequent intermediate results, the proofs of which are provided in the Section 6. □

Theorem 3.3. *Under assumptions (A1)-(A5), we have*

$$\sup_{y\in S_{\mathbb{R}}}\left|\widehat{F}(\theta, y, x) - F(\theta, y, x)\right| = O\left(h_K^a + h_H^b\right) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right),$$

and

$$\sup_{y\in S_{\mathbb{R}}}\left|\widehat{f}(\theta, y, x) - f(\theta, y, x)\right| = O\left(h_K^a + h_H^b\right) + O_{a.co}\left(\sqrt{\frac{\log n}{nh_H\phi_{\theta,x}(h_K)}}\right).$$

Proof. Let $x \in \mathcal{H}$, $y \in \mathbb{R}$, and $i = 1, \ldots, n$. Define

$$K_i(x, \theta) := K(h_K^{-1}\langle x - X_i, \theta\rangle), \text{ for } j = 0, 1;\ H_i^{(j)}(y) := H^{(j)}\left(h_H^{-1}(y - Y_i)\right).$$

Note that the estimator defined in (3) can be written as:

$$\widehat{F}^{(j)}(\theta, y, x) = \frac{\frac{\mu_n}{nh_H^{(j)}\mathbb{E}[K_1(\theta,x)]}\sum_{i=1}^n G_n^{-1}(Y_i)K_i(\theta, x)H_i^{(j)}(y)}{\frac{\mu_n}{n\mathbb{E}[K_1(\theta,x)]}\sum_{i=1}^n G_n^{-1}(Y_i)K_i(\theta, x)} := \frac{\widehat{F}_N^{(j)}(\theta, y, x)}{\widehat{F}_D(\theta, x)}.$$

This proof is based on the following decomposition for all $j = 0, 1$:

$$\begin{aligned}\left|\widehat{F}^{(j)}(\theta,y,x) - F^{(j)}(\theta,y,x)\right| \leq & \frac{1}{\widehat{F}_D(\theta,x)}\left|\widehat{F}_N^{(j)}(\theta,y,x) - \widetilde{F}_N^{(j)}(\theta,y,x)\right| \\ & + \frac{1}{\widehat{F}_D(\theta,x)}\left|\widetilde{F}_N^{(j)}(\theta,y,x) - \mathbb{E}\left(\widetilde{F}_N^{(j)}(\theta,y,x)\right)\right| \\ & + \frac{1}{\widehat{F}_D(\theta,x)}\left|F^{(j)}(\theta,y,x) - \mathbb{E}\left(\widetilde{F}_N^{(j)}(\theta,y,x)\right)\right| \\ & + \frac{F^{(j)}(\theta,y,x)}{\widehat{F}_D(\theta,x)}\left|\widehat{F}_D(\theta,x) - \widetilde{F}_D(\theta,x)\right| \\ & + \frac{F^{(j)}(\theta,y,x)}{\widehat{F}_D(\theta,x)}\left|\widetilde{F}_D(\theta,x) - \mathbb{E}\left(\widetilde{F}_D(\theta,x)\right)\right| \\ & + \frac{F^{(j)}(\theta,y,x)}{\widehat{F}_D(\theta,x)}\left|1 - \mathbb{E}\left(\widetilde{F}_D(\theta,x)\right)\right|, \qquad (5)\end{aligned}$$

where

$$\widetilde{F}_N^{(j)}(\theta,y,x) = \frac{\mu}{nh_H^j\mathbb{E}\left[K_1(\theta,x)\right]}\sum_{i=1}^{n} G^{-1}(Y_i)K_i(\theta,x)H_i^{(j)}(y).$$

So, the result of Theorem 3.3 is a consequence of the following intermediate results, with proofs provided in Section 6. □

Lemma 3.4. *Under assumptions (A1), (A3) and (A5), we have*

$$\mathbb{E}\left(\widetilde{F}_D(\theta,x)\right) = 1,$$

and

$$\left|\widehat{F}_D(\theta,x) - \widetilde{F}_D(\theta,x)\right| = O_{a.s}\left(\sqrt{\frac{1}{n}}\right). \qquad (6)$$

Lemma 3.5. *Under assumptions (A1)-(A5), we have for all $j = 0, 1$*

$$\sup_{y\in S_{\mathbb{R}}}\left|\mathbb{E}\left[\widetilde{F}_N^{(j)}(\theta,y,x)\right] - F^{(j)}(\theta,y,x)\right| = O\left(h_K^a + h_H^b\right), \qquad (7)$$

$$\sup_{y\in S_{\mathbb{R}}}\left|\widehat{F}_N^{(j)}(\theta,y,x) - \widetilde{F}_N^{(j)}(\theta,y,x)\right| = O_{a.s}\left(\sqrt{\frac{1}{n}}\right), \qquad (8)$$

and

$$\sup_{y\in S_{\mathbb{R}}}\left|\widetilde{F}_N^{(j)}(\theta,y,x) - \mathbb{E}\left[\widetilde{F}_N^{(j)}(\theta,y,x)\right]\right| = O_{a.co}\left(\sqrt{\frac{\log n}{nh_H^{(j)}\phi_{\theta,x}(h_K)}}\right). \qquad (9)$$

Lemma 3.6. *Under assumptions (A1), (A3) and (A5), we have*

$$\left|\widetilde{F}_D(\theta,x) - \mathbb{E}\left[\widetilde{F}_D(\theta,x)\right]\right| = O_{a.co}\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right), \qquad (10)$$

and

$$\left|\widehat{F}_D(\theta,x) - 1\right| = O_{a.co}\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right). \qquad (11)$$

Corollary 3.7. *Under assumptions (A1), (A3) and (A5), we have:*

$$\sum_{n\geq 1} \mathbb{P}\left(\widehat{F}_D(\theta,x) \leq \frac{1}{2}\right) < \infty.$$

4. Uniform almost complete convergence

This section is dedicated to deriving the uniform version of Proposition 3.2 and Theorem 3.3. In addition to the conditions introduced previously, we require the following:

Firstly, we assume that $S_{\mathbb{R}}$ is a compact subset of $\mathbb{R}$ and $S_{\mathcal{H}}, \Theta_{\mathcal{H}}$ (the spaces of parameters) are such that $S_{\mathcal{H}} \bigcup_{k=1}^{d_n^{S_n}} B(x_k, r_n)$ and $\Theta_{\mathcal{H}} \subset \bigcup_{m=1}^{d_n^{\Theta_n}} B(\theta_m, r_n)$ with $x_k, \theta_m \in \mathcal{H}$, and $r_n, d_n^{\Theta_n}, d_n^{S_n}$ are sequences of positive real numbers that tend to infinity as $n \to +\infty$. Furthermore, we need the following assumptions:

(U1) There exists a differentiable function $\phi(\cdot)$ such that $\forall\ (x,\theta) \in S_{\mathcal{H}} \times \Theta_{\mathcal{H}}$,

$$0 < C\phi(h) \leq \phi_{\theta,x}(h) \leq C'\phi(h) < \infty \ \text{ and } \ \exists \eta_0 > 0,\ \forall \eta < \eta_0,\ \phi'(\eta) < C.$$

(U2) $\forall (y_1,y_2) \in S_{\mathbb{R}} \times S_{\mathbb{R}}$, $\forall (x_1,x_2) \in S_{\mathcal{H}} \times S_{\mathcal{H}}$ and $\forall \theta \in \Theta_{\mathcal{H}}$ for $j = 0,1$

$$\left|F^{(j)}(\theta,y_1,x_1) - F^{(j)}(\theta,y_2,x_2)\right| \leq C_{\theta,x}\left(\|x_1 - x_2\|^a + \left|y_1 - y_2\right|^b\right).$$

(U3) The kernel K satisfies (A3) and Lipschitz's condition holds

$$\left|K(x) - K(y)\right| \leq C\left\|x - y\right\|.$$

(U4) For $r_n = O\left(\frac{\log n}{n}\right)$, the sequences $d_n^{S_{\mathcal{H}}}$ and $d_n^{\Theta_{\mathcal{H}}}$ satisfy:

$$\begin{cases} (i)\ \dfrac{(\log n)^2}{n\phi(h_K)} < \log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}} < \dfrac{n\phi(h_K)}{\log n}, \\ (ii)\ \sum_{n=1}^{\infty} n^{1/2\gamma}(d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}})^{1-\beta} < \infty, \quad for \quad some \quad \beta > 1, \\ (iii)\ n\phi(h_K) = O\left((\log n)^2\right). \end{cases}$$

(U5) For some $\gamma \in (0,1)$, $\lim_{n\to\infty} n^{\gamma} h_H = \infty$, and for $r_n = O\left(\frac{\log n}{n}\right)$, the sequences $d_n^{S_{\mathcal{H}}}$ and $d_n^{\Theta_{\mathcal{H}}}$ satisfy:

$$\begin{cases} (i)\ \dfrac{(\log n)^2}{nh_H\phi(h_K)} < \log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}} < \dfrac{nh_H\phi(h_K)}{\log n}, \\ (ii)\ \sum_{n=1}^{\infty} n^{(3\gamma+1)/2}(d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}})^{1-\beta} < \infty, \quad for \quad some \quad \beta > 1, \\ (iii)\ nh_H\phi(h_K) = O\left((\log n)^2\right). \end{cases}$$

Proposition 4.1. *Under assumptions (U1)-(U5), we get*

$$\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \sup_{y\in S_{\mathbb{R}}} \left|\widehat{h}(\theta,y,x) - h(\theta,y,x)\right| = O\left(h_K^a + h_H^b\right) + O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right)$$

Proof. The proof relies on the decomposition given in (4). The stated results directly follow from Theorem 4.3 and (17) of Lemma 4.6. Proposition 4.1 and Corollary 4.2 can be deduced from the subsequent intermediate results, which represent the uniform versions of Proposition 3.2. □

Corollary 4.2. *Under assumptions of Proposition 4.1, we have*

$$\sup_{x\in S_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\left|\widehat{h}(\theta,y,x)-h(\theta,y,x)\right| = O\left(h_K^a+h_H^b\right)+O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}}{nh_H\phi\left(h_K\right)}}\right).$$

Theorem 4.3. *Under assumptions (A1), (A3)-(A4) and (U1)-(U5), as n goes to infinity, we have*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\left|\widehat{F}(\theta,y,x)-F(\theta,y,x)\right| = O\left(h_K^a+h_H^b\right)+O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right),$$

and

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\left|\widehat{f}(\theta,y,x)-f(\theta,y,x)\right| = O\left(h_K^a+h_H^b\right)+O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right).$$

Proof. The proof is grounded in the decomposition given by (5) and relies on the subsequent intermediate results. □

Lemma 4.4. *Under assumptions (U1)-(U2) and (A5), we have*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\left|\mathbb{E}\left[\widetilde{F}_N(\theta,y,x)\right]-F(\theta,y,x)\right| = O\left(h_K^a+h_H^b\right),$$

and

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\left|\mathbb{E}\left[\widetilde{f}_N(\theta,y,x)\right]-f(\theta,y,x)\right| = O\left(h_K^a+h_H^b\right).$$

Lemma 4.5. *Under assumptions (A3)-(A5),(U1) and (U4)-(U5), we have*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\left|\widehat{F}_D(\theta,x)-\widetilde{F}_D(\theta,x)\right| = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}+\log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right), \tag{12}$$

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\left|\widehat{F}_N(\theta,y,x)-\widetilde{F}_N(\theta,y,x)\right| = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right), \tag{13}$$

and

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\left|\widehat{f}_N(\theta,y,x)-\widetilde{f}_N(\theta,y,x)\right| = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right). \tag{14}$$

Lemma 4.6. *Under assumptions (U1)-(U5), we have*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\left|\widetilde{F}_N(\theta,y,x)-\mathbb{E}\left[\widetilde{F}_N(\theta,y,x)\right]\right| = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right), \tag{15}$$

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\left|\widetilde{f}_N(\theta,y,x)-\mathbb{E}\left[\widetilde{f}_N(\theta,y,x)\right]\right| = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right), \tag{16}$$

$$\sup_{\theta \in \Theta_{\mathcal{H}}} \sup_{x \in S_{\mathcal{H}}} \left| \widetilde{F}_D(\theta, x) - \mathbb{E}\left[\widetilde{F}_D(\theta, x)\right] \right| = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}} \right), \tag{17}$$

and

$$\sup_{\theta \in \Theta_{\mathcal{H}}} \sup_{x \in S_{\mathcal{H}}} \left| \widehat{F}_D(\theta, x) - 1 \right| = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}} \right). \tag{18}$$

Corollary 4.7. *Under assumptions (A1), (U1) and (U3), we have*

$$\sum_{n \geq 1} \mathbb{P}\left(\inf_{\theta \in \Theta_{\mathcal{H}}} \inf_{x \in S_{\mathcal{H}}} \widehat{F}_D(\theta, x) \leq \frac{1}{2} \right) < \infty.$$

In the particular case, where the functional single index is fixed we get the following result.

Corollary 4.8. *Under the hypotheses of Theorem* 4.3*, as n goes to infinity, we have*

$$\sup_{x \in S_{\mathcal{H}}} \sup_{y \in S_{\mathbb{R}}} \left| \widehat{F}(\theta, y, x) - F(\theta, y, x) \right| = O\left(h_K^a + h_H^b\right) + O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}}{n\phi\left(h_K\right)}} \right),$$

and

$$\sup_{x \in S_{\mathcal{H}}} \sup_{y \in S_{\mathbb{R}}} \left| \widehat{f}(\theta, y, x) - f(\theta, y, x) \right| = O\left(h_K^a + h_H^b\right) + O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}}{n h_H \phi\left(h_K\right)}} \right).$$

These last results are only uniform version of Theorem 3.3.

5. Simulation study

In this section, we present two simulation examples to illustrate the finite-sample characteristics of our proposed estimator. A comparative analysis is conducted with respect to the kernel-type nonparametric conditional hazard estimator (NP) examined in the work of Ferraty *et al.* [14].

The empirical investigation is designed to assess the finite-sample properties of the nonparametric conditional hazard function estimator within the context of functional data in a single functional index model for randomly left-truncated data. Our focus is on evaluating the performance of this innovative estimator, rooted in the single functional index model (SFIM), in contrast to the established kernel-type nonparametric conditional hazard estimator (NP).

In recent years, the single functional index model has emerged as a promising approach for efficiently handling functional data. The objective of this study is to evaluate the performance of the proposed nonparametric conditional hazard function estimator in the context of random left-truncated data. Additionally, we aim to conduct a comparative analysis with the TNPFDA (truncated non-parametric functional data analysis) estimator.

Simulated functional data will be generated according to the single functional index model for various scenarios, including different sample sizes, functional index structures, and left-truncation rates.

In each simulation scenario, we will compute the estimated hazard function using both the proposed estimator and the kernel-type nonparametric estimator (NP). Performance metrics will be calculated to facilitate a thorough comparison of these estimators. Our analysis includes a comparative assessment between our proposed model TFSIM (functional single index model with truncated data) and TNPFDA (truncated non-parametric functional data analysis). Specifically, for TNPFDA, where the distribution of the regression model is known and conventional, we examine the performance of our conditional hazard function estimator with respect to this distribution. To provide insights into the behavior of the estimator, we employ the mean square error (MSE) as a key metric. This section of the paper is dedicated to a detailed comparison of the estimation of the conditional hazard function between our TFSIM model and TNPFDA, as defined as follows:

$$h_n(x|y) = \frac{f_n(y|x)}{1 - F_n(y|x)}, \tag{19}$$

where

$$F_n(y|x) = \frac{\sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(h_K^{-1} d(x, X_i)\right) H\left(h_H^{-1}(y - Y_i)\right)}{\sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(h_K^{-1} d(x, X_i)\right)},$$

$$f_n(y|x) = \frac{\sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(h_K^{-1} \langle x - X_i, \theta \rangle\right) H'\left(h_H^{-1}(y - Y_i)\right)}{h_H \sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(h_K^{-1} d(x, X_i)\right)}.$$

The typical scenario involves the inherent uncertainty surrounding the single functional index, denoted as $\theta \in \mathcal{H}$, necessitating practical estimation. Existing literature on single functional regression models discusses various estimation methods, such as cross-validation or maximum-likelihood approaches, as exemplified in the work of Aït Saidi *et al.* [2] and related references. Another approach, employed in this section, entails selecting $\theta(t)$ from the eigenfunctions of the covariance operator

$$\mathbb{E}\left[(X' - \mathbb{E}(X')) < X', . >_{\mathcal{H}}\right],$$

where $X(t)$ represents, for instance, a diffusion-type process defined on a real interval $[a, b]$, and $X'(t)$ is its first derivative (see Attaoui and Ling [5]). Utilizing a training sample $\mathcal{L}$, the empirical version

$$\frac{1}{|\mathcal{L}|} \sum_{i \in \mathcal{L}} (X'_i - \mathbb{E}X')^t (X'_i - \mathbb{E}X'),$$

allows the estimation of the covariance operator. Subsequently, applying the principle component analysis method provides a discretized form of the eigenfunctions $\theta_i(t)$. Designating $\theta^\star$ as the first eigenfunction corresponding to the highest eigenvalue of the empirical covariance operator, it serves as a surrogate for θ in the simulation steps for computing the estimator of the conditional hazard function.

Now, we employ our methodology to assess the effectiveness of predictors in a finite sample. Specifically, our focus is on the classical nonparametric functional regression model given by

$$Y_i = R(\langle \theta, X_i \rangle) + \epsilon_i, \quad i = 1, \ldots, n,$$

where ϵ_i is normally distributed with variance 0.5. The functional covariate X is assumed to be a diffusion process defined on $[0, 1]$ and generated by the following equation:

$$X(t) = A(2 - \cos(\pi t W)) + (1 - A)\cos(\pi t W), \quad t \in [0, 1],$$

where $W \rightsquigarrow \mathcal{U}(0, 1)$ and $A \rightsquigarrow$ Bernoulli(1/2). The following steps are undertaken: we fix the random size n (bearing in mind that n is known):

Step 1: Compute the inner product: $< \theta^*, X_1 >, \ldots, < \theta^*, X_n >$, generate independently the variables $\epsilon_1, \ldots, \varepsilon_n$, then simulate the response variables $Y_i = r(< \theta^*, X_i >) + \varepsilon_i$, where $r(< \theta^*, X_i >) = \exp(10(< \theta^*, X_i > -0.25))$ and generate independently the variables $\epsilon_1, \ldots, \varepsilon_n$.

Step 2: Generate the random variables T_1, $X_1(t)$, $t \in [0, 1]$, in the following manner: $T_1 \rightsquigarrow \mathcal{N}(\mu, 1)$, and adapt μ to get a different rate of truncation. $X_1(t)$ is generated as indicated before. Furthermore, simulate $\epsilon_1 \rightsquigarrow \mathcal{N}(0, 0.5)$.

Table 1: MSE results for TFSIM and TNPFDA methods according to TR

	$TR = 0\%$	$TR = 47\%$	$TR = 80\%$
MSE (TFSIM)	0.00841	0.06175	0.01084
MSE (TNPFDA)	0.0191	0.091	0.1181

Step 3: Calculate $Y_1 = R(\langle\theta, X_1\rangle) + \epsilon_1$, where $R(X_1(t)) = \frac{1}{4}\int_0^1 \left(X_1'(t)\right)^2 dt$, and ϵ_1 is as indicated before.

Step 4: Test: Start with the configuration $N = 0$, $j = 0$, although $j \leq n$: Put $N = N + 1$. Test: if $Y_1 < T_1$ reject the triple $(X_1(t), Y_1, T_1)$. Otherwise, retain the triple $(X_1(t), Y_1, T_1)$. At the end, get a deterministic N, which allows obtaining the rate of truncation $\tau = n/N$. More precisely, the rate of the observed triplet. Continue the process until $n = 100$. Obtain the random vectors $(X_i(t), Y_i, T_i)$, $i = 1, \ldots, 100$. Then calculate the Lynden-Bell estimate for the observed pair (Y_i, T_i), $i = 1, \ldots, n$.

Step 5: Divide the data into randomly selected subsets $\mathcal{J}$ and $\mathcal{L}$: $(X_i, Y_i, \delta_i)_{i\in\mathcal{L}}$ training sample $(X_j, Y_j, \delta_j)_{j\in\mathcal{J}}$ test sample.

Step 6: For each X_j in the test sample, set: $j_\star := \arg\min_{i\in\mathcal{L}} d_\theta(X_i, X_j)$.

Step 7: To be more precise evaluate the prediction errors given by

$$SSR = \frac{1}{|\mathcal{J}|}\sum_{j\in\mathcal{J}}\left(Y_j - \widehat{Y}_j\right)^2,$$

where $\widehat{Y}_j$ is a predictor of Y_j obtained either semi-parametrically by $\widehat{h}(\theta, y, x)$ or nonparametrically via $h_n(y|x)$.

Moreover, certain tuning parameters need specification. The kernel $K(\cdot)$ is selected as the quadratic function defined by $K(u) = \frac{3}{2}\left(1 - u^2\right)\mathbf{1}_{[0,1]}$, and the cumulative distribution function (CDF) $H(u) = \int_{-\infty}^{u} \frac{3}{4}\left(1 - z^2\right)\mathbf{1}_{[-1,1]}(z)\,dz$.

The semi-metric $d(\cdot,\cdot)$ will be determined based on the choice of the functional space $\mathcal{H}$ discussed in the scenarios below. It is widely acknowledged that one of the pivotal parameters in semi-parametric models is the smoothing parameters, which play a crucial role in shaping the link function between the response and the covariate.

Example 5.1. *We fix the truncation percentage at $\mu = 2$ and explore varying sample sizes of $n = 100$ and 300. In each scenario, the data is partitioned into two subsets: a learning sample and a test sample. Predicted values are computed for all $i \in \mathcal{L}$ using our estimator, which is derived from the training sample. Subsequently, we calculate predictor values using the test sample.*

The ensuing figures depict the plotted predicted values estimated by our estimator against the true values. The continuous line represents the ideal prediction. The efficacy of the prediction method is typically assessed by how closely the plotted points align with this continuous line.

Figures; Fig. 1 and Fig. 2 depict the curves and predictions evaluated using the Mean Squared Error (MSE). It is evident that the quality of fit improves with the sample size n.

We compare our model TFSIM (functional single index model with truncated data) with the TNPFDA (truncated non-parametric functional data analysis) model, where the distribution of the regression model is known and usual. We evaluate the performance of our estimator of the conditional hazard function by computing its mean square error and comparing it with the TNPFDA estimator defined in (19). The obtained results are in Table 1.

It is evident that the TFSIM estimator outperforms the kernel estimator (TNPFDA). Furthermore, the quality of both TNPFDA and TFSIM methods deteriorates as the truncation rate (TR) increases.

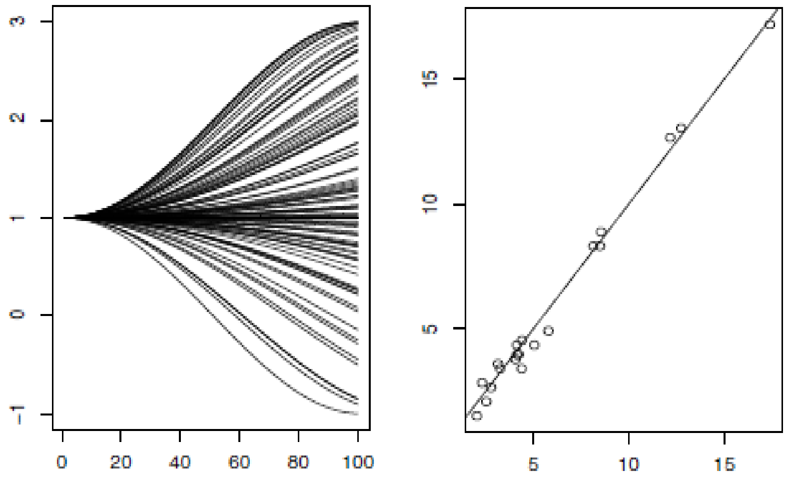

Figure 1: A sample of 100 curves, MSE=0.25

Table 2: MSE comparaison for TFSIM and TNPFDA methods for the three samples sizes (n)

n	MSE (TFSIM)	MSE (TNPFDA)
50	0.43	0.57
150	0.3	0.48
300	0.21	0.35

To check more the quality of our estimator, we compare the two methods in a different way in the next example.

Example 5.2. *We vary the sample size* $n = 50,\ 150,\ 300$ *and we consider the functional covariate* $X_1(t)$ *generated in the following way*

$$X_1(t) = 2 - \cos\left(W\left(t - \frac{2\pi}{3}\right)\right),\ t \in \left[0,\ \frac{2\pi}{3}\right],$$

where $W \rightsquigarrow \mathcal{N}(0,\ 1)$*. The scalar response is defined as*

$$Y_1 = R\left(\langle\theta, X_1\rangle\right) + \epsilon_1,$$

where X_1 *and* ϵ_1 *are independent, the error* $\epsilon_1 \rightsquigarrow \mathcal{N}(0,\ 0.1)$ *and nonlinear regression function is considered such that*

$$R(X) = \frac{1}{4}\exp\left\{2 - \frac{1}{\left(\int_0^1 X'(t)dt\right)^2}\right\}.$$

Figure 6 depicts a sample of 300 *curves representing a realization of the functional random variable X.*

In this model, we implement the truncation mechanism based on the sample $(X_i, Y_i, T_i)_{1\le i\le n}$*, where the truncation variable* T_1 *follows a normal distribution* $\mathcal{N}(0, 2)$*. This choice is made to control the percentage of truncation.*

For each sample size case, we partitioned our data into a learning sample and a test sample, following the methodology employed in the previous example. To assess the performance of our estimator, we computed the mean squared error (MSE) for both TFSIM and TNPFDA methods (see Table 2).

It is evident from Table 2 that the TFSIM estimator consistently outperforms the kernel TNPFDA estimator across various sample sizes. Moreover, the quality of both estimators improves with larger sample sizes.

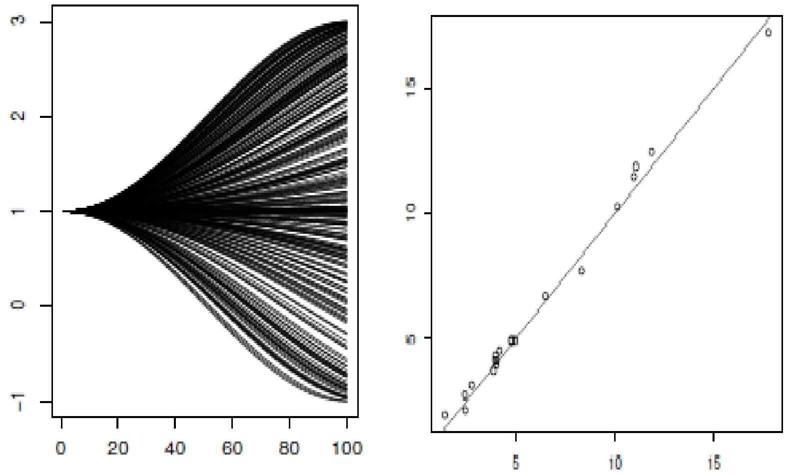

Figure 2: A sample of 300 curves, MSE=0.2

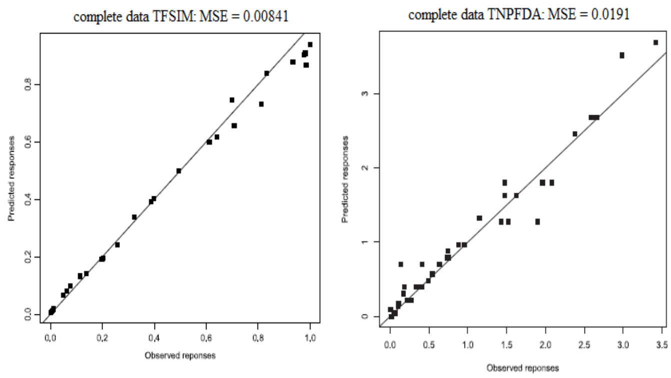

Figure 3: TR=0%, complete data

In summary, the results from both examples demonstrate the effectiveness of our estimator. Notably, the FSIM method exhibits superior performance compared to the NPFDA kernel method, even in the case of truncated data.

6. Proofs of technical lemmas

First of all, we state the following lemma which can be found in the monograph by (Ferraty and Vieu [15]). It's proof thus is omitted.

Lemma 6.1. *Let* $(Z_i)_i$ *a sequence of i.i.d. center random variables such that*

$$\forall\, m \geq 2,\ \exists\, C_m > 0, \mathbb{E}\left(\left|Z_1^m\right|\right) < C_m a^{2(m-1)}.$$

Then

$$\forall \varepsilon > 0,\ \mathbb{P}\left(\left|\sum_{i=1}^{m} Z_i\right| > \varepsilon n\right) \leq 2\exp\left(-\frac{n\varepsilon^2}{2a^2(1+\varepsilon)}\right).$$

Proof. [Proof of Lemma 3.4]

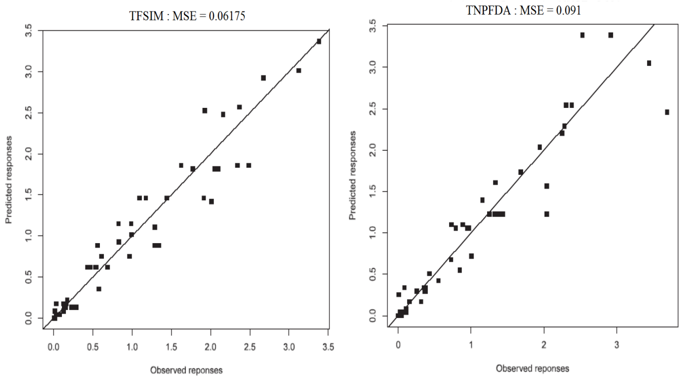

Figure 4: TR=47%

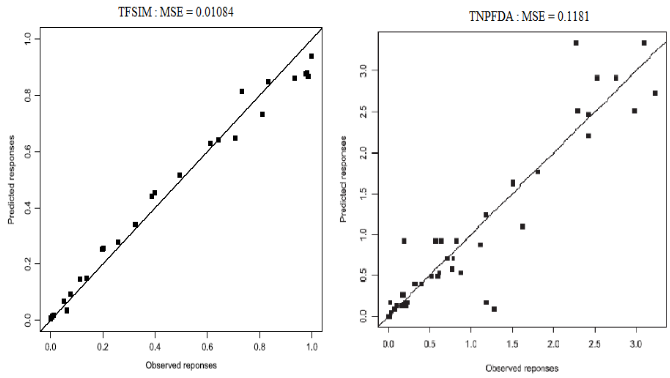

Figure 5: TR=80%

(i) First, we have

$$\begin{aligned}\mathbb{E}\left(\widetilde{F}_D(\theta,x)\right) &= \frac{\mu}{\mathbb{E}\left[K_1(\theta,x)\right]}\mathbb{E}\left(\mathbb{E}\left[K_1(\theta,x)\frac{I_{Y_1\geq T_1}}{\mu G(Y_1)}\right]\mid\langle X_1,\theta\rangle,Y_1\right)\\ &= \frac{1}{\mathbb{E}\left[K_1(\theta,x)\right]}\mathbb{E}\left[K_1(\theta,x)\right]=1.\end{aligned}$$

(ii) For establish (6), we have

$$\begin{aligned}\left|\widehat{F}_D(\theta,x)-\widetilde{F}_D(\theta,x)\right| &= \left|\frac{\mu_n\sum_{i=1}^n G_n^{-1}(Y_i)K_i(\theta,x)}{n\mathbb{E}\left[K_1(\theta,x)\right]}-\frac{\mu\sum_{i=1}^n G^{-1}(Y_i)K_i(\theta,x)}{n\mathbb{E}\left[K_1(\theta,x)\right]}\right|\\ &\leq \left(\frac{\left|\mu_n-\mu\right|}{G_n(a_F)}-\mu\frac{\left|G_n(y)-G(y)\right|}{G(a_F)G_n(a_F)}\right)\left|\frac{1}{n\mathbb{E}\left[K_1(\theta,x)\right]}\sum_{i=1}^n K_i(\theta,x)\right|.\end{aligned}$$

From theorem 3.2 of (He and Yang [21]), we have $|\mu_n-\mu| = O_{a.s}\left(n^{-1/2}\right)$, while Remark 6 of (Woodroofe [45]) gives $|G_n(a_F)-G(a_F)| = O_{a.s}\left(n^{-1/2}\right)$ which are negligible with respect to $O\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right)$. Using Lemma 10 of (Ferraty *et al.* [13]), we have easily $\frac{1}{n\mathbb{E}[K_1(\theta,x)]}\sum_{i=1}^n K_i(\theta,x) = o(1)$ which permits to conclude

$$\left|\widehat{F}_D(\theta,x)-\widetilde{F}_D(\theta,x)\right| = O_{a.s}\left(\sqrt{\frac{1}{n}}\right).$$

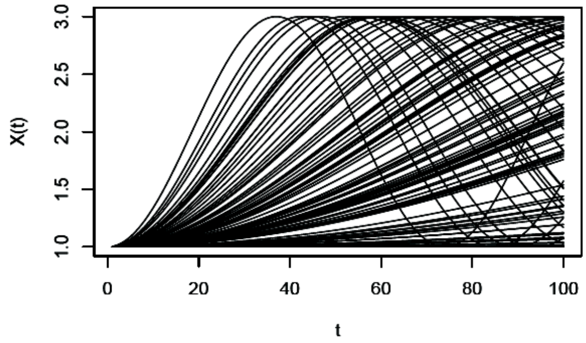

Figure 6: A sample of simulated curves

□

Proof. [Proof of Lemma 3.5] The case where $j = 0$

(i) Firstly,

$$\begin{aligned}
\mathbb{E}\left(\widetilde{F}_N(\theta, y, x)\right) &= \frac{\mu}{n\mathbb{E}\left[K_1(\theta, x)\right]} \sum_{i=1}^{n} \mathbb{E}\left(G^{-1}(Y_i)K(\theta, x_i)H_i(y)\right) \\
&= \frac{\mu}{\mathbb{E}\left[K_1(\theta, x)\right]} \mathbb{E}\left\{G^{-1}(Y_1)K_1(\theta, x)\mathbb{E}\left(\mathbb{E}\left[H_1(y|\langle X_1, \theta\rangle)\right]\right)\right\}.
\end{aligned}$$

By Integration by parts, changing variables and because (A3) and (A4), we have

$$\begin{aligned}
I &:= \mathbb{E}\left[H_1\left(y|\langle X_1, \theta\rangle\right)\right] \\
&= \int_{\mathbb{R}} H_1\left(\frac{y-u}{h_H}\right) f(\theta, u, X_1)du \\
&= \int H_1'(z)(\theta, y - zh_H, X_1) - F(\theta, y, x))dz + F(\theta, y, x)\int_{\mathbb{R}} H_1'(z)dz
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\left[\widetilde{F}_N(\theta, y, x)\right] &= \frac{\mu\mathbb{E}\left[G^{-1}(Y_1)K_1(\theta, x)\right]}{\mathbb{E}\left[K_1(\theta, x)\right]} \int H_1'(z)F(\theta, y - zh_H, X_1) - F(\theta, y, x)dz \\
&\quad + \frac{\mu}{E\left[K_1(\theta, x)\right]} F(\theta, y, x)\mathbb{E}\left[G^{-1}(Y_1)K_1(\theta, x)\right] \\
&:= I_1 + I_2.
\end{aligned}$$

By using Lemma 6.1, we have $I_2 = F(\theta, y, x)\mathbb{E}\left(\widetilde{F}_D(\theta, x)\right) = F(\theta, y, x)$.
Then by assumptions (A3) and (A4), we get

$$\begin{aligned}
\int H_1'(z)F(\theta, y - zh_H, X_1) - F(\theta, y, x))dz &\leq C_{\theta,x} \int H'(z)\left(h_K^a + |z|^b h_H^b\right)dz \\
&\leq C_{\theta,x}\left(h_K^a \int H'(z)dz\right) \\
&\quad + C_{\theta,x}\left(h_H^b \int H^{'}(z)|z|^b dz\right)
\end{aligned}$$

Then, $I_1 = O\left(h_K^a + h_H^b\right)$.

(ii) For establish (8), we use the same arguments as in the proof of (6) of Lemma 3.4. Since, H is bounded and $\frac{1}{n\mathbb{E}[K_1(\theta,x)]}\sum_{i=1}^{n} K_i(\theta,x) = o(1)$, where we use the same techniques used for the proof of the previous lemma.

(iii) Finally, for (9), using the compactness of $S_{\mathbb{R}} \subset \mathbb{R}$, we can write $S_{\mathbb{R}} \subset \bigcup_{k=1}^{\tau_n}(v_k - l_n, v_k + l_n)$ with l_n and τ_n can be chosen such that $l_n = O\left(\tau_n^{-1}\right) = O\left(n^{-\frac{\gamma}{2}-\frac{1}{2}}\right)$. We also put $k_y = \arg\min_{\{v_1,\dots,v_{\tau_n}\}}\left|y - v_k\right|$.

We have the following decomposition, valid for any $y \in S_{\mathbb{R}}$

$$
\begin{aligned}
\left|\widetilde{F}_N(\theta,y,x) - \mathbb{E}\left[\widetilde{F}_N(\theta,y,x)\right]\right| &\leq \left|\widetilde{F}_N(\theta,y,x) - \widetilde{F}_N\left(\theta,y_{k_y},x\right)\right| \\
&\quad + \left|\widetilde{F}_N\left(\theta,y_{k_y},x\right) - \mathbb{E}\left[\widetilde{F}_N\left(\theta,y_{k_y},x\right)\right]\right| \\
&\quad + \left|\mathbb{E}\left[\widetilde{F}_N\left(\theta,y_{k_y},x\right)\right] - \mathbb{E}\left[\widetilde{F}_N(\theta,y,x)\right]\right| \\
&=: L_1 + L_2 + L_3.
\end{aligned}
$$

- Clearly L_1 and L_3 can be treated in the same manner, we deal only with the first term. By the fact that K is bounded and because the Lipschitz's condition of H and $\lim_{n\to\infty} n^{\gamma}h_H = \infty$. Making use of (A2)-(i) and (A3), we get

$$
\begin{aligned}
L_1 &\leq \frac{\mu}{n\mathbb{E}(K_1(\theta,x))}\sum_{i=1}^{n}\left|\frac{K_i(\theta,x)}{G(Y_i)}\right|\left|H\left(\frac{y-Y_i}{h_H}\right) - H\left(\frac{y_{k_y}-Y_i}{h_H}\right)\right| \\
&\leq \frac{C}{G(a_F)\mathbb{E}(K_1(\theta,x))}\sup_{y\in S_{\mathbb{R}}}\frac{|y-y_{k_y}|}{h_H} \\
&\leq \frac{Cl_n}{G(a_F)h_H\phi_{\theta,x}(h_K)} = o\left(\frac{l_n}{h_H}\phi_{\theta,x}(h_K)\right).
\end{aligned}
$$

Using the fact that $\lim_{n\to\infty} n^{\gamma}h_H = +\infty$, and choosing $l_n = n^{-\frac{\gamma-1}{2}}$ implies

$$
\frac{l_n}{h_H} = o\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right).
$$

Thus, for n large enough, we have $L_1 = O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right)$, we can conclude that

$$
L_3 \leq L_1 = O_{a.co}\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right).
$$

- Concerning L_2,

$$
L_2 \leq \frac{1}{n}\sum_{i=1}^{n}\left(\underbrace{\frac{\mu G^{-1}(Y_i)K_i(\theta,x)H_i\left(y_{k_y}\right)}{\mathbb{E}\left[K_1(\theta,x)\right]} - \mathbb{E}\left[\frac{\mu G^{-1}(Y_i)K_i(\theta,x)H_i\left(y_{k_y}\right)}{\mathbb{E}\left[K_1(\theta,x)\right]}\right]}_{\Lambda_i}\right) \tag{20}
$$

$(\Lambda_i)_{i=1,\dots,n}$ is a sequence of i.i.d. center random variables. The application of exponential inequality given in Lemma 6.1 on the latter sequence is based on evaluation of $\mathbb{E}\left(\left|\Lambda_1^m\right|\right)$ for all $m \geq 2$. By using the same arguments as in lemma 6.3 in (Ferraty and Vieu [15]). Indeed, by Newton's Binomial expansion, we get

$$\mathbb{E}\left[\Lambda_1^m\right] \leq C\mathbb{E}\left[K_1(\theta,x)\right]^{1-m} \quad \text{for all} \quad m \geq 2 \tag{21}$$

- **Proof of** (21):

$$\begin{aligned}
\mathbb{E}\left[\left|\Lambda_1^m\right|\right] &= \frac{1}{\mathbb{E}\left[K_1(\theta,x)\right]^m}\sum_{k=1}^{m} C_m^k \mathbb{E}\left(\frac{\mu^k}{G^k(Y_1)}K_1^k(\theta,x)H_i^k\left(y_{k_y}\right)\right)\\
&\quad \times\left(\mathbb{E}\left[\frac{\mu}{G(Y_1)}K_1(\theta,x)H_i^k\left(y_{k_y}\right)\right]\right)^{m-k}.
\end{aligned}$$

Observe that, under (2) and (A4), we have, for all $k > 0$,

$$\mathbb{E}\left(\frac{\mu}{G(Y_i)}K_1(\theta,x)H_i^k\left(y_{k_y}\right)\right)^k \leq \frac{C\mu^k}{G^k\left(a_F\right)}\mathbb{E}\left[K_1(\theta,x)\right]$$

and

$$\mathbb{E}\left(\frac{\mu}{G(Y_i)}K_1(x,\theta)H_i^k\left(y_{k_y}\right)\right)^{m-k} \leq C\left(\mathbb{E}\left[K_1(x,\theta)\right]\right)^{m-k}.$$

It follows that

$$\mathbb{E}\left(\left|\Lambda_1^m\right|\right) \leq C\mathbb{E}\left(K_1(\theta,x)\right)^{1-m}.$$

Then, we apply Lemma 6.1 with $a = \sqrt{\mathbb{E}\left[K_1(\theta,x)\right]^{-1}}$. Thus under the fact that $\mathbb{E}\left[K_1(\theta,x)\right] = O\left(\phi_{\theta,x}(h_K)\right)$ we get

$$\begin{aligned}
\mathbb{P}\left(\left|\widetilde{F}_N(\theta,x) - \mathbb{E}\left(\widetilde{F}_N(\theta,x)\right)\right| > \varepsilon\right) &= \mathbb{P}\left(\left|\sum_{i=1}^{n} L_i\right| > \varepsilon n\right)\\
&\leq \exp\left(-\frac{n\varepsilon^2}{2\left(\mathbb{E}\left(K_1(\theta,x)\right)\right)^{-1}(1+\varepsilon)}\right).
\end{aligned}$$

Consequently, for $\varepsilon = \varepsilon_0\sqrt{\frac{\log n}{n\mathbb{E}(K_1(\theta,x))}}$, we get

$$\begin{aligned}
\mathbb{P}\left(U_2 > \varepsilon\right) &\leq 2\exp\left(-\frac{\varepsilon_0^2\frac{\log n}{\mathbb{E}[K_1(\theta,x)]}}{\frac{2}{\mathbb{E}[K_1(\theta,x)]}\left(1+\varepsilon_0\sqrt{\frac{\log n}{n\mathbb{E}[K_1(\theta,x)]}}\right)}\right)\\
&\leq 2\exp\left(-c\varepsilon_0^2\log n\right)\\
&\leq 2n^{-c\varepsilon_0^2}.
\end{aligned}$$

Finally, an appropriate choice of ε_0 and the Borel-Cantelli's Lemma use completes the proof.

$$\begin{aligned}
\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\Lambda_i\right| > \varepsilon_0\sqrt{\frac{\log n}{n\mathbb{E}\left[K_1(x,\theta)\right]}}\right) &\leq 2\exp\left(-c\varepsilon_0^2\log n\right)\\
&\leq 2n^{-c\varepsilon_0^2}.
\end{aligned}$$

Now, if $j = 1$

(i) Firstly,

$$\begin{aligned}\mathbb{E}\left[\widetilde{f_N}(\theta,y,x)\right] &= \frac{\mu}{nh_H\mathbb{E}\left[K_1(\theta,x)\right]}\sum_{i=1}^{n}\mathbb{E}\left[G^{-1}(Y_i)K(\theta,x_i)H_i'(y)\right]\\ &= \frac{\mu}{h_H\mathbb{E}\left[K_1(\theta,x)\right]}\mathbb{E}\left(G^{-1}(Y_1)K_1(\theta,x)\mathbb{E}\left(\mathbb{E}\left[H_1'\left(y|\left\langle X_1,\theta\right\rangle\right)\right]\right)\right).\end{aligned}$$

Now, by changing variables and using assumptions (A3)-(A4), we have:

$$\begin{aligned}M: &= \mathbb{E}\left[H_1'\left(y|\left\langle X_1,\theta\right\rangle\right)\right] = \int_{\mathbb{R}} H_1'\left(\frac{y-u}{h_H}\right)f(\theta,u,X_1)du\\ &= h_H\int_{\mathbb{R}} H_1'(z)\left(f(\theta,y-zh_H,X_1)-f(\theta,y,x)\right)dz\\ &\quad +h_H' f(\theta,y,x)\int_{\mathbb{R}} H_1(z)dz,\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\left[\widetilde{f_N}(\theta,y,x)\right] &= \frac{\mu\mathbb{E}\left[G^{-1}(Y_1)K_1(\theta,x)\right]}{\mathbb{E}\left[K_1(\theta,x)\right]}\int_{\mathbb{R}} H_1'(z)\left(f(\theta,y-zh_H,X_1)-f(\theta,y,x)\right)dz\\ &\quad +\frac{\mu}{\mathbb{E}\left[K_1(\theta,x)\right]}f(\theta,y,x)\mathbb{E}\left[G^{-1}(Y_1)K_1(\theta,x)\right]\\ &=: M_1+M_2.\end{aligned}$$

By using Lemma 6.1, we have:

$$M_2 = f(\theta,y,x)\mathbb{E}\left[\widetilde{F}_D(\theta,x)\right] = f(\theta,y,x).$$

Then, by assumptions (A3) and (A4), we get

$$\begin{aligned}\int_{\mathbb{R}} H_1'(z)\left(f(\theta,y-zh_H,X_1)-f(\theta,y,x)\right)dz &\leq C_{\theta,x}\int_{\mathbb{R}} H_1'(z)\left(h_K^a+|z|^b h_H^b\right)dz\\ &\leq C_{\theta,x}\left(h_K^a\int_{\mathbb{R}} H_1'(z)dz + h_H^b\int_{\mathbb{R}} H_1'(z)|z|^b dz\right).\end{aligned}$$

Thus, $M_1 = \mathcal{O}\left(h_K^a+h_H^b\right)$.

(ii) For establish (8) in the case where $j=1$ (i.e $\sup_{y\in S_{\mathbb{R}}}\left|\widehat{f_N}(\theta,y,x)-\widetilde{f_N}(\theta,y,x)\right| = \mathcal{O}_{a.s}\left(\sqrt{\frac{1}{n}}\right)$), we use the same arguments as in the proof of Lemma 3.4. Since, H' is bounded and $\frac{1}{nh_H\mathbb{E}[K_1(\theta,x)]}\sum_{i=1}^{n}K_i(\theta,x)=o(1)$, we use the same techniques as the previous lemma.

(iii) Finally, for (9) if $j=1$, using the compactness of $S_{\mathbb{R}}\subset\mathbb{R}$, we can write $S_{\mathbb{R}}\subset\bigcup_{k=1}^{\tau_n}(v_k-l_n,v_k+l_n)$ with l_n and τ_n can be chosen such that $l_n=\mathcal{O}\left(\tau_n^{-1}\right)=\mathcal{O}\left(n^{-\frac{3\gamma}{2}-\frac{1}{2}}\right)$. We also put $k_y=\arg\min_{\{v_1,\dots,v_{\tau_n}\}}\left|y-v_k\right|$.

We have the following decomposition, valid for any $y\in S_{\mathbb{R}}$

$$\begin{aligned}\left|\widetilde{f_N}(\theta,y,x)-\mathbb{E}\left[\widetilde{f_N}(\theta,y,x)\right]\right| &\leq \left|\widetilde{f_N}(\theta,y,x)-\widetilde{f_N}\left(\theta,y_{k_y},x\right)\right|\\ &\quad +\left|\widetilde{f_N}\left(\theta,y_{k_y},x\right)-\mathbb{E}\left[\widetilde{f_N}\left(\theta,y_{k_y},x\right)\right]\right|\\ &\quad +\left|\mathbb{E}\left[\widetilde{f_N}\left(\theta,y_{k_y},x\right)\right]-\mathbb{E}\left[\widetilde{f_N}(\theta,y,x)\right]\right|\\ &=: U_1+U_2+U_3.\end{aligned}$$

- Concerning U_1 and U_3: By the fact that K is bounded and because the Lipschitz's condition of H' and $\lim_{n\to\infty} n^\gamma h_H = \infty$, we get:

$$\begin{aligned} U_1 &\leq \frac{\mu}{nh_H\mathbb{E}(K_1(\theta,x))}\sum_{i=1}^{n}\left|K_i(\theta,x)G^{-1}(Y_i)\right|\left|H'\left(\frac{y-Y_i}{h_H}\right)-H'\left(\frac{y_{k_y}-Y_i}{h_H}\right)\right| \\ &\leq \frac{C}{h_HG(b_F)\mathbb{E}(K_1(\theta,x))}\sup_{y\in S_\mathbb{R}}\frac{|y-y_{k_y}|}{h_H} \\ &\leq \frac{Cl_n}{h_H^2\phi_{\theta,x}(h_K)G(b_F)}. \end{aligned}$$

Since $l_n = O\left(n^{-\frac{3\gamma}{2}-\frac{1}{2}}\right)$ and $\lim_{n\to\infty} n^\gamma = \infty$, we get

$$\frac{l_n}{h_H^2\phi\,(h_K)} = o\left(\sqrt{\frac{\log n}{nh_H\phi\,(h_K)}}\right).$$

Thus, for n large enough, we can conclude that

$$U_3 \leq U_1 = O_{a.co}\left(\sqrt{\frac{\log n}{nh_H\phi\,(h_K)}}\right).$$

- Concerning U_2, it is treated in the same manner as L_2 (formula (20) in the proof of (9) in the case where $j = 0$).

□

Proof. [Proof of Lemma 3.6] The proof of (10) is very close to (9) of Lemma 3.5 (formula (21)). It is based on the exponential inequality given in Lemma 6.1 and the following: $\widetilde{F}_D(\theta,x) - \mathbb{E}\left(\widetilde{F}_D(\theta,x)\right) = \frac{1}{n}\sum_{i=1}^n Z_i$ where

$$Z_i := \frac{\mu G^{-1}(Y_i)K_i(\theta,x)H_i(y)}{\mathbb{E}\left[K_1(\theta,x)\right]} - \mathbb{E}\left(\frac{\mu G^{-1}(Y_i)K_i(\theta,x)H_i(y)}{\mathbb{E}\left[K_1(\theta,x)\right]}\right).$$

We apply Lemma 6.1 with $a = \sqrt{\mathbb{E}(K_1)^{-1}}$. Thus, under the fact that $\mathbb{E}\left[K_1(\theta,x)\right] = O\left(\phi_{\theta,x}(h_K)\right)$ we get,

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i\right| > \varepsilon_0\sqrt{\frac{\log n}{n\mathbb{E}\left[K_1(x,\theta)\right]}}\right) \leq 2\exp\left(-c\varepsilon_0^2\log n\right) \leq 2n^{-c\varepsilon_0^2}.$$

Concerning (11), the proof is based on the following decomposition, (6) of Lemma 3.4 and (9) of Lemma 3.5

$$\begin{aligned} \left|\widehat{F}_D(\theta,x)-1\right| &= \left|\widehat{F}_D(\theta,x)-\mathbb{E}\left(\widetilde{F}_D(\theta,x)\right)\right| \\ &\leq \left|\widehat{F}_D(\theta,x)-\widetilde{F}_D(\theta,y,x)\right| + \left|\widetilde{F}_D(\theta,y,x)-\mathbb{E}\left(\widetilde{F}_D(\theta,y,x)\right)\right|. \end{aligned}$$

□

Proof. [Proof of Corollary 3.7] It is clear that

$$\widehat{F}_D(\theta,x) \leq \frac{1}{2} \Longrightarrow 1-\widehat{F}_D(\theta,x) \geq \frac{1}{2} \Longrightarrow \left|\widehat{F}_D(\theta,x)-1\right| \geq \frac{1}{2}$$

which implies that

$$\mathbb{P}\left(\left|\widehat{F}_D(\theta,x)\right| \leq \frac{1}{2}\right) \leq \mathbb{P}\left(\left|\widehat{F}_D(\theta,x)-1\right| \geq \frac{1}{2}\right) < \infty.$$

□

Proof. [Proof of Lemma 4.4] Because the proof of Lemma 3.5 (formula (7)) is uniform on $(\theta, y, x) \in \Theta_{\mathcal{H}} \times S_{\mathbb{R}} \times S_{\mathcal{H}}$, the proof of Lemma 4.4 is the same of the latter. □

Proof. [Proof of Lemma 4.5] Firstly, for equation (12), using similar tools as those to proof from (6) of Lemma 3.4, one can show that

$$\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \left|\widehat{F}_D(\theta,x) - \widetilde{F}_D(\theta,x)\right| = O_{a.s}\left(n^{-\frac{1}{2}}\right).$$

By using (U4) and (A5), we get

$$O_{a.s}\left(n^{-\frac{1}{2}}\right) = O\left(\frac{\log n}{n\phi(h_K)}\right) = O\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

Finally, since H and H' are bounded, the proof of (12) as well as the proof of (13)-(14) are the same. □

Proof. [Proof of Lemma 4.6]

$\forall\, x \in S_{\mathcal{H}}$ and $\forall \theta \in \Theta_{\mathcal{H}}$, we put:

$$k(x) = \arg\min_{k\in\left\{1,\dots,d_n^{S_{\mathcal{H}}}\right\}} \|x - x_k\|,\ m(\theta) = \arg\min_{m\in\left\{1,\dots,d_n^{\Theta_{\mathcal{H}}}\right\}} \|\theta - \theta_m\|.$$

By the compactness property of $S_{\mathbb{R}} \subset \mathbb{R}$, we have $S_{\mathbb{R}} \subset \bigcup_{k=1}^{\tau_n} (v_k - l_n,\ v_k + l_n)$ with l_n and τ_n chosen as $l_n = O\left(\tau_n^{-1}\right)$. Recall that $k_y = \arg\min_{\{v_1,\dots,v_{\tau_n}\}} \left|y - v_k\right|$.

Let us consider the following decomposition for $j = 0, 1$

$$\begin{aligned}
\left|\widetilde{F}_N^{(j)}(\theta,y,x) - \mathbb{E}\left[\widetilde{F}_N^{(j)}(\theta,y,x)\right]\right| \leq\ & \left|\widetilde{F}_N^{(j)}(\theta,y,x) - \widetilde{F}_N^{(j)}(\theta,y,x_{k(x)})\right| \\
&+ \left|\widetilde{F}_N^{(j)}(\theta,y,x_{k(x)}) - \widetilde{F}_N^{(j)}(\theta_{m(\theta)},y,x_{k(x)})\right| \\
&+ \left|\widetilde{F}_N^{(j)}(\theta_{m(\theta)},y,x_{k(x)}) - \widetilde{F}_N^{(j)}(\theta_{m(\theta)},y_{k_y},x_{k(x)})\right| \\
&+ \left|\widetilde{F}_N^{(j)}(\theta_{m(\theta)},y_{k_y},x_{k(x)}) - \mathbb{E}\left[\widetilde{F}_N^{(j)}(\theta_{m(\theta)},y_{k_y},x_{k(x)})\right]\right| \\
&+ \left|\mathbb{E}\left[\widetilde{F}_N^{(j)}(\theta_{m(\theta)},y_{k_y},x_{k(x)})\right] - \mathbb{E}\left[\widetilde{F}_N^{(j)}(\theta_{m(\theta)},y,x_{k(x)})\right]\right| \\
&+ \left|\mathbb{E}\left[\widetilde{F}_N^{(j)}(\theta_{m(\theta)},y,x_{k(x)})\right] - \mathbb{E}\left[\widetilde{F}_N^{(j)}(\theta,y,x_{k(x)})\right]\right| \\
&+ \left|\mathbb{E}\left[\widetilde{F}_N^{(j)}(\theta,y,x_{k(x)})\right] - \mathbb{E}\left[\widetilde{F}_N^{(j)}(\theta,y,x)\right]\right| \\
=:\ & \Psi_1 + \Psi_2 + \Psi_3 + \Psi_4 + \Psi_5 + \Psi_6 + \Psi_7,
\end{aligned}$$

for (15), chosen as $l_n = O\left(\tau_n^{-1}\right) = O\left(n^{-\frac{1}{2\gamma}}\right)$.

- Concerning Ψ_3 and Ψ_5, by conditions (A3) and (U4), boundness of K and using Lipcshtz's condition on H, we obtain

$$\begin{aligned}
\Psi_3 &\leq \frac{1}{n\mathbb{E}(K_1(\theta,x))} \sum_{i=1}^{n} \left|\frac{\mu}{G(Y_i)} K_i(\theta_{m(\theta)},x_{k(x)})\right| \left|H\left(\frac{y-Y_i}{h_H}\right) - H\left(\frac{y_{k_y}-Y_i}{h_H}\right)\right| \\
&\leq \sup_{y\in S_{\mathbb{R}}} C\frac{\left|y - y_{k_y}\right|}{h_H} \frac{1}{nG(a_F)\mathbb{E}(K_1(\theta_{m(\theta)},x_{k(x)}))} \sum_{i=1}^{n} \left|K_i(\theta_{m(\theta)},x_{k(x)})\right| \\
&\leq \frac{Cl_n}{h_H G(a_F)\phi(h_K)} = O\left(\frac{l_n}{h_H\phi(h_K)}\right).
\end{aligned}$$

Now, the fact that $\lim_{n\to\infty} n^{\gamma} h_H = \infty$, choosing $l_n = n^{-\frac{1}{2\gamma}}$ and using (U4)-(ii), it yields

$$\frac{l_n}{h_H\phi(h_K)} = o\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

Hence, for n large enough, we have

$$\Psi_5 \leq \Psi_3 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

- Concerning Ψ_1 and Ψ_2 we have

$$\begin{aligned}
\Psi_1 &\leq \frac{1}{n\mathbb{E}(K_1(\theta,x))} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \sup_{y\in S_{\mathbb{R}}} \sum_{i=1}^{n} \left|\frac{\mu}{G(Y_i)}\right| \left|H_i(y)\right| \times \left|K_i(\theta,x) - K_i(\theta,x_{k(x)})\right| \\
&\leq \frac{\mu}{nG(a_F)\phi(h_K)} \sup_{x\in S_{\mathcal{H}}} \sup_{\theta\in\Theta_{\mathcal{H}}} \sum_{i=1}^{n} \left|K_i(\theta,x) - K_i(\theta,x_{k(x)})\right| \\
&\leq \frac{\mu}{\phi(h_K)G(a_F)} \sup_{x\in S_{\mathcal{H}}} \sup_{\theta\in\Theta_{\mathcal{H}}} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{B_\theta(x,h_K)\cup B_\theta(x_{k(x)},h_K)}(X_i) \\
&= O\left(\frac{\log n}{n\phi\,(h_K)}\right).
\end{aligned}$$

Then using

$$\frac{\log n}{n\phi(h_K)} = o\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right),$$

we get

$$\Psi_1 = \Psi_2 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

Similarly, one can show that

$$\Psi_6 = \Psi_7 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi\,(h_K)}}\right).$$

Now we deal with Ψ_4. For all $\eta > 0$ we have

$$\mathbb{P}\left(\Psi_4 > \eta\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi\,(h_K)}}\right)$$

$$\leq$$

$$\tau_n d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}} \max_{k\in\left\{1\ldots d_n^{S_{\mathcal{H}}}\right\}} \max_{k\in\left\{1\ldots d_n^{\Theta_{\mathcal{H}}}\right\}} \max_{k\in\{1\ldots\tau_n\}} \mathbb{P}\left(\Psi_4 > \eta\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi\,(h_K)}}\right).$$

Applying Bernstein's exponential inequality to

$\Delta_i = \frac{1}{\phi(h_K)} \left\{ K_i\left(\theta_{m(\theta)}, x_{k(x)}\right) H_i\left(y_{k_y}\right) - \mathbb{E}\left(K_i\left(\theta_{m(\theta)}, x_{k(x)}\right) H_i\left(y_{k_y}\right)\right)\right\}$ one get, under (U4) that

$$\Psi_4 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi\left(h_K\right)}}\right).$$

For (16), chosen as $l_n = O\left(\tau_n^{-1}\right) = O\left(n^{-\frac{(3\gamma+1)}{2}}\right)$.

Using the same ideas previous, permit to get, when n tends to infinity

$$\begin{aligned} F_7 &= \left|\mathbb{E}\left[\widetilde{f_N}(\theta, y, x_{k(x)})\right] - \mathbb{E}\left[\widetilde{f_N}(\theta, y, x)\right]\right| \\ &\leq F_1 = \left|\widetilde{f_N}(\theta, y, x) - \widetilde{f_N}(\theta, y, x_{k(x)})\right|, \end{aligned}$$

and

$$\begin{aligned} F_6 &= \left|\mathbb{E}\left[\widetilde{f_N}(\theta_{m(\theta)}, y, x_{k(x)})\right] - \mathbb{E}\left[\widetilde{f_N}(\theta, y, x_{k(x)})\right]\right| \\ &\leq F_2 = \left|\widetilde{f_N}(\theta, y, x_{k(x)}) - \widetilde{f_N}(\theta_{m(\theta)}, y, x_{k(x)})\right|, \end{aligned}$$

when n tends to infinity

$$F_7 \leq F_1 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right),$$

and

$$F_6 \leq F_2 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right).$$

Concerning the terms F_3 and F_5 using Lipcshtz's condition on H', permits to write

$$\left|\widetilde{f_N}(\theta_{m(\theta)}, y, x_{k(x)}) - \widetilde{f_N}(\theta_{m(\theta)}, y_{k_y}, x_{k(x)})\right| \leq \frac{l_n}{h_H^2\phi(h_K)},$$

where

$$F_3 = \left|\widetilde{f_N}(\theta_{m(\theta)}, y, x_{k(x)}) - \widetilde{f_N}(\theta_{m(\theta)}, y_{k_y}, x_{k(x)})\right|$$

and

$$F_5 = \left|\mathbb{E}\left[\widetilde{f_N}(\theta_{m(\theta)}, y_{k_y}, x_{k(x)})\right] - \mathbb{E}\left[\widetilde{f_N}(\theta_{m(\theta)}, y, x_{k(x)})\right]\right|.$$

Now, the fact that $\lim_{n\to\infty} n^\gamma h_H = \infty$, choosing $l_n = n^{-\frac{(3\gamma+1)}{2}}$ and using (U5)-(ii), it yields

$$\frac{l_n}{h_H^2\phi(h_K)} = o\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right).$$

Hence, for n large enough, we have

$$F_5 \leq F_3 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right).$$

Finally, the evaluation of the term F_4 is very close to Ψ_4 for $j = 0$, where

$$F_4 = \left|\widetilde{f_N}(\theta_{m(\theta)}, y_{k_y}, x_{k(x)}) - \mathbb{E}\left[\widetilde{f_N}(\theta_{m(\theta)}, y_{k_y}, x_{k(x)})\right]\right|.$$

Applying Bernstein's exponential inequality to

$$\Delta_i = \frac{1}{h_H\phi\left(h_K\right)}\left\{K_i\left(\theta_{m(\theta)}, x_{k(x)}\right)H_i'\left(y_{k_y}\right) - \mathbb{E}\left(K_i\left(\theta_{m(\theta)}, x_{k(x)}\right)H_i\left(y_{k_y}\right)\right)\right\},$$

it follows that

$$F_4 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi\left(h_K\right)}}\right).$$

For (17), we have

$$\begin{aligned}\left|\widetilde{F}_D(\theta,x) - \mathbb{E}\left(\widetilde{F}_D(\theta,x)\right)\right| &\leq \left|\widetilde{F}_D(\theta,x) - \widetilde{F}_D\left(\theta, x_{k(x)}\right)\right| \\ &+ \left|\widetilde{F}_D\left(\theta, x_{k(x)}\right) - \widetilde{F}_D\left(t_{j(\theta)}, x_{k(x)}\right)\right| \\ &+ \left|\widetilde{F}_D\left(\theta, x_{k(x)}\right) - \mathbb{E}\left(\widetilde{F}_D\left(t_{j(\theta)}, x_{k(x)}\right)\right)\right| \\ &+ \left|\mathbb{E}\left(\widetilde{F}_D\left(t_{j(\theta)}, x_{k(x)}\right)\right) - \mathbb{E}\left(\widetilde{F}_D\left(\theta, x_{k(x)}\right)\right)\right| \\ &+ \left|\mathbb{E}\left(\widetilde{F}_D\left(\theta, x_{k(x)}\right)\right) - \mathbb{E}\left(\widetilde{F}_D(\theta,x)\right)\right| \\ &:= I_1 + I_2 + I_3 + I_4 + I_5.\end{aligned}$$

Using the same ideas as for Ψ_1, Ψ_2, Ψ_4 and Ψ_5, permits to get, for n tending to infinity

$$I_5 \leq I_1 = O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right)$$

$$I_4 \leq I_2 = O_{a.co}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

Finally, the evaluation of the term I_3 is very close to Ψ_4. Applying Bernstein's exponential inequality to

$$\mathbb{E}_i = \frac{1}{\phi(h_K)}\left\{K_i\left(\theta_{m(\theta)}, x_{k(x)}\right) - \mathbb{E}\left[K_i\left(\theta_{m(\theta)}, x_{k(x)}\right)\right]\right\},$$

it follows that

$$I_3 = O_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

Finally, the proof from (18) is deduced from (12) of Lemma 4.5 and (17) of Lemma 4.6 and the following decomposition.

$$\begin{aligned}\left|\widehat{F}_D(\theta,x) - 1\right| &= \left|\widehat{F}_D(\theta,x) - \mathbb{E}\left(\widetilde{F}_D(\theta,y,x)\right)\right| \\ &\leq \left|\widehat{F}_D(\theta,x) - \widetilde{F}_D(\theta,y,x)\right| \\ &+ \left|\widetilde{F}_D(\theta,y,x) - \mathbb{E}\left(\widetilde{F}_D(\theta,y,x)\right)\right|.\end{aligned}$$

□

Proof. [Proof of Corollary 4.7] It is clear that from the inequality $\inf_{\theta\in\Theta_{\mathcal{H}}}\inf_{x\in S_{\mathcal{H}}}\left|\widehat{F}_D(\theta,x)\right| \le \frac{1}{2}$ it exist $x \in S_{\mathcal{H}}$ and $\theta \in \Theta_{\mathrm{H}}$, such that $1-\widehat{F}_D(\theta,x) \ge \frac{1}{2}$.

So, $\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\left|\widehat{F}_D(\theta,x)-1\right| \ge \frac{1}{2}$ which implies that

$$\mathbb{P}\left(\inf_{\theta\in\Theta_{\mathcal{H}}}\inf_{x\in S_{\mathcal{H}}}\left|\widehat{F}_D(\theta,x)\right| \le \frac{1}{2}\right) \le \mathbb{P}\left(\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\left|\widehat{F}_D(\theta,x)-1\right| \ge \frac{1}{2}\right) < \infty.$$

□

Conclusion

This paper studied the nonparametric estimation of the conditional hazard function in the single functional index model for independent data, when the variable of interest is subject to random left truncation. We established the almost complete convergence and almost uniform complete convergence of the proposed estimators under some standard assumptions in Functional Data Analysis (FDA). Furthermore, we conducted a simulation study to demonstrate the performance of our estimator.

Acknowledgement

The authors are grateful to anonymous referees and the Editor-in-Chief for valuable comments and suggestions towards the improvement of this paper.

Declarations: Conflict of interest

The authors declare that they have no conflict of interest to disclose.

References

[1] A. Aït-Saidi, F. Ferraty and R. Kassa, *Single functional index model for a time series*, Revue Roumaine de Mathématique Pures et Appliquées, **50** (2005), 321–330.

[2] A. Aït-Saidi, F. Ferraty, R. Kassa and P. Vieu, *Cross-validated estimation in the single functional index model*, Statistics, **24** (2008), 475–494.

[3] B. Altendji, B. Rachdi, J. Demongeot and A. Laksaci, *Functional data analysis: estimation of the relative error in functional regression under left-truncation model*, Journal of Nonparametric Statistics, **30(2)** (2018), 472–490.

[4] Attaoui,S., *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** (2014), 257–286

[5] S. Attaoui and N. Ling, *Asymptotic results of a nonparametric conditional cumulative distribution estimator in the single functional index modeling for time series data with applications*, Metrika: International Journal for Theoretical and Applied Statistics, **79(5)** (2016), 485–511.

[6] O. Belabbaci, A. Rabhi and S. Soltani, *Strong uniform consistency of hazard function with functional explicatory variable in single functional index model under censored data*, Applications and Applied Mathematics: An International Journal (AAM), **10(1)** (2015), 114–138.

[7] N. Bellatrach, W. Bouabsa, and M. K. Attouch, *The convergence rate of the hazard function with functional explanatory variable: case of spacial data with k Nearest Neighbor method*, Turkish Journal of Computer and Mathematics Education, **14(2)** (2023), 180-202

[8] A. Cuevas, *A partial overview of the theory of statistics with functional data*, JJournal of Statistical Planning and Inference, **147** (2014), 1–23.

[9] S. Greven, F. Scheipl, *A general framework for functional regression modelling*, Statistical Modelling, **17(1-2)** (2017), 1–35. doi:10.1177/1471082X16681317

[10] S. Derrar, A. Laksaci and E. Ould-Saïd, *On the nonparametric estimation of the functional regression for a random left-truncation model*, Journal of Statistical Theory and Practice, **9** (2015), 823–849.

[11] M. Febrero-Bande, P. Galeano and W. Gonzalez-Manteiga, *Functional principal component regression and functional partial least-squares regression: An overview and a comparative study*, International Statistical Review, **85** (2017), 61–83

[12] F. Ferraty, A. Peuch and P. Vieu, *Modèle à indice fonctionnel simple*, Comptes rendus de l'Académie des sciences, Série 1, Paris., **336** (2003), 1025–1028.

[13] F. Ferraty, A. Lakssaci, A. Tadj and P. Vieu, *Rate of uniform consistency for nonparametric estimates with functional variables* Journal of Statistical Planning and inference, **140** (2010), 335–352.

[14] F. Ferraty, A. Rabhi and P. Vieu, *Estimation nonparamtrique de la fonction de hasard avec variable explicative fonctionnelle*, Romanian Journal of Pure and Applied Mathematics, **52** (2008), 1–18.

[15] F. Ferraty and P. Vieu, *Nonparametric Functional Data Analysis: Theory and Practice*, Springer Series in Statistics, Springer, New York, 2006.

[16] A. Gagui, A. Chouaf, *On the nonparametric estimation of the conditional hazard estimator in a single functional index,* Statistics in transition new series, **23(2)** (2022), 89–105.

[17] L. Gardes and G. Stupfler, *Estimating extreme quantiles under random truncation*, TEST, **24** (2015), 207–227.

[18] A. Goia, P. Vieu, *An introduction to recent advances in high/infinite dimensional statistics,* Journal of Multivariate Analysis, **146** (2016), 1–6.

[19] L. Horvath, G. Rice, *An introduction to functional data analysis and a principal component approach for testing the equality of mean curves,* Revista Matematica Complutense , **28** (2015), 505–548. https://doi.org/10.1007/s13163-015-0169-7

[20] N. Helal and E. Ould-Saïd, *Kernel conditional quantile estimator under lefttruncation for functional regressors*, Opuscula Mathematica, **36** (2016), 25–48.

[21] S. He and G. Yang, *Estimation of the truncation probability in the random truncationmodel*, Annals of statistics, **26** (1998), 1011–1027.

[22] T. Hsing and R. Eubank, *Theoretical Foundations of Functional Data Analysis, With an Introduction to Linear Operators*, Chichester: Wiley, 2015.

[23] J. Jacques, C. Preda, *Functional data clustering: A survey,* Advances in Data Analysis and Classification, **8** (2014), 231–255

[24] J.F. Lawless, *Statistical Models and Methods for Lifetime Data*, Second Edition. Wiley Series in Probability and Statistics, 2002.

[25] H.Kebir, B. Mechab, *Estimation of the conditional hazard function with a recursive Kernel from Censored Functional Ergodic Data*. Computer Sciences & Mathematics Forum, **7(1):16** (2023), 1–5. https://doi.org/10.3390/IOCMA2023-14412

[26] P. Kokoszka, *Dependent functional data*, International Scholarly Research Notices Probability and Statistics, **2012** (2012), 1–30.

[27] P. Kokoszka, H. Oja, B. Park and L. Sangalli, *Special issue on functional data analysis*, Econometrics and Statistics, Elsevier, **1** (2017), 99–100.

[28] P. Kokoszka, M. Reimherr, *Introduction to Functional Data Analysis,* CRC Press, Boca Raton, FL, 2017.

[29] M. Lemdani and E. Ould-Saïd, *Asymptotic behaviour of the hazard rate kernel estimator under truncated and censored data*, Communications in Statistics-Theory and Methods, **36(1)** (2007), 155–173.

[30] N. Ling and P. Vieu, *Nonparametric Modelling for Functional Data: Selected Survey and Tracks for Future*, Statistics, **52** (2018), 934–949.

[31] D. Lynden-Bell, *A method of allowing for known observational selection in small Samples applied to 3CR quasars*, Monthly Notices of the Royal Astronomical Society, **155** (1971), 95–118.

[32] I. Massim and B. Mechab, *Local linear estimation of the conditional hazard function*, International Journal of Statistics and Economics, **17** (2016), 1–11.

[33] J. Mateu and E. Romano, *Advances in spatial functional statistics*, Stochastic Environmental Research and Risk Assessment, **31** (2017), 1–6.

[34] T. Merouan, B. Mechab and I. Massim, *Quadratic error of the conditional hazard function in the local linear estimation for functional data*, Afrika Statistika, **13(3)** (2018), 1759–1777.

[35] A. Nieto-Reyes, H. Battey, Statistical functional depth, In: G. Aneiros, E.G. Bongiorno, R. Cao, P. Vieu (Eds.), *Functional Statistics and Related Fields: Contributions to Statistics*, Springer, Cham, Switzerland, 2017, 197–202

[36] O. Omari, T. Merouan, B. Mechab, *Uniform convergence of nonparametric conditional hazard function in the single functional index modeling for dependent data,* Journal of Science and Arts, **23,(1)** (2023), 91–106.

[37] A. Rabhi, S. Benaissa, E. Hamel and B. Mechab *Mean square error of the estimator of the conditional hazard function*, Applicationes Mathematicae (Warsaw), **40(4)** (2013), 405–430.

[38] A. Rabhi, Y. Hammou and T. Djebbouri, *Non-parametic estimation of the maximum of conditional hazard function under dependence conditions for functional data*, Afrika Statistika, **10(1)** (2015), 777–794.

[39] J.O. Ramsay, B.W. Silverman, Functional Data Analysis, 2nd Ed., Springer, New York, 2005.

[40] G.G. Roussas, *On some properties of nonparametric estimates of probability density functions*, Bulletin de la Société mathématique de Grèce (N.S.), **9(9)** (1968), 29–43.

[41] W. Stute and J. Wang, *The central limit theorem under random truncation*, Bernoulli, **14** (2008), 604–622.

[42] H. Tabti, and A. Aït Saidi, *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**, (2017), 816–838.

[43] P. Vieu, *On dimension reduction models for functional data,* Statistics and Probability Letters, **136** (2018), 134–138.

[44] M.C. Wang, *A semiparametric model for randomly truncated data*, Journal of the American Statistical Association, **84** (1989), 742–748.

[45] M. Woodroofe, *Estimating a distribution function with truncated data*, Annal of Statistics, **13** (1985), 163–177.