Applications and Applied Mathematics: An International Journal (AAM)

Volume 17 | Issue 1 | Article 3

6-2022

(R1483) Local Linear Estimator of the Conditional Hazard Function for Index Model in Case of Missing at Random Data

Khadidja Abdelhak
Salhi Ahmed University Center; University of Sidi Bel Abbes

Abderrahmane Belguerna
Salhi Ahmed University Center

Zeyneb Laala
Salhi Ahmed University Center, zinebla91@hotmail.com

Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466

Applications and Applied Mathematics:
An International Journal
(AAM)

Vol. 17, Issue 1 (June 2022), pp. 33 – 53

Local Linear Estimator of the Conditional Hazard Function For Index Model in Case of Missing at Random Data

[1]Khadidja Abdelhak, [2]Abderrahmane Belguerna and [3,*]Zeyneb Laala

[1,2,3] Department of Mathematics and Computer Sciences
Salhi Ahmed University Center
BP 66, Naâma 45000
Naâma, Algeria

[1]Laboratory of Analysis and Control of P.D.E.
University of Sidi Bel Abbes
University of Sidi Bel Abbes, Algeria
[1]khadidjamekkaoui14@gmail.com; [2]abderrahmane.belguerna@gmail.com;
[3]zinebla91@hotmail.com

*Corresponding Author

Received: February 17, 2021; Accepted: March 22, 2022

Abstract

The estimation of hazard function becomes an important tool in statistics. Also, the single-index model is an essential method that has much application in a lot of domains but the use of this model in nonparametric estimation is very limited, and only a few theoretical results are available in the literature. In this paper, a kernel estimator of the conditional hazard function is proposed for the index model using the local linear model. This estimator is created under some regularity conditions and in the case of missing data at random. Then, the almost complete convergence is established with rate for the proposed estimator. The results prove that it has good asymptotic properties.

Keywords: Almost complete convergence; Conditional hazard function; Functional data; Index model; Local linear; Missing at random; Rate

MSC 2010 No.: 62G05, 62G20

1. Introduction

Nonparametric estimation of the hazard function plays an important role in statistics; it assures building confidence ranges deriving many applications in survival analysis such as seismology, medicine, reliability (Ramsey and Silverman (2005); Ferrary and Vieu (2006); Ferraty et al. (2008)). Furthermore, functional statistics is a research topic that has seen significant progress in recent years, in which several approaches appear (Aldo and Philipe (2016); Aneiros et al. (2019); Antonio and Beatriz (2018)). This field of statistics studies the data obtained from big samples and the functions are collected on fine grids that can be assimilated on curves and surfaces, such as space-time functions. We can refer to the book of Ferrary and Vieu (2006).

The kernel hazard function estimation was firstly introduced in the work of Watson and Leadbetter (1964). In the same topic we can cite also Roussas (1989) in which the asymptotic normality of the hazard rate function with dependence conditions is shown. But, a few studies on functional spatial data of the hazard function are presented, and for more details we can refer the reader to Laksaci and Mechab (2010) in which they studied the almost complete convergence for the kernel estimate of the conditional hazard function.

Another approach used in the functional estimation is the local linear method that was applied by Barrientos-Marin et al. (2010) to estimate the regression function. They established the almost complete convergence (with rate) of the proposed estimator. In the works of Demongeot et al. (2010) and Demongeot et al. (2014) the estimation of the conditional density and the conditional distribution was introduced, the authors based on the local modelling approach in the case of functional explanatory variable. Then, in a recent work of Massim and Mechab (2016) the conditional hazard function was estimated with the almost complete convergence for the proposed estimator. More recently, in the work of Abeidallah et al. (2020) they obtain previous results in the case of spatial model.

When observing data in practice, it is common to find that it is incomplete, and the responses are missing at random (MAR). The works on the multivariate settings for MAR samples are developed in Cheng (1994), Little and Rubin (2002) and Efromovich (2011). In functional data siting, the first work is of Ferraty et al. (2013) where the estimator of the mean response was considered and under missing at random (MAR) assumption it was proven that the infinite dimensionality of the problem did not affect the rates of convergence when stating that the estimator is root-n consistent. Bouiadjra (2017) studied the uniform almost complete convergence of the kernel type estimator of the conditional hazard function when the explanatory variable takes its values in a semi-metric space and the scalar response is missing at random.

Moreover, many recent works present the functional ergodic data missing at random (Ling et al. (2015); Ling et al. (2016); Hamidi and Mechab (2018)). These works are interested on the study of asymptotic properties of the regression function estimator, conditional mode and conditional quantile. Benchiha and Kaid (2018) considered the estimation of the regression function in the case of local linear estimation and established the almost complete convergence. Kenouza et al. (2019) considered the local linear estimation of the conditional hazard function of a real response variable that is not completely observed. They used the local linear estimation of the conditional density and cumulative distribution function. In the case of missing data at random and under some regularity conditions, the almost complete convergence with rate is established for the proposed estimator.

The single-index model is an important method for dimensional reduction. Its applications are found in many domain such as econometrics, biostatistics, medicine and financial econometric. The use of this model in nonparametric estimation is very limited and only few theoretical results are available in literature. This model was introduced firstly by Ferraty et al. (2003) for the kernel regression estimation. The authors established the almost complete convergence, in the i.i.d. case, of the link regression function of this model. Also, Ait Saidi et al. (2008) proposed the estimation of the unknown functional index by using cross-validation technique. We refer also Attaoui et al. (2011) in which point-wise and uniform almost complete convergence for the conditional density of the single-index model are presented. Ferraty et al. (2011) also proposed a new estimator of the single index based on the idea of functional derivative estimation. The work of Ferraty et al. (2003) was extended by Bouchentouf et al. (2014) to the case of the conditional distribution where kernel type estimators for the conditional cumulative distribution function and the successive derivatives of the conditional density are considered. The authors introduced point-wise almost complete convergence and uniform almost complete convergence of the kernel estimate of this model, and this was applied on the estimations of the conditional mode. Recently, the single index model is used in Tabti and Ait Saidi (2018) to estimate the conditional hazard function in the quasi-associated data. Attaoui (2014) investigate nonparametric estimation of the conditional density of scalar response variable given a random variable taking values in separable Hilbert space. They established the uniform almost complete convergence rates and the asymptotic normality of the conditional density kernel estimator, when the variables satisfy the strong mixing dependency, based on the single-index structure. In Belabbaci et al. (2015), they studied the nonparametric estimate of the conditional hazard function, with functional covariate. Kernel type estimators for the conditional hazard function are introduced when the observations are linked with a single-index structure.They introduced the point-wise almost complete convergence and the uniform almost complete convergence of the kernel estimate.

The remainder of this paper is structured as follow. We give and explain the model and necessary hypothesis in Section 2. In Section 3 we present the major results and their proofs. A simulation study in Section 4 illustrates our findings. Finally, we add a conclusion in Section 5.

2. Model and hypothesis

Let $(X_i, Y_i)_{1\leq i\leq n}$ be a sequence of n independent and identically random pair according to the distribution of the pair (X, Y), taking their values in $\mathcal{H}\times\mathbb{R}$, where $\mathcal{H}$ is a separable real Hilbert space with the norm $||\cdot||$ generated by an inner product $<\cdot,\cdot>$. We consider the semi metric d_θ, associated to the single index $\theta\in\mathcal{H}$ defined by $\forall x_1, x_2\in\mathcal{H} : d_\theta(x_1, x_2) := | < x_1 - x_2, \theta > |$. In our context, we assume that the conditional hazard function of Y given X has a single-index structure θ in $\mathcal{H}$, denoted by $h_\theta^x(\cdot)$, given for $y \in \mathbb{R}$, by

$$h_\theta^x(y) := h(y| < x, \theta >).$$

Distinctly, the identifiability of the model is assured such that for all $x \in \mathcal{H}$, we get,

$$h_1(y| <\cdot, \theta_1 >) = h_2(y| <\cdot, \theta_2 >) \Rightarrow h_1 \equiv h_2 \text{ and } \theta_1 = \theta_2.$$

We intend to estimate the conditional hazard function h_θ^x where the regular version F_θ^x of the conditional distribution function of Y given $X = x$ exists for any $x \in N_x$. Moreover, we suppose that F_θ^x has a continuous density f_θ^x with respect to (w.r.t.) Lebesgue's measure over $\mathbb{R}$. We define the function hazard h_θ^x, for $y \in \mathbb{R}$ and $F_\theta^x(y) < 1$, by

$$h_\theta^x(y) = \frac{f_\theta^x(y)}{1 - F_\theta^x(y)}.$$

In this paper, we define the local linear estimator $\widehat{h}_\theta^x(y)$ of $h_\theta^x(y)$ by

$$\widehat{h}_\theta^x(y) = \frac{\widehat{f}_\theta^x(y)}{1 - \widehat{F}_\theta^x(y)}, \; y \in \mathbb{R} \quad \text{and} \quad \widehat{F}_\theta^x(y) < 1.$$

In what follows, we consider the case where the response variables are MAR data, an available incomplete sample of size n from (X, Y, δ) is $\{(X_i, Y_i, \delta_i), 1 \leq i \leq n\}$, where X_i is observed completely, $\delta_i = 1$ if Y_i is observed, and $\delta_i = 0$ otherwise. Meanwhile, the Bernoulli random variable δ is satisfied with

$$\mathbb{P}(\delta = 1|X, Y) = \mathbb{P}(\delta = 1| < \theta, X >) = P(X),$$

where $P(X)$ is an unknown functional operator called the conditional probability of observing response variable Y given an explanatory variable X. Missing at random is a common assumption in statistical analysis with missing data and it is feasible in many practical situations (see Little and Rubin (2002)).

As indicated by Fan and Gijbels (1996), the function $F_\theta^x(\cdot)$ can be considered as a nonparametric regression model with response variable $H(h_H^{-1}(\cdot - Y_i))$ where H is some cumulative distribution function and h_H is a sequence of positive real numbers. This consideration is motivated by the fact that

$$\mathbb{E}[H(h_H^{-1}(y - Y_i))| < \theta, X_i = x >] \to F_\theta^x(y) \text{ as } h_H \to 0.$$

We use a technique that extends the local linear ideas to the infinite dimensional framework (see Barrientos-Marin et al. (2010) and Demongeot et al. (2010)). We combine this idea with the fact that the data are considered missing at random. Here, we adopt the fast functional local modeling, that is, the conditional cumulative distribution function $\widehat{F}^x$ is estimated by $\widehat{a}$ where the couple $(\widehat{a}, \, \widehat{b})$ is found by the following optimization problem:

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} \left(H(h_H^{-1}(y - Y_i)) - a - b\beta_\theta(X_i, \; x)\right)^2 \delta_i K(h_K^{-1} \varrho_\theta(x, X_i)). \tag{1}$$

With direct computations, we get

$$\widehat{F}_{\theta}^{x}(y)=\frac{\sum_{1\leq i,j\leq n}W_{ij}(\theta,x)H({h_H}^{-1}(y-Y_j))}{\sum_{1\leq i,j\leq n}W_{ij}(\theta,x)},\qquad\forall y\in\mathbb{R},\tag{2}$$

and

$$\widehat{f}_{\theta}^{x}(y)=\frac{\sum_{1\leq i,j\leq n}W_{ij}(\theta,x)H^{(1)}({h_H}^{-1}(y-Y_j))}{h_H\sum_{1\leq i,j\leq n}W_{ij}(\theta,x)},\qquad\forall y\in\mathbb{R},\tag{3}$$

where $H^{(1)}$ is the derivative of the distribution function H.

We note that

$$\widehat{F}_{\theta}^{x}(y)=\frac{\widehat{F}_{\theta,N}^{x}(y)}{\widehat{F}_{D}(\theta,x)}\qquad\text{and}\qquad\widehat{f}_{\theta}^{x}(y)=\frac{\widehat{f}_{\theta,N}^{x}(y)}{\widehat{F}_{D}(\theta,x)},$$

where

$$\widehat{F}_{\theta,N}^{x}(y)=\frac{1}{(nh_K\phi_{\theta,x}(h_K))^2}\sum_{1\leq i,j\leq n}W_{ij}(\theta,x)H(h_H^{-1}(Y-y_j)),$$
$$\widehat{f}_{\theta,N}^{x}(y)=\frac{1}{(nh_K\phi_{\theta,x}(h_K))^2h_H}\sum_{1\leq i,j\leq n}W_{ij}(\theta,x)H^{(1)}(h_H^{-1}(Y-y_j)),$$

and

$$\widehat{F}_{D}(\theta,x)=\frac{1}{(nh_K\phi_{\theta,x}(h_K))^2}\sum_{1\leq i,j\leq n}W_{ij}(\theta,x),$$

with

$$W_{ij}(\theta,x)=\beta_\theta(X_i,x)\Big(\beta_\theta(X_i,x)-\beta_\theta(X_j,x)\Big)\delta_i\delta_jK(h_K^{-1}d_\theta(x,X_i))K(h_K^{-1}\varrho_\theta(x,X_j)),$$

and $\beta_\theta(X_i,x)=<x-X_i,\theta>$ is a known bi-functional operator from $\mathcal{H}^2$ into $\mathbb{R}$, such that $\forall x'\in\mathcal{H}$, $\forall\theta\in\mathcal{H}$, $<x-x',\theta>=0$, with the function K is a kernel and $h_K=h_{K,n}$ (respectively, $h_H=h_{H,n}$) is a sequence of positive real numbers. The final form of our estimator in the case of the single functional index with missing data at random is: for $n\geq 1, y\in\mathbb{R}$,

$$\widehat{h}_{\theta}^{x}(y)=\frac{h_H^{-1}\sum_{1\leq i,j\leq n}W_{ij}(\theta,x)H^{(1)}({h_H}^{-1}(y-Y_j))}{\sum_{1\leq i,j\leq n}W_{\theta,ij}(x)-\sum_{1\leq i,j\leq n}W_{ij}(\theta,x)H({h_H}^{-1}(y-Y_j))}.\tag{4}$$

2.1. Pointwise almost complete convergence

In the following, x (respectively y) denotes a fixed point in ($\mathcal{F}$ (respectively $\mathbb{R}$), N_x (respectively N_y)) is a fixed neighborhood of x (respectively of y).

Moreover, we denote by C_1, C_2, C_3 and $C_{\theta,x}$ some strictly positive constants and

$$K_i(\theta,x) := K(h_K^{-1} d_\theta(x,X_i)),\ \forall x \in \mathcal{H}, i = 1,...,n,$$

and

$$H_j(y) := H({h_H}^{-1}(y - Y_j)), \forall y \in \mathbb{R}, j = 1,...,n.$$

In order to establish the almost complete convergence (a.co.) of our estimator $\widehat{h}_\theta^x$, we need to include the following assumptions:

(H1) $\mathbb{P}(| < X - x, \theta > | < h_K) =: \phi_{\theta,x}(h_K) > 0.$

(H2) The operators F_θ^x and f_θ^x satisfy the Hölder condition:

(i) $\forall (x,y) \in \mathcal{H} \times S_\mathbb{R},\ \forall \theta \in \Theta, \exists 0 < \tau < 1,\ F_\theta^x(y) \leq 1 - \tau,$
$\exists b_1, b_2 > 0,\ \forall (y,y') \in S_\mathbb{R}^2, \forall (x,x') \in \mathcal{N}_x \times \mathcal{N}_x,\ \forall \theta \in \Theta,$

$$|F_\theta^x(y) - F_\theta^{x'}(y')| \leq C_{\theta,x}(||x - x'||^{b_1} + |y - y'|^{b_2}).$$

(ii) $\forall (x,y) \in \mathcal{H} \times S_\mathbb{R},\ \forall \theta \in \Theta, \exists \alpha < \infty,\ f_\theta^x(y) \leq \alpha,$

$$|f_\theta^x(y) - f_\theta^{x'}(y')| \leq C_{\theta,x}(||x - x'||^{b_1} + |y - y'|^{b_2}).$$

(H3) The functions $\varrho_\theta(\cdot,\cdot)$ and $\beta_\theta(\cdot,\cdot)$ are such that:

$$\forall z \in \mathcal{F}, \beta_\theta(z,z) = 0, |\varrho_\theta(x,z)| = d_\theta(x,z),$$

and

$$C_1|\varrho_\theta(x,z)| \leq |\beta_\theta(x,z)| \leq C_2|\varrho_\theta(x,z)|.$$

(H4) (i) K is a function with support $[-1,1]$ such that

$$0 < C_1 \mathbb{I}_{[-1,1]} < K(t) < C_2 \mathbb{I}_{[-1,1]} < \infty,$$

where $\mathbb{I}_A$ is the indicator function.

(ii) The kernel H is a differentiable function and $H^{(1)}$ is a positive, bounded, Lipschitzian continuous function such that:

$$\int |t|^{b_2} H^{(1)}(t)dt < \infty \quad \text{and} \int {H^{(1)}}^2(t)dt < \infty.$$

(H5) β satisfies:

(i)

$$\exists \eta_0 \in \mathbb{N},\ \forall \eta > \eta_0,\ \ -\frac{1}{\phi_{\theta,x}(\beta)} \int_{-1}^{1} \phi_{\theta,x}(t\beta,\beta) \frac{\mathrm{d}}{\mathrm{d}t}(t^2 K(t))\mathrm{d}t > C_3 > 0,$$

and

$$\beta \int_{B(x,h_K)} \beta_\theta(u,x)\mathrm{d}P(u) = o\Big(\int_{B(x,\beta)} \beta_\theta^2(u,x)\mathrm{d}P(u)\Big),$$

where $B(x,h)=\{z\in\mathcal{H}|\ d_\theta(z,x)\leq h\}$ denotes the closed ball centered at x and of radius h, and $\mathrm{d}P(u)$ is the probability measure of X.

(ii)

$$\lim_{n\to\infty}\frac{\ln n}{n\phi_{\theta,x}(\beta)}=0.$$

(H6) For some $\lambda>0$, the bandwidth h_H satisfies

$$\lim_{n\to\infty}n^\lambda h_H=\infty,\ \text{and}\ \lim_{n\to\infty}\frac{\ln n}{nh_H\phi_{\theta,x}(h_K)}=0.$$

(H7) $P(x)$ is continuous in a neighborhood of x such that $0<P(x)<1$.

3. Main results

Theorem 3.1.

Under assumptions (H1) - (H7), we have

$$|\widehat{h}_\theta^x(y)-h_\theta^x(y)|=O\left(h_K^{b_1}\right)+O\left(h_H^{b_2}\right)+O\left(\left(\frac{\log n}{nh_H\phi_{\theta,x}\left(h_K\right)}\right)^{\frac{1}{2}}\right),\ a.co.$$

Proof:

We have the following decomposition

$$\widehat{h}_\theta^x(y)-h_\theta^x(y)=\frac{1}{1-\widehat{F}_\theta^x(y)}\left[\widehat{f}_\theta^x(y)-f_\theta^x(y)\right]+\frac{h_\theta^x(y)}{1-\widehat{F}_\theta^x(y)}\left[\widehat{F}_\theta^x(y)-F_\theta^x(y)\right].$$

Then, the proof of Theorem 3.1 is a consequence of Theorem 3.2, Theorem 3.3 and the result (5) such that

$$\exists\eta>0,\sum_{n=1}^{\infty}\mathbb{P}\left\{|1-\widehat{F}_\theta^x(y)|<\eta\right\}<\infty. \tag{5}$$

■

Theorem 3.2.

Under assumptions (H1) - (H7), we have

$$|\widehat{F}_\theta^x(y)-F_\theta^x(y)|=O\left(h_K^{b_1}+h_H^{b_2}\right)+O\left(\left(\frac{\log n}{n\,\phi_{\theta,x}(h_K)}\right)^{\frac{1}{2}}\right),\ a.co.$$

Theorem 3.2 presents the complete convergence of the conditional distribution estimator. We note that, in the result of our work, we extended the complete data presented in Demongeot et al. (2014) to MAR case. The completely observed data is obtained by taking $\delta=1$ in our case of study.

Proof:

The proof of this Theorem is based on the present decomposition

$$\widehat{F}_{\theta}^{x}(y) - F_{\theta}^{x}(y) = \widehat{B}_{n,1}(\theta,x) + \frac{\widehat{R}_{n,1}(\theta,x)}{\widehat{F}_D(\theta,x)} + \frac{\widehat{Q}_{n,1}(\theta,x)}{\widehat{F}_D(\theta,x)}, \tag{6}$$

where

$$\widehat{Q}_{n,1}(\theta,x) = (\widehat{F}_{\theta,N}^{x}(y) - \mathbb{E}\widehat{F}_{\theta,N}^{x}(y)) - F_{\theta}^{x}(y)(\widehat{F}_D(\theta,x) - \mathbb{E}\widehat{F}_D(\theta,x)),$$
$$\widehat{B}_{n,1}(\theta,x) = \frac{\mathbb{E}\widehat{F}_{\theta,N}^{x}(y)}{\mathbb{E}\widehat{F}_D(\theta,x)} - F_{\theta}^{x}(y),$$
$$\widehat{R}_{n,1}(\theta,x) = -\widehat{B}_{n,1}(\theta,x)(\widehat{F}_D(\theta,x) - \mathbb{E}\widehat{F}_D(\theta,x)),$$

with

$$\widehat{F}_{\theta,N}^{x}(y) = \frac{1}{n(n-1)\mathbb{E}[W_{12}(\theta,x)]} \sum_{1\le i,j\le n} W_{ij}(\theta,x) H(h_H^{-1}(Y-y_j)),$$

and

$$\widehat{F}_D(\theta,x) = \frac{1}{n(n-1)\mathbb{E}[W_{12}(\theta,x)]} \sum_{i\neq j} W_{ij}(\theta,x).$$ ∎

Lemma 3.1.

Under assumptions (H1) and (H3) - (H7), we have that

$$|\widehat{F}_D(\theta,x) - \mathbb{E}\widehat{F}_D(\theta,x)| = O\left(\left(\frac{\log n}{n\phi_{\theta,x}(h_K)}\right)^{\frac{1}{2}}\right),\ a.co., \tag{7}$$

and

$$|\widehat{F}_{\theta,N}^{x}(y) - \mathbb{E}[\widehat{F}_{\theta,N}^{x}(y)]| = O\left(\frac{\log n}{n\,\phi_{\theta,x}(h_K)}\right)^{\frac{1}{2}},\ \ a.co. \tag{8}$$

Corollary 3.1.

Under assumptions of Lemma 3.1, we have:

$$\exists C_1 > 0,\ \text{such that} \sum_n \mathbb{P}\left(\widehat{F}_D(\theta,x) < C_1\right) < \infty. \tag{9}$$

Lemma 3.2.

Under assumptions (H1), (H2), (H4), (H5) and (H7), we obtain:

$$|\widehat{B}_{n,1}(\theta,x)| = O\left(h_K^{b_1}\right) + O\left(h_H^{b_2}\right),\ \ a.co. \tag{10}$$

Theorem 3.3.

Under assumptions (H1) - (H7), we have

$$|\widehat{f}^x_\theta(y) - f^x_\theta(y)| = O\left(h_K^{b_1} + h_H^{b_2}\right) + O\left(\left(\frac{\log n}{n\, h_H \phi_{\theta,x}(h_K)}\right)^{\frac{1}{2}}\right), \ a.co.$$

Theorem 3.3 presents the complete convergence of the conditional density estimator. Also, note that the current result is a generalization of the work introduced by Demongeot et al. (2010) for the case of missing at random.

Proof:

The next decomposition given in (11) affirms that the proof of Theorem 3.3 can be deduced from the lemmas and their proofs cited:

$$\widehat{f}^x_\theta(y) - f^x_\theta(y) = \widehat{B}_{n,2}(\theta,x) + \frac{\widehat{R}_{n,2}(\theta,x)}{\widehat{F}_D(\theta,x)} + \frac{\widehat{Q}_{n,2}(\theta,x)}{\widehat{F}_D(\theta,x)}, \tag{11}$$

where

$$\widehat{Q}_{n,2}(\theta,x) = (\widehat{f}^x_{\theta,N}(y) - \mathbb{E}\widehat{f}^x_{\theta,N}(y)) - f^x_\theta(y)(\widehat{F}_D(\theta,x) - \mathbb{E}\widehat{F}_D(\theta,x)),$$
$$\widehat{B}_{n,2}(\theta,x) = \frac{\mathbb{E}\widehat{f}^x_{\theta,N}(y)}{\mathbb{E}\widehat{F}_D(\theta,x)} - f^x_\theta(y),$$
$$\widehat{R}_{n,2}(\theta,x) = -\widehat{B}_{n,2}(\theta,x)(\widehat{F}_D(\theta,x) - \mathbb{E}\widehat{F}_D(\theta,x)),$$

with

$$\widehat{f}^x_{\theta,N}(y) = \frac{1}{(nh_K\phi_{\theta,x}(h_K))^2 h_H}\sum_{1\leq i,j\leq n} W_{ij}(\theta,x)H^{(1)}(h_H^{-1}(Y-y_j)),$$

$$\widehat{f}^x_{\theta,N}(y) = \frac{1}{n(n-1)\mathbb{E}[W_{12}(\theta,x)]}\sum_{i\neq j} W_{ij}(\theta,x)H^{(1)}(h_H^{-1}(Y-y_j)).$$ ■

Lemma 3.3.

Under the assumptions (H1) and (H3) - (H7), we get:

$$|\widehat{f}^x_{\theta,N}(y) - \mathbb{E}[\widehat{f}^x_{\theta,N}(y)]| = O\left(\frac{\log n}{nh_H\phi_{\theta,x}(h_K)}\right)^{\frac{1}{2}}, \ \ a.co. \tag{12}$$

Lemma 3.4.

Under assumptions (H1), (H2), (H4), (H5) and (H7), we obtain:

$$|\widehat{B}_{n,2}(\theta,x)| = O\left(h_K^{b_1}\right) + O\left(h_H^{b_2}\right), \ \ a.co. \tag{13}$$

Further, we define the following quantity, for any $x \in \mathcal{F}$, and for all $i = 1, ..., n$ by:

$$W_{\theta,ij} = W_{ij}(\theta, x).$$

Proof:

We follow the proof of Lemma 2 given in Barrientos-Marin et al. (2010), by writing:

$$\widehat{f}^x_{\theta,N}(y) = T_1(\theta, x)(T_2(\theta, x)T_3(\theta, x) - T_4(\theta, x)T_5(\theta, x)),$$

$$\widehat{F}^x_{\theta,D} = \underbrace{\frac{n^2 h_K^2 \phi^2_{\theta,x}(h_K)}{n(n-1)\mathbb{E}[W_{\theta,12}]}}_{T_1}(\theta, x)\Big[\underbrace{\Big(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j(\theta, x)\delta_j}{\phi_{\theta,x}(h_K)}\Big)}_{T_2(\theta,x)}\underbrace{\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i(\theta, x)\beta_i^2(\theta, x)\delta_i}{h_K^2\phi_{\theta,x}(h_K)}\Big)}_{T_3(\theta,x)}$$
$$-\underbrace{\Big(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j(\theta, x)\beta_j(\theta, x)\delta_j}{h_K\phi_{\theta,x}(h_K)}\Big)}_{T_4(\theta,x)}\underbrace{\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i(\theta, x)\beta_i(\theta, x)\delta_i}{h_K\phi_{\theta,x}(h_K)}\Big)}_{T_5(\theta,x)}\Big],$$

and

$$\widehat{F}^x_{\theta,N}(y) = \underbrace{\frac{n^2 h_K^2 \phi^2_{\theta,x}(h_K)}{n(n-1)\mathbb{E}[W_{\theta,12}]}}_{I_1}(\theta, x)\Big[\underbrace{\Big(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j(\theta, x)\delta_j H_j}{\phi_{\theta,x}(h_K)}\Big)}_{I_2(\theta,x)}\underbrace{\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{K_i(\theta, x)\beta_i^2(\theta, x)\delta_i}{h_K^2\phi_{\theta,x}(h_K)}\Big)}_{I_3(\theta,x)}$$
$$-\underbrace{\Big(\frac{1}{n}\sum_{j=1}^{n}\frac{K_j(\theta, x)\beta_j(\theta, x)\delta_j H_j}{h_K\phi_{\theta,x}(h_K)}\Big)}_{I_4(\theta,x)}\underbrace{\Big(\frac{1}{n}\sum_{i=1}^{n}\frac{K_{\theta,i}(x)\beta_i(\theta, x)\delta_i}{h_K\phi_{\theta,x}(h_K)}\Big)}_{I_5(\theta,x)}\Big].$$

Therefore, we write for all $i \neq j$,

$$\begin{aligned}T_i(\theta, x)T_j(\theta, x) - \mathbb{E}[T_i(\theta, x)T_j(\theta, x)] &= (T_i(\theta, x) - \mathbb{E}[T_i(\theta, x)])\,(T_j(\theta, x) - \mathbb{E}[T_j(\theta, x)]) \\ &+ (T_j(\theta, x) - \mathbb{E}[T_j(\theta, x)])\,\mathbb{E}[T_i(\theta, x)] \\ &+ (T_i(\theta, x) - \mathbb{E}[T_i(\theta, x)])\,\mathbb{E}[T_j(\theta, x)] \\ &+ \mathbb{E}[T_i(\theta, x)]\mathbb{E}[T_j(\theta, x)] - \mathbb{E}[T_i(\theta, x)T_j(\theta, x)],\end{aligned}$$

and

$$\begin{aligned}I_i(\theta, x)I_j(\theta, x) - \mathbb{E}[I_i(\theta, x)I_j(\theta, x)] &= (I_i(\theta, x) - \mathbb{E}[I_i(\theta, x)])\,(I_j(\theta, x) - \mathbb{E}[I_j(\theta, x)]) \\ &+ (I_j(\theta, x) - \mathbb{E}[I_j(\theta, x)])\,\mathbb{E}[I_i(\theta, x)] \\ &+ (I_i(\theta, x) - \mathbb{E}[I_i(\theta, x)])\,\mathbb{E}[I_j(\theta, x)] \\ &+ \mathbb{E}[I_i(\theta, x)]\mathbb{E}[I_j(\theta, x)] - \mathbb{E}[I_i(\theta, x)I_j(\theta, x)].\end{aligned}$$

So, all that remains is, similar to Barrientos-Marin et al. (2010), as soon as the following statements are checked, the claimed result will be obtained:

$$I^x_{\theta,i} - \mathbb{E}(I^x_{\theta,i}) = O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_k)}}\right) \qquad \text{For } i = 2,3,4,5\,, \tag{14}$$

$$\mathbb{E}(I^x_{\theta,i}) = O(1) \quad \text{and} \quad \mathbb{E}(T^x_{\theta,j}) = O(1) \quad \text{For } i = 2,3,4,5\,, \tag{15}$$

$$cov(I^x_{\theta,2}, I^x_{\theta,3}) = o\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_k)}}\right)\ ;\ cov(I^x_{\theta,4}, I^x_{\theta,5}) = o\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_k)}}\right), \tag{16}$$

$$cov(T^x_{\theta,2}, T^x_{\theta,3}) = o\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_k)}}\right)\ ;\ cov(T^x_{\theta,4}, T^x_{\theta,5}) = o\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_k)}}\right). \tag{17}$$

Let us show the result (14). We set

$$I^x_{\theta,l} - \mathbb{E}[I^x_{\theta,l}] = \frac{1}{n}\sum_{i=1}^{n} Z^l_{\theta,i} \text{ for } l = 0,1,2,$$

where

$$Z^l_{\theta,i} = \frac{1}{h^l_K h_H \phi_{\theta,x}(h_K)}\left(\delta_i K_i(\theta,x)H_i\beta^l_i(\theta,x) - \mathbb{E}\left[\delta_i K_i(\theta,x)H_i\beta^l_i(\theta,x)\right]\right).$$

Thus, it remains to check that

$$I^x_{\theta,l} - \mathbb{E}[I^x_{\theta,l}] = O_{a.co.}\left(\sqrt{\frac{\log n}{n\,\phi_{\theta,x}(h_K)}}\right),\ \text{for } l = 0,1,2\,.$$

By (H3), we have $\frac{1}{h^l_K}\left(\delta_i K_i(\theta,x)\beta^l_i(\theta,x)\right) < C_1$, and since $H < 1$, we can write

$$|Z^l_{\theta,i}| \leq \frac{C_1}{\phi_{\theta,x}(h_K)}.$$

For $\mathbb{E}[Z^l_{\theta,i}]^2$, we have

$$\begin{gathered}\mathbb{E}\left(\delta_i K_i(\theta,x)H_i\beta^l_i(\theta,x) - \mathbb{E}\left[\delta_i K_i(\theta,x)H_i\beta^l_i(\theta,x)\right]\right)^2 = \\ \mathbb{E}\left(\delta_i K^2_i(\theta,x)H^2_i\beta^{2l}_i(\theta,x)\right) - \left(\mathbb{E}\left[\delta_i K_i(\theta,x)H_i\beta^l_i(\theta,x)\right]\right)^2 = \\ \mathbb{E}\left(K^2_i(\theta,x)\beta^{2l}_i(\theta,x)\mathbb{E}(\delta_i H^2_i| < \theta, X_i >)\right) - \left(\mathbb{E}\left[K_i(\theta,x)\beta^l_i(\theta,x)\mathbb{E}(\delta_i H_i| < \theta, X_i >)\right]\right)^2.\end{gathered}$$

Since the variables δ_i and Y_i are independent given X_i, then under the hypothesis (H7), we get

$$\mathbb{E}(\delta_i H^2_i| < \theta, X_i >) = (P(X) + o(1))\mathbb{E}(H^2_i| < \theta, X_i >) \leq C_1,$$

and

$$\mathbb{E}(\delta_i H_i| < \theta, X_i >) = (P(X) + o(1))\mathbb{E}(H_i| < \theta, X_i >) \leq C_2.$$

Finally, it obtained that

$$\mathbb{E}[Z^l_{\theta,i}]^2 \leq \frac{C_2}{\phi_{\theta,x}(h_K)}.$$

Therefore, it is sufficient to apply the classical inequality of Bernstein to all $\eta \in (0, \frac{C_2}{C_1})$. We can then write:

$$\mathbb{P}\left\{\left|I^x_{\theta,l} - \mathbb{E}[I^x_{\theta,l}]\right| > \eta\sqrt{\frac{\log n}{n\,\phi_{\theta,x}(h_K)}}\right\} \;\leq\; C_2 n^{-C_1\eta^2}.$$

Finally, an appropriate choice of η permits to deduce that:

$$\mathbb{P}\left\{\left|I^x_{\theta,l} - \mathbb{E}[I^x_{\theta,l}]\right| > \eta\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right\} \;\leq\; C_2 n^{-1-\gamma},\text{ for } l = 0, 1, 2,$$

$$\mathbb{P}\left\{\left|T^x_{\theta,l} - \mathbb{E}[T^x_{\theta,l}]\right| > \eta\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right\} \;\leq\; C_2 n^{-1-\gamma},\text{ for } l = 0, 1, 2.$$

Regarding the proof of (14), it will be stated by the same steps used in Barrientos-Marin et al. (2010). ■

Proof (of the second result):

To prove these results, we focus on the fact that the observations (δ_i, X_i, Y_i) for $i = 1, ..., n$ are identically distributed.

Therefore,

$$\mathbb{E}(T^x_{\theta,2}) = \frac{\mathbb{E}(\delta_2 K_1(\theta,x))}{\phi_{\theta,x}(h_K)}, \tag{18}$$

$$\mathbb{E}(T^x_{\theta,3}) = \mathbb{E}(I^x_{\theta,3}) = \frac{\mathbb{E}(\delta_1 K_1(\theta,x)\beta_1^2(\theta,x))}{h^2\phi_{\theta,x}(h_K)}, \tag{19}$$

$$\mathbb{E}(T^x_{\theta,4}) = \mathbb{E}(T^x_{\theta,5}) = \mathbb{E}(I^x_{\theta,5}) = \frac{\mathbb{E}(\delta_1 K_1(\theta,x)\beta_1(\theta,x))}{h_k\phi_{\theta,x}(h_K)}, \tag{20}$$

$$\mathbb{E}(I^x_{\theta,2}) = \frac{\mathbb{E}(\delta_1 K_1(\theta,x)H_1)}{\phi_{\theta,x}(h_K)}\;;\;\mathbb{E}(I^x_{\theta,4}) = \frac{\mathbb{E}(\delta_1 K_1(\theta,x)H_1\beta_1(\theta,x))}{h_k\phi_{\theta,x}(h_K)}. \tag{21}$$

The proof is reduced to evaluate the following quantities:

$$\mathbb{E}(\delta_1 K_1(\theta,x)H_1^k\beta_1^l(\theta,x)) \quad \text{for } l = 0,1,2 \quad \text{and } k = 0,1.$$

Again, we are using the fact that the variables δ and Y are conditionally independent with respect to the functional variables X. Therefore, for all $l = 0, 1, 2$ and under hypothesis (H7), we get

$$\begin{aligned}\mathbb{E}(\delta_1 K_1(\theta,x)H_1^k\beta_1^l(\theta,x)) &= \mathbb{E}(\mathbb{E}[\delta_1| <\theta, X_1>]K_1(\theta,x)H_1^k\beta_1^l(\theta,x)) \\ &= (P(X_1)+o(1))\mathbb{E}(H_1^kK_1(\theta,x)\beta_1^l(\theta,x)) \\ &\leq C_1 h_k^l\phi_{\theta,x}(h_k) = O(h_k^l\phi_{\theta,x}(h_k)).\end{aligned}$$

Now, we move to the proof of the results of (16) and (17). For both equations, we are using the fact that $(\delta_i, X_i, Y_i)_{i=1\dots n}$ are identically distributed. We find that:

$$\begin{aligned}Cov(I_2, I_3) = \frac{1}{nh_K^2\phi^2_{\theta,x}(h_K)}&\left[\mathbb{E}[\delta_1 K_1(\theta,x)^2H_1\beta_1^2(\theta,x)]-\right. \\ &\left.\mathbb{E}[\delta_1K_1(\theta,x)H_1]\mathbb{E}[\delta_1K_1(\theta,x)\beta_1^2(\theta,x)]\right],\end{aligned}$$

and

$$Cov(I_4, I_5) = \frac{1}{nh_K^2 \phi_{\theta,x}^2(h_K)} \left[\mathbb{E}[\delta_1 K_1^2(\theta, x) H_1 \beta_1^2(\theta, x)] - \right.$$
$$\left. \mathbb{E}[\delta_1 K_1(\theta, x) H_1 \beta_1(\theta, x)] \mathbb{E}[\delta_1 K_1(\theta, x) \beta_1(\theta, x)] \right].$$

By the same procedure used before, we have:

$$\begin{aligned}
\mathbb{E}\left[\delta_i K_i^2(\theta, x) H_i^k \beta_i^l(\theta, x)\right] &= \mathbb{E}\left[\mathbb{E}(\delta_i | <\theta, x>) K_i^2(\theta, x) H_i^k \beta_i^l(\theta, x)\right] \\
&= (P(X) + o(1)) \mathbb{E}\left[H_i^k K_i^2(\theta, x) \beta_i^l(\theta, x)\right] \\
&\leq C_1 h_k^l \phi_{\theta,x}(h_k) = O(h^l \phi_{\theta,x}(h_K)),
\end{aligned}$$

which implies that,

$$Cov(I_2, I_3) = O\left(\frac{1}{n\phi_{\theta,x}(h_K)}\right) = o\left(\frac{\log n}{n\phi_{\theta,x}(h_K)}\right),$$

and

$$Cov(I_4, I_5) = O\left(\frac{1}{n\phi_{\theta,x}(h_K)}\right) = o\left(\frac{\log n}{n\phi_{\theta,x}(h_K)}\right). \qquad \blacksquare$$

Proof (of Corollary 3.1):

We have

$$\begin{aligned}
\mathbb{E}[W_{\theta,12}] &= \mathbb{E}[\beta_{\theta,1}^2 \delta_1 \delta_2 K_{\theta,1}(x) K_{\theta,2}(x) - \beta_{\theta,1} \beta_{\theta,2} \delta_1 \delta_2 K_{\theta,1}(x) K_{\theta,2}(x)] \\
&= \mathbb{E}[\beta_{\theta,1}^2 \delta_1 K_{\theta,1}(x) \delta_2 K_{\theta,2}(x)] - \mathbb{E}[\beta_{\theta,1} \delta_1 K_{\theta,1}(x) \beta_{\theta,2} \delta_2 K_{\theta,2}(x)].
\end{aligned}$$

We also have

$$\mathbb{P}(\delta_2 = 1 | \langle \theta, X_2 \rangle) = P(X_2) \quad \text{and} \quad \mathbb{P}(\delta_1 = 1 | \langle \theta, X_1 \rangle) = P(X_1).$$

Then,

$$\begin{aligned}
\mathbb{E}[\beta_{\theta,1}^2 \delta_1 K_{\theta,1}(x) \delta_2 K_{\theta,2}(x)] &= \underbrace{\Big(P(X_1) + o(1)\Big)\Big(P(X_2) + o(1)\Big)}_{C_1} \\
&\quad \mathbb{E}\big(\beta_{\theta,1}^2(x) K_{\theta,1}(x)\big) \mathbb{E}\big(K_{\theta,2}(x)\big) \\
&\geq C_1 \big(C_1 h^2 \phi_{\theta,x}(h)\big) \phi_{\theta,x}(h) \\
&\geq C_1 h^2 \phi_{\theta,x}^2(h) = o(h^2 \phi_{\theta,x}^2(h)),
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[\beta_{\theta,1} \beta_{\theta,2} \delta_1 \delta_2 K_{\theta,1}(x) K_{\theta,2}(x)] &= \underbrace{\Big(P(X_1) + o(1)\Big)\Big(P(X_2) + o(1)\Big)}_{C_1} \\
&\quad \mathbb{E}\big(\beta_{\theta,1}(x) K_{\theta,1}(x)\big) \mathbb{E}\big(\beta_{\theta,2}(x) K_{\theta,2}(x)\big) \\
&\leq C_1 \mathbb{E}^2\big(\beta_{\theta,1}(x) K_{\theta,1}(x)\big) \\
&\leq C_1 h_K^2 \phi_{\theta,x}^2(K) = o(h_K^2 \phi_{\theta,x}^2(h_K)).
\end{aligned}$$

Therefore,

$$\mathbb{E}[W_{\theta,12}] = \mathbb{E}\big(\beta_{\theta,1}^2(x) K_{\theta,1}(x)\big) \mathbb{E}\big(K_{\theta,2}(x)\big) - \mathbb{E}^2\big(\beta_{\theta,1}(x) K_{\theta,1}(x)\big) \geq C_1 h_K^2 \phi_{\theta,x}^2(h_K).$$

By applying the result we found above, we can deduce that:

$$\sum_{n=1}^{\infty}\mathbb{P}\Big(\hat{F}_D(\theta,x)<C_1\Big)<\infty. \qquad \blacksquare$$

Proof (of Lemma 3.2):

We have

$$\widehat{B}_{n,1}(\theta,x)=\frac{\mathbb{E}[W_{\theta,12}(H_2-F^x_\theta(y))]}{\mathbb{E}[W_{\theta,12}]},$$

and

$$\widehat{B}_{n,1}(\theta,x)=\frac{\mathbb{E}[W_{\theta,12}\left(\mathbb{E}[H_2|<\theta,X_2>]-F^x_\theta(y)\right)]}{\mathbb{E}[W_{\theta,12}]},$$

with

$$\mathbb{E}[\widehat{F}^x_\theta(y)]=\frac{1}{\mathbb{E}[W_{\theta,12}]}\mathbb{E}\left(W_{\theta,12}\mathbb{E}[H_2^{(1)}|<\theta,X_2>]\right).$$

Since (X_i,δ_i,Y_i) are identically distributed, we have that

$$\mathbb{P}(\delta_2=1|<\theta,X_2>)=P(X_2),$$

then,

$$\begin{aligned}\mathbb{E}\left[W_{\theta,12}\right]&=\mathbb{E}\big[\delta_1\beta_1^2(\theta,x)K_1(\theta,x)\delta_2\beta_2(\theta,x)K_2(\theta,x)\\&\quad-\delta_1\beta_1(\theta,x)K_1(\theta,x)\delta_2\beta_2(\theta,x)K_2(\theta,x)\big]\\&=\mathbb{E}\big[\mathbb{E}\big(\delta_1\beta_1^2(\theta,x)K_1(\theta,x)\delta_2\beta_2(\theta,x)K_2(\theta,x)\\&\quad-\delta_1\beta_1(\theta,x)K_1(\theta,x)\delta_2\beta_2(\theta,x)K_2(\theta,x)|<\theta,X_2>)\big]\\&=\mathbb{E}\big[\delta_1\beta_1^2(\theta,x)K_1(\theta,x)\beta_2(\theta,x)K_2(\theta,x)P(X_2)\\&\quad-\delta_1\beta_1(\theta,x)K_1(\theta,x)\beta_2(\theta,x)K_2(\theta,x)P(X_2)\big],\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\left[W_{\theta,12}(\mathbb{E}[H_2|<\theta,X_2>]-F^x_\theta(y))\right]=\\&\mathbb{E}[\delta_1\beta_1^2(\theta,x)K_1(\theta,x)\beta_2(\theta,x)K_2(\theta,x)P(X_2)(\mathbb{E}[H_2|<\theta,X_2>]-F^x_\theta(y))\\&-\delta_1\beta_1(\theta,x)K_1(\theta,x)\beta_2(\theta,x)K_2(\theta,x)P(X_2)(\mathbb{E}[H_2|<\theta,X_2>]-F^x_\theta(y))].\end{aligned}$$

We use an integration by parts to show that

$$\mathbb{E}[H_2|<\theta,X_2>]=h_H^{-1}\int_{\mathbb{R}}H^{(1)}(h_H^{-1}(y-z))F_\theta^{X_2}(z)dz.$$

Now, with the change of variables $t=\frac{y-z}{h_H}$ we can write:

$$\mathbb{E}[H_2|<\theta,X_2>]=\int_{\mathbb{R}}H^{(1)}(t)F_\theta^{X_2}(y-h_Ht)dt.$$

Therefore,

$$|\mathbb{E}[H_2| < \theta, X_2 >] - F_\theta^x(y)| \le \int_{\mathbb{R}} H^{(1)}(t)|F_\theta^{X_2}(y - h_H t) - F_\theta^{X_2}(y)|dt.$$

Further, as $K_{\theta,1}(.)$ has a compact support in (0,1), then, with the Lipschtzian assumption $(H2)$, we can write:

$$\mathbb{I}_{B(x,h_K)}(X_2)|\mathbb{E}[H_2|X_2] - F^x(y)| \le C_{\theta,x}\int_{\mathbb{R}} H^{(1)}(t)\big(h_K^{b_1} + |t|^{b_2}h_H^{b_2}\big)dt = O(h_K^{b_1}) + O(h_H^{b_2}).$$

Since $H^{(1)}$ is a probability density, the claimed result of this lemma is then a direct consequence of the assumption (H5). ∎

Proof (of Lemma 3.3):

With the same concept used in the proof of Lemma 3.1, we prove this lemma. We note that

$$\widehat{f}_{\theta,N}^x(y) = \underbrace{\frac{n^2h_K^2\phi_{\theta,x}^2(h_K)}{n(n-1)\mathbb{E}(W_{\theta,12})}}_{S_1}\Bigg[\underbrace{\left(\frac{1}{n}\sum_{j=1}^n\frac{\delta_J K_j(\theta,x)H_j^{(1)}}{h_H\phi_{\theta,x}(h_K)}\right)}_{S_2}\underbrace{\left(\frac{1}{n}\sum_{i=1}^n\frac{\delta_i K_i(\theta,x)\beta_i^2(\theta,x)}{h_K^2\phi_{\theta,x}(h_K)}\right)}_{I_3}$$
$$-\underbrace{\left(\frac{1}{n}\sum_{j=1}^n\frac{\delta_J K_j(\theta,x)\beta_j(\theta,x)H_j^{(1)}}{h_H h_K\phi_{\theta,x}(h_K)}\right)}_{S_4}\underbrace{\left(\frac{1}{n}\sum_{i=1}^n\frac{\delta_i K_i(\theta,x)\beta_i(\theta,x)}{h_K\phi_{\theta,x}(h_K)}\right)}_{I_5}\Bigg],$$

and with the same method used above, we show that

$$S_{\theta,i}^x - \mathbb{E}[S_{\theta,i}^x] = O_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H\phi_{\theta,x}(h_K)}}\right), \text{ for } i = 2,4,$$
$$\mathbb{E}[S_{\theta,l}^x] = O(1), \text{ for } l = 2,4,$$
$$Cov(S_2, I_3) = o\left(\sqrt{\frac{\log n}{n\,h_H\phi_{\theta,x}(h_K)}}\right),$$
$$\text{and } Cov(S_4, I_5) = o\left(\sqrt{\frac{\log n}{n\,h_H\phi_{\theta,x}(h_K)}}\right).$$

We take the following variable,

$$S_{l,k}(\theta,x) - \mathbb{E}[S_{l,k}(\theta,x)] = \frac{1}{n\phi_{\theta,x}(h_K)}\sum_{i=1}^n Z_{\theta,i}^{l,k} \text{ for } l = 0,1,2, \text{ and } k = 0,1\,,$$

where

$$Z_{\theta,i}^{l,k} = \frac{1}{h_K^l h_H^k\phi_{\theta,x}(h_K)}\left(\delta_i K_i(\theta,x)H_i^{\prime k}\beta_i^l(\theta,x) - \mathbb{E}\left[\delta_i K_i(\theta,x)H_i^{\prime^k}\beta_i^l(\theta,x)\right]\right).$$

With simple computations, we prove that

$$|Z_{\theta,i}^{l,k}| \le \frac{C_1}{h_H\phi_{\theta,x}(h_K)} \text{ and } \mathbb{E}[Z_{\theta,i}^{l,k}]^2 \le \frac{C_2}{h_H\phi_{\theta,x}(h_K)}.$$

For all $l = 0, 1, 2$, and $k = 0, 1$, we have

$$\mathbb{E}\left[\delta_1 K_1(\theta, x) H_1^{'^k} \beta_1^l(\theta, x)\right] = (P(X) + o(1))\mathbb{E}\left[K_1(\theta, x) H_1^{'^k} \beta_1^l(\theta, x)\right]$$
$$= O(h_K^l h_H \phi_{\theta,x}(h_K)).$$

This proves that

$$\mathbb{E}[S_{\theta,2}^x] = \mathbb{E}[S_{\theta,4}^x] = O(1) \text{ and for } i = 2, 4,$$

$$S_{\theta,i}^x - \mathbb{E}[S_{\theta,i}^x] = O\left(\sqrt{\frac{\log n}{n\, h_H \phi_{\theta,x}(h_K)}}\right).$$

Also, we obtain

$$\mathbb{E}\left[\delta_i K_i^2(\theta, x) H_i^{'^k} \beta_i^l(\theta, x)\right] = \mathbb{E}\left[\mathbb{E}(\delta_i| < \theta, x >) K_i^2(\theta, x) H_i^{'^k} \beta_i^l(\theta, x)\right] = O(h_K^l h_H \phi_{\theta,x}(h_K)),$$

which implies that

$$Cov(S_{\theta,2}^x, I_{\theta,3}^x) = O\left(\frac{1}{nh_H \phi_{\theta,x}(h_K)}\right) = o\left(\frac{\log n}{nh_H \phi_{\theta,x}(h_K)}\right),$$

and

$$Cov(S_{\theta,4}^x, I_{\theta,5}^x) = O\left(\frac{1}{nh_H \phi_{\theta,x}(h_K)}\right) = o\left(\frac{\log n}{nh_H \phi_{\theta,x}(h_K)}\right).$$

■

Proof (of Lemma 3.4):

We have

$$\widehat{B}_{n,2}(\theta, x) = \frac{\mathbb{E}\widehat{f}_{\theta,N}^x(y) - \mathbb{E}[\widehat{F}_D(\theta, x)] f_\theta^x(y)}{\mathbb{E}\widehat{F}_D(\theta, x)}.$$

Then,

$$\widehat{B}_{n,2}(\theta, x) = \frac{\mathbb{E}[W_{\theta,12}(h_H^{-1} H_2^{(1)} - f_\theta^x(y))]}{\mathbb{E}[W_{\theta,12}]}.$$

Since (X_i, δ_i, Y_i) are identically distributed, we obtain

$$\widehat{B}_{n,2}(\theta, x) = \frac{\mathbb{E}[W_{\theta,12}(h_H^{-1}\mathbb{E}[H_2^{(1)}|\langle\theta, X_2\rangle] - f_\theta^x(y))]}{\mathbb{E}[W_{\theta,12}]}.$$

Then,

$$\mathbb{E}\left[W_{\theta,12}\left(h_H^{-1}\mathbb{E}[H_2^{(1)}|\langle\theta, X_2\rangle] - f_\theta^x(y)\right)\right] =$$
$$\mathbb{E}\Big[\delta_1 \beta_1^2(\theta, x) K_1(\theta, x) K_2(\theta, x) P(X_2)(h_H^{-1}\mathbb{E}[H_2^{(1)}|\langle\theta, X_2\rangle] - f_\theta^x(y))$$
$$-\delta_1 \beta_1(\theta, x) K_1(\theta, x) \beta_2(\theta, x) K_2(\theta, x) P(X_2)(h_H^{-1}\mathbb{E}[H_2^{(1)}|\langle\theta, X_2\rangle] - f_\theta^x(y))\Big].$$

Under assumption (H4) and by the classical change of variables $t = \frac{y-z}{h_H}$, we obtain:

$$h_H^{-1}\mathbb{E}[H_2^{(1)}|\langle\theta, X_2\rangle] = \int_{\mathbb{R}} H^{(1)}(t) f_\theta^{X_2}(y - h_H t)dt.$$

Therefore,

$$|h_H^{-1}\mathbb{E}[H_2^{(1)}|\langle\theta, X_2\rangle] - f_\theta^x(y)| \leq \int_{\mathbb{R}} H^{(1)}(t)|f_\theta^{X_2}(y - h_H t) - f_\theta^x(y)|dt.$$

Thus, by the assumption (H2) we get:

$$\mathbb{I}_{B(x,h_K)}(X)|\mathbb{E}[H^{(1)}(y)|\langle\theta, X_2\rangle] - f_\theta^x(y)| \leq C_{\theta,x}\int_{\mathbb{R}} H^{(1)}(t)\big(h_K^{b_1} + |t|^{b_2}h_H^{b_2}\big)dt = O(h_K^{b_1}) + O(h_H^{b_2}).$$

Since $H^{(1)}$ is a probability density, then the claimed result of this lemma is a direct consequence of the assumption (H5). ■

Proof (of* (5)*):

We can write

$$|1 - \widehat{F}_\theta^x(y)| \leq (1 - F_\theta,^x(y))/2 \Rightarrow |\widehat{F}_\theta^x(y) - F_\theta^x(y)| \geq F_\theta^x(y)/2,$$

so that we are getting to

$$\mathbb{P}\{|1 - \widehat{F}_\theta^x(y)| < (1 - F_\theta^x(y))/2\} \leq \mathbb{P}\{|\widehat{F}_\theta^x(y) - F_\theta^x(y)| \geq F_\theta^x(y)/2\}.$$

It is sufficient to take $\eta = (1 - F_\theta^x(y))/2$ to demonstrate the result. ■

4. A numerical study

In this section, we illustrate by evaluate the performance of our estimation approach using the single index dimensional reduce model and with respect to the percentage of missing observations. For this goal, we generate the data by the following simulation.

The process $X_i(t)$ are defined, for any $t \in [0, 1]$, by:

$$X_i(t) = a_i t^2 + b_i(1 - \sin(2\pi t)) + c_i,$$

where $a_i \rightsquigarrow \mathcal{N}(0, 1)$, $b_i \rightsquigarrow \mathcal{N}(1, 0.5)$ and $c_i \rightsquigarrow \mathcal{N}(1, 0.5)$. For simplicity, Figure 1 presents a sample of $n = 100$ of the curves $X(t)$.

Now we choose the single index model θ as the eigenvectors of variance-covariance matrix of the simulated data. We consider the following choice of the locating functions $\varrho(.,.)$ and $\beta(.,.)$,

$$\beta(x_1, x_2) =< \theta, x_1 - x_2 >_{\mathcal{F}} = \frac{||\theta||^2_{\mathcal{F}} + ||x_1 - x_2||^2_{\mathcal{F}} - ||\theta - x_1 - x_2||^2_{\mathcal{F}}}{2},$$

and $\varrho(x_1, x_2) = ||\theta - x_1 - x_2||_{\mathcal{F}}$.

We take for our application,

$$\beta(x_1, x_2) =< \theta, x_1 - x_2 >_{\mathcal{F}} \text{ and } \varrho(x_1, x_2) = ||x_1 - x_2||_{\mathcal{F}}.$$

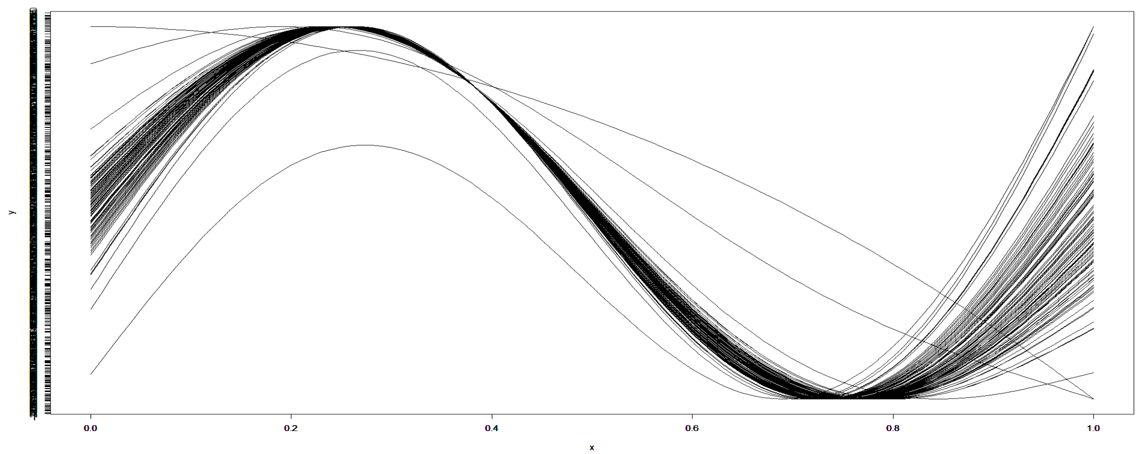

Figure 1. 100 curves of X_i

The concentration of the probability measure is quantified here with respect to the bi-functional operator $\varrho(.,.)$ which can be related to the topological structure of the functional space $\mathcal{F}$ by taking $d = |\varrho(.,.)|$.

We use the semi-metric defined by the L^2-distance between the curves,

$$|d(x_1, x_2)| = \sqrt{\int_0^1 \left(x_1^{(1)}(t) - x_2^{(1)}(t) \right)^2 dt}.$$

The response variable Y is calculated by the following model:

$$Y = R(X) + \varepsilon,$$

where $\varepsilon \sim \mathcal{N}(0, .5)$. The operator $R(\cdot)$ is defined by:

$$R(X) = \frac{1}{1 + \int_0^1 X(t)dt}.$$

The kernel K is chosen to be quadratic on $(-1, 1)$ and $K = H'$.

For the selection of bandwidth parameters h_K and h_H, we propose to choose the optimal bandwidth by using cross-validation procedure. We adopt the selection rule, proposed Ferrary and Vieu (2006).

The missing mechanism is similar to that in Ferraty et al. (2013),

$$P(x) = \mathbb{P}(\delta = 1 | X = x) = \exp it \left(2\alpha \int_0^1 x^2(t)dt \right),$$

where $\exp it(u) = e^u/(1 + e^u)$ for $u \in \mathbb{R}$, and the degree of dependency between the functional covariate X and the missing variable δ is controlled by the parameter α with missing proportion

$$\overline{\delta} = 1 - \frac{1}{n} \sum_{i=1}^{n} \delta_i.$$

We consider the case by taking

$$P(x) = \left| \sin\left(\pi * \int_0^1 x^2(t)dt\right) \right|.$$

We obtain the following results.

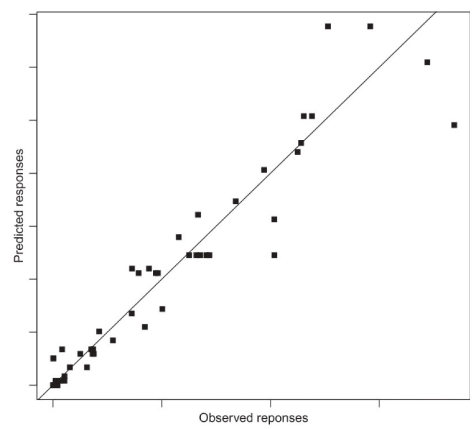

Figure 2. Conditional hazard function estimate

From Figure 2, we can say that even the behavior of the estimator is affected by the percentage of the missing observations. Overall, the estimator of the hazard function using the single index model have a good performance for lower missing observations.

5. Conclusion

In this paper, we create a local linear estimator of the conditional hazard function. Then, we establish the almost complete convergence with rate for the proposed estimator in the case of functional single index model and in presence of missing data. We give some hypothesis and some techniques to show our main results. In our work, the results prove that our estimator has good asymptotic properties. The case when response variable Y is also functional will be refer to future research.

REFERENCES

Abeidallah, M., Mechab, B. and Merouan, T. (2020). Local linear estimate of the point at high risk: Spatial functional data case, Comm. Statist. Theory Methods, Vol. 49, No. 11, pp. 2561–2584.

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

Aldo, G. and Philipe, V. (2016). An introduction to recent advances in high/infinite dimensional statistics, Journal of Multivariate Analysis, Vol. 146, pp. 1–6.

Aneiros, G., Cao, R., Fraiman, R., Genest, C. and Philippe, V. (2019). Recent advances in functional data analysis and high-dimensional statistics, Journal of Multivariate Analysis, Vol. 170, pp. 3–9.

Antonio, C. and Beatriz, PL. (2018). Polynomial volume estimation and its applications, Journal of Statistical Planning and Inference, Vol. 196, pp. 174–184.

Attaoui, S. (2014). Strong uniform consistency rates and asymptotic normality of conditional density estimator in the single functional index modeling for time series data, Advances in Statistical Analysis, Vol. 98, No. 3, pp. 257–286.

Attaoui, S., Laksaci, A. and Ould Said, E. (2011). A note on the conditional density estimate in the single functional index model, Stat. Probab. Lett, Vol. 81, No. 1, pp. 45–53.

Barrientos-Marin, J., Ferraty, F. and Vieu, P. (2010). Locally modelled regression and functional data, Journal of Nonparametric Statistics, Vol. 22, No. 5, pp. 617–632.

Belabbaci, O., Rabhi, A. and Soltani, S. (2015). 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), Vol. 10, No. 1, pp. 114– 138.

Benchiha, A. and Kaid, Z. (2018). Local linear estimate for functional regression with missing data at random, International Journal of Mathematics and Statistics, Vol. 19, No. 3, pp. 22–33.

Bouchentouf, A., Djebbouri, T., Rabhi, A. and Sabri, K. (2014). Strong uniform consistency rates of some characteristics of the conditional distribution estimator in the functional single-index model, Appl. Math., Vol. 41, No. 4, pp. 301–322.

Bouiadjra, H. B. (2017). Conditional hazard function estimate for functional data with missing at random, International Journal of Statistics & Economics, Vol. 18, No. 4, pp. 45–58.

Cheng, P.E. (1994). Nonparametric estimation of mean functionals with data missing at random, Journal of the American Statistical Association, Vol. 89, No. 425, pp. 81–87.

Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2010). Local linear estimation of the conditional density for functional data, C. R., Math., Acad. Sci. Paris, Vol. 348, pp. 931–934.

Demongeot, J., Laksaci, A., Rachdi, M. and Rahmani, S. (2014). On the local linear modelization of the conditional distribution for functional data, Sankhya: The Indian Journal of Statistics, Vol. 76, pp. 328–355.

Efromovich, S. (2011). Nonparametric regression with responses missing at random, Journal of Statistical Planning and Inference, Vol. 141, No. 12, pp. 3744–3752.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and its Applications: Monographs on Statistics and Applied Probability, 66*, CRC Press.

Ferraty, F., Peuch, A. and Vieu, P. (2003) Modèle à indice fonctionnel simple, C. R. Acad. Sci. Paris, Vol. 336, pp. 1025–1028.

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis*, Springer Series in Statistics.

Ferraty, F., Rabhi, A. and Vieu, P. (2008). Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle, Revue de Mathématiques Pures et Appliquées, Vol. 53, No. 1, pp. 1–18.

Ferraty, F., Park, J. and Vieu, P. (2011). Estimation of a Functional Single Index Model. In *Recent Advances in Functional Data Analysis and Related Topics* (F. Ferraty, Editor), Contribution to Statistics, Physica.

Ferraty, F., Sued, M. and Vieu, P. (2013). Mean estimation with data missing at random for functional covariables, Statistics, Vol. 47, No. 4, pp. 688–706.

Hamidi, N. and Mechab, B. (2018). Estimation of the conditional quantile for functional stationary ergodic data with responses missing at random, Journal of Probability and Statistical Science, Vol. 16, No. 2, pp. 131–149.

Kenouza, J., Mechab, B. and Benaissa, S. (2019). Functional local linear estimate of the conditional hazard function with missing at random, International Journal of Applied Mathematics and Statistics, Vol. 58, No. 4, pp. 115–121.

Laksaci, A. and Mechab, B. (2010). Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle: Cas des données spatiales, Revue de Mathématiques Pures et Appliquées, Vol. 55, No. 1, pp. 35–51.

Ling, N., Liang, L. and Vieu, P. (2015). Nonparametric regression estimation for functional stationary ergodic data with missing at random, Journal of Statistical Planning and Inference, Vol. 162, pp. 75–87.

Ling, N., Liu, Y. and Vieu, P. (2016). Conditional mode estimation for functional stationary ergodic data with responses missing at random, Statistics, Vol. 50, No. 5, pp. 991–1013.

Little, R.J.A. and Rubin, D.B. (2002). *Statistical Analysis with Missing Data*, Wiley series in probability and statistics.

Massim, I. and Mechab, B. (2016). Local linear estimation of the conditional hazard function, International Journal of Statistics & Economics, Vol. 17, No. 3, pp. 1–11.

Marcelo, M.T. and Pedro, A.M. (2017). Analysis of single-index models with scale mixture of normals errors by using Bayesian P-splines, Chilean Journal of Statistics, Vol. 8, No. 2, pp. 3–23.

Mechab, B., Hamidi, N. and Benaissa, S. (2019). Nonparametric estimation of the relative error in functional regression and censored data, Chilean Journal of Statistics, Vol. 10, No. 2, pp. 177–195.

Merouan, T., Mechab, B. and Chouaf, A. (2019). Nonparametric local linear estimator of the conditional hazard function in the functional single-index model, International Journal of Mathematics and Computation, Vol. 30, No. 3, pp. 43–54.

Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, Springer series in statistics, Springer-Verlag.

Roussas, G. (1989). Hazard rate estimation under dependence conditions, Journal of Statistical Planning and Inference, Vol. 22, No. 1, pp. 81–93.

Sabri, K., Meghnafi, M., Rabhi, A. and Benafla, D. (2019). A note on asymptotic normality of a Copula function in regression model, Applications and Applied Mathematics, An International Journal (AAM), Vol. 14, No. 1, pp. 76–98.

Tabti, H. and Ait Saidi, A. (2018). Estimation and simulation of conditional hazard function in the quasi-associated framework when the observations are linked via a functional single-index structure, Commu. Stat. Theory and Methods, Vol. 47, No. 4, pp. 816–838.

Watson, G.S. and Leadbetter, M.R. (1964). Hazard analysis, Sankhyia, Vol. 26, No. 1, pp. 101–116.