Applications and Applied Mathematics: An International Journal (AAM)

Volume 16 | Issue 2 | Article 4

12-2021

(R1463) On the Central Limit Theorem for Conditional Density Estimator In the Single Functional Index Model

Abbes Rabhi
University Djillali LIABES of Sidi Bel Abbés

Nadia Kadiri
University Djillali LIABES of Sidi Bel Abbés

Fatima Akkal
University Djillali LIABES of Sidi Bel Abbés

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

Applications and Applied Mathematics:
An International Journal
(AAM)

Vol. 16, Issue 2 (December 2021), pp. 844 – 866

On the Central Limit Theorem for Conditional Density Estimator In the Single Functional Index Model

[1*]Abbes Rabhi, [1,2]Nadia Kadiri, and [1,3]Fatima Akkal

[1]University Djillali LIABES of Sidi Bel Abbés
P.O. Box 89
Sidi Bel Abbés 22000, Algeria
[1*]rabhi_abbes@yahoo.fr; [2]nad.kad06@yahoo.com; [3]fatima.akkal@hotmail.com

*Corresponding Author

Received: December 13, 2020; Accepted: October 26, 2021

Abstract

The main objective of this paper is to investigate the nonparametric estimation of the conditional density of a scalar response variable Y, given the explanatory variable X taking value in a Hilbert space when the sample of observations is considered as an independent random variables with identical distribution (i.i.d.) and are linked with a single functional index structure. First of all, a kernel type estimator for the conditional density function (*cond-df*) is introduced. Afterwards, the asymptotic properties are stated for a conditional density estimator when the observations are linked with a single-index structure from which we derive an central limit theorem (CLT) of the conditional density estimator to show the asymptotic normality of the kernel estimate of this model. As an application the conditional mode in functional single-index model is presented. As an application the conditional mode in functional single-index model is presented as well as the asymptotic $(1-\xi)$ confidence interval of the conditional mode function is given for $0 < \xi < 1$. Simulation study is also presented to illustrate the validity and finite sample performance of the considered estimator.

Keywords: Asymptotic normality; Conditional density; Functional single-index process functional random variable; Nonparametric estimation; Small ball probability

MSC 2010 No.: 62G07, 62G99, 62G20

1. Introduction

Statistical analysis of functional variables has considerably grown over the last two decades. Indeed, an immense innovation on measuring devices has emerged and permitting to monitor several objects in a continuous way, such as stock market indexes, pollution, climatology, satellite images. Thus, a new branch of statistics, called functional statistics, has developed to treat observations as functional random elements. The study of statistical models for functional data has been a subject of several recent works and developments. The first results on the conditional models were obtained by Ferraty et al. (2006), where these authors established the almost complete convergence rate of the kernel estimators for the conditional distribution function, the conditional density and its derivatives, the conditional mode and the conditional quantiles. As a conditional nonparametric model, regression was one of the first predictive analysis tools. Conditional mode estimation is useful in prediction setting, it provides an alternative approach to classical regression estimation. For more recent advances in the topic, see Ezzahrioui and Ould-Saïd (2010)). In functional statistics, this model was introduced by Cardot et al. (2004). The nonparametric study of this model has been considered by Ferraty and Vieu (2006).

The ergodic theory has appeared in statistical mechanics, notably in Maxwell's and Gibbs's theories. It is necessary to make a sort of logical transition between the average behavior of the set of dynamic systems and the temporal average of the behaviors of a single dynamic system. It is derived from an ingenious hypothesis used for a long time without justifying it, and in various forms. In the context of the ergodic functional case with censored observations the literature is very restricted. We refer to Chaouch and Khardani (2015), who studied the asymptotic properties of the kernel-type estimator of the conditional quantiles when the response variable is right-censored and the data are sampled from an underlying stationary ergodic process. The single-index model represents one of the well-known semi-parametric models, which is very popular in the economics community as which allow to reduce the dimensionality of the covariate space while offering a flexibility in describing the relationship between the response and the covariate, through an unknown link function. The statistical study of these models, in the context of vectorial explanatory random variables, was initiated by Härdle and Marron (1985). Hristache et al. (2001) provide both new theoretical and bibliographic elements. Several authors have worked on simple functional index models, we can cite Ferraty et al. (2003), Aït-Saidi et al. (2008), Attaoui and Boudiaf (2014) and Bouchentouf et al. (2014).

The statistical study of single index models have been investigated and developed by several authors from a practical and theoretical point of view. The case of a vector explanatory variable was studied by Härdle et al. (1993) and Hristache et al. (2001). The single index models are very popular in the econometric community because it respond two important preoccupations. The first concerns dimension reduction since this type of model makes it possible to provide a solution to the problem of the curse of dimensionality, in the sense that pure nonparametric models are highly affected by dimensionality effects while semi-parametric ideas are more appealing candidates. The second is related to the interpretability of the index θ introduced in these models, for more details on refer to Cuevas (2014), Goia and Vieu (2016) and Aneiros et al. (2019) for an overview on methodological issues on FDA. Therefore, the single functional index model accumulate the advantages of single index model, and inherits the potential of the functional linear model in terms of applications. The interested reader, for the semi-parametric and the nonparametric functional models, may refer Geenens (2011), Ling and Vieu (2018, 2021), Novo et al. (2019a) and Chowdhury and Chaudhuri (2019) for survey on the topics.

The modelization of functional data, has been developed intensively. The motivation of such statistical analysis is justified by the recent technological development of the measuring instruments that offers the opportunity to observe phenomena in an increasingly accurate way, but this accuracy obviously generates a large amount of data observed over a finer grid, which can be considered as observations varying over a continuum. The most theoretical results are obtained under independence condition. However, in practice, it is rarely that we have an independent identically distributed observations of functional nature. The functional time series presents the more realistic situation. Thus, it is really crucial to study the functional statistical models when the usual independence condition on the statistical sample is relaxed. In this paper, we consider the problem of the nonparametric estimation of the regression function in single functional index model when the data are weakly dependant.

However, in the literature of functional statistics, the single functional index model is strictly limited in the case where the data is functional (a curve). The first result in this context was given by Ferraty et al. (2003). They obtained the almost complete convergence of the regression function $r()$ in the independent and identically distributed (i.i.d.) case. The generalization of this result to the dependent case has been studied by Masry (2005). Shang (2020) uses a Bayesian method to estimate the bandwidths in the kernel form error density and regression function, under an autoregressive error structure, and according to empirical studies, the author considered that the single functional index model gives improved estimation and prediction accuracies compared to any nonparametric functional regression considered. Novo et al. (2019b) have proposed a new automatic and location-adaptive procedure for estimating regression in a Functional Single-Index Model (FSIM) based on k-Nearest Neighbors ideas. Motivated by the analysis of imaging data, Li et al. (2017) proposed a novel functional varying-coefficient single-index model to carry out the regression analysis of functional response data on a set of covariates of interest. This method represents a new extension of varying-coefficient single-index models for scalar responses collected from cross-sectional and longitudinal studies. By simulation and real data analysis, the authors demonstrated the advantages of the proposed estimate. Wang et al. (2016) have considered the problem of predicting the real-valued response variable using explanatory variables containing both multivariate random variable and random curve. The authors considered the functional partial linear single-index model in order to treat the multivariate random variable as linear part and the random curve as functional single-index part, respectively.

These models gave attracted the attention of many researchers as Aït-Saidi et al. (2005, 2008). Bouchentouf et al. (2014) established a nonparametric estimation of some characteristics of the the conditional cumulative distribution function and the successive derivatives of the conditional density of a scalar response variable Y given a Hilbertian random variable X when the observations are linked with a single-index structure. Attaoui et al. (2011) studied the functional single-index model via its conditional density Kernel estimator, and established its pointwise and uniform almost complete convergence rates, their results were extended to dependent case by Attaoui (2014). Furthermore, Ling and Xu (2012) obtained the asymptotic normality of the conditional density estimator and the conditional mode estimator for the α-mixing dependence functional time series data.

The main contribution of this work is to generalize the result of Akkal et al. (2021) in case where a functional parameter θ is present in the model. In this work, we establish the asymptotic properties of the asymptotic normality for the estimators of conditional density function and conditional mode of a randomly scalar response given a functional covariate when the data are sampled from a stationary and ergodic process with single-index structure.

The paper is organized as follows. We present our model and some basic assumptions in Section 2. In Section 3 we state the main results as well as their proofs. As then application, we study the asymptotic normality of the conditional mode in functional single-index model in Section 4. Finally, Section 5 illustrates those asymptotic properties through some simulations.

2. Model and some basic assumptions

All along the paper, we will denote by $\mathcal{C}$, $\mathcal{C}'$ or/and $C_{\theta,x}$ some generic constant in $\mathbb{R}_+^*$. We consider that, given the $(X_i, Y_i)_{i=1,\dots,n}$ be a strictly stationary and ergodic sequence, with the same distribution as (X, Y), where Y is a real-valued random variable and X be a functional random variable (*frv*), which takes its values in a separable real Hilbert space $\mathcal{H}$ with the norm $\|\cdot\|$ generated by an inner product $< \cdot, \cdot >$.

Moreover, we consider $d_\theta(\cdot,\cdot)$ a semi-metric associated with the single index $\theta \in \mathcal{H}$ defined by $d_\theta(x_1, x_2) := | < x_1 - x_2, \theta > |$, for x_1 and x_2 in $\mathcal{H}$.

For a fixed $\mathcal{H}$, let $F(\theta, y, x)$ be the conditional cumulative distribution function (*cond-cdf*) of Y given $< \theta, X > = < \theta, x >$, specifically:

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

Saying that, we are implicitly assuming the existence of a regular version of the conditional distribution and that it's absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$, our aim is to build nonparametric estimates of several functions related with the conditional density of Y given $< X, \theta > = < x, \theta >$. Let

$$\forall y \in \mathbb{R},\ f(y|x) =: f(y| < x, \theta >),$$

be the conditional density of Y given $< X, \theta > = < x, \theta >$, for $x \in \mathcal{H}$.

In the following, we denote by $f(\theta, \cdot, x)$ the conditional density of Y given $< x, \theta >$ and we define the kernel estimator $\widehat{f}(\theta,\cdot,x)$ of $f(\theta,\cdot,x)$ by

$$\widehat{f}(\theta,y,x)=\frac{h_H^{-1}\sum_{i=1}^{n}K(h_K^{-1}(<x-X_i,\theta>))H(h_H^{-1}(y-Y_i))}{\sum_{i=1}^{n}K(h_K^{-1}(<x-X_i,\theta>))},$$

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

Let, for any $x\in\mathcal{H},\ i=1,\ldots,n$ and $y\in\mathbb{R}$,

$$K_i(\theta,x):=K(h_K^{-1}|<x-X_i,\theta>|),\ \ H_i(y):=H(h_H^{-1}(y-Y_i)).$$

Let $\mathcal{F}_i$ and $\mathcal{G}_i$ denote σ-fields generated by $((<X_1,\theta>,Y_1),\ldots,(<X_i,\theta>,Y_i))$ and $((<X_1,\theta>,Y_1),\ldots,(<X_i,\theta>,Y_i),<X_{i+1},\theta>)$, respectively. We denote by $B_\theta(x,h)=\{\chi\in\mathcal{H}/0<|<x-\chi,\theta>|<h\}$ be a ball of center x and radius h, and let $d_\theta(x,X_i)=|<x-X_i,\theta>|$ denote a random variable such that its cumulative distribution function is given by $\phi_{\theta,x}(u)=\mathbb{P}(d_\theta(x,X_i)\le u)=\mathbb{P}(X_i\in B_\theta(x,u))$. We consider that, given the σ-field $\mathcal{F}_{i-1}$, the conditional cumulative distribution function of $d_\theta(x,X_i)$ is defined by $\phi_{\theta,x}^{\mathcal{F}_{i-1}}(u)=\mathbb{P}(d_\theta(x,X_i)\le u\mid\mathcal{F}_{i-1})=\mathbb{P}(X_i\in B_\theta(x,u)\mid\mathcal{F}_{i-1})$.

Let $\mathcal{N}_x$ be a fixed neighborhood of x in $\mathcal{H}$, S_R will be a fixed compact subset of $\mathbb{R}$. Now, consider the following basic assumptions that are necessary in deriving the main result of this paper.

(H1) For $x\in\mathcal{H}$, there exists a sequence of nonnegative random functionals $(f_{i,1})_{i\ge1}$ bounded by a sequence of deterministic quantities $(b_i(\theta,x))_{i\ge1}$, a sequence of random functions $(g_{i,\theta,x})_{i\ge1}$, a deterministic nonnegative bounded functional f_1 and a nonnegative real function ϕ tending to zero, as its argument tends to 0, such that

- **(i)** $F_{\theta,x}(h)=\phi(h)f_1(\theta,x)+o(\phi(h))$ as $h\longrightarrow0$.
- **(ii)** For any $i\in\mathbb{N}$, $F_{\theta,x}^{\mathcal{F}_{i-1}}(h)=\phi(h)f_{i,1}(\theta,x)+g_{i,\theta,x}(h)$ with $g_{i,\theta,x}(h)=o_{a.s.}(\phi(h))$ as $h\longrightarrow0$, $g_{i,\theta,x}(h)/\phi(h)$ almost surely bounded and $n^{-1}\sum_{i=1}^{n}g_{i,\theta,x}^{j}(h)=o_{a.s.}(\phi^j(h))$ as $n\longrightarrow\infty$, for $j=1,2$.
- **(iii)** $n^{-1}\sum_{i=1}^{n}f_{i,1}^{j}(\theta,x)\longrightarrow f_1^j(\theta,x)$, almost surely when $n\longrightarrow\infty$, for $j=1,2$.
- **(iv)** There exists a nondecreasing bounded function τ_0 such that, uniformly in $s\in[0,1]$, $\phi(hs)/\phi(h)=\tau_0(s)+o(1)$, as $h\longrightarrow0$, and for $j=1,2$, $\int_0^1(K^j(t))'\tau_0(t)dt<\infty$.
- **(v)** $n^{-1}\sum_{i=1}^{n}b_i(\theta,x)\longrightarrow D(\theta,x)<\infty$ as $n\longrightarrow\infty$.

(H2) The conditional density $f(\theta,y,x)$ satisfies the Hölder condition, that is:
$\forall(y_1,y_2)\in S_R\times S_R$, $\forall(x_1,x_2)\in\mathcal{N}_x\times\mathcal{N}_x$
$|f(\theta,y_1,x_1)-f(\theta,y_2,x_2)|\le C_{\theta,x}(\|x_1-x_2\|^{b_1}+|y_1-y_2|^{b_2}),\ b_1>0,\ b_2>0.$

(H3) The kernel H is a positive bounded function such that $\forall(t_1,t_2)\in\mathbb{R}^2,\ |H(t_1)-H(t_2)|\le C|t_1-t_2|$, $\int H^2(t)dt<\infty$ and $\int|t|^{b_2}H(t)dt<\infty$.

(H4) The kernel K is a positive bounded function supported on $[0,1]$ and is differentiable on $[0,1]$ with derivative such that: $\exists C_1,\ C_2,\ -\infty < C_1 < K'(t) < C_2 < 0$, for $0 < t < 1$.

(H5) For any $m \geq 1$ and $j = 0,\ 1$,

$$\mathbb{E}\left[\left(h_H^{-1} H^{(j)}\left(h_H^{-1}(y - Y_i)\right)\right)^m \mid \mathcal{G}_{i-1}\right] = \mathbb{E}\left[\left(h_H^{-1} H^{(j)}\left(h_H^{-1}(y - Y_i)\right)\right)^m \mid < X_i, \theta >\right].$$

(H6) For any $x^{'} \in \mathcal{H}$ and $m \geq 2$,

$$\sup_{t \in S_{\mathbb{R}}} |g_m(\theta, x^{'}, t)| := \sup_{t \in S_{\mathbb{R}}} |\mathbb{E}[H^m(h_H^{-1}(t - T_1))| < X_1, \theta > = < x^{'}, \theta >]| < \infty,$$

and $g_m(\theta, x^{'}, t)$ is continuous in N_x uniformly on t:

$$\sup_{t \in S_{\mathbb{R}}} \sup_{x^{'} \in B_\theta(x,h)} |g_m(\theta, x^{'}, t) - g_m(\theta, x, t)| = o(1).$$

(H7) There exists a function $\beta_{\theta,x}(\cdot)$ such that $\lim_{n \to +\infty} \frac{\phi_{\theta,x}(sh_K)}{\phi_{\theta,x}(h_K)} = \beta_{\theta,x}(s)$, for $\forall s \in [0,1]$.

(H8) The bandwidth h_K and h_H, small ball probability $\phi_{\theta,x}(h_K)$ satisfying

- **(i)** $\lim_{n \to +\infty} h_K = 0$, $\lim_{n \to +\infty} h_H = 0$ and $\lim_{n \to +\infty} \frac{\log n}{n h_H \phi_{\theta,x}(h_K)} = 0$.
- **(ii)** $h_H^{b_2} \sqrt{n h_H \phi_{\theta,x}(h_K)} \to 0, \quad$ as $n \to \infty$.
- **(iii)** $n h_H^3 \phi_{\theta,x}^3(h_K) \longrightarrow 0, \quad$ as $n \longrightarrow \infty$.

3. Main result

In this section the asymptotic normality of the estimator $\widehat{f}(\theta, \cdot, x)$ in the single functional index model is established.

Theorem 3.1.

Under Assumptions we have (H1)-(H8)-(ii) for all x $\in \mathcal{H}$,

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

where $\sigma^2(\theta, y, x) = \frac{\alpha_2(\theta,x) f(\theta,y,x)}{(\alpha_1(\theta,x))^2 f_1(\theta,x)} \int H^2(t)dt$ with $\alpha_l(\theta, x) = K^l(1) - \int_0^1 (K^l)'(u) \beta_{\theta,x}(u) du,\ l = 1, 2$.

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

Proof:

In order to establish the asymptotic normality of $\widehat{f}(\theta, y, x)$, we need further notations and definitions. First we consider the following decomposition,

$$\widehat{f}_n(\theta, y, x) - f(\theta, y, x) = \frac{Q_n(\theta, y, x) + R_n(\theta, y, x)}{\widehat{f}_D(\theta, x)} + B_n(\theta, y, x),$$

where

$$Q_n(\theta,y,x) = \left(\widehat{f}_N(\theta,y,x) - \bar{f}_N(\theta,y,x)\right) - f(\theta,y,x)\left(\widehat{f}_D(\theta,x) - \bar{f}_D(\theta,x)\right), \quad (1)$$

and

$$R_n(\theta,y,x) = -B_n(\theta,y,x)\left(\widehat{f}_D(\theta,x) - \bar{f}_D(\theta,x)\right), B_n(\theta,y,x) = \frac{\bar{f}_N(\theta,y,x)}{\bar{f}_D(\theta,x)} - f(\theta,y,x),$$

with

$$\bar{f}_N(\theta,y,x) = \frac{1}{n\, h_H\, \mathbb{E}(K_1(\theta,x))}\sum_{i=1}^{n}\mathbb{E}\left[K_i(\theta,x)H_i(y)|,\mathcal{F}_{i-1}\right],$$

$$\bar{f}_D(\theta,x) = \frac{1}{n\, \mathbb{E}(K_1(\theta,x))}\sum_{i=1}^{n}\mathbb{E}\left[K_i(\theta,x)|\mathcal{F}_{i-1}\right].$$

$$\widehat{f}_N(\theta,y,x) = \frac{\sum_{i=1}^n K_i(\theta,x)H_i(y)}{nh_H\mathbb{E}(K_1(\theta,x))},\ \widehat{f}_D(\theta,x) = \frac{\sum_{i=1}^n K_i(\theta,x)}{n\mathbb{E}(K_1(\theta,x))}.$$

Then, the proof of Theorem 3.1 can be deduced from the following Lemmas. ■

Lemma 3.1.

Assume that assumptions (H2)(i)(ii)(iv) and (H4) hold true. For any real numbers $1 \leq j \leq 2+\delta$ with $\delta > 0$, as $n \to \infty$, we have

(i) $\frac{1}{\phi_\theta(h_K)}\mathbb{E}[K_i^j(\theta,x)|\mathcal{F}_{i-1}] = \alpha_j f_{i,1}(\theta,x) + O_{a.s}\left(\frac{g_{i,\theta,x}(h_K)}{\phi_\theta(h_K)}\right).$

(ii) $\frac{1}{\phi_\theta(h_K)}\mathbb{E}[K_1^j(\theta,x)] = \alpha_j f_1(\theta,x) + o(1).$

(iii) $\frac{1}{\phi_\theta^k(h_K)}(\mathbb{E}(K_1^j(\theta,x)))^k = \alpha_1^k f_1^k(\theta,x) + o(1).$

Proof:

See the proof of Lemma 1 in Laib and Louani (2010). ■

Lemma 3.2.

Under assumptions (H1) and (H4), we have

$$\lim_{n\to\infty}\widehat{f}_D(\theta,x) = \lim_{n\to\infty}\bar{f}_D(\theta,x) = 1 \quad \text{a.s.} \quad (2)$$

Proof:

The proof of this Lemma is the same of Lemma 3 and Lemma 5 in Laib and Louani (2011). ■

Lemma 3.3.

Under assumptions (H1)-(H5), we have

$$\lim_{n\to\infty}B_n(\theta,y,x) = \lim_{n\to\infty}R_n(\theta,y,x) = 0 \quad \text{a.s.} \quad (3)$$

Proof:

Observe that $R_n(\theta, t, x)$ is negligible and $\widehat{f}_D(\theta, x)$ converges almost surely towards 1, as $n \to \infty$.

Concerning $B_n(\theta, y, x)$, it is a direct consequence of Equation (13) given in Lemma 6.2 from Chaouch and Khardani (2015).

Concerning $R_n(\theta, y, x)$, making use of Lemma 6.1(ii) in Chaouch and Khardani (2015), we find that $\widehat{f}_D(\theta, x)$ converges almost surely to 1, as $n \to \infty$. In addition, using the proof of Lemma 6.2 in Chaouch and Khardani (2015), we obtain easily that $R_n(\theta, t, x)$ converges almost surely to zero when n tends to infinity. Therefore, the asymptotic normality of $\frac{Q_n(\theta,y,x)+R_n(\theta,y,x)}{\widehat{f}_D(\theta,x)}$ will be provided by the term $Q_n(\theta, t, x)$, which is treated in Lemma 3.4. ∎

Lemma 3.4.

Under conditions of Theorem 3.1, we have

$$\sqrt{nh_H\phi_\theta(h_K)}Q_n(\theta, y, x) \xrightarrow{D} \mathcal{N}(0, \sigma^2(\theta, y, x)).$$

Proof:

Let's denote

$$\zeta_{ni} = \left(\frac{\phi_\theta(h_K)}{nh_H}\right)^{1/2} (H_i(y) - h_H f(\theta, y, x)) \frac{K_i(\theta, x)}{\mathbb{E}(K_1(\theta, x))},$$

and define

$$\xi_{ni} = \zeta_{ni} - \mathbb{E}[\zeta_{ni}|\mathcal{F}_{i-1}].$$

It is easy to see that

$$(nh_H\phi_\theta(h_K))^{1/2}Q_n(\theta, y, x) = \sum_{i=1}^{n} \xi_{ni}.$$

Thus, the $\xi_{ni}, 1 \leq i \leq n$ forms a triangular array of stationary martingale differences with respect to the σ-field $\mathcal{F}_{i-1}$. By apply the central limit theorem for discrete-time arrays of real-valued martingales (Hall and Heyde (1980)), the asymptotic normality of $Q_n(\theta, y, x)$ can be obtained if we establish the following statements:

(a) $\displaystyle\sum_{i=1}^{n} \mathbb{E}[\xi_{ni}^2|\mathcal{F}_{i-1}] \xrightarrow{\mathbb{P}} \sigma^2(\theta, y, x).$

(b) $n\mathbb{E}\left[\xi_{ni}^2 \mathbf{1}_{[|\xi_{ni}|>\varepsilon]}\right] = o(1) \quad$ for $\quad \forall\varepsilon > 0.$

Proof of part (a) observe that

$$\left|\sum_{i=1}^{n} \mathbb{E}\left[\zeta_{ni}^2|\mathcal{F}_{i-1}\right] - \sum_{i=1}^{n} \mathbb{E}\left[\xi_{ni}^2|\mathcal{F}_{i-1}\right]\right| \leq \sum_{i=1}^{n} \left(\mathbb{E}[\zeta_{ni}|\mathcal{F}_{i-1}]\right)^2.$$

Then, similar to the proof of Equation (3) and using Lemma 3.1, we have

$$\begin{aligned}|\mathbb{E}[\zeta_{ni}|\mathcal{F}_{i-1}]| &= \frac{1}{\mathbb{E}(K_1(\theta,x))}\left(\frac{\phi_\theta(h_K)}{nh_H}\right)^{1/2}|\mathbb{E}\left[K_i(\theta,x)(H_i(y)-h_Hf(\theta,y,x))|\mathcal{F}_{i-1}\right]| \\ &= \frac{1}{\mathbb{E}(K_1(\theta,x))}\left(\frac{\phi_\theta(h_K)}{nh_H}\right)^{1/2}\Big|\mathbb{E}\Big[K_i(\theta,x) \\ &\qquad \mathbb{E}\left[(H_i(y)-h_Hf(\theta,y,x))| <\theta, X_i>\right]|\mathcal{F}_{i-1}\Big]\Big| \\ &\le C\left(h_K^{b_1}+h_H^{b_2}\right)\left(\frac{\phi_\theta(h_K)h_H}{n}\right)^{1/2}\left(\frac{f_{i,1}(\theta,x)}{f_1(\theta,x)}+\mathcal{O}_{a.s.}\left(\frac{g_{i,\theta,x}(h_K)}{\phi(h_K)}\right)\right).\end{aligned}$$

Thus, by (H1)(ii)-(iii), we get

$$\sum_{i=1}^n\left(\mathbb{E}[\zeta_{ni}|\mathcal{F}_{i-1}]\right)^2 = \mathcal{O}_{a.s.}\left(h_H\phi_\theta(h_K)(h_K^{b_1}+h_H^{b_2})^2\right).$$

Hence, the statement of (a) follows if we show that

$$\sum_{i=1}^n\mathbb{E}\left[\zeta_{ni}^2|\mathcal{F}_{i-1}\right]\xrightarrow{\mathbb{P}}\sigma^2(\theta,y,x),\quad \text{as}\quad n\to\infty.$$

By assumption (H5) we have

$$\begin{aligned}\sum_{i=1}^n\mathbb{E}\left(\zeta_{ni}^2|\mathcal{F}_{i-1}\right) &= \frac{\phi_\theta(h_K)}{nh_H(\mathbb{E}(K_1(\theta,x)))^2}\sum_{i=1}^n\mathbb{E}\left\{K_i^2(\theta,x)\left(H_i(y)-h_Hf(\theta,y,x)\right)^2|\mathcal{F}_{i-1}\right\} \\ &= \frac{\phi_\theta(h_K)}{nh_H(\mathbb{E}(K_1(\theta,x)))^2}\sum_{i=1}^n\mathbb{E}\left\{K_i^2(\theta,x)\mathbb{E}\left((H_i(y)-h_Hf(\theta,y,x))^2| <\theta,X_i>\right)|\mathcal{F}_{i-1}\right\}.\end{aligned}$$

Thus, by the properties of conditional expectation and (H5) for $j=0$ and $m=2$, we obtain that

$$\sum_{i=1}^n\mathbb{E}\left[\zeta_{ni}^2|\mathcal{F}_{i-1}\right] = V_{1,n}(\theta,y,x)+V_{2,n}(\theta,y,x),$$

where

$$\begin{aligned}V_{1,n}(\theta,y,x) = \frac{\phi_\theta(h_K)}{n\,h_H\,(\mathbb{E}(K_1(\theta,x)))^2}\sum_{i=1}^n\mathbb{E}\Big[K_i^2(\theta,x)\Big(&\mathbb{E}(H_i^2(y)| <\theta,X_i>) \\ &-(\mathbb{E}(H_i(y)| <\theta,X_i>))^2\Big)|\mathcal{F}_{i-1}\Big],\end{aligned}$$

and

$$V_{2,n}(\theta,y,x) = \tfrac{\phi_\theta(h_K)}{n\,h_H\,(\mathbb{E}(K_1(\theta,x)))^2}\textstyle\sum_{i=1}^n\mathbb{E}\Big[K_i^2(\theta,x)\mathbb{E}[(H_i(y)-h_Hf(\theta,y,x))| <\theta,X_i>]^2|\mathcal{F}_{i-1}\Big].$$

It should be noted that the second term is negligible: $V_{2,n}(\theta,y,x)\longrightarrow 0,\quad$ as $n\to\infty$. Indeed, we used conditions (H2) and (H3), we get

$$\mathbb{E}(H_i(\theta,y,x)-h_Hf(\theta,y,x)| <\theta,X_i>)\le C_{\theta,x}h_H\textstyle\int_{\mathbb{R}}H(v)(h_K^{b_1}+|v|^{b_2}h_H^{b_2})dv,$$

and Lemma 3.1, in order to get our result.

For $V_{1,n}$, notice that by changing variables, and by assumptions (H2)-(H3), we have

$$\begin{aligned}\mathbb{E}(H_i^2(y)| < \theta, X_i >) &= \int_{\mathbb{R}} H^2\left(\frac{y-v}{h_H}\right) f(\theta, v, x)dv \\ &\leq h_H \int_{\mathbb{R}} H^2(u)\big[f(\theta, y-uh_H, x) - f(\theta, y, x)\big]du \\ &\quad + h_H f(\theta, y, x)\int_{\mathbb{R}} H^2(u)du \\ &\leq h_H^{1+b_2}\int_{\mathbb{R}} |u|^{b_2} H^2(u)du + h_H f(\theta, y, x)\int_{\mathbb{R}} H^2(u)du \\ &= h_H\bigg(o(1) + f(\theta, y, x)\bigg(\int_{\mathbb{R}} H^2(u)du\bigg)\bigg),\end{aligned}$$

which implies that

$$\frac{1}{h_H}\mathbb{E}(H_i^2(y)| < \theta, X_i >) \to f(\theta, y, x)\int_{\mathbb{R}} H^2(u)du, \quad \text{as} \quad n \to \infty. \tag{4}$$

Similarly, as $n \to \infty$, we have

$$\begin{aligned}\mathbb{E}(H_i(t)| < \theta, X_i >) &= \frac{1}{h_H}\int H\left(\frac{t-v}{h_H}\right) f(\theta, v, x)dv \\ &= \int H(u) f(\theta, t-uh_H, x)du \to f(\theta, t, x)\int H(u)du. \end{aligned} \tag{5}$$

Then, by Equations (4)-(5) and Lemma 3.1, we arrive at

$$\begin{aligned}V_{1,n}(\theta, y, x) &= \frac{\phi_\theta(h_K)}{n\left(\mathbb{E}(K_1(\theta, x))\right)^2} f(\theta, y, x)\int_{\mathbb{R}} H^2(u)du \sum_{i=1}^{n}\mathbb{E}[K_i^2(\theta, x)|\mathcal{F}_{i-1}] \\ &\to \frac{\alpha_2}{\alpha_1^2}\frac{f(\theta, y, x)}{f_1(\theta, x)}\int_{\mathbb{R}} H^2(u)du, \quad \text{as} \quad n \to \infty.\end{aligned}$$

Proof of part (b). The definition of ξ_{ni} implies that $n\mathbb{E}\left[\xi_{ni}^2\mathbf{1}_{[|\xi_{ni}>\varepsilon|]}\right] \leq 4n\mathbb{E}\left[\zeta_{ni}^2\mathbf{1}_{[|\zeta_{ni}>\varepsilon/2|]}\right]$, where $\mathbf{1}_A$ is an indicator function of a set A. Let $a > 1$ and $b > 1$ such that $1/a + 1/b = 1$. By Hölder and Markov inequalities, one can write, for all $\varepsilon > 0$,

$$\mathbb{E}\left[\zeta_{ni}^2\mathbf{1}_{[|\zeta_{ni}>\varepsilon/2|]}\right] \leq \frac{\mathbb{E}|\zeta_{ni}|^{2a}}{(\varepsilon/2)^{2a/b}}.$$

Taking C_0 a positive constant and $2a = 2+\delta$ (with δ as in (A6)), we obtain

$$\begin{aligned}4n\mathbb{E}\left[\zeta_{ni}^2\mathbf{1}_{[|\zeta_{ni}>\varepsilon/2|]}\right] &\leq C_0\Gamma(\theta, x)\mathbb{E}\left([|H_i(y) - h_H f(\theta, y, x)|K_i(\theta, x)]^{2+\delta}\right) \\ &\leq C_0\Gamma(\theta, x)\mathbb{E}\Big((K_i(\theta, x))^{2+\delta} \\ &\qquad \mathbb{E}\left[|H_i(y) - h_H f(\theta, y, x)|^{2+\delta}| < \theta, X_i >\right]\Big),\end{aligned}$$

where $\Gamma(\theta, x) = \left(\frac{\phi_\theta(h_K)}{nh_H}\right)^{(2+\delta)/2} \frac{n}{(\mathbb{E}(K_1(\theta, x)))^{2+\delta}}$.

By changing variables, we get

$$\begin{aligned}
\mathbb{E}\left[|H_i(y) - h_H f(\theta, y, x)|^{2+\delta}| < \theta, X_i >\right] &= \int_{\mathbb{R}} \left(H\left(\frac{y-v}{h_H}\right) - h_H f(\theta, y, x)\right)^{2+\delta} f(\theta, v, x)dv \\
&\leq C \int_{\mathbb{R}} H^{2+\delta}\left(\frac{y-v}{h_H}\right) f(\theta, v, x)dv \\
&\quad + h_H^{2+\delta} f^{2+\delta}(\theta, y, x) \\
&= C h_H \int_{\mathbb{R}} H^{2+\delta}(u) f(\theta, y - uh_H, x)du \\
&\quad + h_H^{2+\delta} f^{2+\delta}(\theta, y, x) \\
&= h_H \int_{\mathbb{R}} H^{2+\delta}(u) f(\theta, y - uh_H, x)du \\
&\quad + h_H^{2+\delta} f^{2+\delta}(\theta, y, x),
\end{aligned}$$

$$\begin{aligned}
4n\mathbb{E}\left[\zeta_{ni}^2 \mathbf{1}_{[|\zeta_{ni} > \varepsilon/2|]}\right] &\leq C_0 \left(\frac{\phi_\theta(h_K)}{n}\right)^{(2+\delta)/2} \frac{n}{h_H^{\delta/2}(\mathbb{E}(K_1(\theta, x)))^{2+\delta}} \\
&\quad \mathbb{E}\left(K_i^{2+\delta}(\theta, x)\left[\int_{\mathbb{R}} H^{2+\delta}(u) f(\theta, y - uh_H, x)du + h_H^{1+\delta} f^{2+\delta}(\theta, y, x)\right]\right) \\
&\leq C_0 \left(\frac{\phi_\theta(h_K)}{n}\right)^{(2+\delta)/2} \frac{n\mathbb{E}(K_i^{2+\delta}(\theta, x))}{h_H^{\delta/2}(\mathbb{E}(K_1(\theta, x)))^{2+\delta}}.
\end{aligned}$$

Thus, by Lemma 3.1, it follows that

$$\begin{aligned}
4n\mathbb{E}\left[\zeta_{ni}^2 \mathbf{1}_{[|\zeta_{ni} > \varepsilon/2|]}\right] &\leq C_0 (nh_H \phi(h_K))^{-\delta/2} \frac{M_{2+\delta} f_1(\theta, x) + o(1)}{M_1^{2+\delta} f_1^{2+\delta}(\theta, x) + o(1)} \\
&= \mathcal{O}((nh_H \phi(h_K))^{-\delta/2}).
\end{aligned}$$

■

4. Application: The conditional mode in functional single-index model

The main objective of this section is to establish the asymptotic normality of the conditional mode estimator of Y given $< X, \theta > = < x, \theta >$ denoted by $M_\theta(x)$. We estimate the conditional mode $\widehat{M_\theta}(x)$ with a random variable $M_\theta(x)$ such that

$$\widehat{M_\theta}(x) = \arg\sup_{y \in \mathcal{S}_{\mathbb{R}}} \widehat{f}(\theta, y, x), \tag{6}$$

where $M_\theta(x) = \arg\sup_{y \in \mathcal{S}_{\mathbb{R}}} f(\theta, y, x)$, $\mathcal{S}_{\mathbb{R}}$ is a fixed compact subset of $\mathbb{R}$.

Let's note that in the remainder of our paper we will consider any value $\widehat{M_\theta}$ satisfying (6).

In order to present the estimation of the conditional mode in the functional single-index model, we introduce the following additional smoothness condition.

(U1) $f(\theta, \cdot, x)$ is twice continuously differentiable around the point $M_\theta(x)$ with $f^{(1)}(\theta, M_\theta(x), x) = 0$ and $f^{(2)}(\theta, \cdot, x)$ is uniformly continuous on $S_\mathbb{R}$ such that $f^{(2)}(\theta, M_\theta(x), x) \neq 0$, where $f^{(j)}(\theta, \cdot, x)$ $(j = 1, 2)$ is the jth order derivative of the conditional density $f(\theta, y, x)$.

(U2) $\forall \varepsilon > 0$, $\exists \eta > 0$,$\forall y \in S_\mathbb{R}$

$$|M_\theta(x) - y| \geq \varepsilon \Rightarrow |f(\theta, M_\theta(x), x) - f(\theta, y, x)| \geq \eta.$$

(U3) The conditional density function $f(\theta, y, x)$ satisfies: $\exists \beta_0 > 0,\ \forall (y_1, y_2) \in S_\mathbb{R} \times S_\mathbb{R}$,

$$|f^{(j)}(\theta, y_1, x) - f^{(j)}(\theta, y_2, x)| \leq C\big(|y_1 - y_2|^{\beta_0}\big),\ \forall j = 1, 2.$$

(U4) H' and H'' are bounded respectively with

$$\int \big(H'(t)\big)^2 dt < \infty\, , \int |t|^{\beta_0} H(t) dt < \infty.$$

Theorem 4.1.

Suppose that hypotheses (H1)-(H8) and (U1)-(U4) are satisfied. If

$$\lim_{n\to\infty} n h_H^3 \phi_{\theta,x}(h_K) = \infty, \tag{7}$$

we have as $n \to \infty$,

$$\sqrt{\frac{n h_H^3 \phi_{\theta,x}(h_K)}{\nu^2(\theta, x)}}(\widehat{M_\theta}(x) - M_\theta(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \text{ as } n \to \infty, \tag{8}$$

where

$$\nu^2(\theta, x) = \frac{\alpha_2(\theta, x) f(\theta, M_\theta(x), x)}{\Big(\alpha_1(\theta, x) f^{(2)}(\theta, M_\theta(x), x)\Big)^2} \int (H'(t))^2 dt.$$

Proof:

First, by (6) and (U1), it follows that $f^{(1)}(\theta, M_\theta(x), x) = 0$.

Writing the first order Taylor expansion for $f^{(1)}(\theta, y, x)$ at point $M_\theta(x)$ leads to the existence of some $M_\theta^*(x)$ between $\widehat{M_\theta}(x)$ and $M_\theta(x)$ such that

$$\sqrt{n h_H^3 \phi_{\theta,x}(h_K)}(\widehat{M_\theta}(x) - M_\theta(x)) = \frac{-\sqrt{n h_H^3 \phi_{\theta,x}(h_K)}\widehat{f}^{(1)}(\theta, M_\theta(x), x)}{\widehat{f}^{(2)}(\theta, M_\theta^*(x), x)}.$$

In order to prove (8), we only need to show that

$$-\sqrt{n h_H^3 \phi_{\theta,x}(h_K)}\widehat{f}^{(1)}(\theta, M_\theta(x), x) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \nu_0^2(\theta, x)), \tag{9}$$

and

$$\widehat{f}^{(2)}(\theta, M_\theta^*(x), x) \longrightarrow \widehat{f}^{(2)}(\theta, M_\theta(x), x) \neq 0, \text{ in probability,} \tag{10}$$

where

$$\nu_0^2(\theta, x) = \frac{\alpha_2(\theta, x) f(\theta, M_\theta(x), x)}{\Big(\alpha_1(\theta, x)\Big)^2} \int (H^{'}(t))^2 dt.$$

In fact, because the continuity of the function $f(\theta, y, x)$ and by (U2) and the definitions of $\widehat{M_\theta}(x)$ and $M_\theta(x)$, we have, for all $\varepsilon > 0$, $\exists \eta(\varepsilon) > 0$ such that:

$$\begin{aligned}\mathbb{P}\left(|\widehat{M_\theta}(x) - M_\theta(x)| \geq \varepsilon\right) \leq \mathbb{P}\left(|f(\theta, M_\theta(x), x) - \widehat{f}(\theta, M_\theta(x), x)| \geq \frac{\eta(\varepsilon)}{2}\right) \\ +\mathbb{P}\left(|\widehat{f}(\theta, \widehat{M_\theta}(x), x) - f(\theta, \widehat{M_\theta}(x), x)| \geq \frac{\eta(\varepsilon)}{2}\right).\end{aligned} \tag{11}$$

Thus, similar to Ferraty and Vieu (2006), by (H1)-(H4) and (H8)-(i), we have $\widehat{f}(\theta, y, x) \longrightarrow f(\theta, y, x)$ in probability, which implies that $\widehat{M_\theta}(x) \longrightarrow M_\theta(x)$ in probability by (11) as $n \to \infty$. Similarly, the methodology can be also applied to obtain $\widehat{f}^{(2)}(\theta, y, x) \longrightarrow f^{(2)}(\theta, y, x)$ in probability as $n \to \infty$ by (H1), (H4), (H8), (U3) and (U4). Therefore, (10) is valid by the fact that $f^{(2)}(\theta, y, x)$ is uniformly continuous with respect to y on $\mathcal{S}_\mathbb{R}$. Next, we prove (9). In fact, since

$$\begin{aligned}\widehat{f}^{(1)}(\theta, M_\theta(x), x) = \frac{1}{\widehat{f}_D(\theta, x)}\left(\widehat{f}_N^{(1)}(\theta, M_\theta(x), x) - \mathbb{E}\widehat{f}_N^{(1)}(\theta, M_\theta(x), x)\right) \\ -\frac{1}{\widehat{f}_D(\theta, x)}\left(f^{(1)}(\theta, M_\theta(x), x) - \mathbb{E}\widehat{f}_N^{(1)}(\theta, M_\theta(x), x)\right).\end{aligned} \tag{12}$$

By (U1), (U3)-(U4), (7) and (12), similar to the proof of Lemmas, Lemma 3.3 and Lemma 3.4 respectively, (9) follows directly. Then, the proof of Theorem 4.1 is completed. ■

4.1. Application and Confidence bands

The asymptotic variances $\sigma^2(\theta, y, x)$ and $\nu^2(\theta, x)$ in Theorem 3.1 and Corollary 4.1 depend on some unknown quantities including α_1, α_2, $\phi(u)$, $M_\theta(x)$ and $f(\theta, M_\theta(x), x)$. So, $M_\theta(x)$, and $f(\theta, M_\theta(x), x)$ should be replaced by their respective estimators $\widehat{M_\theta}(x)$, and $\widehat{f}(\theta, M_\theta(x), x)$.

Because the unknown functions $\alpha_j := \alpha_j(\theta, x)$ and $f(\theta, y, x)$ intervening in the expression of the variance. So we need to estimate the quantities $\alpha_1(\theta, x)$, $\alpha_2(\theta, x)$ and $f(\theta, y, x)$, respectively.

By the assumptions (H1)-(H4) we know that $\alpha_j(\theta, x)$ can be estimated by $\widehat{\alpha}_j(\theta, x)$ which is defined as:

$$\widehat{\alpha}_j(\theta, x) = \frac{1}{n\widehat{\phi}_{\theta,x}(h)} \sum_{i=1}^n K_i^j(\theta, x), \text{ where } \widehat{\phi}_{\theta,x}(h) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{\{|<x-X_i,\theta>|<h\}},$$

with $\mathbf{1}_{\{\cdot\}}$ being the indicator function.

By applying the kernel estimator of $f(\theta, y, x)$ given above, the quantity $\sigma^2(\theta, y, x)$ can be estimated finally by:

$$\widehat{\sigma}^2(\theta, y, x) = \frac{\widehat{\alpha}_2(\theta, x)\widehat{f}(\theta, y, x)}{\widehat{\alpha}_1^2(\theta, x)} \int H^2(t)dt,$$

so we can derive the following corollary.

Corollary 4.1.

Under the assumptions of Theorem 3.1, we have as $n \to \infty$

$$\sqrt{\frac{nh_H\widehat{\phi}_{\theta,x}(h_K)}{\widehat{\sigma}^2(\theta, y, x)}}\left(\widehat{f}(\theta, y, x) - f(\theta, y, x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

Proof:

Observe that

$$\begin{aligned}
\Sigma &= \frac{\widehat{\alpha}_1(\theta, x)}{\sqrt{\widehat{\alpha}_2(\theta, x)}}\sqrt{\frac{nh_H\widehat{\phi}_{\theta,x}(h_K)}{\widehat{f}(\theta, y, x)}}\left(\widehat{f}(\theta, y, x) - f(\theta, y, x)\right) \\
&= \frac{\widehat{\alpha}_1(\theta, x)\sqrt{\alpha_2(\theta, x)}}{\alpha_1(\theta, x)\sqrt{\widehat{\alpha}_2(\theta, x)}}\sqrt{\frac{nh_H\widehat{\phi}_{\theta,x}(h_K)f(\theta, y, x)}{\widehat{f}(\theta, y, x)nh_H\phi_{\theta,x}(h_K)}} \\
&\quad \times \frac{\alpha_1(\theta, x)}{\sqrt{\alpha_2(\theta, x)}}\sqrt{\frac{nh_H\phi_{\theta,x}(h_K)}{f(\theta, y, x)}}\left(\widehat{f}(\theta, y, x) - f(\theta, y, x)\right).
\end{aligned}$$

Via Theorem 3.1, we have

$$\frac{\alpha_1(\theta, x)}{\sqrt{\alpha_2(\theta, x)}}\sqrt{\frac{nh_H\phi_{\theta,x}(h_K)}{f(\theta, y, x)}}\left(\widehat{f}(\theta, y, x) - f(\theta, y, x)\right) \longrightarrow \mathcal{N}(0,\ 1).$$

Next, by Laib and Louani (2010), we can prove that

$$\widehat{\alpha}_1(\theta, x) \xrightarrow{\mathbb{P}} \alpha_1(\theta, x),\ \ \widehat{\alpha}_2(\theta, x) \xrightarrow{\mathbb{P}} \alpha_2(\theta, x), \text{ and } \frac{\widehat{\phi}_{\theta,x}(h_K)}{\phi_{\theta,x}(h_K)} \xrightarrow{\mathbb{P}} 1, \quad \text{as} \quad n \to \infty.$$

Therefore, we obtain

$$\frac{\widehat{\alpha}_1(\theta, x)\sqrt{\alpha_2(\theta, x)}}{\alpha_1(\theta, x)\sqrt{\widehat{\alpha}_2(\theta, x)}}\sqrt{\frac{nh_H\widehat{\phi}_{\theta,x}(h_K)f(\theta, y, x)}{\widehat{f}(\theta, y, x)nh_H\phi_{\theta,x}(h_K)}} \longrightarrow 1, \text{ as } n \to \infty.$$

This yields the proof of Corollary 4.1. ■

Finally, in order to show the asymptotic $(1 - \xi)$ confidence interval of $M_\theta(x)$, we need to consider the estimator of $\nu^2(\theta, x)$ as follows:

$$\widehat{\nu}^2(\theta, x) = \frac{\widehat{\alpha}_2(\theta, x)\widehat{f}(\theta, \widehat{M}_\theta(x), x)}{\left(\widehat{\alpha}_1(\theta, x)\widehat{f}^{(2)}(\theta, \widehat{M}_\theta(x), x)\right)^2} \int (H^{'}(t))^2dt.$$

Thus, the following corollary is obtained.

Corollary 4.2.

Under conditions of Theorem 4.1, as $n \to \infty$ we have

$$\sqrt{\frac{nh_H^3\widehat{\phi}_{\theta,x}(h_K)}{\widehat{\nu}^2(\theta,x)}}(\widehat{M}_\theta(x) - M_\theta(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Proof:

Observe that

$$\begin{aligned}\Sigma' &= \frac{\widehat{\alpha}_1(\theta,x)\widehat{f}^{(2)}(\theta,\widehat{M}_\theta(x),x)}{\sqrt{\widehat{\alpha}_2(\theta,x)}}\sqrt{\frac{nh_H^3\widehat{\phi}_{\theta,x}(h_K)}{\widehat{f}(\theta,\widehat{M}_\theta(x),x)}}\left(\widehat{M}_\theta(x) - M_\theta(x)\right) \\ &= \frac{\widehat{\alpha}_1(\theta,x)\sqrt{\alpha_2(\theta,x)}}{\alpha_1(\theta,x)\sqrt{\widehat{\alpha}_2(\theta,x)}}\sqrt{\frac{nh_H^3\widehat{\phi}_{\theta,x}(h_K)f(\theta,M_\theta(x),x)}{\widehat{f}(\theta,\widehat{M}_\theta(x),x)nh_H^3\phi_{\theta,x}(h_K)}}\frac{\widehat{f}^{(2)}(\theta,\widehat{M}_\theta(x),x)}{f^{(2)}(\theta,M_\theta(x),x)} \\ &\quad\times\frac{\alpha_1(\theta,x)}{\sqrt{\alpha_2(\theta,x)}}\sqrt{\frac{nh_H^3\phi_{\theta,x}(h_K)}{f(\theta,M_\theta(x),x)}}f^{(2)}(\theta,M_\theta(x),x)\left(\widehat{M}_\theta(x) - M_\theta(x)\right).\end{aligned}$$

Making use of Theorem 4.1, we obtain

$$\frac{\alpha_1(\theta,x)}{\sqrt{\alpha_2(\theta,x)}}\sqrt{\frac{nh_H^3\phi_{\theta,x}(h_K)}{f(\theta,M_\theta(x),x)}}f^{(2)}(\theta,M_\theta(x),x)\left(\widehat{M}_\theta(x) - M_\theta(x)\right) \longrightarrow \mathcal{N}\left(0,\,1\right).$$

Further, by considering Lemma 3.3, (10) and (11), we obtain

$$\frac{\widehat{\alpha}_1(\theta,x)\sqrt{\alpha_2(\theta,x)}}{\alpha_1(\theta,x)\sqrt{\widehat{\alpha}_2(\theta,x)}}\sqrt{\frac{nh_H^3\widehat{\phi}_{\theta,x}(h_K)f(\theta,M_\theta(x),x)}{\widehat{f}(\theta,\widehat{M}_\theta(x),x)nh_H^3\phi_{\theta,x}(h_K)}}\frac{\widehat{f}^{(2)}(\theta,\widehat{M}_\theta(x),x)}{f^{(2)}(\theta,M_\theta(x),x)} \xrightarrow{\mathbb{P}} 1, \text{ as } n \to \infty.$$

Hence, the proof is completed. ■

Remark 4.1.

Thus, following the corollaries, Corollary 4.1 and Corollary 4.2, the asymptotic $(1-\xi)$ confidence interval of $f(\theta,y,x)$ and $M_\theta(x)$ are given by

$$\widehat{f}(\theta,y,x) \pm \tau_{\xi/2} \times \frac{\widehat{\sigma}(\theta,x)}{\sqrt{nh_H\widehat{\phi}_{\theta,x}(h_K)}} \text{ and } \widehat{M}_\theta(x) \pm \tau_{\xi/2} \times \frac{\widehat{\nu}(\theta,x)}{\sqrt{nh_H^3\widehat{\phi}_{\theta,x}(h_K)}},$$

where $\tau_{\xi/2}$ is the upper $\xi/2$ quantile of standard Normal $\mathcal{N}(0,1)$.

5. Simulation study

To study the behavior of our conditional mode estimator, we consider in this part two examples of simulation. In the first one, we compare our model FSIM (fonctional single index model) with that of NPFDA (nonparametric functional data analysis) and in the latter, knowing the distribution of the regression model (the distribution is known and usual), we look to the behavior of our estimator of the conditional density function with respect to this distribution. Therefore, the best way to know the behavior of the estimator of conditional density is to compute its mean square error. So, in this part of paper we compare between the conditional density estimation in the FSIM which is our model and the conditional density estimation in the NPFDA defined in (13),

$$f_n(x|y) = \frac{h_H^{-1}\sum_{i=1}^{n} K\left(h_K^{-1}d(< x, X_i)\right) H\left(h_H^{-1}(y - Y_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}d(x, X_i)\right)}, \tag{13}$$

where we estimate the conditional mode $\widehat{M}(x)$ with a random variable $M(x)$ such that

$$M(x) = \arg\sup_{y\in\mathcal{S}_{\mathbb{R}}} f(x|y) \text{ and } \widetilde{M}(x) = \arg\sup_{y\in\mathcal{S}_{\mathbb{R}}} f_n(x|y).$$

So, we have to compare their respective conditional density estimators by computing and comparing their respective mean square errors for some values of the scalar response Y.

In the following, our purpose consists in assessing the performance, in terms of prediction, of $\widetilde{M_\theta}(x)$ and $\widetilde{M}(x)$. For each given predictor $(X_j)_{j\in\mathcal{J}}$ in the testing subsample, we are interested in the prediction of the response variable $(Y_j)_{j\in\mathcal{J}}$ via the single functional index conditional mode $\widehat{M_\theta}(x)$ and the fully nonparametric conditional mode $\widetilde{M}(x)$ so as to compare the finite-sample behavior of the estimator. As assessment tool we consider the mean square error (MSE) defined as follows:

$$SSR = \frac{1}{|\mathcal{J}|}\sum_{j\in\mathcal{J}}\left(Y_j - \widehat{Y_j}\right)^2, \tag{14}$$

where $\widehat{Y_j}$ is a predictor of Y_j obtained either semi-parametrically by $\widehat{M_\theta}(x)$ or nonparametrically via $\widetilde{M}(x)$.

Furthermore, some tuning parameters have to be specified. The kernel $K(\cdot)$ is chosen to be the quadratic function defined as $K(u) = \frac{3}{2}(1-u^2)\mathbf{1}_{[0,1]}$ and the cumulative df $H(u) = \int_{-\infty}^{u}\frac{3}{4}(1-z^2)\mathbf{1}_{[-1,1]}(z)\,dz$.

The semi-metric $d(\cdot,\cdot)$ will be specified according to the choice of the functional space $\mathcal{H}$ discussed in the scenarios below. It is well-known that one of the crucial parameters in semi-parametric models is the smoothing parameters which are involved in defining the shape of the link function between the response and the covariate.

Using the result given in Theorem 4.1, the variance of our estimator is obtained as

$$CV = \frac{\alpha_2(\theta,x)f(\theta,M_\theta(x),x)}{nh_H^3\phi_{\theta,x}(h_K)\Big(\alpha_1(\theta,x)f^{(2)}(\theta,M_\theta(x),x)\Big)^2}\ .$$

The idea is to choose the parameters h_K and h_H so that the variance is minimal. Since the variance (CV) depends on several unknown parameters that must be estimated, the calculus becomes tedious. Thus, by replacing the unknown parameters by their respective estimators $\widehat{\alpha}_1(\theta,x)$, $\widehat{\alpha}_2(\theta,x)$, $\widehat{M_\theta}(x)$, $\widehat{f}$, and $\widehat{\phi}_{\theta,x}(h_K)$, we obtain

$$(h_K,h_H)=\arg\min_{h_K,h_H} CV\left(h_K,h_H\right)=\arg\min_{h_K,h_H}\frac{\widehat{\alpha}_2(\theta,x)\widehat{f}(\theta,\widehat{M_\theta}(x),x)}{nh_H^3\widehat{\phi}_{\theta,x}(h_K)\Big(\widehat{\alpha}_1(\theta,x)\widehat{f}^{(2)}(\theta,\widehat{M_\theta}(x),x)\Big)^2}\ .$$

Now, for simplifying the implementation of our methodology, we take the bandwidths $h_H\sim h_K=h$, where h will be chosen by the cross-validation method on the k-nearest neighbors (see Ferraty and Vieu (2006), p. 102).

5.1. Simulation 1: case of smooth curves

Let us consider the following regression model, where the covariate is a curve and the response is a scalar:

$$T_i=R\left(X_i\right)+\epsilon_i,\ i=1,\ldots,n,$$

where ϵ_i a sequence of i.i.d. random variables normally distributed with a variance equal to 0.1.

The functional covariate X is assumed to be a diffusion process defined on $[0,\ 1]$ and generated by the following equation:

$$X(t)=a\cos(b+\pi Wt)+c\sin(d+\pi Wt)+(1-A)\sin(\pi tW),\ t\in[0,\ 1],$$

where W, b and d are independent of normal distributions respectively $\rightsquigarrow\mathcal{N}(0,\ 1)$, $\rightsquigarrow\mathcal{N}(0,\ 0.03)$ and $\rightsquigarrow\mathcal{N}(0,\ 0.05)$. The variables a and c are Bernoulli's laws Bernoulli $\mathcal{B}(0.5)$. Figure 1 depicts a sample of 200 curves representing a realization of the functional random variable X.

Take into account of the smoothness of the curves $X_i(t)$ (see Figure 1), we choose the distance $deriv_1$ (the semi-metric based on the first derivatives of the curves) in $\mathcal{H}$ as:

$$d\left(\chi_1,\chi_2\right)=\left(\int_0^1\left(\chi_1'\left(t\right)-\chi_2'\left(t\right)\right)^2dt\right)^{1/2}$$

as semi-metric. Then, we consider a nonlinear regression function defined as

$$R\left(X\right)=4\log\left\{1/\left(\int_0^1\left(X'(t)\right)^2dt+\left[\int_0^1X'(t)dt\right]^2\right)\right\}.$$

Given $X=x$, $Y\rightsquigarrow\mathcal{N}\left(R\left(x\right),\,0.2\right)$, and thus, the conditional median, the conditional mode and the conditional mean functions will coincide and will be equal to $R\left(x\right)$, for any fixed x. The computation of our estimator is based on the observed data $\left(X_i,\ Y_i\right)_{i=1,\ldots,n}$, and the single index θ which is unknown and has to be estimated.

In practice this parameter can be selected by cross-validation approach (see Aït-Saidi et al. (2008)). In this passage it may be that one can select the real-valued function $\theta\left(t\right)$ among the eigenfunctions

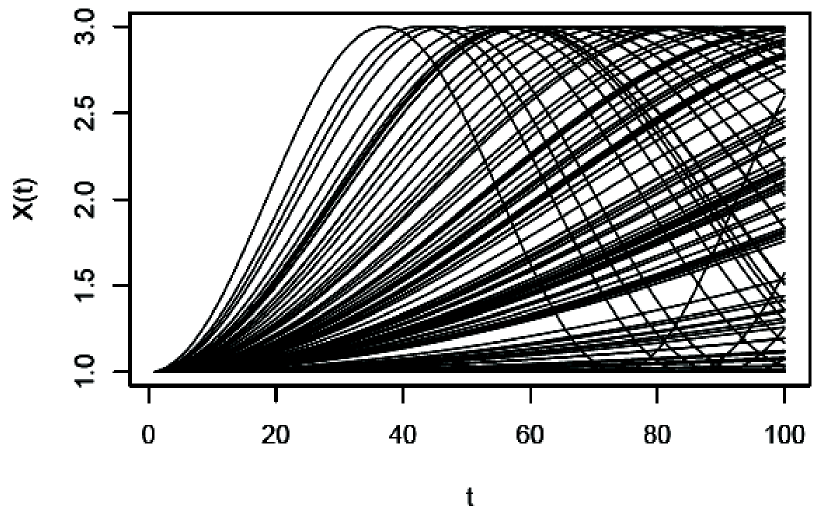

Figure 1. A sample of 200 curves $X_{i=1,\ldots,200}(t_j),\ t_{j=1,\ldots,200} \in [0,1]$

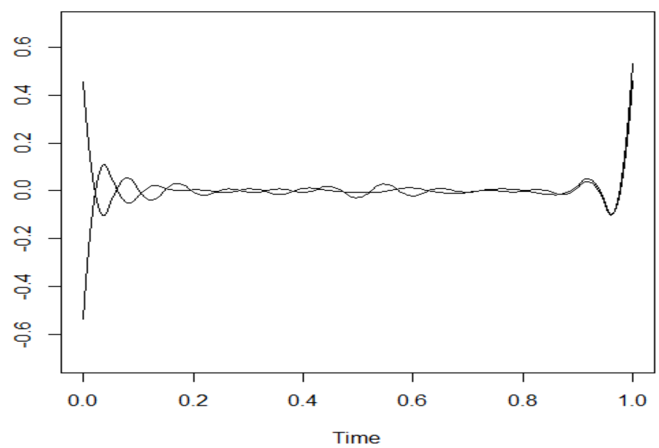

Figure 2. The curves $\theta_{i=1,2}(t_j),\ t_{j=1,\ldots,200} \in [0,\ 1]$

of the covariance operator $\mathbb{E}\left[(X' - \mathbb{E}X') < X', \cdot >_{\mathcal{H}}\right]$ where $X\,(t)$ is a diffusion processes defined on a real interval $[a,\ b]$ and $X'\,(t)$ its first derivative (see Attaoui and Ling (2016)). So for a chosen training sample $\mathcal{L}$, by applying the principal component analysis (PCA) method, the computation of the eigenvectors of the covariance operator estimated by its empirical covariance operator: $\frac{1}{|\mathcal{L}|}\sum_{i\in\mathcal{L}}(X'_i - \mathbb{E}X')\,{}^t(X'_i - \mathbb{E}X')$, will be the one best approximation of our functional parameter θ. Now, let us denote $\theta^\star$ the first eigenfunction corresponding to the first higher eigenvalue of the empirical covariance operator, which will replace θ during the simulation step.

In the following graphs, the covariance operator for $\mathcal{L} = \{1, \ldots, 200\}$ gives the discretization of the first eigenfunction θ (presented by a continuous curve) and all the eigenfunctions $\theta_i(t)$ (Figure 2 and Figure 3). In this simulation part, we divide our sample of size 200 into two parts. The first

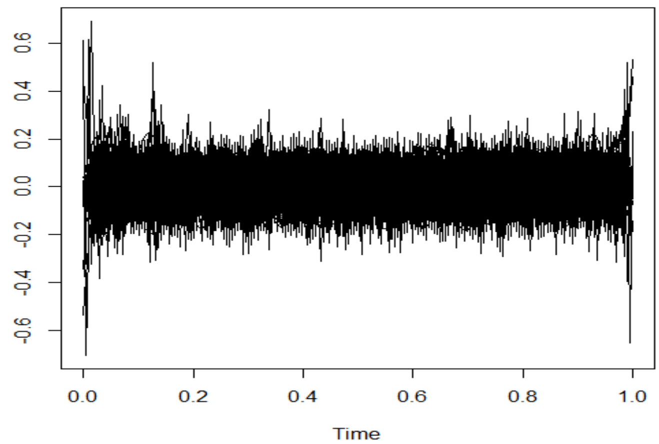

Figure 3. The curves $\theta_{i=1,\dots,200}(t_j),\ t_{j=1,\dots,200} \in [0,\ 1]$

one from 1 to 125 will be used to make the simulation and the second from 126 to 200 will serve us for the prediction.

We follow the following steps:

Step 1. Simulate the responses variables Y_i.
Step 2. For each j in the test sample $\mathcal{J} = \{126, \dots, 200\}$, we compute: $\widehat{Y}_j = \widehat{M_{\theta^\star}}(X_j)$ and $\widehat{Y}_j = \widehat{M}(X_j)$,

Finally, we present the results by plotting the predicted values versus the true values and compute the sum of squared residuals (SSR) defined by (14).

We see that the sum of squared residuals (SSR) of our method Functional-Single-Index-Model (FSIM) is less than the one of the Nonparametric-Functional-Data-Analysis (NPFDA). This is confirmed by the following graphs, when we compare the conditional mode by (FSIM) against the conditional mode by (NPFDA) (Fig. 4). Our estimator is so acceptable. As intuitively expected, it is well observed that the mean square errors of our estimator are smaller than that of NPFDA. Thus, again, the FSIM model produces much more accurate estimation accuracies than NPFDA model in all criteria.

In order to construct conditional confidence bands we proceed by the following algorithm:

Step 1. $< \theta^\star, X_1 >, \dots, < \theta^\star, X_{200} >$, generate independently the variables $\varepsilon_1, \dots, \varepsilon_{200}$, then simulate the response variables $Y_i = r(< \theta^\star, X_i >) + \varepsilon_i$, where $r(< \theta^\star, X_i >) = \exp(10(< \theta^\star, X_i > -0.05))$ and generate independently the variables $\varepsilon_1, \dots, \varepsilon_{200}$.
Step 2. For each i in the training sample, we calculate the estimator: $\widehat{Y}_i = \widehat{M_{\theta^\star}}(X_i)$.
Step 3. For each X_j in the test sample $\mathcal{J} = 126, \dots, 200$, we set: $j_\star := \arg\min_{i \in \mathcal{L}} d_\theta(X_i, X_j)$.

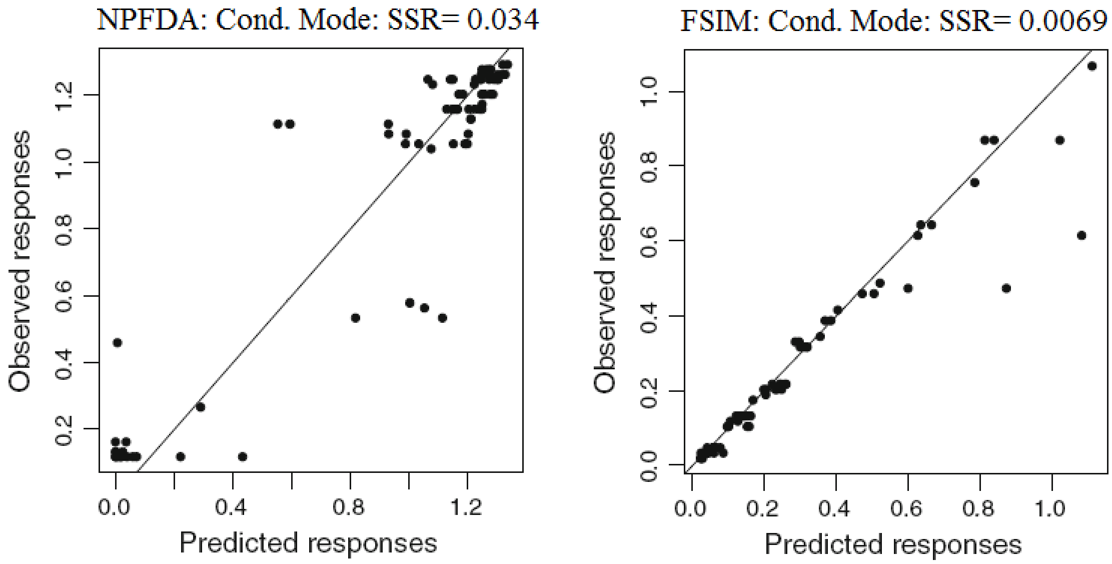

Figure 4. Prediction via the conditional mode by FSIM with error $SSR = 0.0069$ against NPFDA with error $SSR = 0.034$

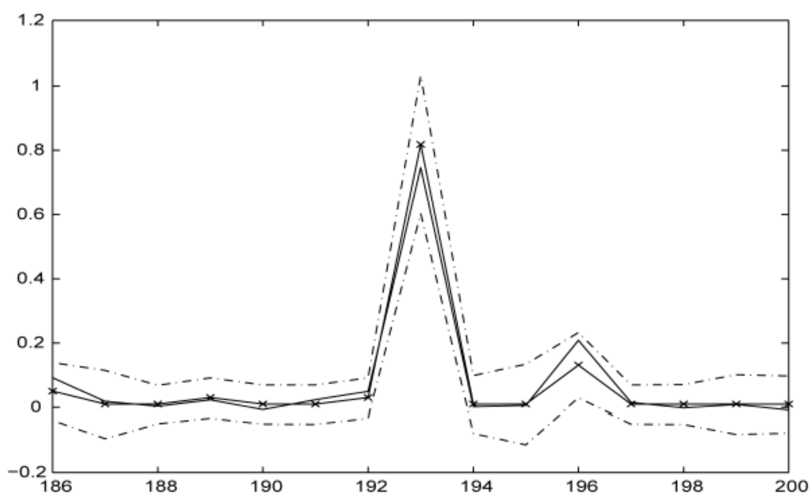

Figure 5. The 95% conditional predictive bands. The solid curve connects the true values. The crossed curve joins the predicted values. The dashed curves connects the lower and upper predicted values

Step 4**.** For each j in the test sample $\mathcal{J} = 126, \ldots, 200$, we define the confidence bands by $\left[\widehat{M_{\theta^\star}}(X_{j_\star}) - \tau_{0.975} \times \left(\frac{\widehat{\nu}(\theta^\star, X_{j_\star})}{\sqrt{\mathcal{L}h_H^3 \widehat{\phi}_{\theta^\star,x}(h_K)}}\right), \widehat{M_{\theta^\star}}(X_{j_\star}) + \tau_{0.975} \times \left(\frac{\widehat{\nu}(\theta^\star, X_{j_\star})}{\sqrt{\mathcal{L}h_H^3 \widehat{\phi}_{\theta^\star,x}(h_K)}}\right)\right]$.

We obtain the following figure which gathers asymptotic confidence bands study.

6. Conclusion

This paper focused on nonparametric estimation of conditional mode in the single functional index model for dependant stationary ergodic data. Both the asymptotic normality as well as a confidence interval of the resulted estimator are derived. The proofs are based on a combination of existing techniques. Our prime aim was to improve the performance of the single-index model for the conditional mode under the ergodic property. Through a series of simulations, our model outperforms nonparametric functional estimator. Note that this approach is more significant in the presence of a simple single functional index. The dimensionality of the model is the bias part while the dimensionality of the functional space of the explanatory variable is in the dispersion part. Then, the estimation and forecast accuracies between our FSIM and NPFDA models have been evaluated and compared. Via empirical analysis, it has been shown that the considered estimator has nice finite sample behavior for the predication and provides improved estimation and prediction accuracy compared to the NPFDA estimator.

REFERENCES

Akkal, F., Rabhi, A. and Keddani, L. (2021). Some asymptotic properties of conditional density function for functional data under random censorship, Applications and Applied Mathematics: An International Journal (AAM), Vol. 16, No. 1, pp. 12–42.

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

Aït-Saidi, A., Ferraty, F., Kassa, R. and Vieu, P. (2008). Cross-validated estimation in the single functional index model, Statistics, Vol. 42, pp. 475–494.

Aneiros, G., Cao, R., Fraiman, R., Genest, C. and Vieu, P. (2019). Recent advances in functional data analysis and high-dimensional statistics, J. Multivariate Anal., Vol. 170, pp. 3–9.

Attaoui, S. (2014). Strong uniform consistency rates and asymptotic normality of conditional density estimator in the single functional index modeling for time series data, J. AStA Adv. Stat. Anal., Vol. 98, pp. 257–286.

Attaoui, S. and Boudiaf, M. (2014). On the nonparametric conditional density and mode estimates in the single functional index model with strongly mixing data, Sankhya, Vol. 76-A, pp. 356-378.

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

Ataoui, S. and Ling, N. (2016). 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, Vol. 79, No. 5, pp. 485–511.

Bouchentouf, A.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. (Warsaw), Vol. 41, No. 4, 2014, pp. 301–322.

Cardot, H., Crambes, C. and Sarda, P. (2004). Estimation spline de quantiles conditionnels pour variables explicatives fonctionnelles, C.R. Acad. Sci. Paris, Vol. 339, pp. 141–144.

Chaouch, M. and Khardani, S. (2015). Randomly censored quantile regression estimation using functional stationary ergodic data, Journal of Nonparametric Statistics, Vol. 27, pp. 65–87.

Chowdhury, J. and Chaudhuri, P. (2019). Nonparametric depth and quantile regression for functional data, Bernoulli, Vol. 25, No. 1, pp. 395–423.

Cuevas, A. (2014). A partial overview of the theory of statistics with functional data, J. Statist. Plann. Inference, Vol. 147, pp. 1–23.

Ezzahrioui, M. and Ould-Saïd, M. (2010). Some asymptotic results of a non-parametric conditional modeestimator for functional time-series data, Statist. Neerlandica, Vol. 64, No. 2, pp. 171–201.

Ferraty, F., Laksaci, A. and Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models, Statist. Inference Stoch. Process., Vol. 9, pp. 47–76.

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

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice*, New York: Springer.

Geenens, G. (2011). *A nonparametric functional method for signature recognition. In Recent advances in functional data analysis and related topics*, Contrib. Statist., pages 141–147. Physica-Verlag/Springer, Heidelberg.

Goia, A. and Vieu, P. (2016). An introduction to recent advances in high/infinite dimensional statistics [Editorial], J. Multivariate Anal., Vol. 146, pp. 1–6.

Härdle, W. and Marron, J.S., (1985). Optimal bandwidth selection in nonparametric regression function estimation, Ann. Statist., Vol. 13, pp. 1465–1481.

Härdle, W., Hall, P. and Ichumira, H. (1993). Optimal smoothing in single-index models, Ann. Statist., Vol. 21, pp. 157–178.

Hristache, M., Juditsky, A. and Spokoiny, V. (2001). Direct estimation of the index coefficient in a single-index model, Ann. Statist., Vol. 29, No. 3, pp. 595–623.

Laib, N. and Louani, D. (2010). Nonparametric kernel regression estimation for functional stationary ergodic data: Asymptotic properties, J. Multivariate Anal., Vol. 101, pp. 2266–2281.

Laib, N. and Louani, D. (2011). Rates of strong consistencies of the regression function estimator for functional stationary ergodic data, Journal of Statistical Planning and Inference, Vol. 141, pp. 359–372.

Li, J., Huang, C. and Zhu, H. (2017). A functional varying-coefficient single-index model for functional response data, J. Amer. Statist. Assoc., Vol. 112, No. 519, pp. 1169–1181.

Ling, N. and Vieu, P. (2018). Nonparametric modeling for functional data: Selected survey and tracks for future, Statistics, Vol. 52, No. 4, pp. 934–949.

Ling, N. and Vieu, P. (2021). On semi-parametric regression in functional data analysis, Wiley Interdiscip. Rev. Comput. Stat., Vol. 13, No. 2, pp. 15–38.

Ling, N. and Xu, Q. (2012). Asymptotic normality of conditional density estimation in the single index model for functional time series data, Statistics and Probability Letters, Vol. 82, pp.

2235–2243.

Masry, E. (2005). Nonparametric regression estimation for dependent functional data: Asymptotic normality, Stochastic Process. Appl., Vol. 115, No. 1, pp. 155–177.

Novo, S., Aneiros, G. and Vieu, P. (2019a). Automatic and location-adaptive estimation in functional single-index regression, J. Nonparametr. Stat., Vol. 31, No. 2, pp. 364–392.

Novo, S., Aneiros, G., and Vieu, P. (2019b). Automatic and location-adaptive estimation in functional single-index regression, J. Nonparametr. Stat., Vol. 31, No. 2, pp. 364–392.

Shang, H. L. (2020). Estimation of a functional single index model with de- pendent errors and unknown error density, Comm. Statist. Simulation Comput., Vol. 49, No. 12, pp. 3111–3133.

Wang, G., Feng, X.-N. and Chen, M. (2016). Functional partial linear single-index model, Scand. J. Stat., Vol. 43, No. 1, pp. 261–274.