ECONOMETRICS. EKONOMETRIA
Advances in Applied Data Analysis
Year 2021, Vol. 25, No. 1
ISSN 1507-3866; e-ISSN 2449-9994

ASYMPTOTIC NORMALITY OF CONDITIONAL DENSITY AND CONDITIONAL MODE IN THE FUNCTIONAL SINGLE INDEX MODEL

Fatima Akkal
Statistics Laboratory Stochastic Processes
University Djillali LIABES of Sidi Bel Abbes, Sidi Bel Abbes, Algeria
e-mail: fatima.akkal@hotmail.com
ORCID: 0000-0001-6423-3127

Nadia Kadiri, Abbes Rabhi
University Djillali LIABES of Sidi Bel Abbes, Sidi Bel Abbes, Algeria
e-mails: nad.kad06@yahoo.com; rabhi abbes@yahoo.fr
ORCID: 0000-0002-3405-7414; 0000-0001-6740-0226

© 2021 Fatima Akkal, Nadia Kadiri, Abbes Rabhi
This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-sa/4.0/

Quote as: Akkal, F., Kadiri, N., and Rabhi A. (2021). Asymptotic normality of conditional density and conditional mode in the functional single index model. *Econometrics. Ekonometria. Advances in Applied Data Analysis*, *25*(1).

DOI: 10.15611/eada.2021.1.01
JEL Classification: C13, C14, C15

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 one derives a 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, and the asymptotic $(1 - \zeta)$ confidence interval of the conditional mode function is given for $0 < \zeta < 1$. A simulation study is also presented to illustrate the validity and finite sample performance of the considered estimator. Finally, the estimation of the functional index via the pseudo-maximum likelihood method is discussed.

Keywords: asymptotic normality, conditional density, functional single index model, functional random variable, nonparametric estimation.

1. Introduction

The statistical analysis of functional variables has grown considerably over the last two decades. Infact, an important innovation in measuring devices has emerged, permitting to monitor several objects in a continuous way, such as stock market index, pollution, climatology, and satellite images, etc.

Thus a new branch of statistics, called functional statistics, has been developed to treat observations as functional random elements.

The first results on the conditional models were obtained by (Ferraty, Laksaci, and Vieu, 2006). They 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. Quantile regression is the common way to describe the dependence structure between a response variable Y and some covariate X. Unlike the regression function (which is defined as the conditional mean) that relies only on the central tendency of the data, the conditional mode function allows the analyst to estimate the functional independence between variables for all portions of the conditional density of the response variable. However, compared with the standard approach based on functional conditional mean prediction that is sensitive to outliers, functional condition mode prediction could be seen as a reasonable alternative to the conditional mean because of its robustness, which allows to consider it as a useful alternative to the regression function.

The conditional model estimator has been widely used to estimate some characteristic features of the data set, such as the conditional mode, the conditional median, and the conditional quantiles. Many authors are interested in the estimation of the conditional mode of a scalar response given a functional covariate. Ferraty, Laksaci and Vieu (2006) introduced nonparametric the kernel-type estimators of some characteristics of the conditional cumulative distribution function and successive derivatives of the conditional density, and some asymptotic properties were established with particular attention to the conditional mode and conditional quantiles. An application to a chemometrical data set coming from the food industry was also presented. The uniform strong consistency with rates and the asymptotic normality for the kernel conditional mode estimator were obtained by Ezzahrioui and Ould-Saïd (2008) in the i.i.d case.

In the case of censoring, Ould-Saïd and Cai (2005) established the strong uniform convergence (with rate) of kernel conditional mode estimator for i.i.d random variables, while Ould-Saïd (2006) constructed a kernel estimator of the conditional quantile and establish its strong uniform convergence rate. Next, (Khardani, Lemdani, and Ould-Saïd, 2010) obtained strong consistency with the rate and asymptotic normality of the conditional mode (Khardani, Lemdani, and Ould-Saïd, 2011) established strong consistency with the rate of the conditional mode for the censored

dependent case, while (Khardani, Lemdani, and Ould-Saïd, 2014) presented asymptotic normality.

For infinite dimensional purpose, the study used the terminology *functional nonparametric*, where the term *functional* refers to the infinite dimensionality of the data, and where *nonparametric* refers to the infinite dimensionality of the model. Such *functional nonparametric* statistics is also called *doubly infinite dimensional* (see Ferraty and Vieu, 2003), for more details). Conditional density function estimation is one of the crucial problems in non-parametric statistics, see (De Gooijer and Zerom, 2003). Ling and Xu (2012) established the asymptotic normality of the conditional density estimator and the conditional mode estimator for the α-mixing dependence functional time series data. (Ling, Li, and Yang, 2014) investigated the pointwise almost complete consistency and the uniform almost complete convergence of the kernel estimation with a rate for the conditional density in the setting of the α-mixing functional data. Attaoui (2014) investigated the nonparametric estimation of the conditional density of a scalar response variable given a random variable taking values in separable Hilbert space. The author established under general conditions 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.

The single index models have been used and studied in both statistical and econometric literature, and are very popular in the economics community as they address two important concerns. The first is the reduction of dimension, since this type of model makes it possible to solve the problem of the scourge of the dimension. The second is related to the interpretability of the index (parameter) introduced in these models. The statistical study of these models, in the context of vectorial explanatory random variables, was initiated by Härdle and Marron (1985). Hristache, Juditsky, and Spokoiny (2001) provided both new theoretical and bibliographic elements. Several authors have worked on simple functional index models, e.g. (Attaoui and Boudiaf, 2014; Aït-Saidi, Ferraty, Kassa, and Vieu, 2008; Belabbaci, Rabhi, and Soltani, 2015; Ferraty, Peuch, and Vieu 2003)).

These models attracted the attention of many researchers, such as Aït-Saidi, Ferraty and Kassa (2005). Bouchentouf, Djebbouri, Rabhi, and Sabri (2014) established a nonparametric estimation of some characteristics of 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, Laksaci, and Ould-Saïd (2011) studied the functional single-index model via its conditional density kernel estimator, and established its pointwise and uniform almost complete convergence rates, and their results were extended to the 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 single-index models are becoming incrementally important and popular, and have been attracting considerable attention in the last few years because of their importance in several areas of science such as econometrics, biostatistics, medicine, etc. The single-index approach is extensively and mostly used in econometrics. Such kind of modelization is extensively studied in the multivariate case, for example in (Härdle, Hall, and Ichumira, 1993, Hristache, Juditsky, and Spokoiny, 2001). Based on the regression function, (Delecroix, Härdle, and Hristache, 2003) studied the estimation of the single-index and established some asymptotic properties. The literature is strictly limited in the case where the explanatory variable is functional (that is a curve). The first asymptotic properties in the fixed functional single-model were obtained by Ferraty and Vieu (2003), who established the almost complete convergence in the i.i.d case, of the link regression function of this model. Their results were extended to the dependent case by Aït-Saidi, Ferraty, and Kassa (2005). Aït-Saidi, Ferraty, Kassa, and Vieu (2008) studied the case where the functional single--index is unknown, and proposed an estimator of this parameter, based on the cross-validation procedure.

The main contribution of this work is to generalize the result of Ezzahrioui and Ould-Saïd (2008), where a functional parameter θ is present in the model. The results can be used to construct prediction intervals, for instance regarding electricity when one wants to construct a maximum interval of demand (or need) forchemometrical data coming from the food industry.

This study established 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 ani.i.d process with a single-index structure.

The paper is organized as follows. The model and some basic assumptions are presented in Section 2. Section 3 shows the main results, and the proofs of some lemmas and of the main result. In Section 4 an application of the conditional mode in functional single-index model is presented. Finally, Section 5 illustrates those asymptotic properties through some simulations.

2. The model and some basic assumptions

In all the paper, the authors denote by C, C_0or/and $C_{\theta,x}$ some generic constant in $\mathbb{R}_+^*$. The authors consider that, given the $(X_i, Y_i)_{i=1,\dots,n}$ a sequence of independent pairs functional samples, with the same distribution as (X, Y), where Y is a real-valued random variable and X is 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, the study considers $d_\theta(\cdot,\cdot)$ 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 in $\mathcal{H}$ and let $F(\theta, y, x)$ be the conditional cumulative distribution function (cond-cdf) of Y given $<\theta, X> = <\theta, x>$, specifically:

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

The authors implicitly assume the existence of a regular version of the conditional distribution and that itis absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$, the aim was 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, the authors denote by $f(\theta,\cdot,x)$ the conditional density of Y given $<x,\theta>$ and define the kernel estimator $\hat{f}(\theta,\cdot,x)$ for $f(\theta,\cdot,x)$ 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 $h_K := h_{n,K}$ (respectively $h_H := h_{n,H}$) is a sequence of bandwidths that decrease to zero as n goes to infinity.

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

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

The authors denote by $B_\theta(x,h) = \{\chi \in$ H/0 $<| < x - \chi,\theta>| < h\}$ be a ball of centre 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 $\phi_{\theta,x}(u) = P(d_\theta(x,X_i) \leq u) = P(X_i \in B_\theta(x,u))$.

Let N_x be a fixed neighbourhood of x in $\mathcal{H}$, $S_{\mathbb{R}}$ will be a fixed compact subset of $\mathbb{R}$, now, consider the following basic assumptions that are necessary to accomplish the main result of this paper.

(H1): $P(X \in B_\theta(x,h_K)) =: \varphi_{\theta,x}(h_K) > 0$; $\varphi_{\theta,x}(h_K) \to 0$ *as* $h_K \to 0$.

(H2): The conditional density $f(\theta,y,x)$ satisfies the Hölder condition, that is:

$$\forall (y_1,y_2) \in S_{\mathbb{R}} \times S_{\mathbb{R}},\ \forall (x_1,x_2) \in N_x \times N_x\ |f(\theta,y_1,x_1) - f(\theta,y_2,x_2)| \leq C_{\theta,x}(||x_1-x_2||^{b_1}+|y_1-y_2|^{b_2}),\ b_1 > 0,\ b_2 > 0.$$

(H3): The kernel H is a positive bounded function such that $\forall (t_1, t_2) \in \mathbb{R}^2$, $|H(t_1) - H(t_2)| \leq C|t_1 - t_2|$, $\int H^2(t)dt < \infty$ 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): 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 $\forall s \in [0,1]$.

(H6): The bandwidth h_K and h_H, small ball probability $\varphi_{\theta,x}(h_K)$ satisfying

(i) $\lim_{n\to\infty} h_K = 0, \lim_{n\to\infty} h_H = 0$ and $\lim_{n\to\infty} \frac{\log n}{nh_H\phi_{\theta,x}(h_K)} = 0.$

(ii) $h_H^{b_2}\sqrt{nh_H\phi_{\theta,x}(h_K)} \longrightarrow 0, as n \to \infty.$

(iii) $nh_H^3\varphi_{\theta,x}^3(h_K) \longrightarrow 0, as n \to \infty.$

3. Main result

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

Theorem 3.1. Under assumptions, there are (H1)-(H6)-(ii) for all $x \in \mathcal{H}$

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

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

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

Proof. In order to establish the asymptotic normality of $\hat{f}(\theta, y, x)$, further notations and definitions were needed. First the study considered the following decomposition

$$\begin{aligned}
\widehat{f}(\theta,y,x) - f(\theta,y,x) &= \frac{\widehat{f}_N(\theta,y,x)}{\widehat{f}_D(\theta,x)} - \frac{a_1(\theta,x)f(\theta,y,x)}{\alpha_1(\theta,x)}, \\
&= \frac{1}{\widehat{f}_D(\theta,x)}\left(\widehat{f}_N(\theta,y,x) - \mathbb{E}\widehat{f}_N(\theta,y,x)\right) \\
&\quad - \frac{1}{\widehat{f}_D(\theta,x)}\left(\alpha_1(\theta,x)f(\theta,y,x) - \mathbb{E}\widehat{f}_N(\theta,y,x)\right) \\
&\quad + \frac{f(\theta,y,x)}{\widehat{f}_D(\theta,x)}\left(\alpha_1(\theta,x) - \mathbb{E}\widehat{f}_D(\theta,x)\right) \\
&\quad - \frac{f(\theta,y,x)}{\widehat{f}_D(\theta,x)}\left(\widehat{f}_D(\theta,x) - \mathbb{E}\widehat{f}_D(\theta,x)\right), \\
&= \frac{1}{\widehat{f}_D(\theta,x)}\left(A_n(\theta,y,x) + B_n(\theta,y,x)\right),
\end{aligned}$$

where

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

and

$$\begin{aligned} A_n(\theta,y,x) &= \frac{1}{nh_H\mathbb{E}K_1(\theta,x)}\sum_{i=1}^{n}\left(H_i(y)-h_Hf(\theta,y,x)\right)K_i(\theta,x)-\mathbb{E}\left[\left(H_i(y)-h_Hf(\theta,y,x)\right)K_i(\theta,x)\right] \\ &= \frac{1}{nh_H\mathbb{E}K_1(\theta,x)}\sum_{i=1}^{n}N_i(\theta,y,x). \end{aligned}$$

It follows that,

$$\begin{aligned} nh_H\phi_{\theta,x}(h_K)Var(A_n(\theta,y,x)) &= \frac{\phi_{\theta,x}(h_K)}{h_H(\mathbb{E}K_1(\theta,x))^2}Var(N_1(\theta,y,x)), \\ &= V_n(\theta,y,x), \end{aligned}$$

and

$$B_n(\theta,y,x)=\alpha_1(\theta,x)f(\theta,y,x)-\mathbb{E}\widehat{f}_N(\theta,y,x)+f(\theta,y,x)(\alpha_1(\theta,x)-\mathbb{E}\widehat{f}_D(\theta,x)).$$

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

Lemma 3.1. Under conditions of Theorem 3.1

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

Proof.

$$\begin{aligned} V_n(\theta,y,x) &= \frac{\phi_{\theta,x}(h_K)}{h_H(\mathbb{E}K_1(\theta,x))^2}\mathbb{E}\left[K_1^2(\theta,x)\left(H_1(y)-h_hf(\theta,y,x)\right)^2\right], \\ &= \frac{\phi_{\theta,x}(h_K)}{h_H(\mathbb{E}K_1(\theta,x))^2}\mathbb{E}\left[K_1^2(\theta,x)\mathbb{E}\left(\left(H_1(y)-h_Hf(\theta,y,x)\right)^2\,|<\theta,X_1>\right)\right]. \end{aligned} \tag{3.1}$$

Using the definition of conditional variance

$$\mathbb{E}\left[\left(H_1(y)-h_H f(\theta,t,x)\right)^2\middle|<\theta,\,X_1>\right]=J_{1n}+J_{2n},$$

where

$$J_{1n}=V\,ar(H_1(y)|<\theta,\,X_1>),$$
$$J_{2n}=[\mathrm{E}(H_1(y)|<\theta,\,X_1>)-h_Hf(\theta,y,x)]^2.$$

• Concerning J_{1n}

$$J1n=\mathrm{E}\,(H_1^2(y)|<\theta,\,X_1>)-\mathrm{E}(H_1(y)|<\theta,\,X_1>)^2=J_1+J_2.$$

As for J_1, by the property of conditional expectation and by changing variables, one obtains as $n\to\infty$

$$\begin{aligned} J_1 &= \mathbb{E}\left(H_1^2\left(\frac{y-Y_1}{h_H}\right) | < \theta, X_1 >\right), \\ &= \int_{\mathbb{R}} H_1^2\left(\frac{y-v}{h_H}\right) f(\theta, v, X_1) dv, \\ &= \int_{\mathbb{R}} H_1^2(u) dF(\theta, y - uh_H, X_1). \end{aligned}$$

On the other hand, by applying (H2) and (H3)

$$\begin{aligned} J_1 &= \int_{\mathbb{R}} H_1^2(u) dF(\theta, y - uh_H, X_1) \\ &= h_H \int_{\mathbb{R}} H_1^2(u) f(\theta, y - uh_H, X_1) du, \\ &\leq h_H \int_{\mathbb{R}} H_1^2(u)\big(f(\theta, y - uh_H, X_1) - f(\theta, y, x)\big) du \\ &\quad + h_H \int_{\mathbb{R}} H_1^2(u) f(\theta, y, x) du, \\ &\leq h_H \left(C_{x,\theta} \int_{\mathbb{R}} H^2(u)\big(h_K^{b_1} + |v|^{b_2} h_H^{b_2}\big) du + f(\theta, y, x) \int_{\mathbb{R}} H^2(u) du\right), \\ &= \mathcal{O}\big(h_K^{b_1} + h_H^{b_2}\big) + h_H f(\theta, y, x) \int_{\mathbb{R}} H^2(u) du, \end{aligned} \tag{3.2}$$

as for J_2, $J_2' = \mathbb{E}\left(H_1(y) | < \theta, X_1 >\right) = \int_{\mathbb{R}} H\left(\frac{y-v}{h_H}\right) f(\theta, y, X_1) dv.$

Moreover, by changing variables one obtains:

$$J_2' = h_H \int_{\mathbb{R}} H(u)(f(\theta, y - uh_H, X_1 - f(\theta, y, x)) du + h_H f(\theta, y, x) \int_{\mathbb{R}} H(u) du,$$

the last equality is due to the fact that H is a probability density, and thus:

$$J_2' = \mathcal{O}\left(h_k^{b_1} + h_H^{b_2}\right) + h_H f(\theta, y, x)$$

Finally, as $n \to \infty$, $J_2 \longrightarrow \infty$. As for J_{2n}, by (H1)-(H3), one obtains that $n \to \infty$

$$J_{2n} \longrightarrow \infty.$$

Meanwhile, by (H1)-(H3), it follows that

$$\frac{\phi_{\theta,x}(h_K)\mathbb{E}K_1^2(\theta, x)}{\mathbb{E}^2 K_1(\theta, x)} \underset{n \longrightarrow \infty}{\longrightarrow} \frac{\alpha_2(\theta, x)}{(\alpha_1(\theta, x))^2}$$

which leads to combining equations (3.1) and (3.2)

$$V_n(\theta, t, x) \longrightarrow \frac{\alpha_2(\theta, x) f(\theta, y, x)}{(\alpha_1(\theta, x))^2} \int_{\mathbb{R}} H^2(u) du.$$

Lemma 3.2. If the assumptions (H1)-(H6) are satisfied, then$n \to \infty$

$$\sqrt{nh_H\phi_{\theta,x}(h_K)}B_n(\theta,y,x) \to 0, \text{ in probability.}$$

Proof. One has

$$\sqrt{nh_H\phi_{\theta,x}(h_K)}B_n(\theta,y,x) = \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}}{\widehat{f}_D(\theta,x)}\Big\{\mathbb{E}\widehat{f}_N(\theta,y,x) - \alpha_1(\theta,x)f(\theta,y,x) + f(\theta,y,x)(\alpha_1(\theta,x) - \mathbb{E}\widehat{f}_D(\theta,x))\Big\}.$$

Firstly, the authors observed that the results below

$$\frac{1}{\phi_{\theta,x}(h_K)}\mathbb{E}\left[K^l\left(\frac{<x-X_i,\theta>}{h_K}\right)\right] \longrightarrow \alpha_l(\theta,x), \text{ as } n\longrightarrow\infty, \text{ for } l=1,2, \tag{3.3}$$

$$\mathbb{E}\left[\widehat{f}_D(\theta,x)\right] \longrightarrow \alpha_1(\theta,x) \text{ and } \mathbb{E}\left[\widehat{f}_n(\theta,y,x)\right] \longrightarrow \alpha_1(\theta,x)f(\theta,y,x), \text{ as } n\longrightarrow\infty, \tag{3.4}$$

can be proved in the same way as in Ezzahrioui and Ould-Saïd (2008) corresponding to their lemmas Lemma 1 and Lemma 3, and then their proofs were omitted.

Secondly, on the one hand, making use of (3.3) and (3.4).

$$\{\mathbb{E}\widehat{f}_N(\theta,y,x) - \alpha_1(\theta,x)f(\theta,y,x) + f(\theta,y,x)(\alpha_1(\theta,x) - \mathbb{E}\widehat{f}_D(\theta,x))\} \underset{n\longrightarrow\infty}{\longrightarrow} 0$$

On otherhand,

$$\frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}}{\widehat{f}_D(\theta,x)} = \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\widehat{f}(\theta,y,x)}{\widehat{f}_D(\theta,x)\widehat{f}(\theta,y,x)} = \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}\widehat{f}(\theta,y,x)}{\widehat{f}_N(\theta,y,x)}.$$

As K(.)H(.) is continuous with support on [0,1], then by (H3) and (H4) $\exists m = \inf_{[0,1]} K(t)H(t)$ it follows that

$$\widehat{f}_N(\theta,t,x) \geq \frac{m}{h_H\phi_{\theta,x}(h_K)}, \text{ which gives } \frac{\sqrt{nh_H\phi_{\theta,x}(h_K)}}{\widehat{f}_N(\theta,t,x)} \leq \frac{\sqrt{nh_H^3\phi_{\theta,x}^3(h_K)}}{m}.$$

Finally, using (H6)-(iii), completes the proof of Lemma 3.2.

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

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

$$\widehat{M}(x) = \arg\sup_{y\in S_{\mathbb{R}}} \widehat{f^x}\ (y) \tag{4.1}$$

Note that the rest of the paper considered any value $\widehat{M_\theta}$ satisfying (4.1).

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

(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 j-th 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: $\exists\beta_0 > 0$, $\forall(y_1, y_2) \in S_{\mathbb{R}} \times S_{\mathbb{R}}$,

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

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

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

Theorem 4.1. Suppose that hypotheses (H1)-(H6) and (U1)-(U4) are satisfied. If

$$\lim_{n\to\infty} nh_H^3 \phi_{\theta,x}(h_K) = \infty, \text{ we have } as\ n \to \infty$$

$$\sqrt{\frac{nh_H^3 \phi_{\theta,x}(h_K)}{\nu^2(\theta,x)}}(\widehat{M_\theta}(x) - M_\theta(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \tag{4.2}$$

where

$$\nu^2(\theta,x) = \frac{\alpha_2(\theta,x) f(\theta, M_\theta(x), x)}{\left(\alpha_1(\theta,x) f^{(2)}(\theta, M_\theta(x), x)\right)^2} \int (H'(t))^2 dt. \tag{4.3}$$

Proof. Firstly, by (4.1) 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

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

In order to prove (4.3), one only needs to show that

$$-\sqrt{nh_H^3 \phi(h_K)}\hat{f}^{(1)}(\theta,M_\theta(x),x) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\nu_0^2(\theta,x)), \text{ and} \tag{4.4}$$

$$\widehat{f}^{(2)}(\theta, M_\theta^*(x), x) \longrightarrow \widehat{f}^{(2)}(\theta, M_\theta(x), x) \neq 0, \tag{4.5}$$

in probability,
where

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

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

$$\begin{aligned}\mathbb{P}\left(|\widehat{M_\theta}(x)-M_\theta(x)|\geq\varepsilon\right) &\leq \mathbb{P}\left(|f(\theta,M_\theta(x),x)-\widehat{f}(\theta,M_\theta(x),x)|\geq\frac{\eta(\varepsilon)}{2}\right)\\ &+\mathbb{P}\left(|\widehat{f}(\theta,\widehat{M_\theta}(x),x)-f(\theta,\widehat{M_\theta}(x),x)|\geq\frac{\eta(\varepsilon)}{2}\right).\end{aligned}\tag{4.6}$$

Thus, similarly to Ferraty and Vieu (2006), by (H1)-(H4) and (H6)-(i), there is $\hat{f}(\theta,y,x)\longrightarrow f(\theta,y,x)$ in probability, which implies that $\widehat{M_\theta}(x)\to M_\theta(x)$ in probability by (4.6) as $n\to\infty$. Similarly, the methodology can be also applied to obtain $\hat{f}^{(2)}(\theta,y,x)\longrightarrow f^{(2)}(\theta,y,x)$ in probability as $n\to\infty$ by (H1), (H4), (H6), (U3) and (U4). Therefore, (4.5) is valid by the fact that $f^{(2)}(\theta,y,x)$ is uniformly continuous with respect to y on $S_{\mathbb{R}}$. Next, (4.4) is proved. In fact, since

$$\begin{aligned}\widehat{f}^{(1)}(\theta,M_\theta(x),x) &= \frac{1}{\widehat{f}_D(\theta,x)}\left(\widehat{f}_N^{(1)}(\theta,M_\theta(x),x)-\mathbb{E}\widehat{f}_N^{(1)}(\theta,M_\theta(x),x)\right)\\ &-\frac{1}{\widehat{f}_D(\theta,x)}\left(f^{(1)}(\theta,M_\theta(x),x)-\mathbb{E}\widehat{f}_N^{(1)}(\theta,M_\theta(x),x)\right).\end{aligned}\tag{4.7}$$

By (U1), (U3)-(U4), (4.2) and (4.7), similar to the proof of Lemmas, Lemma 3.1 and Lemma 3.2 respectively, (4.4) 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 $f(\theta,M_\theta(x),x)$. Hence $M_\theta(x)$, and $f(\theta,M_\theta(x),x)$ should be replaced by their respective estimators $\widehat{M}_\theta(x)$, and $\hat{f}(\theta,M_\theta(x),x)$.

Due to the unknown functions $\alpha_j := \alpha_j(\theta,x)$ and $f(\theta,y,x)$ 0intervening in the expression of the variance, it is necessary to estimate the quantities $\alpha_1(\theta,x)$, $\alpha_2(\theta,x)$ and $f(\theta,y,x)$, respectively.

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

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

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

By applying the kernel estimator $f(\theta,y,x)$ 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, one obtains 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, one obtains

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

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

$\alpha b_1(\theta,x)\xrightarrow{P}\alpha_1(\theta,x)$, $\alpha b_2(\theta,x)\xrightarrow{P}\alpha_2(\theta,x)$, and , $\dfrac{\widehat{\phi}_{\theta,x}(h_K)}{\phi_{\theta,x}(h_K)}\xrightarrow{\mathbb{P}}1,$ as $n\to\infty$.

Therefore, one obtains

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

This yields the proof of Corollary 4.1.

Finally, in order to show the asymptotic $(1-\xi)$ confidence interval of $M_\theta(x)$, one needs 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 $n\to\infty$ one has

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

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, one obtains

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

Further, by considering Lemma 3.2, (4.5) and (4.6), one obtains as $n\to\infty$,

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

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 $f(\theta,y,x)$ and $M_\theta(x)$ are given by

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

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

5. Simulation study

To study the behaviour of the conditional mode estimator, in this part two examples of simulation were considered. In the first one, the authors compared the model FSIM (functional single index model) with that of NPFDA (non-parametric functional data analysis) and in the latter, knowing the distribution of the regression model (the distribution is known and usual), looked at the behaviour of this estimator of the conditional density function with respect to this distribution. Therefore, the best way to know the behaviour of the estimator of conditional density is to compute its mean square error. Thus this part of paper compared the conditional density estimation in the FSIM which is the authors' model and the conditional density estimation in the NPFDA defined in (5.1).

$$\widehat{f}_n(x|y) = \frac{h_H^{-1}\sum_{i=1}^{n} K\left(h_K^{-1}d(<x,X_i)\right) H\left(h_H^{-1}(y-Y_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}d(x,X_i)\right)}, \tag{5.1}$$

where, the authors estimated the conditional mode $\widehat{M}(x)$ with a random variable $M(x)$ such that

$$M(x) = \arg\sup_{y\in S_{\mathbb{R}}} f(x|y) \text{ and } \widehat{M}(x) = \arg\sup_{y\in S_{\mathbb{R}}} \hat{f}_n(x|y).$$

Therefore one has 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, the purpose consists in assessing the performance, in terms of prediction, of $\widehat{M}_\theta(x)$ and $\widehat{M}(x)$. For each given predictor $(X_j)_{j\in J}$ in the testing subsample, the authors were interested in the prediction of the response variable $(Y_j)_{j\in J}$ via the single functional index conditional mode $\widehat{M}_\theta(x)$ and the fully nonparametric conditional mode $\widehat{M}(x)$ so as to compare the finite-sample behaviour of the estimator. As the assessment tool, the authors considered the mean square error (MSE) defined as follows:

$$SSR = \frac{1}{|\mathcal{J}|}\sum_{j\in\mathcal{J}}\left(Y_j - \widehat{Y}_j\right)^2, \tag{5.2}$$

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

Furthermore, some tuning parameters had to be specified. The kernel $K(\cdot)$ was chosen to be the quadratic function defined as $K(u) = \frac{3}{2}(1-u^2)\mathbf{1}_{[0,1]}$ and the cumulative df $H(u) = \int_{-\infty}^{u} \frac{3}{4}(1-z^2)\,1_{[-1.1]}(z)dz$.

The semi-metric $d(\cdot,\cdot)$ is specified according to the choice of the functional space 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 this 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 $\widehat{\alpha_1}(\theta,x)$, $\widehat{\alpha_2}(\theta,x)$, $\widehat{M_\theta}(x)$, $\hat{f}$, and $\widehat{\emptyset}_{\theta,x}(h_K)$ one obtains

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

Now in order to simplify the implementation of the methodology, the authors took the bandwidths $h_H \sim h_K = h$, where h is chosen by the cross-validation method on the k-nearest neighbours (see Ferraty and Vieu, 2006, p. 102).

5.1. Simulation 1: the case of smooth curves

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

$$T_i = R(X_i) + \varepsilon_i\ ,i = 1, \ldots, n,$$

where ε_i isa 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 $\hookrightarrow \mathcal{N}(0, 1)$,$\hookrightarrow \mathcal{N}(0, 0.03)$ and $\hookrightarrow \mathcal{N}(0, 0.05)$. The variables a and c are Bernoulli's laws Bernoulli $\mathrm{B}(0.5)$. Figure 1 shows a sample of 200 curves representing a realization of the functional random variable X.

Taking into account the smoothness of the curves $X_i(t)$ (see Figure 1), the authors chose the distance $deriv_1$ (the semi-metric based on the first derivatives of the curves) in H as:

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

as semi-metric.

Then, the study considered a nonlinear regression function defined as

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

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

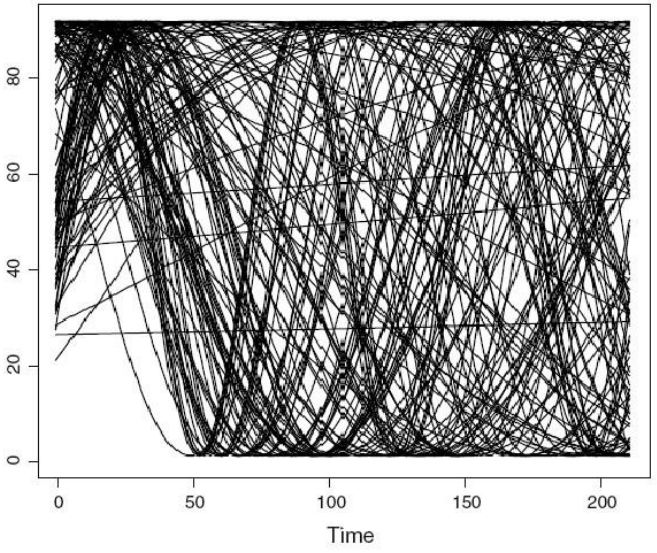

Fig. 1. A sample of 200 curves $X_{i=1,\dots,200}(t_j)$, $t_{j=1,\dots,200} \in [0,1]$

Source: own calculations.

In practice this parameter can be selected by the cross-validation approach (see Aït--Saidi et al., 2008). In this step it may be that one can select the real-valued function $\theta(t)$ among the eigenfunctions of the covariance operator $\mathbb{E}[(X' - \mathbb{E}X') < X', . >_{\mathcal{H}}]$ where $X(t)$ is a diffusion process defined on a real interval [a, b] and $X'(t)$ its first derivative (see Attaoui and Ling, 2016). Hence for the 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: $\frac{1}{\mathcal{L}}\sum_{i\in\mathcal{L}}(X'_i - \mathbb{E}X')^t(X'_i - \mathbb{E}X')$, is the one best approximation of the functional parameter θ. Now, let us denote θ^* the first eigenfunction corresponding to the first higher eigenvalue of the empirical covariance operator, which replaces θ during the simulation step.

In the following graphs, the covariance operator for $\mathcal{L} = \{1, \dots, 200\}$ gives the discretization of the first eigenfunction θ(represented by a continuous curve), twenty and all the eigenfunctions $\theta_i(t)$ (see Figure 2, Figure 3 and Figure 4).

In the simulation part, the sample of 200 was divided into two parts. The first one from 1 to 125 was used to make the simulation, and the second from 126 to 200 served for the prediction.

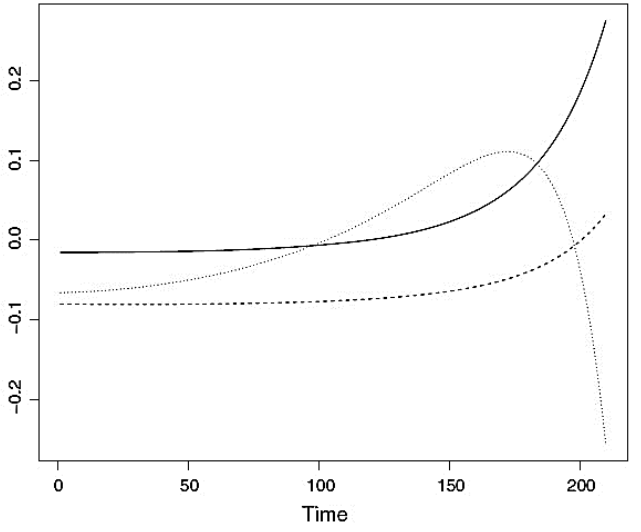

Fig. 2. The curves θi = 1,2,3(tj), tj = 1, ...,200 ∈ [0, 1]

Source: own calculations on ground (Attaoui and Ling, 2016).

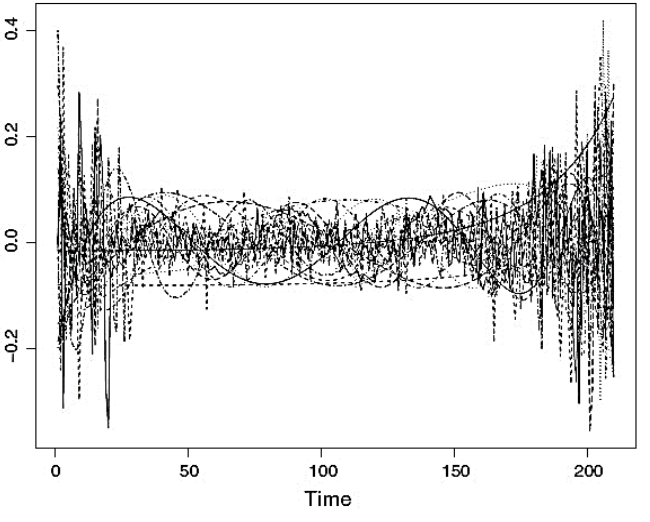

Fig. 3. The curves θi =1, ..., 20(tj), tj = 1 ,..., 200 ∈ [0, 1]

Source: own calculations on ground (Attaoui and Ling, 2016).

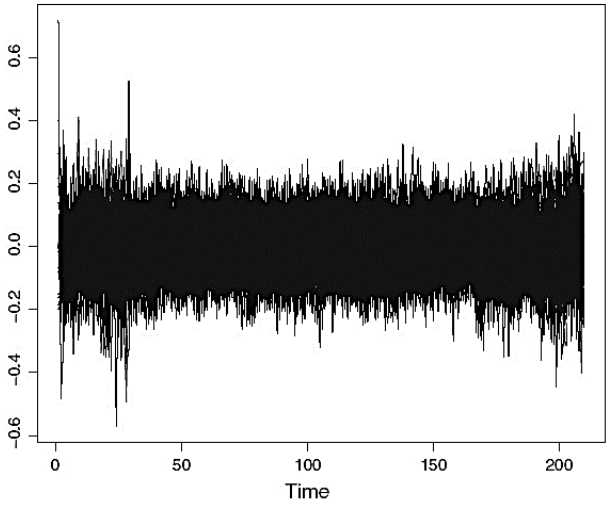

Fig. 4. The curves $\theta i = 1, ..., 200(tj)$, $tj = 1, ..., 200 \in [0, 1]$

Source: own calculations on ground (Attaoui and Ling, 2016).

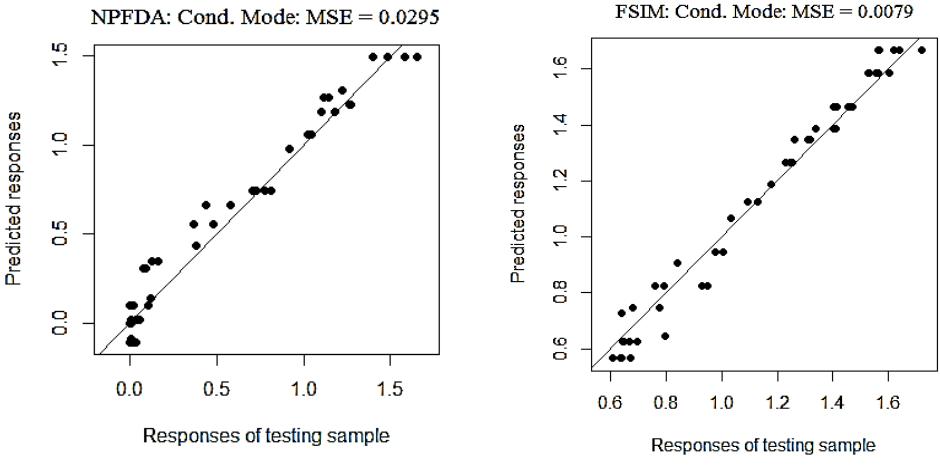

Fig. 5. Prediction via the conditional mode by FSIM with error $SSR = 0.0069$ against NPFDA with error $SSR = 0.0295$

Source: own calculations.

The following steps were taken:

Step 1. Simulate the responses variables Y_i.

Step 2. For each j in the test sample $\mathcal{J} = \{126, ..., 200\}$, compute: $\hat{Y}_j = \widehat{M_{\theta^*}}(X_j)$ and $\hat{Y}_j = \widehat{M}(X_j)$.

Finally, the authors presented the results by plotting the predicted values versus the true values and computed the sum of squared residuals (SSR) defined by (5.2).

One can see that the sum of squared residuals (SSR) of the method of Functional--Single-Index-Model (FSIM) is less than the one of the Non-Parametric-Functional--Data-Analysis (NPFDA). This is confirmed by the following graphs, which compare the conditional mode by (FSIM) against the conditional mode by (NPFDA) (Figure 1). Thus the estimator is acceptable. As was intuitively expected, it can be observed that the mean square errors of the estimator are smaller than that of NPFDA. Therefore, the FSIM model produces much more accurate estimations than the NPFDA model in all the criteria.

In order to construct conditional confidence bands the study proceeded by the following algorithm:

Step 1. $< \theta^*, X_i >, \ldots, < \theta^*, X_{200} >$, generate independently the variables $\varepsilon_1, \ldots, \varepsilon_{200}$, then simulate the response variables $Y_i = r(< \theta^*, X_i >) + \varepsilon_i$, where $r(< \theta^*, X_i >) = \exp(10(< \theta^*, X_i > -0.05))$ and generate independently the variables $\varepsilon_1, \ldots, \varepsilon_{200}$.

Step 2. For each *i* in the training sample, calculate the estimator: $\hat{Y}_i = \widehat{M_{\theta^*}}(X_i)$.

Step 3. For each X_j in the test sample $\mathcal{J}$ = 126, ..., 200, set: $j_* := arg \min_{i \in \mathcal{L}} d_\theta(Xi, Xj)$.

Step 4. For each *j* in the test sample $\mathcal{J}$ = 126, ..., 200, define the confidence bands by

$$\left[\widehat{M_{\theta^*}}(X_{j_*}) - \tau_{0.975} \times \left(\frac{\hat{\nu}(\theta^*, X_{j_*})}{\sqrt{\mathcal{L} h_H^3 \hat{\phi}_{\theta^*, x}(h_K)}}\right), \widehat{M_{\theta^*}}(X_{j_*}) + \tau_{0.975} \times \left(\frac{\hat{\nu}(\theta^*, X_{j_*})}{\sqrt{\mathcal{L} h_H^3 \hat{\phi}_{\theta^*, x}(h_K)}}\right)\right].$$

One obtains the following figure which joins the asymptotic confidence bands study.

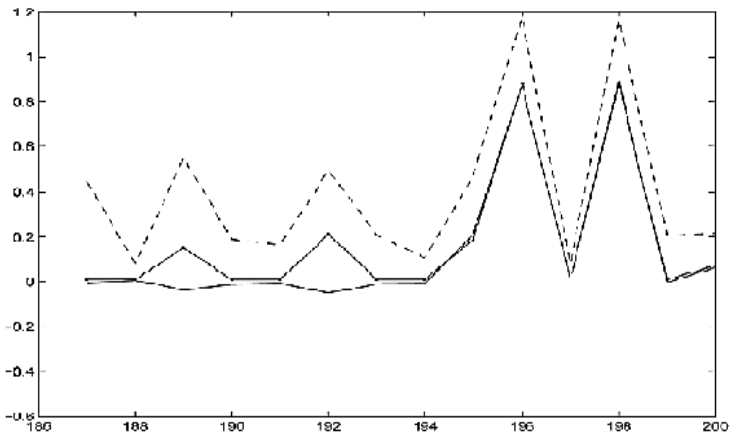

Fig. 6. The 95% conditional predictive bands. The solid curve connects the true values. The crossed curve joins the predicted values. The dashed curve connects the predicted values

Source: own calculations.

For the purpose of making a decision, the authors chose another Example (5.1) in which the distribution of the model is known and usual.

Example 5.1. Let $X_1, \ldots, X_n$ be a standard Brownian movement in [0,1], with $n = 100$. This study focused on the linear model with defined functional index by:

$$Y_i = \frac{|<\theta, X_i>|}{150} + 0.5\mathcal{N}(0,1).$$

The study kept the values of θ^* and $(X_i)_{i=1,\ldots,100}$ of the precedent example (θ is replaced by θ^*).

According to this model, it is clear that, when $X = x$, the variable $Y \hookrightarrow \mathcal{N}(\frac{|<\theta,x>|}{150}, 4)$. In this study, as the curves are rough (see Figure 7) the study used the semi-metric pca.

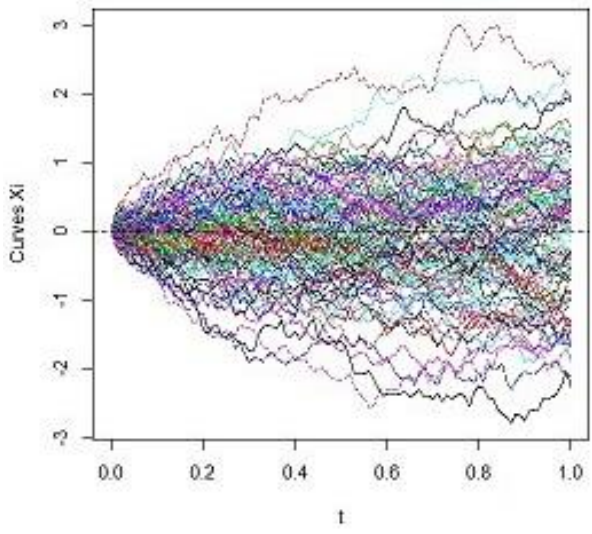

Fig. 7. Standard Brownian motions

Source: own calculations.

Table 1 presents the MSE of FSIM and NPFDA models by considering different values of ξ,$ with 100 replications.

Table 1. Estimation accuracy of the conditional mode function between the functional single index model and the nonparametric functional model for different values of ξ

Error	Model	Semi-metric	$n = 100$		
			$\xi = 0.05$	$\xi = 0.50$	$\xi = 0.95$
MSE	FSIM	pca	0.0116	0.0112	0.0127
	NPFDA	pca	0.0634	0.0621	0.0641

Source: own calculation.

From the obtained results presented in Table 1, one can confirm that the FSIM estimator of conditional mode is better than that of NPFDA. It gives a smaller mean square error, hence it allows for a more accurate estimation.

After the calculation of the errors, one find for this method an error SSR = 0.091. The NPFDA method gives an error SSR = 0.1181, while the real error (knowing that $Y \hookrightarrow \mathcal{N}(\frac{|<\theta,x>|}{150},4)$ is equal to $SSR = 1.672 \times 10^{-29}$SSR = 1.938. This confirms once again that this estimator is much better than that in the NPFDA case. Therefore, in the context of i.i.d data, this estimator is more preferable.

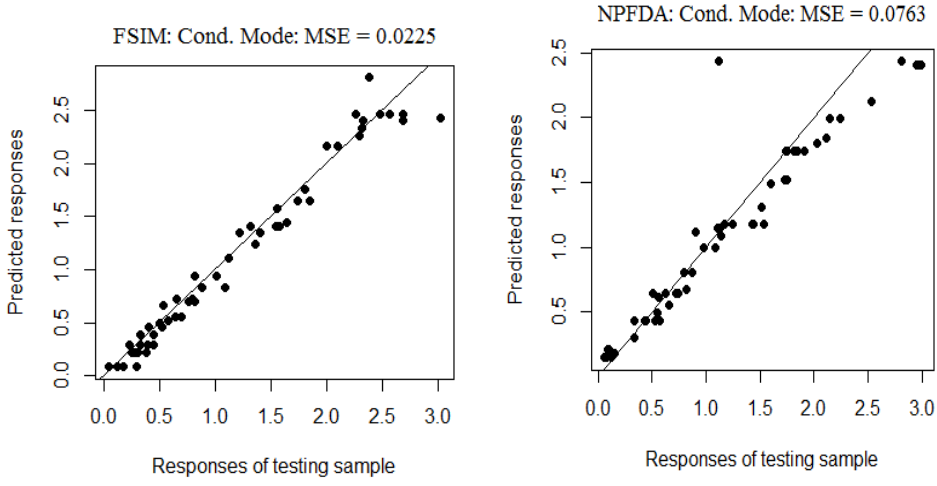

Fig. 8. Prediction via the conditional mode by NPFDA with error *SSR* = 0.0763 against FSIM with error *SSR* = 0.0225

Source: own calculations.

6. Conclusions

This paper focused on the nonparametric estimation of the conditional mode in the single functional index model for independent data. Both the asymptotic normality as well as a confidence interval of the resulted estimator were derived. The proofs are based on a combination of existing techniques. The study's prime aim was to improve the performance of the single-index model for the conditional mode with a scalar response variable conditioned by a functional Hilbertian regressor under the independent property. Through a series of simulations, this model out performs the nonparametric functional estimator. The contribution in this study is focused on the estimation of the conditional density function as well on the estimation of the regression for complete data in a functional framework. The first approach is used for

the estimation of the conditional mode. Then on parametric aspect is properly exploited in the first two sections by the given hypotheses. The proposed estimators are consistent an asymptotically distributed under appropriate conditions. 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 the FSIM and NPFDA models were evaluated and compared, and via empirical analysis, it was shown that the considered estimator has good finite sample behaviour for the prediction, and provides improved estimation and prediction accuracy compared to the NPFDA estimator. Research in the non--parametric field remains an open matter that will be the subject of several future studies in order to improve and high light the results obtained in this study.

In addition, in order to explore the effectiveness of the authors' method in real situations, one can apply this approach to data constituting hourly electricity demand as well as spectrometric data. An other real example is forecasting the daily peak in electricity demand, as the accurate prediction of daily peak load demand is very important for decisions made in the energy sector. In fact, short-term load forecastsenable effective load shifting between transmission substations, scheduling of the startup times of peak stations, load flow analysis and power system security studies. Other real data applications (Maximum Ozone Concentration, Peak electricity demand) can be highlighted, asseveral attractive features of a functional prediction context, with unknown scale parameter estimator.

Research in the nonparametric field remains an open matter which will be the subject of several future studies in order to improve and highlight the results obtained in this work. To extend this study of estimation of the conditional mode to the estimation of the conditional models of a MAR (missing at random) response to the independent case and the dependent case, another type of dependency could be considered such as the quasi-partner.

Developing the asymptotic properties of a kernel estimator of the k-nearest neighbors, and generalize the results obtained by using other families of semi-metrics in order to improve the prediction performance of the estimators means that the choice of the smoothing window is important.

References

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

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

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*., (98), 257-286.

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

Attaoui, S., Laksaci, A., and Ould-Saïd, E. (2011). A note on the conditional density estimatein the single functional index model. *Statist. Probab. Lett*., *81*(1), 45-53.

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

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)*, *10*(1), 114-138.

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.*, *41*(4), 301-322.

De Gooijer, J. G., and Zerom, D. (2003). On conditional density estimation. *Statistica Neerlandica*, (57), 159-76.

Delecroix, M., Härdle, W., and Hristache, M. (2003). Efficient estimation in conditional singleindex regression. *J. Multivariate Anal.*, (86), 213-226.

Ezzahrioui, M., and Ould-Saïd, M. (2008). Asymptotic normality of a non parametric estimator of the conditional mode function for functional data. *Nonparametr. Stat.*, (20), 3-18.

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

Ferraty, F., Peuch, A., and Vieu, P. (2003). Modèle à indice fonctionnel simple. *C.R. Mathématiques,* (336), 1025-1028.

Ferraty, F., and Vieu, P. (2003). Curves discrimination: A nonparametric functional approach. *Computat. Statist. Data Anal*., (44), 161-173.

Ferraty, F., and Vieu, P. (2006). *Nonparametric functional data analysis: Theory and practice*. New York: Springer.

Härdle, W., and Marron, J. S. (1985). Optimal bandwid the selection in nonparametric regression function estimation. *Ann. Statist*., 13, 1465-1481.

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

Hristache, M., Juditsky, A., and Spokoiny, V. (2001). Direct estimation of the index coefficient in a single-index model. *Ann. Statist.*, *29*(3), 595-623.

Khardani, S., Lemdani, M., and Ould-Saïd, E. (2010). Some asymptotic properties for a Smooth kernel estimator of the conditional mode under random censorship. *Journal of Korean Statistical Society*, (39), 455-469.

Khardani, S., Lemdani, M., and Ould Saïd, E. (2011). Uniform rate of strong consistency for a smooth kernel estimator of the conditional mode under random censorship. *J. Statist. Plann. and Inf.*, (141), 3426-3436.

Khardani, S., Lemdani, M. and Ould Saïd, E. (2014). On the central limit theorem for a conditional mode estimator of a randomly censored time series. *Journal of Statistical Theory and Practice*, (8), 722-742.

Kołodko, G. (2010). Neoliberalizm i światowy kryzys gospodarczy. *Ekonomista*, (1), 23-30.

Kowalski, J., Nowak, T., and Pisarek, W. (red.). (2018). *Aspekty zarządzania*. Warszawa: Wydawnictwo Naukowe PWN.

Kowalski, J. (red.). (2013). *Rola polityki logistycznej*. Wrocław: Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu.

Laib, N., and Louani, D. (2010). Nonparametric kernel regression estimation for functional stationaryergodic data: Asymptotic properties. *J. Multivariate Anal.*, (101), 2266-2281.

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*, (82), 2235-2243.

Ling, N., Li, Z., and Yang, W. (2014). Conditional density estimation in the single functional index model for α-mixing functional data. *Communications in Statistics – Theory and Methods*, *43*(3), 441-454. doi: 10.1080/03610926.2012.664236

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*, (82), 2235-2243.

OECD. (2010). *Sprawozdanie dotyczące przygotowania Strategii Zielonego Wzrostu*. Retrieved from http://www.oecd-ilibrary.org

Ould-Saïd, E. (2006). A strong uniform convergence rate of kernel conditional quantile estimator under random censorship. *Statistics and Probability Letters*, (76), 579-586.

Ould-Saïd, E., and Cai, Z. (2005). Strong uniform consistency of nonparametric estimation of the censored conditional mode function. *Nonparametric Statistics*, *17*(7), 797-806.

Sturman, M. C., Shao, L., and Katz, J. H. (Eds.). (2012). *The effect of culture on the curvilinear relationship between performance turnover*. Berlin: Springer.

The World Bank. (2012). *Inclusive green growth: The pathway to sustainable development*. Washington, DC: The World Bank.

Ustawa z dnia 17 grudnia 2004 r. o odpowiedzialności za naruszenie dyscypliny finansów publicznych (Dz. U. z 2004 r. Nr 14, poz. 114 ze zm.)

ASYMPTOTYCZNA NORMALNOŚĆ ROZKŁADU WARUNKOWEJ GĘSTOŚCI I WARUNKOWEJ DOMINANTY MODELU JEDNOWSKAŹNIKOWEGO

Streszczenie: Celem niniejszego artykułu jest zbadanie nieparametrycznej estymacji warunkowej gęstości skalarnej zmiennej zależnej *Y*, przy założeniu, że zmienna objaśniająca *X* przyjmuje wartość w przestrzeni Hilberta, gdy próbka obserwacji jest traktowana jako niezależne zmienne losowe o identycznym rozkładzie i są one połączone jedną funkcjonalną strukturą indeksu. Przede wszystkim wprowadzono estymator typu jądrowego dla warunkowej funkcji gęstości (*cond-df*). Następnie określono asymptotyczne właściwości warunkowego estymatora gęstości, gdy obserwacje są połączone ze strukturą pojedynczego indeksu, i wyprowadzano centralne twierdzenie graniczne (CLT) warunkowego estymatora gęstości w celu zaprezentowania asymptotycznej normalności estymacji jądrowej tego modelu. W aplikacji przedstawiono dominantę warunkową w funkcjonalnym modelu z pojedynczym indeksem, a także asymptotyczny (1-ζ) przedział ufności funkcji dominanty warunkowej dla $0 < \zeta < 1$. Na koniec omówiono estymację indeksu funkcjonalnego metodą pseudomaksymalnej wiarygodności.

Słowa kluczowe: asymptotyczna normalność, gęstość warunkowa, funkcjonalny model pojedynczego wskaźnika, funkcjonalna zmienna losowa, estymacja nieparametryczna.