ECONOMETRICS
EKONOMETRIA
Advances in Applied Data Analysis

Econometrics. Ekonometria.
Advances in Applied Data Analysis
Year 2024, Vol. 28, No. 1
ISSN 2449-9994

Central Limit Theorem for Conditional Mode in the Single Functional Index Model with Data Missing at Random

Anis Allal

University Djillali Liabes of Sidi Bel Abbes, Algeria

e-mail: anis.allal@univ-sba.dz

ORCID: 0000-0002-1553-8132

Abdessamad Dib

University Djillali Liabes of Sidi Bel Abbes, Algeria

e-mail: dibaddassamad@gmail.com

ORCID: 0000-0003-3654-8185

Abbes Rabhi

University Djillali Liabes of Sidi Bel Abbes, Algeria

e-mail: rabhi_abbes@yahoo.fr

ORCID: 0000-0001-6740-0226

Quote as: Allal, A., Dib, A., and Rabhi, A. (2024). Central Limit Theorem for Conditional Mode in the Single Functional Index Model with Data Missing at Random. *Econometrics. Ekonometria. Advances in Applied Data Analysis*, *28*(1), 39-60.

DOI: 10.15611/eada.2024.1.04

JEL Classification: C10, C13, C14, C15, C24

Abstract: This paper concentrates on nonparametrically estimating the conditional density function and conditional mode within the single functional index model for independent data, particularly when the variable of interest is affected by randomly missing data. This involves a semi-parametric single model structure and a censoring process on the variables. The estimator's consistency (with rates) in a variety of situations, such as the framework of the single functional index model (SFIM) under the assumption of independent and identically distributed (i.i.d) data with randomly missing entries, as well as its performance under the assumption that the covariate is functional, are the main areas of focus.

For this model, the nearly almost complete uniform convergence and rate of convergence are established.

The rates of convergence highlight the critical part that the probability of concentration play in the law of the explanatory functional variable. Additionally, we establish the asymptotic normality of the derived estimators proposed under specific mild conditions, relying on standard assumptions in Functional Data Analysis (FDA) for the proofs. Finally, we explore the practical application of our findings in constructing confidence intervals for our estimators. The rates of convergence highlight the critical part that the probability of concentration play in the law of the explanatory functional variable.

Keywords: functional data analysis, functional single-index process, kernel estimator, missing at random, nonparametric estimation, small ball probability

1. Introduction

The field of nonparametric statistics deals with estimating unknown functions from data without assuming a specific parametric form for the underlying distribution. Functional data analysis (FDA) is a subfield of statistics that deals with data that can be viewed as functions, often represented as curves or smooth functions. These could be, for example, time series data or spatial data where each observation is a curve or function.

In this context, the asymptotic results refer to the behaviour of estimators as the sample size approaches infinity. Some common asymptotic results include:

Consistency: an estimator is consistent if it converges in probability to the true value of the parameter being estimated as the sample size grows. In other words, the estimator becomes more and more accurate as more data becomes available.

Asymptotic normality: under certain conditions, the estimator's distribution approaches a normal distribution as the sample size increases. This property is essential for constructing confidence intervals and conducting hypothesis tests.

Asymptotic efficiency: an efficient estimator achieves the smallest possible variance among all consistent estimators, as the sample size becomes large.

Rates of convergence: this refers to how fast an estimator converges to the true parameter value as the sample size increases. Different estimators may have different rates of convergence.

The specific asymptotic results for estimators in nonparametric models with functional data would depend on the specific estimation method and the assumptions made in the analysis. Researchers typically derive these results by establishing mathematical properties of the estimators and their behaviour as the sample size grows. These results provide valuable insights into the performance and limitations of different estimation approaches in functional data analysis.

Estimating the conditional mode for functional data in the Single Index Model (SIM) with Missing Data at Random (MAR) is a challenging problem in nonparametric statistics and functional data analysis. The Single Index Model is a popular framework used for reducing dimensionality and modelling complex relations between covariates and responses in a simplified way. When dealing with functional data, where each observation is a curve or function, the SIM is extended to handle functional predictors and responses.

In the context of Missing Data at Random (MAR), the ‘missingness’ in the data is assumed to be dependent on the observed variables but not on the unobserved (missing) data. This assumption allows for various imputation methods to handle missing values in the functional data effectively.

Estimating the conditional mode involves finding the most likely value of the response variable given the observed functional predictors under the SIM, which can be a challenging task, especially with missing data. Several approaches and algorithms have been proposed in the literature to address this problem. Some common methods include:

Multiple Imputation: one approach is to use multiple imputation techniques to impute the missing functional data values. Several imputed datasets are created, and the conditional mode is estimated for each imputed dataset separately. The results are then combined to obtain an overall estimate.

EM Algorithm: the Expectation-Maximisation (EM) algorithm can be used to iteratively impute the missing values and estimate the conditional mode in a SIM setting with missing data.

Nonparametric Smoothing: nonparametric smoothing techniques, such as kernel smoothing or spline methods, can be employed to estimate the conditional mode for functional data.

Local Polynomial Estimation: local polynomial estimators can be used to estimate the conditional mode by fitting polynomials to local segments of the data.

Profile Likelihood: profile likelihood-based approaches can also be used to estimate the conditional mode for functional data in the SIM with missing data.

It is important to note that the choice of the estimation method may depend on the specific characteristics of the data and the nature of the missingness. Additionally, assessing the performance of these methods often involves simulations and comparisons on synthetic and real-world datasets with known missingness patterns.

The Single Index Model (SIM) is a financial modelling technique used to analyse the risk and return of a portfolio. It assumes that the returns of individual assets can be explained by their exposure to a common factor or market index. When dealing with missing data in the SIM framework, the missingness is assumed to be at random (MAR). This means that the missing values are related to the observed data but not to the missing values themselves. It is important to note that the choice of approach depends on the specifics of the data, the extent of missingness, and the assumptions one is willing to make. It is always recommended to carefully consider the nature of the data and consult with domain experts when handling missing data in the SIM or any other modelling framework.

The asymptotic properties of semi-parametric estimators of the conditional mode for functional data in the Single Index Model (SIM) with missing data at random (MAR) are an active area of research, and specific results may depend on the particular assumptions and estimation methods employed. However, this study can provide a general overview of some relevant concepts and approaches in this context. In the SIM framework, functional data refers to observations that are functions rather than scalar values. The goal is to estimate the conditional mode of a functional response variable given a set of functional predictors and a single index variable.

To establish the asymptotic properties of the semi-parametric estimators of the conditional mode for functional data in the SIM with missing data at random, various theoretical conditions need to be satisfied. These conditions often involve assumptions about the functional data, the missing data mechanism, and the model specification. Some common conditions include consistency and efficiency. Specific results in this area may depend on the assumptions and estimation techniques employed in each study. Therefore, one should refer to the literature and research articles that focus on the specific estimation method and the relevant assumptions to obtain more detailed and precise asymptotic properties of the estimators.

One of the most frequently encountered problems in non-parametric statistics is the question of forecasting. In some situations, regression is the main tool adapted to answer this question. However, in other situations, such as in the case where the conditional density is asymmetrical or multimodal, this tool is inadequate. Therefore, the conditional quantile better predicts the impact of the variable of interest Y on explanatory variable X. When the explanatory variable is infinite dimensional or it is of functional nature, only very few studies were reported to investigate the statistical properties of functional nonparametric regression model for missing data. Relatively recently, Ferraty, Sued and Vieu (2013) first proposed to estimate the mean of a scalar response based on an i.i.d. functional sample in which explanatory variables are observed for every subject, while part of the responses are missing at random (MAR) for some of them. This generalised the results obtained in Cheng (1994) to the case where the explanatory variables are of functional nature.

To the best of the authors' knowledge, the estimation of the nonparametric conditional distribution in the functional single index structure, combining missing data and stationary processes with functional nature, has not been studied in the statistical literature. Thus, in theed present work, the authors investigated conditional quantile estimation when the data are MAR. The aim was to develop a functional methodology for dealing with MAR samples in non-parametric problems (namely – in conditional quantile estimation). Then, the asymptotic properties of the estimator were obtained under some mild conditions, Hence the study considers a model in which the response variable is missing. Alongside the infinite dimensional character of the data, the authors avoided the strong mixing condition and its variants to measure the dependency and the very involved probabilistic calculations that it implies.

Therefore, the study considers, in this setting, the independent concept. As far as is known, the estimation of conditional quantile combining censored data, independent theory, and functional data with single-index structure has not been studied in the statistical literature. This work extends, to the functional single index model case, the work of Ling, Liang and Vieu (2015), Ling, Liu and Vieu (2016) and Rabhi, Kadiri and Mekki (2021). For the above mentioned theoretical and application reasons, the statistical community has shown great interest in estimating conditional quantiles, specifically the conditional median function, as an interesting alternative predictor to the conditional mean, thanks to its robustness to the presence of outliers (see Chaudhuri et al., 1997). The estimation of the conditional mode of a scalar response given a functional covariate has attracted the attention of many researchers. Ferraty, Rabhi and Vieu (2005) introduced a nonparametric estimator of the conditional quantile, defined as the inverse of the conditional distribution function, when data are dependent. Ezzahrioui and Ould-Saïd (2008) established the asymptotic normality of the kernel conditional mode estimator. In the censored scope, Ould-Saïd and Cai (2005) stated the uniform strong consistency with rates of the kernel estimator of the conditional mode function, and in this context the study refers to Lemdani, Ould-Saïd and Poulin (2009) for the estimation of conditional quantiles. Other authors were interested in the estimation of conditional models when the observations are censored or truncated (see e.g. Hamri et al., 2022; Liang and de Uña-Alvarez, 2010; Ould-Saïd and Djabrane, 2011; Ould-Saïd and Tatachak, 2011; Rabhi et al., 2021, etc.).

For instance, Aït-Saidi, Ferraty, Kassa and Vieu (2008) were interested in using SFIM to estimate the regression operator and suggested using a cross-validation procedure allowing the estimated unknown link function as well as the unknown functional index. Attaoui and Boudiaf (2014) and Attaoui and Ling (2016) studied, respectively, the estimation of the conditional density and the conditional cumulative distribution function based on a SFIM and assuming that the data satisfy a strong mixing condition. Kadiri, Bouchentouf and Rabhi (2018) examined the asymptotic properties of the kernel-type estimator of the conditional quantiles when the response is right-censored and the data is sampled from a strong mixing process.

The rest of the paper is arranged as follows: Section 2 presents the non-parametric estimator of the functional conditional model, when the data are MAR; Section 3 poses useful assumptions for this theoretical study, followed by the point-wise almost complete convergence, and the uniform almost-complete convergence of the kernel estimator for the models (with rates) is established in Section 4.

2. Model and Estimator

2.1. The Functional Nonparametric Framework

Consider a random pair (X, Y) where *Y* is valued in $\mathbb{R}$ and *X* is valued in some infinite dimensional Hilbertian space $\mathcal{H}$ with scalar product $<\cdot, \cdot>$ considering that, given that $(X_i, Y_i)_{i=1,\dots,n}$ is the statistical sample of pairs which are identically distributed like (X, Y), but not necessarily independent. Henceforward, *X* is called a functional random variable *f.r.v*. Let *x* be fixed in $\mathcal{H}$ and let $F(\theta, y, x)$ be the conditional cumulative distribution function *(cond-cdf)* of *T* given $< \vartheta, X > = < \vartheta, x >$ specifically:

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

In saying that, one is implicitly assuming the existence of a regular version of conditional distribution *Y* given $< \vartheta, X > = < \vartheta, x >$.

For the infinite dimensional purpose, the authors used the term *functional nonparametric*, where *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). The authors also used the term *operational statistics* since the target object to be estimated (the *cond-df* $f(\theta, ., x)$) can be viewed as a nonlinear operator.

2.2. The Estimators

In the case of complete data, the kernel estimator $\tilde{f}_n(\theta, ., x)$ of $f(\theta, ., x)$ is presented as follows:

$$\tilde{f}(\theta, t, x) = \frac{g_n^{-1} \sum_{i=1}^n K\left(h_n^{-1}(|< x - X_i, \theta >|)\right) H\left(g_n^{-1}(y - Y_i)\right)}{\sum_{i=1}^n K\left(h_n^{-1}(< x - X_i, \theta >)\right)}, \tag{2.1}$$

where K and H are kernel functions and h_n (resp. g_n) a sequence of positive real numbers. Note that using similar ideas, Roussas (1969) introduced some related estimates, but only in the special case when *X* is real, while Samanta (1989) produced an earlier asymptotic study.

Meanwhile, in an incomplete case with missing at random for the response variable, we examine $(X_i, Y_i, \delta_i)_{1 \leq i \leq n}$ where X_i is observed completely, and $\delta_i = 1$ if Y_i is observed and $\delta_i = 0$ otherwise. The Bernoulli random variable δ is defined by

$$\mathbb{P}(\delta = 1 | \langle X, \theta \rangle = \langle x, \theta \rangle, Y = y) = \mathbb{P}(\delta = 1 | \langle X, \theta \rangle = \langle x, \theta \rangle) = p(x, \theta),$$

where $p(x, \theta)$ is a functional operator which is conditional only on X. Therefore, the estimator of $f(\theta, y, x)$ in the single index model with response MAR is given by

$$\hat{f}(\theta, t, x) = \frac{g_n^{-1} \sum_{i=1}^n \delta_i K\left(h_n^{-1}(|< x - X_i, \theta >|)\right) H\left(g_n^{-1}(y - Y_i)\right)}{\sum_{i=1}^n \delta_i K\left(h_n^{-1}(< x - X_i, \theta >)\right)} = \frac{\hat{f}_N(\theta, y, x)}{\hat{f}_D(\theta, x)}, \tag{2.2}$$

where $K_i(\theta, x) := \mathrm{K}(h_n^{-1} |< x - X_i, \theta >|)$, $H_i(y) = H\left(g_n^{-1}(y - Y_i)\right)$,

$$\hat{f}_D(\theta, x) = \frac{\sum_{i=1}^n \delta_i K_i(\theta, x)}{n \mathbb{E}\left(K_1(\theta, x)\right)} \text{ and } \hat{f}_N(\theta, y, x) = \frac{\sum_{i=1}^n \delta_i K_i(\theta, x) H_i(y)}{n g_n \mathbb{E}\left(K_1(\theta, x)\right)}.$$

2.3. Assumptions on the Functional Variable

Let N_x be a fixed neighbourhood of *x* in $\mathcal{H}$ and let $B_\vartheta(x,h)$ be the ball of centre *x* and radius *h*, namely $B_\vartheta(x,h) = \{ \chi \in \mathcal{H}: 0 <| < x - \chi,\vartheta >| < h\}$, $d_\vartheta (x,X_i) = | < x - X_i,\vartheta >|$ denote a random variable such that its cumulative distribution function is given by $\phi_{\theta,x}(u) = \mathbb{P}(d_\theta(x,X_i) \leq u) = \mathbb{P}(X_i \in B_\theta(x,u))$.

Now, consider the following basic assumptions that are necessary in deriving the main result of this paper.

(H1) $\mathbb{P}(X \in B_\theta(x,h_n)) =: \phi_{\theta,x}(h_n) > 0; \phi_{\theta,x}(h_n) \longrightarrow 0$ as $h_n \rightarrow 0$.

3. Asymptotic Study

Here, the objective was to adapt these ideas to the framework of a functional explanatory variable, and to construct a kernel-type estimator of the conditional density function $f(\theta,y,x)$ adapted to MAR response samples. Thus, the aim was to establish a uniform almost complete convergence of kernel estimator $\hat{f}(\theta,y,x)$ when one considers a model in which the response variable is missing. The presented results are accompanied by the data on the rate of convergence. Therefore, *C* and C' denote generic strictly positive real constants, where h_n (resp. g_n) is a sequence which tends to 0 with n.

3.1. Uniform Almost Complete Convergence and Rate of Convergence

In this section, the authors adapted these ideas to the framework of a functional explanatory variable, to construct a kernel-type estimator of conditional density function $f(\theta,y,x)$ adapted to the MAR response samples. The objective was to establish an almost complete uniform convergence of kernel estimator $\hat{f}(\theta,y,x)$ when one considers a model in which the response variable is missing, which is standard extensions of the pointwise results. Clearly, achieving these outcomes necessitates more intricate technical advancements compared to those outlined in the context of standard results pertaining to almost complete pointwise convergence. The results presented are accompanied by the data on the rate of convergence. Thus, *C* and *C'* denote generic strictly positive real constants, and h_n (resp. g_n) is a sequence which tends to 0 with n.

To enhance the clarity of this concept, it was necessary to employ additional tools and consider certain topological conditions (see Hamri et al., 2022). Initially, due to the compactness o sets $S_\mathcal{H}$ and $\Theta_\mathcal{H}$, it was possible to cover them using a finite number of disjoint intervals. Let $d_n^{S_\mathcal{H}}$ and $d_n^{\Theta_\mathcal{H}}$denote the minimal numbers of open balls with radius r_n in $\mathcal{H}$ that are required to cover $S_\mathcal{H}$ and $\Theta_\mathcal{H}$, respectively. Within these intervals, points x_k (resp. t_j) $\in \mathcal{H}$.

$$S_\mathcal{H} \subset \bigcup_{k=1}^{d_n^{S_\mathcal{H}}} B_\theta\,(x_k,r_n)\ and\ \Theta_\mathcal{H} \subset \bigcup_{j=1}^{d_n^{\Theta_\mathcal{H}}} B_\theta\,(t_j,r_n).$$

3.2. Conditional Density Estimation

The objective of this part is to demonstrate almost complete uniform convergence. In order to extend the results to the uniform case, it was essential to introduce a topological framework for the functional space of the observations and the functional character of the proposed model. The asymptotic conclusions made use of the topological properties in the functional space of the study's observations. It is worth mentioning that all the convergence rates rely on the assumption of probability measure concentration of the functional variable within small balls, as well as the concept of Kolmogorov's entropy, which quantifies the number of balls required to cover a given set. To achieve this objective, the authors introduced the following conditions:

(H2) Kernel H is a positive bounded function such that

(i) $\forall(y_1,y_2)\in\mathbb{R}^2, |H(y_1)-H(y_2)|\le C|y_1-y_2|$, $\int|y|^{\alpha_2}H(y)dy<\infty$, $\int H^2(t)dt<\infty$.

(ii) $H^{(1)}$ and $H^{(2)}$ are bounded with $\int\left(H^{(1)}(t)\right)^2dt<\infty$.

(H3) K is a positive bounded kernel function with support [0,1]: $\forall u\in(0,1)$, $0<\mathrm{K}(u)$ and the Lipschitz condition holds $|K(x)-K(y)|\le\|x-y\|$, and derivative K' exists on [0,1] with $K'(t)<0$ for all $t\in[0,1]$ and $\int_0^1\left(K^j\right)'(t)dt<\infty$ for $j=1,2$.

(H4) There exists function $\phi(\cdot)$ that is differentiable, $\forall x\in S_{\mathcal{H}}$ and $\forall\theta\in\Theta_{\mathcal{H}}$,

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

(H5) Conditional density $f(\theta,y,x)$ satisfies the uniform Hölder condition, i.e. there exist some $\alpha_1,\alpha_2>0$ and $C>0$ such that for $\forall(x_1,x_2)\in S_{\mathcal{H}}\times S_{\mathcal{H}}$, $\forall(y_1,y_2)\in S_{\mathbb{R}}\times S_{\mathbb{R}}$ and $\forall\theta\in\Theta_{\mathcal{H}}$,

$$|f(\theta,y_1,x_1)-f(\theta,y_2,x_2)|\le C(\|x_1-x_2\|^{\alpha_1}+|y_1-y_2|^{\alpha_2}).$$

(H6) $p(x,\theta)$ is continuous in a neighbourhood of x, such that $0<p(x,\theta)<1$.

(H7) for some $\nu\in(0,1)$, $\lim_{n\to\infty}n^\nu g_n=\infty$, and for $r_n=\mathcal{O}\left(\frac{\log n}{n}\right)$, sequences $d_n^{S_{\mathcal{H}}}$ and $d_n^{\Theta_{\mathcal{H}}}$ satisfy:

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

In what follows, denote

$$\Upsilon_i(\theta,x)=\frac{1}{g_n\phi(h_n)}1_{B_\theta(x,h)\cup B_\theta(x_{k(x)},h)}(X_i),$$
$$\Omega_i(\theta,x)=\frac{1}{g_n\phi(h_n)}1_{B_\theta(x_{k(x)},h)\cup B_{t_{j(\theta)}}(x_{k(x)},h)}(X_i),$$
$$\Psi_i\left(t_{j(\theta)},x_{k(x)}\right)=\frac{\delta_iK\left(h_n^{-1}(<x_{k(x)}-X_i,t_{j(\theta)}>)\right)}{\mathbb{E}\left(K\left(h_n^{-1}(<x_{k(x)}-X_i,t_{j(\theta)}>)\right)\right)}.$$

$$\text{and } \Sigma_i(\theta,x)=\frac{1}{g_n}\frac{\delta_iK\left(h_n^{-1}(\langle t_{j(\theta)},x_{k(x)}-X_i\rangle)\right)}{\mathbb{E}\left(K\left(h_n^{-1}(\langle t_{j(\theta)},x_{k(x)}-X_i\rangle)\right)\right)}H\left(g_n^{-1}(y_{k(y)}-Y_i)\right)$$
$$-\frac{1}{g_n}\mathbb{E}\left(\frac{\delta_iK\left(h_n^{-1}(\langle t_{j(\theta)},x_{k(x)}-X_i\rangle)\right)}{\mathbb{E}\left(K\left(h_n^{-1}(\langle t_{j(\theta)},x_{k(x)}-X_i\rangle)\right)\right)}H\left(g_n^{-1}(y_{k(y)}-Y_i)\right)\right).$$

Theorem 3.1. Assuming hypotheses (H1) to (H7), if $\exists\beta > 0, n^{\beta} g_n^2 \underset{n\to\infty}{\longrightarrow} \infty$, and if $\lim_{n\to\infty} \frac{\log n}{n g_n^2 \phi_{\theta,x}(h_n)} = 0$, one has

$$\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{\theta y\in S_{\mathbb{R}}} \sup_{x\in S_{\mathcal{H}}} \left|\hat{f}(\theta,y,x) - f(\theta,y,x)\right| = \mathcal{O}\left(h_n^{\alpha_1} + g_n^{\alpha_2}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{n g_n^2 \phi(h_n)}}\right).$$

Proof. Clearly, the proof was constructed formed on the following decomposition and the following intermediate results, which holds true for any $y \in S_{\mathbb{R}}$, $\theta \in \Theta_{\mathcal{H}}$, and $x \in S_{\mathcal{H}}$:

$$\begin{aligned}
\Xi(\theta,y,x) &= \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{\theta y\in S_{\mathbb{R}}} \sup_{x\in S_{\mathcal{H}}} \left|\hat{f}(\theta,y,x) - f(\theta,y,x)\right| \\
&\le \frac{1}{\hat{f}_D(\theta,x)} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{\theta y\in S_{\mathbb{R}}} \sup_{x\in S_{\mathcal{H}}} \left|\hat{f}_N(\theta,y,x) - \mathbb{E}\hat{f}_N(\theta,y,x)\right| \\
&\quad + \frac{1}{\hat{f}_D(\theta,x)} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{\theta y\in S_{\mathbb{R}}} \sup_{x\in S_{\mathcal{H}}} \left|\mathbb{E}\hat{f}_N(\theta,y,x) - f(\theta,y,x)\right| \\
&\quad + \frac{f(\theta,y,x)}{\hat{f}_{\mathrm{D}}(\theta,x)} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \left|\hat{f}_{\mathrm{D}}(\theta,x) - \mathbb{E}\hat{f}_{\mathrm{D}}(\theta,x)\right|. \qquad (3.1)
\end{aligned}$$

Lemma 3.1. Under conditions (H1), (H2) and (H5), (H6), one has:

$$\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{\theta y\in S_{\mathbb{R}}} \sup_{x\in S_{\mathcal{H}}} \left|f(\theta,y,x) - \mathbb{E}\left[\hat{f}_N(\theta,y,x)\right]\right| = \mathcal{O}\left(h_n^{\alpha_1} + g_n^{\alpha_2}\right).$$

Proof. One has:

$$\begin{aligned}
I = \mathbb{E}\hat{f}_N(\theta,y,x) - f(\theta,y,x) &= \mathbb{E}\left(\frac{1}{n g_n \mathbb{E}\left(K_1(\theta,x)\right)} \sum_{i=1}^{n} \delta_i K_i(\theta,x) H_i(y)\right) - f(\theta,y,x) \\
&= \frac{1}{n g_n \mathbb{E}\left(K_1(\theta,x)\right)} \sum_{i=1}^{n} \mathbb{E}([\mathbb{E}(\delta_i K_i(\theta,x) H_i(y) | < \theta, X_i >)]) - f(\theta,y,x), \\
&= \frac{1}{g_n \mathbb{E}\left(K_1(\theta,x)\right)} \mathbb{E}\left(p(x,\theta) K_1(\theta,x) \mathbb{E}(H_1(y))\right) - f(\theta,y,x);
\end{aligned}$$

Moreover, by changing variables and using the fact that H is a *df* and using a double conditioning with respect to Y_1, one can easily obtain

$$\begin{aligned}
\mathbb{E}\left(H\left(g_n^{-1}(y - Y_1)\right) | < \theta, X_1 >\right) &= \int_{\mathbb{R}} H\left(\frac{y-u}{g_n}\right) f(\theta,u,X_1) du \\
&= \int_{\mathbb{R}} H(v) f(\theta, y - v g_n, X_1) dv \\
&= g_n \int_{\mathbb{R}} H(v)\left(f(\theta, y - v g_n, X_1) - f(\theta,u,x)\right) dv + g_n f(\theta,u,x) \int_{\mathbb{R}} H(v) dv
\end{aligned}$$

hence one can write, because of (H2) and (H5):

$$\begin{aligned}
I &= \frac{1}{\mathbb{E}K_1} \mathbb{E}\left(p(x,\theta) K_1(\theta,x) \int_{\mathbb{R}} H(v)\left(f(\theta, y - v g_n, X_1) - f(\theta,y,x)\right) dv\right) \\
&\le C_{\theta,x}(p(x,\theta) + o(1)) \int_{\mathbb{R}} H(v)\left(h_n^{\alpha_1} + |v|^{\alpha_2} g_n^{\alpha_2}\right) dv \le \mathcal{O}\left(h_n^{\alpha_1} + g_n^{\alpha_2}\right).
\end{aligned}$$

Finally, the proof is achieved.

Lemma 3.2. Under the assumptions of Theorem 3.1:

1. $\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in S_{\mathcal{H}}}\left|\hat{f}_D(\theta,x)-p(\theta,x)\right| = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_n)}}\right).$

2. $\sum_{n\geq 1}\mathbb{P}\left(\hat{f}_{\mathrm{D}}(\theta,x) < 1/2\right) < \infty.$

Proof. To demonstrate the first part of this lemma, and following a similar methodology as shown in the proof of Lemma 4.4 in (Kadiri et al., 2018), the proof can be readily completed. However, for the sake of brevity, the authors omitted the detailed proof in this context.

For the proof of the second part, one only needed to establish $\mathbb{E}\hat{f}_{\mathrm{D}}(\theta,x) \underset{n\to\infty}{\longrightarrow} p(x,\theta)\ a.co.$

By the properties of conditional expectation and the mechanism of MAR and (H6), it follows that:

$$\begin{aligned}\mathbb{E}\hat{f}_{\mathrm{D}}(\theta,x) &= \frac{1}{n\mathbb{E}\left(K_1(\theta,x)\right)}\sum_{i=1}^{n}\mathbb{E}\left(\delta_i K_i(\theta,x)\right)\\ &= \frac{1}{n\mathbb{E}\left(K_1(\theta,x)\right)}\sum_{i=1}^{n}\mathbb{E}[\mathbb{E}(\delta_i| < \theta, X_i >)K_i(\theta,x)]\\ &= \frac{(p(x,\theta)+o(1))}{n\mathbb{E}\left(K_1(\theta,x)\right)}\sum_{i=1}^{n}\mathbb{E}\left(K_i(\theta,x)\right)\underset{n\to\infty}{\longrightarrow} p(x,\theta)\ a.co.\end{aligned}$$

Therefore (2) of Lemma 3.4 follows from (1) and because $\hat{f}_{\mathrm{D}}(\theta,x) \underset{n\to\infty}{\longrightarrow} p(x,\theta)\ a.co.$

Concerning the last part, one has

$$\begin{aligned}\left\{\hat{f}_{\mathrm{D}}(\theta,x) < p(x,\theta)/2\right\} &\subseteq \left\{\left|\hat{f}_{\mathrm{D}}(\theta,x) - p(x,\theta)\right| > p(x,\theta)/2\right\} \Rightarrow\\ \mathbb{P}\left\{\hat{f}_{\mathrm{D}}(\theta,x) < p(x,\theta)/2\right\} &\leq \mathbb{P}\left\{\left|\hat{f}_{\mathrm{D}}(\theta,x) - p(x,\theta)\right| > p(x,\theta)/2\right\}\\ &\leq \mathbb{P}\left\{\left|\hat{f}_{\mathrm{D}}(\theta,x) - \mathbb{E}\hat{f}_{\mathrm{D}}(\theta,x)\right| > 1/2\right\},\end{aligned}$$

because $\lim_{n\to\infty}\hat{f}_{\mathrm{D}}(\theta,x) = p(x,\theta)$, hence

$$\sum_{n\geq 1}\mathbb{P}\left(\hat{f}_{\mathrm{D}}(\theta,x) < p(x,\theta)/2\right) < \infty.$$

Lemma 3.3. Considering the assumptions of Theorem 3.1:

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{\theta y\in S_{\mathbb{R}}}\sup_{x\in S_{\mathcal{H}}}\left|\hat{f}_N(\theta,y,x) - \mathbb{E}\left[\hat{f}_N(\theta,y,x)\right]\right| = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{n g_n^2\phi(h_n)}}\right).$$

Proof. For all $x \in S_{\mathcal{H}}$ and $\forall\theta \in \Theta_{\mathcal{H}}$, it was set

$$k(x) = \arg\min_{k\in\left\{1,\cdots,d_n^{S_{\mathcal{H}}}\right\}}\|x - x_k\|, j(\theta) = \arg\min_{\left\{1,\cdots,d_n^{\Theta_{\mathcal{H}}}\right\}}\left\|\theta - t_j\right\|$$

and by the compact property of $S_{\mathbb{R}} \subset \mathbb{R}$, one obtains $S_{\mathbb{R}} \subset \bigcup_{m=1}^{\tau_n}(y_m - l_n, y_m + l_n)$ with l_n and τ_n can be chosen such that $l_n = \mathcal{O}(\tau_n^{-1}) = \mathcal{O}\left(n^{-(3\nu+1)/2}\right)$. In the context of abstract semi-metric spaces, it is usually assumed that $d_n^{S_{\mathcal{H}}} r_n$ $(d_n^{\Theta_{\mathcal{H}}} r_n)$ is bounded. For more discussion, refer to Ferraty and Vieu (2006). Taking $k(y) = \arg\min_{k\in\{1,\cdots,\tau_n\}}|y - y_k|$. Let us consider the following decomposition

$$
\begin{aligned}
\widehat{\Lambda}_N(\theta,y,x) &= \sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\sup_{x\in S_{\mathcal{H}}}\left|\hat{f}_N(\theta,y,x)-\mathbb{E}\left(\hat{f}_N(\theta,y,x)\right)\right| \\
&\leq \sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\sup_{x\in S_{\mathcal{H}}}\left|\hat{f}_N(\theta,y,x)-\hat{f}_N(\theta,y,x_{k(x)})\right| \\
&+ \sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\sup_{x\in S_{\mathcal{H}}}\left|\hat{f}_N(\theta,y,x_{k(x)})-\hat{f}_N(t_{j(\theta)},y,x_{k(x)})\right| \\
&+ \sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\sup_{x\in S_{\mathcal{H}}}\left|\hat{f}_N(t_{j(\theta)},y,x_{k(x)})-\hat{f}_N(t_{j(\theta)},y_{m(y)},x_{k(x)})\right| \\
&+ \sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\sup_{x\in S_{\mathcal{H}}}\left|\hat{f}_N(t_{j(\theta)},y_{m(y)},x_{k(x)})-\mathbb{E}\left(\hat{f}_N(t_{j(\theta)},y_{m(y)},x_{k(x)})\right)\right| \\
&+ \sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\sup_{x\in S_{\mathcal{H}}}\left|\mathbb{E}\left(\hat{f}_N(t_{j(\theta)},y_{m(y)},x_{k(x)})\right)-\mathbb{E}\left(\hat{f}_N(t_{j(\theta)},y,x_{k(x)})\right)\right| \\
&+ \sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\sup_{x\in S_{\mathcal{H}}}\left|\mathbb{E}\left(\hat{f}_N(t_{j(\theta)},y,x_{k(x)})\right)-\mathbb{E}\left(\hat{f}_N(\theta,y,x_{k(x)})\right)\right| \\
&+ \sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\sup_{x\in S_{\mathcal{H}}}\left|\mathbb{E}\left(\hat{f}_N(\theta,y,x_{k(x)})\right)-\mathbb{E}\left(\hat{f}_N(\theta,y,x)\right)\right| \\
&\leq D_1+D_2+D_3+D_4+D_5+D_6+D_7.
\end{aligned}
$$

- Concerning D_3 and D_5 by satisfying conditions (H2)-(ii) and (H7), as well as the boundedness of K, one obtains

$$
\begin{aligned}
\left|\hat{f}_N(t_{j(\theta)},y,x_{k(x)})-\hat{f}_N(t_{j(\theta)},y_{k(y)},x_{k(x)})\right| &\leq \frac{1}{ng_n\mathbb{E}(K_1(\theta,x))}\sup_{t\in S_{\mathbb{R}}}\sum_{i=1}^{n}\left|\delta_iK_i(t_{j(\theta)},x_{k(x)})\right| \\
&+\left|H(g_n^{-1}(y-Y_i))\delta_iH\left(g_n^{-1}(y_{m(y)}-Y_i)\right)\right| \\
\leq \sup_{y\in S_{\mathbb{R}}}C\frac{|y-y_{m(y)}|}{g_n^2}\left(\frac{\sum_{i=1}^n\left|\delta_iK_i(t_{j(\theta)},x_{k(x)})\right|}{n\mathbb{E}\left(K_1(t_{j(\theta)},x_{k(x)})\right)}\right) &\leq \frac{Cl_n}{g_n^2\phi(h_n)}=\mathcal{O}\left(\frac{l_n}{g_n^2\phi(h_n)}\right).
\end{aligned}
$$

Now, the fact that $\lim_{n\to\infty} n^{\nu}g_n^2=\infty$, and choosing $l_n=n^{-(3\nu+1)/2}$ and employing the second part of (H7), as $n\to\infty$, it follows that

$$
\frac{l_n}{g_n^2\phi(h_n)}=o\left(\sqrt{\frac{\log n}{ng_n^2\phi(h_n)}}\right),\ \ D_5\leq D_3=\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{ng_n^2\phi(h_n)}}\right).
$$

Concerning D_4 let us consider $\varepsilon=\varepsilon_0\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{ng_n^2\phi(h_n)}}$. Since

$$
\begin{aligned}
\mathbb{P}\left(D_4>\varepsilon_0\sqrt{\frac{\log d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{ng_n^2\phi(h_n)}}\right) &\leq \mathbb{P}\left(\max_{j\in\{1,\cdots,d_n^{\Theta_{\mathcal{H}}}\}}\max_{k\in\{1,\cdots,d_n^{S_{\mathcal{H}}}\}}\max_{m\in\{1,\cdots,\tau_n\}}|\Sigma_i-\mathbb{E}\Sigma_i|>\varepsilon\right) \\
&\leq \tau_n d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}\ \mathbb{P}(|\Sigma_i-\mathbb{E}\Sigma_i|>\varepsilon).
\end{aligned}
$$

Applying Bernstein's exponential inequality, under (H2) and (H5), to get

$$
\forall j\leq d_n^{\Theta_{\mathcal{H}}},\ \forall k\leq d_n^{S_{\mathcal{H}}} \text{ and } \forall m\leq\tau_n,
$$

$$
\mathbb{P}(|\Sigma_i-\mathbb{E}\Sigma_i|>\varepsilon)\leq 2\left(d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}\right)^{-C\varepsilon_0^2}.
$$

Choosing $\tau_n\leq Cn^{(3\nu+1)/2}$, one obtains

$$
\mathbb{P}(D_4>\varepsilon)\leq C\left(d_n^{S_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}\right)^{1-C\varepsilon_0^2}.
$$

Putting $C\varepsilon_0^2 = \xi$ and using (H7), one obtains

$$D_4 = \mathcal{O}_{a.co.}\left(\sqrt{\frac{log\, d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{ng_n^2\phi(h_n)}}\right). \tag{3.2}$$

- Concerning D_1 and D_2

$$\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{y\in S_{\mathbb{R}}} \sup_{x\in S_{\mathcal{H}}} \left|\hat{f}_N(\theta,y,x) - \hat{f}_N(\theta,y,x_{k(x)})\right| \leq$$

$$\frac{1}{ng_n\mathbb{E}(K_1(\theta,x))} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \sup_{y\in S_{\mathbb{R}}} \sum_{i=1}^{n} \left|\delta_i\left(K_i(\theta,x) - K_i(\theta,x_{k(x)})\right)\right| |H_i(y)|$$

$$\leq \frac{C}{ng_n\phi(h_n)} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \sum_{i=1}^{n} \left|\Psi_i(\theta,x) - \Psi_i(\theta,x_{k(x)})\right|$$

$$\leq \frac{1}{g_n\phi(h_n)} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \frac{1}{n}\sum_{i=1}^{n} 1_{B_\theta(x,h)\cup B_\theta(x_{k(x)},h)}(X_i)$$

$$\leq \frac{C}{g_n} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \frac{1}{n}\sum_{i=1}^{n} \Upsilon_i(\theta,x)\,.$$

Therefore, similarly to the arguments for (3.2), one can obtain

$$D_1 = \mathcal{O}_{a.co.}\left(\sqrt{\frac{log\, d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{ng_n^2\phi(h_n)}}\right).$$

$$\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \sup_{y\in S_{\mathbb{R}}} \left|\hat{f}_N(\theta,y,x_{k(x)}) - \hat{f}_N(t_{j(\theta)},y,x_{k(x)})\right| \leq$$

$$\frac{g_n^{-1}}{n\mathbb{E}(K_1(\theta,x))} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \sup_{y\in S_{\mathbb{R}}} \sum_{k=1}^{n} \left|\delta_i\left(K_i(\theta,x_{k(x)}) - K_i(t_{j(\theta)},x_{k(x)})\right)\right| |H_i(y)|$$

$$\leq \frac{C'g_n^{-1}}{\phi(h_n)} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \frac{1}{n}\sum_{i=1}^{n} \left|\Psi_i(\theta,x_{k(x)}) - \Psi_i(t_{j(\theta)},x_{k(x)})\right|$$

$$\leq \frac{C'}{g_n\phi(h_n)} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \frac{1}{n}\sum_{i=1}^{n} 1_{B_\theta(x_{k(x)},h)\cup B_{t_{j(\theta)}}(x_{k(x)},h)}(X_i)$$

$$\frac{C'}{g_n} \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in S_{\mathcal{H}}} \frac{1}{n}\sum_{i=1}^{n} \Omega_i(\theta,x)\,.$$

Similarly to the deductions of (3.2), this results in

$$D_2 = \mathcal{O}_{a.co.}\left(\sqrt{\frac{log\, d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{ng_n^2\phi(h_n)}}\right).$$

On the other hand, since $D_7 \leq D_1$ and $D_6 \leq D_2$, it also leads to

$$D_6 = \mathcal{O}_{a.co.}\left(\sqrt{\frac{log\, d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{ng_n^2\phi(h_n)}}\right) \text{ and } D_7 = \mathcal{O}_{a.co.}\left(\sqrt{\frac{log\, d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{ng_n^2\phi(h_n)}}\right).$$

Then the proof of Lemma 3.3 can be concluded.

The authors concluded the proof of the Theorem 3.1 by making use of the inequality (3.1), in conjunction with Lemma 3.1 to Lemma 3.3.

3.3. Conditional Mode Estimation

In this part, the study rated convergence of the conditional mode estimator $\widehat{M}_\theta(x)$. Obviously, obtaining these results required more sophisticated technical developments than those presented previously. To ensure greater clarity, the authors introduced conditions related to the flatness of the *cond-df* $f(\theta,.,x)$ around the conditional quantile $M_\theta(x)$.

Then a natural estimator of conditional mode $M_\theta(x)$ was defined as,

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

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

However, a complementary way to take this local shape constraint into account was to suppose that:

(H8) Conditional density $f(\theta,.,x)$ satisfies:

(i) $\exists\epsilon_0$, such that $f(\theta,.,x)$ is strictly increasing on $\left(M_\theta(x)-\epsilon_0, M_\theta(x)\right)$ and strictly decreasing on $(M_\theta(x), M_\theta(x)+\epsilon_0)$, with respect to x.

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

(H9) $\forall\varepsilon > 0$, $\exists\eta > 0$, $\forall\varphi$

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

The difficulty of the problem is naturally linked with the flatness of the function $f(\theta,y,x)$ around mode M_θ. This flatness can be controlled by the number of vanishing derivatives at point M_θ, and this parameter would also have a great influence on the asymptotic rates of the estimates; more precisely, the following additional smoothness condition was introduced.

(H10) There exists a certain integer $j > 1$ such that $\forall x$, and function $f(\theta,.,x)$ is j-times continuously differentiable w.r.t y on $S_\mathbb{R}$ with

$$\begin{cases} f^{(j)}(\theta, M_\theta(x), x) = 0, \quad if; 1 \leq j < l \\ f^{(j)}(\theta,.,x) \text{ is uniformly continuous on } S_\mathbb{R} \\ \text{such that} \quad f^{(j)}(\theta, M_\theta(x), x) \neq 0. \end{cases}$$

Proposition 3.1. Under the assumptions of Theorem 3.1, one has

$$\sup_{\theta\in\Theta_\mathcal{H}} \sup_{y\in\mathcal{S}_\mathbb{R}} \sup_{x\in S_\mathcal{H}} \left|\widehat{M}_\theta(\gamma,x) - M_\theta(\gamma,x)\right| = \mathcal{O}\left(\left(h_n^{\alpha_1} + g_n^{\alpha_2}\right)^{\frac{1}{j}}\right) + \mathcal{O}_{a.co.}\left(\left(\frac{\log d_n^{S_\mathcal{H}} d_n^{\Theta_\mathcal{H}}}{n g_n^2 \phi_{\theta,x}(h_n)}\right)^{\frac{1}{2j}}\right).$$

Proof. The proof is based on the Taylor expansion of $f(\theta,.,x)$ in the neighbourhood of $M_\theta(\gamma,x)$, obtaining

$$\hat{f}\left(\theta, \widehat{M}_\theta(x), x\right) = f(\theta, M_\theta(x), x) + \frac{f^{(j)}(\theta, M_\theta^*(x), x)}{j!}\left(\widehat{M}_\theta(x) - M_\theta(x)\right)^j,$$

where $M_\theta^*(x)$ is between $M_\theta(x)$ and $\widehat{M}_\theta(x)$, combining the last equality with the fact that

$$\left|\hat{f}\left(\theta,\widehat{M}_\theta(x),x\right)-\ f(\theta,M_\theta(x),x)\right|\le 2\sup_{y\in S_{\mathbb{R}}}\left|\ \hat{f}(\theta,y,x)-\ f(\theta,y,x)\right|,$$

allowing to write:

$$\left|M_\theta(x)-\widehat{M}_\theta(x)\right|^j\le\frac{j!}{f^{(j)}\left(\theta,M_\theta^*(x),x\right)}\sup_{y\in S_{\mathbb{R}}}\left|\ \hat{f}(\theta,y,x)-\ f(\theta,y,x)\right|.$$

Using the second part of (H8), one obtains

$$\exists\delta>0,\sum_{n\ge1}\mathbb{P}\left(f^{(j)}(\theta,M_\theta^*(x),x)\ge\delta\right)<\infty.$$

Thus, one would obtain

$$\left|\widehat{M}_\theta(\gamma,x)-M_\theta(\gamma,x)\right|^j=\mathcal{O}_{a.co.}\left(\sup_{y\in S_{\mathbb{R}}}\left|\ \hat{f}(\theta,y,x)-\ f(\theta,y,x)\right|\right).$$

Finally, Proposition 3.1 can be deduced from Theorem 3.1.

Corollary 3.1. Under the hypotheses of Theorem 3.1, one has

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{y\in S_{\mathbb{R}}}\sup_{x\in S_{\mathcal{H}}}\left|\widehat{M}_\theta(x)-M_\theta(x)\right|\underset{n\to\infty}{\longrightarrow}0,\qquad a.co.$$

Proof. The proof was based on the point-wise convergence of $\hat{f}(\theta,.,x)$ and the Lipschitz property introduced in (H2)-(i) and hypothesis (H9), where $f(\theta,t,x)$ is a continuous function. Thus:

$\forall\,\epsilon>0,\exists\eta(\epsilon)>0$, such that

$$\left|f(\theta,y,x)-\ \hat{f}(\theta,M_\theta(x),x)\right|\le\ \eta(\epsilon)\ \Rightarrow|y\ -\ M_\theta(x)|\le\epsilon.$$

Therefore, for $y=\widehat{M}_\theta(x)$,

$$\mathbb{P}\left(\left|\widehat{M}_\theta(x)-M_\theta(x)\right|>\ \epsilon\right)\le\mathbb{P}\left(\left|f\left(\theta,\widehat{M}_\theta(x),x\right)-f(\theta,M_\theta(x),x)\right|\ge\ \eta(\epsilon)\right).\qquad(3.2)$$

Then, according to theorem, $\widehat{M}_\theta-M_\theta$ go almost completely to 0, as *n* goes to infinity.

4. Asymptotic Normality

The asymptotic normality of the semi-parametric estimators of the conditional mode for functional data in the Single Index Model (SIM) with missing data at random (MAR) is an important property that establishes the limiting distribution of the estimators as the sample size increases. Although specific results may vary depending on the assumptions and estimation methods used, it allows to construct confidence intervals and hypothesis tests for the estimated mode. In this section, the asymptotic normality of estimator $\hat{f}(\theta,.,x)$ in the single functional index model was established.

(N1) There exists function $\beta_{\vartheta,x}(\cdot)$ such that $\lim_{n\to\infty}\frac{\phi_{\theta,x}(sh_n)}{\phi_{\theta,x}(h_n)}=\beta_{\theta,x}(s)$, for $\forall s\in[0,1]$.

(N2) Bandwidth h_n and g_n, small ball probability $\phi_{\theta,x}(h_n)$ satisfying

(i) $ng_n^3\,\phi_{\theta,x}^3(h_n)\longrightarrow\ 0$ and $\frac{ng_n^3\,\phi_{\theta,x}(h_n)\log n}{\log^2 n}\longrightarrow\infty,\,as\ n\to\infty.$
(ii) $ng_n^2\phi_{\theta,x}^3(h_n)\longrightarrow\ 0,as\ n\to\infty.$

(N3) The conditional density *f(ϑ,y,x)* satisfies: $\exists\alpha>0,\ \forall(y1,y2)\ \in S_{\mathbb{R}}\times S_{\mathbb{R}}$,

$$\left|f^{(j)}(\theta,y_1,x_1)-f^{(j)}(\theta,y_2,x_2)\right|\le C(|y_1-y_2|^\alpha),\qquad j=1,2.$$

Theorem 4.1 Under the assumptions of Theorem 3.1 and (N1)-(N3) for all $x \in \mathcal{H}$, and in addition if

$$\sqrt{ng_n\phi_{\theta,x}(h_n)}\left(h_n^{\alpha_1}+g_n^{\alpha_2}\right)\underset{n\to\infty}{\longrightarrow}0,$$

then one has

$$\sqrt{\frac{ng_n\phi_{\theta,x}(h_n)}{\sigma^2(\theta,y,x)}}\left(\hat{f}(\theta,y,x)-f(\theta,y,x)\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,1),$$

where $\sigma^2(\theta,y,x)=\frac{M_2(\theta,x)}{(M_1(\theta,x))^2}\frac{f(\theta,y,x)}{p(\theta,x)}\int H^2(u)du$ with $M_l(\theta,x)=K^l(1)-\int_0^1\left(K^l\right)'(u)\,\beta_{\theta,x}(u)du$, $l=1,2$, $\xrightarrow{\mathcal{D}}$ means the convergence in distribution.

Proof.

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

$$\hat{f}(\theta,y,x)-f(\theta,y,x)=\frac{\hat{f}_N(\theta,y,x)}{\hat{f}_{\mathrm{D}}(\theta,x)}-\frac{M_1(\theta,x)f(\theta,y,x)}{M_1(\theta,x)}=\frac{1}{\hat{f}_{\mathrm{D}}(\theta,x)}\left(\hat{f}_N(\theta,y,x)-\mathbb{E}\hat{f}_N(\theta,y,x)\right)$$
$$-\frac{1}{\hat{f}_{\mathrm{D}}(\theta,x)}\left(M_1(\theta,x)f(\theta,y,x)-\mathbb{E}\hat{f}_N(\theta,y,x)\right)+\frac{f(\theta,y,x)}{\hat{f}_{\mathrm{D}}(\theta,x)}\left(M_1(\theta,x)-\mathbb{E}\hat{F}_D(\theta,x)\right)$$
$$-\frac{f(\theta,y,x)}{\hat{f}_{\mathrm{D}}(\theta,x)}\left(\hat{F}_D(\theta,x)-\mathbb{E}\hat{F}_D(\theta,x)\right)=\frac{1}{\hat{f}_{\mathrm{D}}(\theta,x)}A_n(\theta,y,x)+B_n(\theta,y,x),$$

where

$$A_n(\theta,y,x)=\frac{1}{ng_n\mathbb{E}\left(K_1(\theta,x)\right)}\sum_{i=1}^{n}\left\{\left(H_i(t)-g_nf(\theta,y,x)\right)\delta_iK_i(\theta,x)\right.$$
$$\left.-\mathbb{E}\left[\left(H_i(t)-g_nf(\theta,y,x)\right)\delta_iK_i(\theta,x)\right]\right\}=\frac{1}{ng_n\mathbb{E}\left(K_1(\theta,x)\right)}\sum_{i=1}^{n}N_i\,(\theta,t,x)$$

and $B_n(\theta,y,x)=M_1(\theta,x)f(\theta,y,x)-\mathbb{E}\hat{f}_N(\theta,y,x)+f(\theta,y,x)\left(M_1(\theta,x)-\mathbb{E}\hat{f}_D(\theta,x)\right)$.

It follows that,

$$ng_n\phi_{\theta,x}(h_n)Var\left(A_n(\theta,y,x)\right)=\frac{\phi_{\theta,x}(h_n)}{g_n\mathbb{E}^2\left(K_1(\theta,x)\right)}Var\left(N_1(\theta,y,x)\right)=V_n(\theta,y,x).$$

Then, the proof of Theorem 4.1 can be deduced from the following lemmas.

Lemma 4.1. Under the assumptions of Theorem 4.1, one has

$$\sqrt{ng_n\phi_{\theta,x}(h_n)}A_n(\theta,y,x)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\sigma^2(\theta,y,x)\right).$$

Proof.

$$\begin{aligned}V_n(\theta,y,x)&=\frac{\phi_{\theta,x}(h_n)}{g_n\mathbb{E}^2\left(K_1(\theta,x)\right)}\mathbb{E}\left[\delta_1K_1^2(\theta,x)\left(H_1(y)-g_nf(\theta,y,x)\right)^2\right]\\&=\frac{\phi_{\theta,x}(h_n)}{g_n\mathbb{E}^2\left(K_1(\theta,x)\right)}\mathbb{E}\left[K_1^2(\theta,x)\mathbb{E}\left(\delta_1\left(H_1(y)-g_nf(\theta,y,x)\right)^2\Big|\langle\theta,X_1\rangle\right)\right].\end{aligned}\quad(4.1)$$

Using the definition of conditional variance, one has

$$\mathbb{E}\left(\delta_1\left(H_1(t)-g_nf(\theta,y,x)\right)^2\Big|\langle\theta,X_1\rangle\right)=J_{1n}+J_{2n},$$

where $J_{1n}=Var(\delta_1 H_1(y)|\langle\theta,X_1\rangle)$, $J_{2n}=[\mathbb{E}(H_1(y)|\langle\theta,X_1\rangle)-g_nf(\theta,y,x)]^2$.

- Concerning J_{1n},

$$J_{1n}=\mathbb{E}\left(H^2\left(\frac{y-Y_1}{g_n}\right)\Big|\langle\theta,X_1\rangle\right)-\left[\mathbb{E}\left(\delta_1\,H_1\left(\frac{y-Y_1}{g_n}\right)\Big|\langle\theta,X_1\rangle\right)\right]^2=\mathcal{J}_1+\mathcal{J}_2.$$

As for $\mathcal{J}_1$, by the property of double conditional expectation, one obtains that

$$\begin{aligned}\mathcal{J}_1=\mathbb{E}\left(\delta_1 H^2\left(\frac{y-Y_1}{g_n}\right)\Big|\langle\theta,X_1\rangle\right)&=p(x,\theta)\int H^2\left(\frac{y-v}{g_n}\right)f(\theta,v,X_1)dv\\ &=p(x,\theta)\int H^2(u)\,dF(\theta,y-ug_n,X_1)\end{aligned}.$$

On the other hand, under assumption (H2) and (H3)

$$\begin{aligned}\mathcal{J}_1&=\int H^2(u)\,dF(\theta,y-ug_n,X_1)=h_n\int H^2(u)\,f(\theta,y-ug_n,X_1)du\\ &\le g_n\int H^2(u)\big(f(\theta,y-ug_n,X_1)-f(\theta,y,x)\big)du\;+g_n\int H^2(u)\,f(\theta,y,x)du\\ &\le g_n\left(C_{\theta,x}\int H^2(u)\big(h_n^{\alpha_1}+|v|^{\alpha_2}g_n^{\alpha_2}\big)du+f(\theta,y,x)\int H^2(u)\,du\right)\\ &=O\big(h_n^{\alpha_1}+g_n^{\alpha_2}\big)+g_nf(\theta,y,x)\int H^2(u)\,du\end{aligned}. \qquad (4.2)$$

As for J_2, $\mathcal{J}_2'=\mathbb{E}(\delta_1 H_1(y)|\langle\theta,X_1\rangle)=p(x,\theta)\int H\left(\frac{y-v}{g_n}\right)f(\theta,y,X_1)dv$.

Moreover, by changing the variables, one has

$$\mathcal{J}'_2=h_n\int H(u)\big(f(\theta,y-ug_n,x)-f(\theta,y,x)\big)du+g_nf(\theta,y,x)\int H(u)du,$$

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

$$\mathcal{J}'_2=O\big(h_n^{\alpha_1}+g_n^{\alpha_2}\big)+g_nf(\theta,y,x).$$

Finally, one obtains $J_2\underset{n\to\infty}{\longrightarrow}\infty$. As for J_{2n}, by (H1),(H2) and (H5), one obtains $J_{2n}\underset{n\to\infty}{\longrightarrow}\infty$.

Meanwhile, by (H1)-(H2) and (H5), it follows that

$$\frac{\phi_{\theta,x}(h_n)\mathbb{E}K_1^2(\theta,x)}{\mathbb{E}^2\big(K_1(\theta,x)\big)}\underset{n\to\infty}{\longrightarrow}\frac{M_2(\theta,x)}{(M_1(\theta,x))^2},$$

which leads to combining equations (4.1) and (4.2)

$$V_n(\theta,t,x)\underset{n\to\infty}{\longrightarrow}\frac{M_2(\theta,x)f(\theta,y,x)}{(M_1(\theta,x))^2p(x,\theta)}.$$

Lemma 4.2. If the assumptions (H1) to (H9) are satisfied, one has

$$\sqrt{ng_n\phi_{\theta,x}(h_n)}B_n(\theta,t,x)\to 0, \text{ in probability.}$$

Proof.

One has

$$\sqrt{ng_n\phi_{\theta,x}(h_n)}B_n(\theta,t,x)=\frac{\sqrt{ng_n\phi_{\theta,x}(h_n)}}{\hat{f}_D(\theta,x)}\{\mathbb{E}\hat{f}_N(\theta,y,x)-M_1(\theta,x)f(\theta,y,x)$$
$$+f(\theta,y,x)\left(M_1(\theta,x)-\mathbb{E}\hat{f}_D(\theta,x)\right).$$

Firstly, observe that the results below as $n\to\infty$

$$\frac{1}{\phi_{\theta,x}(h_n)}\mathbb{E}\left[K^l\left(\frac{\langle\theta,x-X_i\rangle}{h_n}\right)\right]\to M_l(\theta,x),\text{for } l=1,2, \tag{4.3}$$

$$\mathbb{E}\hat{f}_D(\theta,x)\to M_1(\theta,x)p(x,\theta)\ \ and\ \ \mathbb{E}\hat{f}_N(\theta,y,x)\to M_1(\theta,x)f(\theta,y,x), \tag{4.4}$$

can be proved in the same way as in (Ezzahrioui and Ould-Saïd, 2008) corresponding to their Lemmas 5.1 and 5.2, and then their proofs are omitted.

Secondly, on the one hand, making use of (4.3) and (4.4), one has as $n\to\infty$

$$\left\{\mathbb{E}\hat{f}_N(\theta,y,x)-M_1(\theta,x)f(\theta,y,x)+f(\theta,y,x)\left(M_1(\theta,x)-\mathbb{E}\hat{f}_D(\theta,x)\right)\right\}\to 0.$$

On the other hand,

$$\frac{\sqrt{ng_n\phi_{\theta,x}(h_n)}}{\hat{f}_D(\theta,x)}=\frac{\sqrt{ng_n\phi_{\theta,x}(h_n)}\hat{f}(\theta,y,x)}{\hat{f}_D(\theta,x)\hat{f}(\theta,y,x)}=\frac{\sqrt{ng_n\phi_{\theta,x}(h_n)}\hat{f}(\theta,y,x)}{\hat{f}_N(\theta,y,x)}.$$

As K and H are continuous with support on [0,1], then by (H2) and (H3) $\exists\, m=\min_{[0,1]}K(t)H(t)$ it follows that

$$\hat{f}_N(\theta,y,x)\geq\frac{m}{g_n\phi_{\theta,x}(h_n)},$$

which gives

$$\frac{\sqrt{ng_n\phi_{\theta,x}(h_n)}}{\hat{f}_N(\theta,y,x)}\leq\frac{\sqrt{ng_n^3\,\phi_{\theta,x}^3(h_n)}}{m}.$$

Finally, using (N2)-(i), completes the proof of Lemma 4.2.

4.1. Application: The Conditional Mode in Functional Single-Index Model

The main objective of this part was to establish the asymptotic normality of the conditional mode estimator of Y given $<\theta,X>=<\theta,x>$ denoted by $M_\theta(x)$.

Corollary 4.1 Under the assumptions of Theorem 4.1, and if (H8) holds true, and in addition if

$$ng_n^3\phi_{\theta,x}(h_n)\underset{n\to\infty}{\longrightarrow}0,$$

then with $n\to\infty$,

$$\sqrt{ng_n^3\phi_{\theta,x}(h_n)}\left(\widehat{M}_\theta(x)-M_\theta(x)\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\varrho^2(\theta,M_\theta(x),x)\right),$$

where

$$\varrho^2(\theta,M_\theta(x),x)=\frac{M_2(\theta,x)f(\theta,M_\theta(x),x)}{p(\theta,x)\left(M_1(\theta,x)f^{(2)}(\theta,M_\theta(x),x)\right)^2}\int\left(H'(u)\right)^2du.$$

Proof.

By the first order Taylor expansion for $\hat{f}^{(1)}(\theta, y, x)$ at point $M_\theta(x)$, and the fact that $\hat{f}^{(1)}\left(\theta, \widehat{M}_\theta(x), x\right) = 0$, it follows that

$$\sqrt{n g_n^3 \phi_{\theta,x}(h_n)}\left(\widehat{M}_\theta(x) - M_\theta(x)\right) = -\sqrt{n g_n^3 \phi_{\theta,x}(h_n)}\frac{\hat{f}^{(1)}(\theta, M_\theta(x), x)}{\hat{f}^{(2)}\left(\theta, M_\theta^*(x), x\right)},$$

where $M_\theta^*(x)$ is between $\widehat{M}_\theta(x)$ and $M_\theta(x)$. Similarly to the proof of Theorem 4.1, it follows that

$$-\sqrt{n g_n^3 \phi_{\theta,x}(h_n)}\hat{f}^{(1)}(\theta, M_\theta(x), x) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \varrho_0^2(\theta, M_\theta(x), x)\right), \tag{4.4}$$

where

$$\varrho_0^2(\theta, M_\theta(x), x) = \frac{M_2(\theta, x)}{(M_1(\theta, x))^2}\frac{f(\theta, M_\theta(x), x)}{p(\theta, x)}\int \left(H'(u)\right)^2 du.$$

Thus, as above, similarly to Ferraty and Vieu (2006), one can obtain $\hat{f}^{(2)}(\theta, y, x) \xrightarrow{\mathbb{P}} f^{(2)}(\theta, y, x)$, as $n \to \infty$, which implies that $\widehat{M}_\theta(x) \to M_\theta(x)$. Therefore

$$\hat{f}^{(2)}(\theta, M_\theta^*(x), x) \xrightarrow[n\to\infty]{} f^{(2)}(\theta, M_\theta(x), x) \neq 0. \tag{4.5}$$

By (H2), (H8) and (N3), similarly to the proof of Lemma 4.1 and Lemma 4.2, respectively, (4.4) follows directly. Then, the proof of Corollary 4.1 is completed.

4.2. Application: The Conditional Mode in Functional Single-Index Model

The asymptotic variances $\sigma^2(\theta, t, x)$ and $\varrho^2(\theta, M_\theta(x), x)$ in Theorem 4.1 and Corollary 4.1 depend on some unknown quantities including M_1, M_2, $\phi(u)$, $M_\theta(x)$, $p(\theta, x)$ and $f(\theta, M_\theta(x), x)$. Therefore, $p(\theta, x)$, $M_\theta(x)$, and $f(\theta, M_\theta(x), x)$ can be estimated respectively by $P_n(\theta, x)$, $\widehat{M}_\theta(x)$ and $\hat{f}(\theta, M_\theta(x), x)$ and $\widehat{M}_\theta(x)$. Moreover, using the decomposition given by assumption (H1), one can estimate $\phi_{\theta,x}(.)$ by $\hat{\phi}_{\theta,x}(.)$, with the unknown functions $M_j := M_j(\theta, x)$ and $f(\theta, y, x)$ intervening in the expression of the variance. Therefore, it was necessary to estimate the mode $M_1(\theta, x)$, $M_2(\theta, x)$ and $f(\theta, y, x)$, respectively.

By assumptions (H1) to (H4), one knows that $M_j(\theta, x)$ can be estimated by $\widehat{M}_j(\theta, x)$ which is defined as

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

with $\mathbf{1}_{\{.\}}$ being the indicator function. Finally, the estimator of $p(\theta, x)$ is denoted by

$$P_n(\theta, x) = \frac{\sum_{i=1}^{n} \delta_i K\left(h_n^{-1}(<x - X_i, \theta>)\right)}{\sum_{i=1}^{n} K\left(h_n^{-1}(<x - X_i, \theta>)\right)}.$$

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

$$\hat{\sigma}^2(\theta, y, x) = \frac{\widehat{M}_2(\theta, x)}{\left(\widehat{M}_1(\theta, x)\right)^2}\frac{\hat{f}(\theta, y, x)}{P_n(\theta, x)}\int H^2(u) du.$$

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

$$\hat{\varrho}^2(\theta, M_\theta(x), x) = \frac{\widehat{M}_2(\theta,x)}{\left(\widehat{M}_1(\theta,x)\right)^2} \frac{\hat{f}(\theta, \widehat{M}_\theta(x), x)}{P_n(\theta,x)\left(\hat{f}^{(2)}(\theta, \widehat{M}_\theta(x), x)\right)^2} \int \left(H'(u)\right)^2 du.$$

Hence one can derive the following corollary:

Corollary 4.2. Under the assumptions of Theorem 4.1, K' and $(K^2)'$ are integrable functions, $with$ $n \to \infty$.

1.

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

2.

$$\sqrt{\frac{ng_n^3\hat{\phi}_{\theta,x}(h_n)}{\hat{\varrho}^2(\theta, M_\theta(x), x)}}\left(\widehat{M}_\theta(x) - M_\theta(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Proof.

Observe that

1.

$$\Sigma(\theta,y,x) = \frac{\widehat{M}_1}{M_1}\sqrt{\frac{M_2}{\widehat{M}_2}}\sqrt{\frac{ng_n\hat{\phi}_{\theta,x}(h_n)P_n(\theta,x)f(\theta,y,x)}{p(\theta,x)\hat{f}(\theta,y,x)ng_n\phi_{\theta,x}(h_n)}} \times \frac{M_1}{\sqrt{M_2}}\sqrt{\frac{ng_n\phi_{\theta,x}(h_n)}{\sigma^2(\theta,y,x)}}\left(\hat{f}(\theta,y,x) - f(\theta,y,x)\right),$$

where $\Sigma(\theta,y,x) = \sqrt{\frac{ng_n\hat{\phi}_{\theta,x}(h_n)}{\hat{\sigma}^2(\theta,y,x)}}\left(\hat{f}(\theta,y,x) - f(\theta,y,x)\right)$, by Theorem 4.1, with $n \to \infty$, one has

$$\sqrt{\frac{ng_n\phi_{\theta,x}(h_n)}{\sigma^2(\theta,y,x)}}\left(\hat{f}(\theta,y,x) - f(\theta,y,x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

In order to prove (1), one has to show that

$$\frac{\widehat{M}_1}{M_1}\sqrt{\frac{M_2}{\widehat{M}_2}}\sqrt{\frac{ng_n\hat{\phi}_{\theta,x}(h_n)P_n(\theta,x)f(\theta,y,x)}{p(\theta,x)\hat{f}(\theta,y,x)ng_n\phi_{\theta,x}(h_n)}}\left(\hat{f}(\theta,y,x) - f(\theta,y,x)\right) \xrightarrow{\mathbb{P}} \mathcal{N}(0,1),$$

using the results given from (Laib and Louani, 2010), one obtains

$$\widehat{M}_1 \xrightarrow{\mathbb{P}} M_1, \widehat{M}_2 \xrightarrow{\mathbb{P}} M_2 \text{ and } \frac{\hat{\phi}_{\theta,x}(h_n)}{\sqrt{\phi_{\theta,x}(h_n)}} \xrightarrow{\mathbb{P}} 1 \text{ as } n \to \infty.$$

On the other hand, by Proposition 2 in (Laib and Louani, 2010), it follows that

$$P_n(\theta,x) \xrightarrow[n\to\infty]{} \mathbb{E}(\delta|\langle X,\theta\rangle = \langle x,\theta\rangle) = \mathbb{P}(\delta = 1|\langle X,\theta\rangle = \langle x,\theta\rangle) = p(x,\theta).$$

In addition, by Theorem 3.1, one has $\hat{f}(\theta,y,x) \longrightarrow f(\theta,y,x)$ with $n \to \infty$. This yields the proof of the first part of Corollary 4.2.

2.

$$\frac{\widehat{M}_1 \hat{f}^{(2)}(\theta, \widehat{M}_\theta(x), x)}{\sqrt{\widehat{M}_2}} \sqrt{\frac{n g_n^3 \hat{\phi}_{\theta,x}(h_n) P_n(\theta, x)}{\hat{f}(\theta, \widehat{M}_\theta(x), x)}} \left(\widehat{M}_\theta(x) - M_\theta(x)\right)$$
$$= \frac{\widehat{M}_1 \sqrt{M_2}}{M_1 \sqrt{\widehat{M}_2}} \sqrt{\frac{n g_n^3 \hat{\phi}_{\theta,x}(h_n) P_n(\theta, x) f(\theta, M_\theta(x), x)}{n g_n^3 \phi_{\theta,x}(h_n) p(\theta, x)}} \frac{\hat{f}^{(2)}(\theta, \widehat{M}_\theta(x), x)}{f^{(2)}(\theta, M_\theta(x), x)}$$
$$\times \frac{M_1}{\sqrt{M_2}} \sqrt{\frac{n g_n^3 \phi_{\theta,x}(h_n) p(\theta, x)}{f(\theta, M_\theta(x), x)}} f^{(2)}(\theta, M_\theta(x), x) \left(\widehat{M}_\theta(x) - M_\theta(x)\right).$$

Making use of Corollary 4.1, one obtains

$$\frac{M_1}{\sqrt{M_2}} \sqrt{\frac{n g_n^3 \phi_{\theta,x}(h_n) p(\theta, x)}{f(\theta, M_\theta(x), x)}} f^{(2)}(\theta, M_\theta(x), x) \left(\widehat{M}_\theta(x) - M_\theta(x)\right) \longrightarrow \mathcal{N}(0,1).$$

Further, by considering Lemma 4.2, (3.2) and (4.5), one obtains with $n \to \infty$, that

$$\frac{\widehat{M}_1 \sqrt{M_2}}{M_1 \sqrt{\widehat{M}_2}} \sqrt{\frac{n g_n^3 \hat{\phi}_{\theta,x}(h_n) P_n(\theta, x) f(\theta, M_\theta(x), x)}{n g_n^3 \phi_{\theta,x}(h_n) p(\theta, x)}} \frac{\hat{f}^{(2)}(\theta, \widehat{M}_\theta(x), x)}{f^{(2)}(\theta, M_\theta(x), x)} \xrightarrow{\mathbb{P}} 1.$$

Hence, the proof is completed.

Remark 4.1.

Thus, following Corollary 4.2, the asymptotic $(1-\xi)$ confidence interval of conditional density $f(\theta, y, x)$ and conditional mode $M_\theta(x)$, respectively, which are expressed as follows:

$$\hat{f}(\theta, y, x) \pm \eta_{\gamma/2} \sqrt{\frac{\hat{\sigma}^2(\theta, y, x)}{n g_n \hat{\phi}_{\theta,x}(h_n)}},$$

and

$$\widehat{M}_\theta(x) \pm \eta_{\gamma/2} \sqrt{\frac{\hat{\varrho}^2(\theta, M_\theta(x), x)}{n g_n^3 \hat{\phi}_{\theta,x}(h_n)}},$$

where $\hat{\sigma}^2(\theta, y, x)$ and $\hat{\varrho}^2(\theta, M_\theta(x), x)$ are defined in Corollary 4.2 respectively and $\eta_{\gamma/2}$ is the upper $\gamma/2$ quantile of the normal distribution $\mathcal{N}(0,1)$.

5. Examples

This section is devoted to specific examples demonstrating the application value of the research.

The examples illustrate the broad applicability of research that involves the estimation of conditional models with a functional single index structure while considering missing data. The specific details and examples would depend on the particularities of the research in question and its subsequent applications in various domains.

To understand the application value of this research, one may want to look into how the methodology or findings of the research were utilised in subsequent studies or practical applications. Here are some hypothetical examples of how such research could be applied:

- ***Biomedical Research:*** In the context of medical studies, the research might be applied to model and estimate the relation between a functional index and health outcomes. Functional data analysis is often applied in biomedical research, where data collected over time or space can have a functional structure. For example, analysing medical images over time or studying physiological responses over a period can benefit from models that incorporate missing data and functional single index structures. Dealing with missing data is common in clinical studies, and incorporating methodologies to address this improves the robustness of the analysis. Stationary processes with a functional nature could be relevant in analysing time-dependent physiological data.
- ***Finance:*** The research could find applications in modelling financial time series data where a functional single index could represent a key factor influencing market dynamics.
- Dealing with missing financial data is common due to holidays, weekends, or other reasons, making the incorporation of missing data handling techniques important.
- The consideration of stationary processes with a functional nature could be valuable for risk assessment and portfolio management.
- In finance, time-dependent data such as stock prices, interest rates, or economic indicators can be analysed using functional data models. Estimating conditional models with missing data can provide insights into the relation between various financial variables, helping in risk assessment and investment strategies.
- ***Environmental Studies:*** Environmental data, such as air quality measurements, temperature variations, and ecological patterns, often exhibit functional structures. Research in this domain might involve estimating conditional models to understand how different variables interact over time, considering missing data and the functional nature of the processes involved.
- ***Social Sciences:*** Functional data analysis is applicable in social sciences when dealing with time--dependent or spatial data. For instance, studying the evolution of social trends or economic indicators over time could benefit from models that account for missing data and incorporate a functional single index structure.
- ***Manufacturing and Engineering:*** In manufacturing processes, understanding the variation and performance of equipment over time is crucial. Functional data analysis can be applied to model the functional relations within the context of a single index structure, especially when dealing with processes that might have missing data due to equipment failure or bad maintenance.
- ***Psychology and Education:*** Longitudinal studies in psychology and education often involve collecting data over time, and the application of functional data models can aid in understanding how certain factors influence outcomes. Models that handle missing data are essential to account for participants dropping out or missing certain assessment points.
- ***Economic Forecasting:*** The functional single index structure could be applied in economic modelling where variables evolve over time. Incorporating missing data handling techniques can be crucial in dealing with incomplete economic datasets. The stationary processes with a functional nature could be relevant for understanding economic stability or trends.

6. Conclusion

This paper concentrated on the non-parametrical estimation of the conditional density function and conditional mode within the single functional index model for independent data, particularly when the variable of interest is affected by randomly missing data. This involved a semi-parametric single model structure and a censoring process on the variables. The study established both the almost complete uniform convergence of the proposed estimators and the asymptotic normality of the derived estimates under specific mild conditions, relying on standard assumptions in Functional Data Analysis

(FDA) for the proofs. Additionally, the authors explored the practical application of their findings in constructing confidence intervals for the proposed estimators.

This research has distinct characteristics from a theoretical standpoint, aiming to elucidate the model, process, and asymptotic outcomes of the primary focus of this analysis. The adaptive regression model considered for the presence of randomly missing data eliminated the deviations effect, and the nonparametric method employed in the single functional index model served as an alternative smoothing solution to the Nadaraya-Watson one, overcoming the technical complexity associated with bandwidth selection. Consequently, even in the presence of randomly missing data, the accuracy of the estimator was maintained, allowing for the integration of more practical scenarios.

References

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., and Boudiaf, M. (2014). On the Nonparametric Conditional Density and Mode Estimates in the Single Functional Index Model with Strongly Mixing Data. *Sankhyã Indian J. Stat.*, *76*(2), 356-378.

Attaoui, 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*, *79*(5), 485-511.

Chaudhuri, P., Doksum, K., and Samarov, A. (1997). On Average Derivative Quantile Regression. *Ann. Statist.*, (25), 715-744.

Cheng, P. E. (1994). Nonparametric Estimation of Mean Functional with Data Missing at Random. *Journal of the American Statistical Association*, *89*(425), 81-87.

Ezzahrioui, M., and Ould-Saïd, E. (2008). Asymptotic Normality of a Nonparametric Estimator of the Conditional Mode Function for Functional Data. *Journal of Nonparametric Statistics*, (20), 3-18.

Ferraty, F., Rabhi, A., and Vieu, P. (2005). Conditional Quantiles for Functional Dependent Data with Application to the Climatic El Niño Phenomenon. *Sankhyã*, (67), 378-398.

Ferraty, F., Sued, F., and Vieu, P. (2013). Mean Estimation with Data Missing at Random for Functional Covariables. *Statistics*, *47*(4), 688-706.

Ferraty, F., and Vieu, P. (2003). *Functional Nonparametric Statistics: A Double Infinite Dimensional Framework. Recent Advances and Trends in Nonparametric Statistics*, M. Akritas and D. Politis (Ed.). Elsevier.

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

Hamri, M. M., Mekki, S. D., Rabhi, A., and Kadiri, N. (2022). Single Functional Index Quantile Regression for Independent Functional Data Under Right-Censoring. *Econometrics*, *26*(1), 31-62. https://doi.org/10.15611/eada.2022.1.03

Kadiri, N., Rabhi, A., and Bouchentouf, A. A. (2018). Strong Uniform Consistency Rates of Conditional Quantile Estimation in the Single Functional Index Model under Random Censorship. *Journal Dependence Modeling*, *6*(1), 197-227.

Laib, N., and Louani, D. (2010). Nonparametric Kernel Regression Estimation for Functional Stationary Ergodic Data: Asymptotic Properties. *Journal of Multivariate Analysis*, (101), 2266-2281.

Lemdani, M., Ould-Saïd, E., and Poulin, N. (2009). Asymptotic Properties of a Conditional Quantile Estimator with Randomly Truncated. *Journal of Multivariate Analysis*, *100*(3), 546-559.

Liang, H., and de Uña-Alvarez, J. (2010). Asymptotic Normality for Estimator of Conditional Mode under Left-truncated and Dependent Observations. *Metrika*, *72*(1), 1-19.

Ling, N., Liang, L., and Vieu, P. (2015). Nonparametric Regression Estimation for Functional Stationary Ergodic Data with Missing at Random. *Journal of Statistical Planning and Inference*, (162), 75-87.

Ling, N., Liu, Y., and Vieu, P. (2016). Conditional Mode Estimation for Functional Stationary Ergodic Data with Responses Missing at Random. *Statistics*, *50*(5), 991-1013.

Ould-Saïd, E., and Cai, Z. (2005). Strong Uniform Consistency of Nonparametric Estimation of the Censored Conditional Mode Function. *Journal of Nonparametric Statistics*, *17*(7), 797-806.

Ould-Saïd, E., and Djabrane, Y. (2011). Asymptotic Normality of a Kernel Conditional Quantile Estimator Under Strong Mixing Hypothesis and Left-truncation. *Communications in Statistics. Theory and Methods*, *40*(14), 2605-2627.

Ould-Saïd, E., and Tatachak, A. (2011). A Nonparametric Conditional Mode Estimate under RLT model and Strong Mixing Conditions. *International Journal of Statistics and Economics*, (6), 76-92.

Rabhi, A., Kadiri, N., and Akkal, F. (2021). On the Central Limit Theorem for Conditional Density Estimator in the Single Functional Index Model. *Applications and Applied Mathematics: An International Journal (AAM)*, *16*(2), 844-866.

Rabhi, A., Kadiri, N., and Mekki, S. D. (2021). Asymptotic Properties of the Semi-Parametric Estimators of the Conditional Density for Functional Data in the Single Index Model with Missing Data at Random. *Statistica*, *81*(4), 399-422.
Roussas, G. (1969). Nonparametric Estimation of the Transition Distribution Function of a Markov Process. *Ann. Statist*., (40), 1386-1400.
Samanta, M. (1989). Nonparametric Estimation of Conditional Quantiles. *Stat. Prob. Lett*., (7), 407-412.

Centralne twierdzenie graniczne dla trybu warunkowego w jednolitym funkcjonalnym modelu indeksowym z losowym brakiem danych

Streszczenie: W artykule skoncentrowano się na nieparametrycznym estymowaniu warunkowej funkcji gęstości i warunkowej dominanty w modelu pojedynczego wskaźnika funkcjonalnego dla niezależnych danych, szczególnie gdy na interesującą zmienną wpływają losowo brakujące dane. Obejmuje to strukturę półparametrycznego pojedynczego modelu i proces cenzurowania zmiennych. Zgodność estymatora (ze współczynnikami) w różnych sytuacjach, np. w ramach modelu pojedynczego wskaźnika funkcjonalnego przy założeniu niezależnych i z identycznym rozkładem danych z losowymi brakami, a także jego działanie w warunkach, gdy zmienna towarzysząca jest funkcjonałem, to główne obszary zainteresowania. Dla tego modelu wyznacza się prawie całkowicie jednolitą zbieżność i wskaźnik zbieżności. Wskaźniki zbieżności podkreślają kluczową rolę, jaką prawdopodobieństwo koncentracji odgrywa w założeniach dotyczących objaśniającej zmiennej funkcjonalnej. Dodatkowo ustala się asymptotyczną normalność wyprowadzonych estymatorów zaproponowanych w określonych łagodnych warunkach, opierając się na standardowych założeniach z analizy danych funkcjonalnych dla dowodów. Na koniec zbadano praktyczne zastosowanie ustaleń w konstruowaniu przedziałów ufności dla naszych estymatorów. Wskaźniki zbieżności podkreślają kluczową rolę, jaką prawdopodobieństwo koncentracji odgrywa w założeniach dotyczących objaśniającej zmiennej funkcjonalnej.

Słowa kluczowe: funkcjonalna analiza danych, funkcjonalny proces pojedynczego indeksu, estymator jądra, losowe braki, estymacja nieparametryczna, prawdopodobieństwo małej kuli