République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieure et de la recherche scientifique

N Attribué par la bibliothèque

Année univ.: 2018/2019

ON THE CONSISTENCY AND THE ASYMPTOTIC NORMALITY OF SEVERAL CONDITIONAL MODELS, DEPENDENT CASE

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université de Saida - Dr Moulay Tahar
Discipline : MATHEMATIQUES
Spécialité: Analyse stochastique, statistique des processus et applications (ASSPA)

par

Seba **D**jillali[1]

Sous la direction de

Docteur Saâdia Rahmani

Soutenu le 17/07/2019 devant le jury composé de

Mr. **M**okhtar **K**adi	Université Dr. Moulay Tahar. Saïda	Président
Melle. **S**aâdia **R**ahmani	Université Dr. Moulay Tahar. Saïda	Encadreur
Mme. **R**achida **R**ouane	Université Dr. Moulay Tahar. Saïda	Examinatrice
Mme. **F**atima **B**enziadi	Université Dr. Moulay Tahar. Saïda	Examinatrice

[1]e-mail: sebadjillali833@gmail.com

Acknowledgements

First of all, I thank ALLAH for giving me the courage and the determination, as well as guidance in conducting this research study despite all difficulties.

I want to express my great appreciation and deepest gratitude to my supervisor, Miss Saâdia Rahmani, without whom this dissertation would not have been emerged. It was a great honor and an immense pleasure to be able to accomplish this work under her direction. I could benefit from her sound advice and careful guidance, her availability. Again thank you !

I want to thank very warmly Mr. kadi Mokhtar for chairing my committee members.

I wish to extend my sincere thanks to Miss Benziadi Fatima and Miss Rouane Rachida for agreeing to judge and criticize this work assuming roles of examiner.

I am grateful to my family for their continuous support, their encouragement, and especially for their putting up with me during all my study period.

Last but certainly not least, I would like to thank everybody who was important to the successful realization of this dissertation.

Contents

Chapter 1

General introduction

There is actually an increasing number of estimation coming from different fields of applied sciences, in which the collected data are curves, indeed the progress of the computing tools both in term of memory and computational capacities, allows us to deal with large sets of data. Since the middle of nineties, the different situations when functional variables can be observes has motivated the statistical development that we called "statistics for functional data".
Traditional statistical methods fail as soon as we deal with the functional data, if for instance we consider a sample of finely discretized curves two crucial statistical problems appear: the first comes from the relation between the size of the sample and the number of variables, the second is due to the existence of the strong correlations between variables and becomes an ill-conditional problem in the context of multivariate linear model, so there is a real necessity to develop statistical models.

A well-known statistical problem consists in studying the link between two variables in order to predict one of them, this problem has been widely studied for real or multivariate variables, but it also obviously occurs with functional variables.
There are several ways to approach the prediction setting, and one of the most popular is certainly the regression method which is based on conditional expectation. For robustness purposes we have two alternative techniques : conditional quantile and conditional mode.

The disadvantage of classical regression is that the estimation of the regression function is sensitive to outliers and may be inappropriate in some cases when the distribution is multimodal or strongly asymmetrical, the problem of robustness can be solved by the prediction using conditional mode.

The conditional quantile which can reveal an entire distributional relationship between the covariates and the response variable is another alternative predictor to classical regression. Moreover, conditional quantiles are well-known for their robustness with respect to heavy-tailed error distributions and outliers which allows to consider them as a useful alternative to the regression function. For the above reasons, conditional quantiles are used in many areas of applied research and are frequently used in a regression setup, called quantile regression.

1.1 Bibliographic context

We consider the estimation of the conditional mode and conditional quantile function when the covariates take values in some abstract functional space.
The main goal of this work is to establish the consistency and the asymptotic normality of the kernel estimator under α-mixing assumption and on the concentration properties on small balls of the probability measure of the functional regressors.
The two models has taken considerable attention in the past for both dependent and independent data.

The conditional mode, by its importance in the nonparametric forecasting field, has motivated a number of researchers in the investigation of mode estimators, and constitutes an alternative method to estimate the conditional regression (see *Ould Saïd* [25] for more discussion and examples).
In finite dimension spaces, there exists an extensive bibliography for independent and dependent data cases. In the independent case, strong consistency and asymptotic normality using the kernel method estimation of the conditional mode is given in *Samanta* and *Thavaneswaran* [31]. In the dependent case, the strong consistency of conditional mode estimator was obtained by *Collomb* et *al.* [3].

In our work, we pay attention to non-parametric conditional mode estimation via the functional conditional density. This problem has been interesting in the last few years. For example, *Ferraty* et *al.* [12] focused on kernel methods and almost sure (with rate) convergence was stated. This precursor work has been extended in many directions, including asymptotic normality (see *Ezzahrioui* and *Ould Saïd* [7] and [9] for both independent and dependant functional data). The consistency in L^p-norm of the conditional mode function estimator is given in *Dabo-Niang* and *Laksaci* [5]. The asymptotic normality, under α-mixing conditions was established by *Louani* and *Ould Saïd* [22]. Local linear estimation (see *Rachdi* et *al.* [26] and

Bouanani et *al.* [2]) or semi-parametric extensions to single index setting (see *Ling* and *Xu* [21]). For more discussions on non-parametric functional estimation context via some conditional features including the estimation of conditional mean, conditional median and conditional mode, one can refer to the monograph by *Ferraty* and *Vieu* [14] and the references therein.

In fact, in this dissertation, we prove, under certain standard conditions, the consistency and the asymptotic normality of the kernel estimator when we approach the problem of the prediction of a functional time series by estimating the conditional mode.

Since the seminal paper by *Koenker* and *Bassett* [20], the quantile regression method has been widely used in many disciplines such as economics, finance and other science fields. The quantile method serves as a robust alternative to the mean regression method and there is an extensive literature on the conditional quantile function estimation when the data are (i.i.d) or dependent and in finite dimensional spaces, for example, *Fan* et al [6], *Jones* and *Hall* [19], *Mehra* et *al.* [24], *Samanta* [30], *Welsh* [33]. As an introduction to this field (for parametric models), we refer the reader to the monographs of *Bosq* [1] and *Ramsay* and *Silverman* [27], [28].

There are many results for nonparametric models. For instance, *Ferraty* and *Vieu* [11] established the strong consistency of kernel estimators of the regression function when the explanatory variable is functional and the response is scalar. Their study is extended to nonstandard regression problems such as time series prediction or curves discrimination. The asymptotic normality result for the same estimator in the α-mixing case has been obtained by *Masry* [23]. *Gasser* et *al.* [16] gave an approach to introduce a nonparametric estimation of the quantile. They highlighted the issue of the curse of dimensionality for functional data and gave methods to overcome the problem. *Dabo-Niang* [4] studied density estimation in a *Banach* space with an application to the density estimation of a diffusion process with respect to *Wiener*'s measure. *Ferraty* et *al.* [14] studied the estimation of the conditional quantile for dependent functional data. The almost complete convergence with rates for the kernel type estimates is established and illustrated by an application to *El Nino* data. Recently, the kernel conditional quantile estimator has been treated under left truncation for functional regressors by *Helal* and *Ould-Saïd* [18].
Our aim in this work is to study the joint asymptotic properties of the kernel estimation of the conditional quantiles under an α-mixing condition. This mixing condition ensures asymptotically vanishing memory of the strictly stationary process.

1.2 Preliminaries

In what follows, we regroup some definitions and some tools that are going to be necessary established by this dissertation.

1.2.1 Definitions

In order to clarify many notions, we propose some basic definitions.
To fix the ideas, we give the following general definitions of a functional variable, functional data and functional model.

Definition 1.1. [14] (Functional variable) A random variable x is called functional variable if it takes values in an infinite dimensional space (or functional space). An observation X of x is called a functional data.

Definition 1.2. [14] (Functional dataset) A functional dataset $X_1, \ldots, X_n$ is the observations of n functional variables $x_1, \ldots, x_n$ identically distributed as x.

Definition 1.3. [14] (Functional model) Let x be a random variable valued in some infinite dimensional space $\mathcal{S}$ and let Υ be a mapping defined on the functional space $\mathcal{S}$ and depending on the distribution of x. A model for the estimation of Υ consists in introducing some constraint of the form:

$$\Upsilon \in \mathcal{C}.$$

The model is called a functional parametric model for the estimation of Υ if Λ is indexed by a finite number of elements of $\mathcal{S}$. Otherwise, the model is called a functional nonparametric model.

The appellation ***Functional Nonparametric Statistics*** covers all statistical backgrounds involving a nonparametric functional model.

Definition 1.4. (Mode function) The mode of a set of data values is the value that appears most often. For probability density f is the value which maximizes the function f, denoted by Θ,

$$\Theta = arg \max_{x \in \mathcal{C}} f(x),$$

where $\mathcal{C}$ is a compact set.

Definition 1.5. (Quantile function) The quantile function is the synonym of the percentile, we say that the quantile function of y of order p is defined by

$$\xi(p) = F^{-1}(p) = \inf\{y : \; F(y) \geq p\},$$

where $p \in [0, 1]$ and $F^{-1}(.)$ is the inverse of $F(.)$.

In the functional context, the choice of the preliminary norm becomes crucial. Even more, considering normed or metric spaces can become too restrictive. In some situations and this is the case for our data-sets, it appears that semi-metric spaces are better adapted than metric spaces. Before going on, let us just recall the definition of semi-metric.
Proximities measures between mathematical objects play a major role in all statistical methods. In many situations, a classical norm can be used to measure the closeness between two objects. Because in a finite dimensional Euclidean space (typically R^p) there is an equivalence between all norms. In the infinite dimensional space, the equivalence between norms fails and the problem has to be solved in a different way.

Definition 1.6. [14] (Semi-metric) d is a semi-metric on some space $\mathcal{S}$ as soon as

(1) $\forall x \in S \quad d(x, x) = 0,$

(2) $\forall (x, y, z) \in \mathcal{S} \times \mathcal{S} \times \mathcal{S} \quad d(x, y) \leq d(x, z) + d(z, y).$

Now we focus on some important modes of convergence.

Definition 1.7. [32] (Almost sure convergence) Let X_n be a sequence of random variables, the sequence is said to converge almost surely to the random variable if: $P(\lim_{n\to\infty} X_n = X) = 1$ in short $X_n \xrightarrow{a.s.} X$.

Definition 1.8. [32] (Almost complete convergence) We say that $(X_n)_n$ converges to X almost completely if and only if:

$$\forall \varepsilon > 0 \quad \sum_{n\geq 0} P(|X_n - X| > \varepsilon) < \infty,$$

and the almost complete convergence of $(X_n)_{n\in\mathbb{N}}$ to X is denoted by

$$\lim_{n\to\infty} X_n = X \quad \text{a.co.}$$

Remark 1 *The almost complete convergence implies other modes of convergence.*

Mixing conditions are usual structures for modeling dependence for a sequence of random variables, this notion is defined in the following way.

Definition 1.9. [1] (Mixing Processes) Mixing conditions are usual structures for modeling dependence for a sequence of random variables. Let (Ω, A, P) be a probability space, let $\mathcal{B}$ and $\mathcal{C}$ be two sub $\sigma-$field of $\mathcal{A}$.
Where $\mathcal{B} \in \sigma(X_s, s \leq t)$ and $\mathcal{C} \in \sigma(X_s, s \geq t+k)$.
In order to estimate the correlation between $\mathcal{B}$ and $\mathcal{C}$ various coefficient are used

(1) $\alpha = \alpha(\mathcal{B},\mathcal{C}) = \sup\limits_{\substack{B\in\mathcal{B}\\ C\in\mathcal{C}}} \mid P(B\cap C) - P(B)P(C) \mid,$

(2) $\beta = \beta(\mathcal{B},\mathcal{C}) = \sup\limits_{C\in\mathcal{C}} \mid P(C) - P(C|B) \mid,$

(3) $\varphi = \varphi(\mathcal{B},\mathcal{C}) = \sup\limits_{\substack{B\in\mathcal{B}, P(B)>0\\ C\in\mathcal{C}}} \mid P(B\cap C) - P(B)P(C) \mid,$

(4) $\rho = \rho(\mathcal{B},\mathcal{C}) = \sup\limits_{\substack{X\in L^2(\mathcal{B})\\ Y\in L^2(\mathcal{C})}} \mid corr(X,Y) \mid.$

These coefficients satisfy the following inequalities:

$$2\alpha \leq \beta \leq \varphi,$$

$$4\alpha \leq \rho \leq 2\varphi^{\frac{1}{2}}.$$

Then

$$\varphi\text{-mixing} \Rightarrow \beta\text{-mixing} \Rightarrow \alpha\text{-mixing},$$
$$\varphi\text{-mixing} \Rightarrow \rho\text{-mixing} \Rightarrow \alpha\text{-mixing}.$$

In this work we use the α-mixing (or strong mixing) notion, which is one of the most general among the different mixing structures introduced in the literature.
A process $(X_t, t\in\mathbb{Z})$ is said to be $\alpha-$mixing if

$$\sup_{t\in\mathbb{Z}} \alpha(\sigma(X_s, s\leq t),\ (\sigma(X_s, s\leq t+k)) \to 0 \quad \text{as} \quad k\to\infty,$$

and

$$\lim_{k\to\infty} \alpha(k) = 0.$$

Definition 1.10. [1](Geometrical and arithmetical α-mixing) We will mainly consider both of the following subclasses of mixing processes.
The sequence $X_n, n \in \mathbb{Z}$, is said to be arithmetically α-mixing with rate $a > 0$ if

$$\exists\, C > 0, \quad \alpha(n) \leq Cn^{-a}$$

It is called geometrically α-mixing if

$$\exists\, C > 0, \quad \exists\, t \in (0,1) \quad \alpha(n) \leq Ct^{n}$$

We will use the geometrical α-mixing process in the remaining of the work.

Definition 1.11. [32] (stochastic o and O symbols) For a given sequence of random variables R_n,

$$X_n = o(R_n) \quad \text{means} \quad X_n = Y_n R_n \quad \text{and} \quad Y_n \xrightarrow{P} 0,$$
$$X_n = O(R_n) \quad \text{means} \quad X_n = Y_n R_n \quad \text{and} \quad Y_n = O(1).$$

There are many rules of calculus with o and O symbols, which we apply without comment. For instance

$$o(1) + o(1) = o(1),$$
$$o(1) + O(1) = O(1),$$
$$o(1)O(1) = o(1),$$
$$(1 + o(1))^{-1} = O(1),$$
$$o(R_n) = R_n o(1),$$
$$O(R_n) = R_n O(1),$$
$$o(O(1)) = o(1).$$

1.2.2 Tools

Here are some tools that are going to be used to prove the main results.

Lemma 1.1. *[1] (Bernstein's inequality) This inequality is very useful to prove the almost sure and almost complete convergence.*

Let $(X_t, t \in \mathbb{Z})$ be a zero-mean real-valued strictly stationary bounded process. Then for each integer $q \in [1, \frac{n}{2}]$ and each $\varepsilon > 0$

$$\mathbb{P}(\sum_{t=1}^{n} X_t) \leq 4\exp\left(\frac{-\varepsilon^2}{8v^2(q)}q\right) + 22\left(1+\frac{4||X_0||\infty}{\varepsilon}\right)^{\frac{1}{2}} q\alpha\left(\left\lfloor\frac{n}{2q}\right\rfloor\right),$$

where

$$v^2(q) = \frac{8q^2}{n}V(\sum_{t=1}^{[\frac{n}{2q}]+1}) + \frac{\varepsilon||X_0||\infty}{2}$$

and $\alpha\left(\left\lfloor\frac{n}{2q}\right\rfloor\right)$ is the strong mixing coefficient of order $\left\lfloor\frac{n}{2q}\right\rfloor$[1].

Now, we present some lemmas that can be very important in the remainder of this work.

Lemma 1.2. *[17] (*Davydov's *Lemma) Let $\{\Delta_i, i \in \mathbb{N}\}$ be a sequences of real-valued random variables that verify strong mixing prcess, and let $\parallel \Delta \parallel_\infty < \infty$, $\forall i$ we have for all $i \neq j$,*

$$|cov(\Delta_i, \Delta_j)| < 4 \parallel \Delta_i \parallel\parallel \Delta_j \parallel \alpha(|i-j|).$$

Lemma 1.3. *[23] (*Volkonskii *and* Rozanov's *Lemma) Let $V_1, \ldots, V_L$ be strongly mixing random variables measurable with respect to the σ-algebras $\mathcal{F}_{i_1}^{j_1}, \ldots, \mathcal{F}_{i_L}^{j_L}$ respectively with $1 \leq i_1 < j_1 < i_2 < \ldots j_L \leq n, i_{l+1} - j_l > w > 1$ and $|V_j| \leq 1$ for $j = 1 \ldots L$, then*

$$\left|\mathbb{E}\left(\prod_{j=1}^{L}(V_j)\right) - \prod_{j=1}^{L}\mathbb{E}(V_j)\right| \leq 16(L-1)\alpha(w),$$

where $\alpha(w)$ is the strongly mixing coefficient.

Theorem 1.1. *[32] (*Slutsky's *Theorem) This theorem prepare the way to obtain the asymptotic normality of the model.*
Let X_n, X and Y_n be random variables, if $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} C$ for a constant C, then

(i) $X_n + Y_n \xrightarrow{D} X + C$,

(ii) $X_n Y_n \xrightarrow{D} XC$,

[1] $\lfloor x \rfloor$ integer part of x.

(iii) $X_n Y_n^{-1} \xrightarrow{D} XC^{-1}$.

Central limit theorems are theorems concerning convergence in distribution of sums of random variables. There are versions for dependent observations and normal limit distributions. The *Lindeberg-Feller* Theorem is the simplest extension of the classical central limit theorem and is applicable to independent observations with finite variances.

Theorem 1.2. *[32] (*Lindeberg-Feller *Theorem) For each n let $Y_{n,1}, \dots, Y_{n,k_n}$ be independent random vectors with finite variances such that*

$$\forall \varepsilon > 0, \quad \sum_{i=1}^{k_n} \mathbb{E}(\| Y_{n,i} \|^2 \mathbb{1}_{\|Y_{n,i}\| > \varepsilon}) \to 0,$$

$$\sum_{i=1}^{k_n} Cov(Y_{n,i}) \to \Sigma,$$

then the sequence $\sum_{i=1}^{k_n}(Y_{n,i} - \mathbb{E}(Y_{n,i}))$ converges in distribution to a normal $N(0, \Sigma)$ distribution.

1.2.3 Brief outline of the dissertation

This dissertation is dedicated to the survey of the asymptotic properties of conditional functional parameters in nonparametric statistics, the functional conditional mode and conditional quantile, when the explanatory variable takes its values in infinite dimensional space. The study of our functional estimators deals with the dependent data especially the strong mixing data. This dissertation is divided into four chapters, we begin by an introductory chapter which contains the bibliographic context and the most important definitions and tools which clarify many notions and will be used in the proofs of the main results. In the second chapter we focus on the functional mode, we study the consistency of the estimator and some results about asymptotic normality. In the third chapter we deal with conditional quantile then we show its asymptotic properties. In the last chapter, some applications are studied, we apply the conditional quantile approach to the prediction and the building of confidence bands. Finally we conclude the dissertation by a general conclusion.

Chapter 2

Functional Conditional Mode

The main goal of this chapter is to treat some asymptotic results of a conditional model "Mode function". This chapter is organized as follows. The first section is concerned with the presentation of our model and its estimator. In the second section we give the assumptions, some notations and remarks on the assumptions. In the third section we provide results about the consistency of the estimator and the asymptotic normality. In the fourth and last section we prove the results elaborated in the previous section.

2.1 The model and its estimator

Let $\{(X_i, Y_i), i = 1, \ldots, n\}$ be n copies of a random vector identically distributed as (X, Y) where X is valued in infinite dimensional semi-metric vector space $(\mathcal{S}, d(\ldots))$ and the variables Y are valued in $\mathbb{R}$. In most practical applications, $\mathcal{S}$ is a normed space (e.g. *Hilbert* or *Banach* space) with norm $\| . \|$ so that $d(x, x') = \| x - x' \|$. For $x \in \mathcal{S}$, we denote by $g(.|x)$ the conditional density function of Y given $X = x$. We assume that $g(.|x)$ has a unique mode and the conditional mode of Y given $X = x$, denoted by $\Theta(x)$ and defined by

$$g(\Theta(x)|x) = \max_{y \in \mathbb{R}} g(y|x). \tag{2.1}$$

A kernel estimator of the conditional mode $\Theta(x)$ is defined as the random variable $\Theta_n(x)$ which maximizes the kernel estimator $g_n(.|x)$ of $g(.|x)$ that is

$$g_n(\Theta_n(x)|x) = \max_{y \in \mathbb{R}} g_n(y|x), \tag{2.2}$$

here

$$g_n(y|x) = \frac{f_n(x,y)}{\ell_n(x)}, \tag{2.3}$$

where

$$f_n(x,y) = \frac{1}{nh_H\phi_x(h_K)}\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{h_K}\right) H^{(1)}\left(\frac{y-Y_i}{h_H}\right), \tag{2.4}$$

and

$$\ell_n(x) = \frac{1}{n\phi_x(h_K)}\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{h_K}\right). \tag{2.5}$$

Where K is a real valued kernel function h_K (resp. h_H) (depending on n) is a sequence of real positive numbers which goes to zero as n goes to infinity and $\phi_x(.)$ is a function which will be described later. $H^{(1)}$ is the first derivative of a given distribution function H.

For any $j \geq 1$, we define the j$^{\text{th}}$ partial derivative with respect to second component of $f_n(.,.)$ by

$$\frac{\partial f_n(x,y)}{\partial y^{(j)}} = f_n^{(j)}(x,y) = \frac{1}{nh_H^{j+1}\phi_x(h_K)}\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{h_K}\right) H^{(j+1)}\left(\frac{y-Y_i}{h_H}\right) \tag{2.6}$$

By the definition of the conditional mode function, we have

$$g^{(1)}(\Theta(x) \mid x) = 0.$$

It follows that

$$g_n^{(1)}(\Theta_n(x) \mid x) = 0.$$

Furthermore, we assume that $g^{(2)}(\Theta(x) \mid x) \neq 0$ and $g_n^{(2)}(\Theta(x) \mid x) \neq 0$.

By a *Taylor* expansion of $g_n^{(1)}(. \mid x)$ in the neighborhood of $\Theta(x)$, we have

$$\Theta_n(x) - \Theta(x) = -\frac{g_n^{(1)}(\Theta(x) \mid x)}{g_n^{(2)}(\Theta^*(x) \mid x)}. \tag{2.7}$$

Where $\Theta^*(x)$ lies between $\Theta_n(x)$ and $\Theta(x)$.

Using (2.3), we can write

$$\Theta_n(x) - \Theta(x) = -\frac{f_n^{(1)}(\Theta(x) \mid x)}{f_n^{(2)}(\Theta^*(x) \mid x)}. \tag{2.8}$$

If the denominator does not vanish.

Finally we put

$$a_l^x = K^l(1) - \int_0^1 ((K^l(u))'\zeta_0^x(u)du \quad \text{for} \quad l = 1, 2.$$

2.2 Assumptions and some remarks

2.2.1 Assumptions and notations

To formulate our assumptions, some additional notations are required. Let $B(x, h_K)$ be the ball of center x and radius h_K and consider the random variable $W := d(x, X)$ with $\phi_x(h_K) = \mathbb{P}(W \leq h_K) = \mathbb{P}(X \in B(x, h_K))$ for any fixed $x \in \mathcal{S}$ and $h_K > 0$. Furthermore let $\mathcal{C}$ be a compact set of $\mathbb{R}$ such that $\Theta(x) \in \mathcal{C}^{\circ}$, where $\mathcal{C}^{\circ}$ is the interior of $\mathcal{C}$.

(**H1**) There exist two functions $\Gamma_x(.)$ and $\zeta_0^x(.)$ such that

(i) for all $h_K > 0$, $\phi_x(h_K) > 0$ and $\lim_{h_K \to 0} \phi_x(h_K) = 0$,

(ii) for all $u \in [0, 1]$, $\lim_{h_K \to 0} \dfrac{\phi_x(uh_K)}{\phi_x(h_K)} = \lim_{h_K \to 0} \zeta_{h_K}^x(u) = \zeta_0^x(u)$,

(iii) $\sup_{i \neq j} \mathbb{P}[(X_i, X_j) \in B(x, h_K), B(x, h_K)] = \sup_{i \neq j} \mathbb{P}[W_i \leq h_K, W_j \leq h_K] = \Gamma_x(h_K)$, where $\lim_{h_K \to 0} \Psi_x(h_K) = 0$. Furthermore, we assume that $\Gamma_x(h_K) = o(\phi_x^2(h_K))$.

(**H2**) The conditional joint probability density and the conditional density are continuous with respect to each variable such that

(i) $\forall x \in \mathcal{S}, \forall (y, t) \in \mathbb{R}^2$,
$\sup_{\|X - x\| \leq h_K} \mid g^{(j)}(y - th_K \mid X) - g^{(j)}(y \mid X) \mid = o(1)$, as $h_K \to 0$ for $j = 1, 2$.

(ii) $\forall x \in \mathcal{S}, \forall (y, u, v) \in \mathbb{R}^3$,
$\sup_{\substack{\|X_i - x\| \leq h_K \\ \|X_j - x\| \leq h_K}} \mid g_{ij}(y - h_K u, y - h_K v \mid (X_i, X_j)) - g_{ij}(y, y \mid (x, x)) \mid = o(1)$,
where $g_{ij}((.,.), \mid (x, x))$ is the conditional density of (Y_i, Y_j) given $(X_i, X_j) = (x, x)$.

(**H3**) The mixing coefficient satisfies:

$$\sum_{k=1}^{+\infty} k^{\delta} (\alpha(k))^{1 - (\frac{2}{\tau})} \leq +\infty, \quad \text{for some } \tau > 2 \quad \text{and} \quad \delta > 1 - \frac{2}{\tau}$$

(**H4**) The kernel K has a compact support $[0, 1]$ and is of class C^1 on $[0, 1]$. Moreover

$K(0) > 0,\ K(1) > 0$ and $K' < 0$.

(**H5**) $H^{(1)}$ is twice differentiable such that

(i) $\forall (u_1, u_2) \in \mathbb{R}^2, \mid H^{(j)}(u_1) - H^{(j)}(u_2) \mid \leq C \mid u_1 - u_2 \mid$ for $j = 1, 3$.

(ii) $\int_{\mathbb{R}} \mid H^{(j)}(u) \mid^k du < +\infty$ for $j = 1, 3, k = 1$ and for $j = 2, k = 1, 2, \tau$.

(iii) $H^{(j)}$ is bounded for $j = 1, 2, 3$, furthermore, we assume that $m := \inf_{[0,1]} K(t) H^{(3)}(t) \neq 0$.

(**H6**) The bandwidths h_K and h_H satisfies, for $0 \leq j \leq 2$

(i) $n h_K^{2j+1} \phi_x^2(h_K) \to +\infty$ and $\dfrac{n h_K^{2j+1} \phi_x^2(h_K)}{\log^2 n} \to +\infty$ as $n \to +\infty$,

(ii) $n^\zeta h_H^{j+2} \to +\infty$ as $n \to +\infty$ for some $\zeta > 0$,

(iii) $n h_K^{2j+2} \phi_x^2(h_K) \to 0$ as $n \to +\infty$.

(**H7**) There exist sequence of integers (u_n) and (v_n) increasing to infinity such that $(u_n + v_n) \leq n$ satisfying

(i) $v_n = o((n h_K \phi_x(h_K))^{\frac{1}{2}}$ and $\left(\dfrac{n}{h_K \phi_x(h_K)}\right)^{\frac{1}{2}} \alpha(v_n) \to 0$ as $n \to +\infty$,

(ii) $q_n v_n = o((n h_K \phi_x(h_K))^{\frac{1}{2}}$ and $q_n \left(\dfrac{n}{h_K \phi_x(h_K)}\right)^{\frac{1}{2}} \alpha(v_n) \to 0$, as $n \to +\infty$, where q_n is the largest integer such that $q_n(u_n + v_n) \leq n$.

2.2.2 Remarks on the assumptions

Remark 2 (**H1**) *plays an important role in our methodology. It is known as the "concentration hypothesis acting on the distribution of X" in infinite dimensional space. This assumption is not at all restrictive and overcomes the problem of the non existence of the probability density function. In many examples, around zero the small ball probability $\phi_x(h_K)$ can be written approximately as the product of two independent functions $\jmath(x)$ and $\varphi(h_K)$ as $\phi_x(h_K)) = \jmath(x)\varphi(h_K) + o(\varphi(h_K))$. Furthermore, in infinite dimension, there exist many examples fulfilling the decomposition mentioned above. We quote the following decomposition (which can be found in* Ferraty, Mas *and* Vieu *[15]):*

(1) $\phi_x(h_K) \approx \jmath(x)h_K^{\gamma}$ [1] *for some* $\gamma > 0$*,*

(2) $\phi_x(h_K) \approx \jmath(x)h_K^{\gamma} \exp \dfrac{-C}{h_K^p}$ *for some* $\gamma > 0$ *and* $p > 0$*,*

(3) $\phi_x(h_K) \approx \jmath(x)/\mid \ln h_K \mid$.

The function $\zeta^x_{h_K}(.)$ *which intervenes in* (**H1**)*(ii) is increasing for all fixed* h_K*. Its pointwise limit* $\zeta^x_0(.)$ *also plays a determinant role. It intervenes in all asymptotic properties, in particular in the asymptotic variance term. With simple algebra, it is possible to specify this function (with* $\zeta_0(u) := \zeta^x_0(u)$*) in the above examples by*

(1) $\zeta_0(u) = u^{\gamma}$*,*

(2) $\zeta_0(u) = \delta_1(u)$ *where* $\delta_1(.)$ *is a Dirac function,*

(3) $\zeta_0(u) = 1\!\!1_{]0,1]}(u)$*.*

(**H1**)*(iii) is a classical and permits to make the covariance term negligible.*

Remark 3 (**H2**) *is smoothness condition, continuity type and it is the only condition involving the conditional probability density of* Y *given* X *and the joint conditional probability density of* (Y_1, Y_2) *given* (X_1, X_2)*. it means that* $g(. \mid .)$ *and* $g(.. \mid ..)$ *are continuous with respect to each variable.*
(**H4**) *and* (**H5**) *are technical and deal with the kernels* K *and* H*.*

Remark 4 (**H6**) *is classical in the functional estimation in finite and infinite-dimensional space, this assumption is used to balance between the bias term and the variance term.*
The choice of sequences (u_n)*,*(v_n) *and* (q_n) *in* (**H7**) *is classical and almost the same as in* Masry *[23], another choice can be found in* Roussas *[29].*

2.3 Main results

The first result of this section is Proposition 2.1 and Proposition 2.2 which focus on the consistency of the estimator Θ_n.

[1] $\approx$ approximately equal.

2.3.1 Consistency of the estimator

Our first result deals with the almost sure convergence of the kernel estimator $f_n(x, y)$.

Proposition 2.1. *Under* (**H4**), (**H5**) *(iii) and* (**H6**) *(i) and if for some* $\gamma \geq 2$

$$\sum_{p=1}^{+\infty} P^{\frac{2\gamma}{2\gamma+1}+3} \alpha^{\frac{2\gamma}{2\gamma+1}}(p) < \infty, \tag{2.9}$$

then for all $\varepsilon > 0$, *we have*

$$\sum_{n=1}^{+\infty} \mathbb{P}\{| f_n^{(j)}(x, y) - \mathbb{E}[f_n^{(j)}(x, y)] |> \varepsilon\} < +\infty \quad \text{for} \quad j = 0, 1, 2.$$

Demonstration

To prove Proposition 2.1, we use a version of *Bernstein's* exponential Inequality for the strong mixing case which is quoted in Lemma 2.1.
For any fixed $i \geq 1$, let

$$T_i^{(j)} = K\left(\frac{d(x, X_i)}{h_K}\right) H^{(j+1)}\left(\frac{y - Y_i}{h_H}\right) =: K_i H_i^{(j+1)}.$$

Lemma 2.1. *Let* $(T_i)_{i\geq 1}$ *be a sequence of the strong mixing sequence of real random variable with* Laplace *transform uniformly bounded on some interval* $[-\delta, +\delta]$ *then for every* $n \geq 2, \gamma > 0$ *and* $p \leq \frac{n}{2}$ *we have*

$$\mathbb{P}\left\{\frac{1}{n}\sum_{i=1}^{n} | T_i - \mathbb{E}[T_i] |> \varepsilon\right\} \leq 6\exp\left(\frac{-nt\varepsilon}{30p}\right) + 6\frac{n}{p}\left(\frac{10M_\gamma}{\varepsilon} + 1\right)^{\lambda} \alpha^{2\lambda}(p).$$

Where,

$$\lambda = \frac{\gamma}{2\gamma + 1},\ M_\gamma = \sup_{i\geq 1} \| T_i \|_\gamma,\ t = \min(\frac{\delta}{2}, \frac{\varepsilon}{3c}),\ c = 4\sup_{i\geq 1}\sum_{l=2}^{\infty} \frac{\delta^{l-2}}{l!}\mathbb{E} | T_i |^l .$$

This proposition leads us to obtain

Proposition 2.2. *Under the assumptions of Proposition 2.1,* (**H1**) *(i), (ii),* (**H2**) *(i),* (**H5**) *(i) and* (**H6**) *(iii), if the function* $\Theta(.)$ *satisfies* $\forall\varepsilon > 0$ *and* $\mu(x)$*, there exists* $\xi > 0$ *such that*

$$| \Theta(x) - \mu(x) |\geq \varepsilon \Rightarrow | g(\Theta(x) \mid x) - g(\mu(x) \mid x) |\geq \xi,$$

then, for all $x \in \mathcal{S}$ we have

$$\Theta_n(x) - \Theta(x) \xrightarrow{a.co.} 0 \quad \textit{as} \quad n \to \infty.$$

Demonstration

The proof of this proposition is based on the following decomposition:

$$\begin{aligned}
\mid g(\Theta_n(x) \mid x) - g(\Theta(x) \mid x) \mid &= \mid g(\Theta_n(x) \mid x) - g_n(\Theta_n(x) \mid x) + g_n(\Theta_n(x) \mid x) - g(\Theta(x) \mid x) \mid \\
&\leq \mid g(\Theta_n(x) \mid x) - g_n(\Theta_n(x) \mid x)| + |g_n(\Theta_n(x) \mid x) - g(\Theta(x) \mid x) \mid \\
&\leq \sup_{y \in \mathcal{C}} \mid g_n(y \mid x) - g(y \mid x) \mid + \sup_{y \in \mathcal{C}} \mid g_n(y \mid x) - g(y \mid x) \mid \\
&\leq 2 \sup_{y \in \mathcal{C}} \mid g_n(y \mid x) - g(y \mid x) \mid . \qquad (2.10)
\end{aligned}$$

The uniqueness hypothesis of the conditional mode gives us the result provided, we prove that the right hand side of (2.10) converges almost completely to zero.
For $0 \leq j \leq 2$, we have

$$\begin{aligned}
\sup_{y \in \mathcal{C}} \mid g_n^{(j)}(y \mid x) - g^{(j)}(y \mid x) \mid \;\; &\leq \frac{1}{\ell_n(x)} \{ \sup_{y \in \mathcal{C}} \mid f_n^{(j)}(x, y) - \mathbb{E}[f_n^{(j)}(x, y)] \mid \\
&+ \sup_{y \in \mathcal{C}} \mid \mathbb{E}[f_n^{(j)}(x, y)] - a_1^x g^{(j)}(y \mid x) \mid \} \\
&+ \frac{1}{\ell_n(x)} \{ \mid \ell_n(x) - \mathbb{E}[\ell_n(x)] \mid + \mid \mathbb{E}[\ell_n(x)] - a_1^x \mid \}. \qquad (2.11)
\end{aligned}$$

To show the almost complete convergence of the estimator we have to establish the following results.
The Lemmas 2.2 and 2.4 show the asymptotic bias term of $f_n^{(j)}(x, y)$ and $\ell_n(x)$ as n tends to infinity.

Lemma 2.2. *Under assumptions* (**H1**) *(i), (ii),* (**H2**) *(i),* (**H4**) *and* (**H5**) *(i) we have for* $0 < j < 2$

$$\mathbb{E}[f_n^{(j)}(x, y)] \to a_1^x g^{(j)}(y \mid x) \quad \textit{as} \quad n \to \infty.$$

Lemma 2.3. *Under assumptions* (**H1**), (**H2**), *for x fixed we have*

$$\frac{1}{\phi_x(h_K)} \mathbb{E}\left[K^j \left(\frac{d(x, X_i)}{h_K} \right) \right] \to a_l^x \quad n \to \infty \quad \textit{for} \quad l = 1, 2,$$

where

$$a_l^x = K^l(1) - \int_0^1 (K^l)'(u) \zeta_{h_K}(u) du$$

This lemma plays the same role as the classical *Bochner's* Lemma in finite dimension and gives us the convergence of the last term of the right hand side of (2.11).

Lemma 2.4. *Under assumptions (H1) (i), (ii) and (H4), we have*

$$\mathbb{E}[\ell_n(x)] = \frac{1}{\phi_x(h_K)}\mathbb{E}\left[K\left(\frac{d(x,X_i)}{h_K}\right)\right] \to a_1^x,$$

$$\frac{1}{\phi_x(h_K)}\mathbb{E}\left[K^2\left(\frac{d\,(x,X_i)}{h_K}\right)\right] \to a_2^x.$$

Now, we prove that $\ell_n(x) - \mathbb{E}[\ell_n(x)]$ converges almost completely to zero.

Lemma 2.5. *Under Assumptions* (**H1**) *(i), (ii),* (**H2**) *(i),* (**H4**) *and* (**H5**) *(i) we have* $0 \le j \le 2$

$$\sum_{n\ge1}\mathbb{P}\{\mid \ell_n(x) - \mathbb{E}[\ell_n(x)]\} < \infty.$$

This lemma allows us to obtain the following corollary.

Corollary 5 *Under Assumptions of Lemma 2.5, for all fixed $x \in \mathcal{S}$ we have*

$$\exists\, \tilde{\delta} > 0, \qquad \sum_{n=1}^{\infty}\mathbb{P}\{\mid \ell_n(x) < \tilde{\delta}\} < +\infty. \tag{2.12}$$

To conclude the result, we deal with the following lemma.

Lemma 2.6. *Under assumptions of Proposition 2.1, assumptions* (**H5**) *(i) and* (**H6**) *(i), (ii) we have*

$$\frac{1}{\ell_n(x)}\sup_{y\in\mathcal{C}}\mid f_n^{(j)}(x,y) - \mathbb{E}[f_n^{(j)}(x,y)]\mid \xrightarrow{a.co.} 0 \quad \text{as} \quad n\to\infty \quad \text{for} \quad 0\le j\le 2. \tag{2.13}$$

2.3.2 Asymptotic normality

The following result deals with the asymptotic normality of (2.8)

Theorem 2.3. *Under* (**H1**)*-*(**H7**) *we have for any $x \in \Xi$*

$$\left(\frac{nh_H^3\phi_x(h)}{\sigma^2(x,\Theta(x))}\right)^{\frac{1}{2}}(\Theta_n(x) - \Theta(x)) \xrightarrow{D} \mathcal{N}(0,1) \quad \text{as} \quad n\to\infty, \tag{2.14}$$

where $\Xi = \{x : x \in \mathcal{S}, g(\Theta(x)\mid x) \neq 0\}$ and

$$\sigma^2(x,\Theta(x)) := \frac{a_2^x g(\Theta(x)\mid x)}{[a_1^x g^{(2)}(\Theta(x)\mid x)]^2}\int_{\mathbb{R}}(H^{(2)}(t))^2dt.$$

Demonstration

Now, we give a decomposition which help us in the next steps of the proof. Recall that

$$\Theta_n(x)-\Theta(x)=-\frac{f_n^{(1)}(x,\Theta(x))-\mathbb{E}[f_n^{(1)}(x,\Theta(x))]}{f_n^{(2)}(x,\Theta_n^*(x))}-\frac{\mathbb{E}[f_n^{(1)}(x,\Theta(x))]}{f_n^{(2)}(x,\Theta_n^*(x))}. \tag{2.15}$$

According to *Slutsky's Theorem*, we will prove the asymptotic normality of the dominant term of the first term in the right hand side of (2.15), then we show that the denominator converges in probability to a constant c, Finally we check that the second term of the right hand side term is negligible.
We begin by the first step:

$$f_n^{(1)}(x,\Theta(x))-\mathbb{E}[f_n^{(1)}(x,\Theta(x))]=:\frac{1}{nh_H^2\phi_x(h_K)}\sum_{i=1}^{n}Z_i(x,\Theta(x)), \tag{2.16}$$

after suitable normalization where

$$Z_i(x,y):=K_iH_i^{(2)}-\mathbb{E}[K_iH_i^{(2)}],$$

denoted by

$$R_n(x,y):=\frac{1}{nh_H^2\phi_x(h_K)}\sum_{i=1}^{n}Z_i(x,y),$$

with $Z_i(x,y)=Z_i$ we have

$$Var(R_n(x,y))=\frac{1}{nh_H^4\phi_x^2(h_K)}Var(Z_1)+\frac{2}{n^2h_H^4\phi_x^2(h_K)}\sum_{1\le i<j\le n}^{n}Cov(Z_i,Z_j),$$

note that

$$Var(R_n(x,y))=Var(f_n^{(1)}(x,y)),$$

then

$$\begin{aligned}nh_H^3Var(f_n^{(1)}(x,y)) &= \frac{1}{h_H\phi_x^2(h_K)}Var(Z_1)+\frac{2}{nh_H\phi_x(h_K)}\sum_{1\le i<j\le n}^{n}Cov(Z_i,Z_j)\\ &= V_n(x,y)+\frac{2}{nh_H\phi_x(h_K)}\sum_{1\le i<j\le n}^{n}Cov(Z_i,Z_j).\end{aligned}$$

Now, we are in a position to establish the following lemma.

Lemma 2.7. *If* (**H1**)-(**H4**) *and* (**H5**) *(i), (ii) are fulfilled, then*

(i) $\lim_{n\to\infty} V_n(x,y) = V(x,y)$,

(ii) $\lim_{n\to\infty} \frac{2}{nh_H\phi_x(h_K)} \mid \sum_{1\le i<j\le n}^{n} Cov(Z_i, Z_j) \mid = 0,$

where

$$V(x,y) = a_2^x g(y \mid x) \int_{\mathbb{R}^2} (H^{(2)}(u))^2 du.$$

Therefore, we can write

$$\sqrt{nh_H^3\phi_x(h_K)}R_n(x,y) =: \frac{1}{\sqrt{nh_H\phi_x(h_K)}} \sum_{i=1}^{n} Z_i =: \frac{1}{\sqrt{n}}\widetilde{Z}_i =: \frac{1}{\sqrt{n}}S_n.$$

Then we show that

$$\frac{1}{\sqrt{n}}S_n \xrightarrow{D} \mathcal{N}(0, V(x,y)).$$

The second step is treated in the following lemma.

Lemma 2.8. *Under assumptions of Lemma 2.2, Lemma 2.6 and Proposition 2.1, for* $j = 2$

$$f_n^{(2)}(x, \Theta_n^*(x)) \xrightarrow{P} a_1^x g^{(2)}(\Theta(x) \mid x).$$

Now, we move to the second term in the right hand side of (2.15).

Lemma 2.9. *Under assumptions* (**H1**) *(i), (ii),* (**H2**)*(i),* (**H4**)*,* (**H5**)*(i), we have*

$$\frac{\sqrt{nh_H^3\phi_x(h_K)}}{f_n^{(2)}(x, \Theta_n^*(x))}(\mathbb{E}[f_n^{(1)}(x, \Theta_n(x))]) \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

In the last section we give technical proofs of the results elaborated in the previous section

2.4 Technical proofs and auxiliary results

2.4.1 Proof of Proposition 2.1

Using the Lemma 2.1, and by (**H4**) and (**H5**) (iii), we have that $T_i^{(j)}$ is bounded, therefore, its Laplace transform exists on any interval $[-\delta, +\delta]$. Furthermore, we have $M_\gamma < \infty$ and $c < \infty$. Now, we have

$$\mathbb{P}\{\mid f_n^{(j)}(x,y) - \mathbb{E}[f_n^{(j)}(x,y)] \mid > \varepsilon\} = \mathbb{P}\left\{\frac{1}{n}\sum_{i=1}^{n} \mid T_i^j - \mathbb{E}[T_i^{(j)}] \mid > \varepsilon\phi_x(h_K)h_H^{j+1}\right\}.$$

Applying (2.1) , we get

$$\begin{aligned}\mathbb{P}\{\mid f_n^{(j)}(x,y)-\mathbb{E}[f_n^{(j)}(x,y)]\mid>\varepsilon\} &\leq 6\exp\left(\frac{-nt_j\varepsilon\phi_x(h_K)h_H^{j+1}}{30p_j}\right)\\ &+ 6\frac{n}{p_j}\left(\frac{10M_\gamma^j}{\varepsilon\phi_x(h_K)h_H^{j+1}}+1\right)^\lambda\alpha^{2\lambda}(p_j).\end{aligned} \tag{2.17}$$

On the one hand, let
$t_j=\dfrac{\varepsilon\phi_x(h_K)h_H^{j+1}}{3c_j}$, where $c_j=4\sum_{l=2}^{\infty}\dfrac{\delta^{l-2}}{l!}\mathbb{E}\mid Ti^{(j)}\mid^l$.
Put $M_j=\sup_{[0,1]}\mid H^{(j+1)}\mid$,
clearly we have

$$\begin{aligned}c_j &= \frac{4}{\delta^2}\sum_{l=2}^{\infty}\frac{\delta^l M_j}{l!}\mathbb{E}\left[K^l\mathbb{E}\left[\mid\frac{H^{(j+1)}}{M_j}\mid^l\mid X_1\right]\right]\\ &\leq \frac{4}{\delta^2 M_j}\sum_{l=2}^{\infty}\frac{\delta^l M_j}{l!}\mathbb{E}\left[K^l\mathbb{E}\left[|H^{(j+1)}|\,|X_1\right]\right].\end{aligned}$$

Now for n large enough and by (**H5**) (ii), there exist a constant C_j such that:
$\mathbb{E}[\mid H^{(j+1)}\mid\mid X_1]]\leq C_jh_H^{j+1}$
We obtain that $c_j\leq C_j\phi(h)h_H^{j+1}$ then, $-t_j\leq -C_j\varepsilon$.
choosing $p_j=[\sqrt{nh_H^j\phi_x(h_K)}]$, we get that the first term of the right hand side of (2.17) satisfies

$$\exp\left(\frac{-nt_j\varepsilon\phi_x(h_K)h_H^{j+1}}{30p_j}\right)\leq\exp(C_j\varepsilon^2\sqrt{nh_H^{j+2}\phi_x(h_K)}).$$

Under the last part of (**H6**) (i), we have

$$\sum_{n\leq 1}\exp\left(\frac{-nt_j\varepsilon\phi_x(h_K)h^{j+1}}{30p_j}\right)<\infty.$$

On the other hand, we have

$$M_\gamma^j\parallel T_i^{(j)}\parallel_\gamma=(\mathbb{E}\mid T_i^{(j)}\mid^\gamma)^{\frac{1}{\gamma}}\leq C.$$

Now, we show that

$$\frac{n}{p_j}\leq\frac{4p_j}{h_H^j\phi_x(h_K)}\leq\frac{Cp_j^3}{nh_H^{2j}\phi_x^2(h_K)}\leq Cp_j^3.$$

Then, for n large enough we get

$$6\frac{n}{p_j}\left(\frac{10M_\gamma^j}{\varepsilon\phi_x(h_K)h_H^{j+1}}+1\right)^\lambda\cong C\frac{n}{p_j}\left(\frac{1}{\varepsilon\phi_x(h_K)h_H^{j+1}}\right)^\lambda\leq Cp_j^3\left(\frac{p_j^2}{\varepsilon p_j^2\phi_x(h_K)h_H^{j+1}}\right)^\lambda. \tag{2.18}$$

By the first part of (**H6**) (i), we have that $p_j^2\phi_x(h_K)h_H^{j+1} \to \infty$ then from (2.18)

$$6\sum_{p_j\geq 1}\frac{n}{p_j}\left(\frac{10M_\gamma^j}{\varepsilon\phi_x(h_K)h_H^{j+1}}+1\right)^{\lambda}\alpha^{2\lambda}(p_j) < \sum_{p_j\geq 1} Cp_j^{2\lambda+3} < \infty,$$

if and only if (2.9) holds, which complete the proof of Proposition 2.1. ■

Remark 6 *If $\alpha(p) = p^{-a}$ for $a > 0$, then for any $\frac{2}{5} \leq \lambda \leq \frac{1}{2}$ Proposition 2.1 holds for $a > 2\lambda + 1$.*

2.4.2 Proof of Proposition 2.2

Proof of Lemma 2.2

As the proof is identical for $j = 1$ and $j = 2$, we give only the first case and we suppose $X_1 \in B(X, h_K)$.

$$\begin{aligned}
\mathbb{E}[f_n^{(1)}(x,y)] &= \frac{1}{h_H^2\phi_x(h_K)}\mathbb{E}\left(K\left(\frac{d(x,X_i)}{h_K}\right)H^{(2)}\left(\frac{\theta(x)-Y_1}{h_H}\right)\right)\\
&= \frac{1}{h_H^2\phi_x(h_K)}\mathbb{E}\left(K\left(\frac{d(x,X_i)}{h_K}\right)\right)\mathbb{E}\left(H^{(2)}\left(\frac{\theta(x)-Y^1}{h_H}\right) \mid X_1\right).
\end{aligned}$$

Now we calculate

$$\begin{aligned}
\mathbb{E}\left(H^{(2)}\left(\frac{y-Y_1}{h_H}\right) \mid X_1\right) &= \int_{\mathbb{R}^2} H^{(2)}\left(\frac{y-Y^1}{h_H}\right) g(z \mid X_1)\\
&= h_H^2\int_{\mathbb{R}^2} H^{(1)}(t)[g^{(1)}(y-th_H \mid X_1) - g^{(1)}(y \mid X_1)]dt\\
&\quad + H^{(1)}(t)g^{(1)}(y \mid X_1)dt\\
&= h_H^2\int_{\mathbb{R}^2} H^{(1)}(t)[g^{(1)}(y-th_H \mid X_1) - g^{(1)}(y \mid X_1)] + h_H^2 g^{(1)}(y \mid X_1).
\end{aligned}$$

Then

$$\begin{aligned}
\mathbb{E}[f_n^{(1)}(x,y)] &= \frac{1}{\phi_x(h_K)}\mathbb{E}\left(K\left(\frac{d(x,X_i)}{h_K}\right)\right)\int_{\mathbb{R}^2}[g^{(1)}(y-th_H \mid X) - g^{(1)}(y \mid X)]\\
&\quad + \frac{1}{\phi_x(h_K)}\mathbb{E}\left(K\left(\frac{d(x,X_i)}{h_K}\right)\right)g^{(1)}(y \mid X).
\end{aligned}$$

Finally we get

$$\frac{1}{\phi_x(h_K)}\mathbb{E}\left(K\left(\frac{d(x,X_i)}{h_K}\right)\right)g^{(1)}(y \mid X) \to a_1^x g^{(1)}(y \mid X) \quad \text{as} \quad n\to\infty.$$

■

Proof of Lemma 2.4

To prove this lemma we calculate

$$\begin{aligned}
\frac{1}{\phi_x(h_K)}\mathbb{E}\left[K\left(\frac{d(x,X_i)}{h_K}\right)\right] &= \frac{1}{\phi_x(h_K)}\int_0^{h_K} K^l\left(\frac{u}{h_K}\right) dP^{d(x,X_i)} \\
&= \frac{1}{\phi_x(h_K)}\left[K^l\left(\frac{u}{h_K}\right)\phi_x(u)\right]_0^{h_K} - \frac{1}{\phi_x(h_K)}\int_0^{h_K} K^l\left(\frac{u}{h_K}\right)\phi_x(u)du \\
&= \frac{1}{\phi_x(h_K)}[K^l(1)\phi_x(h_K)] - \frac{1}{\phi_x(h_K)}\int_0^1 (K^l)'(u)\phi_x(uh_K)du \\
&= K^l(1) - \int_0^1 (K^l)'(u)\zeta_{h_K}(u)du \\
&= K^l(1) - \int_0^1 (K^l)'(u)\zeta_{h_K}(u)du \\
&= a_l^x \quad \text{as} \quad n \to \infty.
\end{aligned}$$

We replace l by 1 and 2 to obtain the result. ■

Proof of Lemma 2.5

The proof of this lemma is the same as in Proposition 2.1 with slight difference, to get the result we make use of the following choice:

$$K\left(\frac{d(x,X_i)}{h_K}\right), \quad t = \frac{\varepsilon\phi_x(h_K)}{3c}, \quad p = [\sqrt{n\phi_x(h_K)}].$$

■

Proof of Corollary 5

To obtain the claimed this result, we combine Lemma 2.5 below with the fact that $\ell_n(x) \geq \mathbb{E}[\ell_n(x)] - \mid \ell_n(x) - \mathbb{E}[\ell_n(x)] \mid$. Furthermore, we have $\mathbb{E}[\ell_n(x)] = \frac{a_1^x}{2}$, which completes the proof. ■

Now to end the proof of Proposition 2.2, it suffices to show the uniformity over $y \in \mathcal{C}$ of Proposition 2.1. This is the object Lemma 2.6.

Proof of Lemma 2.6

Consider a coverage of $\mathcal{C}$ by a finite number l_n of intervals $\mathcal{C}_k$ of the folowing form $\mathcal{C}_k = (s_k - w_n, s_k + w_n)$. We have that $\mathcal{C} \subset \cup_{k=1}^{l_n}\mathcal{C}_k$. Put $s_y = arg \min_{s\in\{s_1,\ldots,s_y\}} \mid y - s \mid$, then we have

$$\begin{aligned}\frac{1}{\ell_n(x)}\sup_{y\in\mathcal{C}} \mid f_n^{(j)}(x,y)-\mathbb{E}[f_n^{(j)}(x,y)] \mid &\leq \frac{1}{\ell_n(x)}\sup_{y\in\mathcal{C}} \mid f_n^{(j)}(x,y)-f_n^{(j)}(x,s_y) \mid \\ &+\frac{1}{\ell_n(x)}\sup_{y\in\mathcal{C}} \mid [f_n^{(j)}(x,s_y)]-\mathbb{E}[f_n^{(j)}(x,s_y)] \mid \\ &+\frac{1}{\ell_n(x)}\sup_{y\in\mathcal{C}} \mid \mathbb{E}[f_n^{(j)}(x,s_y)-\mathbb{E}[f_n^{(j)}(x,y)]) \\ &=: I_1+I_2+I_3.\end{aligned}$$

Regarding I_1, using the last part (**H5**) (i), we have

$$\begin{aligned}\frac{1}{\ell_n(x)}\sup_{y\in\mathcal{C}} \mid f_n^{(j)}(x,y)-f_n^{(j)}(x,s_y) \mid &\leq \frac{1}{\ell_n(x)}\sup_{y\in\mathcal{C}}\frac{1}{nh_H^{j+1}\phi_x(h_K)} \\ &\quad \sum_{i=1}^{n} \mid H_i^{(j+1)}(y)-H_i^{(j+1)}(s_y) \mid K_i(x) \\ &\leq \sup_{y\in\mathcal{C}}\frac{C\mid y-s_y|}{h_H^{2+j}\ell_n(x)}\left(\frac{1}{nh_H\phi_x(h_K)}\sum_{i=1}^{n}K_i(x)\right) \\ &\leq C\frac{w_n}{h_H^{j+2}}.\end{aligned}$$

Now take $w_n=n^{-\varsigma}$, then by (**H6**) (ii) we obtain that for n large enough

$$\mathbb{P}\{\frac{1}{\ell_n(x)}\sup_{y\in\mathcal{C}} \mid f_n^{(j)}(x,y)-\mathbb{E}[f_n^{(j)}(x,y)] \mid > \frac{\varepsilon}{3}\}=0. \tag{2.19}$$

Now we deal with I_2, we have

$$\begin{aligned}\mathbb{P}\{\sup_{y\in\mathcal{C}} \mid f_n^{(j)}(x,y)-\mathbb{E}[f_n^{(j)}(x,s_y)] \mid > \frac{\varepsilon}{3}\} &= \mathbb{P}\{\max_{s_y\in\{s_1,\dots,s_y\}} \mid f_n^{(j)}(x,y)-\mathbb{E}[f_n^{(j)}(x,s_y)] \mid > \frac{\varepsilon}{3}\} \\ &\leq l_n \max_{s_y\in\{s_1,\dots,s_y\}} \mathbb{P} \mid f_n^{(j)}(x,y)-\mathbb{E}[f_n^{(j)}(x,s_y)] \mid > \frac{\varepsilon}{3}\}.\end{aligned}$$

Now, using the same arguments as in Proposition 2.1, we take $l_n \leq Cw_n^{-1}$ with the same choice of w_n we get that for n large enough

$$\mathbb{P}\{\sup_{y\in\mathcal{C}} \mid f_n^{(j)}(x,y)-\mathbb{E}[f_n^{(j)}(x,s_y)] \mid = o(1) \quad \text{a.co.},$$

thus

$$\frac{1}{\ell_n(x)}\sup_{y\in\mathcal{C}} \mid f_n^{(j)}(x,y)-\mathbb{E}[f_n^{(j)}(x,s_y)] \mid = o(1) \quad \text{a.co.} \tag{2.20}$$

Finally, we use the same idea as for I_1, by Lemma 2.2 and for n large enough, we have

$$\begin{aligned}\sup_{y\in\mathcal{C}} \mid \mathbb{E}[f_n^{(j)}(x,y)] - \mathbb{E}[f_n^{(j)}(x,s_y)] \mid &\leq \sup_{y\in\mathcal{C}} \frac{1}{nh_H^{j+1}\phi_x(h)} \\ &\quad \sum_{i=1}^{n} \mid \mathbb{E}[H_i^{(j+1)}(y) - H_i^{(j+1)}(s_y) \mid K_i(x)] \\ &\leq \sup_{y\in\mathcal{C}} \frac{C \mid y - s_y \mid}{h_H^{j+2}} \left(\frac{1}{\phi_x(h_K)} \mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]\right) \\ &\leq C\frac{w_n}{h_H^{j+2}}.\end{aligned}$$

By the same arguments as for I_1 and for n large enough, we get

$$\mathbb{P}\{\frac{1}{l_n(x)}\mathbb{E}[f_n^{(j)}(x,y)] - \mathbb{E}[f_n^{(j)}(x,s_y)] \mid > \frac{\varepsilon}{3}\} = 0. \tag{2.21}$$

We conclude Lemma 2.6 from (2.19)-(2.21) and hence Proposition 2.2 is proved. ∎

2.4.3 Proof of Theorem 2.3

This theorem will be proved as long as the following lemmas can be checked.

Proof of Lemma 2.7

Part(i): We prove that $\lim_{n\to\infty} V_n(x,y) = V(x,y)$.

$$\begin{aligned}nh_H^3\phi_x(h)Var(f_n^{(1)}(x,y)) &= \frac{1}{h_H\phi_x(h_K)}\mathbb{E}\left[K^2\left(\frac{d(x,X_i)}{h_K}\right)\left(H^{(2)}\left(\frac{y-y_i}{h_H}\right)\right)^2\right] \\ &\quad -\frac{1}{h_H\phi_x(h_K}\mathbb{E}^2\left[K\left(\frac{d(x,X_i)}{h_K}\right)H^{(2)}\left(\frac{y-y_i}{h_H}\right)\right].\end{aligned}$$

First, we deal with D_n : $\quad D_n := h_H^3\phi_x(h_K)\left[\frac{\mathbb{E}\left(K\left(\frac{d(x,X_i)}{h_K}\right)H^{(2)}\left(\frac{y-y_i}{h_H}\right)\right)}{h_H^2\phi_x(h)}\right]^2$

Using (**H2**) we obtain:

$$h_H^3\phi_x(h_K)\mathbb{E}^2(f_n^{(1)}(x,y)) = 0 \quad \text{as} \quad n\to\infty.$$

Now, we deal with C_n

$$C_n := \frac{1}{h_H\phi_x(h_K)}\mathbb{E}\left(K^2\left(\frac{d(x,X_i)}{h_K}\right)\right)\mathbb{E}\left(H^{(2)}\left(\frac{y-y_i}{h_K} \mid X_1\right)\right).$$

We calculate

$$\begin{aligned}\mathbb{E}\left(H^{(2)}\left(\frac{y-Y_i}{h_H}\right) \mid X_1\right) &= h_H\int_{\mathbb{R}}(H^{(2)}(t))^2[g(y-th_H \mid x_1)-g(y\mid x)]dt \\ &+h_H g(y\mid x)\int_{\mathbb{R}}(H^{(2)}(t))^2dt.\end{aligned}$$

Then

$$\begin{aligned}\lim_{n\to\infty} C_n &= g(y\mid x)\int_{\mathbb{R}}(H^{(2)}(t))^2dt\,\underbrace{\frac{1}{\phi_x(h_K)}\mathbb{E}K^2\left(\frac{d(x,X_i)}{h_K}\right)}_{a_2^x} \\ &+a_2^x g(y\mid x)\int_{\mathbb{R}}(H^{(2)}(t))^2dt.\end{aligned}$$

By combining all these results, we get

$$\lim_{n\to\infty} V_n(x,y)=V(x,y).$$

Let us turn to Part (ii)
Using the decomposition of *Masry* [23], we define:

$$\{(i,j);1\leq |i-j|\leq l_n\} \quad \text{and} \quad \{(i,j);l_n\leq |i-j|\leq n-1\},$$

where $l_n=o(n)$, so

$$\begin{aligned}\frac{1}{nh_H\phi_x(h_K)}\sum_{1\leq i<j\leq n}^{n} Cov(Z_i,Z_j) &= \frac{1}{nh_H\phi_x(h_K)}\sum_{|i-j|\leq l_n} Cov(Z_i,Z_j) \\ &+ \frac{1}{nh_H\phi_x(h_K)}\sum_{|i-j|\geq l_n} Cov(Z_i,Z_j) \\ &=: A_n+B_n.\end{aligned}$$

On the one hand, by stationarity we have

$$\begin{aligned}Cov(Z_i,Z_j) &= \mathbb{E}[Z_i,Z_j] \\ &= \mathbb{E}[K_iH_i^{(2)}K_jH_j^{(2)}]-\mathbb{E}^2[K_1H_1^{(2)}] \\ &= \mathbb{E}[K_iK_j\mathbb{E}[H_i^{(2)}H_j^{(2)}\mid (X_iX_j)]]-\mathbb{E}^2[K_1H_1^{(2)}].\end{aligned}$$

By Assumption (**H2**) (ii) and changing variables, we get

$$\begin{aligned}\mathbb{E}[H_i^{(2)}H_j^{(2)}\mid (X_iX_j)] &= h_H^2\int_{\mathbb{R}}\int_{\mathbb{R}}H^{(2)}(u)H^{(2)}(v)[g_{i,j}(y-h_Hu,y-h_Hv\mid (X_i,X_i)) \\ &- g_{i,j}(y,y\mid (x,x))]dudv+h_H^2 g_{i,j}(y,y\mid (x,x))\left[\int_{\mathbb{R}}H^{(2)}(u)du\right]^2 \\ &= h_H^2\left(o(1)+g_{i,j}(y,y\mid (x,x))+\left[\int_{\mathbb{R}}H^{(2)}(u)du\right]^2\right) \\ &= h_H^2(o(1)+C).\end{aligned}$$

Then, by Assumption (**H1**) (iii)

$$\begin{aligned} \mid Cov(Z_i,Z_j)\mid &\leq Ch_H^2(o(1)+C)\sup_{i\neq j}\mathbb{P}\{(X_i,X_j)\in B(x,h_K),B(x,h_K)\}+\mathbb{E}^2[K_1H_1^{(2)}] \\ &\leq h_H^2(o(1)+C)\Gamma_x(h_K)+\mathbb{E}^2[K_1H_1^{(2)}]. \end{aligned}$$

Then

$$\begin{aligned} \mid A_n\mid &\leq \sum_{|i-j|\leq l_n}\frac{1}{n\phi_x(h_K)}(h_H^2(o(1)+C)\Gamma_x(h_K)+\mathbb{E}^2[K_1H_1^{(2)}]) \\ &\leq \frac{(o(1)+C)\Gamma_x(h_K)}{\phi_x(h_K)}h_Hl_n+h_H^3\phi_x(h_K)l_n\left[\frac{\mathbb{E}^2[K_1H_1^{(2)}]}{h_H^2\phi_x(h_K)}\right] \\ &= (o(1)+C)\left[\frac{\Gamma_x(h_K)}{\phi_x^2(h_K)}\right](h_H\phi_x(h)l_n)+h_H^3\phi_x(h_K)l_n(\mathbb{E}[f_n^{(1)}(x,y)])^2. \end{aligned}$$

Now choosing $\quad l_n=\left(\frac{1}{[h_H\phi_x(h_K)]^{1-\frac{2}{\delta}}}\right)^{\frac{1}{\delta}}$,
with assumption (**H1**) (iii) and Lemma 2.4 we get $A_n=o(1)$ as $n\to\infty$.
On the other hand, by *Davydov*'s Lemma [17] we have:

$$\mid Cov(Z_i,Z_j)\mid\leq 8(\mathbb{E}[\mid Z_i^\tau\mid])^{\frac{2}{\tau}}(\alpha(\mid i-j\mid))^{1-\frac{2}{\tau}}.$$

Using the conditional property, Assumptions (**H2**) (i) and (**H5**) (ii), we get

$$\begin{aligned} \mathbb{E}[\mid Z_i^\tau\mid] &\leq C\mathbb{E}[(K_i\mid H_i^2)^\tau] \\ &= C\mathbb{E}[K_i^\tau\mathbb{E}[(\mid H_i^2\mid)^\tau\mid X]] \\ &= Ch_H\mathbb{E}[K_i^\tau\{\int_{\mathbb{R}}(\mid H_i^2(u)\mid)^\tau[g(y-hu\mid X)-g(y\mid X)]du \\ &\quad +g(y\mid X)\int_{\mathbb{R}}(\mid H_i^2(u)\mid)^\tau du\}] \\ &= ch_H\mathbb{E}[K_i^\tau(o(1)+C)] \\ &= (o(1)+C)h_H\phi_x(h_K). \end{aligned}$$

Then

$$\mid Cov(Z_i,Z_j)\mid\leq(o(1)+C)^{\frac{2}{\tau}}(\alpha(\mid i-j\mid))^{1-\frac{2}{\tau}}.$$

By reducing the above double sum to a single sum and with the same choice of l_n as before, we obtain

$$\mid B_n\mid\leq\frac{(o(1)+C)^{\frac{2}{\tau}}}{h_H\phi_x(h_K)l_n^\delta(h_H\phi_x(h_K))^{\frac{2}{\tau}}}\sum_{p=l_n+1}^{\infty}p^\delta\alpha^{1-\frac{2}{\tau}}(p)=(o(1)+C)^{\frac{2}{\tau}}\sum_{p=l_n+1}^{\infty}p^\delta\alpha^{1-\frac{2}{\tau}}(p).$$

Finally, by (**H3**), we get $B_n = (o(1))$ as $n \to \infty$, which completes the proof of the lemma. ■

Now we have

$$\sqrt{nh_H^3\phi_x(h_K)}R_n(x,y) =: \frac{1}{\sqrt{nh_H\phi_x(h_K)}}\sum_{i=1}^{n} Z_i =: \frac{1}{\sqrt{n}}\widetilde{Z}_i =: \frac{1}{\sqrt{n}}S_n.$$

Where

$$\widetilde{Z}_i := \frac{Z_i}{\sqrt{h_H\phi_x(h_K)}} \quad \text{and} \quad S_n := \sum_{i=1}^{n}\widetilde{Z}_i$$

By the second part of Lemma 2.6, we have

$$\sum_{1\leq i<j\leq n}^{n} Cov(\widetilde{Z}_i, \widetilde{Z}_j) = o(n) \tag{2.22}$$

It remains to show that

$$\frac{1}{\sqrt{n}}S_n \xrightarrow{D} \mathcal{N}(0, V(x,y)). \tag{2.23}$$

To prove the asymptotic normality for S_n of dependent variables, make use of *Doob's* technique, used in *Masry* [23].

We consider the classical big and small block decomposition. We split the set $\{1, 2, \ldots, n\}$ into $2k_n+1$ subsets with large blocks of size u_n and small blocks of size v_n and put $k = \left\lfloor \frac{n}{u_n+v_n} \right\rfloor$.

Assumption (**H7**) (ii) allows us to define the large block size $u_n = \left\lfloor \left(\frac{nh\phi_x(h_K)}{q_n}\right)^{\frac{1}{2}} \right\rfloor$.

We use (**H7**) to prove that

$$\frac{u_n}{v_n} \to 0, \quad \frac{u_n}{n} \to 0, \quad \frac{u_n}{\sqrt{nh\phi_x(h_K)}} \to 0, \quad \frac{n}{u_n}\alpha(v_n) \to 0. \tag{2.24}$$

Now let N_j, N_j' and $N''(j)$ be defined as follows

$$N_j = \sum_{j(u+v)+1}^{j(u+v)+u} \widetilde{Z}_i \quad , \quad 0 \leq j \leq k+1,$$

$$N_j' = \sum_{j(u+v)+u+1}^{(j+1)(u+v)+u} \widetilde{Z}_i \quad , \quad 0 \leq j \leq k+1,$$

$$N''_j = \sum_{k(u+v)+1}^{n} \widetilde{Z}_i \quad , \quad 0 \leq j \leq k+1.$$

We can write

$$S_n = \sum_{j=0}^{k-1} N_j + \sum_{j=0}^{k-1} N'_j + rN''_k = S'_n + S''_n + S'''_n \tag{2.25}$$

Now, we make some claims which leads us to obtain the result.
Equation (2.26) proves that S''_n and S'''_n are asymptotically negligible.

$$(i) \quad \frac{1}{n}\mathbb{E}[S''_n] \to 0, \qquad (ii) \quad \frac{1}{n}\mathbb{E}[S'''_n] \to 0. \tag{2.26}$$

We begin by the first part (i), we have

$$\begin{aligned}
\mathbb{E}[S''_n]^2 &= var\left(\sum_{j=0}^{K-1} N'_j\right) \\
&= \sum_{j=0}^{k-1} Var(N'_j) + \sum_{0\leq i\leq j\leq K-1} Cov(N'_i, N'_j) \\
&= L_1 + L_2.
\end{aligned}$$

By the second-order stationarity we get

$$\begin{aligned}
Var(N'_j) &= Var\left(\sum_{j(u+v)+u+1}^{(j+1)(u+v)+u} \widetilde{Z}_i\right) \\
&= v_n Var(\widetilde{Z}_1) + \sum_{i\neq j}^{v_n} Cov(\widetilde{Z}_i, \widetilde{Z}_j),
\end{aligned}$$

then

$$\begin{aligned}
\frac{L_1}{n} &= \frac{kv_n}{n} Var(\widetilde{Z}_1) + \frac{1}{n}\sum_{j=0}^{k-1}\sum_{i\neq j}^{v_n} Cov(\widetilde{Z}_i, \widetilde{Z}_j) \\
&\leq \frac{kv_n}{n}\left[\frac{1}{n\phi_x(h_K)} Var(Z_1)\right] + \frac{1}{n}\sum_{i\neq j}^{n} \mid Cov(\widetilde{Z}_i, \widetilde{Z}_j) \mid .
\end{aligned}$$

The following algebra gives us

$$\frac{kv_n}{n} \cong \left(\frac{n}{u_n + v_n}\right)\frac{v_n}{n} \cong \frac{v_n}{u_n + v_n} \cong \frac{v_n}{u_n} \to 0 \quad \text{as} \quad n \to \infty.$$

Using (2.22) we have

$$\lim_{n\to\infty} \frac{L_1}{n} = 0. \tag{2.27}$$

Let us turn to $\frac{L_2}{n}$

$$\begin{aligned}\frac{L_2}{n} &= \frac{1}{n}\sum_{i=0}^{k-1}\sum_{j=0}^{k-1} Cov(N_i', N_j') \\ &= \frac{1}{n}\sum_{i=0}^{k-1}\sum_{j=0}^{k-1}\sum_{l_1=1}^{v_n}\sum_{l_2=1}^{v_n} Cov(\tilde{Z}_{\lambda_1+l_1}, \tilde{Z}_{\lambda_1+l_1})\end{aligned}$$

with $\lambda_i = i(u_n + v_n) + v_n$.as $i \neq j$, we have $\mid \lambda_i - \lambda_j + l_1 + l_2 \mid$, it follows that

$$\frac{L_2}{n} \leq \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} Cov(\mid \tilde{Z}_i, \tilde{Z}_j \mid),$$

then

$$\lim_{n\to\infty} \frac{L_2}{n} = 0. \tag{2.28}$$

By (2.27) and (2.28) we get part (i) of the lemma.
Now we deal with (ii), we have

$$\begin{aligned}\frac{1}{n}\mathbb{E}[S_n''']^2 &= \frac{1}{n}Var(N_k'') \\ &= \frac{\mu_n}{n}Var(\tilde{Z}_1) + \frac{1}{n}Cov(\tilde{Z}_i, \tilde{Z}_j),\end{aligned}$$

where $\mu_n = n - k_n(u_n + v_n)$ by the definition of k_n, we have $\mu_n \leq u_n + v_n$, then

$$\frac{1}{n}\mathbb{E}[S_n''']^2 \leq \frac{u_n + v_n}{n} var(\tilde{Z}_1) + \underbrace{\frac{1}{n}\sum_{i=1}^{\mu_n}\sum_{j=1}^{\mu_n} Cov(\tilde{Z}_i, \tilde{Z}_j)}_{o(1)}$$

and by the definition of u_n and v_n, we achieve the proof of (ii) ■
Equation (2.29) shows the asymptotic independence of $\{N_j\}$ in $S_n^{'}$

$$\mid \mathbb{E}[\exp(itn^{\frac{-1}{2}} S_n')] - \prod_{j=0}^{K-1} [\exp(itn^{\frac{-1}{2}} N_j)] \mid \to 0. \tag{2.29}$$

We make use of *Volkonskii* and *Rozanov's* Lemma [23] and the fact that the process (X_i, Y_i) is strong mixing.
Note that N_a is $\mathcal{F}_{i_a}^{i_a}$-measurable with $i_a = a(u_n + v_n) + 1$ and $j_a = a(u_n + v_n) + u_n$ hence, with $V_j = \exp(itn^{\frac{-1}{2}} N_j)$ we have

$$\mid \mathbb{E}[\exp(itn^{\frac{-1}{2}} S_n')] - \prod_{j=0}^{K-1} [\exp(itn^{\frac{-1}{2}} N_j)] \mid \leq 16 k_n \alpha(v_n + 1) \cong \frac{n}{u_n}\alpha(v_n + 1),$$

which goes to zero by the last part of (2.24). ■

Equations (2.30) and (2.31) are standard *Lindeberg-Feller conditions* for asymptotic normality under independence.

$$\frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}[N_j^2] \to V(x,y), \tag{2.30}$$

Proceeding as in the proof of (2.26) by replacing v_n by u_n we get

$$\frac{1}{n}\sum_{i=0}^{k-1}\mathbb{E}[N_j^2] = \frac{ku_n}{v_n}Var(\widetilde{Z}_1) + o(1).$$

As $var(\tilde{Z}_1) = V(x,y)$ and $\dfrac{ku_n}{n} \to 1$, we get the result. ■

$$\frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}[N_j^2 1\!\!1_{\{|N_j|>\varepsilon\sqrt{nV(x,y)}\}}] \to 0, \tag{2.31}$$

for every ε, $n > 0$.

Recall that

$$N_j = \sum_{i=j(u+v)+1}^{j(u+v)+u}\widetilde{Z}_i \quad \text{where} \quad \widetilde{Z}_i := \frac{1}{\sqrt{h_H\phi_x(h_K)}}(K_iH_i^{(2)} - \mathbb{E}[K_iH_i^{(2)}]).$$

Making use (**H4**) and (**H5**) (iii), we have

$$|\ \widetilde{Z}_i\ | \leq \frac{C}{\sqrt{h_H\phi_x(h_K)}} \quad \text{thus} \quad |\ \frac{N_j}{\sqrt{n}}\ | \leq \frac{Cu_n}{\sqrt{nh_H\phi_x(h_K)}},$$

which goes to zero as n goes to infinity by (2.24), if we take n large enough, the set $\{|\ N_j\ | > \varepsilon\sqrt{nV(x,y)}\}$ becomes empty, this completes the proof and therefore that of (2.23). ■

Lemma 2.9 proves the convergence in probability of the decomposition given in equation (2.15).

Proof of Lemma 2.9

By the last part of assumption (**H5**) (iii) we have

$$\frac{\sqrt{nh_H^3\phi_x(h_K)}}{f_n^{(2)}(x,\Theta_n^*(x))} \leq \frac{\sqrt{nh_H^3\phi_x(h_K)}}{m}.$$

Finally (**H6**) (iii) and Lemma 2.2 allows us to conclude. ■

Now, we come back to the proof of Theorem 2.3.

From (2.15), we have

$$\sqrt{nh_H^3\phi_x(h_K)}(\Theta_n(x)-\Theta(x)) = \frac{\sqrt{nh_H^3\phi_x(h_K)}(f_n^{(1)}(x,\Theta_n^*(x))\mathbb{E}[f_n^{(1)}(x,\Theta_n^*(x))])}{f_n^{(2)}(x,\Theta_n^*(x))} - \frac{\sqrt{nh_H^3\phi_x(h_K)}\mathbb{E}[f_n^{(1)}(x,\Theta_n^*(x))])}{f_n^{(2)}(x,\Theta_n^*(x))}. \tag{2.32}$$

Using Lemma 2.8, and then by Lemma 2.9, we conclude that the last term of the right-hand side of (2.32) goes to zero as n goes to infinity.

Now, by (2.23) and *Slutsky's* Theorem we have that the first term of right hand side (2.32) is asymptotically normal with variance term equal to $\frac{V(x,y)}{[a_1^x g^{(2)}(\Theta(x)\mid x)]^2}$.

Which completes the proof of Theorem (2.3). ■

Chapter 3

Functional Conditional Quantile

In this chapter we will treat some asymptotic properties of a non-parametric conditional quantile estimator, this chapter is divided into four sections. Firstly we describe our model then we construct its estimator. Secondly we give the assumptions, notations and some important remarks. Then we provide some important results about the consistency of the estimator and the asymptotic normality. The last chapter is concerned with technical proofs of the main results elaborated in the previous section.

3.1 The model and its estimator

Let $\{(X_i, Y_i), i = 1, \ldots, n\}$ be n copies of a random vector identically distributed as (X, Y) where X is valued in infinite dimensional semi-metric vector space $(\mathcal{S}, d(.,.))$ and the variables Y are valued in $\mathbb{R}$. For $x \in \mathcal{S}$ we denote the conditional probability distribution of Y given $X = x$ by

$$\forall y \in \mathbb{R} \quad F(y|x) = \mathbb{P}(Y \leq y | X = x). \tag{3.1}$$

A natural way to model a conditional quantile function is to invert a conditional cumulative distribution function at the desired quantile. So, we denote the conditional quantile of order p in $[0, 1]$ by

$$\xi_p = \inf\{y : F(y|x) \geq p\}. \tag{3.2}$$

It is clear that an estimator of ξ_p can easily be deduced from an estimator of $F(y|x)$. A kernel of $F(y|x)$ is given by

$$F_n(y|x) = \frac{\sum_{i=1}^n K\left(\frac{d(x,X_i)}{h_K}\right) H\left(\frac{y-Y_i}{h_H}\right)}{\sum_{i=1}^n K\left(\frac{d(x,X_i)}{h_K}\right)}. \quad (3.3)$$

Where K, H, h_H and h_K are defined in Chapter 2.
Suppose that, for any fixed x, $F(.|x)$ be continuously differentiable real function, and admits a unique conditional quantile.
Let $p \in [0,1]$, we will consider the problem of estimating the parameter $\xi_p(x)$ which satisfies

$$F(\xi_p(x)|x) = p. \quad (3.4)$$

Then a natural estimator of $\xi_p(.)$ is given by

$$\xi_{p,n} = \inf\{y : F_n(y|x) \geq p\}, \quad (3.5)$$

which satisfies

$$F_n(\xi_{p,n}(x)|x) = p. \quad (3.6)$$

We can write (3.3) as

$$F_n^{(j)}(y|x) = \frac{\Psi_n^{(j)}(x,y)}{g_n(x)}, \quad (3.7)$$

where

$$\Psi_n^{(j)}(x,y) = \frac{1}{nh_H^j \phi_x(h_K)} \sum_{i=1}^n K\left(\frac{d(x,X_i)}{h_K}\right) H^{(j)}\left(\frac{y-Y_i}{h_H}\right),$$

and

$$g_n(x) = \frac{1}{n\phi_x(h_K)} \sum_{i=1}^n K\left(\frac{d(x,X_i)}{h_K}\right).$$

In what follows, for $j \geq 0$ we denote by $F_n^j(.|x)$ the j^{th} derivative of $F(.|x)$ (resp $F_n(.|x)$).
By a *Taylor* expansion of $F_n(.|x)$ around $\xi_p(x)$ we get

$$F_n(\xi_p(x)|x) - F(\xi_p(x)|x) = (\xi_p(x) - \xi_{p,n}(x)) f_n(\xi_{p,n}^*|x). \quad (3.8)$$

Where $f_n(.|x) = \dfrac{\partial F_n}{\partial y}(.|x)$ is the estimate of the conditional density of Y given $X = x$ and $\xi_{p,n}^*(x)$ lies between $\xi_p(x)$ and $\xi_{p,n}(x)$.
Equation (3.8) shows that from the asymptotic behavior of $F_n(\xi_p(x)|x) - F(\xi_p(x)|x)$ as n goes to infinity, it is easy to obtain the asymptotic results for the sequence

$(\xi_{p,n}(x) - \xi_p(x))$ and prove why the rates of convergence are the same for $\xi_{p,n}(x)$ and $F_n(\xi_{p,n}(x)|x)$.
In the next section, the assumptions that allow us to derive the main results are stated.

3.2 Assumptions and some remarks

3.2.1 Assumptions and notations

We start by introducing the following notation

$$a_l^x = K^l(1) - \int_0^1 (K^l(u))'\zeta_0^x(u)du \quad \text{for} \quad l = 1, 2.$$

The first three assumptions are the same used in Chapter 2: (**H1**)-(**H3**).
The fourth assumption is modified by adding a second condition
(**Q4**) The kernel functions satisfies

(ii) For $j = 0, 1$, $H^{(j)}$ satisfy the *Lipschitz* Condition and furthermore $m = \inf_{[0,1]} K(t)H^1(t) > 0$.

Then we impose another assumption
(**Q5**) The bandwidths h_K and h_H satisfy the following, for $0 \le j \le 1$

(i) $nh_K^{2j+1}\phi_x^2(h_K) \to +\infty$ and $\dfrac{nh_K^{2j+1}\phi_x^2(h_K)}{\log^2 n} \to +\infty$ as $n \to +\infty$,

(ii) $nh_H^{2j}\phi_x^3(h_K) \to +\infty$ as $n \to +\infty$,

(iii) there exists a positive sequence v_n such that $v_n \to +\infty$, $v_n = o(n\phi_x(h_K))$ and $\left(\dfrac{n}{\phi_x(h_K)}\right)\alpha(v_{n)} \to 0$ as n goes to infinity.

3.2.2 Remarks on the assumptions

Remark 7 (**Q4**) *is classical in functional estimation for finite and infinite dimensional space, this condition deals with the kernel H. Concerning* (**Q5**)*, we use this assumption to balance between variance and bias terms.*

3.3 Main results

Our first result states the almost complete convergence of the conditional distribution function and its first derivative.

3.3.1 Consistency of the estimator

Theorem 3.1. *Under assumptions* (**H1**) *(i), (ii),* (**H3**), (**Q4**) *and* (**Q5**) *(i), we have for* $j = 0, 1$

$$\sup_{y\in\mathcal{C}} |F_n^{(j)}(y|x) - F^{(j)}(y|x)| \to 0 \quad a.co. \quad as \quad n \to \infty.$$

Demonstration

The first step is to give the decomposition of $\sup_{y\in\mathcal{C}} |F_n^{(j)}(y|x) - F^{(j)}(y|x)|$ as follows

$$\begin{aligned}\sup_{y\in\mathcal{C}} |F_n^{(j)}(y|x) - F^{(j)}(y|x)| \leq & \frac{1}{g_n(x)}\{\sup_{y\in\mathcal{C}} |\Psi_n^{(j)}(x,y) - \mathbb{E}[\Psi_n^{(j)}(x,y)]| \\ & + \sup_{y\in\mathcal{C}} |\mathbb{E}[\Psi_n^{(j)}(x,y)] - a_1^x \Psi^{(j)}(x,y)| \\ & + \sup_{y\in\mathcal{C}} F^{(j)}(y|x)|a_1^x g(x) - \mathbb{E}[g_n(x)]| \\ & + \sup_{y\in\mathcal{C}} F^{(j)}(y|x)|\mathbb{E}[g_n(x)] - g_n(x)|\}.\end{aligned}$$

The following lemma shows the asymptotic bias term of $\Psi_n(x,y)$ and $g_n(x)$ as n tends to infinity

Lemma 3.1. *Under assumptions* (**H1**) *(i), (ii),* (**H2**) *and* (**Q4**), *we have*

$$(i) \quad \mathbb{E}[g_n(x)] \to a_1^x g(x) \quad as \quad n \to \infty$$

$$(ii) \quad \mathbb{E}[\Psi_n^{(j)}(x,y)] \to a_1^x \Psi^{(j)}(y|x) \quad as \quad n \to \infty \tag{3.9}$$

The following lemma deals with the variance term $\sup_{y\in\mathcal{C}} |\Psi_n^{(j)}(x,y) - \mathbb{E}[\Psi_n^{(j)}(x,y)]$

Lemma 3.2. *Under assumptions* (**Q4**) *and* (**Q5**) *(i) then if* $\sum_k k^{2\lambda+3}(\alpha(k))^{2\lambda} < \infty$ *for some* $\lambda \in [\frac{2}{5}, \frac{1}{2}]$, *we have*

$$\sup_{y\in\mathcal{C}} |\Psi_n^{(j)}(x,y) - \mathbb{E}[\Psi_n^{(j)}(x,y)]| \to 0, \quad a.co. \quad for\ j = 0, 1.$$

The almost complete convergence of the conditional quantile is given below.

Theorem 3.2. *Under assumptions of Theorem 3.1, we suppose that the conditional quantile* $\xi_p(x)$ *of the order* $p \in [0, 1]$ *is unique then*

$$|\xi_{p,n}(x) - \xi_p(x)| \to 0 \quad a.co. \quad as \quad n \to \infty \tag{3.10}$$

3.3.2 Asymptotic normality

The following result gives the asymptotic normality of the conditional distribution function.
Let $\mathfrak{A} = \{(x,y) : (x,y) \in \mathcal{S} \times \mathbb{R}, V(x,y) \neq 0\}$ and $\Xi = \{x : x \in \mathcal{S}, \Sigma(x, \xi_p(x)) \neq 0\}$

Theorem 3.3. *Under assumptions* (**H1**)-(**H3**), (**Q4**) *and* (**Q5**), *we have*

$$\left(\frac{n\phi_x(h_K)}{\sigma^2(x,y)}\right)^{\frac{1}{2}} (F_n(y|x) - F(y|x)) \xrightarrow{D} \mathcal{N}(0,1) \quad \text{as} \quad n \to \infty. \tag{3.11}$$

Where $\sigma^2(x,y) := \dfrac{V(x,y)}{(a_1^x)^2}$

Demonstration

Now, to prove the asymptotic normality of $F_n(y|x)$ and $\xi_{p,n}(x)$, we need to define some notations.
Firstly, we give the following decomposition:

$$\begin{aligned} F_n(y(x)|x) - F(y(x)|x) &= \frac{\Psi_n(x,y)}{g_n(x)} - \frac{a_1^x F(y|x)}{a_1^x} \\ &= \frac{1}{g_n(x)}\{\Psi_n(x,y) - \mathbb{E}[\Psi_n(x,y)] - F(y|x)[g_n(x) - \mathbb{E}[g_n(x)]] \\ &\quad - \frac{1}{g_n(x)}\{a_1^x F(y|x) - \mathbb{E}[\Psi_n(x,y)] - F(y|x)[a_1^x - \mathbb{E}[g_n(x]]\} \\ &=: \frac{1}{g_n(x)}(R_n(x,y) + B_n(x,y)). \end{aligned} \tag{3.12}$$

After that, we deal with $R_n(x,y)$ then $B_n(x,y)$. Thus

$$\begin{aligned} R_n(x,y) &= \frac{1}{n\phi_x(h_K)}\sum_{i=1}^{n}[(H_i(y) - F(y|x))K_i(x) - \mathbb{E}[(H_i(y) - F(y|x))K_i(x)]] \\ &= \frac{1}{n\phi_x(h_K)}\sum_{i=1}^{n} N_i(x,y). \end{aligned}$$

It follows that

$$\begin{aligned} n\phi_x(h_K)Var[R_n(x,y)] &= \frac{1}{\phi_x(h_K)}Var(N_1) + \frac{1}{n\phi_x(h_K)}\sum_{|i-j|>0}^{n}\sum^{n} Cov(N_i, N_j) \\ &= V_n(x,y) + \frac{1}{n\phi_x(h_K)}\sum_{|i-j|>0}^{n}\sum^{n} Cov(N_i, N_j). \end{aligned}$$

To show the claimed asymptotic normality, the following results have to be checked.

Lemma 3.3. *Under assumptions* (**H2**) *and* (**Q4**) *(ii), we have*

$$Var\left[H\left(\frac{y-Y_1}{h_H}\right)\right] \to F(y|x)[1-F(y|x)] \quad as \quad n \to \infty.$$

Now, we deal with $n\phi_x(h_K)Var[R_n(x,y)]$

Lemma 3.4. *Under assumptions* (**H1**)-(**H3**) *and* (**Q4**), *we have*

$$V_n(x,y) \to V(x,y) \quad as \quad n \to \infty, \tag{3.13}$$

$$\frac{1}{n\phi_x(h_K)} \sum_{|i-j|>0}^{n} \sum^{n} |Cov(N_i, N_j)| = o(1), \quad as \quad n \to \infty, \tag{3.14}$$

where $V(x,y) = a_2^x F(y|x)(1-F(y|x))$.

Lemma 3.5. *Under assumptions* (**H1**)-(**H3**), (**Q4**) *and* (**Q5**)*(iii), we have for any* $(x,y) \in \mathcal{S} \times \mathbb{R}$

$$(n\phi_x(h_K))^{\frac{1}{2}} Var[R_n(x,y)] \xrightarrow{D} \mathcal{N}(0, V(x,y)) \quad as \quad n \to \infty.$$

Corollary 8 *Under assumptions of Lemma 3.4, we have*

$$n\phi_x(h_K)Var[R_n(x,y)] \to V(x,y) \quad as \quad n \to \infty.$$

Now, we move to the second term and we prove that $(n\phi_x(h_K))^{\frac{1}{2}} B_n(x,y)$ converges to zero in probability.

Lemma 3.6. *Under assumptions* (**H1**), (**Q5**) *(i) (ii), we have*

$$(n\phi_x(h_K))^{\frac{1}{2}} B_n(x,y) \xrightarrow{P} 0 \quad as \quad n \to \infty.$$

Corollary 9 *Under assumptions of Lemma 3.6, we have*

$$\frac{(n\phi_x(h_K))^{\frac{1}{2}} B_n(x,y)}{f_n(\xi_{p,n}^*(x)|x)} \xrightarrow{P} 0 \quad as \quad n \to \infty.$$

Theorem 3.4. *Let p is the unique order of the conditional quantile such that* $p = F(\xi_p(x)|x) = F_n(\xi_{p,n}(x)|x)$, *if the Assumptions* (**H1**)-(**H3**), (**Q4**) *and* (**Q5**) *we have for any* $x \in \Xi$

$$\left(\frac{n\phi_x(h_K)}{\Sigma^2(x,\xi_p(x))}\right)^{\frac{1}{2}} (\xi_{p,n}(x) - \xi_p(x)) \xrightarrow{D} \mathcal{N}(0,1) \quad as \quad n \to \infty, \tag{3.15}$$

where $\Sigma(x,\xi_p(x)) := \dfrac{\sigma(x,\xi_p(x))}{f(\xi_p(x)|x)}$.

3.4 Technical proofs and auxiliary result

3.4.1 Proof of Theorem 3.1

Proof of Lemma 3.1

i) Using assumption (**H1**) (iii) and integration by parts, we get

$$\begin{aligned}\frac{1}{\phi_x(h_K)}\mathbb{E}\left(K^l\left(\frac{d(x,X_i)}{h_K}\right)\right) &= \frac{1}{\phi_x(h_K)}\int_0^h K^l\left(\frac{u}{h_K}\right)dP^{d(x,X_1)}(u)\\ &= \frac{K^l(1)\phi_x(h_K)}{\phi_x(h_K)} - \frac{1}{\phi_x(h_K)}\int_0^1 (K(u))'\phi_x(uh_K)du\\ &= \left(K^l(1) - \int_0^1 (K(u))'\zeta_0^x(u)du\right)[g(x)+o(1)]\\ &= a_l^x g(x).\end{aligned}$$

When $l=1$, we conclude the result. ■

ii) Now, we turn to the second part of this lemma

$$\begin{aligned}\mathbb{E}(\Psi_n^{(j)}(x,y)) &= \frac{1}{nh_H^j\phi_x(h_K)}\sum_{i=1}^n \mathbb{E}\left(K\left(\frac{d(x,X_i)}{h_K}\right)H^{(j)}\left(\frac{y-Y_i}{h_H}\right)\right)\\ &= \frac{1}{h_H^j\phi_x(h_K)}\mathbb{E}\left(K\left(\frac{d(x,X_i)}{h_K}\right)\right)\mathbb{E}\left(H^{(j)}\left(\frac{y-Y_i}{h_H}|X_1\right)\right)\\ \mathbb{E}\left(H^{(j)}\left(\frac{y-Y_i}{h_H}|X_1\right)\right) &= \int_{\mathbb{R}} H\left(\frac{y-z}{h_H}\right)f(z|x)dz \quad (\text{ Integrating by parts})\\ &= \int_{\mathbb{R}} H'(t)[F(y-th_H|X_1) - F(y|x) + F(y|x)]dt\\ \mathbb{E}(\Psi_n(x,y)) &= \frac{1}{\phi_x(h_K)}\mathbb{E}\left(K\left(\frac{d(x,X_i)}{h_K}\right)\right)\int_{\mathbb{R}} H'(t)[F(y-th_H|X_1) - F(y|x)]\\ &\quad + \frac{F(y|x)\mathbb{E}\left(K\left(\frac{d(x,X_i)}{h_K}\right)\right)}{\phi_x(h_K)}\\ &= a_1^x F(y|x)g(x)\\ &= a_1^x\Psi(x,y).\end{aligned}$$

Where $\Psi(x,y) = F(y|x)g(x)$. ■

Now, we deal with $\sup_{y\in\mathcal{C}}|\Psi_n(x,y) - \mathbb{E}[\Psi_n(x,y)]|$, for $g_n(x) - \mathbb{E}(g_n(x))$ we use the same method with a slight difference.

Proof of Lemma 3.2

We give a decomposition then we deal with $j = 0$, but firstly we consider a cover of $\mathcal{C}$ by a finite number l_n of intervals $\mathcal{C}_k$ of the form $\mathcal{C}_k = (y_k - h_H^\eta, y_k + h_H^\eta)$ where $\eta > 1$, we have that $\mathcal{C} \subset \bigcup_{k=1}^{l_n} \mathcal{C}_k$, since $\mathcal{C}$ is bounded there exists $M > 0$ such that $l_n \leq M h_H^\eta$, clearly we have

$$\begin{aligned}\sup_{y\in\mathcal{C}} |\Psi_n(x,y) - \mathbb{E}[\Psi_n(x,y)]| &\leq \sup_{y\in\mathcal{C}} |\Psi_n(x,y) - \Psi_n(x,y_k)| \\ &\quad + \sup_{y\in\mathcal{C}} |\Psi_n(x,y_k) - \mathbb{E}[\Psi_n(x,y_k)]| \\ &\quad + \sup_{y\in\mathcal{C}} |\mathbb{E}[\Psi_n(x,y_k)] - \mathbb{E}[\Psi_n(x,y)]|.\end{aligned}$$

We deal with the first term using (**Q4**) (ii), we get

$$\begin{aligned}\sup_{y\in\mathcal{C}} |\Psi_n(x,y) - \Psi_n(x,y_k)| &\leq \frac{1}{n\phi_x(h_K)} \sup_{y\in\mathcal{C}} \sum_{i=1}^n |H\left(\frac{y-Y_i}{h_H}\right) - H\left(\frac{y_k-Y_i}{h_H}\right)| K\left(\frac{d(x,X_i)}{h_K}\right) \\ &\leq \sup_{y\in\mathcal{C}} \frac{C|y-y_k|}{h_H}\left(\frac{1}{n\phi_x(h_K)}\sum_{i=1}^n K\left(\frac{d(x,X_i)}{h_K}\right)\right) \\ &\leq C h_H^{\eta-1} g_n(x).\end{aligned}$$

We have also $g_n(x) = (g_n(x) - \mathbb{E}(g_n(x))) + \mathbb{E}(g_n(x))$ by Lemma 3.1 (i) the first term is bounded and $n\phi_x(h_K)$ goes to infinity, we conclude that the first difference converges almost completely to zero.
Now we deal with the term expressed by $\sup_{y\in\mathcal{C}} |\Psi_n(x,y_k) - \mathbb{E}[\Psi_n(x,y_k)]|$
As the Assumptions imply that laplace transforms of the real random variables $U_i = K_i(x)H_i(y)$, where U_i, $i > 0$ is strong mixing, exist on every interval $[-\delta, \delta]$. So, $\forall n \geq 2$, $\forall \gamma \geq 2$ and $\forall k \leq \frac{n}{2}$ we have

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n |U_i - \mathbb{E}(U_i)| > \varepsilon\right) \leq 6\exp\left(\frac{-ns\varepsilon}{30k}\right) + 6\frac{n}{k}\left(\frac{10D_\gamma}{\varepsilon}+1\right)^{\frac{2\gamma}{2\gamma+1}} (\alpha(k))^{\frac{2\gamma}{2\gamma+1}},$$

where

$$D_\gamma = \sup_{i\geq 1} \| U_i \|_\gamma, \quad s = \min\left(\frac{\delta}{2}, \frac{\varepsilon}{3c}\right), \quad c = 4\sup_{i\geq 1}\sum_{j=2}^\infty \frac{\delta^{l-2}}{j!}\mathbb{E}|U_i|.$$

We obtain

$$\begin{aligned}\mathbb{P}(|\Psi_n(x,y_k) - \mathbb{E}[\Psi_n(x,y_k)] > \varepsilon) &= \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n |U_i - \mathbb{E}(U_i)| > \varepsilon\phi_x(h)\right) \\ &\leq 6\exp\left(\frac{-ns\varepsilon}{30k}\right) + 6\frac{n}{k}\left(\frac{10D_\gamma}{\varepsilon}+1\right)^{\frac{2\gamma}{2\gamma+1}} (\alpha(k))^{\frac{2\gamma}{2\gamma+1}} \\ &=: J_1 + J_2.\end{aligned}$$

Firstly, we deal with J_1,
we know that $K(t) < K(0)$ and $H(t) \leq 1$, using (**Q1**) (i), then we get $D_\gamma \leq K(0)$, $\mathbb{E}|U_i|^j \leq (K(0))^j \phi_x(h_K)$ and for n large enough $s = \dfrac{\varepsilon\phi_x(h_K)}{3c}$, then we deduce that $-s \leq -C\varepsilon$, where $C = \dfrac{\delta^2}{12}\exp(-\delta K(0))$.
Now choosing $k = [(n\phi_x(h_K)^{\frac{1}{2}})]$, by (**Q5**) (i), we have

$$J_1 \leq \exp\left(\frac{-C\varepsilon}{30}(n\phi_x(h_K))^{\frac{1}{2}}\right) < \infty. \tag{3.16}$$

Let us turn to J_2
$J_2 \approx A_n = C\dfrac{n}{k}\left(\dfrac{1}{\phi_x(h_K)}\right)^\lambda (\alpha(k))^{2\lambda}$, for n large enough, where $C > 0$.
Now, use the definition of k such that $k = [(n\phi_x(h_K))^{\frac{1}{2}}]$, by an algebra we obtain $n > \dfrac{4k^2}{\phi_x(h_K)}$
As $k^2\phi_x(h_K) \to \infty$ then for n large enough , we get $A_n \leq Ck^{2\lambda+3}(\alpha(k))^{2\lambda}$.
Now we have

$$\sum_n k^{2\lambda+3}(\alpha(k))^{2\lambda} < \infty \quad \textit{then} \quad \sum_n J_2 < \infty. \tag{3.17}$$

We combine the results (3.16), (3.17) to conclude that $\sup_{y\in\mathcal{C}} |\Psi_n(x, y_k) - \mathbb{E}[\Psi_n(x, y_k)]|$ converges almost completely to zero.
Finally, we deal with $\sup_{y\in\mathcal{C}} |\mathbb{E}\Psi_n(x, y_k) - \mathbb{E}[\Psi_n(x, y)]|$.
Using (**Q1**) (ii) we get

$$\begin{aligned}
\sup_{y\in\mathcal{C}} |\mathbb{E}\Psi_n(x, y_k) - \mathbb{E}\Psi_n(x, y)| &\leq \frac{1}{n\phi_x(h_K)} \sup_{y\in\mathcal{C}} \sum_{i=1}^n \mathbb{E}\left(|H\left(\frac{y - Y_i}{h_H}\right) - H\left(\frac{y_k - Y_i}{h_H}\right)|K\left(\frac{d(x, X_i)}{h_K}\right)\right) \\
&\leq \sup_{y\in\mathcal{C}} \frac{C|y - y_k|}{h_H} \times \mathbb{E}\left(\frac{1}{n\phi_x(h_K)}\sum_{i=1}^n K\left(\frac{d(x, X_i)}{h_K}\right)\right) \\
&\leq Ch_H^{\eta-1}\mathbb{E}(g_n(x)).
\end{aligned}$$

We know that $\mathbb{E}(g_n(x))$ is bounded and $n\phi_x(h_K)$ goes to infinity, we conclude that the last term converges almost completely to zero and Lemma 3.2 is checked.
Using the previous results and Lemma 3.1 we conclude the proof of Theorem 3.1. ■

3.4.2 Proof of Theorem 3.2

This proof is based on the following idea:
As $F(.|x)$ is a distribution function with a unique quantile of order p then for any

$\varepsilon > 0$ let

$$\eta(\varepsilon) = \min\{F(\xi_p(x) + \varepsilon|x) - F(\xi_p(x)|x), F(\xi_p(x)|x) - F(\xi_p(x) - \varepsilon|x)\}.$$

Then

$$\forall\varepsilon_0, \forall y > 0, \quad |\xi_p(x) - y| > \varepsilon \Rightarrow |F(\xi_p(x)|x) - F(y(x)|x)| > \eta(\varepsilon).$$

Now, we use (3.4) and (3.6) we have

$$F(\xi_{p,n}(x)|x) - |F(\xi_p(x)|x)| \leq \sup_{y\in\mathcal{C}} |F_n(y(x)|x) - F(y(x)|x)|. \tag{3.18}$$

Making use (3.1) and the continuity of $F(.|x)$, we obtain

$$\sum_n^\infty \mathbb{P}(\xi_{p,n}(x) - \xi_p(x) \geq \varepsilon) \leq \sum_n^\infty \mathbb{P}(\sup_{y\in\mathcal{C}} |F_n(y(x)|x) - F(y(x)|x)| \geq \eta(\varepsilon)).$$ ■

3.4.3 Proof of Theorem 3.3

Proof of Lemma 3.3

Using the definition of the conditional variance

$$Var\left[H\left(\frac{y-Y_1}{h_H}|X_1\right)\right] = \mathbb{E}\left[H^2\left(\frac{y-Y_1}{h_H}|X_1\right)\right] - \left[\mathbb{E}\left[H\frac{y-Y_1}{h_H}|X_1\right]\right]^2$$

$$\begin{aligned}
\mathbb{E}\left(H^2\left(\frac{y-Y_1}{h_H}|X_1\right)\right) &= \int_{\mathbb{R}} H^2\left(\frac{y-z}{h_H}\right) f(z|X_1)dz \\
&= \int_{\mathbb{R}} H^2\left(\frac{y-z}{h_H}\right) dF(y - th_H|X_1)dt \quad \text{(integrating by parts)} \\
&= \int_{\mathbb{R}} 2H'(t)H(t)[F(y - th_H|X_1) - F(y|x)]dt + \int_{\mathbb{R}} 2H'(t)H(t)F(y|x)dt \\
&= F(y|x) \quad \text{as} \quad n \to \infty.
\end{aligned}$$

Because $\int_{\mathbb{R}} 2H'(t)H(t)F(y|x)dt = F(y|x)$ and using (**Q3**) we obtain the result. Now the second term on the right hand side of this last equality tends to $F^2(y|x)$ as n tends to infinity, which completes the proof of the lemma. ■

Proof of Lemma 3.4

We begin by the proof of (3.13)

$$\begin{aligned}
V_n(x,y) &= \frac{1}{\phi_x(h_K)} Var(N_1(x,y)) \\
&= \frac{1}{\phi_x(h_K)} \mathbb{E}(N_1^2(x,y)) \\
&= \frac{1}{\phi_x(h_K)} \mathbb{E}[(H_i(y) - F(y|x))^2 K_i^2(x)] - \frac{1}{\phi_x(h_K)} [\mathbb{E}[(H_i(y) - F(y|x))^2 K_i^2(x)]]^2 \\
&=: T_1 + T_2.
\end{aligned}$$

Then

$$\begin{aligned}
T_2 &= \phi_x(h_K)[\frac{1}{\phi_x(h_K)} \mathbb{E}[(H_i(y) - F(y|x))^2 K_i^2(x)]]^2 \\
&= \phi_x(h_K)[\mathbb{E}[\Psi_n(x,y)] - F(y|x)\mathbb{E}[g_n(x)]^2].
\end{aligned}$$

Making use of (3.1) the last term goes to zero as n goes to infinity.
Now we turn to T_1

$$\begin{aligned}
T_1 &= \frac{1}{\phi_x(h_K)} \mathbb{E}[\mathbb{E}[(H_i(y) - F(y|x))^2 | X_1] K_i^2(x)] \\
&= \frac{1}{\phi_x(h_K)} \mathbb{E}[Var(H_i(y)|X_1) K_i^2(x)] \\
&\quad + \frac{1}{\phi_x(h_K)} \mathbb{E}[[\mathbb{E}(H_i(y)|X_1) - F(y|x))]^2 K_i^2(x)].
\end{aligned}$$

Using (3.9), we have

$$\frac{1}{\phi_x(h_K)} \mathbb{E}[[\mathbb{E}(H_i(y)|X_1) - F(y|x))]^2 K_i^2(x)] \to 0 \quad \text{as} \quad n \to \infty.$$

Using also (3.3), we have

$$\frac{1}{\phi_x(h_K)} \mathbb{E}[Var(H_i(y)|X_1) K_i^2(x)] \to V(x,y) \quad \text{as} \quad n \to \infty.$$

■

Now, we deal with (3.14), and to prove this part we have to split the sum as follows:

$$\begin{aligned}
\frac{1}{n\phi_x(h_K)} \sum_i^n \sum_j^n Cov(N_i, N_j) &= \frac{1}{n\phi_x(h_K)} \sum_{0<|i-j|\le m_n}^n Cov(N_i, N_j) \\
&\quad + \frac{1}{n\phi_x(h_K)} \sum_{|i-j|\ge m_n} Cov(N_i, N_j) \\
&=: L_1 + L_2,
\end{aligned}$$

where $m_n = o(n)$ at a rate specified below, for L_1 we have

$$|Cov(N_i, N_j)| = \mathbb{E}(N_i, N_i) = \mathbb{E}[(H_i(y) - F(y|x))(H_j(y) - F(y|x))K_i(x)K_j(x)]$$

Conditioning on (X_i, X_j), we have

$$|Cov(N_i, N_j)| = \mathbb{E}[\mathbb{E}[(H_i(y) - F(y|x))(H_j(y) - F(y|x))|(X_i, X_j)]K_i(x)K_j(x)]$$

By the assumption made on $H(.)$ and $F(.|x)$ we have $|H_i(y) - F(y|x)| \leq 1$, it follows that

$$\begin{aligned} |Cov(N_i, N_j)| &\leq \mathbb{E}[K_i(x), K_j(x)] \\ &\leq C\mathbb{P}[(X_i, X_i) \in B(x, h_K) \times B(x, h_K)]. \end{aligned}$$

Using (**H1**) (iii) we obtain

$$Cov(N_i, N_j) \leq C\Gamma_x(h_K)$$

Therefore

$$\begin{aligned} L_1 &\leq \frac{1}{n\phi_x(h_K)}[C\Gamma_x(h_K)]nm_n \\ &\leq Cm_n\phi_x(h_K)\left[\frac{\Gamma_x(h_K)}{\phi_x^2(h_K)}\right]. \end{aligned}$$

We choose m_n such that $m_n\phi_x(h_K)$ goes to zero as n goes to infinity.
Making use of (3.1) (i) and $\dfrac{\Gamma_x(h_K)}{\phi_x^2(h_K)}$ is bounded, we get $L_1 = o(1)$.
We turn to L_2, by *Davydov's* Inequality, we get

$$|Cov(N_i, N_i)| \leq 8[\mathbb{E}[|(H_i(y) - F(y|x))K_i(x)|^v]]^{\frac{2}{v}}[\alpha(|i-j|)]^{1-\frac{2}{v}}.$$

Conditioning on X_i and $|H_i(y) - F(y|x)| \leq 1$, we get

$$\mathbb{E}[|(H_i(y) - F(y|x))K_i(x)|^v] \leq C\phi_x(h_K)$$

Therefore

$$\begin{aligned} |Cov(N_i, N_i)| &\leq C[\phi_x(h_K)]^{\frac{2}{v}}[\alpha(|i-j|)]^{1-\frac{2}{v}} \\ &\leq \frac{C[\phi_x(h_K)]^{\frac{2}{v}}}{(m_n)^\delta}|i-j|^\delta[\alpha(|i-j|)]^{1-\frac{2}{v}}. \end{aligned}$$

Thus

$$\begin{aligned} L_2 &\leq \frac{C}{n(m)^{\delta}[\phi_x(h_K)]^{1-\frac{2}{v}}} \sum_{|i-j|>m_n}^{n} |i-j|^{\delta}[\alpha(|i-j|)]^{1-\frac{2}{v}} \\ &\leq \frac{C}{(m_n)^{\delta}[\phi_x(h_K)]^{1-\frac{2}{v}}} \sum_{k>m_n}^{n} k^{\delta}[\alpha(k)]^{1-\frac{2}{v}}. \end{aligned}$$

We choose m_n such that $m_n = [\phi_x(h_K)]^{-\frac{1-2/v}{\delta}}$.

Finally, we use (**Q4**) to show that $L_2 = o(1)$, which complete the proofs ■

Now we prove the asymptotic normality of $F_n(y|x)$ and $\xi_{p,n}(x)$ making use of the decomposition of $F_n(y|x) - F(y|x)$.

So, the next results prove that $(n\phi_x(h_K))^{\frac{1}{2}}R_n(x,y)$ is asymptotically normal and $(n\phi_x(h_K))^{\frac{1}{2}}B_n(x,y)$ converges to zero in probability .

Proof of Lemma 3.5

To prove this lemma we need to introduce the following modification

We normalize N_i: $\widetilde{N_i} := \dfrac{N_i}{\sqrt{\phi_x(h_K)}}$, then $Var(\tilde{N}_i) = V_n(x,y)$.

So, we can write

$$Var(\widetilde{N_i}) \to V(x,y) \quad \text{as } n \text{ goes to infinity,}$$

we have also

$$\sum_{|i-j|>0}^{n} |Cov(\widetilde{N_i},\widetilde{N_j})| = \frac{1}{\phi_x(h)} \sum_{|i-j|>0}^{n} |Cov(N_i,N_j)| = o(n). \tag{3.19}$$

We can write

$$(n\phi_x(h_K))^{\frac{1}{2}}R_n(x,y) =: \frac{1}{\sqrt{n}}S_n,$$

because $S_n = \sum_{i=1}^{n} \widetilde{N_i}$, now we have to show that

$$\frac{1}{\sqrt{n}}S_n \xrightarrow{D} \mathcal{N}(0, V(x,y)) \quad \text{as} \quad n \to \infty. \tag{3.20}$$

Now, we want to prove the asymptotic normality for S_n of dependent variables, make use of *Doob's* Technique which uses the *Bernstein's* large-block and small-block procedure. Partition of the set $[1,\ldots,n]$ into $2k_n+1$ subsets with large-blocks of size $u = u_n$ and small-block $v = v_n$ with $k := k_n = \left\lfloor \dfrac{n}{u_n + v_n} \right\rfloor$.

Let v_n be the sequence described in Assumption (**Q5**) (iii) and the same hypothesis implies that there exists a sequence of positive numbers $q_n \to \infty$ such that:

$$q_n v_n = o((n\phi_x(h_K))^{\frac{1}{2}}) \quad and \quad q_n \left(\frac{n}{\phi_x(h_K)}\right)^{\frac{1}{2}} \alpha(v_n) \to 0 \quad \text{as} n \text{ goes to infinity.} \tag{3.21}$$

We define the large-block size as $u_n = \left\lfloor \left(\frac{n\phi_x(h_K)}{q_n}\right)^{\frac{1}{2}} \right\rfloor$, with simple algebra when n going to infinity we get

$$\frac{v_n}{u_n} \to 0, \quad \frac{u_n}{n} \to 0, \quad \frac{u_n}{(n\phi_x(h_K))^{\frac{1}{2}}} \to 0 \quad and \quad \frac{n}{u_n}\alpha(v_n) \to 0.$$

Now we split S_n as follows

$$S_n(x,y) = S_n = \sum_{j=1}^{k-1} Z_j^{'} + \sum_{j=1}^{k-1} Z_j^{''} + Z_j^{'''} = S_n^{'} + S_n^{''} + S_n^{'''}.$$

Where

$$Z_j^{'} = \sum_{i=t_j}^{t_j+u-1} \widetilde{N}_i \quad 0 \le j \le k-1,$$

$$Z_j^{''} = \sum_{i=t_j+u}^{t_j+u+v-1} \widetilde{N}_i \quad 0 \le j \le k-1,$$

$$Z_j^{'''} = \sum_{i=t_k}^{n} \widetilde{N}_i \quad 0 \le j \le k-1.$$

Where $t_j = j(u+v)+1$.

Now we present some claims to obtain the result.

$$(i) \quad \frac{1}{n}\mathbb{E}[S_n^{''}] \to 0, \qquad (ii) \quad \frac{1}{n}\mathbb{E}[S_n^{'''}] \to 0, \tag{3.22}$$

For this claim, we show that the first part (i) and second part (ii) are negligible as n tends to infinity.

We begin by the first part (i)

$$\begin{aligned}
\mathbb{E}[S_n^{''}]^2 &= \sum_{j=1}^{k-1} Var[Z_j^{'}] \\
&= \sum_{j=1}^{k-1} Var(Z_j^{''}) + \sum_{|i-j|>0}^{k-1} Cov(Z_i^{''}, Z_j^{''}) \\
&=: A_1 + A_2.
\end{aligned}$$

By stationarity and (3.19), we obtain

$$\begin{aligned} Var(Z_j^{''}) &= Var[\sum_{i=t_j+u}^{t_j+u+v-1} \tilde{N}_1] \\ &= v_n Var(\tilde{N}_1) + \sum_{|i-j|>0}^{v_n} Cov(\tilde{N}_i, \tilde{N}_i) \\ &= v_n Var(\tilde{N}_1) + o(v_n), \end{aligned}$$

thus

$$A_1 = \sum_{j=1}^{k-1} Var(Z_j^{''}) = k_n v_n Var(\widetilde{N}_1) + k_n o(v_n).$$

We know that

$$k_n v_n \approx \frac{n v_n}{u_n + v_n} \approx \frac{n v_n}{u_n} = o(n).$$

This show that $A_1 = o(n)$. Now, we turn to A_2

$$A_2 = \sum_{|i-j|>0}^{k-1} Cov(Z_i^{''}, Z_j^{''}) = \sum_{|i-j|>0}^{k-1} \sum_{l_1=1}^{v_n} \sum_{l_2=1}^{v_n} Cov(\tilde{N}_{\mu_i}, \widetilde{N}_{\mu_j}),$$

where $\mu_i = t_j + u_n + l_1$. Since for $i \neq j$ we have
$|(t_j + u_n + l_1) - (t_j + u_n + l_2) = t_j - t_i + l_1 + l_2 \geq u_n$, it follows that

$$A_2 \leq \sum_{|i-j|>u_n}^{n} Cov(\widetilde{N}_i, \widetilde{N}_j) = o(n).$$

For the second part (ii)

$$\begin{aligned} \frac{1}{n}\mathbb{E}[S_n^{'''}]^2 &\leq \frac{1}{n}((n - t_k + 1)Var(\widetilde{N}_1)) + \frac{1}{n}\sum_{|i-j|>0}^{n-t_k+1} Cov(\widetilde{N}_i, \widetilde{N}_j) \\ &\leq \frac{1}{n}((u_n + v_n)Var(\widetilde{N}_1)) + \frac{1}{n}\sum_{|i-j|>0}^{n} Cov(\widetilde{N}_i, \widetilde{N}_j). \end{aligned}$$

Using (3.19) and the definition of u_n and v_n, we get the result which completes the proof of (3.22). ■

$$|\mathbb{E}[(\exp(itn^{-\frac{1}{2}}S_n^{'}) - \prod_{j=0}^{k-1} \mathbb{E}[\exp(itn^{-\frac{1}{2}}Z_j^{'})]| \rightarrow 0, \tag{3.23}$$

Now, we prove (3.23) using the *Volkonskii and Rosanov's* Lemma to show the asymptotically independence of $S_n^{'}$ in $Z_j^{'}$ Note that $Z_p^{'}$ is $\mathcal{F}_{i_p}^{i_p}$-measurable with $i_p = p(u+v)+1$, and $j_p = p(u+v)+u$ and $U_i = \exp(itn^{-\frac{1}{2}}S_n^{'})$ we have

$$|\mathbb{E}[(\exp(itn^{-\frac{1}{2}}S_n^{'}) - \prod_{j=0}^{k-1} \mathbb{E}[\exp(itn^{-\frac{1}{2}}Z_j^{'})]| \leq 16 k_n \alpha(v_n+1) \approx 16\frac{n}{u_n}\alpha(v_n+1),$$

tends to zero, because $\frac{n}{u_n}\alpha(v_n) \to 0$. ■

The following two claims are concerned by the conditions of *lindeberg-Feller* Theorem.

$$\frac{1}{n}\sum_{j=1}^{k-1}\mathbb{E}[Z_j^{'}]^2 \to 0, \tag{3.24}$$

Let us turn to (3.25), we replace u_n by v_n, we have $Var(Z_j^{(1)}) = u_n Var(\tilde{N}_1) + o(u_n)$, hence

$$\frac{1}{n}\sum_{j=1}^{k-1}\mathbb{E}[Z_{j=1}^{'}]^2 = \frac{k_n u_n}{n}Var(\widetilde{N}_1) + \frac{k_n}{n}o(u_n),$$

since $\frac{k_n u_n}{n} \to 1$ and $Var(\tilde{N}_1) \to V(x,y)$. ■

Finally, it remains to establish (3.25)

$$\frac{1}{n}\sum_{j=1}^{k-1}\mathbb{E}[(Z_j^{'})^2 1\!\!1_{\{|Z_j^{'}|>\varepsilon\sqrt{nV(x,y)}\}}]. \tag{3.25}$$

Now we have to show that $\{|Z_j^{'}| > \varepsilon\sqrt{nV(x,y)}\}$ is empty for n large enough.

$$\begin{aligned} |Z_j^{'}| &\leq \frac{u_n|N_j|}{\sqrt{\phi_x(h_K)}} \\ &\leq \frac{K(0)u_n}{\sqrt{\phi_x(h_K)}}. \end{aligned}$$

Therefore

$$\frac{1}{n}|Z_j^{'}| \leq \frac{K(0)u_n}{\sqrt{n\phi_x(h_K)}}.$$

Because $|H_i(y) - F(y|x)| \leq 1$ and $K_i(x) < K(0)$. This show that $\{|Z_j^{'}| > \varepsilon\sqrt{nV(x,y)}\}$ is empty for n large enough which complete the proof. ■

Proof of Lemma 3.6

We have

$$(n\phi_x(h_K))^{\frac{1}{2}}B_n(x,y) = \frac{(n\phi_x(h_K))^{\frac{1}{2}}}{g_n(x)}[a_1^x F(y|x) - \mathbb{E}(\Psi_n(x,y)) - F(y|x)[a_1^x - \mathbb{E}(g_n(x))]].$$

Using Lemmas 2.3 and 3.1, we have

$$[a_1^x F(y|x) - \mathbb{E}(\Psi_n(x,y)) - F(y|x)[a_1^x - \mathbb{E}(g_n(x))]] \to 0 \quad \text{as} \quad n \to \infty.$$

On the other hand

$$\frac{(n\phi_x(h_K))^{\frac{1}{2}}}{g_n(x)} = \frac{(n\phi_x(h_K))^{\frac{1}{2}} f_n(y|x)}{g_n(x) f_n(y|x)} = \frac{(n\phi_x(h_K))^{\frac{1}{2}} f_n(y|x)}{\Psi'_n(x,y)}$$

Then we use Theorem 3.1, it suffices to show that $\dfrac{(n\phi_x(h_K))^{\frac{1}{2}}}{\Psi'_n(x,y)}$ tends to zero as n goes to infinity.

Because $K(.)H'(.)$ is continuous with support on $[0,1]$ then by the last part of (**Q4**) (ii) it follows that

$$\Psi'_n(x,y) \geq \frac{m}{h_H \phi_x(h_K)},$$

then

$$\frac{(n\phi_x(h_K))^{\frac{1}{2}}}{\Psi'_n(x,y)} \leq \frac{(nh_H^2 \phi_x^3(h_K))^{\frac{1}{2}}}{m}.$$

Finally, using (**Q5**) (ii) completes the proof of Lemma 3.6. ∎

Now, to prove Theorem 3.3, we apply the Lemmas 3.4, 3.5 and 3.6 to obtain the claimed result.

3.4.4 Proof of Theorem 3.4

Using (3.8) and (3.12), we get:

$$\begin{aligned}
(n\phi_x(h_K))^{\frac{1}{2}}(\xi_p(x) - \xi_{p,n}(x)) &= (n\phi_x(h_K))^{\frac{1}{2}} \frac{F_n(\xi_p(x)) - F(\xi_p(x))}{f_n(\xi^*_{p,n}(x)|x)} \\
&= \frac{(n\phi_x(h_K))^{\frac{1}{2}} R_n(x,y)}{f_n(\xi^*_{p,n}(x)|x)} + \frac{(n\phi_x(h))^{\frac{1}{2}} B_n(x,y)}{f_n(\xi^*_{p,n}(x)|x)}.
\end{aligned}$$

Combining Theorem 3.1 for $j = 1$, with Theorems 3.2 and 3.3 and the corollary 9, we obtain the result. ∎

Chapter 4

Application and conclusion

This chapter is devoted to the implementation of a functional non-parametric prediction method. (We deal with conditional quantile). In the first section we focus on the building of confidence bands then we illustrate an application to prediction. Secondly we apply our methodology on simulated data in order to show the utility of the asymptotic normality in the determination of the confidence bands, then we test the performance of the conditional quantile estimator via an application to a real data study (*El Nino* data). In the second and the last section, we conclude this dissertation by a general conclusion.

4.1 Application

4.1.1 Confidence bands

It is worth to note that the conditional quantile method can be used for prediction confidence band construction. As we know the central limit theorem is usually used in the determination of the confidence bands for the estimates. In non-parametric estimation, the asymptotic variance depends on certain unknown functions. Here, we have

$$\Sigma^2(x, \xi_p(x)) = \frac{a_2^x P(1-P)}{(a_1^x)^2 f_n^2(\xi_p(x)|x)},$$

where $f(y|x)$, $\xi_p(x)$, a_1^x and a_2^x are unknown a priori and have to be estimated in practice. Then we can obtain a confidence band even if $\Sigma^2(x, y)$ is functionally specified.

The normalization constants a_l^x for $l = 1, 2$ can be estimated by

$$a_{l,n}^x = \frac{1}{n\widehat{\phi}_x(h_K)} \sum_{i=1}^{n} K^l\left(\frac{d(x, X_i)}{h_K}\right),$$

where $\widehat{\phi}_x(h_K) = \dfrac{1}{n}\sum_{i=1}^{n} 1\!\!1_{\{W_i \le x\}}$.

Now, a plug-in estimate for the asymptotic standard deviation $\Sigma(x, \xi_p(x))$, can be obtained using the estimators $\xi_{p,n}(.)$, $f_n(.|.)$ and $a_{l,n}^x$ of $\xi_p(x)$, $f(.|.)$ and a_l^x respectively, that is

$$\Sigma_n^2(x, \xi_{p,n}(x)) = \frac{a_{2,n}^x P(1-P)}{(a_{1,n}^x)^2 f_n^2(\xi_{p,n}(x)|x)}.$$

We have

$$a_{l,n}^x = \frac{\phi_x(h_K)}{\widehat{\phi}_x(h_K)} \times \frac{1}{n\phi_x(h_K)} \sum_{i=1}^{n} K^l\left(\frac{d(x, X_i)}{h_K}\right).$$

By *Glivenko-Cantelli*[1] type result the first ratio goes to 1 in probability as n goes to infinity

$$\frac{\phi_x(h_K)}{\widehat{\phi}_x(h_K)} \xrightarrow{P} 1$$

For $l = 1$ Lemma 3.1 shows that

$$a_{l,n}^x \xrightarrow{P} a_1^x.$$

In the same manner, we show that $a_{l,n}^x$ converges in probability to a_2^x for $l = 2$.
Theorem 3.1 gives us that $f_n(.|x)$ converges almost completely to $f(.|x)$, then in probability. Now Theorem 3.2 allows us to conclude that $\Sigma_n(x, \xi_{p,n}(x))$ converges in probability to $\Sigma(x, \xi_p(x))$, which leads us to use an asymptotic approximation provided by the following Corollary where :

Corollary 10 *Under assumptions* **(H1)**-**(H3)**, **(Q4)** *and* **(Q5)**

$$\left(\frac{n\widehat{\phi}_x(h_K)}{\Sigma_n^2(x, \xi_{p,n}(x))}\right)^{\frac{1}{2}} (\xi_{p,n}(x) - \xi_p(x)) \xrightarrow{D} \mathcal{N}(0,1) \quad \text{as } n \to \infty$$

This corollary permits us to built $(1-\zeta)$ confidence bands for $\xi_p(x)$ which is given by

$$-u_{1-\frac{\zeta}{2}} \le \left(\frac{n\widehat{\phi}_x(h_K)}{\Sigma_n^2(x, \xi_{p,n}(x))}\right)^{\frac{1}{2}} (\xi_{p,n}(x) - \xi_p(x)) \le u_{1-\frac{\zeta}{2}},$$

[1] F_n uniformly convergent to F, where F is a distribution function $\quad \mathbb{P}(\lim_n ||F_n - F||_\infty = 0) = 1$

then

$$\xi_{p,n}(x) - u_{1-\frac{\zeta}{2}} \left(\frac{\Sigma_n^2(x, \xi_{p,n}(x))}{n\widehat{\phi}_x(h_K)} \right)^{\frac{1}{2}} \leq \xi_p(x) \leq \xi_{p,n}(x) + u_{1-\frac{\zeta}{2}} \left(\frac{\Sigma_n^2(x, \xi_{p,n}(x))}{n\widehat{\phi}_x(h_K)} \right)^{\frac{1}{2}}.$$

Finally we can represent the confidence bands as follows

$$CB = \left[\xi_{p,n}(x) - u_{1-\frac{\zeta}{2}} \left(\frac{\Sigma_n^2(x, \xi_{p,n}(x))}{n\widehat{\phi}_x(h_K)} \right)^{\frac{1}{2}} , \; \xi_{p,n}(x) + u_{1-\frac{\zeta}{2}} \left(\frac{\Sigma_n^2(x, \xi_{p,n}(x))}{n\widehat{\phi}_x(h_K)} \right)^{\frac{1}{2}} \right],$$

where $u_{1-\frac{\zeta}{2}}$ denotes the $(1-\zeta)$ quantile of the standard normal distribution.

4.1.2 Application to prediction

The term quantile is synonymous with percentile, the median is the best-known example of a quantile. We know that the sample median can be defined as the middle value (or the value half-way between the two middle values).
In prediction problem, the use of the conditional median $\mu(x)$ is good alternative to standard method based on the conditional mean for its robustness. Note that the estimation of $\mu(x)$ is given by $\xi_{\frac{1}{2},n}(x)$.
For each $n \in \mathbb{N}^*$, let $X_i(t),\ i = 1, \ldots, n$ be functional random variables with $t \in \mathbb{R}$. For each curve $X_i(t)$, we have a real response variable Y_i which is corresponding to some modality of our problem. Now, we say that we can predict the corresponding response variable $y_{new} = y_{n+1}$ given a new curve $x_{new} = x_{n+1}$.
The predictor estimator is obtained by calculating the quantity:

$$y_{new} = \mu_n(x_{new}) = \xi_{\frac{1}{2},n}(x_{new})$$

applying Theorem 3.4, we have

Corollary 11 *Under assumptions* (**H1**)-(**H3**), (**Q4**) *and* (**Q5**)

$$\left(\frac{n\widehat{\phi}_x(h_K)}{\Sigma^2(x_{new}, \xi_{\frac{1}{2}}(x_{new}))} \right)^{\frac{1}{2}} (\xi_{\frac{1}{2},n}(x_{new}) - \xi_{\frac{1}{2}}(x_{new})) \xrightarrow{D} \mathcal{N}(0,1) \quad as\ n \to \infty.$$

4.1.3 Simulation study

In this section we deal with simulated data then with real data to show how our estimator behaves in α-mixing contexts through asymptotic normality and confidence bands.

Simulated data

We now see the role of the asymptotic normality in the determination of confidence bands, we consider the classical non-parametric functional regression model

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

where ε is α-mixing generated by the following model:

$$\varepsilon_i = \frac{1}{\sqrt{2}}(\varepsilon_{i-1} + \eta_i) \quad i = 1, \dots, 200$$

with η_i and ε_0 are independent and generated as $\mathcal{N}(0,1)$.
For our functional data, we consider two diffusion processes on the interval $[0,1]$

$$Z_1(t) = 2 - cos(\pi t W) \quad \text{and} \quad Z_2(t) = cos(\pi t W),$$

where W is also normally distributed $\mathcal{N}(0,1)$. We take $X(t) = AZ_1(t) + (1-A)Z_2(t)$, where A is a random variable Bernoulli distributed. We carried out a 200 sample simulation of the curve X which is represented by the following graph 4.1.

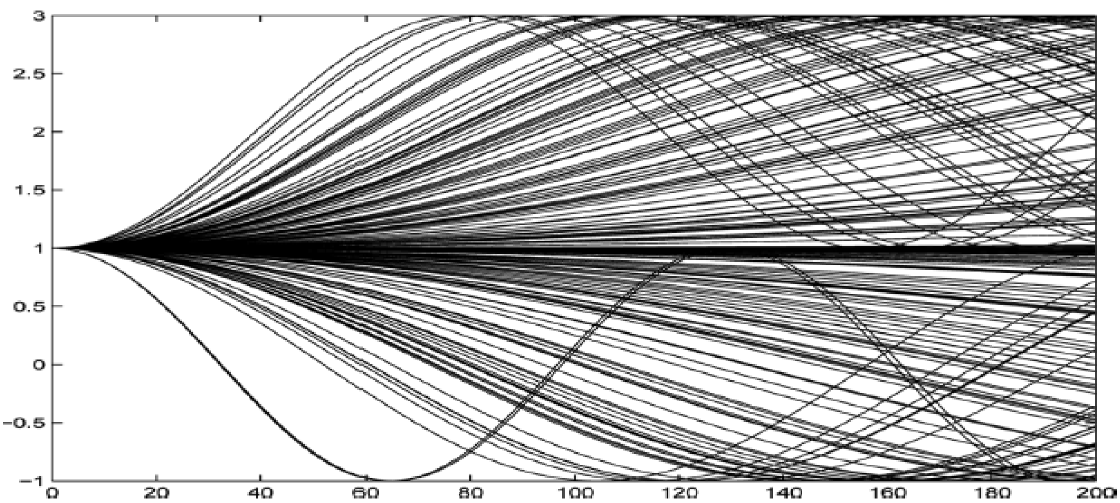

Figure 4.1: curves of $X_i(t)$ $i = 1, \dots, 200$ $t = 1, \dots, 100$.

Here we choose $R(X) = Ar_1(X) + (1-A)r_2(X)$ where r_1 (resp. r_2) is the nonlinear regression model $r_1(X) = 0.25 \times \int_0^1 (X'(t))^2 dt$ $(resp.\ r_2(X))$ is the null function. The quadratic kernel choice is given by

$$K(x) = \frac{3}{2}(1-x^2)1\!\!1_{[0,1)}$$

and the distribution function $H(.)$ is defined by

$$H(x) = \int_{-\infty}^{x} \frac{3}{4}(1-t^2)1\!\!1_{[-1,1]}(t)dt.$$

We ensure the good behavior of our method by using a norm well adapted to the kind of the data we have to deal with. Because of the regularity of the curves we choose the norm defined by the L^2 distance between the seconde derivatives of the curves.

In this simulation, the optimal bandwidth is obtained by minimizing the mean square error between the predicted and the true values over a set of known bandwidths values.

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

- We split our data into randomly chosen subsets: $(X_j, Y_j)_{j\in J}$ training sample, $(X_i, Y_i)_{i\in I}$ test sample.

- We use the training sample to calculate the estimator $\xi_{\frac{1}{2}}(X_j)$ for all $j \in J$.

- We set $i_* = arg \min_{j\in J} d(X_i, X_j)$, for each X_i in the test sample.

- For all $i \in I$ we define confidence bands by

$$\left[\xi_{\frac{1}{2},n}(X_{i_*}) - u_{0.975}\left(\frac{\Sigma_n^2(X_{i_*}, \xi_{\frac{1}{2},n}(x))}{J\hat{\phi}_{X_{i_*}}(h_K)}\right)^{\frac{1}{2}},\ \xi_{\frac{1}{2},n}(X_{i_*}) + u_{0.975}\left(\frac{\Sigma_n^2(X_{i_*}, \xi_{\frac{1}{2},n}(x))}{J\hat{\phi}_{X_{i_*}}(h_K)}\right)^{\frac{1}{2}}\right],$$

  where $u_{0.975}$ denotes 5 percent quantile of the standard normal distribution.

- Finally, we present our results in graph 4.2.

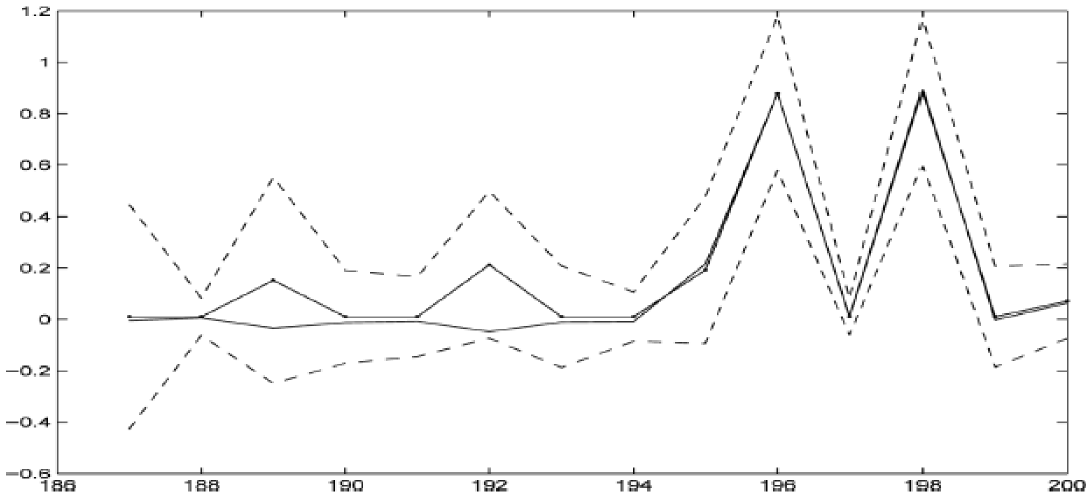

Figure 4.2: the extremities of the predicted values, the true values and the confidence bands.

Real data study

This study deals with a functional approach for time-series forecasting, based on the splitting of the observed time series into several continuous functional data.
We start by describing the *El Nino* time series, then we show how it can be viewed as a set of functional dependent variables. Thereafter, we will explain how a forecasting method can be built from the estimation of the conditional quantiles, Finally, we will show how this forecasting approach behaves on the real *El Nino* data-set.
El Nino data : The aim of this part is to apply the conditional quantile estimator on real functional data. Our study concerns the monthly times series of the Sea Surface Temperature (SST) from December 1950 up to November 2004. This dataset is a part of the original one which is available on line [2]. These temperatures are measured by moored buyos in the *El Nino* region defined by the coordinates south and west. Our (SST) time-series comes from the average of the monthly temperatures over the moored buoys in this area. Finally, the statistical sample is of size 648. The graphical display is given in graph 4.3.
This time series (z_i) $i = 1, \ldots, 648$ is loaded into a 54×12 matrix.

[2]http://www.cpc.ncep.noaa.gov/data/indices/

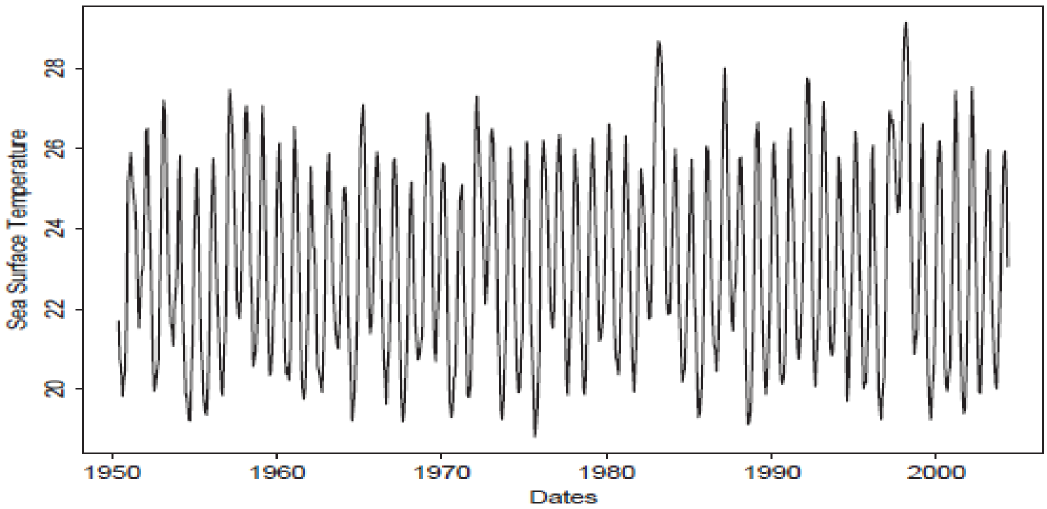

Figure 4.3: El Nino monthly Sea Surface Temperature.

Splitting El Nino : We split time-series into functional data. A useful way to display such a time series consists in cutting it into 54 pieces or 54 annual curves (see graph 4.4). More precisely, let $\{Z_{(k)}\}$ $k = 1, \ldots, 648$ be our *El Nino* time-series. We can build, for $i = 1, \ldots, 54$, the following subsequences:

$$\forall t\{1, 2, \ldots, 12\}, \qquad , \quad Z_i(t) = Z(12 \times (i-1) + t),$$

$Z_i = (Z_i(1), \ldots, Z_i(12))$ corresponding to the variations of the (SST) at the i^{th} year.

Because the climatic phenomenon is changing continuously over time, there is evidence for considering each annual curve as a continuous path. Of course, this continuous yearly curve will be observed only at some discretized points (here, at 12 discretized points). Finally, the time series can be viewed as a sample of 54 dependent functional data, namely $Z_1, \ldots, Z_{54}$. the main advantage of using such continuous path for the past of the time series is to be unsensitive to the curse of dimensionality.

Our approach is able to capture much information in the past of the time series, but still using for the past a single continuous object and avoiding the dimensionality effects.

As the kernel K does not affect the quality of the estimation in the mean square

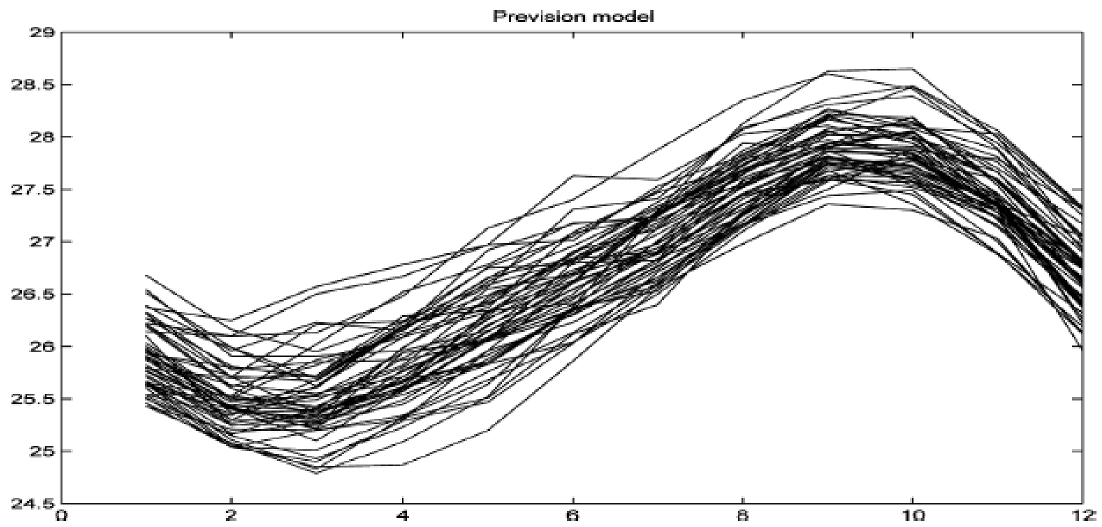

Figure 4.4: The 54 curves.

error (MSE) criterion, here we choice the quadratic kernel:

$$K(x) = \frac{3}{2}(1 - x^2)\mathbb{1}_{[0,1)}$$

and the distribution function $H(.)$ is defined by

$$H(x) = \int_{-\infty}^{x} \frac{3}{4}(1 - u^2)\mathbb{1}_{[-1,1]}(u)du$$

Another important point for ensuring a good behavior of the method, is to use a semi-metric that is well adapted to the kind of data we have to deal with. Here we used some semi-metric based on the q first term of the functional principal components analysis (FPCA) of the data. Indeed we do not choose a semi-metric at the beginning of the study but only a family of semi-metrics, and the key question is more to select the best semi-metric inside of the family than to choose the family itself. The key parameter is the order q of the (FPCA) expansion [3], which should also be chosen in a data-driven way.

Prediction procedure : We have at hand a set of 54 functional data. However, to show the performance of our method we will ignore the 54th year and we will predict it from the 53's ones. We will build our statistical method only on the 52

[3] Other choices of semi-metrics are possible, like for instance those based on some L^2 errors between higher order derivatives of the curves, and in this case the key parameter would be the order of the derivative.

previous data, the 53rd being used as a learning step to select the parameters using cross validation. $(X_i, X_{i+1}(j))_{(i=1,\ldots,52)}$ training sample.
Then, for the curve X_{54} we set $i_* = arg \min_{i=1,\ldots,53} d(X_{54}, X_i)$.
Thus, we calculate our predictor by $\widehat{X}_{54}(j) = \xi^j_{\frac{1}{2},54}(X_{i_*})$. Moreover we define the confidence bands

$$\left[\widehat{X}_{54}(j) - u_{0.975}\left(\frac{\Sigma_n^2(X_{i_*}, \xi_{\frac{1}{2},54}(X_{i_*}))}{53\hat{\phi}_{X_{i_*}}(h_K)}\right)^{\frac{1}{2}}, \ (\widehat{X}_{54}(j)) + u_{0.975}\left(\frac{\Sigma_n^2(X_{i_*}, \xi_{\frac{1}{2},54}(X_{i_*}))}{53\hat{\phi}_{X_{i_*}}(h_K)}\right)^{\frac{1}{2}}\right],$$

where $u_{0.975}$ denotes 5 percent quantile of the standard normal distribution.

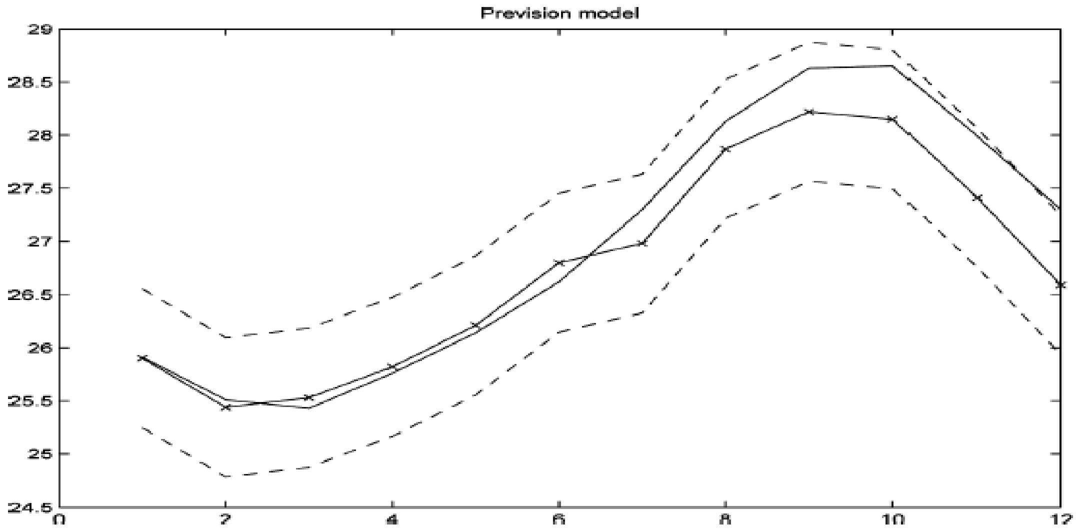

Figure 4.5: The extremities of the predicted values versus the true values and the confidence bands.

Finally, we present our results in the graph 4.5.

Real data study with R

Entering and organizing El Nino:
dataset ELNINODAT <- as.matrix(read.table("npfda-elnino.dat"))
attributes(ELNINODAT)$dimnames[[1]] <- character(0)
learning <- 1:52
testing <- 53

```
elnino.past.learn <- ELNINODAT[learning,]
elnino.past.testing<-LNINODAT[testing,]
s <- 1
elnino.futur.s <- ELNINODAT[2:53,s]
```

The functional nonparametric prediction :

```
result.pred.quantile.step.s <- funopare.quantile.lcv( elnino.futur.s,
elnino.past.learn,elnino.past.testing,2, Knearest=NULL,
kind.of.kernel="quadratic", semimetric="pca").
result.pred.quantile.step.s$Predicted.values .
```

predict the 54th year :

♯ To do that, it suffices to repeat the previous stages for s = 1,. . . ,12 (horizons):

```
pred.mode <- 0
for(s in 1:12){
elnino.futur.s <- ELNINODAT[2:53,s]
result.pred.quantile.step.s <- funopare.quantile.lcv(elnino.futur.s,
elnino.past.learn,elnino.past.testing,3,Knearest=NULL,
kind.of.kernel="quadratic",semimetric="pca")
pred.quantile[s]<- result.pred.quantile.step.s$Predicted.values
}
```

Plotting the predicted values :

♯ The following command lines allow to display the forecasted 54th year obtained (Figure 1) by the various functional prediction methods and we compare them with the observed values (54th year)

```
year54 <-ELNINODAT[54,]
mse.quantile<- round(sum((pred.reg-year54)^2)/12,2)
plot(1:12,year54,type='l',lty=1,axes=F,xlab=′ ′, ylab=”, ylim=range(c(pred.median,year54)))
plot(1:12,pred.quantile,xlab='54th year', ylab=′ ′, main=paste('quantile: MSE=',mse.quantile,sep=′
type='l',lty=2,ylim=range(c(pred.quantile,year54)))
par(new=T)
```

4.2 Conclusion

There are many situations in which we study the link between two variables in order to be able to predict new values of one of them given the other one, this problem occurs with real, multivariate variables and functional variables. There are several ways to approach the prediction setting, in this dissertation we have been interested in two important models: conditional mode and conditional quantile which are studied when the explanatory variables are functional and the response variable still real. We have provided some theoretical supports by showing how the dependence is acting on the asymptotic behavior of the non-parametric functional method. In the last chapter, we have illustrated an application in which we have applied the conditional quantile approach in time-series analysis to the prediction and the building of confidence bands, then we have implement our methodology with *el Nino* data which is a real data study that test the performance of the conditional quantile estimator.

Bibliography

[1] Bosq, D. (2000). Linear processs in function spaces. Theory and Application. *Lectures Notes in Statistics*. **129**.

[2] Bouanani, O., Laksaci, A., Rachdi, M. and Rahmani, S. (2018). Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics. *Behaviormetrika*. **46**,

[3] Collomb, G., Hardle, W. and Hassani, S. (1987). A note on prediction via conditional mode estimation. *J. Statist. Plann. and Inf* **15**, 227-236.

[4] Dabo-Niang, S. (2004). Kernel density estimator in an infinite dimensional space with a rate of convergence in the case of diffusion process. *Appl. Math. Lett.* **17**, 381-386.

[5] Dabo-Niang, S. and Laksaci, A. (2007). Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle. *C. R. Math. Acad. Sci. Paris*, **344**, 49–52.

[6] Fan, J., Hu, T.C. and Truong, Y.K. (1994). Y.K., Robust nonparametric function estimation, *Scand. J. Statist.* **21**, 433-446.

[7] Ezzahrioui, M and Ould saïd, E. (2008). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional models, *Journal Of Nonparametric statistics*. **20**, 3-18.

[8] Ezzahrioui, M and Ould saïd, E. (2008). Asymptotic normality of a nonparametric estimator of the conditional quantile function in the normed space, *Far East J. Theorest Statistics*. **25**, 15-38.

[9] Ezzahrioui, M and Ould saïd, E. (2010). Some asymptotic results of a nonparametric conditional mode estimator for functional time-series, *Statistica Neerlandica.* **64**, 171-201.

[10] Ezzahrioui, M and Ould saïd, E. (2010). Asymptotic results of a nonparametric conditional quantile estimator for functional time-series, *Communications in statistics - Theory and methods.* **37**, 2735-2759.

[11] Ferraty, F. and Vieu, P. (2004). Nonparametric models for functional data, with application in regression times series prediction and curves discrimination. *J. Nonparametric Statist.* **16**, 111–127.

[12] Ferraty, F., Laksaci, A. and Vieu, P. (2005). Functional time series prediction via conditional mode estimation. *C R Math.*, **340**, 389–392.

[13] Ferraty, F. Laksaci, A. and Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statist. Inf. for Stoch. Processes.* **9**, 47-76.

[14] Ferraty, F. and Vieu, P. (2006). Nonparametric functional data analysis. Theory and Practice. *Springer Verlag.*

[15] Ferraty, F. Mas, A. and Vieu, P. (2007). Advances in nonparametric regressionfor functional variables. *Austral. New. Zeal. J. Statist.* **49**, 1-20.

[16] Gasser, T., Hall, P. and Presnell, B. (1998). Nonparametric estimation of the mode of a distribution of random curves. *J. Roy. Statist. Soc.* **60**, 681-691.

[17] Hall, P. and Heyde, C.C. (1980). Martingale Limit Theory and its Applications. *Academic Press. New York.*

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

[19] Jones, M.c. and Hall, P. (1990). Mean squared error properties of kernel estimates of regression quantiles. *Statist. Probab. Letters.* **10**, 283-289.

[20] Koenker, R. and Bassett, G. W. (1978). Regression quantiles. *Econometrica*, **46**, 33-50.

[21] Ling, N. and Xu, Q. (2012). Asymptotic normality of conditional density estimation in the single index model for functional time series data. *Statist Probab Lett.*, **82**, 2235–2243.

[22] Louani, D. and Ould-Said, E. (1999). Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. Nonparametric Statist.* **11**, 413-442.

[23] Masry, E. (2005). Nonparametric regression estimation for dependent functional data: Asymptotic normality. *Stoch. Proc. and their Applications.* **115**, 155-177.

[24] Mehra, K.L., Rao, M.s. and Upadrasta, S.P. (1991). A smooth conditional quantile estimator and related applications of conditional empirical processes. *J. Multivariate Anal.* **37**, 151-179.

[25] Ould-Saïd, E. (1997). A note on ergodic processes prediction via estimation of the conditional mode function. *Scand. J. of Statistics.* **24**, 231-239.

[26] Rachdi, M., Laksaci, A., Demongeot, J., Abdali, A. and Madani, F. (2014). Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data. *Comput Statist Data Anal.* **73**, 53–68.

[27] Ramsay, J. and Silverman,B. (2005). Applied Functional Data Analysis : Methods and Case Studies. *Springer, New-York.*

[28] Ramsay, J. and Silverman, B. (2002). Functional Data Analysis. *Springer, New-York.*

[29] Roussas, G.G. (1991). Estimation of a transition distribution function and its quantiles in Markov processes : Strong consistency and asymptotic normality, Nonparametric Functional Estimation and Related Topics. *Kluwer Academic Publishers, Dordrecht.* **335**, 443-362.

[30] Samanta, M. (1989). Non-parametric estimation of conditional quantiles. *Statist. Probab.* **7**, 407-412.

[31] Samanta, M. and Thavaneswaran, A. (1990). Nonparametric estimation of conditional mode. *Comm. Statist. Theory and Meth.* **16**, 4515-4524.

[32] Van der vaart, A.W. (1998). Asymptotic statistics. *Cambridge Series in Statistical and Probabilistic Mathematics.*

[33] Welsh, A.H. (1996). Robust estimation of smooth regression and spread functions and their derivatives. *Statist. Sinica.* **6**, 347-366.