République algérienne démocratique et populaire
Ministère de l'enseignement supérieur et de la recherche scientifique

Université Frères Mentouri Constantine
Faculté des sciences exactes
Département de mathématiques

N° d'ordre: 48/DS/2018.
N° de série: 05/Math/2018.

THÈSE

PRÉSENTÉE POUR L'OBTENTION DU DIPLÔME
DE DOCTORAT EN SCIENCES
EN MATHÉMATIQUES

"Estimation non paramétrique pour des données fonctionnelles"
"Nonparametric estimation for functional data"

Par Sara Leulmi

OPTION : Probabilités et Statistique

Devant le jury :

Président	M.	Z. Mohdeb	Prof.	Ecole Polytechnique de Constantine
Directrice de thèse	Mme	F. Messaci	Prof.	Université frères Mentouri, Constantine 1
Examinatrice	Mme	A. Chadli	Prof.	Université Bordj Badji Mokhtar, Annaba
Examinatrice	Mme	N. Nemouchi	Prof.	Université frères Mentouri, Constantine 1
Examinatrice	Mme	N. Seddik Ameur	Prof.	Université Bordj Badji Mokhtar, Annaba
Examinatrice	Mme	Kharfouchi	M.C.A.	Université Salah Boubnider, Constantine 3

Soutenue le: 21/03/2018.

Dedication

I dedicate this thesis ...

TO MY DEAR PARENTS, no dedication can express my respect, my eternal love and my consideration for the sacrifices you have made for my education and my well-being.

My mother, "Abed Fatima" who worked hard for my success, by her love, her support, all the sacrifices she made and her precious advices. For all her assistance and her presence in my life, receive through this work as modest as- it can be, the expression of my feelings and my eternal gratitude...Thank you for helping with my son, without you my mother, I would never reach this level0...frankly I do not think these little words grant her all what she deserves.

My father, "Salah" who can be proud and find here the result of long years of sacrifices and hardships to help me move forward in life. May God make this work bear fruit; Thank you for the noble values, the education and the constant support you offered to me.

TO MY DEAR AND ADORABLE SISTERS AND BROTHERS, who have not ceased to be for me examples of perseverance, courage and generosity. As a witness to my fraternal affection, my deepest tenderness and gratitude, I wish you all a life full of happiness and success and that God, the Almighty, protect you. (Specially my beautiful little sister Noussaiba).

TO MY DEAR HUSBAND "Sofiane Laahmar" and I would like him to thank for standing beside me throughout my time of writing this thesis. He has been my inspiration and motivation for continuing to improve my knowledge and move my career forward. He is my rock, and I dedicate this thesis to him.

TO MY WONDERFUL SON "Abdessamad", I also thank him for always making me smile and for understanding on those weekend mornings when I was writing this thesis instead of playing games. I hope that one day he can read this thesis and understand why I spent so much time in front of my computer.

TO the family "LEULMI", "ABED" and "LAAHMAR". I owe so much to all for their undying support, their unwavering belief that I can achieve so much. Unfortunately, I cannot thank everyone by name because it would take a lifetime but, I just want you all to know that you count so much; were it not for your sincere love and help, I would never have completed this thesis. So thank you all.

TO all those COLLEAGUES and FRIENDS of science faculty.

TO all those WHO LOVE ME, to all those WHOM I LOVE, I dedicate the fruit of MY DOCTORAT's thesis.

Acknowledgment

First and foremost, My deepest gratitude goes to Allah for allowing us to arrive where we are, and to successfully conclude this modest work, and for having such affective and caring parents.

I especially want to express my gratitude to Prof. Messaci Fatiha at the Mentouri Brothers University of Constantine who had the heavy task of being my supervisor. This work would never have succeeded without its unswerving dynamism and unwavering conviction. In the interest of my work, her constant encouragement has always allowed me to go beyond the different periods of encouragement which have marked this period of thesis and the experience you have helped me gain by working with you and the precious time you spent making sure my thesis is always on track, I hope you find some kind of satisfaction in this modest paper....frankly I do not think these little words grant her all what she deserves.

I extend my sincere thanks to Mr. Z. Mohdeb for having accepted to present the jury, and for all the good times we shared during my thesis's years.

Thank you very much to Mrs. N. Nemouchi , Mrs. A. Chadli, Mrs. N. Seddik Ameur and Mrs. S. Kharfouchi, for doing me the honor of being part of the jury, and for the time given to the attentive reading of my thesis.

Thanks also to Mr. A. Laksaci, Mrs. K. Kebabi and Mr. M. Boukeloua, for their relevant remarks, ideas and contributions to my programs computer.

Articles

1. Leulmi, S. et Messaci, F. (2017). Local linear estimation of a generalized regression function with functional dependent data. *Communications in Statistics- Theory and Methods*, DOI: 10.1080/03610926.2017.1402048

2. Leulmi, S. et Messaci, F. (2017). A class of local linear estimators with functional data. Submitted for publication.

Proceeding

1. Leulmi, S., et Messaci, F. Forecasting with Functional Time Series. ITISE'2017. Granada (SPAIN), 18-20 septembre 2017.

Communications

1. Leulmi, S. et Messaci, F. Estimation locale linéaire de la fonction de régression pour des données fonctionnelles et dépendantes (convergence ponctuelle). CMA'2016. Batna, 08-09 Nouvembre 2016.

2. Leulmi, S. et Messaci, F. Estimation locale linéaire de la fonction de régression généralisée pour des données fonctionnelles et dépendantes (convergence uniforme). ICM'2017. Alger, 12-15 Juillet 2017.

Contents

Introduction

The problem of the regression function estimation is often sought out by practitioners to solve their problems. Both parametric and nonparametric technics have been developed on the basis of a set of observations. In all these research works, the authors focus on two essential points.

- The nature of the data set.
- The link between a first variable of interest and a covariate X.

These variables can take different forms. It can be unidimensional or multidimensional if it contains an exact number of information. We can also encounter the case where the quantity of information is illimitated. In this case we define a functional variable, we mainly talk about "Data curves".
Functional variables can be only observed on a finite grid of discretization points, the estimation can then be viewed as a multidimensional problem. This technic fails because of the great number of discretization points which leads to the well-known problem of curse of dimensionality, linked to the sparseness of the data. This motivates the extension of the finite dimensional statistical technics to the infinite dimensional data setting. The nonparametric methods are then reasonable ways to deal with this type of data sets.
There is nowadays a large number of fields where functional data are collected such as environmetrics, medicine, finance and pattern recognition. A classical statistical problem is that of regression which consists to study the relationship between two observed variables with the aim to predict the value

of the response variable when a new value of the explanatory one is observed. Note that the modelization of functional variable is becoming more and more popular since the publication of the monograph of Ramsay and Silverman (1997) on functional data analysis. However, the first results concerning the nonparametric models were obtained by Ferraty and Vieu (2000) who established the almost complete pointwise consistency of kernel regression estimators when the observations are independent and identically distributed (i.i.d.). These results have been extended in Ferraty et al. (2002) by treating the time series prediction. Dabo and Rhomari (2003) stated the convergence in $\mathbb{L}^p$ norm of the kernel estimator of this model and Delsol (2007) states the asymptotic expression for the $\mathbb{L}^p$ errors. The reader can found in Ferraty and Vieu (2006) more discussions on nonparametric methods for functional data. The asymptotic results including the mean squared convergence, with rates, as well as the asymptotic normality of kernel estimators of regression function have been obtained by Ferraty et al. (2007); many other recent related references about the nonparametric functional data analysis include Amiri et al. (2014), Ezzahrioui and Ould-Said (2008) , Rachdi and Vieu (2007) and so on.

Meanwhile, the nonparametric k-Nearest-Neighbours (kNN) estimator for functional data has also been investigated. For example, Burba et al. (2009) established the pointwise consistency for independent data and Kudraszow and Vieu (2013) gave the rate of the almost complete uniform convergence of the regression KNN estimator.

In the most of the aforementioned works the authors used the kernel method, whereas it is known, in the finite dimensional data case, that the latter produces high bias compared to the local linear method. We can found in Chu and Marron (1991), Fan (1992) and Fan and Gijbels (1996) an interesting comparison between both methods. Since the open question (cf. Ferraty and Vieu (2006)) "How can the local polynomial ideas be adapted to infinite-dimensional settings?", the local linear smoothing in the functional data setting has been considered by many authors in several versions.

The first one was considered by Baillo and Grane (2009) who studied the consistency in mean square of the constructed local linear estimator when the covariates are of Hilbertian nature (see also the paper by El Methni and Rachdi (2011)). Another version of a functional local linear regression estimator was given by Barrientos et al. (2010) in the case where the explanatory variable is valued in a functional semi metric space. Then, Berlinet et al. (2011) stated the asymptotic mean square error of a functional local linear estimator of the regression operator which is constructed by inverting the local covariance operator of the functional explanatory variable. The mean-square convergences of the locally modelled regression estimation for conditional density function and conditional cumulative distribution function have also been established in Rachdi et al. (2014) and Demongeot et al. (2014), respectively for independent functional data. Zhiyong and Zhengyan (2016) established the mean-square convergence as well as the asymptotic normality for the regression function, they also adapt the empirical likelihood method to construct the pointwise confidence intervals for the regression function and derived the Wilk's phenomenon for the empirical likelihood inference. Attaoui et al. (2017) considered the problem of the local linear estimation of the regression operator when the regressor is functional, they constructed an estimator by the kNN method and established its almost complete consistency with rate.

Notice that Barrientos et al. (2010) obtained a rate of the pointwise almost-complete convergence for the local linear estimator of the regression function. But, as pointed out in Ferraty et al. (2010) "the uniform consistency results are indispensable tools for the study of more sophisticated models in which multi-stage procedures are involved". Under uniform convergence, one can make prediction even if the data are not well observed. We also can solve

some problems such as data-driven bandwidth choice (see Benhenni et al. (2007)), or bootstrapping (see Ferraty et al. (2008)). Uniform convergence of other local linear nonparametric estimators has been investigated in some papers as Demongeot et al. (2010) and Demongeot et al. (2011) for the conditional density and Messaci et al. (2015) for the conditional quantile. As for us, one of our principal aims is to establish the uniform almost complete convergence of the local linear estimator of a generalized regression function which generalizes the regression estimator studied in Barrientos et al. (2010) and to focus on a robust tool of prediction (a conditional quantile estimator). First, the researchers considered a functional explanatory variable and a real response variable (see the references previously cited). Then, the case when the response variable is also functional is treated , see for example Ferraty et al. (2011), (2012a). Moreover, Demongeot et al. (2017) generalized the results established by Baillo and Grane (2009), considering both the response and the explanatory variables of functional kind. In this direction, they stated the rate of uniform almost-complete convergence of the local linear estimator of the regression operator.
However, in practice, observed data can exhibit a dependence form. A large studied example is the case of the α-mixing dependence. We cite Laksaci et al. (2011) and Attaoui et al. (2014) for papers dealing with such functional dependent data. In the last works, the pointwise almost complete convergence has been studied, while Laib and Louani (2010), and Ling et al. (2015) obtained the asymptotic properties of the nonparametric kernel estimator for functional stationary ergodic data, Benhenni et al. (2008) for the long memory dependent case. In 2005, Masry (2005) investigated the asymptotic normality of the nonparametric kernel estimator for α-mixing functional data. Demongeot et al. (2013) established the pointwise almost-complete consistency of a fast functional local linear estimator of the conditional density when the explanatory variable is functional and the observations are

dependent and Ferraty et al. (2012b) treated the case when the response variable is also functional for the β-mixing observations.

The uniform almost sure convergence has been proved in Ferraty and Vieu (2004) for kernel estimators in the situation of dependent functional data for α-mixing functional data. It is known that the local linear method can improve the quality of the estimation. But, despite the importance of the uniform convergence, we are not aware of results dedicated to this topic, for local linear estimates, in the setting of dependent functional data. In this thesis, we address this problem. More precisely, we establish both pointwise and uniform almost complete convergence of the local linear estimator of the generalized regression function based on dependent functional data.

Organization of the thesis

This thesis is organized as follows.

Chapter 1 : It consists to study a local modelling approach when one regresses a scalar response on an explanatory functional variable via a regression estimator proposed in Barrientos et al. (2010). The pointwise almost complete convergence of this estimator is given in Section 1.1.

Chapter 2 : Our principal aim, in this part, is to establish the uniform almost complete convergence of the local linear estimator of a generalized regression function which generalizes the regression estimator studied in (Barrientos et al., 2010) and to focus on a robust tool of prediction (a conditional quantile estimator). More precisely, Section 2.1 is devoted to introduce the generalized regression function estimator and to state its pointwise convergence. Section 2.2 contains the principal result of this section which consists to establish the rate of the uniform almost convergence of the last estimator. Then, we focus on the particular case of the conditional distribution function estimation from which we deduce a rate of the uniform consistency of a conditional quantile estimator. In Section 2.3 using a real data set, the prediction obtained from this last estimator is compared to those of two other known estimators.

Chapter 3 : In this chapter, we establish the almost complete convergence of a local linear nonparametric estimator of the conditional distribution function of a scalar response variable given a random variable taking values in a semi metric space (the functional variable) when the collected observations are α-mixing (see Sections 3.1 and 3.2). Then, we derive the consistency of a conditional median estimator which is a prediction tool. Finally, in Section 3.3 a real data study shows that our estimator performs well with respect to other known conditional median estimators.

Chapter 4 : We establish, in this chapter, the pointwise and the uniform almost complete convergence (see Sections 4.1 and 4.2) of the local linear estimator of the generalized regression function presented in Chapter 2, except that the data are here assumed to be α-mixing. This dependence complicates considerably the theoretical study. A comparison between kernel and local linear estimators, based on functional dependent data, is conducted from two real datasets in Section 4.3.

Chapter 5 : For the sake of easy references, we briefly recall, in this annex, some basic definitions and probabilistic tools needed in this thesis .

Chapter 1

Local linear estimation

A very widely studied problem in statistics is the link between two variables, the main goal of which is to predict one of the variables (the response variable) given a new value of the other (the explanatory variable). One way to deal with this problem is by means of the regression method which is based on the conditional expectation.

Since the pioneer works in Ferraty and Vieu (2006), various studies dealt with the nonparametric functional estimation. This research field is motivated by the fact that several data collected in practice, are given in the form of curves and that the progress of the digital computing tools allows the treatment of such observations.

Here, we focuse on the nonparametric estimation of the regression operator defined by

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

where the explanatory variable X is valued in some infinite-dimensional space $\mathcal{F}$, Y is a scalar response, ε is a random noise independent from X.

To do that, one way consists in using a functional kernel estimator (see Ferraty and Vieu (2000) and Ferraty and Vieu (2006), for a deep study), which is an extension to this functional framework of the Nadaraya-Watson kernel estimator, based on n pairs $(Xi, Yi)_{i=1,\dots,n}$ identically and independently distributed as (X, Y). The functional kernel estimator is defined as follows

$$\widehat{m}_0(x) = \frac{\sum_{i=1}^{n} Y_i K(h^{-1}|\delta(x, X_i)|)}{\sum_{i=1}^{n} K(h^{-1}|\delta(x, X_i)|)}, \tag{1.1}$$

where K is a standard univariate kernel function, $\delta(.,.)$ locates one element of $\mathcal{F}$ with respect to another one, and the bandwidth $h := h_n$ is a sequence of strictly positive real numbers which plays a smoothing parameter role. This kernel estimator $\widehat{m}_0(x)$ can be seen as the solution of the minimization problem

$$\min_{a \in \mathbb{R}} (WSE_x(a)) \text{ with } WSE_x(a) = \sum_{i=1}^{n} (Y_i - a)^2 K(h^{-1}|\delta(X_i, x)|),$$

since it is easy to check that the derivative of WSE_x vanishes at $a = \widehat{m}_0(x)$. Actually, the kernel estimator given by (1.1) is locally approximating m by a constant (a zero-degree polynomial). So, to increase the effictive of the functional nonparametric regression estimator, we use a local approximation which is more accurate than a constant one. In particular, we consider a polynomial of degree one, which is called "local linear estimator" and has been extended to the functional framework (see Barrientos et al. (2010), Baillo and Grane (2009) for example).

Here we are interested in estimating the regression function in a nonparametric fashion. Barrientos et al. (2010), proposed an estimator $\hat{m}$, as the solution for a of the following minimization problem

$$\min_{(a,b) \in \mathbb{R}^2} (WSE'_x(a, b)) \text{ with } WSE'_x(a, b) = \sum_{i=1}^{n} \left[Y_i - a - b\beta(X_i, x))\right]^2 K(h^{-1}|\delta(X_i, x)|),$$

where $\beta(.,.)$ is a known operator from $\mathcal{F} \times \mathcal{F}$ into $\mathbb{R}$ such that, $\forall x \in \mathcal{F}, \beta(x, x) = 0$. So, we can write

$$\widehat{m} = e'_1 (C'_\beta W C_\beta)^{-1} C'_\beta W Y, \tag{1.2}$$

where

$$\text{C'}=\begin{bmatrix} 1 & ... & 1 \\ \beta(X_1,x) & ... & \beta(X_n,x) \end{bmatrix}, \quad \text{Y}=\begin{bmatrix} Y_1 \\ ... \\ Y_n \end{bmatrix}$$

$W = diag\,(K(h^{-1}|\delta(X_1,x)|), ..., K(h^{-1}|\delta(X_n,x)|))$ and $e_1 = (1,0) \in \mathbb{R}^2$.
By a simple calculus, one's can derive the following explicit estimator

$$\widehat{m}(x) = \frac{\sum_{i,j=1}^{n} W_{ij}(x) Y_j}{\sum_{i,j=1}^{n} W_{ij}(x)} \quad \left(\frac{0}{0} := 0\right),$$

where

$$W_{ij}(x) = \beta(X_i,x)\,(\beta(X_i,x) - \beta(X_j,x))\,K(h^{-1}|\delta(X_i,x)|)K(h^{-1}|\delta(X_j,x)|).$$

Notice that for $l \in \{0,1\}$, we have

$$\sum_{i,j=1}^{n} W_{ij}(x) Y_j^l = \sum_{i<j} \{ \left[\beta(X_i,x) - \beta(X_j,x)\right] \left[\beta(X_i,x) Y_j^l - \beta(X_j,x) Y_i^l\right]$$
$$K(h^{-1}|\delta(X_i,x)|)K(h^{-1}|\delta(X_j,x)|\},$$

so, if the denominator of the estimator $\widehat{m}(x)$ is zero, it is the same for its numerator. Moreover, under appropriate assumptions to be assumed later we get $EW_{12}(x) > 0$ (see (1.8) in section 1.2).
This approach assumes that $a + b\beta(.,x)$ is a good approximation of $m(.)$ around x. As $\beta(x,x) = 0$, a will be a suitable estimate for $m(x)$. Notice that the expression of $\widehat{m}$ allows fast computational issue and that the choices of β and d will be crucial.

1.1 Asymptotic properties

Barrientos et al. (2010) have studied the almost complete convergence of the locally modelled estimator of the regression operator $m(x) = E(Y|X = x)$, with x being a fixed element of a semimetric space $(\mathcal{F}, d)$ where $d = |\delta(x,y)|$.

First of all, for any positive real h, let $B(x,h) := \{y \in \mathcal{F}/,\ |\delta(x,y)| \leq h\}$ be a closed ball in $\mathcal{F}$ of center x and radius h , $\Phi_x(h,h') := P(h \leq |\delta(x,X)| \leq h')$ and $\Phi_x(h) := \Phi_x(0,h)$.

The pointwise almost complete convergence of the local linear estimator $\widehat{m}$ of the regression function m will be established under the following assumptions.

(H1) for any $h > 0$; $\Phi_x(h) > 0$.

$(H2_C)$ $m \in \{f : \mathcal{F} \to \mathbb{R}, \lim_{|\delta(x,x')| \to 0} f(x') = f(x)\}$.

(H2$_L$) $m \in \{f : \mathcal{F} \to \mathbb{R}, \forall x' \in \mathcal{F}; |f(x) - f(x')| \leq C|\delta(x,x')|^b\}$ where b and C are fixed in $\mathbb{R}^+$.

(H3) The function $\beta(.,.)$ is such that: $\exists 0 < M_1 < M_2, \forall x' \in \mathcal{F}$,

$$M_1|\delta(x,x')| \leq |\beta(x,x')| \leq M_2|\delta(x,x')|.$$

(H4) The kernel K is a positive and differentiable function on its support $[0,1]$.

(H5) The bandwidth h satisfies:

$$\lim_{n\to\infty} h = 0 \ \ and \ \lim_{n\to\infty} \left(\frac{\ln n}{n\Phi_x(h)} \right) = 0.$$

(H6) There exist an integer n_0, such that:

$$\forall n > n_0, \forall x \in \mathcal{F}, \frac{1}{\Phi_x(h)} \int_0^1 \Phi_x(zh,h) \frac{d}{dz} \left(z^2 K(z)\right) > C > 0.$$

(H7)

$$h \int_{B(x,h)} \beta(u,x) dP_X(u) = o\left(\int_{B(x,h)} \beta^2(u,x) dP_X(u) \right),$$

where dP_X is the distribution of X.

(H8) $\forall m \geq 2, \sigma_m : x \longmapsto E(Y^m/X)$ is a continuous operator on $\mathcal{F}$.

Hypotheses (H2)-(H5) and (H8) are standard in the nonparametric functional regression setting and extend what is usually assumed in the classical p-dimensional nonparametric literature (see Ferraty and Vieu (2006) for a large discussion). The kind of kernels in (H4) contains the standard kernels used in the literature (uniform, triangle, quadratic,...). Hypothesis (H6) precises the behaviour of the bandwidth h in relation with the small ball probabilities and the kernel function K. The key new hypothesis is (H7) about the local behaviour of the operator β which models the local shape of the regression. For instance, in the special case where $\beta = \delta$, this assumption means that the local expectation of β is small enough with respect to its moment of second order. If, in addition, the real rodom variable (r.r.v.) $\beta(X, x)$ admits a differentiable density (around 0) with respect to the Lebesgue measure then (H7) is satisfied (see Appendix 1 for more details). Let's state two important theorems that consolidate the pointwise almost-complete convergence (a.co.) and its rate.

Theorem 1.1. *(Barrientos et al. (2010))*
Under assumptions (H1), ($H2_C$); (H3)-(H8), we have

$$\widehat{m}(x) - m(x) = o_{a.co.}(1). \tag{1.3}$$

Theorem 1.2. *(Barrientos et al. (2010))*
Under assumptions (H1), ($H2_L$); (H3)-(H8), we have

$$\widehat{m}(x) - m(x) = O(h^b) + O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right). \tag{1.4}$$

Barrientos et al. (2010) have introduced the following decomposition on which the proofs of the above theorems are based.
For all $x \in \mathcal{F}$,

$$\widehat{m}(x) - m(x) = \frac{1}{r_0(x)}\left(r_1(x) - Er_1(x)\right) - \left(m(x) - Er_1(x)\right) - \frac{m(x)(r_0(x) - 1)}{r_0(x)} \tag{1.5}$$

where, for $l = 0, 1$

$$r_l(x) = \frac{1}{n(n-1)EW_{12}(x)} \sum_{i \neq j} W_{ij}(x) Y_j^l. \tag{1.6}$$

Further, to complete the proof, we need to apply the two following lemmas.

Lemma 1.1. *Under the assumptions (H1),(H3)-(H5),*
(i) The hypothesis ($H2_C$) permits to write

$$m(x) - Er_1(x) = o(1).$$

(ii) and the assumption ($H2_L$) gives

$$m(x) - Er_1(x) = O(h^b).$$

Lemma 1.2. *Under the assumptions of (H1), ($H2_C$),(H3)-(H7), we obtain that*
(i)

$$r_0(x) - 1 = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

(ii) In addition, if (H8) holds, one gets

$$r_1(x) - Er_1(x) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

We give the proofs of these results because we need them in the next chapter.

1.2 Proofs

In what follows, let C be some strictly positive generic constant and for any $x \in \mathcal{F}$, and for all $i = 1, \ldots, n$

$$K_i(x) := K(h^{-1}|\delta(X_i, x)|) \text{ and } \beta_i(x) := \beta(X_i, x).$$

To treat the pointwise almost-complete convergence of $\widehat{m}(x)$, we need the following preliminary technical lemma.

Lemma 1.3. *(see Lemma A.1 in (Barrientos et al. (2010))*
Under assumptions (H1), (H3)–(H7), we obtain
i) $\forall (p,l) \in \mathbb{N}^\star \times \mathbb{N},\ E\left(K_1^p(x)|\beta_1(x)|^l\right) \leq Ch^l \Phi_x(h)$.
ii) $E\left[K_1(x)\beta_1^2(x)\right] > Ch^2\left[\Phi_x(h)\right]$ *for n sufficiently large.*

Proof 1.1. i) *One starts by using (H3) which implies*

$$K_1^p(x)|\beta_1(x)|^l h^{-l} \leq C K_1^p(x)|\delta(X_1,x)|^l h^{-l},$$

and because the kernel K is bounded on $[0,1]$ (see (H4)), one gets

$$K_1^p(x)|\beta_1^l(x)|h^{-l} \leq C|\delta(X_1,x)|^l h^{-l} 1_{[0,1]}\left(h^{-l}|\delta(X_1,x)|\right),$$

and thus, we have

$$E\left(K_1^p(x)|\beta_1(x)|^l h^{-l}\right) \leq C\Phi_x(h),$$

which is the claimed result.
ii) *By using (H3), it is easy to see that*

$$E\left[K_1(x)\beta_1^2(x)\right] > CE\left[K_1(x)\delta^2(X_1,x)\right].$$

Moreover, one can write

$$\begin{aligned} E\left(K_1(x)\frac{\delta^2(X_1,x)}{h^2}\right) &= \int_0^1 t^2 K(t) dP^{|\delta(X,x)|/h}(t) \\ &= \int_0^1 \left[\int_0^t \left(\frac{d}{du}(u^2K(u))\right)du\right] dP^{|\delta(X,x)|/h}(t) \\ &= \int_0^1 \left[\int_0^1 1_{[u,1]}(t) dP^{|\delta(X,x)|/h}(t)\right] \frac{d}{du}(u^2K(u)), \end{aligned}$$

the last equation comes from the Fubini's theorem. In addition, it is easy to check that

$$\int_0^1 1_{[u,1]}(t) dP^{|\delta(X,x)|/h}(t) = P(uh \leq |\delta(X,x)| \leq h).$$

So,

$$E\left(K_1(x)\frac{\delta^2(X_1,x)}{h^2}\right)=\int_0^1 \Phi_x(uh,h)\frac{d}{du}(u^2K(u))du.$$

It remains to use (H6) to obtain the desired lower bound, which ends the proof of Lemma 1.3-(ii).

Proof of lemma 1.1 We have

$$Er_l(x)=\frac{1}{E(W_{12}(x))}E(W_{12}(x)Y_2^l)),$$

and $Er_1(x)$ can also be written as

$$Er_1(x)=E\left(E(r_1(x)|X_2)\right)=\frac{1}{E(W_{12}(x)}E\left(W_{12}(x)E(Y_2|X_2)\right).$$

which allows us to write, under assumption (H4)
$|m(x)-Er_1(x)|=\frac{1}{|E(W_{12}(x))|}\left|E\left(W_{12}(x)(m(x)-m(X_2))\right)\right|\leq \sup_{x'\in B(x,h)}|m(x)-m(x')|.$
We need to take into account hypothesis ($\mathrm{H2}_C$) to obtain Lemma 1.1-(i). However, if one uses ($\mathrm{H2}_L$) instead of ($\mathrm{H2}_C$), it is clear that

$$\sup_{x'\in B(x,h)}|m(x)-m(x')|=O(h^b).$$

which leads us to Lemma 1.1-(ii).
Proof of lemma 1.2 ii) Remark that

$$r_1(x)=Q(x)\left[M_{2,1}(x)M_{4,0}(x)-M_{3,1}(x)M_{3,0}(x)\right], \qquad (1.7)$$

where, for $p=2,3,4$, and $l=0,1$,

$$M_{p,l}(x)=\frac{1}{n\Phi_x(h)}\sum_{i=1}^n\frac{K_i(x)\beta_i^{p-2}(x)Y_i^l}{h^{p-2}}$$

and

$$Q(x)=\frac{n^2h^2\Phi_x^2(h)}{n(n-1)E\left(W_{12}(x)\right)}.$$

So, one has

$$r_1(x) - E(r_1(x)) = Q(x)\{M_{2,1}(x)M_{4,0}(x) - E(M_{2,1}(x)M_{4,0}(x)) \\ - [M_{3,1}(x)M_{3,0}(x) - E(M_{3,1}(x)M_{3,0}(x))\},$$

and since

$$\begin{aligned} M_{2,1}(x)M_{4,0}(x) - E(M_{2,1}(x)M_{4,0}(x) &= (M_{2,1}(x) - E(M_{2,1}(x))(M_{4,0}(x) - E(M_{4,0}(x)) \\ &+ (M_{2,1}(x) - E(M_{2,1}(x))E(M_{4,0}(x)) \\ &+ (M_{4,0}(x) - E(M_{4,0}(x))E(M_{2,1}(x)) \\ &+ E(M_{2,1}(x))E(M_{4,0}(x)) - E(M_{2,1}(x)M_{4,0}(x)), \end{aligned}$$

$$\begin{aligned} M_{3,1}(x)M_{3,0}(x) - E(M_{3,1}(x)M_{3,0}(x) &= (M_{3,1}(x) - E(M_{3,1}(x))(M_{3,0}(x) - E(M_{3,0}(x)) \\ &+ (M_{3,1}(x) - E(M_{3,1}(x))E(M_{3,0}(x)) \\ &+ (M_{3,0}(x) - E(M_{3,0}(x))E(M_{3,1}(x)) \\ &+ E(M_{3,1}(x))E(M_{3,0}(x)) - E(M_{3,1}(x)M_{3,0}(x)). \end{aligned}$$

We have to show that for $p \in \{2, 3, 4\}$ and $l \in \{0, 1\}$

$$\begin{aligned} &Q(x) = O(1), \\ &EM_{p,l}(x) = O(1), \\ &E(M_{2,1}(x))E(M_{4,0}(x)) - E(M_{2,1}(x)M_{4,0}(x)) = O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right), \\ &E(M_{3,1}(x))E(M_{3,0}(x)) - E(M_{3,1}(x)M_{3,0}(x)) = O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right), \\ &M_{p,l}(x) - EM_{p,l}(x) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right). \end{aligned}$$

• Treatment of the term $Q(x)$

We have

$$\begin{aligned} EW_{12}(x) &= E\left[\beta_1^2(x)K_1(x)K_2(x)\right] - E\left[\beta_1(x)\beta_2(x)K_1(x)K_2(x)\right] \\ &= E\left[\beta_1^2(x)K_1(x)\right]E(K_2(x)) - \left(E\left[\beta_1(x)K_1(x)\right]\right)^2, \end{aligned}$$

together with

$$hE\left[\beta_1(x)K_1(x)\right] \le Ch\int_{B(x,h)} \beta(u,x)dP_{X_1}(u)$$

and (H7) implies that

$$hE\left[\beta_1(x)K_1(x)\right] = o\left(\int_{B(x,h)} \beta^2(u,x)dP_{X_1}(u)\right).$$

By applying Lemma 1.3-(i), with $K = 1_{[0,1]}$, $p = 1$ and $l = 2$ one gets

$$\int_{B(x,h)} \beta^2(u,x)dP_{X_1}(u,t) \le Ch^2\left[\Phi_x(h)\right]^2,$$

which implies that

$$E\left[\beta_1(x)K_1(x)\right] = o\left(h\left[\Phi_x(h)\right]^2\right).$$

Now, Lemma 1.3-(ii) and the last result allow to write

$$EW_{12}(x) > Ch^2\left[\Phi_x(h)\right]^2. \tag{1.8}$$

So, for n sufficiently large

$$Q(x) = O(1)$$

• It is easy to see that under (H1)–(H4), for $p \in \{2,3,4\}$ and $l \in \{0,1\}$, we have

$$\begin{aligned} EM_{p,l}(x) &= h^{2-p}\Phi_x(h)^{-1}E[K_1(x)Y_1^l\beta_1^{p-2}] \\ &= h^{2-p}\Phi_x(h)^{-1}E[K_1(x)m^l(X_1)\beta_1^{p-2}], \end{aligned}$$

and because $m(X_1) = m(x) + o(1)$ (under (H2$_C$)), one gets $EM_{p,l}(x) = O(1)$.
• Treatment of the term $E(M_{2,1}(x))E(M_{4,0}(x)) - E(M_{2,1}(x)M_{4,0}(x))$
On one side, we have

$$\begin{aligned} E(M_{2,1}(x))E(M_{4,0}(x)) &= \frac{1}{n^2h^2[\Phi_x(h)]^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E(K_i(x)\beta_i^2(x))E(K_j(x)Y_j) \\ &= \frac{1}{h^2[\Phi_x(h)]^2}E(K_1(x)\beta_1^2(x))E(K_1(x)Y_1), \end{aligned}$$

and on the other side, we get

$$
\begin{aligned}
E(M_{2,1}(x)M_{4,0}(x)) &= \frac{1}{n^2h^2[\Phi_x(h)]^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E(K_i(x)\beta_i^2(x)K_j(x)Y_j)\\
&= \frac{1}{n^2h^2[\Phi_x(h)]^2}\left(\sum_{i=j=1}^{n}E(K_i^2(x)\beta_i^2(x)Y_i)+\sum_{i\neq j}E(K_i(x)\beta_i^2(x)K_j(x)Y_j)\right)\\
&= O\left((n\Phi_x(h))^{-1}\right)+\frac{n^2-n}{n^2h^2[\Phi_x(h)]^2}E(K_1(x)\beta_1^2(x))E(K_1(x)Y_1),
\end{aligned}
$$

which allows us to write

$$
\begin{aligned}
E(M_{2,1}(x))E(M_{4,0}(x))-E(M_{2,1}(x)M_{4,0}(x)) = (1-\frac{n(n-1)}{n})h^{-2}\Phi_x(h)^{-2}E[K_1(x)\beta_1^2(x)]\\
E[K_1(x)Y_1]+O\left((n\Phi_x(h))^{-1}\right).
\end{aligned}
$$

Using similar arguments as previously, it is easy to see that

$$
E(M_{2,1}(x))E(M_{4,0}(x))-E(M_{2,1}(x)M_{4,0}(x)) = O\left((n\Phi_x(h))^{-1}\right),
$$

which is negligible with respect to $O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$, under (H5).

• By similar arguments, one can state

$$
E(M_{3,1}(x))E(M_{3,0}(x))-E(M_{3,1}(x)M_{3,0}(x)) = O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).
$$

• Treatment of the term $M_{p,l}(x)-EM_{p,l}(x)$

We have

$$
M_{p,l}(x)-EM_{p,l}(x)=\frac{1}{n}\sum_{i=1}^{n}Z_i^{(p,l)}(x),
$$

where

$$
Z_i^{(p,l)}(x)=\frac{1}{h^{p-2}\Phi_x(h)}\left\{K_i(x)\beta_i^{p-2}(x)Y_i^l-E\left[K_i(x)\beta_i^{p-2}(x)Y_i^l\right]\right\}. \qquad (1.9)
$$

In order to apply an exponential inequality, we focus on the absolute moments of the r.r.v. $Z_i(x)$

$$E|\{Z_i^{(p,l)}(x)\}^m| = h^{(-p+2)m}\Phi_x(h)^{-m}E|\sum_{k=0}^{m} c_k^m(-1)^{m-k}(K_i(x)\beta_i^{p-2}(x)Y_i^l)^k(E\left[K_i(x)\beta_i^{p-2}(x)Y_i^l\right])^{m-k}|$$
$$\leq h^{(-p+2)m}\Phi_x(h)^{-m}\sum_{k=0}^{m} c_k^m E[K_i^k(x)\beta_i^{(p-2)k}(x)\sigma_k^l(X_i)]|E\left[K_i(x)\beta_i^{p-2}(x)m^l(X_i)\right]|^{m-k}, \quad (1.10)$$

the last inequality is obtained by conditionning on X_1. In addition, (H2$_C$) implies that $m(X_1) = m(x) + o(1)$ whereas one gets $\sigma_k(X_1) = \sigma_k(x) + o(1)$ as soon as (H8) is checked. This, combined with (1.10) and Lemma 1.3-(i), allows us to write

$$E|\{Z_i^{(p,l)}(x)\}^m| = O\left(h^{(-p+2)m}[\Phi_x(h)]^{-m}\sum_{k=0}^{m} E[K_i^k(x)\beta_i^{(p-2)k}(x)]E\left[K_i(x)\beta_i^{p-2}(x)\right]|^{m-k}\right)$$
$$= O\left(max_{k\in\{0,\ldots,m\}}[\Phi_x(h)]^{-k+1}\right)$$
$$= O\left([\Phi_x(h)]^{-m+1}\right).$$

Finally, it suffices to apply Proposition 5.3–(ii) with $a_n^2 = [\Phi_x(h)]^{-1}$to get, for $p \in \{2,3,4\}$ and $l \in \{0,1\}$

$$M_{p,l}(x) - EM_{p,l}(x) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right),$$

1.3 Appendix 1

Remark on (H6)

Let us investigate here the special case where the fuctional variable X is a fractal process of order k ($k > 0$) such that

$$\lim_{\varepsilon\to 0}\sup_{t\in[0,1]}\left|\frac{\Phi_x(t\varepsilon)}{t^k\varepsilon^k} - C_x\right| = 0,$$

where C_x is a constant which does not depent on t and ε. This implies that, for any ε small enough, $\Phi_x(\varepsilon) \sim C_x \varepsilon^k$. We have

$$\begin{aligned}
\Phi_x(uh, h) &= P(uh \leq |\delta(x, X)| \leq h) \\
&= P(|\delta(x, X)| \leq h) - P(|\delta(x, X)| \leq uh)) \\
&= \Phi_x(h) - \Phi_x(uh) \\
&= C_x h^k (1 - u^k) + o(1).
\end{aligned}$$

Then, it is easy to state

$$\int_0^1 \Phi_x(uh, h) \frac{d}{du}(u^2 K(u)) du = C_x h^k \int_0^1 (1 - u^k) \frac{d}{du}(u^2 K(u)) du + o(h^k).$$

Now, one considers the family of kernels indexed by $\alpha > 0$ and defined by $K_\alpha(u) = \frac{\alpha+1}{\alpha}(1 - u^k) 1_{[0,1]}(u)$. Tis family of kernel contain standard asymmetric ones (triangle, quadratic). It comes with trivial calculus that

$$\int_0^1 \Phi_x(uh, h) \frac{d}{du}(u^2 K(u)) du = \frac{(\alpha + 1)k}{(k + 2)(\alpha + k + 2)} C_x h^k + o(h^k),$$

which leads us to assumption (H6) as soon as h is small enough (i.e. as soon as n is large enough). In the same way, (H6) holds when ones considers the uniform kenel $1_{[0,1]}(.)$.

(H6) is satisfied for much wider class of functional random variable (i.e. Hilbertian squared integrable ones) as soon as one considers stuitable semi metric δ (for more details, see Lemma 13.6 in Ferraty and Vieu Ferraty and Vieu (2006), p.213).

Remark on (H7)

In the special case where $\beta = \delta$, and the r.r.v. $Z := \beta(x, X)$ admits a differentiable density f_Z (around 0) with respect to the Lebesgue measure and such that $f_Z(0) \neq 0$, which implies that

$$\exists \alpha > 0 \; f_Z(z) \neq 0; \; \forall z \in [-\alpha, \alpha],$$

then hypothesis (H7) is satisfied. Indeed, we have, for any $x \in \mathcal{F}$

$$\begin{aligned} h\int_{B(x,h)} \beta(u,x)dP_X(u) &= h\int_{\mathcal{F}} \beta(u,x)1_{B(x,h)}(u)dP_X(u) \\ &= h\int_{\Omega} Z(w)1_{\{|Z(w)|\leqslant h\}}dP(w) \\ &= h\int_{\mathbb{R}} z1_{\{|z|\leqslant h\}}dP_Z(z) \\ &= h\int_{-h}^{h} zf_Z(z)dz \\ &= h\int_0^h z(f_Z(z)-f_Z(-z))dz, \end{aligned}$$

f_Z is a differentiable density around 0 such that $f_Z(0) \neq 0$ and by using the Taylor-Young formula ($f_Z(z) = f_Z(0) + zf'_Z(0) + z\varepsilon(z)$ where $\varepsilon(z) \to 0$ as $z \to 0$), we get, for some $\alpha > 0$ such that $\alpha > h$

$$\begin{aligned} h\left|\int_{B(x,h)} \beta(u,x)dP_X(u)\right| &= \left|2hf'_Z(0)\int_0^h zdz + h\int_0^h z\varepsilon(z)dz\right| \\ &\leqslant \frac{2h\left(|f'_Z(0)|+1\right)}{\inf_{z\in[0,\alpha]} f_Z(z)}\int_0^h z^2 f_Z(z)dz \\ &\leqslant \frac{2h\left(|f'_Z(0)|+1\right)}{\inf_{z\in[0,\alpha]} f_Z(z)}\int_{B(x,h)} \beta^2(u,x)dP_X(u). \end{aligned}$$

Chapter 2

A class of local linear estimators with functional data

As an alternative to the well-known Nadaraya-Watson estimator for regression function, in the framework of functional data, locally modelled regression estimators perform well (see Baillo and Grane (2009), Barrientos et al. (2010)). In this chapter, using the last method, we investigate a nonparametric estimation of some functionals of the conditional distribution of a scalar response variable Y given a random variable X taking values in a semi-metric space. These functionals include the regression function, the conditional cumulative distribution and some other ones.

The paper of Barrientos et al. (2010) is only concerning pointwise consistency results and our main aim is to prove the uniform almost complete convergence of estimators including those studied in the for-mentioned paper.

2.1 The estimation and the pointwise almost-complete convergence

Throughout this chapter, we consider a sample of independent pairs $(X_i, Y_i)_{i=1,\ldots,n}$ identically distributedas (X, Y) which is a random vector valued in $\mathcal{F} \times \mathbb{R}$, where $(\mathcal{F}, d)$ is a semi-metric space. Our goal is to estimate the generalized regression function, defined for all x in $\mathcal{F}$, by

$$m_\varphi(x) = E(\varphi(Y)|X = x),$$

where φ is a known real-valued borel function.
It is clear that m_φ generalizes the classical regression function (set $\varphi(t) = t$) as well as the conditional distribution function (set for any $y \in \mathbb{R}, \varphi(t) = 1_{]-\infty,y]}(t)$).
Following Barrientos et al. (2010) who proved the pointwise almost complete convergence of the classical regression function estimator, the local linear estimate of m_φ is obtained as the solution for a of the following minimization problem

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} \left(\varphi(Y_i) - a - b\beta(X_i, x)\right)^2 K(h^{-1}d(X_i, x)),$$

where $\beta(.,.)$ is a known operator from $\mathcal{F} \times \mathcal{F}$ into $\mathbb{R}$ such that, $\forall x \in \mathcal{F}, \beta(x, x) = 0$, the function K is a kernel and $h := h_n$ is a sequence of strictly positive real numbers which plays a smoothing parameter role.
This approach assumes that $a + b\beta(., x)$ is a good approximation of $m_\varphi(.)$ around x. As $\beta(x, x) = 0$, a will be a suitable estimate for $m_\varphi(x)$.
By a simple calculus, one's can derive the following explicit estimator

$$\widehat{m}_\varphi(x) = \frac{\sum_{i,j=1}^{n} W_{ij}(x)\varphi(Y_j)}{\sum_{i,j=1}^{n} W_{ij}(x)} \quad \left(\frac{0}{0} := 0\right), \tag{2.1}$$

with the convention $0/0 := 0$, where

$$W_{ij}(x) = \beta(X_i, x)\left(\beta(X_i, x) - \beta(X_j, x)\right) K(h^{-1}d(X_i, x))K(h^{-1}d(X_j, x)).$$

We investigate the asymptotic behaviour of the local linear estimator $\widehat{m}_\varphi(x)$ for a fixed point x in $\mathcal{F}$, under the assumptions (H1), (H3)–(H7) in chapter 1 and the addition followings assumptions.

(H2) $m_\varphi \in \{f : \mathcal{F} \to \mathbb{R}, \lim_{d(x,x')\to 0} f(x') = f(x)\}$

(H2') $m_\varphi \in \{f : \mathcal{F} \to \mathbb{R}, \exists b > 0, \forall x' \in \mathcal{F}; \quad |f(x) - f(x')| \leq C_x d^b(x,x')\}$, where C_x is a positive constant depending on x.

(H8') $\forall m \geq 2, \sigma_m : x \longmapsto E(\varphi(Y)^m/X)$ is a continuous operator on $\mathcal{F}$.

Remark that our hypotheses are very similar to the assumed conditions (H2$_C$), (H2$_L$) and (H8) in Section 1.1.

Let us state the pointwise almost-complete convergence (a.co.) of $\widehat{m}_\varphi(x)$, along with a rate.

Theorem 2.1. *Assume that assumptions (H1), (H3)–(H7) and (H8') are satisfied.*
(i) Under the additional hypothesis (H2), we have

$$\widehat{m}_\varphi(x) - m_\varphi(x) = o_{a.co.}(1).$$

(ii) If in addition (H2') is satisfied, we get

$$\widehat{m}_\varphi(x) - m_\varphi(x) = O(h^b) + O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

Notice that the proof of this theorem is based on a standard decomposition given for all $x \in \mathcal{F}$, by

$$\widehat{m}_\varphi(x) - m_\varphi(x) = \frac{1}{m_0(x)}\left[(m_1(x) - Em_1(x)) - (m_\varphi(x) - Em_1(x))\right] - \frac{m_\varphi(x)(m_0(x)-1)}{m_0(x)} \quad (2.2)$$

where, for all $l = 0, 1$

$$m_l(x) = \frac{1}{n(n-1)EW_{12}(x)}\sum_{i\neq j} W_{ij}(x)\varphi^l(Y_j).$$

The study of each term of this decomposition can be carried out exactly as done in the proof of Lemma 1.1 and Lemma 1.2 of the previous chapter with replacing Y by $\varphi(Y)$, so for the sake of avoiding repetitions, we omit the proof.
Now, we will focus on the uniform consistency.

2.2 The uniform almost-complete convergence

2.2.1 The estimator $\widehat{m}_\varphi$

We will establish the uniform almost-complete convergence of $\widehat{m}_\varphi$ on some subset $S_\mathcal{F}$ of $\mathcal{F}$ which can be covered by a finite number of balls. This number has to be related to the radius of these balls (see hypothesis (U5)).
We suppose that $x_1, \ldots, x_{N_{r_n}(S_\mathcal{F})}$ is an r_n-net for $S_\mathcal{F}$ where for all $k \in \{1, \ldots, N_{r_n}(S_\mathcal{F})\}$, $x_k \in S_\mathcal{F}$ and (r_n) is a sequence of positive real numbers.
In this study, we need the following assumptions.

(U1) There exist a differentiable function Φ and strictly positive constants C, C_1 and C_2 such that

$$\forall x \in S_\mathcal{F}, \forall h > 0;\ 0 < C_1\Phi(h) \leq \Phi_x(h) \leq C_2\Phi(h) < \infty$$

and

$$\exists \eta_0 > 0, \forall \eta < \eta_0, \Phi'(\eta) < C,$$

where Φ' denotes the first derivative of Φ with $\Phi(0) = 0$.

(U2) The generalized regression function m_φ satisfies:

$$\exists C > 0, \exists b > 0, \forall x \in S_\mathcal{F}, x' \in B(x,h), |m_\varphi(x) - m_\varphi(x')| \leq Cd^b(x,x').$$

(U3) The function $\beta(.,.)$ satisfies (H3) uniformly on x and the following Lipschitz's condition

$$\exists C > 0, \forall x_1 \in S_{\mathcal{F}}, x_2 \in S_{\mathcal{F}}, x \in \mathcal{F}, |\beta(x, x_1) - \beta(x, x_2)| \leq Cd(x_1, x_2).$$

(U4) The kernel K fulfills (H4) and is Lipschitzian on $[0, 1]$.

(U5) $\lim_{n\to\infty} h = 0$, and for $r_n = O\left(\frac{\ln n}{n}\right)$, the function $\psi_{S_{\mathcal{F}}}$ satisfies for n large enough:

$$\frac{(\ln n)^2}{n\Phi(h)} < \psi_{S_{\mathcal{F}}}(\left(\frac{\ln n}{n}\right) < \frac{n\Phi(h)}{\ln n},$$

and

$$\sum_{n=1}^{\infty} \exp\{(1-\beta)\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)\} < \infty,$$

for some $\beta > 1$.

(U6) The bandwidth h satisfies: $\exists n_0 \in \mathbb{N}$, $\exists C > 0$, such that

$$\forall n > n_0, \forall x \in S_{\mathcal{F}}, \frac{1}{\Phi_x(h)} \int_0^1 \Phi_x(zh, h) \frac{d}{dz}\left(z^2 K(z)\right) > C > 0$$

and

$$h \int_{B(x,h)} \beta(u, x) dP_X(u) = o\left(\int_{B(x,h)} \beta^2(u, x) dP_X(u)\right)$$

uniformly on x.

(U7) $\exists C > 0$ such that $\forall m \geq 2 : E(|\varphi(Y)|^m / X = x) < \delta_m(x) < C < \infty$ with $\delta_m(.)$ continuous on $S_{\mathcal{F}}$.

Roughly speaking, these hypotheses are uniform version of the assumed conditions in the pointwise case and have already been used in the literature. We refer to Messaci et al. (2015) for conditions (U1), (U3), (U4) and (U6) and to Ferraty et al. (2010) for assumptions (U2), (U5) and (U7).
The claimed result is as follows.

Theorem 2.2. *Under assumptions (U1)–(U7), we have*

$$\sup_{x\in S_{\mathcal{F}}} |\widehat{m}_{\varphi}(x) - m_{\varphi}(x)| = O(h^b) + O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right).$$

We can readily deduce the uniform consistency of the estimator studied in Barrientos et al. (2010) for which, to the best of our knowledge, only the pointwise convergence is available.
This result shows that, contrary to the finite case, the rate of convergence obtained may differ from that of the pointwise consistency, it is function of the entropy of the subset on which the uniform convergence states.
It is easy to see that the proof of Theorem 2.2 is a direct consequence of the decomposition (2.2) and of the following lemmas for which the proofs are relegated to the Appendix 2.4.

Lemma 2.1. *Assume that hypotheses (U1), (U2) and (U4) hold, then:*

$$\sup_{x\in S_{\mathcal{F}}} |m_{\varphi}(x) - Em_1(x)| = O(h^b).$$

Lemma 2.2. *Under assumptions of Theorem 2.2, we obtain that:*

$$\sup_{x\in S_{\mathcal{F}}} |m_1(x) - Em_1(x)| = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right).$$

Lemma 2.3. *If assumptions (U1),(U3)–(U6) are satisfied, we get:*

$$\sup_{x\in S_{\mathcal{F}}} |m_0(x) - 1| = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right)$$

and

$$\sum_{n=1}^{\infty} P\left(\inf_{x\in S_{\mathcal{F}}} m_0(x) < \frac{1}{2}\right) < \infty.$$

2.2.2 A conditional quantile estimator

Let $F_x(y) = P(Y \leq y | X = x)$ be the conditional distribution function of Y given $X = x$ where y is real and x is a fixed object in $\mathcal{F}$. To estimate it, we treat this function as a particular case of m_φ with $\varphi(t) = 1_{]-\infty, y]}(t)$ for $y \in \mathbb{R}$. Thus, we estimate $F^x(y)$ by

$$\widehat{F}^x(y) = \frac{\sum_{i,j=1}^n W_{ij}(x) 1_{\{Y_j \leqslant y\}}}{\sum_{i,j=1}^n W_{ij}(x)}, \tag{2.3}$$

where $W_{ij}(x)$ is defined in (2.1).
The conditional quantile of order α ($\alpha \in (0,1)$) is $t_\alpha(x) = \inf\{y \in \mathbb{R}, F^x(y) \geqslant \alpha\}$. So, we deduce from $\widehat{F}^x$ a natural conditional quantile estimator as,

$$\widehat{t}_\alpha(x) = \inf\{y \in \mathbb{R}, \widehat{F}^x(y) \geqslant \alpha\}. \tag{2.4}$$

Notice that $t_{1/2}(x)$ is the so called conditional median.
To investigate the asymptotic convergence of $\widehat{F}^x(y)$, we introduce the following standard conditions.

(U2)' There exist $\delta > 0$, $C > 0$ and $b > 0$, such that for any $x \in S_\mathcal{F}, x' \in B(x, h)$ and $y \in [t_\alpha(x) - \delta, t_\alpha(x) + \delta]$, we have

$$|F^{x'}(y) - F^x(y)| \leq C d^b(x, x').$$

(U5)' $\lim_{n\to\infty} h = 0$, and for $r_n = O\left(\frac{\ln n}{n}\right)$, the function $\psi_{S_\mathcal{F}}$ satisfies for n large enough:

$$\frac{(\ln n)^2}{n\Phi(h)} < \psi_{S_\mathcal{F}}(\left(\frac{\ln n}{n}\right) < \frac{n\Phi(h)}{\ln n},$$

and

$$\sum_{n=1}^{\infty} n^{(\xi+1/2)} \exp\{(1-\beta)\psi_{S_\mathcal{F}}\left(\frac{\ln n}{n}\right)\} < \infty,$$

for some $\beta > 1$ and $\xi > 0$.

The following result concerns the uniform almost complete convergence of $\widehat{F}^x(y)$.

Theorem 2.3. *Under assumptions (U1), (U2)', (U3), (U4), (U5)' and (U6), we have*

$$\sup_{x\in S_{\mathcal{F}}}\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]}|\widehat{F}^x(y)-F^x(y)| = O(h^b)+O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right).$$

To prove this theorem we make use of the decomposition given, for all x and y, by

$$\widehat{F}^x(y)-F^x(y) = \frac{1}{m_0(x)}\left[\left(\widehat{F}^x_N(y)-E\widehat{F}^x_N(y)\right)\left(F^x(y)-E\widehat{F}^x_N(y)\right)\right]-\frac{F^x(y)}{m_0(x)}(m_0(x)-1) \quad (2.5)$$

where $\widehat{F}^x_N(y) = \frac{1}{n(n-1)EW_{12}(x)}\sum_{i\neq j}W_{ij}(x)1_{\{Yj\leqslant y\}}$ and $m_0(x)$ is defined in (1.5). Now, it sufficies to apply Lemma 2.3 together with the following lemmas.

Lemma 2.4. *Assume that hypotheses (U1), (U2)' and (U4) hold, then*

$$\sup_{x\in S_{\mathcal{F}}}\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]}\left|F^x(y)-E\widehat{F}^x_N(y)\right| = O(h^b).$$

Lemma 2.5. *Under assumptions of Theorem 2.3, we obtain that*

$$\sup_{x\in S_{\mathcal{F}}}\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]}\left|\widehat{F}^x_N(y)-E\widehat{F}^x_N(y)\right| = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right).$$

To obtain the uniform consistency of the conditional quantile estimator, we introduce the following conditions used for example in Messaci et al. (2015).

(U8) $\forall\epsilon>0$, $\exists\xi>0$ such that for any function g_α from $S_{\mathcal{F}}$ into $[t_\alpha(x)-\delta,t_\alpha(x)+\delta]$ we have

$$\sup_{x\in S_{\mathcal{F}}}|t_\alpha(x)-g_\alpha(x)|\geqslant\epsilon \quad \textit{implies} \quad \sup_{x\in S_{\mathcal{F}}}|F^x(t_\alpha(x))-F^x(g_\alpha(x))|\geqslant\xi.$$

(U9) $\exists j > 1$, $\forall x \in S_{\mathcal{F}}$, F^x is j-times continuously differentiable on $]t_\alpha(x) - \delta, t_\alpha(x) + \delta[$ with respect to y and satisfies $F^{x(l)}(t_\alpha(x)) = 0$ if $0 \leq l < j$, $F^{x(j)}(t_\alpha(x)) > C > 0$ and $F^{x(j)}$ is uniformly continuous on $[t_\alpha(x) - \delta, t_\alpha(x) + \delta]$ where $F^{x(l)}$ stands for the lth-order derivative of F^x .

A known method can be applied to derive the following result from Theorem 2.3, see for example the proof of Corollary 3.1 in Messaci et al. (2015).

Corollary 2.1. *Under the hypotheses of Theorem 2.3 and if (U8)and (U9) are satisfied, we obtain*

$$\sup_{x \in S_{\mathcal{F}}} \left| \widehat{t}_\alpha(x) - t_\alpha(x) \right| = O(h^b) + O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}} \right).$$

2.3 A Real data application

In this section, we use a real data set to illustrate the efficacy of the studied method through our conditional median estimator $\widehat{t}_{1/2}$. More precisely, we compare this last estimator to two other conditional median estimators: the first is based on the kernel method (denoted KM) and is studied in Ferraty and Vieu (2006) and the second is based on the local linear method (denoted LLM) and is introduced in Messaci et al. (2015).

For this purpose, we use the spectrometric data set which can be found at http ://lib.stat. cmu.edu/datasets/tecator. These data consist of 215 pairs $(Xi, Yi)_{i=1,\ldots,215}$. For each i, the spectrometric curve X_i is the spectra of a finely chopped meat and Y_i is the the corresponding fat content obtained by an analytical chemical process. Our goal is to predict the fat content in a piece of meat from its spectrometric curve. For this, we estimate the median $t_{1/2}(x)$ of the conditional distribution by $\widehat{t}_{1/2}(x)$.

We split these real data into a learning sample containing the first 160 units used to build the estimator and a test sample containing the last 55 units used to predict the fat content and to make a comparison.

The KM (resp. the LLM) estimator is computed with the same parameters as at subsection 12.4 in Ferraty and Vieu (2006) (resp. at section 4 in Messaci et al. (2015)). For the computation of the estimator $\widehat{t}_{1/2}(x)$, we use the quadratic kernel $K(x) = \frac{3}{2}(1-x^2)1_{[0,1]}(x)$, the bandwith h is chosen by a 2-fold cross-validation method, the semi-metric d is based on the derivative described in Ferraty and Vieu (2006) (see routines "semimetric.deriv" in the website http://www.lsp.ups-tlse.fr/staph/npfda) and $\beta = d$.

To illustrate the performance of our estimator, we first plot the true values (provided in the test sample) against the predicted ones by means of the three estimators (one in each graph). This is displayed in Figure 2.1. Secondly, to be more precise we evaluate their empirical Mean Square Errors (MSE), defined by

$$MSE := \frac{1}{55}\sum_{i=1}^{55}\left(\widehat{Y}_i - Y_i\right)^2,$$

where Yi (resp. $\widehat{Y}_i$) is the true (resp. the estimated) value.

The obtained results are

MSE($\widehat{t}_{1/2}$)=3.22, MSE(LLM)=3.8 and MSE(KM)=4.8.

This shows that the estimator $\widehat{t}_{1/2}$ performs well and that the local linear method seems to improve the quality of the prediction even for functional data.

2.4 Appendix 2

In what follows, let C be some strictly positive generic constant.

To treat the uniform convergence of $\hat{m}_\varphi(x)$, we need to make use of Lemma 4.1 introduced in Messaci et al. (2015) and stated here as follows.

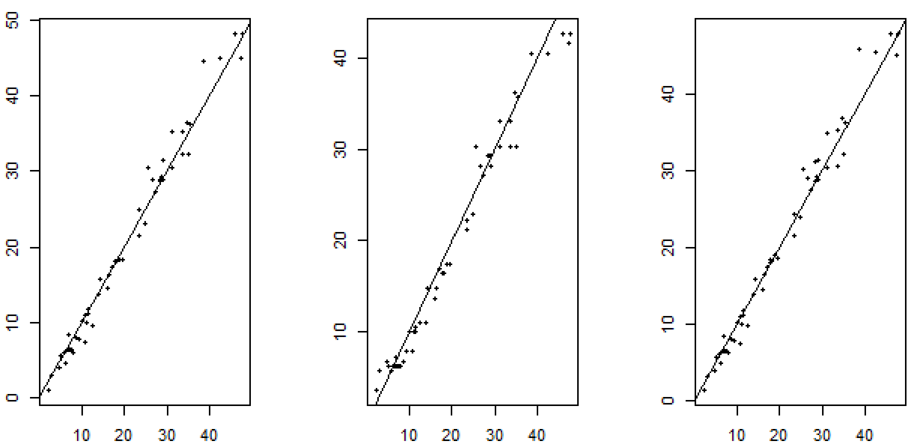

Figure 2.1: From left to right: the estimator $\widehat{t}_{1/2}$, the KM estimator and the LLM estimator for the spectrometric data.

Lemma 2.6. *Under assumptions (U1),(U3),(U4) and (U6), we obtain that:*
i) $\forall (p,l) \in \mathbb{N}^\star \times \mathbb{N}$, $\sup_{x\in S_\mathcal{F}} E\left(K_1^p(x)|\beta_1^l(x)|\right) \leq Ch^l\Phi(h)$.
ii) $\inf_{x\in S_\mathcal{F}} E\left(K_1(x)\beta_1^2(x)\right) > Ch^2\Phi(h)$.

Proof of Lemma 2.1 We have

$$Em_l(x) = \frac{1}{E(W_{12}(x))}E(W_{12}(x)\varphi^l(Y_2)),$$

and $Em_1(x)$ can also be written as

$$Em_1(x) = E\left(E(m_1(x)|X_2)\right) = \frac{1}{E(W_{12}(x)}E\left(W_{12}(x)E(\varphi(Y_2)|X_2)\right).$$

So, we get under assumption (U4)
$|m_\varphi(x) - Em_1(x)| = \frac{1}{|E(W_{12}(x))|} \mid E\left(W_{12}(x)(m_\varphi(x) - m_\varphi(X_2))\right)| \leq \sup_{x\prime\in B(x,h)} |m_\varphi(x) - m_\varphi(x')|$.
We need to take into account hypothesis (U2) to obtain

$$\sup_{x\in S_\mathcal{F}} |m_\varphi(x) - Em_1(x)| = O(h^b).$$

Proof of Lemma 2.2 We use again the follwing decomposition

$$m_1(x) = Q(x) \left[S_{2,1}(x) S_{4,0}(x) - S_{3,1}(x) S_{3,0}(x) \right], \tag{2.6}$$

where, for $p = 2, 3, 4$, and $l = 0, 1$,

$$S_{p,l}(x) = \frac{1}{n\Phi_x(h)} \sum_{i=1}^{n} \frac{K_i(x)\beta_i^{p-2}(x)\varphi^l(Y_i)}{h^{p-2}}$$

and

$$Q(x) = \frac{n^2 h^2 \Phi_x^2(h)}{n(n-1)E\left(W_{12}(x)\right)}.$$

By following the same steps as in the proof of lemma 1.2, and using lemma 2.6 instead of lemma 1.3, we obtain under the assumptions (U1)–(U4) and (U6),

$$\sup_{x \in S_{\mathcal{F}}} Q(x) = O(1), \quad \sup_{x \in S_{\mathcal{F}}} E(S_{p,l}(x)) = O(1),$$

uniformly on x, for $p = 2, 3, 4$, $l = 0, 1$,

$$\sup_{x \in S_{\mathcal{F}}} \left| E(S_{2,1}(x)) E(S_{4,0}(x)) - E(S_{2,1}(x) S_{4,0}(x)) \right| = O\left(\frac{1}{n\Phi(h)}\right),$$

and

$$\sup_{x \in S_{\mathcal{F}}} \left| E(S_{3,1}(x)) E(S_{3,0}(x)) - E(S_{3,1}(x) S_{3,0}(x)) \right| = O\left(\frac{1}{n\Phi(h)}\right),$$

which is, in view of hypothesis $(U5)$, equals to $O\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right)$.

We need to check that for $p = 2, 3, 4$ and $l = 0, 1$,

$$\sup_{x \in S_{\mathcal{F}}} \left| S_{p,l}(x) - E(S_{p,l}(x)) \right| = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right).$$

To satisfy this aim, let us set

$$j(x) = \arg \min_{j \in \{1, 2, \ldots, N_{r_n}(S_{\mathcal{F}})\}} d(x, x_j),$$

and consider the following decomposition

$$\sup_{x\in S_{\mathcal{F}}} |S_{p,l}(x) - ES_{p,l}(x)| \le \sup_{x\in S_{\mathcal{F}}} \left|S_{p,l}(x) - S_{p,l}(x_{j(x)})\right| + \sup_{x\in S_{\mathcal{F}}} \left|S_{p,l}(x_{j(x)}) - ES_{p,l}(x_{j(x)})\right| + \sup_{x\in S_{\mathcal{F}}} \left|ES_{p,l}(x_{j(x)}) - ES_{p,l}(x)\right| := F_1^{p,l} + F_2^{p,l} + F_3^{p,l}.$$

Let's, now, study each term $F_k^{p,l}$ for $k = 1, 2, 3$.

Study of the terms $F_1^{p,l}$ and $F_3^{p,l}$.
First, let us analyze the term $F_1^{p,l}$. Since K is supported in $[0, 1]$ and according to (U1), we can write for all $p = 2, 3, 4$

$$\begin{aligned} F_1^{p,l} &\le \frac{C}{nh^{p-2}\Phi(h)} \sup_{x\in S_{\mathcal{F}}} \sum_{i=1}^{n} \Big| K_i(x)\beta_i^{p-2}(x)\varphi^l(Y_i)1_{B(x,h)}(X_i) - K_i(x_{j(x)})\beta_i^{p-2}(x_{j(x)})\varphi^l(Y_i)1_{B(x_{j(x)},h)}(X_i) \\ &\le \frac{C}{nh^{p-2}\Phi(h)} \sup_{x\in S_{\mathcal{F}}} \sum_{i=1}^{n} K_i(x)1_{B(x,h)}(X_i)|\varphi^l(Y_i)| \left|\beta_i^{p-2}(x) - \beta_i^{p-2}(x_{j(x)})1_{B(x_{j(x)},h)}(X_i)\right| \\ &+ \frac{C}{nh^{p-2}\Phi(h)} \sup_{x\in S_{\mathcal{F}}} \sum_{i=1}^{n} \beta_i^{p-2}(x_{j(x)})1_{B(x_{j(x)},h)}(X_i)|\varphi^l(Y_i)| \left|K_i(x)1_{B(x,h)}(X_i) - K_i(x_{j(x)})\right| \\ &:= F_{1.1}^{p,l} + F_{1.2}^{p,l}. \end{aligned}$$

Analysis of the term $F_{1.1}^{p,l}$.
According to assumption (U3), we get

$$\begin{aligned} & 1_{B(x,h)}(X_i) \left|\beta_i(x) - \beta_i(x_{j(x)})1_{B(x_{j(x)},h)}(X_i)\right| \\ & \le \; Cr_n 1_{B(x,h)\bigcap B(x_{j(x)},h)}(X_i) + Ch 1_{B(x,h)\bigcap \overline{B(x_{j(x)},h)}}(X_i) \end{aligned}$$

and

$$\begin{aligned} & 1_{B(x,h)}(X_i) \left|\beta_i^2(x) - \beta_i^2(x_{j(x)})1_{B(x_{j(x)},h)}(X_i)\right| \\ & \le \; Cr_n h 1_{B(x_{j(x)},h)\bigcap B(x,h)}(X_i) + Ch^2 1_{B(x,h)\bigcap \overline{B(x_{j(x)},h)}}(X_i). \end{aligned}$$

By grouping the cases $p = 3$ and $p = 4$, we found

$$
\begin{aligned}
&1_{B(x,h)}(X_i)\left|\beta_i^{p-2}(x) - \beta_i^{p-2}(x_{j(x)})1_{B(x_{j(x)},h)}(X_i)\right| \\
&\leq \; Cr_n h^{p-3} 1_{B(x_{j(x)},h)\bigcap B(x,h)}(X_i) + Ch^{p-2} 1_{B(x,h)\bigcap \overline{B(x_{j(x)},h)}}(X_i).
\end{aligned}
$$

which gives the following inequality

$$
\begin{aligned}
F_{1.1}^{P,l} \leq \frac{Cr_n}{nh\Phi(h)} \sup_{x\in\mathcal{S}_\mathcal{F}} \sum_{i=1}^{n} |\varphi^l(Y_i)| K_i(x) 1_{B(x,h)\bigcap B(x_{j(x)},h)}(X_i) \\
+ \frac{C}{n\Phi(h)} \sup_{x\in\mathcal{S}_\mathcal{F}} \sum_{i=1}^{n} |\varphi^l(Y_i)| K_i(x) 1_{B(x,h)\bigcap \overline{B(x_{j(x)},h)}}(X_i). \qquad (2.7)
\end{aligned}
$$

Analysis of the term $F_{1.2}^{p,l}$.
Using the following inequality

$$
\begin{aligned}
&1_{B(x_j(x),h)}(X_i)\left|K_i(x)1_{B(x,h)}(X_i) - K_i(x_{j(x)})1_{B(x,h)\bigcup\overline{B(x,h)}}(X_i)\right| \\
&\leq \; 1_{B(x,h)\bigcap B(x_{j(x)},h)}(X_i)|K_i(x) - K_i(x_j(x))| + K_i(x_j(x))1_{B(x_j(x),h)\cap\overline{B(x,h)}}(X_i)
\end{aligned}
$$

and by hypotheses (U3) and (U4), we obtain

$$
\begin{aligned}
&|\beta_i^{p-2}(x_{j(x)})| 1_{B(x_{j(x),h})}(X_i)\left|K_i(x)1_{B(x,h)}(X_i) - K_i(x_{j(x)})\right| \\
&\leq \; Ch^{p-2}\left[\frac{r_n}{h} 1_{B(x,h)\cap B(x_{j(x)},h)}(X_i) + K_i(x_{j(x)}) 1_{B(x_{j(x)},h)\cap\overline{B(x,h)}}(X_i)\right],
\end{aligned}
$$

which leads to

$$
\begin{aligned}
F_{1.2}^{p,l} \leq \frac{Cr_n}{nh\Phi(h)} \sup_{x\in S_\mathcal{F}} \sum_{i=1}^{n} |\varphi^l(Y_i)| 1_{B(x,h)\cap B(x_{j(x)},h)}(X_i) \\
+ \frac{C}{n\Phi(h)} \sup_{x\in S_\mathcal{F}} \sum_{i=1}^{n} |\varphi^l(Y_i)| K_i(x_{j(x)}) 1_{\overline{B(x,h)}\cap B(x_{j(x)},h)}(X_i).
\end{aligned}
$$

This last inequality combined with (2.7) allow us to write

$$\begin{aligned} F_1^{p,l} &\le \frac{Cr_n}{nh\Phi(h)} \sup_{x\in S_{\mathcal{F}}} \sum_{i=1}^{n} |\varphi^l(Y_i)| 1_{B(x,h)\cap B(x_{j(x)},h)}(X_i) \\ &+ \frac{C}{n\Phi(h)} \sup_{x\in S_{\mathcal{F}}} \sum_{i=1}^{n} |\varphi^l(Y_i)| K_i(x_{j(x)}) 1_{B(x_{j(x)},h)\cap \overline{B(x,h)}}(X_i) \\ &+ \frac{C}{n\Phi(h)} \sup_{x\in S_{\mathcal{F}}} \sum_{i=1}^{n} |\varphi^l(Y_i)| K_i(x) 1_{B(x,h)\cap \overline{B(x_{j(x)},h)}}(X_i). \end{aligned}$$

Taking into account hypothesis (U4), we find

$$F_1^{p,l} \le \frac{Cr_n}{nh\Phi(h)} \sup_{x\in S_{\mathcal{F}}} \sum_{i=1}^{n} |\varphi^l(Y_i)| 1_{B(x,h)\cup B(x_{j(x)},h)}(X_i). \tag{2.8}$$

Let

$$Z_i = \frac{Cr_n|\varphi^l(Y_i)|}{h\Phi(h)} \sup_{x\in S_{\mathcal{F}}} 1_{B(x,h)\cup B(x_{j(x)},h)}(X_i).$$

The assumption (U7) implies that

$$E|Z_1^m| \le \frac{Cr_n^m}{h^m\Phi(h)^{m-1}}, \tag{2.9}$$

so, by applying Proposition 5.3–(ii), with $a_n^2 = \frac{r_n}{h\Phi(h)}$,

$$\frac{1}{n}\sum_{i=1}^{n} Z_i = EZ_1 + O_{a.co.}\left(\sqrt{\frac{r_n \ln n}{nh\Phi(h)}}\right).$$

Applying (2.9) again (for $m = 1$), one gets

$$F_1^{p,l} = O(\frac{r_n}{h}) + O_{a.co.}\left(\sqrt{\frac{r_n \ln n}{nh\Phi(h)}}\right).$$

Combining this with assumption (U5) and the second part of the assumption (U1), we obtain

$$F_1^{p,l} = O_{a.co.}\left(\sqrt{\frac{\psi_{S_\mathcal{F}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right). \qquad (2.10)$$

Second, since

$$F_3^{p,l} \leq E\left(\sup_{x\in S_\mathcal{F}}\left|S_{p,l}(x) - S_{p,l}(x_{j(x)})\right|\right),$$

we deduce that

$$F_3^{p,l} = O_{a.co.}\left(\sqrt{\frac{\psi_{S_\mathcal{F}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right). \qquad (2.11)$$

Study of the term $F_2^{p,l}$.

For all $\eta > 0$, we have that

$$P\left(F_2^{p,l} > \eta\sqrt{\frac{\psi_{S_\mathcal{F}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right) = P\left(\max_{j\in\{1,\ldots,N_{r_n}(S_\mathcal{F})\}}\left|S_{p,l}(x_{j(x)}) - E(S_{p,l}(x_{j(x)})\right| > \eta\sqrt{\frac{\psi_{S_\mathcal{F}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right)$$
$$\leq N_{r_n}(S_\mathcal{F})\max_{j\in\{1,\ldots,N_{r_n}(S_\mathcal{F}\})} P\left(\left|S_{p,l}(x_{j(x)}) - E(S_{p,l}(x_{j(x)})\right| > \eta\sqrt{\frac{\psi_{S_\mathcal{F}}\left(\frac{\ln n}{n}\right.}{n\Phi(h)}}\right.$$

Let us set for $p = 2,3,4$ that

$$\Delta_{p,i} = \frac{1}{h^{p-2}\Phi_x(h)}\left[K_i(x_{j(x)})\beta_i^{p-2}(x_{j(x)})\varphi^l(Y_i) - E(K_i(x_{j(x)})\beta_i^{p-2}(x_{j(x)})\varphi^l(Y_i))\right].$$

Using the binomial Theorem, Lemma 4.8 and hypothesis (U1), (U2) and (U7), gives for $p = 2,3,4$,

$$E\left|\Delta_{p,i}\right|^m = O\left(\Phi^{-m+1}(h)\right).$$

Therefore, we can apply a Bernstein- type inequality as done in the Proposition 5.3–(i), to obtain

$$P\left(\frac{1}{n}\left|\sum_{i=1}^{n}\Delta_{p,i}\right| > \eta\sqrt{\frac{\psi_{S_\mathcal{F}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right) \leq 2\exp\left(-C\eta^2\psi_{S_\mathcal{F}}\left(\frac{\ln n}{n}\right)\right).$$

Thus, by choosing β such that $C\eta^2 = \beta$, we get

$$P\left(F_2^{p,l} > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right) \leq CN_{r_n}(S_{\mathcal{F}})^{1-\beta}. \tag{2.12}$$

Then, hypothesis (U5) allows us to write

$$F_2^{p,l} = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right). \tag{2.13}$$

Finally, the result of Lemma 2.2 follows from the relations (2.10), (2.13) and (2.15).

Proof of lemma 2.3 The first part of the claimed results can be directly deduced from the proof of Lemma 2.2 by taking, for all i, $\varphi(Y_i) = 1$.
For the second part, It comes straightforward that

$$\inf_{x\in S_{\mathcal{F}}} m_0(x) < \frac{1}{2} \Rightarrow \exists\, x \in S_{\mathcal{F}} \text{ such that } 1 - m_0(x) > \frac{1}{2} \Rightarrow \sup_{x\in S_{\mathcal{F}}} |1 - m_0(x)| > \frac{1}{2}$$

$$\Rightarrow \sum_{n=0}^{\infty} P\left(\inf_{x\in S_{\mathcal{F}}} m_0(x) < \frac{1}{2}\right) < \infty,$$

. **Proof of Lemma 2.4** We have

$$E\widehat{F}_N^x(y) = \frac{1}{EW_{12}(x)} E\left[W_{12}(x) 1_{\{Y_2 \leqslant y\}}\right]$$

and $E\widehat{F}_N^x(y)$ can also be written as

$$E\widehat{F}_N^x(y) = E\left[E(\widehat{F}_N^x(y)|X_2)\right] = \frac{1}{EW_{12}(x)} E\left[W_{12}(x) E(1_{\{Y_2 \leqslant y\}}|X_2)\right].$$

So, we get under assumption (U4)
$$\left|F^x(y) - E\widehat{F}_N^x(y)\right| = \frac{1}{|EW_{12}(x)|}\left|E\left\{W_{12}(x)\left[F^x(y) - F^{X_2}(y)\right]\right\}\right| \leq \sup_{x'\in B(x,h)} \left|F^x(y) - F^{x'}(y)\right|.$$
We need to take into account hypothesis (U2)' to obtain

$$\sup_{x\in S_{\mathcal{F}}} \sup_{y\in[t_\alpha(x)-\delta, t_\alpha(x)+\delta]} \left|F^x(y) - E\widehat{F}_N^x(y)\right| = O(h^b).$$

Proof of lemma 2.5 First, we write

$$m_1(x) = Q(x)\left[T^x_{2,1}(y)T^x_{4,0}(y) - T^x_{3,1}(y)T^x_{3,0}(y)\right], \tag{2.14}$$

where, for $p = 2, 3, 4$, and $l = 0, 1$,

$$T^x_{p,l}(y) = \frac{1}{n\Phi_x(h)}\sum_{i=1}^{n}\frac{K_i(x)\beta_i^{p-2}(x)1^l_{\{Y_i\leq y\}}}{h^{p-2}}$$

and $Q(x)$ is defined in (1.7).

By following the same steps as in the proof of lemma 1.2, and using lemma 2.6 instead of lemma 1.3, we obtain under the assumptions (U1)–(U4) and (U6),

$$\sup_{x\in S_{\mathcal{F}}} Q(x) = O(1), \quad \sup_{x\in S_{\mathcal{F}}}\ \sup_{y\in[t_\alpha(x)-\delta, t_\alpha(x)+\delta]} E(T^x_{p,l}(y)) = O(1),$$

uniformly on x, for $p = 2, 3, 4$, $l = 0, 1$,

$$\sup_{x\in S_{\mathcal{F}}}\ \sup_{y\in[t_\alpha(x)-\delta, t_\alpha(x)+\delta]} |E(T^x_{2,1}(y))E(T^x_{4,0}(y)) - E(T^x_{2,1}(y)T^x_{4,0}(y))| = O\left(\frac{1}{n\Phi(h)}\right),$$

and

$$\sup_{x\in S_{\mathcal{F}}}\ \sup_{y\in[t_\alpha(x)-\delta, t_\alpha(x)+\delta]} |E(T^x_{3,1}(y))E(T^x_{3,0}(y)) - E(T^x_{3,1}(y)T^x_{3,0}(y))| = O\left(\frac{1}{n\Phi(h)}\right),$$

which is, in view of hypothesis $(U5)'$, equals to $O\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right)$.

We need to check that for $p = 2, 3, 4$ and $l = 0, 1$,

$$\sup_{x\in S_{\mathcal{F}}}\ \sup_{y\in[t_\alpha(x)-\delta, t_\alpha(x)+\delta]} \left|T^x_{p,l}(y) - E(T^x_{p,l}(y))\right| = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right).$$

To satisfy this aim, we can cover the compact $[t_\alpha(x) - \delta, t_\alpha(x) + \delta]$ by d_n open intervals centered at y_j with radius $l_n = n^{-\xi-1/2}$ and $d_n = Cl_n^{-1}$ as the following

$$[t_\alpha(x) - \delta, t_\alpha(x) + \delta] = \cup_{j=1}^{d_n}]y_j - l_n, y_j + l_n[.$$

Set $t_y = \min_{t\in\{1,\dots,d_n\}} |y-t|$ and consider the following decomposition

$$
\begin{aligned}
\sup_{x\in S_{\mathcal{F}}}\ \sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} \left|T_{p,l}^x(y) - ET_{p,l}^x(y)\right| \le{}& \sup_{x\in S_{\mathcal{F}}}\ \sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} \left|T_{p,l}^x(y) - T_{p,l}^{x_{j(x)}}(y)\right| \\
&+ \sup_{x\in S_{\mathcal{F}}}\ \sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} \left|T_{p,l}^{x_{j(x)}}(y) - T_{p,l}^{x_{j(x)}}(t_y)\right| \\
&+ \sup_{x\in S_{\mathcal{F}}}\ \sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} \left|T_{p,l}^{x_{j(x)}}(t_y) - ET_{p,l}^{x}(t_y)\right| \\
&+ \sup_{x\in S_{\mathcal{F}}}\ \sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} \left|ET_{p,l}^{x_{j(x)}}(t_y) - ET_{p,l}^{x_{j(x)}}(y)\right| \\
&+ \sup_{x\in S_{\mathcal{F}}}\ \sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} \left|ET_{p,l}^{x_{j(x)}}(y) - ET_{p,l}^{x}(y)\right| \\
&:= L_1^{p,l} + L_2^{p,l} + L_3^{p,l} + L_4^{p,l} + L_5^{p,l}.
\end{aligned}
$$

Let's, now, study each term $L_k^{p,l}$ for $k \in \{1, ..., 5\}$.
Study of the term $L_3^{p,l}$.
For all $\varepsilon > 0$, we have

$$
\begin{aligned}
P\left(L_3^{p,l} > \eta\right) &= \sup_{x\in S_{\mathcal{F}}}\ \sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} P\left(\left|T_{p,l}^{x_{j(x)}}(t_y) - E(T_{p,l}^{x_{j(x)}}(t_y)\right| > \varepsilon\right) \\
&\le d_n N_{r_n}(S_{\mathcal{F}}) \max_{j\in\{1,\dots,N_{r_n}(S_{\mathcal{F}})\}}\ \max_{t_y\in\{t_1,\dots,t_{d_n}\}} P\left(\left|T_{p,l}^{x_{j(x)}}(t_y) - E(T_{p,l}^{x_{j(x)}}(t_y)\right| > \varepsilon\right) \\
&\le d_n N_{r_n}(S_{\mathcal{F}}) \max_{j\in\{1,\dots,N_{r_n}(S_{\mathcal{F}})\}}\ \max_{t_y\in\{t_1,\dots,t_{d_n}\}} P\left(\left|\sum_{i=1}^n \Delta_{i,p,l}^{x_{j(x)}}(t_y)\right| > n\varepsilon\Phi(h)\right),
\end{aligned}
$$

where

$$
\Delta_{i,p,l}^{x_{j(x)}} = \frac{1}{h^{p-2}}\left[K_i(x_{j(x)})\beta_i^{p-2}(x_{j(x)})1_{\{Y_i\le y\}}^l - E(K_i(x_{j(x)})\beta_i^{p-2}(x_{j(x)})1_{\{Y_i\le y\}}^l)\right].
$$

We set $\varepsilon = \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}$ with $\eta > 0$. By similar arguments as those invoked for studying $F_2^{p,l}$ and combined with (U5)' and $d_n = n^{\xi+1/2}$, ones has

$$
L_3^{p,l} = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right).
$$

Study of the terms $L_2^{p,l}$ and $L_4^{p,l}$.
First, let us analyze the term $L_2^{p,l}$, we have

$$
\begin{aligned}
L_2^{p,l} &= \sup_{x\in S_{\mathcal{F}}} \sup_{y\in[t_\alpha(x)-\delta, t_\alpha(x)+\delta]} \frac{1}{n\Phi_{x_{j(x)}}(h)} \sum_{i=1}^{n} \frac{K_i(x_{x_{j(x)}})\beta_i^{p-2}(x_{x_{j(x)}})}{h^{p-2}} |1^l_{\{Y_i\le y\}} - 1^l_{\{Y_i\le t_y\}}| \\
&\le l_n \sup_{x\in S_{\mathcal{F}}} \frac{1}{n\Phi_{x_{j(x)}}(h)} \sum_{i=1}^{n} \frac{K_i(x_{x_{j(x)}})\beta_i^{p-2}(x_{x_{j(x)}})}{h^{p-2}} \\
&\le l_n \sup_{x\in S_{\mathcal{F}}} S_{p,0}(x_{j(x)}).
\end{aligned}
$$

In view of relations (2.12) and $ES_{p,0}(x_{j(x)}) = O\left(\frac{r_n}{h}\right)$ and combined with (U5)', the fact that $l_n = n^{-\xi-1/2}$, we can derive

$$
L_2^{p,l} = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right).
$$

Second, since

$$
L_4^{p,l} \le E\left(\sup_{x\in S_{\mathcal{F}}} \left|T_{p,l}^{x_{j(x)}}(y) - T_{p,l}^{x_{j(x)}}(t_y)\right|\right),
$$

we deduce that

$$
L_4^{p,l} = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right). \tag{2.15}
$$

Study of the terms $L_1^{p,l}$ and $L_5^{p,l}$.
Because of the boundless of $1^l_{\{Y_i\le y\}}$, the study of the term $L_1^{p,l}$ is exactly the same as that of $F_1^{p,l}$ (see the proof of Lemma 2.2). So we obtain

$$
L_1^{p,l} = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right),
$$

which entails that

$$
L_5^{p,l} = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right).
$$

Proof of corollary 2.1 Since

$$\sup_{x\in S_{\mathcal{F}}}|F^x(t_\alpha(x))-F^x(\widehat{t}_\alpha(x))| = \sup_{x\in S_{\mathcal{F}}}|\widehat{F}^x(\widehat{t}_\alpha(x))-F^x(\widehat{t}_\alpha(x))| \leq \sup_{x\in S_{\mathcal{F}}}\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]}|\widehat{F}^x(y)-F^x(y)|,$$

then the condition (U8), together with Theorem 2.3, imply that

$$\lim_{n\to\infty}|\widehat{t}_\alpha(x)-t_\alpha(x)| = 0, \quad a.co. \tag{2.16}$$

Now using the Taylor expansion of the function F^x, we get under hypothesis (U9), that

$$\begin{aligned}
F^x(\widehat{t}_\alpha(x)) - F^x(t_\alpha(x)) &= \sum_{l=1}^{j-1}\frac{F^{x(l)}(t_\alpha(x))}{l!}\left[\widehat{t}_\alpha(x)-t_\alpha(x)\right]^l + \frac{F^x(t'_\alpha(x))}{j!}\left[\widehat{t}_\alpha(x)-t_\alpha(x)\right]^j \\
&= \frac{F^x(t'_\alpha(x))}{j!}\left[\widehat{t}_\alpha(x)-t_\alpha(x)\right]^j,
\end{aligned}$$

where $t'_\alpha(x)$ lies between $t_\alpha(x)$ and $\widehat{t}_\alpha(x)$.
Because of (2.16) and the uniform continuity of $F^{x(j)}$, we get that

$$\lim_{n\to\infty}\sup_{x\in S_{\mathcal{F}}}|F^{x(j)}(t'_\alpha(x)) - F^{x(j)}(t_\alpha(x))| = 0, \quad a.co. \tag{2.17}$$

So, there exists a positive real number τ such that

$$\sum_{n=0}^{\infty} P\left(\inf_{x\in S_{\mathcal{F}}} F^{x(j)}(t'_\alpha(x)) < \tau\right) < \infty.$$

Then

$$\sup_{x\in S_{\mathcal{F}}}|\widehat{t}_\alpha(x)-t_\alpha(x)|^j \leq \sup_{x\in S_{\mathcal{F}}}\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]}|\widehat{F}^x(y)-F^x(y)|.$$

It remains to apply the result of Theorem 2.3 to obtain the claimed result.

Chapter 3

Forecasting with Functional Time Series

A very widely studied problem in statistics is the link between two variables, the main goal of which is to predict one of the variables (the response variable) given a new value of the other (the explanatory variable). One way to deal with this problem is by means of the regression method which is based on the conditional expectation. In others, one alternative technique used is the conditional quantile which involves the conditional distribution function.

Notice that the nonparametric estimation based on the local linear approach for functional independant data was, for example, studied in Messaci et al. (2015) and subsection 2.2.2 for the conditional quantile.

Moreover, observed data can exhibit a dependence form. A large studied example in Time Series is the case of the α-mixing dependence. We cite for example Attaoui et al. (2014) and Laksaci et al. (2011) for papers dealing with such functional dependent data.

This chapter takes place within this field. We establish strong consistency of a local linear nonparametric estimator of the conditional distribution function of a scalar response variable given a random variable taking values in a semi metric space (the functional variable) when the collected observations are α-mixing. Then, we derive the consistency of a conditional median estimator which is a prediction tool.

3.1 Estimation and hypotheses

Let us consider n pairs of random variables $(X_i, Y_i)_{i=1,\ldots,n}$ identically distributed as the pair (X, Y) which is valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is an infinite-dimensional space equipped with a semi-metric d.
Let x, $\mathcal{N}_x$ and y be resbectively a fixed point in $\mathcal{F}$, a neighbourhood of x and a real, we estimate the conditional cumulative distribution function $F^x(y) = P(Y \leqslant y \mid X = x)$ by $\widehat{F}^x(y)$, given in (2.3). Remark that a double kernel local linear estimator was been introduced in Messaci et al. (2015) and studied for independent data.
As the conditional quantile of order α ($\alpha \in (0,1)$) is $t_\alpha(x) = \inf\{y \in \mathbb{R}, F^x(y) \geqslant \alpha\}$, we deduce from $\widehat{F}^x$ a natural conditional quantile estimator given by,

$$\widehat{t}_\alpha(x) = \inf\{y \in \mathbb{R}, \widehat{F}^x(y) \geqslant \alpha\}. \tag{3.1}$$

Recall that $t_{1/2}(x)$ is the so called conditional median.
To study the asymptotic behaviour of the local linear estimator $\widehat{F}^x$, we need the following assumptions.

(D1) There exist $\delta > 0$, $C > 0$, $b > 0$ such that: $\forall x' \in \mathcal{N}_x$, $\forall y \in [t_\alpha(x) - \delta, t_\alpha(x) + \delta]$, $|F^x(y) - F^{x'}(y)| \leq C(d^b(x, x'))$.

(D2) The kernel K is a positive and differentiable function on its support $[0,1]$ and $\exists\, C, C'$ such that

$$0 < C 1_{[0,1]}(t) \leq K(t) \leq C' 1_{[0,1]}(t) < \infty.$$

(D3) The sequence (X_i, Y_i) is a stationary α-mixing sequence with coefficient $\alpha(n)$, moreover (D3a) and (D3b) are satisfied, where

(D3a): $\exists\, C > 0, \exists a > 3, \forall n \in \mathbb{N}; \alpha(n) \le C n^{-a}$,

(D3b): $\exists\, C, C' > 0$ such that: $C' \left[\Phi_x(h)\right]^{a/(a-1)} < \psi_x(h) \le C \left[\Phi_x(h)\right]^{a/(a-1)}$,

with $\psi_x(h) := \psi_x(0,h)$ and $\psi_x(h_1,h_2) := P\left(h_1 \le d(X_1,x) \le h_2, 0 \le d(X_2,x) \le h_2\right)$.

(D4) The bandwidth h satisfies

$$\exists\, n_0 \in \mathbb{N}, \forall n > n_0,\ \frac{1}{\psi_x(h)} \int_0^1 \psi_x(zh,h) \frac{d}{dz}\left(z^2 K(z)\right) dz > C > 0$$

and

$$h^2 \int_{B(x,h)} \int_{B(x,h)} \beta(u,x)\beta(t,x) dP_{(X_1,X_2)}(u,t)$$
$$= o\left(\int_{B(x,h)} \int_{B(x,h)} \beta^2(u,x)\beta^2(t,x) dP_{(X_1,X_2)}(u,t) \right),$$

where $dP_{(X_1,X_2)}$ is the joint distribution of (X_1, X_2).

(D5) $\lim_{n\to\infty} h = 0$ and $\exists\, 0 < \eta_0 < \frac{a-3}{a+1}, \exists\, C_1 > 0$ such that $C_1 n^{\frac{3-a}{a+1}+\eta_0} \le \Phi_x(h)$.

Hypothese (D1) is a standard regularity condition allowing to deal with the bias. (D2) is a technical condition. (D3a) means that (X_i, Y_i) is arithmetically mixing and is extensively used in the literature as in Ferraty and Vieu (2006) and in Laksaci et al. (2011). (D4) is of the same kind as (H6) together with (H7) in the section 1 (Appendix 3.5 is an example of this codition). The choice of bandwidth is given by (D5), in particular it implies that $\ln n / n\Phi_x(h) \to 0$ as $n \to \infty$.

3.2 Results

Our first result concerns the asymptotic behaviour of $\widehat{F}^x(y)$.

Proposition 3.1. *Under assumptions (H1), (H3) and (D1)–(D5), we have*

$$\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} |\widehat{F}^x(y) - F^x(y)| = O(h^b) + O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

It is easy to see that the proof of Proposition 3.1 is a direct consequence of the standard decomposition given in (2.5) and of the following lemmas whose proofs are relegated to the Appendix 3.4.

Lemma 3.1. *Assume that hypotheses (H1), (H3), (D1)–(D4) hold, then*

$$\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} \left|F^x(y) - E\widehat{F}_N^x(y)\right| = O(h^b).$$

Lemma 3.2. *Under assumptions of Proposition 3.1, we obtain that*

$$\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} \left|\widehat{F}_N^x(y) - E\widehat{F}_N^x(y)\right| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

Lemma 3.3. *If assumptions (H1),(H3) and (D2)–(D5) are satisfied, we get*

$$\left|\widehat{F}_D^x - 1\right| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$$

and

$$\sum_{n=1}^{\infty} P\left(\widehat{F}_D^x < \frac{1}{2}\right) < \infty.$$

To obtain the consistency of the conditional quantile estimator, we add the following assumption.

(D6) F^x is differentiable with a continuous density f^x satisfying $f^x[t_\alpha(x)] > 0$.

A known method can be applied to derive the following result from Proposition 3.1, see for example Theorem 3.1 in Laksaci et al. (2011).

Theorem 3.1. *Under the hypotheses of Proposition 3.1 and if (D6) is satisfied, we obtain*

$$\left|\widehat{t}_\alpha(x) - t_\alpha(x)\right| = O(h^b) + O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

3.3 Real data application

In this section, a real data set will permit us to illustrate the efficacy of our studied estimator $\widehat{t}_{1/2}$ with respect to other conditional median estimators: The kernel one (denoted KM) studied in Ferraty and Vieu (2006) and the local linear estimator (denoted LLM) introduced in Messaci et al. (2015).
The KM (resp. LLM) estimator is computed with the same parameters as at subsection 12.4 in Ferraty and Vieu (2006) (resp. at section 4 in Messaci et al. (2015)). For the computation of the estimator $\widehat{t}_{1/2}$, we use the kernel $K(x) = \left[\frac{3}{2}(1-x^2) + 0,001\right] 1_{[0,1]}(x)$ (close to the quadratic kenel), the bandwidth h is chosen by the cross-validation method and the semimetric d is the PCA one described in Ferraty and Vieu (2006) (see routines "semimetric.pca" in the website http://www.lsp.ups-tlse.fr/staph/npfda with $q = 4$) and $\beta := d$.
Our aim is to study the US monthly electricity consumption observed during 338 months (from January 1973 up to February 2001) which can be found at *http://www.economagic.com.* As pointed out in Ferraty and Vieu (2006), this time series can be viewed as dependent functional data.
The consumption of a year is the explanatory variable and the consumption of each month of the following year is the response one. We eliminate the 337 and 338 months and we retain the remaining 28 years.
Fix $s \in \{1, 2, \ldots, 12\}$, in order to predict the electricity consumption of the s^{th} month of the last year (the 28^{th}) by each cited method, we use the 27 first years to define the training sample $(X_i, Y_i^s)_{(i=1,\ldots,26)}$ used to build the

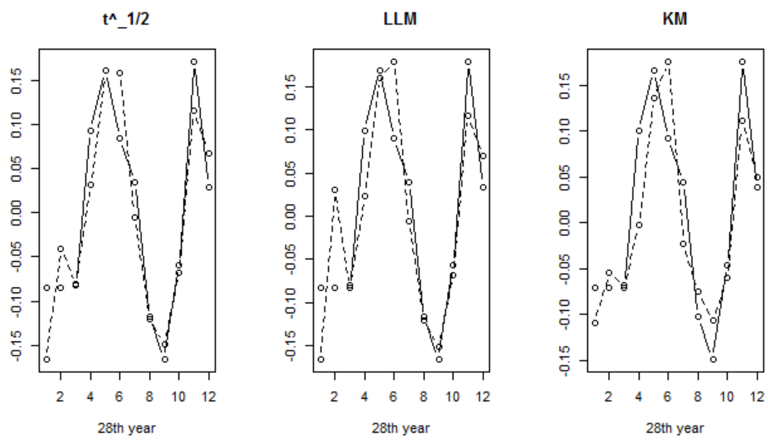

Figure 3.1: Performance of the three methods for the Electricity data.

estimators under investigation, where X_i stands for the consumption of the whole i^{th} year and Y_i^s is the consumption of the s^{th} month of the $(i+1)^{th}$ year. Then, for all $s \in \{1, 2, \ldots, 12\}$, we predict Y_{27}^s, which is the consumption of the s^{th} month of the 28^{th} year, given X_{27}.
The criteria allowing us to compare between the three estimators is the empirical Mean Square Error (MSE), defined by

$$MSE := \frac{1}{12}\sum_{i=1}^{12}\left(Y_i - \widehat{Y_i}\right)^2,$$

where Y_i (resp. $\widehat{Y_i}$) is the real (resp. the forecasted) value of the i^{th} month of the last year.
The obtained results are:
MSE($\widehat{t}_{1/2}$)=0.00235, MSE(LLM)=0.00333 and MSE(KM)=0.00253.
Based on this data set, we see that our estimator provides an acceptable performance.

In Figure 3.1 and for each mentioned method, the solid (resp. dotted) lines stand for the true (resp. forecasted) values.

3.4 Appendix 3

In what follows, let C be some strictly positive generic constant. To treat the almost-complete convergence of $\widehat{F}^x(y)$, we need the following preliminary technical lemma.

Lemma 3.4. *Under assumptions (H1), (H3), (D2), (D3b) and (D4), we obtain*

i) $\forall (p,l) \in \mathbb{N}^\star \times \mathbb{N},\ E\left(K_1^p(x)|\beta_1^l(x)|\right) \leq Ch^l \Phi_x(h)$.

ii) $\forall (p_1,p_2,l_1,l_2) \in \mathbb{N}^\star \times \mathbb{N}^\star \times \mathbb{N} \times \mathbb{N}$,

$E\left[K_1^{p_1}(x)K_2^{p_2}(x)|\beta_1^{l_1}(x)||\beta_2^{l_2}(x)|\right] \leq Ch^{(l_1+l_2)}\left[\Phi_x(h)\right]^{a/(a-1)}$.

iii) $E\left[K_1(x)K_2(x)\beta_1^2(x)\right] > Ch^2\left[\Phi_x(h)\right]^{a/(a-1)}$ *for n sufficiently large.*

Proof 3.1. **i)** *(see Lemma 1.3-i).*

ii) *In view of hypotheses (H3) and (D2), we get*

$$\begin{aligned} & E\left(K_1^{p_1}(x)K_2^{p_2}(x)|\beta_1^{l_1}(x)||\beta_2^{l_2}(x)|\right) \\ & \leq Ch^{(l_1+l_2)}E\left[1_{[0,1]}(h^{-1}d(X_1,x))1_{[0,1]}(h^{-1}d(X_2,x))\right] \\ & \leq Ch^{(l_1+l_2)}P\left[(X_1,X_2) \in B(x,h) \times B(x,h)\right]. \end{aligned}$$

So, we derive the claimed result by using (D3b).

iii) *Applying (H3), it is easy to see that*

$$E\left[K_1(x)K_2(x)\beta_1^2(x)\right] > CE\left[K_1(x)d^2(X_1,x)K_2(x)\right].$$

Combining hypothesis (D2) with Fubini's theorem, we obtain

$$E\left[K_1(x)d^2(X_1,x)K_2(x)\right] = h^2\int_0^1\int_0^1 t^2K(t)K(u)dP_{(h^{-1}d(X_1,x),h^{-1}d(X_2,x))}(t,u)$$

$$> Ch^2\int_0^1\left(\int_0^1\int_0^1 1_{[z,1]}(t)dP_{(h^{-1}d(X_1,x),h^{-1}d(X_2,x))}(t,u)\right)\frac{d}{dz}(z^2K(z))dz.$$

Moreover, we have

$$\int_0^1 \int_0^1 1_{[z,1]}(t) dP_{(h^{-1}d(X_1,x),h^{-1}d(X_2,x))}(t,u) = P\left(zh \leq d(X_1,x) \leq h, 0 \leq d(X_2,x) \leq h\right) = \psi_x(zh,h).$$

Finally (D4) permits us to end the proof.

As the dependence assumption reveals covariance terms, let us define for $p \in \{2,3,4\}$ and $l \in \{0,1\}$

$$(S^x)^2_{n,l,p}(y) = \sum_{i=1}^{n}\sum_{j=1}^{n} |Cov(\zeta^x_{i,p,l}(y), \zeta^x_{j,p,l}(y))|, \tag{3.2}$$

where, for $i \in \{1,\ldots,n\}$

$$\zeta^x_{i,p,l}(y) = \frac{1}{h^{p-2}}\left\{K_i(x)\beta_i^{p-2}(x)1^l_{\{Y_i \leqslant y\}} - E[K_i(x)\beta_i^{p-2}(x)1^l_{\{Y_i \leqslant y\}}]\right\}. \tag{3.3}$$

We now focus on these covariances terms in the following result.

Lemma 3.5. *Under assumptions (H1), (H3), (D1)–(D4) we have*

$$(S^x)^2_{n,l,k}(y) = O(n\Phi_x(h)). \tag{3.4}$$

Proof 3.2. *for $k \in \{0,2\}$ and $l \in \{0,1\}$, we set*

$$(S^x)^2_{n,l,k}(y) = \sum_{i=1}^{n}\sum_{j=1}^{n} |Cov(\zeta^x_{i,p,l}(y), \zeta^x_{j,p,l}(y))| = R_{1,n}(x) + R_{2,n}(x) + nVar(\zeta^x_{1,p,l}(y)) \tag{3.5}$$

with

$$R_{1,n}(x) = \sum_{S_1} |Cov(\zeta^x_{i,p,l}(y), \zeta^x_{j,p,l}(y))|; \quad S_1 = \left\{(i,j) : 1 \leq |i-j| \leq m_n\right\}.$$

and

$$R_{2,n}(x) = \sum_{S_2} |Cov(\zeta^x_{i,p,l}(y), \zeta^x_{i,p,l}(y))|; \quad S_2 = \left\{(i,j) : m_n + 1 \leq |i-j| \leq n-1\right\},$$

where the sequence (m_n) *will be specified below.*
Since for all i *in* $\{1, ..., n\}$, $E(\zeta^x_{i,p,l}(y)) = 0$, *we get*

$$\begin{aligned}R_{1,n}(x) &= \sum_{S_1} |E[\zeta^x_{i,p,l}(y)\zeta^x_{j,p,l}(y)]| \\ &\leq \frac{1}{h^{2(p-2)}} \sum_{S_1} \{E[K_i(x)\beta_i^{(p-2)}(x)K_j(x)\beta_j^{(p-2)}(x)] + |E[K_i(x)\beta_i^{(p-2)}(x)]||E[K_j(x)\beta_j^{(p-2)}(x)]|\}.\end{aligned}$$

Under (D3) in view of hypothesis (H3), together with the application of Lemma 1.3, we obtain

$$\begin{aligned}R_{1,n}(x) &\leq Cnm_n \left[(\Phi_x(h))^{a/a-1} + (\Phi_x(h))^2\right] \\ &\leq Cnm_n \left(\Phi_x(h)\right)^{a/a-1}.\end{aligned}$$

To apply a covariance inequality for bounded mixing sequences, we must calculate the absolute moments of the r.r.v. $\zeta^x_{i,p,l}(y)$.

$$E|\zeta^x_{i,p,l}(y)|^q \leq h^{-q(p-2)} \sum_{j=0}^{q} C_{j,q} E|K_i^j(x)\beta_i^{(p-2)j}(x)||EK_i(x)\beta_i^{(p-2)}(x)|^{q-j}.$$

By using Lemma 1.3, we get

$$\begin{aligned}E|\zeta^x_{i,p,l}(y)|^q &= O\left(max_{0\leqslant j\leqslant q}(\Phi_x(h))^{1+q-j}\right) \\ &= O\left(\Phi_x(h)\right).\end{aligned}$$

Now, we can use Rio inequality (see Proposition 5.4–(i)) and

$$\sum_{j\geq x+1} j^{-q} \leq [(a-1)x^{a-1}]^{-1},$$

together with hypothesis (D3) to obtain

$$\begin{aligned}R_{2,n}(x) &= \sum_{S_2} |Cov(\zeta^x_{i,p,l}(y)), \zeta^x_{j,p,l}(y))| \\ &\leq C \sum_{S_2} [\alpha(|i-j|)] \\ &\leq \frac{C}{a-1} nm_n^{-a+1}.\end{aligned}$$

Choosing $m_n = (\Phi_x(h))^{-1/a-1}$ *, we obtain*

$$R_{1,n}(x) = O(n\Phi_x(h)) \tag{3.6}$$

and

$$R_{2,n}(x) = O(n\Phi_x(h)). \tag{3.7}$$

For the variance term, Lemma 1.3 permit to write

$$\begin{aligned} Var(\zeta^x_{1,p,l}(y)) &\leq C\left[\Phi_x(h) + (\Phi_x(h))^2\right] \\ &\leq C\Phi_x(h). \end{aligned} \tag{3.8}$$

We readily derive the claimed result from (3.5), (3.6), (3.7) and (3.8).

Proof of Lemma 3.1 We have

$$E\widehat{F}^x_N(y) = \frac{1}{EW_{12}(x)} E\left[W_{12}(x) 1_{\{Y_2 \leqslant y\}}\right]$$

and $E\widehat{F}^x_N(y)$ can also be written as

$$E\widehat{F}^x_N(y) = E\left[E(\widehat{F}^x_N(y)|X_2)\right] = \frac{1}{EW_{12}(x)} E\left[W_{12}(x) E(1_{\{Y_2 \leqslant y\}}|X_2)\right].$$

So, we get under assumption (D2)
$\left|F^x(y) - E\widehat{F}^x_N(y)\right| = \frac{1}{|EW_{12}(x)|}\left|E\left\{W_{12}(x)\left[F^x(y) - F^{X_2}(y)\right]\right\}\right| \leq \sup_{x' \in B(x,h)} \left|F^x(y) - F^{x'}(y)\right|.$
It sufficies to take into account hypothesis (D1) to obtain the result.

Proof of Lemma 3.2 Inspired by the decomposition (1.7), we set

$$\widehat{F}^x_N(y) = Q(x)\left[T^x_{2,1}(y) T^x_{4,0}(y) - T^x_{3,1}(y) T^x_{3,0}(y)\right],$$

where

$$T^x_{p,l}(y) = \frac{1}{n\Phi_x(h)} \sum_{i=1}^{n} \frac{K_i(x)\beta_i^{p-2}(x) 1^l_{\{Yj \leqslant y\}}}{h^{p-2}}$$

and

$$Q(x) = \frac{n^2 h^2 \Phi_x^2(h)}{n(n-1)EW_{12}(x)}.$$

So, it suffices to show that, for $p \in \{2,3,4\}$ and $l \in \{0,1\}$, we have

$$\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} |ET^x_{p,l}(y)| = O(1) \text{ and } Q(x) = O(1),$$

$$\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} |T^x_{p,l}(y) - ET^x_{p,l}(y)| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right),$$

$$\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} \left|Cov\left[T^x_{2,1}(y), T^x_{4,0}(y)\right]\right| = O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$$

$$\text{and} \quad \sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} \left|Cov\left[T^x_{3,1}(y), T^x_{3,0}(y)\right]\right| = O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

• Applying Lemma 3.4 i), we readily obtain

$$\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} |ET^x_p(y)| = O(1). \tag{3.9}$$

• Treatment of the term $Q(x)$

On the one hand, we have

$$h^2 E\left[\beta_1(x)\beta_2(x)K_1(x)K_2(x)\right] \le Ch^2 \int_{B(x,h)}\int_{B(x,h)} \beta(u,x)\beta(t,x)dP_{(X_1,X_2)}(u,t).$$

On the other hand and in view of (H3) and (D3b), we obtain

$$E\left[\beta_1(x)\beta_2(x)K_1(x)K_2(x)\right] = o\left(h^2\left[\Phi_x(h)\right]^{a/(a-1)}\right).$$

Now, Lemma 3.4-(iii) and the last result allow to write, for n sufficiently large

$$Q(x) = \frac{n^2h^2\Phi_x^2(h)}{n(n-1)EW_{12}(x)} \le C\frac{\left[\Phi_x(h)\right]^2}{\left[\Phi_x(h)\right]^{a/(a-1)}} \le C.$$

• Treatment of the term $\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} \left|T^x_{p,l}(y) - ET^x_{p,l}(y)\right|$, for $p \in \{2,3,4\}$ and $l \in \{0,1\}$.

We have for any $y \in [t_\alpha(x)-\delta, t_\alpha(x)+\delta]$,

$$T^x_{p,l}(y) - ET^x_{p,l}(y) = \frac{1}{n\Phi_x(h)}\sum_{i=1}^n \zeta^x_{i,p,l}(y),$$

where $\zeta^x_{i,p,l}(y)$ is defined in relation (3.3).
By applying Proposition 5.5–(ii), we get for any $\varepsilon > 0$, $r \geq 1$ and for some $0 < C < \infty$

$$P\left(|T^x_{p,l}(y) - ET^x_{p,l}(y)| > \epsilon\right) \leq P\left(|\sum_{i=1}^{n} \zeta^x_{i,p,l}(y)| > n\epsilon\Phi_x(h)\right) \leq C\left[B_1(x) + B_2(x)\right], \tag{3.10}$$

where

$$B_1(x) = \left(1 + \frac{\epsilon^2 n^2 \left[\Phi_x(h)\right]^2}{r(S^x)^2_{n,l,k}(y)}\right)^{-r/2} \quad and \quad B_2(x) = nr^{-1}\left(\frac{r}{\epsilon n\Phi_x(h)}\right)^{a+1}.$$

Now, taking for $\eta > 0$

$$\varepsilon = \eta\sqrt{\frac{\ln n}{n\Phi_x(h)}} \text{ and } r = (\ln n)^2,$$

we obtain

$$B_2(x) \leq Cn^{1-(a+1)/2}(\ln n)^{2a-\frac{(a+1)}{2}}\left[\Phi_x(h)\right]^{-(a+1)/2},$$

and using (D5), one gets

$$B_2(x) \leq Cn^{-1-\eta_0(a+1)/2}(\ln n)^{2a-\frac{(a+1)}{2}}. \tag{3.11}$$

Moreover, in view of equation (3.4) and the fact that $\ln(x+1) = x - x^2/2 + o(x^2/2)$ where x tends to zero, we can write

$$B_1(x) \leq Cn^{-\eta^2/2}, \tag{3.12}$$

which shows that $B_1(x)$ is the general term of a convergent series for an appropriate choice of η.
Hence, by combining relations (3.10), (3.11) and (3.12), we derive

$$|T^x_{p,l}(y) - ET^x_{p,l}(y)| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

From this last result, it is easy to obtain the uniformity on the compact $[t_\alpha(x)-\delta, t_\alpha(x)+\delta]$. We omit the details because they are well known, we can see for instance the second part of the proof of Lemme 2.4 in Messaci et al. (2015).

- Finally, by following similar arguments used to prove (3.4), we obtain

$$\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} |Cov\left[T^x_{2,1}(y), T^x_{4,0}(y)\right]| = O\left(\frac{1}{n\Phi_x(h)}\right)$$

and

$$\sup_{y\in[t_\alpha(x)-\delta,t_\alpha(x)+\delta]} |Cov\left[T^x_{3,1}(y), T^x_{3,0}(y)\right]| = O\left(\frac{1}{n\Phi_x(h)}\right).$$

In view of (D5), this last rate is negligible with respect to $O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$.

Proof of Lemma 3.3 The first part of the claimed results can be directly deduced from the proof of Lemma 3.2 by taking $l = 0$ in all its proof and this easily yields to the second part.

Proof of Theorem 3.1 Following the proof of Theorem 3.1 in Laksaci et al. (2009), for any $\varepsilon > 0$ small enough, there exist δ_0 in $(0, \delta]$, such that

$$\inf_{y\in[t_\alpha(x)-\delta_0,t_\alpha(x)+\delta_0]} f^x(y) \geq C > 0,$$

so, we get for a large enough n_0

$$\sum_{n\geq n_0} P\left(|\widehat{t}_\alpha(x) - t_\alpha(x)| > \varepsilon\right) \leq \sum_{n\geq n_0} P\left(\sup_{y\in[t_\alpha(x)-\delta_0,t_\alpha(x)+\delta_0]} |\widehat{F}^x(y) - F^x(y)| > C\varepsilon\right).$$

The result is then an easy consequence of Proposition 3.1.

3.5 Appendix 4 : Remark on (D4)

In the following, we give an exemple of a random variable that satisfies the condition (D4).

Let X be a functional squared integrable random element of a separable Hilbert space $\mathcal{F}$ with orthonormal basis $\{e_j,\ \ j = 1, ..., \infty\}$. Assume that $Y = (X^1, ..., X^k)$ (where $k \in \mathbb{N}^*$) be absolutely continuous with respect to the Lebesgues measure on $\mathbb{R}^k$ and let be $B(x, h) := \{u \in \mathcal{F}/\ d_k(x, u) \leq h\}$ be a closed ball for $x \in \mathcal{F}$, where the semi-metrics d_k are usually known as projections type semi-metrics (they described in Lemma 13.6 in Ferraty and Vieu (2006)).

We assume that the density function $f_{1,2}$ of (X_1, X_2) being continuous at (x_1, x_2) and such that $f_{1,2}(x, x) > 0$, we arrive at

$$
\begin{aligned}
\psi_x(uh, h) &= P\left(uh \leq d_k(X_1, x) \leq h, 0 \leq d_k(X_2, x) \leq h\right) \\
&= P\left(d_k(X_1, x) \leq h, d_k(X_2, x) \leq h\right) - P\left(d_k(X_1, x) \leq uh, d_k(X_2, x) \leq h\right).
\end{aligned}
$$

Remark that if $x = \sum_{j=1}^{\infty} x^j e_j$ in $\mathcal{F}$ then $\exists y = (x^1, ..., x^k) \in \mathbb{R}^k$. So, On one side, we have

$$
\begin{aligned}
\psi_x(h) &:= P\left(d_k(X_1, x) \leq h, d_k(X_2, x) \leq h\right) \\
&= \int_{B(y,h)} \int_{B(y,h)} f_{1,2}(t, u) dt du \\
&= \int_{B(y,h)} \int_{B(y,h)} (f_{1,2}(y, u) + \delta_1(y, t)) dt du \\
&= \int_{B(y,h)} \int_{B(y,h)} (f_{1,2}(y, y) + \delta_2(y, u) + \delta_1(y, t)) dt du \\
&= (f_{1,2}(y, y) + O(h)) \int_{B(y,h)} \int_{B(y,h)} dt du \\
&= f_{1,2}(y, y) h^{2k} V^2(k) + O(h) h^{2k} V^2(k) \sim C_x h^{2k},
\end{aligned}
$$

where the notation $V(k)$ is the volume of the unit ball in $\mathbb{R}^k$ such that $V(k) = \frac{\pi^{\frac{k}{2}}}{\frac{k}{2}\Gamma(\frac{k}{2})}$. On the other side, we get

$$\begin{aligned} P\left(d_k(X_1,x) \leq uh, d_k(X_2,x) \leq h\right) &= \int_{B(y,uh)} \int_{B(y,h)} f_{1,2}(t,v)dtdu \\ &= \left(f_{1,2}(y,y) + O(h)\right) \int_{B(y,uh)} \int_{B(y,h)} dtdu \\ &= f_{1,2}(y,y)u^k h^{2k} V^2(k) + O(h)h^{2k}V^2(k) \sim C_x u^k h^{2k}, \end{aligned}$$

which allows us to write

$$\psi_x(uh,h) \sim C_x h^{2k}(1-u^k). \tag{3.13}$$

Now, one considers the family of kernels indexed by $\alpha > 0$ and defined by $K_\alpha(u) = \frac{\alpha+1}{\alpha}(1-u^k)1_{[0,1]}(u)$. It comes with trivial calculs that

$$\int_0^1 \psi_x(uh,h)\frac{d}{du}(u^2K(u))du = \frac{(\alpha+1)k}{(k+2)(\alpha+k+2)}C_x h^{2k} + o(h^{2k}),$$

which leads us to assumption (D4) as soon as h is small enough (i.e. as soon as n is large enough). In the same way, (D4) holds when ones considers the uniform kenel $1_{[0,1]}(.)$.

Chapter 4

Local linear estimation of a generalized regression function with functional dependent data

In Chapter 2, we study a generalized regression estimator in the case where the data are independent. The present work gives an extension to the dependent case (α-mixing) and this fact complicates considerably the study. The interest comes mainly from the fact that some application fields, for functional methods, need to analyze time series. These motivation is illustrated using two real-data sets.

Let us consider n pairs of random variables $(X_i, Y_i)_{i=1,\dots,n}$ identically distributed as the pair (X, Y) which is valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is an infinite-dimensional space equipped with a semi-metric d.

We estimate the generalized regression function m_φ, by the following explicit estimator

$$\widehat{m}_\varphi(x) = \frac{\sum_{i,j=1}^{n} W_{ij}(x)\varphi(Y_j)}{\sum_{i,j=1}^{n} W_{ij}(x)} \quad \left(\frac{0}{0} := 0\right),$$

where

$$W_{ij}(x) = \beta(X_i, x)\left(\beta(X_i, x) - \beta(X_j, x)\right) K(h^{-1}d(X_i, x))K(h^{-1}d(X_j, x)).$$

4.1 The pointwise almost-complete convergence

Let x be a fixed point in $\mathcal{F}$, We investigate the asymptotic behavior of the local linear estimator $\widehat{m}_\varphi(x)$, under the assumptions (H1), (H3), (D2), (D4) and the following addition assumptions.

(M1) There exists $b > 0$ such that for all $x_1, x_2 \in B(x, h)$, $\mid m_\varphi(x_1) - m_\varphi(x_2)| \leq C_x d^b(x_1, x_2)$, where C_x is a positive constant depending on x.

(M2) The sequence (X_i, Y_i) is an α-mixing sequence with coefficient $\alpha(n)$, moreover (M2a) and (M2b) are satisfied.

(M2a) There exist $C > 0,\ a > \sup(3, \frac{1+u}{ud})$ satisfying: $\forall n \in \mathbb{N}; \alpha(n) \leq Cn^{-a}$, where d and u are defined in (M2b) and (M4) respectively.

(M2b) There exist $0 < d \leq 1,\ C > 0,\ C' > 0$ such that $C' \left[\Phi_x(h)\right]^{1+d} < \psi_x(h) \leq C \left[\Phi_x(h)\right]^{1+d}$, where $\psi_x(h) := \psi_x(0, h)$ and $\psi_x(h_1, h_2) := P\left(h_1 \leq d(X_1, x) \leq h_2, 0 \leq d(X_2, x) \leq h_2\right)$.

(M3) For all $m \geq 2, \delta_m : x \mapsto E(|\varphi(Y)|^m / X = x)$ is a continuous operator at x and $\exists C > 0$, such that $\sup_{i \neq j} E\left(|\varphi(Y_i)\varphi(Y_j)| / (X_i, X_j)\right) \leq C < \infty$.

(M4) The bandwidth h satisfies $\lim_{n\to\infty} h = 0$ and $\exists\ \eta_0 > 0,\ 0 < u < 1$, $C_1 > 0,\ C_2 > 0$ such that $C_1 n^{\frac{3-a}{a+1}+\eta_0} \leq \Phi_x(h) \leq C_2 n^{-u}$, with $\eta_0 < \frac{a-3}{a+1}$.

Hypothesis (M1) is standard and has been assumed in the independent case (see Barrientos et al. (2010)). (M2a) means that (X_i, Y_i) is arithmetically mixing which is a standard choice of the mixing coefficient in the time

series as well as in the context of functional data. Concerning (M3), similar conditions have already been imposed in the literature to deal with the regression estimation problem: see for example assumptions (6.4) and (11.10) in Ferraty and Vieu (2006). The choice of the bandwidth is given by (M4) which implies that $n\Phi_x(h)/\ln n \to \infty$ as $n \to \infty$.
Now, let us state the rate of the pointwise almost-complete convergence of $\widehat{m}_\varphi(x)$.

Theorem 4.1. *Assume that assumptions (H1), (H3), (D2), (D4) and (M1)–(M4) are satisfied, then*

$$\widehat{m}_\varphi(x) - m_\varphi(x) = O(h^b) + O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

Notice that this rate of convergence is the same as that of Barrientos et al. (2010) (independent observations) as well as that of Laksaci et al. (2011) (dependent observations).
The proof follows directly from the standard decomposition given in (2.2), then, we apply the following lemmas for which the proofs are relegated to the Appendix 4.4.

Lemma 4.1. *Assume that hypotheses (H1), (H3), (D2), (D4), (M1) and (M2) hold, then*

$$m_\varphi(x) - Em_1(x) = O(h^b).$$

Lemma 4.2. *Under assumptions of Theorem 4.1, we have*

$$m_1(x) - Em_1(x) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

Lemma 4.3. *If assumptions (H1),(H3), (D2), (D4), (M2a), (M2b) and (M4) are satisfied, we obtain*

$$m_0(x) - 1 = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$$

and

$$\sum_{n=1}^{\infty} P\left(m_0(x) < \frac{1}{2}\right) < \infty.$$

4.2 The uniform almost-complete convergence

In this section, we establish the uniform almost-complete convergence of $\widehat{m}_\varphi$ on some subset $S_\mathcal{F}$ of $\mathcal{F}$ which can be covered by a finite number of balls. This number has to be related to the radius of these balls.
To this goal, we suppose that $x_1, \ldots, x_{N_{r_n}(S_\mathcal{F})}$ is an r_n-net for $S_\mathcal{F}$ where for all $k \in \{1, \ldots, N_{r_n}(S_\mathcal{F})\}$, $x_k \in S_\mathcal{F}$ and (r_n) is a sequence of positive real numbers.
In this study, we need the following assumptions.

(E1) The kernel K fulfills (D2) and is Lipschitzian on $[0, 1]$.

(E2) The sequence (X_i, Y_i) satisfies (M2a) and

(E2b) There exist $0 < d \leq 1$, $C_1 > 0$, $C_2 > 0$ such that for all $x_1,\ x_2 \in S_\mathcal{F}$, $0 < C_1\left[\Phi(h)\right]^{1+d} \leq P\left[(X_1, X_2) \in B(x_1, h)\times \in B(x_2, h)\right] \leq C_2\left[\Phi(h)\right]^{1+d}.$

(E3) $\forall m \geq 2, \exists\, C_1 > 0,\ E(|\varphi(Y)|^m/X) \leq C_1$ and $\exists\, C_2 > 0$, $\sup_{i\neq j} E\left(|\varphi(Y_i)\varphi(Y_j)|/(X_i, X_j)\right) \leq C_2 < \infty$.

(E4) The hypothesis (D4) is satisfied uniformly on $x \in S_\mathcal{F}$.

(E5) The bandwidth h satisfies (M4) and for $r_n = O\left(\frac{\ln n}{n}\right)$, the function $\psi_{S_\mathcal{F}}$ satisfies for n large enough $\psi_{S_\mathcal{F}}\left(\frac{\ln n}{n}\right) \sim C \ln n$.

Roughly speaking, most of these hypotheses are uniform version of the corresponding conditions in the pointwise case. (E2) and (E5) allow to treat the dependence terms and were inspired by imposed conditions in Laksaci et al. (2011) and Attaoui et al. (2014). (E1) is a technical assumption. The last

condition on entropy in (E5) is satisfied in some common cases (see example 4 on page 338 in Ferraty et al. (2010)) and leads to find again the same rate as in the pointwise case but uniformly on x.
Our result is as follows.

Theorem 4.2. *Under assumptions (U1)–(U3), (E1)–(E5) we have*

$$\sup_{x\in S_{\mathcal{F}}} |\widehat{m}_{\varphi}(x) - m_{\varphi}(x)| = O(h^b) + O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi(h)}}\right).$$

Notice that this rate of convergence is the same as that of Ferraty et al. (2010) (independent observations) under our hypothesis (E5).
It is easy to see that the proof of Theorem 4.2 is a direct consequence of the decomposition (1.5) and of the following lemmas for which the proofs are postponed to the Appendix 4.4.

Lemma 4.4. *Assume that hypotheses (U1)–(U3), (E1), (E2) and (E4) hold, then*

$$\sup_{x\in S_{\mathcal{F}}} |m_{\varphi}(x) - Em_1(x)| = O(h^b).$$

Lemma 4.5. *Under assumptions of Theorem 4.2, we obtain that*

$$\sup_{x\in S_{\mathcal{F}}} |m_1(x) - Em_1(x)| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi(h)}}\right).$$

Lemma 4.6. *If assumptions (U1),(U3), (E1), (M2a), (E2b), (E4) and (E5) are satisfied, we get*

$$\sup_{x\in S_{\mathcal{F}}} |m_0(x) - 1| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi(h)}}\right)$$

and

$$\sum_{n=1}^{\infty} P\left(\inf_{x\in S_{\mathcal{F}}} m_0(x) < \frac{1}{2}\right) < \infty.$$

4.3 Real data application

In this section, we use two real data sets (the electricity consumption data given in Section 3.3 and El Nino data) to illustrate the efficacy of the local linear estimator (LLR) corresponding to the studied estimator $\widehat{m}_\varphi$ for $\varphi(t) = t$. More precisely, we compare it to the conditional expectation kernel estimator (KR) studied in Ferraty and Vieu (2006).

The (KR) estimator is computed with the same parameters as at subsection 12.4 in Ferraty and Vieu (2006). For the computation of the (LLR) estimator, we use the quadratic kernel $K(x) = \frac{3}{2}(1-x^2)1_{[0,1]}(x)$, the bandwidth h is chosen by the cross-validation method and the semimetric d is the PCA one described in Ferraty and Vieu (2006) (see routines "semimetric.pca" in the website http://www.lsp.ups-tlse.fr/staph/npfda with $q = 5$ for the electricity consumption data and $q = 2$ for El Nino time series) and $\beta := d$.

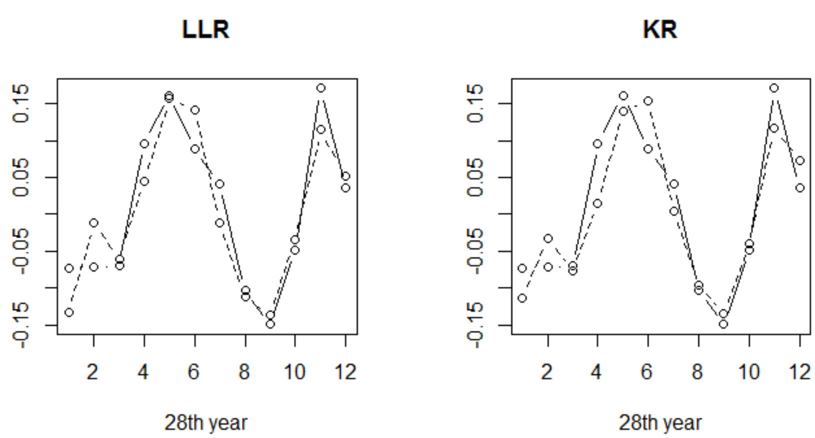

Figure 4.1: Performance of the two methods for the Electricity data.

The study of El Nino time series is the same as the lectricity consumption data but here we study the monthly Sea surface Temperature from June, 1950 up to May, 2004 (available online at http://www.cpc.ncep.noaa.gov/data/indices/)). Our aim is to predict the Sea surface Temperature of each month of the 54^{th} year given the 53^{th} one.

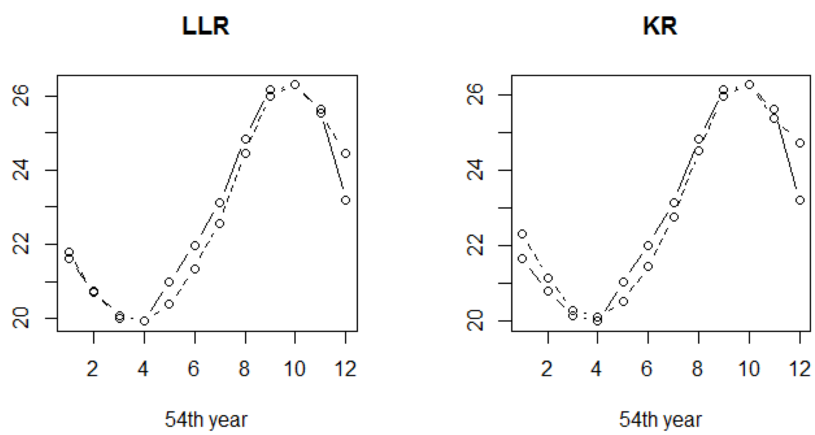

Figure 4.2: Performance of the two methods for El Nino data.

In Figures 4.1 and 4.2, the solid lines (resp. the dotted ones) correspond to the true observations (resp. the forecasted ones obtained by the mentioned method). In order to get a more precise comparison between the two estimators, we evaluate their empirical Mean Square Errors (MSE), with

$$MSE := \frac{1}{12}\sum_{i=1}^{12}\left(Y_i - \widehat{Y_i}\right)^2,$$

where Y_i (resp. $\widehat{Y_i}$) is the real (resp. the estimated) value of the i^{th} month of the last year.

We find

• **For the Electricity data**

MSE(LLR)=0.0016 and MSE(KR)=0.0024.

• **For El Nino time series**

MSE(LLR)=0.2558 and MSE(KR)=0.3500.

In conclusion, as for independent data, the local linear method seems to improve the quality of prediction compared to the kernel one. The (LLR) estimator which have been introduced and studied, for independent data, in Barrientos et al. (2010) also gives good results for dependent observations.

4.4 Appendix 5

In what follows, let C be some strictly positive generic constant.

1. To treat the pointwise almost-complete convergence of $\hat{m}_\varphi(x)$, we need the following preliminary technical lemma.

Lemma 4.7. *Under assumptions (H1), (H3), (D2), (M2b) and (D4), we obtain*

i) $\forall (p,l) \in \mathbb{N}^\star \times \mathbb{N},\ E\left(K_1^p(x)|\beta_1^l(x)|\right) \leq Ch^l \Phi_x(h).$

ii) $\forall (p_1,p_2,l_1,l_2) \in \mathbb{N}^\star \times \mathbb{N}^\star \times \mathbb{N} \times \mathbb{N},$

$E\left[K_1^{p_1}(x)K_2^{p_2}(x)|\beta_1^{l_1}(x)||\beta_2^{l_2}(x)|\right] \leq Ch^{(l_1+l_2)}\left[\Phi_x(h)\right]^{1+d}.$

iii) $E\left[K_1(x)K_2(x)\beta_1^2(x)\right] > Ch^2\left[\Phi_x(h)\right]^{1+d}$ *for n sufficiently large.*

The proof of this Lemma works as that of Lemma 3.4.

As the dependence assumption reveals covariances terms, let us define for $k \in \{0,2\}$ and $l \in \{0,1\}$

$$S_{n,l,k}^2(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} |Cov(\Lambda_i^{(k,l)}(x), \Lambda_j^{(k,l)}(x))|, \tag{4.1}$$

where, for $i \in \{1, \ldots, n\}$

$$\Lambda_i^{(k,l)}(x) = \frac{1}{h^k} \left\{ K_i(x)\beta_i^k(x)\varphi^l(Y_i) - E[K_i(x)\beta_i^k(x)\varphi^l(Y_i)] \right\}. \tag{4.2}$$

We now focus on these covariances terms in the following result.

Lemma 4.8. *Under assumptions (H1), (H3), (D2), (D4), (M1)–(M3) we have*

$$S^2_{n,l,k}(x) = O(n\Phi_x(h)). \tag{4.3}$$

Proof 4.1. *for $k \in \{0, 2\}$ and $l \in \{0, 1\}$, we set*

$$S^2_{n,l,k}(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} |Cov(\Lambda_i^{(k,l)}(x), \Lambda_j^{(k,l)}(x))| = J_{1,n}(x) + J_{2,n}(x) + nVar(\Lambda_1^{(k,l)}(x)) \tag{4.4}$$

with

$$J_{1,n}(x) = \sum_{S_1} |Cov(\Lambda_i^{(k,l)}(x), \Lambda_j^{(k,l)}(x))|; \quad S_1 = \{(i,j) : 1 \leq |i-j| \leq m_n\}.$$

and

$$J_{2,n}(x) = \sum_{S_2} |Cov(\Lambda_i^{(k,l)}(x), \Lambda_j^{(k,l)}(x))|; \quad S_2 = \{(i,j) : m_n + 1 \leq |i-j| \leq n-1\},$$

where the sequence (m_n) will be specified below.
Since for all i in $\{1, ..., n\}$, $E(\Lambda_i^{(k,l)}(x)) = 0$, we get

$$\begin{aligned}
J_{1,n}(x) &= \sum_{S_1} |E[\Lambda_i^{(k,l)}(x)\Lambda_j^{(k,l)}(x)]| \\
&\leq \frac{1}{h^{2k}} \sum_{S_1} \{E[K_i(x)\beta_i^k(x)K_j(x)\beta_j^k(x)E(|\varphi^l(Y_i)\varphi^l(Y_j)||(X_i, X_j)] \\
&+ |E[K_i(x)\beta_i^k(x)E(\varphi^l(Y_i)|X_i)]||E[K_j(x)\beta_j^k(x)E(\varphi^l(Y_j)|X_j]|\}.
\end{aligned}$$

Under (M3) and because $E[\varphi(Y)|X] = m_\varphi(X) = m_\varphi(x) + o(1)$ in view of hypothesis (M1), together with the application of Lemma 4.7, we obtain

$$\begin{aligned}
J_{1,n}(x) &\leq Cnm_n \left[(\Phi_x(h))^{1+d} + (\Phi_x(h))^2\right] \\
&\leq Cnm_n (\Phi_x(h))^{1+d}.
\end{aligned}$$

To apply a covariance inequality for unbounded mixing sequences, we must calculate the absolute moments of the r.r.v. $\Lambda_i^{(k,l)}(x)$.

$$E|\Lambda_i^{(k,l)}(x)|^q \leq h^{-qk} \sum_{j=0}^{q} C_{j,q} E|K_i^j(x)\beta_i^{kj}(x)\varphi^{lj}(Y_i)||EK_i(x)\beta_i^k(x)\varphi^l(Y_i)|^{q-j}$$
$$\leq h^{-qk} \sum_{j=0}^{q} C_{j,q} E[K_i^j(x)\beta_i^{kj}(x)E(|\varphi(Y_i)|^{lj}|X_i)][E(K_i(x)\beta_i^k(x)E(|\varphi^l(Y_i)||X_i))]^{q-j},$$

the last inequality is obtained by conditionning on X_i. *In addition,(M3) implies that* $E(|\varphi(Y)|^j|X) = \delta_j(X) = \delta_j(x) + o(1)$ *and using Lemma 4.7, we get*

$$E|\Lambda_i^{(k,l)}(x)|^q = O\left(max_{0\leqslant j\leqslant q}(\Phi_x(h))^{1+q-j}\right)$$
$$= O\left(\Phi_x(h)\right).$$

Now, we can use Rio inequality (see Proposition 5.5–(ii)) together with hypothesis (M2a) to obtain

$$J_{2,n}(x) = \sum_{S_2} |Cov(\Lambda_i^{(k,l)}(x), \Lambda_j^{(k,l)}(x))|$$
$$\leq C\left[E|\Lambda_1^{(k,l)}(x)|^q\right]^{2/q} \sum_{S_2} [\alpha(|i-j|)]^{1-\frac{2}{q}}$$
$$\leq C\left[\Phi_x(h)\right]^{\frac{2}{q}} \sum_{S_2} |i-j|^{-a(1-\frac{2}{q})}$$
$$\leq C\left[\Phi_x(h)\right]^{\frac{2}{q}} n^2 m_n^{-a(1-\frac{2}{q})}.$$

Choosing $m_n = (\Phi_x(h))^{-d}$ *, we obtain*

$$J_{1,n}(x) = O(n\Phi_x(h)) \tag{4.5}$$

and

$$J_{2,n}(x) \leq C\left(n\Phi_x(h)\right)\left[n\left(\Phi_x(h)\right)^{\frac{(q-2)(ad-1)}{q}}\right]$$
$$\leq C\left(n\Phi_x(h)\right) n^{1-\frac{u(q-2)(ad-1)}{q}},$$

the last result coming from the condition (M4). Now, in view of (M2a) we can choose q such that $u\frac{(q-2)(ad-1)}{q} > 1$. *So, we obtain*

$$J_{2,n}(x) = O(n\Phi_x(h)) \tag{4.6}$$

For the variance term, Lemma 4.7 and hypothesis (M3) permit to write

$$\begin{aligned} Var(\Lambda_1^{(k,l)}(x)) &\le C\left[\Phi_x(h) + (\Phi_x(h))^2\right] \\ &\le C\Phi_x(h). \end{aligned} \tag{4.7}$$

We readily derive the claimed result from (4.4), (4.5), (4.6) and (4.7).

Proof of Lemma 4.1 The bias terms is not affected by the dependence condition. So, the proof works exactly as that of Lemma 1.1 with replacing Y by $\varphi(Y)$. Remark that $EW_{1,2}(x) > 0$ under the assumed hypotheses (see relation (4.8)).

Proof of Lemma 4.2 Inspired by the decomposition given in 2.14, we set

$$m_1(x) = Q(x)\left[S_{2,1}(x)S_{4,0}(x) - S_{3,1}(x)S_{3,0}(x)\right],$$

where, for $p \in \{2,3,4\}$ and $l \in \{0,1\}$,

$$S_{p,l}(x) = \frac{1}{n\Phi_x(h)}\sum_{i=1}^{n}\frac{K_i(x)\beta_i^{p-2}(x)\varphi^l(Y_i)}{h^{p-2}} \text{ and } Q(x) = \frac{n^2h^2\Phi_x^2(h)}{n(n-1)EW_{12}(x)}.$$

So, we need to show taking into account the dependence assumption of the observations, if necessary, that for $p \in \{2,3,4\}$ and $l \in \{0,1\}$

$$\begin{aligned} &ES_{p,l}(x) = O(1), \\ &Q(x) = O(1), \\ &S_{p,l}(x) - ES_{p,l}(x) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right), \\ &Cov\left[S_{2,1}(x), S_{4,0}(x)\right] = O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right), \\ &Cov\left[S_{3,1}(x), S_{3,0}(x)\right] = O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right). \end{aligned}$$

• It is easy to see that under (H1), (H3), (D2) and (M1), for $p \in \{2,3,4\}$ and $l \in \{0,1\}$,we have
$ES_{p,l}(x) = O(1)$.
• Treatment of the term $Q(x)$
We have

$$EW_{12}(x) = E\left[\beta_1^2(x)K_1(x)K_2(x)\right] - E\left[\beta_1(x)\beta_2(x)K_1(x)K_2(x)\right],$$

together with

$$h^2 E\left[\beta_1(x)\beta_2(x)K_1(x)K_2(x)\right] \le Ch^2 \int_{B(x,h)}\int_{B(x,h)} \beta(u,x)\beta(t,x)dP_{(X_1,X_2)}(u,t)$$

and (D4) implies that

$$h^2 E\left[\beta_1(x)\beta_2(x)K_1(x)K_2(x)\right] = o\left(\int_{B(x,h)}\int_{B(x,h)} \beta^2(u,x)\beta^2(t,x)dP_{(X_1,X_2)}(u,t)\right).$$

By applying (H3) and (M2b), we get

$$\int_{B(x,h)}\int_{B(x,h)} \beta^2(u,x)\beta^2(t,x)dP_{(X_1,X_2)}(u,t) \le Ch^4\left[\Phi_x(h)\right]^{1+d},$$

which implies that

$$E\left[\beta_1(x)\beta_2(x)K_1(x)K_2(x)\right] = o\left(h^2\left[\Phi_x(h)\right]^{1+d}\right).$$

Now, Lemma 4.7-(iii) and the last result allow to write

$$EW_{12}(x) > Ch^2\left[\Phi_x(h)\right]^{1+d}. \tag{4.8}$$

So, for n sufficiently large

$$Q(x) = \frac{n^2h^2\Phi_x^2(h)}{n(n-1)EW_{12}(x)} \le C\frac{\left[\Phi_x(h)\right]^2}{\left[\Phi_x(h)\right]^{1+d}} \le C.$$

• Treatment of the term $S_{p,l}(x) - ES_{p,l}(x)$ We have

$$S_{p,l}(x) - ES_{p,l}(x) = \frac{1}{n\Phi_x(h)}\sum_{i=1}^{n}\Gamma_i^{(p,l)}(x),$$

where

$$\Gamma_i^{(p,l)}(x) = \Lambda_i^{(p-2,l)}(x) = \frac{1}{h^{p-2}} \left\{ K_i(x)\beta_i^{p-2}(x)\varphi^l(Y_i) - E\left[K_i(x)\beta_i^{p-2}(x)\varphi^l(Y_i)\right] \right\}, \tag{4.9}$$

with $\Lambda_i^{(k,l)}(x)$ is defined in (4.2).
Note that, because $E(\Gamma_1^{(k,l)}(x)) = 0$, $E|\Gamma_1^{(k,l)}(x)|^q = O(\Phi_x(h))$ for $q > 2$ and using Tchebychev's inequality, we can apply Proposition 5.5–(i), to get for any $q > 2$, $\varepsilon > 0$, $r \geq 1$ and for some $0 < C < \infty$

$$\begin{aligned} P\left(|S_{p,l}(x) - E\left[S_{p,l}(x)\right]| > \varepsilon\right) &= P\left(|\sum_{i=1}^{n} \Gamma_i^{(p,l)}(x)| > n\varepsilon\Phi_x(h)\right) \\ &\leq C\left[A_1(x) + A_2(x)\right], \end{aligned} \tag{4.10}$$

where

$$A_1(x) = \left(1 + \frac{\varepsilon^2 n^2 (\Phi_x(h))^2}{r S_{n,l,k}^2(x)}\right)^{-r/2} \quad and \quad A_2(x) = n r^{-1} \left(\frac{r}{\varepsilon n \Phi_x(h)}\right)^{(a+1)q/(q+a)}.$$

Now, taking for $\eta > 0$

$$\varepsilon = \eta\sqrt{\frac{\ln n}{n\Phi_x(h)}} \quad \text{and} \quad r = (\ln n)^2,$$

we obtain

$$A_2(x) \leq C n^{1-\frac{(a+1)q}{2(q+a)}} (\ln n)^{-2+\frac{3(a+1)q}{2(q+a)}} (\Phi_x(h))^{-\frac{(a+1)q}{2(q+a)}},$$

Next, using (H8), it exists some real number $\nu > 0$ such that

$$A_2(x) = O(n^{-1-\nu}). \tag{4.11}$$

Moreover, in view of equation (4.3) and the fact that $\ln(x+1) = x - x^2/2 + o(x^2/2)$ where x tends to zero, we can write

$$A_1(x) \leq C n^{-\eta^2/2}, \tag{4.12}$$

which shows that $A_1(x)$ is the general term of a convergent series for an appropriate choice of η.
Hence, by combining relations (4.10), (4.11) and (4.12), we derive

$$S_{p,l}(x) - ES_{p,l}(x) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

• Finally, by following similar arguments used to prove (4.3), we obtain

$$Cov\left[S_{2,1}(x), S_{4,0}(x)\right] = O\left(\frac{1}{n\Phi_x(h)}\right)$$

and

$$Cov\left[S_{3,1}(x), S_{3,0}(x)\right] = O\left(\frac{1}{n\Phi_x(h)}\right).$$

In view of (M4), this last rate is negligible with respect to $O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$.
The proof is then completed.
Proof of Lemma 4.3 The first part of the claimed results can be directly deduced from the proof of Lemma 4.2 by taking for all i, $\varphi(Y_i) = 1$.
We can deduce that $m_0(x)$ converges almost completely to 1 and this involves that

$$\sum_{n=1}^{\infty} P\left(m_0(x) < \frac{1}{2}\right) < \infty.$$

2. To treat the uniform convergence of $\hat{m}_\varphi(x)$, we need the following preliminary technical lemma. This is the uniform version of Lemma 4.7 and its proof works in the same manner.

Lemma 4.9. *Under assumptions (U1), (U3), (E1), (E2b) and (E4), we obtain*
i) $\forall (p,l) \in \mathbb{N}^\star \times \mathbb{N},\ \sup_{x\in S_{\mathcal{F}}} E\left(K_1^p(x)|\beta_1^l(x)|\right) \leq Ch^l\Phi(h).$
ii) $\forall (p_1, p_2, l_1, l_2) \in \mathbb{N}^\star \times \mathbb{N}^\star \times \mathbb{N} \times \mathbb{N},$
$\sup_{x\in S_{\mathcal{F}}} E\left(K_1^{p_1}(x)K_2^{p_2}(x)|\beta_1^{l_1}(x)||\beta_2^{l_2}(x)|\right) \leq Ch^{(l_1+l_2)}\left[\Phi(h)\right]^{1+d}.$
iii) $\exists n_0 \in \mathbb{N}, \forall n > n_0,\ \ \inf_{x\in S_{\mathcal{F}}} E\left[K_1(x)K_2(x)\beta_1^2(x)\right] > Ch^2\left[\Phi(h)\right]^{1+d}.$

Proof of Lemma 4.4 We have

$$Em_l(x) = \frac{1}{EW_{12}(x)} E\left[W_{12}(x)\varphi^l(Y_2)\right]$$

and $Em_1(x)$ can also be written as

$$Em_1(x) = E\left[E(m_1(x)|X_2)\right] = \frac{1}{EW_{12}(x)} E\left[W_{12}(x)E(\varphi(Y_2)|X_2)\right].$$

So, we get under assumption (E1)
$|m_\varphi(x) - Em_1(x)| = \frac{1}{|EW_{12}(x)|}\left|E\left\{W_{12}(x)\left[m_\varphi(x) - m_\varphi(X_2)\right]\right\}\right| \leq \sup_{x' \in B(x,h)} |m_\varphi(x) - m_\varphi(x')|.$
We need to take into account hypothesis (U2) to obtain

$$\sup_{x \in S_{\mathcal{F}}} |m_\varphi(x) - Em_1(x)| = O(h^b).$$

Proof of Lemma 4.5 Following the same steps as in the proof of Lemma 4.2, but using Lemma 4.9 instead of Lemma 4.7, we obtain under assumptions (U1), (U3) and (E1)–(E5), for $p \in \{2,3,4\}$ and $l \in \{0,1\}$

$$\sup_{x \in S_{\mathcal{F}}} Q(x) = O(1), \quad \sup_{x \in S_{\mathcal{F}}} ES_{p,l}(x) = O(1) \tag{4.13}$$

and

$$\sup_{x \in S_{\mathcal{F}}} Cov\left[S_{2,1}(x), S_{4,0}(x)\right] = O\left(\frac{1}{n\Phi(h)}\right), \quad \sup_{x \in S_{\mathcal{F}}} Cov\left[S_{3,1}(x), S_{3,0}(x)\right] = O\left(\frac{1}{n\Phi(h)}\right). \tag{4.14}$$

It remains to show that, for $p \in \{2,3,4\}$ and $l \in \{0,1\}$,

$$\sup_{x \in S_{\mathcal{F}}} |S_{p,l}(x) - ES_{p,l}(x)| = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right). \tag{4.15}$$

To this aim, let us set

$$j(x) = \arg\min_{j \in \{1,2,\ldots,N_{r_n}(S_{\mathcal{F}})\}} d(x, x_j),$$

and consider the following decomposition

$$\sup_{x\in S_{\mathcal{F}}} |S_{p,l}(x) - ES_{p,l}(x)| \leq \sup_{x\in S_{\mathcal{F}}} \left|S_{p,l}(x) - S_{p,l}(x_{j(x)})\right|$$
$$+ \sup_{x\in S_{\mathcal{F}}} \left|S_{p,l}(x_{j(x)}) - ES_{p,l}(x_{j(x)})\right|$$
$$+ \sup_{x\in S_{\mathcal{F}}} \left|ES_{p,l}(x_{j(x)}) - ES_{p,l}(x)\right| := F_1^{p,l} + F_2^{p,l} + F_3^{p,l}.$$

Let us now study each term $F_k^{p,l}$ for $k \in \{1,2,3\}$.

Study of the term $F_2^{p,l}$

For all $\varepsilon > 0$, we have that

$$\begin{aligned}
P\left(F_2^{p,l} > \varepsilon\right) &= P\left(\max_{j\in\{1,\dots,N_{r_n}(S_{\mathcal{F}})\}} |S_{p,l}(x_j) - ES_{p,l}(x_j| > \varepsilon\right) \\
&\leq N_{r_n}(S_{\mathcal{F}}) \max_{j\in\{1,\dots,N_{r_n}(S_{\mathcal{F}})\}} P\left(|S_{p,l}(x_j) - ES_{p,l}(x_j| > \varepsilon\right) \\
&\leq N_{r_n}(S_{\mathcal{F}}) \max_{j\in\{1,\dots,N_{r_n}(S_{\mathcal{F}})\}} P\left(\left|\sum_{i=1}^{n} \Gamma_i^{p,l}(x_j)\right| > n\Phi(h)\varepsilon\right),
\end{aligned}$$

where $\Gamma_i^{p,l}(x)$ is defined in (4.9). By applying Proposition 5.5–(i) in Ferraty and Vieu (2006) and since $E|\Gamma_1^{(k,l)}(x)|^q = O(\Phi_x(h))$ for $q > 2$, we have for any $q > 2$, $\varepsilon > 0$, $r \geq 1$ and for some $0 < C < \infty$

$$P\left(F_2^{p,l} > \varepsilon\right) \leq C(A_1 + A_2),$$

where

$$A_1 = N_{r_n}(S_{\mathcal{F}})\left(1 + \frac{\varepsilon^2 n^2 \Phi^2(h)}{r S_{n,l,p}^2}\right)^{-r/2} \quad , \quad A_2 = N_{r_n}(S_{\mathcal{F}}) n r^{-1}\left(\frac{r}{\varepsilon n \Phi(h)}\right)^{(a+1)q/(q+a)}$$

and $S_{n,l,p}^2 := \sup_{x\in S_{\mathcal{F}}} S_{n,l,p}^2(x) = O(n\Phi(h))$ in view of relation (4.3) together with hypothesis (U1).

Choosing for $\eta > 0$

$$\varepsilon = \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}} \;\; and \;\; r = \left(\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)\right)^2,$$

we obtain

$$A_1 = O(n^{-1-\nu}) \ \ and \ \ A_2 = O(n^{-1-\nu'}),$$

where $\nu, \nu' > 0$.

Hence, we get for η large enough

$$P\left(F_2^{p,l} > \eta\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right) \leq Cn^{-1-\xi},$$

where $\xi > 0$.

Study of the terms $F_1^{p,l}$ and $F_3^{p,l}$

First, let us analyse the term $F_1^{p,l}$. Since K is supported in $[0,1]$ and according to (U1), we have the relation (2.8) which give by

$$F_1^{p,l} \leq \frac{Cr_n}{nh\Phi(h)} \sup_{x\in S_{\mathcal{F}}} \sum_{i=1}^{n} |\varphi^l(Y_i)| 1_{B(x,h)\cup B(x_{j(x)},h)}(X_i).$$

Let

$$Z_i = \frac{Cr_n|\varphi^l(Y_i)|}{h} \sup_{x\in S_{\mathcal{F}}} 1_{B(x,h)\cup B(x_{j(x)},h)}(X_i).$$

In the same manner as for proving (4.3), we have under hypotheses (U1), (E2b), (E3) and (E5)

$$S_n^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} |Cov(Z_i, Z_j)| = O(n\Phi(h)).$$

It remains to use similar arguments as to treat $F_2^{p,l}$ to obtain

$$F_1^{p,l} = O_{a.co.}\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right).$$

Second, since

$$F_3^{p,l} \leq E\left(\sup_{x\in S_{\mathcal{F}}} \left|S_{p,l}(x) - S_{p,l}(x_{j(x)})\right|\right),$$

we deduce that

$$F_3^{p,l} = O\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\ln n}{n}\right)}{n\Phi(h)}}\right).$$

Applying (4.13), (4.14) and (4.15) together with the last condition of hypothesis (E5), the result of Lemma 4.5 is immediately obtained.

Proof of Lemma 4.6 The first part of the claimed results can be directly deduced from the proof of Lemma 4.5 by taking for all i, $\varphi(Y_i) = 1$ and this yields easily to the second part.

Chapter 5

Annex : Some probabilistic tools

In this Annex, we briefly present some probabilistic tools we need in this thesis.

The almost Complete Convergence

The concept of the almost complete convergence was introduced by Hsu and Robbins (1947), this convergence is in some sense easier to state than the almost sure one. Moreover, this mode of convergence implies other standard modes of convergence, such that the almost sure convergence and the convergence in probability.

Definition 5.1. *Let $(Z_n)_{n\in\mathbb{N}^*}$ be a sequence of real random variables (r.r.v.). We say that $(Z_n)_{n\in\mathbb{N}^*}$ converges almost completely to some r.r.v. Z, and we note $Z_n \overset{a.co.}{\longmapsto} Z$, if and only if*

$$\forall \varepsilon > 0, \ \sum_{n=1}^{\infty} P(|Z_n - Z| > \varepsilon) < \infty.$$

Moreover, let $(u_n)_{n\in\mathbb{N}^*}$ be a sequence of positive real numbers going to zero; we say that the rate of the almost complete convergence of $(Z_n)_{n\in\mathbb{N}^*}$ to Z is *of order* (u_n) *and we note* $Z_n - Z = O_{a.co.}(u_n)$*, if and only if*

$$\exists\, \varepsilon_0 > 0, \ \sum_{n=1}^{\infty} P(|Z_n - Z| > \varepsilon_0 u_n) < \infty.$$

In the following proposition, we recall some results extensively used in this thesis. For more details, the reader can see Bosq and Lecoutre (1987) and Ferraty and Vieu (2006).

Proposition 5.1. *Let* l_x *and* l_y *be two deterministic real numbers and let* $(u_n)_{n\in\mathbb{N}^*}$ *be a sequence of real numbers going to zero.*

i). If $lim_{n\to+\infty} X_n = l_x$*, a.co. and* $lim_{n\to+\infty} Y_n = l_y$*, a.co., we have*
a) $lim_{n\to+\infty}(X_n + Y_n) = l_x + l_y$ *a.co.,*
b) $lim_{n\to+\infty}(X_n \times Y_n) = l_x \times l_y$ *a.co.,*
c) $lim_{n\to+\infty} \frac{1}{Y_n} = \frac{1}{l_y}$ *a.co. as long* $l_y \neq 0$.

ii). If $X_n - l_x = O_{a.co.}(U_n)$ *and* $Y_n - l_y = O_{a.co.}(U_n)$*, we have*
a) $(X_n + Y_n) - (l_x + l_y) = O_{a.co.}(U_n)$*,*
b) $(X_n \times Y_n) - l_x \times l_y = O_{a.co.}(U_n)$*,*
c) $\frac{1}{Y_n} - \frac{1}{l_y} = O_{a.co.}(U_n)$ *as long* $l_y \neq 0$.

The strong mixing

The field of mixing conditions is of great interest in statistics. This comes mainly from the fact that it opens the door for application involving time series. Notice that, there are many ways of modelling the dependence of a sequence of random variables in the case of mixing. But, In this section we focus on the α-mixing (or strong mixing) notion, which is one of the most general among the different mixing structures introduced in the literature

(see for instance Roussas and Ioannides (1987) or Chapter 1 in Yoshihara (1994) for definitions of various other mixing structures and links between them). For the strong mixing in the functional context, we refer to Ferraty and Vieu (2006), especially sections 10.3 and 10.4.
All that can be done here is to give a narrow snapshot of part of the strong mixing in the functional context which applied in the theoretical advances in Chapters 3 and 4.
To start with, some notations are introduced. Let $(Z_n)_{n\in\mathbb{Z}}$ be a sequence of random variables on the probability space $(\Omega, \mathcal{A}, P)$, which takes values in the measurable space $(\Omega', \mathcal{A}')$. Denote σ_j^k, $-\infty \leqslant j \leqslant k \leqslant +\infty$, the σ-algebra, which is generated by the random variables $\{Z_j, ..., Z_k\}$.

Definition 5.2. *The strong mixing coefficient of a sequence $(Z_n)_{n\in\mathbb{Z}}$ of random variables is defined as*

$$\alpha(n) = \sup_{\{k\in\mathbb{Z}, A\in\sigma_{-\infty}^k, B\in\sigma_{n+k}^{+\infty}\}} |P(A\cap B) - P(A)P(B)|.$$

The sequence $(Z_n)_{n\in\mathbb{Z}}$ is called α-mixing (or strong mixing), if

$$\alpha(n) \to 0 \text{ as } n \to \infty.$$

Depending on the rate of convergence of $\alpha(n)$ one considers two cases.

- arithmetic (or algebraic) α-mixing.
- geometric α-mixing..

Definition 5.3. *The sequence $(Z_n)_{n\in\mathbb{Z}}$ is said to be arithmetically α-mixing with rate $a > 0$ if*

$$\alpha(n) \leq Cn^{-a}.$$

It is called geometrically α-mixing if

$$\exists C \in \mathbb{R}_+^*, \; \exists t \in]0, 1[, \; \alpha(n) \leq Ct^n.$$

To study the nonparametric kernel functional statistical methods (see our chapters 3 and 4), we need the following proposition

Proposition 5.2. *Assume that* Ω' *is a semi-metric space with semi-metric* d*, and that* A *is the* σ*-algebra spanned by the open balls for this semi-metric. Let* x *be a fixed element of* Ω'*. Then we have*
i) $(Z_n)_{n\in\mathbb{R}}$ *is* α*-mixing then* $(d(Z_n,x))_{n\in\mathbb{Z}}$ *is* α*-mixing.*
ii) In addition, if the coefficients of $(Z_n)_{n\in\mathbb{Z}}$ *are geometric (resp. arithmetic) then those of* $(d(Z_n,x))_{n\in\mathbb{Z}}$ *are also geometric (resp. arithmetic with the same order).*

Exponential Inequalities

The literature contains various versions of exponential inequalities. These inequalities differ according to the various hypotheses checked by the random variables.
This section instructs the exponential inequality taking into account two situations: the case of independent observations (Bernstein's inequality) and the case of dependent observations (Rio's inequality or the Fuk-Nagaev inequality), for more detail see Ferraty and Vieu (2006). It is the main tool for proving our asymptotic results that are examined in chapters 1, 2, 3 and 4.

Independent case

In all this subsection, let $(Z_n)_{n\in\mathbb{Z}}$ be a sequence of centered random variables.

Proposition 5.3. *(See Corollary A.8 in Ferraty and Vieu (2006))*
i). if $\forall m \geq 2$, $\exists C_m > 0$; $E|Z_1^m| \leq C_m a^{2(m-1)}$, *we have* $\forall \epsilon > 0$

$$P\left(|\sum_{i=1}^{n} Z_i| > \epsilon n\right) \leq 2exp\left(-\frac{\epsilon^2 n}{2a^2(1+\epsilon)}\right).$$

ii). Assume that the variables depend on n (that is, assume that $Z_i := Z_{i,n}$. if $\forall m \geq 2$, $\exists C_m > 0$; $E|Z_1^m| \leq C_m a_n^{2(m-1)}$ and if $u_n = n^{-1} a_n^2 \log n$ verifies $\lim_{n\to\infty} u_n = 0$, we have

$$\frac{1}{n}\sum_{i=1}^{n} Z_i = O_{a.co.}\left(\sqrt{u_n}\,\right).$$

Mixing case

There is a wide literature concerning covariance inequalities for mixing variables. For this, we us first start with some covariance inequality.
Let $(T_n)_{n\in\mathbb{Z}}$ be a stationary sequence of real random variables

Proposition 5.4. *(See Proposition A.10 in Ferraty and Vieu (2006)) Assume that $(T_n)_{n\in\mathbb{Z}}$ is α-mixing. Let us, for some $k \in \mathbb{Z}$, consider a real variable τ (resp. τ') which is $\sigma_{-\infty}^{k}$-measurable (resp. $\sigma_{n+k}^{+\infty}$-measurable).*
i). If τ and τ' are bounded, then

$$\exists C, 0 < C < +\infty, Cov(\tau, \tau') \leqslant C\alpha(n).$$

ii). If, for some positive numbers p, q and r such that $p^{-1} + q^{-1} + r^{-1} = 1$, we have $E(\tau)^p < \infty$ and $E(\tau')^q < \infty$, then

$$\exists C, 0 < C < +\infty, Cov(\tau, \tau') \leqslant C(E(\tau)^p)^{1/p}(E(\tau')^q)^{1/q}(\alpha(n))^{1/r}.$$

Secondly, we present two Rio's exponential inequalities for partial sums of a sequence $(Z_n)_{n\in\mathbb{Z}}$ of stationary and centered arithmetically mixing real random variables. Assume that $(Z_n)_{n\in\mathbb{N}^*}$ are identically distributed and are arithmetically α-mixing with rate $a > 1$ and let us introduce the notation

$$S_n^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} |cov(Z_i, Z_j)|$$

Proposition 5.5. *(See Proposition A.11 in Ferraty and Vieu (2006))*
i). If $\exists p > 2$ and $M > 0$ such that $\forall t > M$; $P(|Z_1| > t) \leq t^{-p}$, then we have

for any $r \geqslant 1$, $\epsilon > 0$ *and for some* $C < +\infty$

$$P\left(|\sum_{i=1}^{n} Z_i| > \epsilon\right) \leq C\left\{\left(1+\frac{\epsilon^2}{rS_n^2}\right)^{-r/2} + \left(\frac{r}{\epsilon}\right)^{(a+1)p/(a+p)}\right\}.$$

ii). If $\exists M < \infty$ *such that* $|Z_1| \leq M$, *then we have for any* $r \geqslant 1$, $\epsilon > 0$ *and for some* $C < +\infty$

$$P\left(|\sum_{i=1}^{n} Z_i| > \epsilon\right) \leq C\left\{\left(1+\frac{\epsilon^2}{rS_n^2}\right)^{-r/2} + \left(\frac{r}{\epsilon}\right)^{(a+1)}\right\}.$$

Kolmogorov's entropy

For the uniform consistency, where the main tool is to cover a subset $S_{\mathcal{F}}$ with a finite number of balls,one introduces a topological concept defined asfollows

Definition 5.4. *Let* S *be a subset of a semi-metric space* $\mathcal{F}$, *and let* $\varepsilon > 0$ *be given. A finite set of points* $x_1, x_2, ..., x_N$ *in* $\mathcal{F}$ *is called an* ε*-net for* S *if* $S \subset \bigcup_{k=1}^{N} B(x_k, \varepsilon)$. *The quantity* $\psi_S(\varepsilon) = \ln(N_\varepsilon(S))$, *where* $N_\varepsilon(S)$ *is the minimal number of open balls in* $\mathcal{F}$ *of radius* ε *which is necessary to cover* S, *is called Kolmogorov's* ε*-entropy of the set* S.

It is known that the entropy of a set measures its complexity. We refer to Kolmogorov and Tikhomirov (1959) and Ferraty et al. (2010) for more details and examples on this topic.

Perspectives

To conclude this thesis we raise some perspectives that may be the object of future works.

• Establish the quadratic mean convergence and the asymptotic normality of the generalized regression estimator with the local linear method, for α-mixing observations.

• Show the almost complete convergence results (pointwise and uniform) similar to those of Chapter 2 when both the response variable and the explanatory one are functional.

• Study the almost complete convergence (pointwise and uniform) similar to those of chapter 4 in the doubly functional case when the sample considered is a strong mixing sequence.

• Spatial modelization : Local linear estimation of the generalized regression for functional data.

Bibliography

Amiri, A., Crambes, C., and Thiam, B. (2014). Recursive Estimation of Nonparametric Regression with Functional Covariate. *Computational Statistics & Data Analysis*, 69, 154–172.

Attaoui, S.(2014). On the Nonparametric Conditional Density and Mode Estimates in the Single Functional Index Model with Strongly Mixing Data. *Sankhya: The Indian Journal of Statistics*, 76-A(Part 2), 356–378.

Attouch, M., Laksaci, A., and Rafaa, F.(2017). Local linear estimate of the regression operator by the kNN method. *Comptes Rendus Mathematique*, 7, 824–829.

Baillo, A., Grane, A. (2009). Local linear regression for functional predictor and scalar response. *J. Multivariate Anal.* 100, 102–111.

Barrientos-Marin, J., Ferraty, F., Vieu, P.(2010). Locally modelled regression and functional data. *J.ournal of Nonparametric Statistics*, 22, 617–632.

Benhenni, K., Ferraty, F., Rachdi, M., and Vieu, P. (2007). Local smoothing regression with functional data. *Computational Statistics*, 22(3):353–369.

Benhenni, K., Hedli-Griche, S., Rachdi, M., and Vieu, P. (2008). Consistency of The Regression Estimator with Functional Data Under Long Memory Conditions. *Statistics and Probability Letters*, 78, 1043–1049.

Berlinet, A., Elamine, A., and Mas, A. (2011). Local Linear Regression for Functional Data. *Annals of the Institute of Statistical Mathematics*, 63, 1047–1075.

Boj, E., Delicado, P., and Fortiana, J. (2010). Distance-Based Local Linear Regression for Functional Predictors. *Computational Statistics & Data Analysis*, 54, 429–437.

Bradley , R.C. (2007). *Introduction to strong mixing conditions*, Vol I-III. Kendrick Press, Utah.

Bosq, D., and Lecoutre, J.P. *Theorie de l'estimation fonctionnelle (in french)*. Economica.

Burba, F., Ferraty, F., and Vieu, P. (2009). k-Nearest Neighbour Method in Functional Nonparametric Regression. *Journal of Nonparametric Statistics*, 21, 453–469.

Cardot, H., Ferraty, F., and Sarda, P. (2003). Spline estimators for the functional linear model. *Statist. Sin.* 13(3), 571–591.

Carl, B., and Stephani, I. *Entropy, compactness and the approximation of operators*. Cambridge University Press, 1990.

Chate, H.,Courbage,M.,(1997). Lattice systems. *Physica D103*,1–612.

Chu, C.-K., Marron, J.-S.,(1991). Choosing a kernel regression estimator. *Statist. Sci*, 6(4), 404–436.

Dabo, N., and Rhomari, N.(2003). Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *C. R., Math., Acad. Sci. Paris*, 336, 75–80.

Delsol (2007). Regression non-parametrique fonctionnelle: expressions asymptotiques des moments. *Ann. I.S.U.P. LI*, (3)(2007) 43–67.

Demongeot, J., Laksaci, A., Madani, F., Rachdi, M.(2010). Estimation locale lineaire de la densite conditionnelle pour des donnees fonctionnelles. *C. R. Acad. Sci. Paris, Ser.* I 348 (15–16), 931–934..

Demongeot, J., Laksaci, A., Madani, F., Rachdi, M.(2011). A fast functional locally modeled conditional density and mode for functional time-series. *In: Recent Advances in Functional Data Analysis and Related Topics. In: Contributions to Statistics, Physica-Verlag, Springer*, pp. 85–90.

Demongeot, J., Laksaci, A., Madani, F., Rachdi, M.(2013). Functional data analysis: conditional density estimation and its application. *Statistics*, 47, No. 1, 26–44.

Demongeot, J., Laksaci, A., Rachdi, M., and Rahmani, S. (2014). On The Local Linear Modelization of The Conditional Distribution for Functional Data. *Sankhya*, A, 76, 328–355.

Demongeot, J., Laksaci, A., Naceri, A., and Rachdi, M. (2017). Local linear regression modelization when all variables are curves. *Statistics ans probability letters*, 121, 37–44.

El Methni, M. and Rachdi, M., (2011). Local weighted average estimation of the regression operator for functional data. *Journal of Nonparametric Statistics*, 40 (17), 3141–3153.

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

Fan, J. (1992). Design-Adaptive Nonparametric Regression. *Comm. Statist. Theory Methods*, 87, 998–1004.

Fan, J. and Gijbels, I., (1996). *Local Polynomial Modelling and its Applications. In: Monographs on Statistics and Applied Probability.* vol. 66. Chapman & Hall.

Ferraty, F., Goia, A., and Vieu, P. (2002) Functional nonparametric model for time series: a fractal approach to dimension reduction. *Test*, 11, 317–344.

Ferraty, F., and Vieu, P. (2007). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. *C. R. Acad. Sci., Paris* 330(2000) 403–406.

Ferraty, F., Mas, A., and Vieu, P. (2007). Nonparametric Regression on Functional Data: Inference and Practical Aspects. *Australian & New Zealand Journal of Statistics*, 49, 267–286.

Ferraty, F., Laksaci, A., Tadj, A., and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *Journal of Statistical planning and inference*, 140:335–352.

Ferraty, F., Laksaci, A., Tadj, A., and Vieu, P. (2011). Kernel regression with functional response. *Electron. J. Stat*, 5 (2011) 159–171.

F. Ferraty, I. Van Keilegom and P. Vieu (2012). Regression when both response and predictor are functions. *Electron. J. Stat*, 5 (2011) 159–171. *Multivariate Anal*, 109 (2012) 10–28.

Ferraty, F., Laksaci, A., Tadj, A., and Vieu, P. (2010). Nonparametric regression for functional response and functional regressor under dependance *C. R. Acad. Sci. Paris, Ser. I*, 350 (2012) 717–720.

Ferraty, F., Van Keilegom, I., and Vieu, P. (2008). On the validity of the bootstrap in nonparametric functional regression. *Scandinavian Journal of Statistics*, 37(2):286–306.

Ferraty, F. and Vieu, P. (2004). Non-parametric models for functional data, with application in regression, time-series prediction and curve discrimination. *Journal of Nonparametric Statistics*, 16(1-2):111–125.

Ferraty, F. and Vieu, P. (2004). Erratum of : "Non-parametric models for functional data, with application in regression, time-series prediction and curve discrimination". *Journal of Nonparametric Statistics*, 20(2):187–189.

Ferraty, F., Vieu, P.(2006). *Nonparametric functional data analysis. Theory and Practice.* Springer Series in Statistics, New York.

Hsu, P. L. and Robbins,H. (1947). *Complete convergence and the law of large numbers.* Proceedings of the national academy of sciences, 33(2).

Kolmogorov, A. N. and Tikhomirov, V. M. (1959). *ε-entropy and "-capacity.* UspekhiMatematicheskikh Nauk, 14:3–86.

Kudraszow, N.L. Vieu P. (2013). Uniform consistency of kNN regressors for functional variables. *Statistics and Probability Letters.* 83:1863–1870.

Kuelbs, J., and Li, W. V. (1993) Metric Entropy and the Small Ball Problem for Gaussian Measures. *Journal of Functional Analysis* 116, 1, 133–157.

Laksaci, A. Lemdani, M. Ould Said, E.(2009). A generalized L^1-approach for a kernel estimator of conditional quantile with functional regressors: Consistency and asymptotic normality. *Statistics and Probability Letters*, 79, 1065–1073.

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

Laksaci, A. Lemdani, M. Ould Said, E.(2011). Asymptotic Results for an L^1-norm Kernel Estimator of the Conditional Quantile for Functional Dependent Data with Application to Climatology. *Sankhya : The Indian Journal of Statistics*, 73-A (Part 1), 125–141.

Li, W., V., and Linde, W. (1999). Approximation, Metric Entropy and Small Ball Estimates for Gaussian Measures. *The Annals of Probability*, 27(3), 1556–1578.

Lin, Z. Lu, C. (1996). *Limit theory of mixing dependent random variables.* Mathematics and its applications, Sciences Press, Kluwer Academic Publishers, Beijing.

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

Masry, E. (2005). Nonparametric Regression Estimation for Dependent Functional Data: Asymptotic Normality. *Stochastic Processes and their Applications*, 115, 155–177.

Messaci, F. Nemouchi, N. Ouassou, I. Rachdi, M.(2015). Local polynomial modelling of the conditional quantile for functional data. *Statistics Methods and Applications*, 24, No. 4, 597–622.

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. *Computational Statistics & Data Analysis,*, 73, 53–68.

Rachdi, M., and Vieu, P. (2007). Nonparametric Regression for Functional Data: Automatic Smoothing Parameter Selection. *Journal of Statistical Planning and Inference*, 137(9), 2784–2801.

Ramsay, J.O., and Silverman, B.W. (2002). *Applied functional data analysis. Methods and case studies*. Springer Series in Statistics, New York.

Ramsay, J., and Silverman, B. (1997). *Functional Data Analysis*. Springer, NewYork.

Roussas, G. and Ioannides, D. Moment inequalities for mixing sequences of random variables. *Stoch. proc. and Appl.*, 5(1), 61–120 (1987)..

Steinwart, I., and Christmann, A. (2008). *Support Vector Machines*. Springer.

van der Vaart, A., and van Zanten, H. (2007) Bayesian inference with rescaled Gaussian process priors. *Electronic Journal of Statistics 1*, 433–448.

Yoshihara, K. (1994) *Weakly dependent stochastic sequences and their applications. IV: Curve estimation based on weakly dependent data*. Sanseido.

Zhiyong Z. and Zhengyan L. (2016). Asymptotic normality of locally modelled regression estimator for functional data. *Journal of Nonparametric Statistics*, Vol. 28, No. 1, 116–131.

Abstract

In this thesis, we consider the problem of the nonparametric estimation of the generalized regression function when the response variable is real and the regressor is valued in a functional space (space of infinite dimension), by using the local linear method. These study includes, among others, those of the regression function and the conditional cumulative distribution function. Firstly, we consider a sequence of independent and identically distributed observations and we generalize the work of Barrientos et al. (2010). Indeed, we are concerned with a class of estimators, including their estimator, for which both the pointwise and the uniform almost-complete convergence, with rates, are established, while only the pointwise one is proved in the fore-mentioned paper. Then, a real data set study illustrates the performance of our methodology with respect to other known estimators.
Secondly, we suppose that the observations are strongly mixing (α mixing) and we establish the pointwise and the uniform almost complete convergence, with rates, for the previously introduced estimators. These results can be used to solve the prediction problem for functional time series. This is illustrated through two real datasets wich, in addition, permit us to compare the local linear method, adopted in this thesis, with respect to the kernel one.

Keywords: Functional data; Nonparametric estimation; Generalized regression function; Local linear method; Uniform almost complete convergence; Rate of convergence; Entropy; α mixing.

Résumé

La problématique abordée dans cette thèse est l'estimation non paramétrique de la fonction de régression généralisée d'une variable réponse réelle conditionnée par une variable explicative fonctionnelle (à valeurs dans un espace de dimension infinie), par utilisation de la méthode locale linéaire. Cette étude inclut, entre autres, celles des fonctions de régression et de répartition conditionnelle.
Dans un premier temps, nous considérons une suite d'observations indépendantes et identiquement distribuées et nous généralisons l'étude proposée par Barrientos et al. (2010). En effet, nous étudions une classe d'estimateurs, incluant le leur, pour lesquels nous établissons la convergence uniforme presque complète, avec taux, alors que seule la convergence ponctuelle est établie dans le travail cité. Ensuite, une étude sur des données réelles illustre la performance de notre méthodologie par rapport à d'autres estimateurs connus.
Puis, dans un second temps, nous traitons le cas où les observations sont fortement mélangeantes et nous étudions aussi bien la convergence presque complète ponctuelle qu'uniforme, avec taux, des estimateurs précédemment introduits. Ces résultats peuvent être utilisés pour le problème de la prévision de séries chronologiques. Fait que nous illustrons sur deux exemples de données réelles qui ont aussi permis de comparer la méthode locale linéaire, adoptée dans cette thèse, à celle plus ancienne du noyau.

Mots-clés: Données fonctionnelles ; Estimation non paramétrique ; Fonction de régression généralisée ; Méthode locale linéaire ; Convergence presque compète uniforme, Taux de convergence ; Entropie ; α mélange.

ملخص

المشكلة التي تم تناولها في هذه الأطروحة هي التقدير الغير الوسيطي لبعض وظائف التوزيع الشرطي لمتغير الاستجابة الحقيقي المشروط بمتغير تفسري وظيفي (في فضاء ذو بعد غير منتهي)، وذلك باستخدام الطريقة الخطية محليا. وتشمل هذه الوظائف وظيفة الانحدار، والتوزيع التراكمي المشروط، والكثافة المشروطة.

أولا، نعتبر سلسلة من الملاحظات المستقلة وموزعة بشكل موحد ثم نقوم بتعميم تقدير وظيفة الانحدار التي اقترحها باريونتوس وآخرون2010 . حقيقة: ندرس مجموعة من التقديرات التي تشمله و لهذا ندرس التقارب الشبه الكامل المنتظم لهذا التقدير من خلال تحديد سرعة التقارب، اما بالنسبة للتقارب الشبه الكامل النقطي فهو مدروس في المرجع المعطى سابقا. ثم ننهي المرحلة الاولى، بدراسة البيانات الحقيقية التي توضح فعالية منهاجيتنا مقارنة بالتقديرات المعروفة الأخرى. في الخطوة الثانية، نتعامل مع الحالة حيث الملاحظات هي مرتبطة بقوة ثم ندرس التقارب الشبه الكامل النقطي و المنتظم، بإعطاء سرعة التقارب .ويمكن استخدام هذه النتيجة لمشكلة التنبؤ بالسلاسل الزمنية و اخيرا نستخدم امثلة للبيانات الحقيقية لتوضيح فعالية التقدير بالمقارنة مع طريقة النواة.

الكلمات المفتاحية :البيانات الوظيفية؛ التقدير غير الوسيطي؛ دالة الانحدار المعمم؛ الطريقة الخطية المحلية؛ التقارب الشبه الكامل المنتظم؛ سرعة التقارب؛ الانتروبي؛ الخلط القوي.