N° d'ordre :

Université de Saida– Dr. Moulay Tahar

Faculté des Sciences

Thèse

Présentée pour obtenir le diplôme de

Doctorat 3ème Cycle

Spécialité: Probabilités et Statistique

Filière : Mathématiques

Par :

AYAD SOMIA

Thème :

Sur l'estimation locale linéaire pour des données ergodiques fonctionnelles

Thèse soutenue le 31/03/2022 devant le jury composé de :

N°	Nom et prénom	Grade	Etablissement	Qualité
01	**GUENDOUZI Toufik**	Prof.	Université de Saida – Dr. Moulay Tahar	Président
02	**RAHMANI Saâdia**	Prof.	Université de Saida – Dr. Moulay Tahar	Rapporteur
03	**BOUCHENTOUF Amina Angelika**	Prof.	Université de Sidi Bel-Abbès – Djilali liabes	Examinatrice
04	**MECHAB Boubaker**	Prof.	Université de Sidi Bel-Abbès – Djilali liabes	Examinateur
05	**MADANI Fethi**	Prof.	Université de Saida – Dr. Moulay Tahar	Examinateur
06	**ROUANE Rachida**	MCB	Université de Saida – Dr. Moulay Tahar	Invitée

Dedication

This Thesis is dedicated to:

My dear parents, **Zouaoui** and **Khadidja** for the sacrifice they made, their love, their tenderness and their prayers during all the years of research. Special thanks go also to every single person who helped me and supported me in different ways to accomplish my goals.

My sisters **Fatima**, **Nacera**, **Rabia** and my best sister **Hadil** for their support**.**

My brothers **Islem** and **Yasser.**

My aunts and my cousins including **Zineb** who helped me since the beginning of my studies.

All my friends, my colleagues and all those who love me especially my best friend **Fatima** for her encouragement and support.

Looking for glory and pride in Islam and nothing else.

Acknowledgments

In the Name of Allah, the Most Merciful, the Most Compassionate, all praise is due to Allah and may prayers and peace be upon Mohammed, His servant and messenger.

First and foremost, I would like to thank ALLAH for helping me and giving me the strength, patience and courage throughout my research work to complete the research successfully.

Then, I would like to express my sincere gratitude to my supervisor, Professor *Saâdia RAHMANI*, whose expertise was invaluable in formulating the research questions and the methodology used in this work. Besides, I would like to express my deep appreciation and sincere gratitude to her for her multiple pieces of advice and all the time she devoted to supervise me despite her diverse responsibilities. She succeeded in transmitting the pleasure of research to me; this helped me a lot conduct this research work and to present it as clearly as possible. Moreover, her dynamism, vision, sincerity and motivation have deeply inspired me. Her knowledge, remarks and insightful comments have enormously shaped my thought and led to the blossoming of this thesis. It was a great honor and pleasure to finish this work under her supervision.

My thanks also go to my research co-supervisor, Doctor *Rachida ROUANE* for her continued encouragement, her availability, her kindness and her enormous help during this journey towards knowledge. I am extremely grateful for all what she has offered me. Once again thank you!

It was a great honor and privilege for me to be one of Professor *LAKSACI*'s students. I would like to thank him for his valuable guidance throughout my studies. He provided me with the necessary tools that I needed to choose the right direction and to successfully complete my dissertation.

I would like to thank Professor *Toufik GUENDOUZI* for agreeing to chair my doctoral thesis defence. Besides, I am grateful to Professor *Amina Angelika BOUCHENTOUF* who showed me how great I can be as a mathematician when I taught with her. I wlould like also to express my deepest appreciation to the other jury members, including my dear teacher Professor *Boubaker MECHAB*, the exeptional teacher and source of knowledge. In addition, I am deeply indebted to Professor *Fethi MADANI* for agreeing to evaluate this humble work despite the fact that they are always occupied with research. I would like to thank them all for their thorough temarks and constructive criticism that will help to improve this research work.

In addition, I would like to express my deepest thanks to Professor *Abdeldjebbar KANDOUCI* the director of the LMSSA's laboratory. I thank him for his warm welcome every time I ask for help. I will never forget his encouragement and kindness.

Furthermore, I am incredibly thankful to all the members of the Laboratory of stochastic models, Statistics and Applications, the University of Moulay Tahar saida. I would like also to thank Doctor *Ali RIGHI*, Doctor *Abdelkader MOULAY* and all the members of the laboratory of Mathematics at the University of Sidi Bel Abbes.

Finally, I extend special thanks to all those who have participated both directly and indirectly in the success of this thesis defense.

ملخص

في هذه الأطروحة، نأخذ في الاعتبار مشكل التقدير المحلي الخطي لوظيفة التوزيع الشرطي ومشتقاتها عندما يكون الانحدار ذو قيمة في فضاء ذو بعد غير محدود، الاستجابة عددية (تمت ملاحظتها بالكامل أو خضعت للرقابة) ويتم ملاحظة البيانات على أنها سلسلة أوقات وظيفية أرجوديك.

أولا، نبني تحت هيكل التبعية مقدرا محليا خطيا لوظيفة التوزيع الشرطي، وندرس في ظل افتراضات عامة معينة خصائصه المقاربة، مثل التقارب النقطي شبه الكامل (مع السرعة) والتقارب الطبيعي. يتم التحقق من ملاءمة المقدر المقترح من خلال دراسة المحاكاة.

ثانيا، في ظل نفس الشروط، نبني مقدرا محليا خطيا للكثافة الشرطية. ثم ندرس التقارب شبه الكامل، مع السرعة، لهذا المقدر، ونستنتج من ذلك خصائص مقاربة مشابهة للمقدر المحلي الخطي للمنوال الشرطي. يتم توضيح فائدة نتائجنا على بيانات حقيقية.

أخيرًا، نقوم بتعميم النتائج التي تم الحصول عليها مسبقًا في سياق خاضع للرقابة. نبني مرة أخرى مقدرًا للكثافة الشرطية بالطريقة المحلية الخطية وندرس سرعة التقارب شبه الكامل للمقدر المبني.

كلمات مفتاحية: البيانات الوظيفية، بيانات أرجوديك، البيانات الخاضعة للرقابة، المقدر المحلي الخطي، دالة التوزيع الشرطي، الكثافة الشرطية، المنوال الشرطي، التقدير اللامعلمي، التقارب شبه الكامل، التقارب النقطي، التقارب الطبيعي.

Résumé

Dans cette thèse, nous considérons le problème de l'estimation locale linéaire de la fonction de répartition conditionnelle et de ses dérivées lorsque le régresseur est évalué dans un espace de dimension infinie, la réponse est un scalaire (complétement observée ou censurée) et les données sont observées comme séries temporelles fonctionnelles ergodiques.

Tout d'abord, nous construisons sous cette structure de dépendance un estimateur local linéaire de la fonction de répartition conditionnelle, et nous établissons sous certaines hypothèses générales ses propritées asymptotiques, telles que la convergence uniforme presque complète (avec taux) et la normalité asymptotique. La pertinence de l'estimateur proposé est vérifié par une étude de simulation.

Deuxièmement, et sous les mêmes conditions, nous construisons un estimateur local linéaire de la densité conditionnelle. Ensuite, on établit la convergence presque complète, avec des taux, de cet estimateur, et on en déduit des propriétées asymptotiques similaires pour un estimateur linéaire local du mode conditionnel. L'utilité de nos résultats est illustrée sur des données réelles.

Enfin, nous généralisons les résultats précédemment obtenus dans un contexte de censure. On construit de nouveau un estimateur de la densité conditionnelle par la méthode locale linéaire et on établit la vitesse de convergence presque complète de l'estimateur construit.

Mots clés: Données fonctionnelles, données ergodiques, données censurées, estimation locale linéaire, fonction de répartition conditionnelle, densité conditionnelle, mode conditionnel, estimation non-paramétrique, convergence presque complète, convergence uniforme, normalité asymptotique .

Abstract

In this thesis, we consider the problem of the local linear estimation of the conditional cumulative distribution function and its derivatives when the regressor is valued in an infinite dimensional space, the response is a scalar (completely observed or censored) and the data are observed as ergodic functional times series.

Firstly, we build under this dependence structure a local linear estimator of the conditional distribution function, and we establish under a general assumptions its asymptotic properties, such as the uniform almost complete convergence (with rate) and the asymptotic distribution. The relevance of the proposed estimator is verified through a simulation study.

Secondly, under the same conditions, we construct a local linear estimator of the conditional density function. Afterward, we establish the almost-complete convergence, with rates, of this estimator, and we deduce similar asymptotic properties of the local linear estimator of the conditional mode. The usefulness of our results is illustrated on some real data.

Finally, we generalize the results previously obtained in a censored context. We build again an estimator of the conditional density by the local linear method and we establish the strong consistency rate of the constructed estimator.

Key words: Functional data, ergodic data, censored data, local linear estimator, conditional distribution function, conditional density, conditional mode, nonparametric estimation, almost complete convergence, uniform convergence, asymptotic normality.

LIST OF WORKS

Publications in international peer-reviewed journals:

1. Ayad, S., Laksaci, A., Rahmani, S. and Rouane, R. On the local linear modelization of the conditional density for functional and ergodic data. *METRON.*, **78**, 237–254 (2020).

2. Ayad, S., Laksaci, A., Rahmani, S. and Rouane, R. Local linear modelling of the conditional distribution function for functional ergodic data. *Mathematical Modelling and Analysis*. Accepted. (2021).

3. Ayad, S., Laksaci, A., Rahmani, S. and Rouane, R. On the local linear estimation of the conditional density function for functional ergodic data under random censorship model. Submitted for publication.

International communications:

1) Participation in the MFOA'2019 international conference of Financial Mathematical, Tools and Applications " On the quadratic error in the local linear estimation of the conditional cumulative distribution and its first derivative for functional data " at the University of Bejaia (Algeria) from 28 to 29 October 2019.

2) Participation in the first International Conference on Pure and Applied Mathematics IC-PAM'21 "On The Local Linear Modelization Of The Conditional Mode For Functional And Ergodic Data". May 26-27, 2021, Ouargla, Algeria.

National communications:

1) Participation in the first national conference of pure and applied mathematics "CIMPA 2021" "On The Local Linear Modelization Of The Conditional Density For Functional

And Ergodic Data". 15 April 2021, Tebessa, Algeria.

2) Participation in the Doctoral days JDFSE 2019 "On the local linear modelization of the conditional distribution and its first derivative for functional data". 14–15 December 2019, Djilali Liabes University, Sidi Bel Abbès, Algeria.

Conferences proceedings:

1) Participation in the CIMPA research school "Stochastic Analysis and Applications" in Saïda, Algeria, 01–09 March 2019.

2) Participation in the eleventh analysis meeting Mathematics and applications (RAMA11) at Djilali Liabes University, Sidi Bel Abbès, 21–24 November 2019.

3) Participation in the Applied Statistics Day at Dr. Moulay Tahar University, Saida, 24 February 2020.

CONTENTS

LIST OF FIGURES

CHAPTER 1

GENERAL INTRODUCTION

1.1 Parametric versus nonparametric models

A basic problem in statistics is to develop models based on a sample of observations so that further analysis can be carried out with statistical techniques using the model so developed. During the past two decades, parametric modeling has been a subject of investigation by various researchers. A disadvantage of parametric modeling is that it may not be adequate in the sense that slight infection of the data by observations not following the particular parametric family might lead to incorrect conclusions. Further, the data might be of such a type that there is no suitable parametric family that gives a good fit. Specifically, time series have been considered mainly from the parametric viewpoint. This viewpoint has the advantage that if the observed time series is sufficiently described by a parametric model; then a relatively small set of structural parameters serves as tools for interpretation and inference. A disadvantage of this point of view is a statistical situation where the observed data do not follow a specific parametric model. A consequence of imposing such a non-appropriate parametric model results in a bias that dominates the statistical error asymptotically. Under these situations, one might take recourse to nonparametric modeling. It is the strength of the nonparametric approach to consider a richer collection of models and functions with general shape rather than a relatively small set of parameterized curves. Indeed, the nonparametric fits exhibit several benefits compared to parametric fits and have been widely used as exploratory techniques. It is well known that methods of nonparametric estimation allow one to analyze and, present data at hand without any prior opinion about the data. These approaches do not make any assumption about the law or its parameters. Our knowledge about the model is not precise, which is often the case in practice. In this situation, it is natural to want to estimate one of the functions describing the model, usually the distribution function or its derivatives such that the density, the hazard functions, the mode and the quantile; this is the objective of the functional estimation.

The literature on the area of nonparametric functional estimation goes back to the 19th century, precisely in the middle of the 50s, and it has known a very important development, in particular on nonparametric density estimation with the famous paper of Rosenblatt [86] who discussed both the naive estimator and the more general kernel one. Since the appearance of this paper several methods have been developed for the nonparametric estimation of the density function (Tukey and Parzen [80]), the distribution functions as well as the regression functions which, first, considered by Nadaraya [79] then by Watson [91]. While, the failure rate is due to Rice and Rosenblatt [88]. Recently, Bosq [19] gave some indications about implementation of nonparametric methods and comparison with parametric ones including numerical results. The purpose of the book of Györfi et *al.* [63] was on studying nonparametric smoothing techniques for time series and to provide mathematical tools for nonparametric estimation under general dependence assumptions.

1.2 Functional framework

1.2.1 Nonparametric Curve Estimation

Historically, the first developments on continuous/functional variables go back to Deville [40], Besse and Ramsay [14] and Besse [15] was interested in the approximation of factor analysis in the functional case; in particular, principal components analysis of curves. Later, in 1997, Ramsay and Silverman [84] treated factor analysis for the regression models. The effectiveness and adaptation of functional regression compared to the vectorial approach was shown by Besse and Cardot [17].

Functional data analysis take an important place in statistical research. It has experienced very important development in recent years, this branch of statistics is related to the study of observations that are not real or vectorial, it is closely associated to the study of data sets appearing in continuous form (curves, images,. . .) which can be considered as discretized functions (functions observed on a discretization scale quite fine). This comes back to the technological progress, particularly concerning computer tools and their storage capacities, which are helping to record increasingly large amounts of data. The need to consider this type of data, now commonly encountered under the name of functional data in the literature, is primarily all a practical need. Indeed, this kind of data naturally arises in nearly every branch of science, ranging from engineering to geology, biology, medicine, and chemistry.

The assumption on which Functional Data Analysis is based is that the data to be processed has a more or less apparent underlying structure, and that the identification and explicit consideration of this structure can be used in order to extend effectively traditional data analysis techniques. More precisely, Functional Data Analysis applies to data whose structure is correctly represented by one or more functions. This modeling is particularly fruitful in the case where the data present for example a temporal variability. As Ramsay and Silverman underline in their book [85], which constitutes an excellent introduction to the field, this area of research has found a real echo with the community of statisticians, and has therefore been the subject of numerous works, both theoretical as practical. Indeed, the authors gave a large list of examples which show the wide application potential of the different methods linked to Functional Data analysis.

More generally, it should be noted that the monograph of Ramsay and Silverman [87] offered an excellent and accessible introduction to many topics on functional data analysis. Later on, the pioneer book of Ferraty and Vieu [54] exposed the characteristics and difficulties of functional data from a new mathematical point of view. The authors propose a non-parametric approach to Functional Data Analysis problems. They focused on various statistical topics such as predicting from a functional variable, classifying a sample of functional data and estimation based on an independent and an $\alpha-$mixing statistical sample. Furthermore, the authors of this monograph established the almost complete convergence of the proposed estimators. They also precised the convergence rates of each estimator which is linked both with the nonparametric model and the semi-metric.

The pioneer book of Ferraty and Romain [57] contains contributions by leading researchers in the field which summarize a number of recent developments in FDA and point towards future research directions. As a more recent work, we can cite the book of Horváth and Kokoszka [64] which combined between the theoretical and the practical point of view by presenting a general introduction to the mathematical FDA framework and then branches into several directions with the most novel exposition pertaining to functional data which exhibit dependence over time or space. In the same year, Mas [75] derived the minimax rate of convergence for nonparametric estimation of the regression function with independent and identically distributed covariates.

There is a consistent literature both around nonparametric prediction and functional data. Indeed, in 2000, new developments have been carried out by Ferraty and Vieu [48] in order to propose nonparametric statistical methods for dealing with such functional data. Two years latter, Ferraty et *al.* [49], proposed an approximation of the functional regression problem using the fractal dimension.

In [51], the same authors constructed a kernel estimator for the regression operator and obtained convergence rates for their estimator, thus, a solution to the problem of curse of dimensionality, this phenomenon well known in nonparametric statistics concerns the considerable degradation of the quality of the estimation when the dimension increases, so it makes the convergence rates very low. The probability measure of small balls or the concentration property is the solution of this problem which intervenes in the rates of convergence. Considering the concentration on small balls of the functional explanatory variable, Dabo Niang and Rhomani [31] obtained a norm convergence L^p of a kernel estimator of the nonparametric regression. Masry [76] provided a rigorous treatment of nonparametric regression with $\alpha-$mixing data in which the explanatory functional variable lies in a general semi-metric space, establishing the asymptotic normality of the constructed estimator. This property has been studied for the robust regression context by Attouch et *al.* [3], where some numerical studies in chemiometrical real data are carried out to compare the sensitivity to outliers between the classical and robust regression. Different approaches have been used for the the study of functional data, including the nonparametric methods proposed by Müller [78].

1.2.2 Important Fields of application for functional data

Since several decades, many statisticians developed applications allowing the treatment of functional random variables. On the one hand, this treatment allows the use or development of high-performance theoretical tools, and on the other hand, it offers enormous potential in terms of applications (in imaging, agro-industry, geology, econometrics,...). In this thesis, we have chosen example studies to cover a wide range of fields of application of this important field, and our aims is to demonstrate how large is the potential scope of functional data analysis.

In food industry: In order to respond to a quality control problem in the agri-food industry, Ferraty and Vieu [49, 50] studied an example that focuses on estimating the fat content of meat samples based on near-infrared (NIR) absorbance spectra. These data were obtained from http://lib.stat.cmu.edu/datasets/tecator. Each sample contains finely chopped pure meat with different percentages of the fat, protein, and moisture contents. For each unit i (among 215 pieces of finely chopped meat), they observed one spectrometric curve, which corresponds to the absorbance measured at a grid of 100 wavelengths distributed between 850 and 1050 nanometers. Then, they observe for each piece of meat i, the functional variable $X_i(t), t \in [850, 1050]$ which is the spectrometric curve of the piece of meat i. Figure 1.1 displays of the original spectrometric curves.

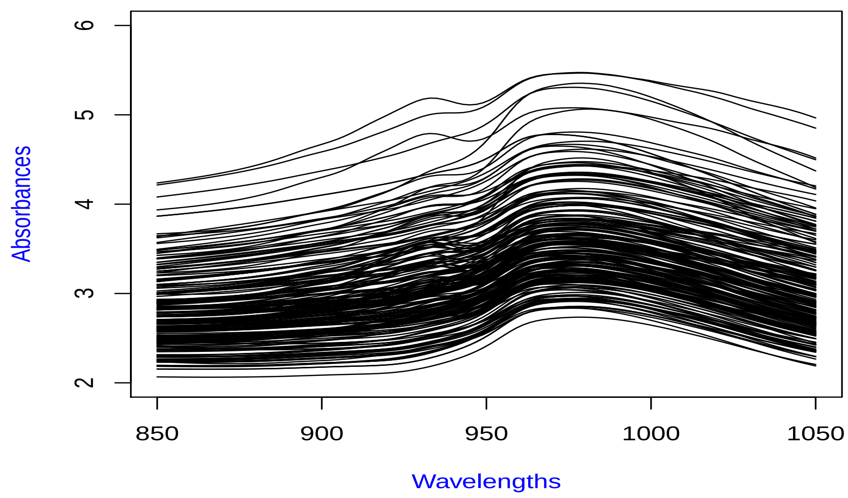

Figure 1.1: spectrometric curves

Study of the phenomenon of El Nino : An El Nino happens when a huge plate of warm water accumulates in the central Pacific and moves east, slackening or reversing the northeast trade winds and bringing warm, humid air to the west coast of South America.
The dataset is composed of sea temperature curves. We have monthly measures covering 54 years. In order to study this time series, one cuts data as (Figure 1.2) and obtains 54 curves. Each curve corresponds to the temperature evolution during one year. These data and their description are available on the website of the U.S. Climate Prediction Center: http://www.cpc.ncep.noaa.gov/data/indices/.

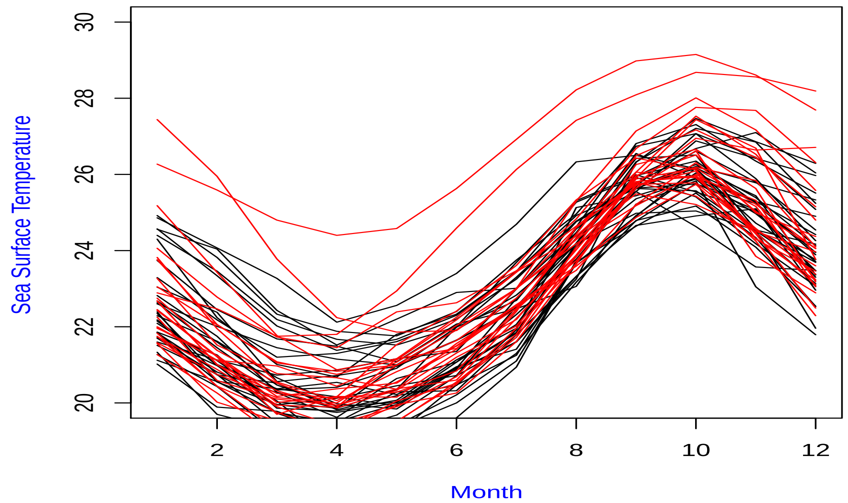

Figure 1.2: The curves corresponding to the current of the El Niño

Economic time series: In the context of dependent data, the second example concerns an economic time series of annual electricity consumption. More precisely, we dispose of both the USA monthly electricity consumed by residential and commercial sectors from January 1973 to February 2001 (338 months). Each of the 28 curves (Figure 1.3) is then made up of 12 points representing the readings for each year. These dependent real data and their descriptions are available on the websites `http://www.economagic.com` and `http://www.eia.doe.gov/emeu/aer, respectively`.

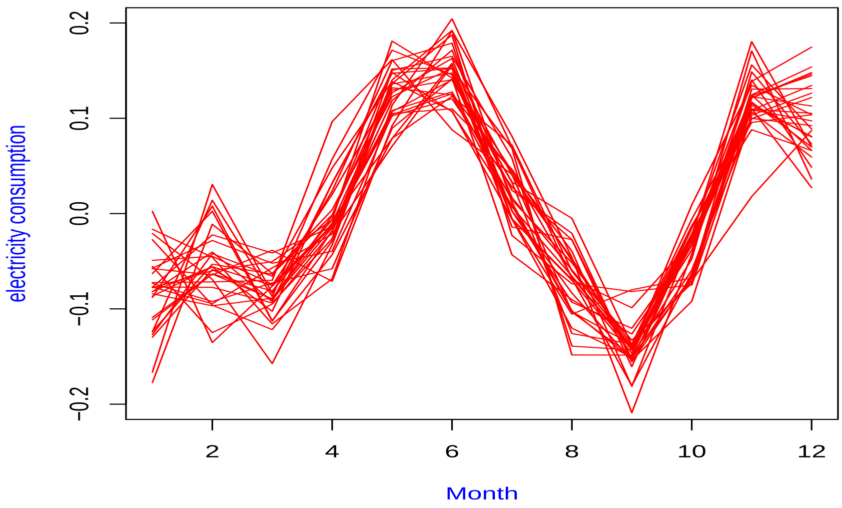

Figure 1.3: Annual electricity consumption curves in the USA

Biological growth curves: Functional data appears fundamentally in biological systems that measure some aspect of growth. Consider, for example, the height evolution of subjects in the famous Berkeley growth data. Figure 1.4 shows the growth patterns of 45 girls between 1 and 18 years old. This example uses a smoothed version of the time-derivative of the height functions, instead of the height functions themselves, as functional data.

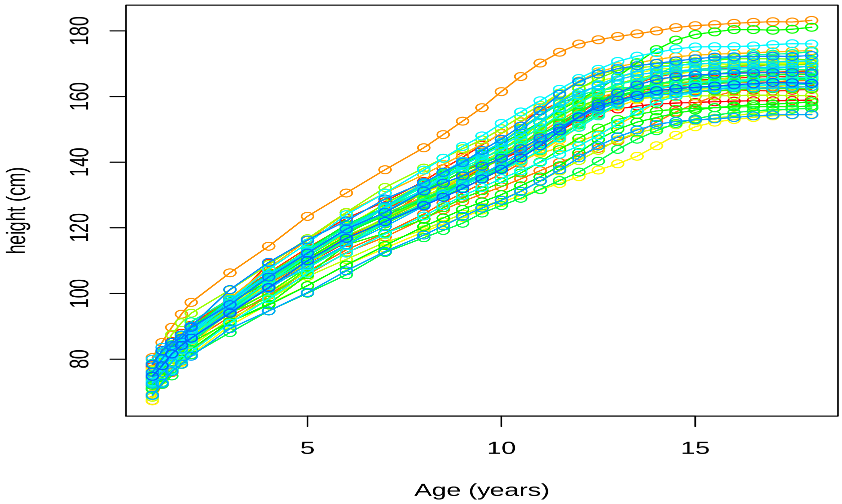

Figure 1.4: Growth curves for 45 girls

1.2.3 Semi metric and concentration measurement

It is known that the convergence rates obtained are optimal in finite dimension for some nonparametric conditional models but as soon as the dimension increases, it become relatively bad. This problem is known in the literature as the curse of dimensionality; introduced, for the first time, by Geenens [60]. This phenomena reflects the evident difficulties in nonparametric estimation of infinite-dimensional objects due to extreme data sparsity, resulting in a decrease in fastest achievable rates of convergence of regression function estimators toward their target curve as the dimension of the regressor vector increases. Consequently, it is not surprising to obtain dramatically bad theoretical properties for the nonparametric functional regression estimators. In most cases, it can be absolutely legitimate to measure the proximity between two elements of the infinite dimensional space by using a semi-metric, which could prevent those estimators suffering from the curse of dimensionality. In the finite dimensional case, the convergence rate is expressed in terms of h^d, where h is smoothing parameter and d is the space dimention. However, in the functional context (the explanatory variable takes its values in a semi-metric space of infinite dimension (E, d)), the asymptotic results are expressed from more general quantities called probabilities of small balls and defined by the function ϕ_x such that

$$\phi_x(h) := \mathbb{P}(d(X, x) \leq h). \tag{1.1}$$

It can be easily observed in the literature that the convergence rate of the regression estimator is linked both to the law of the explanatory variable, to the topology under consideration and consequently, on the manner under which these probabilities decreases to 0. In the literature,

small ball probabilities of various types are studied and applied to many problems of interest under different names such as small ball probability, lower tail behaviors, etc. In addition, there is a fairly large number of probabilistic results that study the manner under which these probabilities tend towards 0 in the case where d is a norm (see for instance, Bogachev [18], Shmileva [89], Gao and Li [59]). We can cite, also, the survey paper of Li and Shao [71] for Gaussian processes, together with its extended references, which covers much of the recent progress in this area.
The problem of small ball measurement $\mathbb{P}(\|X - x\| < h)$, comes down to the problem of the measurement of small ball $\mathbb{P}(\|X\|| < h)$ centered in 0. The measurement of this quantity can be specified in certain situations. Mayer-Wolf and Zeitouni (1993) [77] studied the one-dimensional case of diffusion processes, under conditions on a point x, they also studied the non-Gaussian case for the same process.

Examples of continuous time processes:

We will cite some examples of processes, including the property of concentration which is written in the following form:

$$\mathbb{P}(\|X - x\| \leq h) \approx C_x h^{\alpha} exp\Big(- \frac{C}{h^{\beta}}\Big),$$

where α, β, C_x and C are positive constants and $||.||$ can be the uniform norm, L^P. The applications below show for some functional variables the effect of measuring small balls.

General diffusion processes:

Let us consider the space $C\left([0,1], \mathbb{R}^p\right)$ and its Cameron–Martin associated space: $\mathcal{F} = C\left([0,1], \mathbb{R}^p\right)^{CM}$, where the metric $d(.,.)$ is still the one associated with the supremum norm. Let us consider some diffusion process ζ^{Diff} that can be written on the following usual form:

$$\zeta_t^{Diff} = \int_0^t \beta\left(s, \zeta^{Diff}\right) ds + w_t,$$

where w is the standard Wiener process and β is such that the solution of the above equation has an unique solution ζ^{Diff} (examples of such functions β can be found for instance in Dabo-Niang [30] or in Banon [8]). In Lipster and Shiryayev (1977), it is shown that if the condition

$$\int_0^1 \beta^2\left(s, \zeta^{Diff}\right) dt < \infty, \quad a.s$$

holds, then the process ζ^{Diff} is absolutely continuous with respect to the Wiener measure P^W. This shows once again that

$$\forall x_0 \in \mathcal{F},\, \mathbb{P}\left(\zeta^{Diff} \in B(x_0, h)\right) \approx C_{x_0} exp(-\frac{C}{h^2}).$$

Fractional Brownian Motion:

Let us always consider the space $C\left([0,1],\mathbb{R}\right)$, endowed with its uniform norm, and let $\mathcal{F}$ be its Cameron-Martin space.
Let ζ^{FBM}, the Fractional Brownian Motion [1] of parameter δ, $0 < \delta < 2$. Li and shao [71] studied the measurement of small balls for Brownian motion, according to Theorems 3.1 and 4.6 of Li and shao [71] we have the following property:

$$\forall x_0 \in \mathcal{F},\, C'_{x_0} exp\left(h^{-\frac{2}{\delta}}\right) \le \mathbb{P}\left(\zeta^{FBM} \in B(x_0, h)\right) \le C_{x_0} exp\left(h^{-\frac{2}{\delta}}\right).$$

This allows to write:

$$\forall x_0 \in \mathcal{F},\, \mathbb{P}\left(\zeta^{FBM} \in B(x_0, h)\right) \approx C_{x_0} exp\left(h^{-\frac{2}{\delta}}\right).$$

Semi-metrics:

A crucial problem when dealing with functional predictors is the choice of the semi-metric d, contrary to models with predictors that take values in finite dimensional space, since all norms are equivalent. This concept fails for functional predictors since they take values in an infinite dimensional space. Even more, restricting d to be a metric is sometimes too restrictive in the functional framework. That is why semi-metrics are considered. In other words, the choice of semi metrics allows to extract as much information possible from the functional variable and constitutes an alternative to problems related to large dimensions. In general, the choice which semi-metric to take depends on the shape of the data and the goal of the statistical analysis. There are three main families of semi-metrics which are more usable, but, of course, others can be constructed. These three semi-metrics respectively are based on derivation, principal component analysis (see Jolliffe [62]) and functional index model semi-metrics.

1. **Semi-metrics based on derivatives** One way of constructing a family of semi metrics between curves consists in considering a distance between their derivatives. More precisely,

[1] A Fractional Brownian Motion β of order δ is centered Gaussien process such that

$$\beta_0 = 0 \text{ and } \forall t \neq s \in [0,1], \quad \mathbb{E}\left(|t-s|\right) = |t-s|^{\delta}$$

the semi-metric given two curves X_1 and X_2 is:

$$d_q^{deriv}(X_1, X_2) = \sqrt{\int \left(X_1^{(q)}(t) - X_2^{(q)}(t)\right)^2 dt} \quad q \in \mathbb{N},$$

where $X^{(q)}$ denotes the q^{th} derivative of X. This class of semi metrics is suitable when we are dealing with smooth data.

2. **Semi-metrics based on functional principal components analysis**
This family of semi metrics is used for irregular data, it is introduced by Besse et *al.* [16]; and it computes proximities between curves based on the functional principal components analysis indexed by an integer q indicating the first q eigenvectors of the empirical covariance operator, associated with the q largest eigenvalues. This method reduces the functional data in a reduced dimensional space, but, this kind of semi-metric can be used only if the curves are observed at the same discretized points and in a grid sufficiently fine. Such semi-metric is then defined by:

$$d_q^{PCA}(X_1, X_2) = \sqrt{\sum_{k=1}^{q} \left(\sum_{j=1}^{J} w_j (f_1(t_j) - f_2(t_j)) [v_k]_j \right)^2},$$

where where v_k is the k^{th} orthonormal eigenvector of the covariance matrix $W = diag(w_1, \ldots, w_J)$ with quadrature weights, and f_i is the score of the principal component $\int X_i v$.

3. **Projections type semi-metrics.** This kind of semi-metric can be constructed in various different ways according to the Hilbert space and to its selected orthonormal basis. For instance, it concerns Fourier basis, as various wavelet bases, as well as the Functional PCA projection. More precisely,

$$\forall (x, x') \in \mathcal{H} \times \mathcal{H}, d_k(x, x') = \sqrt{\sum_{j=1}^{k} \langle x - x', e_j \rangle^2},$$

where $\mathcal{H}$ is a separable Hilbert space with inner product $\langle ., . \rangle$ and $\{e_j, j = 1, \ldots \infty\}$ an orthonormal basis.

1.3 Ergodic data

The term "ergodic" comes from the Greek words (ergon, odos) which mean (work, path). Ergodic theory is a fundamental hypothesis of statistical physics, it models the thermodynamic properties of gases, atoms, electrons or plasmas, this condition is also used in signal processing and for the study of the evolution of a kinetic gas signal. It has undergone many developments closely related to dynamic systems theory and chaos theory. The first one who posed this theory is Ludwing Boltzmann in 1871 for the needs of his kinetic theory of gases.

The ergodic theory constitutes a recent and important research area in the study of stochastic processes. This study has a very wide range of applications, because most of the random phenomena we encounter around us are not independent. The ergodic processes are a class of stochastic processes that have the property that one sample of the process represents all the set. This theory represent now a very fashionable research area. From a practical point of view, the strongly mixing dependence suffers from many undesirable features. In fact, there are so many models that have been given in the literature where the mixing properties still need to be verified or even fail to hold for the processes they induce. Moreover, Davidson [35] and Laïb and Louani [67] pointed out that it is difficult to verify the strongly mixing condition in practice as for example, the processes AR(1). Indeed, even very simple autoregressive processes cannot be strongly mixing for some cases. Chernick [27] and Andrews [2] have given the $AR(1)$ linear real process with discrete valued random innovation which is not strongly mixing but ergodic. In the cas of functional data, we can take the example of Xiong and Lin [93], in the example the authors take $\mathcal{H} = C[-1,1]$, which is the space composed of all of the continuous functions defined in $[-1,1]$ and the corresponding semi-metric $d(y,z) = |\int_{-1}^{1}[y(t) - z(t)]dt|, \quad \forall y,z \in \mathcal{H}$. Consider the autoregressive model of order one defined, for any $i \in \mathbb{Z}$, by

$$2T_{i+1} = T_i + E_{i+1}, \tag{1.2}$$

where $E_{i+1} = e_{i+1}h$, e_{i+1} is independent of T_i, $h \in \mathcal{H}, h(t) = t^2(t \in [-1,1])$, and the e_i are independent real random variables with common distribution $B(1,\frac{1}{2})$. Then the stationary solution of (1.2) is:

$$T_i = \sum_{j=0}^{\infty} 2^{-j-1} e_{i-j} t^2, \qquad t \in [-1,1], i \in \mathbb{Z}.$$

It follows from the theorem of ergodic A.3.1 that the process $T = \{T_i : i \in \mathbb{Z}\}$ is ergodic. In addition, (1.2) and the distribution of e_{i+1} implis that $\sigma(T_i) \subseteq \sigma(T_{i+1})$. By iteration we get

$$\sigma(T_i) \subseteq \sigma(T_k,\, k \geq i+1),$$

thus, we have that the mixing coefficient α_n [2]

$$\frac{1}{4} \geq \alpha_n \geq \alpha\left(\sigma(T_t), \sigma(T_t)\right) = \frac{1}{4},$$

which shows that the process T is not strongly mixing.
The literature on the ergodic functional data is still limited. The first intersting result on this subject was obtained by Laïb and Louani [67]. They considered the regression function estimation when the data are functional and assumed to be sampled from a stationary and ergodic process, they established the consistency in probability, with a rate and the asymptotic normality of the estimator. In (2011) [68], those same authors studied the strong pointwise and uniform consistencies with rates of the same estimator. In the same domain of the ergodic data, Gheriballah et *al.*[61] established the asymptotic properties of an alternative estimator of the nonparametric regression function. Chaouch and Khardani [25] introduced a kernel-type estimator of the conditional quantile function of a randomly censored scalar response variable given a functional random covariate whenever a stationary ergodic data are considered. The authors established a strong consistency rate as well as the asymptotic distribution of their estimator. One year later, a nonparametric M-estimation for right censored regression model with stationary ergodic data was investigated by Chaouch et *al.* [26]. They derived under mild assumption the strong consistency (with rate) and the asymptotic distribution of the estimator. In the same year, Benziadi et *al.* [11] considered two recursive estimators of the conditionals quantiles when the explanatory variable is of the ergodic type. Indeed the first one is obtained by using the robust approach while the second estimator is given by inverting the double-kernel estimate of the conditional distribution function. The recursive conditional mode was treated by Ardjoun et *al.* [4].

1.4 Censored data

The existence of incomplete observations is one of the characteristics of survival data such as in epidemiological surveys, data is often collected incompletely. The censoring scheme is an important concept and the most common phenomenon in survival analysis in that one can observe partial information associated with the survival random variable. This is due to some limitations such as loss to follow-up, drop-out, termination of the study, and others. In all these

[2] Let $T = (T_t, t \in \mathbb{Z})$ be a strictly stationary process, its strong mixing coefficient of order n is defined as

$$\alpha_n = \sup_{\substack{B \in \sigma(T_s, s \leq t) \\ C \in \sigma(T_s, s \geq t+n)}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|, \quad n \geq 1.$$

For such a process α_n does not depend on t. Now T is said to be **strongly mixing (or** $\alpha-$**mixing)** if $\lim \alpha_n = 0$. A process $(T_t, t \in \mathbb{Z})$ is said to be **strongly mixing(or** $\alpha-$**mixing** if
$\alpha_n = \sup_{t \in \mathbb{Z}} \alpha\left(\sigma(T_s, s \leq t), \quad \sigma(T_s, s \geq t+n)\right) \longrightarrow_{n\to\infty} 0$, where the "sup" may be omitted if (T_t) is stationary.

cases, the exact time of the event is not observed because the event did not occur. Therefore, censored data arises when a subject's time until the event of interest is known only to occur in a certain period of time.
Depending on the direction of the censoring, censored data can be classified into right censored when the survival time exceeds the observed one, and left censored when the survival time is less than the observed one. Left censoring is particularly important in studies on infectious diseases such hepatitis or human immunodeficiency.

Right Censoring

There is right censoring if instead of observing the variables $\{T_1, T_2, \ldots, T_n\}$ that interests us, we do not observe T_i that when $T_i < C$ (censorship variable). In the real of right censored data, a distinction can be made among three different types of censoring:

Type I censorship

The subjects enter the study at the same time, at a given date the study ends and some of them are lost to follow up or the event is not occurred. This that means, instead of observing the variables $T_1, T_2, \ldots, T_n$ which interest us, we observe T_i when it is less than a fixed duration C, otherwise we only know that T_i is greater than C. We therefore observe a variable Y_i such that $Y_i = \min(T_i, C)$.
This model is often used in epidemiological studies.

Type II censorship:

The subjects enter the study at the same time, the end of the study is not initially fixed and it is carried on until the event occurs for a certain proportion of subjects. More precisely, this type of censoring is present when we decide to observe the survival times of n patients until r of them have died and to stop study at that time.
Let $T_{(i)}$ and $Y_{(i)}$ be the order statistics of the variables T_i and Y_i. The censoring date is so $Y_{(r)}$ and we observe the following variables:
$Y_{(1)} = T_{(1)}, Y_{(2)} = T_{(2)}, \ldots, Y_{(r)} = T_{(r)}, Y_{(r+1)} = T_{(r)}, \ldots, Y_{(n)} = T_{(r)}$.
This model is often used in reliability studies.

Type III censorship (or type I random censorship)

The subjects enter the study at different times. Let $C_1, C_2, \ldots, C_n$ be i.i.d. random variables. We observe the variables $T_i = Y_i \wedge C_i$ and $\delta_i = \mathbb{1}_{\{T_i \leq C_i\}}$. Y_i is the duration actually observed, and

$$\delta_i = \begin{cases} 1 & \text{if the event is observed } (T_i = Y_i) \\ 0 & \text{if the individual is censored } (T_i = C_i). \end{cases}$$

Left Censoring

We say that there is left censorship if instead of observing $Y_1, \ldots, Y_n$, we observe (T_i, δ_i) where $T_i = \max(Y_i, \delta_i)$ and $\delta_i = 1\!\!1_{\{T_i \geq C_i\}}$ for $i = 1, \ldots, n$ and C_i is a random censoring.

1.5 Conditional models in functional statistics

Generally, prediction of a scalar response given an explanatory variable is obtained by estimating the conditional expectation (the regression function). However, this method is not adequate in some situations. For instance, this the case when the conditional density function is either unsymmetrical or has several modes. In this cases, a relevant predictor is obtained by the nonparametric estimation of the conditional mode which is a direct consequence of estimating the conditional density.

The conditional density presents a good alternative of the regression operator, and it has know of great interests in statistics, as this functional parameter is involved in the estimators of the mode, the hazard function ... ect. Moreover, the conditional density provides a very informative summary on response variables because it allows us to examine the overall shape of the conditional distribution (see Fan and Yao [47] and references therein).
In the infinite dimensional setting, the behavior of the nonparametric estimators of the conditional density is extensively studied. In particular, the consistency has been investigated by many authors. The first important results on this topic have been established by Ferraty et al. [55]. They proved the almost complete convergence of the kernel estimator of the conditional density and its derivatives. An application of their results to data from the food industry was presented. The uniform almost complete convergence was studied by Ferraty et *al.* [33], the authors specified the rate of convergence. In 2007, Laksaci [69] studied the quadratic error of this estimator and gave the asymptotic expansion of the exact expression involved in the leading terms of the quadratic error of the considered estimator. In the same year, Dabo-Niang and Laksaci [34] added some results on the convergence in L^p norm of the kernel estimator of the conditional mode in the case where the data are i.i.d. Three years later, Ferraty et al. [58] established the uniform almost complete convergence of the kernel estimator for some nonparametric conditional parameters, in particular, for the conditional density function. Concerning the asymptotic normality of kernel estimators of the conditional modes , Many authors were interested in this asymptotic property, we refer to Ezzahrioui and Ould-saïd [44] who treated the case i.i.d.

The estimation of the conditional distribution function plays an important role in the estimation of other functional parameters. In the functional framework, the estimation of this operator was introduced by Ferraty et *al.* [55], where they constructed a double kernel estimator of the conditional distribution function and studied the rate of the almost complete convergence of the proposed estimator in the case where the observations are independent and identically distributed. Ferraty et *al.* [53], Mahiddine et *al.* [74], Rabhi et *al.* and Bouchentouf et *al.* [83, 20] studied the estimation of conditional distribution function in the case where observations are functional and $\alpha-$ mixing. On the other hand, several authors treated the estimation of the conditional distribution function as a preliminary study of the estimation of conditional quantiles. Let us quote for example, Ferraty et *al.* [53]. In that paper, authors established the convergence almost complete consistency with a rate of a conditional quantile estimator under $\alpha-$mixing conditions. An example of application to prediction via the conditional median, as well as the determination of prediction intervals was considered in Ferraty et *al.* [53]. Concerning the asymptotic normality, Ezzahrioui et Ould-Saïd [42, 43] studied this asymptotic property of the kernel estimator of the conditional quantile in the both case (i.i.d and dependent).

The hazard function, sometimes called the risk function, is very frequently used in the study of statistical reliability and it is a functional parameter of great importance in many practical problems. We note that the use of this parameter is applied to many branches of research under slightly different names, including reliability analysis (engineering), duration analysis (economics), and the analysis of the history of the event (sociology). The literature on the conditional hazard estimation in functional statistics is very restricted. The paper of Ferraty et *al.* [56] is a precursor work on the subject. The authors of this paper studied the almost complete convergence of a kernel estimator of the conditional hazard function as part of i.i.d. complete data (respectively dependent) data, as well as in the framework of censored i.i.d. (respectively dependent) data. Quantela-del Rio [81] studied the almost complete convergence and the mean square error. Ferraty et *al.* [56] and Quantela-del Rio [81] studied the asymptotic normality of a kernel estimator of the conditional hazard function. We can also cite the paper of Laksaci and Mechab [70] on the estimation of the conditional hazard function for spatially dependent functional data. An other point of view of this functional parameter was studied by Rabhi et *al.* [82] by establishing an estimate of the maximum of the conditional hazard function under dependency conditions.

1.6 Local linear estimation for functional data

The simplicity of the classical kernel estimate and its availability in many statistical software packages, like R and Matlab make it easy to understand and implement. However, its simplicity leads to some weaknesses, the most obvious of them is boundary bias effect. And it is well known that among the smoothing procedures, the local polynomial approach has various advan-

tages over the classical kernel method. This last approach was the subject of an open question in the monograph Ferraty and Vieu [54], and several authors have tried to provide answers. The first answer was given by Baìllo and Grané [7], who gave the first functional version of local linear estimation. This is based on the Hilbert structure of space. Baìllo and Grané [7] introduced a local linear regression estimator and they determined its asymptotic behavior by determining its bias as well as its variance. In the same case of Hilbertien space, Berlinet et *al.* [13] constructed another local linear estimator of the regression function based on the inverse of the covariance operator, they proposed a pointwise estimate and they derived its asymptotic mean square error. Barrientos et *al.* [9] constructed a fast local linear estimator of the regression operator, they established their results in semi metric space to establish the almost complete convergence of their estimator, they have even specified its convergence rate which is very optimal. In 2013, Demangeot et *al.* [37] studied the local linear estimation of the conditional density and they established the pointwise almost complete convergence of the constructed estimator, they also deduced asymptotic properties of local linear estimator of conditional mode. The case of the conditional distribution function was addressed by Demongeot et *al.* [38]. The authors studied the almost complete consistency as well as the mean square error with rate of the constructed estimator. All the results which we mentioned have been established in the case when the data are independent. The case of functional spatially dependent data was studied by Chouaf and Laksaci [29]. The authors extended the results of Barrientos [9] by giving the almost complete consistency with rate of the spatial version of the mentioned estimator.
Concerning the asymptotic normality, we can cite the work of Zhou and Lin [94]. In this paper, the mean squared consistency and the asymptotic normality of the locally modelled regression estimation is obtained when the data are independent and identically distributed.
Recently, Bouanani et *al.* [21, 22] established the asymptotic normality of several conditional models in the both dependent and independent cases. We can cite also, Chikr-Elmezouar et *al.* [28] who introduced the Kernel Nearest Neighbor (KNN) method, and presented an estimator of the conditional density and mode when the co-variables are functional They established the almost complete convergence of these two functional parameters. Belarbi et *al.* [10] studied the robust estimation of the functional local linear regression model and they established the almost complete convergence as well as the asymptotic normality of the constructed estimator. In 2017, Xiong et *al.* [92] studied the asymptotic normality of a local linear estimator of the conditional density in the case where the observations are alpha-mixing. In the same year, Demongeot et *al.* [39] proposed a local linear estimator of the regression function where the both variables (response and explanatory) are functional by giving the almost complete convergence of the proposed estimator, whereas Chahad et *al.* [24] studied the pointwise and uniform almost complete convergence of the relative regression estimator when the observations are (i.i.d).

1.6.1 Construction of the local linear estimator

In the complete data case

Let $Z_i = (X_i, Y_i)_{i=1,\dots,n}$ be an $\mathcal{F}\times\mathbb{R}$-valued measurable strictly stationary process, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, where $\mathcal{F}$ is a semi-metric space, and d denotes the semi-metric. Furthermore, we assume that there exists a regular version of the conditional distribution of Y given X, which is absolutely continuous with respect to the *Lebesgue* measure on $\mathbb{R}$, and has a twice continuously differentiable probability density function denoted by $f^X(Y)$.
We focus on the estimation of the conditional distribution(respectively conditional density) of Y given $X = x$ via the local linear method. For this purpose, it is well known that the main idea, in the local linear smoothing, is based on the fact that the function $F^x(y)$ (respectively $f^x(y)$) admits a linear approximation in the neighborhood of the conditioning point. This consideration is motivated by the fact that the conditional distribution function (respectively the conditional density function) can be expressed as a regression model with the response variable $J(\frac{\cdot-Y}{h_J})$ (respectively $\frac{1}{h_J}J(\frac{\cdot-Y}{h_J})$) instead Y, where J is a distribution function(respectively J is a kernel function) and $h_J = h_{J,n}$ is a sequence of positive real numbers under the condition $h_J \longrightarrow 0$.
For this aim, we assume that the underlying process Z_i is functional stationary ergodic, and we propose to construct the estimator $\widehat{F^x}$ of F^x by $\widehat{F^x} = \widehat{a_0}$ (respectively $\widehat{f^x}$ of f^x by $\widehat{f^x} = \widehat{a_0}$) which is obtained from the following minimization procedure:

$$\min_{(a_0,a_1)\in\mathbb{R}^2} \sum_{i=1}^{n} \left(\frac{1}{h_J^l} J\left(\frac{y-Y_i}{h_J}\right) - a_0 - a_1 \rho(X_i, x) \right)^2 K\left(\frac{\delta(x, X_i)}{h_K}\right), \qquad l = 0, 1 \qquad (1.3)$$

with $\rho(.,.)$ and $\delta(.,.)$ are known bi-functional operators defined from $\mathcal{F}^2$ into $\mathbb{R}$ such that $|\delta(x,z)| = d(x,z)$ and $\rho(z,z) = 0, \forall z \in \mathcal{F}$. h_K is the smoothing parameter associated with the kernel K.
$\widehat{g}_l^{(x)}(y)$ is the solution of the problem of minimization (1.3) and we have:

$$\widehat{g}_l^{(x)}(y) = h_J^{-l}\, {}^{\mathbf{t}}\mathbf{u_1} \left({}^{\mathbf{t}}\mathbf{Q}_\rho \mathbf{K} \mathbf{Q}_\rho\right)^{-1} \mathbf{Q}_\rho \mathbf{K} \mathbf{J}$$

where ${}^t\mathbf{Q}_\rho$ is the matrix defined by:

$${}^t\mathbf{Q}_\rho = \begin{bmatrix} 1 & \cdots & 1 \\ \rho(X_1, x) & \cdots & \rho(X_n, x) \end{bmatrix}$$

and

$$\mathbf{J} =^{\mathbf{t}} \left[\mathbf{J}\left(\frac{\mathbf{y}-\mathbf{Y_1}}{\mathbf{h_J}}\right), \dots, \mathbf{J}\left(\frac{\mathbf{y}-\mathbf{Y_n}}{\mathbf{h_J}}\right)\right] \quad \text{and} \quad {}^{\mathbf{t}}\mathbf{u_1} =^{\mathbf{t}} [\mathbf{1}, \mathbf{0}] \in \mathbb{R}^{\mathbf{2}},$$

(where A^t designs the transposed matrix of matrix A).
One designates by K the diagonal weight matrix:

$$\mathbf{K} = \begin{pmatrix} K\left(\frac{\delta(x,X_1)}{h_K}\right) & 0 & \cdots & 0 \\ 0 & K\left(\frac{\delta(x,X_2)}{h_K}\right) & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & K\left(\frac{\delta(x,X_n)}{h_K}\right) \end{pmatrix}$$

To calculate $\widehat{g}_l^{(x)}(y)$, we note that:

$$\rho_i = \rho(X_i, x), K_i = K\left(\frac{\delta(x, X_i)}{h_K}\right) \text{ and } \quad J_j = J\left(\frac{y - Y_j}{h_J}\right).$$

$$\widehat{g}_l^{(x)}(y) = \frac{1}{{h_J}^l}(1,0)\left[\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \rho_1 & \rho_2 & \cdots & \rho_n \end{pmatrix}\begin{pmatrix} K_1 & 0 & \cdots & 0 \\ 0 & K_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & K_n \end{pmatrix}\begin{pmatrix} 1 & \rho_1 \\ 1 & \rho_2 \\ \vdots & \vdots \\ 1 & \rho_n \end{pmatrix}\right]^{-1}$$
$$\times \begin{pmatrix} 1 & \rho_1 \\ 1 & \rho_2 \\ \vdots & \vdots \\ 1 & \rho_n \end{pmatrix}\begin{pmatrix} K_1 & 0 & \cdots & 0 \\ 0 & K_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & K_n \end{pmatrix}\begin{pmatrix} J_1 \\ J_2 \\ \vdots \\ J_n \end{pmatrix}$$

with

$$\left({}^t\mathbf{Q}_\rho \mathbf{K} \mathbf{Q}_\rho\right)^{-1} = \frac{\begin{pmatrix} \sum_{i=1}^n \rho_i^2 K_i & \sum_{i=1}^n \rho_i K_i \\ -\sum_{i=1}^n \rho_i K_i & \sum_{i=1}^n K_i \end{pmatrix}}{\sum_{i=1}^n \sum_{j=1}^n \rho_i(\rho_i - \rho_j) K_i K_j}.$$

Let

$$\sum_{i=1}^n \sum_{j=1}^n W_{ij} = \sum_{i=1}^n \sum_{j=1}^n \rho_i(\rho_i - \rho_j) K_i K_j,$$

thus

$${}^t\mathbf{u_1}\left({}^t\mathbf{Q}_\rho \mathbf{K} \mathbf{Q}_\rho\right)^{-1} = \frac{\left(\sum_{i=1}^n \rho_i^2 \mathbf{K_i}, \quad -\sum_{i=1}^n \rho_i \mathbf{K_i}\right)}{\sum_{i=1}^n \sum_{j=1}^n \mathbf{W_{ij}}},$$

and

$$ {}^t\mathbf{Q}_\rho \mathbf{K}\mathbf{J} = \begin{pmatrix} \sum_{i=1}^n K_i J_i \\ \sum_{i=1}^n \rho_i K_i J_i \end{pmatrix}. $$

Therefore,

$$ \begin{aligned} \widehat{g}_l^{(x)}(y) &= \frac{1}{h_J^l \sum_{i=1}^n \sum_{j=1}^n W_{ij}} \left(\sum_{i=1}^n \rho_i^2 K_i - \sum_{i=1}^n \rho_i K_i \right) \begin{pmatrix} \sum_{i=1}^n K_i J_i \\ \sum_{i=1}^n \rho_i K_i J_i \end{pmatrix} \\ &= \frac{1}{h_J^l \sum_{i=1}^n \sum_{j=1}^n W_{ij}} \left[\left(\sum_{i=1}^n \rho_i^2 K_i \right) \left(\sum_{i=1}^n K_i J_i \right) - \left(\sum_{i=1}^n \rho_i K_i \right) \left(\sum_{i=1}^n \rho_i K_i J_i \right) \right]. \end{aligned} $$

This leads to obtaining the following solution:

$$ \widehat{g}_l^{(x)}(y) = \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} J_j}{h_J^l \sum_{i=1}^n \sum_{j=1}^n W_{ij}}. $$

Then, it is obvious that (1.6.1) can be rewritten as

$$ \widehat{g}_l^{(x)}(y) = \frac{\sum_{j=1}^n \Gamma_j K_j J_j}{h_J^l \sum_{j=1}^n \Gamma_j K_j}, $$

with

$$ \Gamma_j = K_j^{-1} \left(\sum_{i=1}^n W_{ij} \right) = \sum_{i=1}^n \rho_i^2 K_i - \left(\sum_{i=1}^n \rho_i K_i \right) \rho_j. $$

To obtain the estimator of the conditional distribution function , we take $l = 0$ and we get :

$$ \widehat{F}^x(y) = \widehat{g}_0^{(x)}(y) = \frac{\sum_{j=1}^n \Gamma_j K_j J_j}{\sum_{j=1}^n \Gamma_j K_j}. $$

And to obtain the estimator of the conditional density function , we take $l = 1$ and we get :

$$\widehat{f}^x(y) = \widehat{g}_1^{(x)}(y) = \frac{\sum_{j=1}^{n} \Gamma_j K_j J_j}{h_J \sum_{j=1}^{n} \Gamma_j K_j}.$$

Concerning the estimate $\widehat{\Theta}(x)$ of the conditional mode $\Theta(x)$ by the local linear approach. We assume that there exist a compact $\mathscr{C}$ where the conditional density f^x has a unique mode $\Theta(x)$ on $\mathscr{C}$. A natural and usual estimator of $\Theta(x)$ is defined as the random variable $\widehat{\Theta}(x)$ which maximizes the local linear estimator $\widehat{f}^x(.)$ of $f^x(.)$ that is:

$$\widehat{\Theta}(x) = \arg\sup_{y \in \mathscr{C}} \widehat{f}^x(y).$$

In the censored data case

In the censoring case, we can only observe the triplets $(X_i, T_i, \delta_i)_{1 \le i \le n}$, where

$$T_i = Y_i \wedge C_i \quad \text{and} \quad \delta_i = \mathbb{1}_{\{Y_i \le C_i\}} \qquad 1 \le i \le n,$$

with $\mathbb{1}_A$ denotes the indicator function on a set A and C_i is the censoring random variable with unknown continous distribution function G.
We assume that $(C_i)_{1 \le i \le n}$ and $(X_i, Y_i)_{1 \le i \le n}$ are independent. Based on the same idea as in Carbonez et *al.* [23] and Khardani et *al.* [66], we give the "pseudo" estimator of $f^x(y)$:

$$\tilde{f}^x(y) = \frac{\sum_{j=1}^{n} \delta_j \bar{G}^{-1}(T_j) \Gamma_j K_j J_j}{h_J \sum_{j=1}^{n} \Gamma_j K_j}.$$

Since G is unknown in practice, it is not possible to use the estimator (1.6.1), we use the Kaplan Meier [65] estimator of G given by:

$$\bar{G}_n(y) = \begin{cases} \prod_{j}^{n} \left(1 - \frac{1 - \delta_{(j)}}{n - j + 1}\right)^{\mathbf{1}_{\{T_{(j)} \le y\}}} & \text{if} \quad y < T_{(n)}, \\ 0 & \text{otherwise,} \end{cases}$$

where $T_{(1)} < T_{(2)} < \ldots < T_{(n)}$ are order statistics of T_j and $\delta_{(j)}$ is concomitant with $T_{(j)}$.

Thus a feasible estimator of $f^x(y)$ is given by

$$\widehat{f^x}(y) = \frac{\sum_{j=1}^{n} \delta_j \bar{G}_n^{-1}(T_j)\Gamma_j K_j J_j}{h_J \sum_{j=1}^{n} \Gamma_j K_j}$$

Then, We assume that there exists a certain compact set $\mathcal{C}_{\mathbb{R}} \subset \mathbb{R}$, such that $f^x(y)$ has an unique mode $\theta(x)$ on $\mathcal{C}_{\mathbb{R}}$. A natural and usual estimator of $\theta(x)$ is defined by:

$$\widehat{\theta}(x) = \arg\sup_{y \in \mathcal{C}_{\mathbb{R}}} \widehat{f^x}(y).$$

1.7 Brief presentation of the results

We give in this section a short presentation of the results obtained in this thesis.

1.7.1 Results: The complete case

In this part, we consider a subset $\mathcal{C}_{\mathcal{F}}$ of $\mathcal{F}$ such that $\mathcal{C}_{\mathcal{F}} \subset \bigcup_{k=1}^{d_n} B(x_k, r_n)$ where $x_k \in \mathcal{F},\ r_n$ and d_n are two sequences of positive real numbers and $B(x_k, r_n) = \{x'_k \in \mathcal{F}/|\delta(x'_k, x_k)| < r_n\}$. For any fixed y in $\mathbb{R}$, $\mathcal{N}_y$ denotes a fixed neighborhood of y and let $\phi_x(h_1, h_2) = \mathbb{P}(h_2 \leq \delta(X, x) \leq h_1)$ the small ball probability function.
The following theorems give the uniform almost complete convergence (with rate), then the asymptotic normality of $\widehat{F^x}(y)$.

Theorem 1.7.1. *Under some assumptions, we have*

$$\sup_{x \in \mathcal{C}_{\mathcal{F}}} |\widehat{F^x}(y) - F^x(y)| = O\left(h_K^{b_1} + h_J^{b_2}\right) + O\left(\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right), \quad a.co.$$

Theorem 1.7.2. *Under some assumptions and if the smoothing parameters h_K and h_J satisfy* $\sqrt{n\phi(h_K)}\left(h_K^{b_1} + h_J^{b_2}\right) \to 0$ *as* $n \to \infty$, *we have*

$$\sqrt{n\phi(h_K)}\left(\widehat{F^x}(y) - F^x(y)\right) \xrightarrow{\mathcal{D}} N\left(0, V_{JK}(x, y)\right),$$

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

$$V_{JK}(x, y) = \frac{M_2}{M_1^2} F^x(y)\left(1 - F^x(y)\right),$$

with

$$\Psi(t) = \lim_{h_K \to 0} \frac{\phi(-h_K, th_K)}{\phi(h_K)}, \ \forall t \in [-1, 1].$$

and

$$M_c = K^c(1) - \int_{-1}^{1} (K^c(u))' \Psi(u) du \quad \text{where } c = 1, 2.$$

Proof of these results and details of the conditions imposed are given in the Chapter 2.

In the following, we introduce some notations:
For any fixed x in $\mathcal{F}$, $\mathcal{N}_x$ denotes a fixed neighborhood of x; and let $\mathscr{C}$ be a fixed compact subset of $\mathbb{R}$.
In this part, we study the almost complete consistency (a.co.) with rates of the local linear estimator of the conditional density and we derive some asymptotic properties for the local linear estimator of the conditional mode.
The first Theorem states the pointwise almost complete convergence, whereas the second one precises the rate of convergence.

Theorem 1.7.3. *[6] Under some strictrurel regularity and technical assumptions, we have*

$$\sup_{y \in \mathscr{C}} |\widehat{f^x}(y) - f^x(y)| = o(1), \qquad a.co.$$

Theorem 1.7.4. *[6] Under the same assumption of Theorem 1.7.3 with a little change in the regularity assumption, we obtain*

$$\sup_{y \in \mathscr{C}} |\widehat{f^x}(y) - f^x(y)| = O\left(h_K^{b_1}\right) + O\left(h_J^{b_2}\right) + O\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right), \quad a.co,$$

where b_1, b_2 are positive constants linked to the Lipchitz condition and $\varphi_x(h_K) = \sum_{i=1}^{n} \phi_{i,x}(h_K)$.
Next, this corollary concerns the almost comlete covergence with rate of the local linear estimator of the conditional mode.

Corollary 1.7.1. *[6] Under some assumptions, we have:*

$$|\widehat{\Theta}(x) - \Theta(x)| = O\left(h_K^{\frac{b_1}{j}}\right) + O\left(h_J^{\frac{b_2}{j}}\right) + O\left(\left(\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}\right)^{\frac{1}{2j}}\right), \quad a.co.$$

Proof of these results and details of the conditions imposed are given in the Chapter 3.

1.7.2 Results: The censored case

We establish the almost complete convergence with rate of the conditional density function which extends the results established in Ayad et *al.* [6] to the censored case.

Theorem 1.7.5. *Under some assumptions, we have:*

$$\sup_{y\in\mathcal{C}_{\mathbb{R}}}|\widehat{f}^{x}(y)-f^{x}(y)|=O\left(h_K^{b_1}\right)+O\left(h_J^{b_2}\right)+O\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2h_J\phi_x^2(h_K)}}\right),\quad a.co.$$

This result gives the convergence rate of the local linear estimator of the conditional mode.

Theorem 1.7.6. *Under some assumptions, we have:*

$$|\widehat{\theta}(x)-\theta(x)|=O\left(h_K^{\frac{b_1}{2}}\right)+O\left(h_J^{\frac{b_2}{2}}\right)+O\left(\left(\frac{\varphi_x(h_K)\log n}{n^2h_J\phi_x^2(h_K)}\right)^{\frac{1}{4}}\right),\quad a.co.$$

Proof of these results and details of the conditions imposed are given in the Chapter 4.

1.8 Structure of the thesis

As in the literature, there are no asymptotic results on the local linear estimation for the functional and ergodic data, we are interested to study the original results of some conditional models under the only assumption that the process generating the functional data is stationary ergodic.
Our asymptotic results are stated in terms of almost complete convergence which is known to imply both almost sure convergence and the convergence in probability and uniform convergence of the different conditional models.
This thesis is divided into four chapters:

The first chapter is an introductory chapter, where we present the different themes addressed in our research axis. We start with a brief history on the nonparametric statistic for functional data. This part was followed by a presentation of some some filds of application of functional data. We offer numerous bibliographic references. Then, we discuss about the ergodic data and incomplete data essentially the censored data.
Finally, in this introduction, we expose a brief historical on the local linear method and we have given the consruction of our local linear estimators.

In the second chapter, we consider the estimation of the conditional distribution function of a scalar response variable given by a random variable taking values in semi-metric space using local linear approach. This chapter consists of five sections, we start with an introduction which is the first section. At first, we construct an estimator of onditional distribution function with local linear method. In section three, we present and discuss necessary assumptions and study the uniform almost-complete convergence (with rate), as well as the asymptotic normality of the constructed estimator. Detailed proofs of technical lemmas of our main results are given in the last section. This work has been accepted in the Journal of MATHEMATICAL MODELLING AND ANALYSIS.

In the third chapter, we consider the local linear estimation of the conditional density for functional ergodic data when the regressor is valued in a semi-metric space and the response is a scalar. The organisation of this chapter is as follow. We construct the local linear estimator of the conditional density. Under ergodicity condition we study the almost complete convergence of our estimator (with rate). Then, we use the constructed estimator to estimate the conditional mode estimation and we derive the same asymptotic proprieties. This work has been published in the Journal *METRON*.

The fourth chapter consists in extending the previous chapter to the case of censored data. To do this, we construct a new local linear estimator of the density function in which censorship effects are taken into account during our observations. The organisation of this chapter is as follow. Section two introduces the construction of our local linear estimator. In the third Section, we introduce notations and hypothesis, and state the main results. Finally, the detailed proofs of our theoretical results and all technical lemmas needed are gathered in Appendix. This workhas been submitted.

Finally, the appendix is devoted to the mathematical tools and techniques used throughout this thesis. Our thesis, ends with a conclusion and some perspectives of research.

BIBLIOGRAPHY

[1] Ait-Saïdi, A., Kassa, R. and Ferraty, F. Choix optimal du paramètre fonctionnel dans le modèle à indice fonctionnel simple. *Comptes Rendus Mathematique*, **346**, 217–220 (2008).

[2] Andrews, D. WK Non-strong mixing autoregressive processes. *Journal of Applied Probability*, 930–934(1984).

[3] Attouch, M., Laksaci, A., and Ould-Said, E. Asymptotic distribution of robust estimator for functional nonparametric models. *Communications in Statistics—Theory and Methods*, **38**, 1317–1335 (2009).

[4] Ardjoun, F. Z., and Hennani, L. A. and Laksaci, A. A recursive kernel estimate of the functional modal regression under ergodic dependence condition. *J. of Statistical Theory and Practice*, **10**, 475–496(2016).

[5] Aurzada, F. and Simon, T. Small ball probabilities for stable convolutions. *ESAIM: Probability and Statistics*, **11**,327–343(2007).

[6] Ayad, S., Laksaci, A., Rahmani, S. and Rouane, R. On the local linear modelization of the conditional density for functional and ergodic data. *METRON,* **78**, 237–254 (2020).

[7] Baìllo, A. and Grané, A. Local linear regression for functional predictor and scalar response. *J. of Multivariate Anal.*, **100**, 102–111(2009).

[8] Banon, G. Nonparametric identification for diffusion processes. *SIAM Journal on Control and Optimization.*, **16**, 380–395 (1978).

[9] Barrientos, J., Ferraty, F. and Vieu, P. Locally Modelled Regression and Functional Data. *J. Nonparametr. Statist.*, **22**, 617-632 (2010).

[10] Belarbi, F.,Chemikh, S. and Laksaci, A. Local linear estimate of the nonparametric robust regression in functional data. *Statistics & Probability Letters.*, **134**, 128–133 (2018).

[11] Benziadi, F., Laksaci, A. and Tebboune, F. Recursive kernel estimate of the conditional quantile for functional ergodic data. *Communications in Statistics-Theory and Methods,* **45**, 3097–3113 (2016).

[12] Benziadi, F., Gheriballah, A. and Laksaci, A. Asymptotic normality of kernel estimator of ψ-regression function for functional ergodic data. *New Trends in Mathematical Sciences,* **4**, 268–282 (2016).

[13] Berlinet, A., Elamine, A. and Mas, A. Local linear regression for functional data, *Annals of the Institute of Statistical Mathematics*, **63**, 1047–1075 (2011)

[14] Besse, P., and Ramsay, J. O. Principal components analysis of sampled functions. *J. of Psychometrika***51**, 285–311 (1986).

[15] Besse, P. Approximation spline de l'analyse en composantes principales d'une variable aléatoire hilbertienne. *Annales de la Faculté des sciences de Toulouse: Mathématiques*, **12**, 329–349 (1991).

[16] Besse, P., Cardot, H. and Ferraty, F. Simultaneous non-parametric regressions of unbalanced longitudinal data. *Computational Statistics & Data Analysis*, **24**, 255–270 (1997).

[17] Besse, P., and Cardot, H. Modélisation statistique de données fonctionnelles. *Analyse de données. Hermes*(2003).

[18] Bogachev, VI. Gaussian Measures (Mathematical Surveys and Monographs. *Providence, RI: American Mathematical Society*, **62**(1998).

[19] Bosq, D. Nonparametric statistics for stochastic processes: estimation and prediction (Vol. 110). Springer Science & Business Media. (2012).

[20] Bouchentouf, A. A, Djebbouri, T., Rabhi, A. and Sabri, K. Strong uniform consistency rates of some characteristics of the conditional distribution estimator in the functional single-index model. *Applicationes Mathematicae* **4**, 301–322 (2014).

[21] Bouanani, O., Laksaci, A., Rachdi, M. and Rahmani, S. Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics, *Behaviormetrika*, **46**, 1–35 (2018)

[22] Bouanani, O., Rahmani, S. and Ait-Hennani, L. Local linear conditional cumulative distribution function with mixing data. *Arabian Journal of Mathematics,* , 1–19 (2019).

[23] Carbonez, A., Györfi, L., and vander Meulen, EC. Partitioning-estimates of a regression function under random censoring. *Statistics & Risk Modeling* **13**, 21–38 (1995).

[24] Chahad, A., Ait-Hennani, L. and Laksaci, A. Functional local linear estimate for functional relative-error regression. *Journal of statistical theory and practice,* **11**, 771–789 (2017).

[25] Chaouch, M., & Khardani, S. Randomly censored quantile regression estimation using functional stationary ergodic data. Journal of nonparametric statistics, **27**(1), 65-87. (2015).

[26] Chaouch, M., Laïb, N. and Ould Saïd, E. Nonparametric M-estimation for right censored regression model with stationary ergodic data. *Statistical Methodology,* **33**, 234–255 (2016).

[27] Chernick, M. R. A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. *The Annals of Probability,* **9**, 145–149 (1981).

[28] Chikr-Elmezouar, Z., Almanjahie, IM., Laksaci, A and Rachdi, M. FDA: strong consistency of the k NN local linear estimation of the functional conditional density and mode. *Journal of Nonparametric Statistics,* **31**, 175–195 (2019).

[29] Chouaf, A. and Laksaci, A. On the functional local linear estimate for spatial regression. *Statistics & Risk Modeling,* **29**, 189–214 (2012).

[30] Dabo-Niang, S. Sur l'estimation fonctionnelle en dimension infinie: application aux diffusions. thesis, Paris 6, France, (2002).

[31] Dabo-Niang, S.and Rhomari, N. Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *Comptes Rendus Mathematique,* **336**, 75–80 (2003).

[32] Dabo-Niang, S. Kernel density estimator in an infinite-dimensional space with a rate of convergence in the case of diffusion process. *Applied mathematics letters,* **17**, 381–386 (2004).

[33] Dabo-Niang, S., Ferraty, F. and Vieu, P. Mode estimation for functional random variable and its application for curves classification. *Far East Journal of Theoretical Statistics,* **18**, 93–119 (2006).

[34] Dabo-Niang, S. and Laksaci, A. Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle. *Comptes Rendus Mathematique,* **344**, 49–52 (2007).

[35] Davidson, J. *Stochastic limit theory: An introduction for econometricians.* OUP Oxford. (1994).

[36] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. *A fast functional locally modeled conditional density and mode for functional time-series.* Recent Advances in Functional Data Analysis and Related Topics, Contributions to Statistics, Pages 85–90, DOI: 10.1007/978-3-7908-2736-1_13, Physica-Verlag/Springer (2011).

[37] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. Functional data: local linear estimation of the conditional density and its application. *Statistics*, **47**, 26–44 (2013).

[38] Demongeot, J., Laksaci, A., Rachdi, M., and Rahmani, S. On the local linear modelization of the conditional distribution for functional data. *Sankhya A*, **76**, 328–355 (2014).

[39] Demongeot, J., Naceri, A., Laksaci, A., and Rachdi, M. Local linear regression modelization when all variables are curves. *Statistics & Probability Letters*, **121**, 37–44 (2017).

[40] Deville, J. C. Méthodes statistiques et numériques de l'analyse harmonique. *Ann. Insee,*, **15**, 3–101 (1974).

[41] Didi, S. and Louani, D. Asymptotic results for the regression function estimate on continuous time stationary and ergodic data. *Statistics and Risk Modeling*, **31**, 129–150 (2014).

[42] Ezzahrioui, M. and Ould-Saïd, E. Asymptotic normality of a nonparametric estimator of the conditional quantile in the normed space. *Far East J. Theoretical Statist*, **25**, 15–38 (2008a).

[43] Ezzahrioui, M. and Ould-Saïd, E. Asymptotic results of a nonparametric conditional quantile estimator for functional time series. *Communications in Statistics—Theory and Methods*, **37**, 2735–2759 (2008b).

[44] Ezzahrioui, M. and Ould-Saïd, E. Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *Journal of Nonparametric Statistics*, **20**, 3–18 (2008).

[45] Fan, J. Design-adaptive nonparametric regression. *J. Amer. Statist. Assoc.*, **87**, 998–1004 (1992).

[46] Fan, J., Gijbels, I. *Local Polynomial Modelling and its Applications*. London, Chapman and Hall. (1996).

[47] Fan, J. and Yao, Q. *Nolinear Time Series: Nonparametric and Parametric Methods.* Springer-Verlag, New York.(2003).

[48] Ferraty, F. and Vieu, P. Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics* **330**, 139–142 (2000).

[49] Ferraty, F. and Vieu, P. The functional nonparametric model and application to spectrometric data. *Computational Statistics* **17**, 545–564 (2002).

[50] Ferraty, F. and Vieu, P. Curves discrimination: a nonparametric functional approach. *Computational Statistics & Data Analysis* **44**, 161–173 (2003).

[51] Ferraty, F. and Vieu, P. Nonparametric models for functional data, with application in regression, time series prediction and curve discrimination. *Nonparametric Statistics* **16**, :111-125 (2004).

[52] Ferraty, F., Laksaci, A. and Vieu, P. Functional time series prediction via conditional mode estimation. *Comptes Rendus Mathematique* **340**, 389–392 (2005).

[53] Ferraty, F., Rabhi, A. and Vieu, P. Conditional quantiles for dependent functional data with application to the climatic El Niño phenomenon. *Mathematical Reviews (MathSciNet): MR2208895* **67**, 378–398 (2005).

[54] Ferraty, F. and Vieu, P. Nonparametric functional data analysis theory and practice. Springer-Verlag (2006).

[55] Ferraty, F., Laksaci, A. and Vieu, P. Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statistical Inference for Stochastic Processes.*,**9** , 47–76 (2006).

[56] Ferraty, F., Rahbi, A. and Vieu, P. Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle. *Statistical Inference for Stochastic Processes.*,**53** , 1–18 (2008).

[57] Ferraty, F. and Romain, Y. The Oxford Handbook of Fuctional Data Analysis, *Oxford University Press* (2011).

[58] Ferraty, F., Laksaci, A., Tadj. A. and Vieu, P. Rate of uniform consistency for nonparametric estimates with functional variables. *J. of Statistical Planning and Inference.*, **140**, 335–352 (2010).

[59] Gao, F. and Li, W. Small ball probabilities for the Slepian Gaussian fields. *Transactions of the American Mathematical Society.*, **359**, 1339–1350 (2007).

[60] Geenens, G. Curse of dimensionality and related issues in nonparametric functional regression. Statistics Surveys, **5**, 30-43. (2011).

[61] Gheriballah, A., Laksaci, A. and Sekkal, S. Nonparametric M-regression for functional ergodic data. *Statistics and Probability Letters.*, **83**, p. 902–908 (2013).

[62] Jolliffe, I. Principal component analysis: Wiley Online Library. *Kaiser, HF (1960). The application of electronic computers to factor analysis.*, (2002).

[63] Györfi, L., Härdle, W., Sarda, P., & Vieu, P. Nonparametric curve estimation from time series (Vol. 60). Springer. (2013).

[64] Horváth, L., & Kokoszka, P. Inference for functional data with applications (Vol. 200). Springer Science & Business Media. (2012).

[65] Kaplan, EL., and Meier, P. Nonparametric estimation from incomplete observations. *Journal of the American statistical association* **53**, 457–481 (1958).

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

[67] Laïb, N. and Louani D. Nonparametric kernel regression estimation for functional stationary ergodic data: asymptotic properties. *J. of Multivariate analysis*, **101**, 2266–2281(2010).

[68] Laïb, N. and Louani D. Rates of strong consistencies of the regression function estimator for functional stationary ergodic data. *J. Stat. Plann. and Inf.*, **141**, p. 359–372 (2011).

[69] Laksaci, A. Convergence en moyenne quadratique de l'estimateur à noyau de la densité conditionnelle avec variable explicative fonctionnelle. *Annales de l'ISUP*, **51**, 69–80 (2007).

[70] Laksaci, A. and Mechab, B Estimation non paramétrique de la fonction de hasard avec variable explicative fonctionnelle: cas des données spatiales. *Revue Roumaine de Mathématiques Pures et Appliquées*, **55**, 35–51 (2010).

[71] Li, W. V and Shao, Q.M. Gaussian processes: inequalities, small ball probabilities and applications. *Handbook of Statistics.*, **19**, 533–597 (2001).

[72] Lifshits, M. A., Linde, W. and Shi, Z. Small Deviations of Riemann-Liouville Processes In lq Spaces With Respect To Fractal Measures. *Proceedings of the London Mathematical Society.*, **92**, 224–250, (2006).

[73] Lifshits, M. A., and Simon, T. Small deviations for fractional stable processes. *Annales de l'IHP Probabilités et statistiques.*, **41**, 725–752, (2005).

[74] Mahiddine, A and Bouchentouf, A. A and Rabhi, A. Nonparametric estimation of some characteristics of the conditional distribution in single functional index model. *Malaya Journal of Matematik (MJM)*, **2**, 392–410 (2014).

[75] Mas, A. (2012). Lower bound in regression for functional data by representation of small ball probabilities. Electronic Journal of Statistics, 6, 1745-1778.

[76] Masry, E., Nonparametric regression estimation for dependent functional data: asymptotic normality. *Stochastic Processes and their Applications.*, **115**, 155–177 (2005).

[77] Mayer-Wolf, E and Zeitouni, O. The probability of small Gaussian ellipsoids and associated conditional moments. *The Annals of Probability.*, 14–24 (1993).

[78] Müller HG Functional modeling of longitudinal data. In Longitudinal data analysis (pp. 237-266). Chapman and Hall/CRC, New York. (2008).

[79] Nadaraya, E., On estimating regression. *Theory of Probability & Its Applications.*, **9**, 141–142 (1964).

[80] Parzen, E. On estimation of a probability density function and mode, *The annals of mathematical statistics* **33**, 1065–1076 (1962).

[81] Quintela-del-Rio, A. On non-parametric techniques for area-characteristic seismic hazard parameters, *Geophysical Journal International* **180**, 339–346 (2010).

[82] Rabhi, A., Hammou, Y. and Djebbouri, T. Nonparametric estimation of the maximum of conditional hazard function under dependence conditions for functional data. *Afrika Statistika,* **10**, 777–794 (2015).

[83] Rabhi, A., Soltani, S. and Traore, A. Conditional risk estimate for functional data under strong mixing conditions. *Journal of Statistical Theory and Applications,* **14**, 301–323 (2015).

[84] Ramsay, J. O and Silverman, B.W. Functional Data Analysis Springer-Verlag, *New York,* (1997).

[85] Ramsay, J., Silverman, B.W. Applied functional data analysis; Methods and case studies. *Springer-Verlag, New York* (2002).

[86] Rosenblatt, M. Remarks on some nonparametric estimates of a density function, *The Annals of Mathematical Statistics,* 832–837 (1956).

[87] Ramsay, J. O. and Silverman, B. W. Functional Data Analysis. 2nd edition Springer-Verlag, New York. (2005).

[88] Rice, J. and Rosenblatt, M. Estimation of the log survivor function and hazard function. Sankhya Ser. A 38, 60–78. (1976).

[89] Shmileva, E. Small ball probabilities for jump Lévy processes from the Wiener domain of attraction. *Statistics & probability letters.*, **76**, 1873–1881 (2006).

[90] Stone, C. J. Optimal global rates of convergence for nonparametric regression. *The annals of statistics.*, 1040–1053 (1982).

[91] Watson, G.S. Smooth regression analysis. Sankhya A, 26, pp. 359–372 (1964).

[92] Xiong, X., Zhou, P,. and Ailian, C. Asymptotic normality of the local linear estimation of the conditional density for functional time-series data. *Communications in Statistics-Theory and Methods*, **47**, 3418–3440 (2018).

[93] Xiong, X,. and Lin, Z. Nonparametric M-estimation for functional stationary ergodic data. *J. Acta Mathematicae Applicatae Sinica, English Series.*, **35**, 491–512 (2019).

[94] Zhou, Z. and Lin, Z.-Y. Asymptotic normality of locally modelled regression estimator for functional data. *J. Nonparametr. Statist.*, **28**, 116–131 (2016).

CHAPTER 2

LOCAL LINEAR MODELLING OF THE CONDITIONAL DISTRIBUTION FUNCTION FOR FUNCTIONAL ERGODIC DATA

This chapter is the subject of a paper accepted in Journal of MATHEMATICAL MODELLING AND ANALYSIS.

Local linear modelling of the conditional distribution function for functional ergodic data.

Somia Ayad [1], **Ali Laksaci** [2], **Saâdia Rahmani** [1] and **Rachida Rouane** [1].

[1] Laboratory of Stochastic Models, Statistics and Applications,University of Saida–Dr. Moulay Tahar, 20000, Algeria. E-mail: somia.ayad@univ-saida.dz. E-mail: saadia.rahmani@gmail.com. E-mail: rouane09@yahoo.fr

[2] Department of Mathematics, College of Science, King Khalid University, Abha, 61413, Saudi Arabia. E-mail:alilak@yahoo.fr.

Abstract : The focus of functional data analysis has been mostly on independent functional observations. It is therefore hoped that the present contribution will provide an informative account of a useful approach that merges the ideas of the ergodic theory and the functional data analysis by using the local linear approach. More precisely, we aim, in this paper, to estimate the conditional distribution function (CDF) of a scalar response variable given a random variable taking values in a semi-metric space. Under the ergodicity assumption, we study the uniform almost complete convergence (with a rate), as well as the asymptotic normality of the constructed estimator. The relevance of the proposed estimator is verified through a simulation study. **Keywords :** Ergodic data, functional data, local linear estimator, conditional distribution function, nonparametric estimation, Asymptotic properties.

Mathematics Subject Classification: 62G05, 62G08, 62G20, 62G35.

Secondary: 62H12

2.1 Introduction and motivations

Over the almost last two decades, functional data analysis (FDA) has established itself as a dynamic and important field of statistical research. It has become very broad, with many specialized directions of research. This statistic area offers effective new tools and has stimulated novel methodological developments. With the availability of large amounts of data as well as the development of the computer instruments, (FDA) swept across various fields of applied sciences (for instance biometrics, geophysics and econometrics). There are many nonparametrc problems for functional data which have attracted a growing interest; one may refer to the famous work of Ferraty and Romain [17], the monograph of Ferraty and Vieu [16] and the pioneer book of Kokoszka and Reimherr [21] as well as the references therein.

Despite the simplicity of the classical kernel estimate and its availability in many statistical software packages, like R and Matlab make that it easy to understand and implement, its simplicity leads to some weaknesses; the most obvious of which is boundary bias effect. Moreover, it is well known that among the smoothing procedures, the local polynomial approach has various advantages over the classical kernel method. In particular, this method has better properties concerning the bias terms on the other one (cf. Fan and Gijbels [14] for an extensive discussion).

In the context of the finite dimensional space, the local linear method is well established, frequently used and it has been the subject of considerable studies, and key references on this topic are Chen et *al.* [7], Fan and Yao [15] and references therein. However, before the pioneer work of Barrientos-Marin et *al.* [3], only few results are available for the local linear modeling in the functional statistics setup. Indeed, the first results, in this direction, were established by Baìllo and Grané [2]. This paper focuses on the local linear estimation of the regression operator when the explanatory variable takes values in a Hilbert space. The general case, where the regressors do not belong to a Hilbert space but to a semi-metric space, has been considered not only by Barrientos-Marin et *al.* [3] but also by El Methni and Rachdi [12], Demongeot et *al.* [11] and Laksaci et *al.*[24].

Recently, the paper of Bouanani et *al.* [4] has completed the theoretical advances presented by Laksaci et *al.* [24] by establishing the asymptotic normality of the local linear estimates for several conditional models.

Weak dependencies have been considered by many authors in the context of both discrete and continuous-time processes. We consider, in this paper, the ergodic framework which is more general than the weak dependencies. More precisely, we examine the local linear estimator's properties of the (CDF) when the data of our constructed estimator are ergodic.

In the literature, several real examples have been studied in order to emphasize the usefulness of such dependency. For instance, the ergodicity assumption models several phenomena in physics like the thermodynamic properties of gases, atoms or plasma. In a more general way, the ergodic theory becomes crucial because there are many phenomena which are neither independent nor α-mixing either.

In the past three decades, the study of statistical models adapted to such kind of dependency has been impressively large but mostly restricted to the standard multivariate situation where both the response and the explanatory variables are real or multivariate (see, Delecroix and Rosa [8] and Laïb and Ould-Saïd [22]). However, there are very few advances in this direction when the regressor is functional. One may refer to the work of Laib and Louani [23]. The authors studied under ergodicity assumption the asymptotic properties of an estimator of the regression operator. Related works can be found in the paper of Laib and Louani [23] when the data are completely observed and Chaouch et *al.* [6] for the right censored ones.

As it is mentioned above, the main aim of this paper is to construct and study under general conditions, the uniform almost complete convergence rate as well as the asymptotic normality of a local linear estimator of the CDF. For this purpose, it is assumed that the covariate takes its values in an infinite dimensional space and the data are sampled from a stationary ergodic process. Recall that uniform consistency results have been successfully used in the standard nonparametric setting (see for instance, Ferraty et *al.* [18], Ling et *al.* [25] and Kara-Zaitri et *al.* [20]). Each of these papers considers the case of the local constant method. However, in this contribution, we consider a more efficient estimate of the CDF by the local linear method.

To make this paper as much self-contained as possible, the nonparametric model and its associated local linear estimator are constructed in Section 2.2. In the same section, we report some notations required for this contribution. The assumptions, under which the main results are valid, are stated and discussed in Section 2.3. Then, we derive theoretical results by giving a deep asymptotic study of the behaviour of the estimate, including the almost complete convergence of the CDF uniformly in the functional argument x as well as the asymptotic gaussian distribution. The relevance of the proposed estimator is verified through a simulation study in Section 2.4. Finally, the paper is ended with a technical appendix.

2.2 Local linear estimator construction

Let $(X_i, Y_i)_{i=1,\ldots,n}$ be a strictly stationary (in an ergodic sense) process of $\mathcal{F} \times \mathbb{R}$-valued random elements, where $\mathcal{F}$ is a semi-metric space with semi-metric d. We assume that there exists a regular version of the conditional distribution of Y given X, which is absolutely continuous with respect to the *Lebesgue* measure on $\mathbb{R}$.

Interest centers on the conditional behavior of Y given X. To this end it is convenient to consider

$$F^x(y) = \mathbb{P}(Y_i \le y | X_i = x),$$

Since the local linear approach requires a smoothing assumption that allows us to approximate locally the nonparametric CDF, we estimate the function $F^x(.)$ by assuming that it is smoothed enough to be locally approximated by a linear function. For this aim, we introduce two locating functions δ and ρ (see Barrientos et *al.* [3] for more discussion on these bilinear continuous operators) and we consider a subset $\mathcal{C}_\mathcal{F}$ of $\mathcal{F}$ such that for $x_k \in \mathcal{C}_\mathcal{F}, \mathcal{C}_\mathcal{F} \subset \bigcup_{k=1}^{d_n} B(x_k, r_n)$ where r_n (resp. d_n) is a sequence of positive real (resp. integer) numbers and $B(x_k, r_n) = \{x'_k \in \mathcal{F} / |\delta(x'_k, x_k)| < r_n\}$. Such approximation can be expressed, for any $z \in \mathcal{C}_\mathcal{F}$ in the neighborhood of x by:

$$F^z(y) = a_0 + a_1 \rho(z, x) + o(\rho(z, x)). \tag{2.1}$$

We assume that the underlying process (X_i, Y_i) is functional stationary ergodic. Then the estimator $\widehat{F}^x$ of F^x can be seen as the solution of the following minimization problem

$$\min_{(a_0,a_1\in\mathbb{R}^2}\sum_{i=1}^{n}\left(J\left(\frac{y-Y_i}{h_J}\right)-\alpha-\beta\rho(X_i,x)\right)^2 K\left(\frac{\delta(x,X_i)}{h_K}\right), \tag{2.2}$$

where the bi-functional $\delta(.,.)$ is lied with the topological structure of the functional space $\mathcal{F}$, that means $|\delta(x,z)| = d(x,z)$, whereas, ρ controls the local sharp of the model (see formula (2.1)). K is a kernel, J is a distribution function and $h_K = h_{K,n}$ (respectively $h_J = h_{J,n}$) is a sequence of positive real numbers. More precisely, the functional local linear estimator $\widehat{F}^x(y)$ of $F^x(y)$ is then $\widehat{a_0}$ which is the first component of the pair (a_0, a_1) solution of the minimization problem (2.2). However, if the bi-functional operator ρ is such that $\rho(z,z) = 0, \forall z \in \mathcal{F}$, then the quantity $\widehat{F}^x\,(y)$ is explicitly defined by:

$$\widehat{F}^x(y) = \frac{\sum_{j=1}^{n}\Gamma_j K\left(\frac{\delta(x,X_j)}{h_K}\right)J\left(\frac{y-Y_j}{h_J}\right)}{\sum_{j=1}^{n}\Gamma_j(x)K\left(\frac{\delta(x,X_j)}{h_K}\right)}, \tag{2.3}$$

with

$$\Gamma_j = \sum_{i=1}^{n}\rho_i^2(x)K_i(x) - \left(\sum_{i=1}^{n}\rho_i(x)K_i(x)\right)\rho_j(x),$$

$$\text{where}\quad \rho_i(x) = \rho(X_i,x),\ \text{and}\ K_i(x) = K\left(\frac{\delta(x,X_i)}{h_K}\right).$$

2.3 Main results

2.3.1 Uniform almost complete convergence

Assumptions and notations

First we need to introduce some further notations. For $i = 1,\dots,n$, let $\mathfrak{F}_i$ and $\mathcal{G}_i$ denote, respectively, the σ-field generated by $((X_1,Y_1),\dots,(X_i,Y_i))$, and $((X_1,Y_1),\dots,(X_i,Y_i),X_{i+1})$. For any fixed y in $\mathbb{R}$, $\mathcal{N}_y$ denotes a fixed neighborhood of y. In the sequel, we will also need to define the small ball probability function by $\phi_x(h_1,h_2) = \mathbb{P}(h_2 \le \delta(X,x) \le h_1)$ and we will denote by C and C' some strictly positive generic constants. Finally, with some abuse of notations, we write $J_j(y)$ for $J\left(\frac{y-Y_j}{h_J}\right)$ and we shall write ϕ instead of ϕ_x.
Our consistency results are summarized in Theorem 2.3.1 and relie on the following seven assumptions:

➢ **Structural hypotheses:**

On the ergodic functional variables:

We suppose that the strictly stationary ergodic process $(X_i, Y_i)_{i\in\mathbb{N}^*}$ satisfies: For all $r > 0$

(H1) $\forall x \in \mathcal{C}_{\mathcal{F}}, 0 < C\phi(r) \leq \mathbb{P}(X \in B(x,r)) \leq C'\phi(r)$. Furthermore $\exists\, \eta_0 > 0, \forall \eta < \eta_0, \phi'(\eta) < C$, where ϕ' denotes the first order derivative of ϕ.

(H2) For all $i = 1, \dots, n$, there exist a determinist function ϕ_i such that:

i) $0 < C\phi_i(r) < \mathbb{P}(X_i \in B(x,r)|\mathfrak{F}_{i-1}) \leq C'\phi_i(r)$,

ii) $\frac{1}{n\phi(r)} \sum_{i=1}^{n} \phi_i(r) \longrightarrow 1\, a.co.$

➢ **Technical and regularity conditions:**

(H3) **On the regularity of the model:**

There exist some positive constants b_1 and b_2 such that:

$\forall (x_1, x_2) \in \mathcal{C}_{\mathcal{F}} \times B(x_1, h_K)$ and $\forall (y_1, y_2) \in \mathcal{N}_y \times \mathcal{N}_y$:

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C\left(|\delta(x_1,x_2)|^{b_1} + |y_1 - y_2|^{b_2}\right),$$

(H4) **On the bi-functional operators ρ and δ:**

(i) $\forall z \in \mathcal{F}$, $C|\delta\,(x,z)\,| \leq |\rho\,(x,z)\,| \leq C'|\delta\,(x,z)\,|$,

(ii) $\forall (x_1, x_2) \in \mathcal{C}_{\mathcal{F}} \times \mathcal{C}_{\mathcal{F}}$,
$|\rho\,(x_1, x)-\rho\,(x_2, x)\,| \leq C'|\delta(x_1, x_2)|$.

(H5) **On the kernel K and the distribution function J:**

(i) K is a nonnegative bounded and Lipschitz kernel on its support $[-1; 1]$.

(ii) The kernel J is a differentiable function such that:

$$\int_{\mathbb{R}} |t|^{b_2} J^{(1)}(t)dt < \infty.$$

(iii) $\mathbb{E}(J_j(y)|\mathcal{G}_{j-1}) = \mathbb{E}(J_j(y)|X_j)$

(H6) Taking $r_n = O\left(\frac{\log n}{n}\right)$, the sequence d_n satisfies:

$$\frac{(\log n)^2}{n\phi(h_K)} < \log d_n < \frac{n\phi(h_K)}{\log n} \text{ and } \sum_{n=1}^{\infty} d_n^{(1-\varrho)} < \infty \quad \text{for some} \quad \varrho > 1.$$

(H7) **On the bandwidth h_K with respect to ρ and ϕ :**

i) There is a positive integer n_0, such that, $\forall n > n_0$:

$$-\frac{1}{\phi(h_K)}\int_{-1}^{1}\phi(zh_K,h_K)\frac{d}{dz}\left(z^2K(z)\right)dz > C,$$

ii) $\lim_{n\to\infty} h_K = 0, \lim_{n\to\infty} h_H = 0,$ and $\lim_{n\to\infty}\frac{\log n}{n\phi(h_K)} = 0$

iii) $h_K\int_{B(x,h_K)}\rho(u,x)\,dP(u) = o\left(\int_{B(x,h_K)}\rho^2(u,x)\,dP(u)\right),$
where $dP(x)$ is the cumulative distribution of X.

Comments on the hypotheses

(H1) involves the small ball techniques. It is clearly unrestrictive, since it is the same as that frequently used in the FDA context.
We precise that the ergodic nature of the data is exploited by (H2) which is a very mild condition in comparison of that imposed by Laib and Louani [23].
Concerning (H3), this condition is necessary to evaluate the bias term in our asymptotic result. The constants b_1 and b_2 control the model's regularity and the only condition imposed is their positivity. In other words, this assumption guarantees slower variance of the operator compared to X. More the parameters b_1 and b_2 are small, more the curves of the operator's evolution as a function of X are smooth, and less the estimate obtained is biased.
Condition (H4)(i) is unrestrictive condition and it is verified if $\rho(.,.) = \delta(.,.)$ (In this special case, (H7) (iii) means that the local expectation of ρ is small enough with respect to its moment of second order); or if $\lim_{\delta(z,x)\to 0}\left|\frac{\rho(z,x)}{\delta(z,x)} - 1\right| = 0$. Indeed, $\forall z \in B(x,h_K)$, we have:

$$\left|\frac{\rho(z,x)-\delta(z,x)}{h_K}\right| \leq \left|\frac{\rho(z,x)}{\delta(z,x)} - 1\right| \longrightarrow 0 \text{ as } \delta(z,x)\to 0.$$

The Lipschitz condition (H4) (ii) on the locating function ρ is the same used by Barrientos-Marin et *al.* [3] and it is typical in the context of local polynomial smoothing.
(H5)(i) could be replaced by another assumption such as the boundness of the kernel K. The slightly stronger assumption (H5) (i) just makes the proof of uniform convergence simpler. (H5)(ii) and (iii) are technical conditions imposed for brevity of proofs.
In (H6), the covering hypothesis on the subset $\mathcal{C}_{\mathcal{F}}$ is linked to the topological structure of our functional space $\mathcal{F}$. It controls Kolmogorov's entropy of the set $\mathcal{C}_{\mathcal{F}}$. Such consideration has been discussed and commented by Ferraty et *al.* [18]. The authors give several examples for which this condition is satisfied.
The assumption (H7) (i) precise the behaviour of the smoothing parameter h_K in relation with the small ball probabilities and the kernel function K. The local behaviour of ρ which models the local shape of our model is controlled by (H7)(iii).

We are now ready to state our first result which is the uniform almost complete convergence of the estimator $\widehat{F}^x(y)$ on the subset $\mathcal{C}_{\mathcal{F}}$.

Theorem 2.3.1. *As soon as assumptions (H1)–(H7) are fulfilled, we have*

$$\sup_{x\in\mathcal{C}_{\mathcal{F}}} |\widehat{F}^x(y) - F^x(y)| = O\left(h_K^{b_1} + h_J^{b_2}\right) + O\left(\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right), \quad a.co.$$

Before starting the proof of this Theorem, we introduce the following further notations:

$$\widehat{F}_N^x(y) := \frac{1}{n\mathbb{E}(\Gamma_1 K_1)}\sum_{j=1}^{n}\Gamma_j K_j(x) J_j, \quad \bar{F}_N^x(y) := \frac{1}{n\mathbb{E}(\Gamma_1 K_1)}\sum_{j=1}^{n}\mathbb{E}(\Gamma_j K_j(x) J_j|\mathfrak{F}_{j-1}),$$

$$\widehat{F}_D(x) := \frac{1}{n\mathbb{E}(\Gamma_1 K_1)}\sum_{j=1}^{n}\Gamma_j K_j(x) \qquad \text{and } \bar{F}_D(x) := \frac{1}{n\mathbb{E}(\Gamma_1 K_1)}\sum_{j=1}^{n}\mathbb{E}(\Gamma_j K_j(x)|\mathfrak{F}_{j-1}).$$

Then, the proof of Theorem 2.3.1 is based on the following decomposition:

$$\widehat{F}^x(y) - F^x(y) = B_n(x,y) + \frac{1}{\widehat{F}_D(x)}\left[(B_n(x,y) + F^x(y))A_n(x,y) + R_n(x,y)\right], \qquad (2.4)$$

where

$$B_n(x,y) = \frac{\bar{F}_N^x(y)}{\bar{F}_D(x)} - F^x(y),$$

$$A_n(x,y) = \bar{F}_D(x) - \widehat{F}_D(x)$$

and

$$R_n(x,y) = \widehat{F}_N^x(y) - \bar{F}_N^x(y).$$

As immediate consequence of the decomposition (2.4), we need to prove the following lemmas:

Lemma 2.3.1. *Under the assumptions (H1)–(H5) and (H7), we have*

$$\sup_{x\in\mathcal{C}_{\mathcal{F}}} |B_n(x,y)| = O\left(h_K^{b_1} + h_J^{b_2}\right).$$

Lemma 2.3.2. *Under the hypotheses of Theorem 2.3.1, we obtain*

$$\sup_{x\in\mathcal{C}_{\mathcal{F}}} |R_n(x,y)| = O_{a.co}\left(\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right).$$

Lemma 2.3.3. *Under the assumptions (H1)–(H5) (i), (H6) and (H7), we have*

i) $$\sup_{x\in\mathcal{C}_{\mathcal{F}}} |A_n(x,y)| = O_{a.co}\left(\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right).$$

ii) $\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{x \in \mathcal{C}_{\mathcal{F}}} \widehat{F}_D(x) < \frac{1}{2}\right) < \infty.$

Thus, the proof of our main result is based on the previous lemmas combined with Lemma 5 of Ayad et *al.* [1] and the technical lemma 1 of Laib and Louani [23].

2.3.2 Asymptotic normality

Let us first focus on the supplementary assumptions we need to derive the asymptotic normality of our estimator.

(B1) The hypothesis (H1) holds and there exists a function $\Psi(\cdot)$ such that:

$$\forall t \in [-1,1], \lim_{h_K \to 0} \frac{\phi(-h_K, th_K)}{\phi(h_K)} = \Psi(t).$$

(B2) The hypothesis (H3) holds and for all $(x_1, x_2, y_1, y_2) \in \mathcal{C}_{\mathcal{F}} \times \mathcal{C}_{\mathcal{F}} \times \mathcal{N}_y \times \mathcal{N}_y$:

$$\begin{cases} F : \mathcal{F} \times \mathbb{R} \longrightarrow \mathbb{R}, \lim_{|\delta(x_1,x_2)| \to 0} F^{x_1}(y) = F^{x_2}(y), \\ \quad \text{and} \\ \lim_{|y_1 - y_2| \to 0} F^x(y_1) = F^x(y_2). \end{cases}$$

(B3) The hypothesis (H4) holds and

$$\sup_{u \in B(x,r)} |\rho(u,x) - \delta(x,u)| = o(r)$$

.

(B4) The hypothesis (H5) holds and the first derivative K' of the kernel K satisfies:

$$K^2(1) - \int_{-1}^{1} (K^2(u))' \Psi(u) du > 0.$$

(B5) The hypothesis (H7) holds and $\lim_{n \to \infty} (n-1)^k h_K^l \phi(h_K) = 0, \quad \text{for}\, k = 1,2\, \text{and}\, l = 4,5.$

In addition, we need to introduce the quantities M_c and $N(a,b)$ which will appear in the computation of $\mathbb{E}\left(K_j^c | \mathfrak{F}_{j-1}\right)$.

$$M_c = K^c(1) - \int_{-1}^{1} (K^c(u))' \Psi(u) du \quad \text{where } c = 1, 2,$$

and for all $a > 0$ and $b = (2,4)$, $\quad N(a,b) = K^a(1) - \int_{-1}^{1} (u^b K^a(u))' \Psi(u) du.$

Theorem 2.3.2. *Under assumptions (B1)–(B5), (H1) and (H6) and if the smoothing parameters h_K and h_J satisfy $\sqrt{n\phi(h_K)}\left(h_K^{b_1}+h_J^{b_2}\right)\to 0 \quad as\ n\to\infty$, We have*

$$\sqrt{n\phi(h_K)}\left(\widehat{F}^x(y)-F^x(y)\right)\xrightarrow{\mathcal{D}} N(0,V_{JK}(x,y)),$$

where $V_{JK}(x,y)=\dfrac{M_2}{M_1^2}F^x(y)(1-F^x(y))$, *and* $\xrightarrow{\mathcal{D}}$ *means the convergence in distribution.*

The first step of the proof consists in rewriting the decomposition (2.4) in the following way:

$$\widehat{F}^x(y)-F^x(y)=B_n(x,y)+\frac{C_n(x,y)+Q_n(x,y)}{\widehat{F}_D(x)}, \tag{2.5}$$

where

$$C_n(x,y)=B_n(x,y)A_n(x,y), \quad \text{and} \quad Q_n(x,y)=R_n(x,y)+F^x(y)A_n(x,y).$$

Then, to state asymptotic normality, we remark that the hypothesis (B2) ensures the asymptotic negligence of $B_n(x,y)$. Moreover, according to Lemma 2.3.3 (i), C_n converges almost completely to zero when n goes to infinity. Consequently, the proof of Theorem 2.3.2 can be deduced from the following lemmas for which the proofs are relegated to the Appendix.

Lemma 2.3.4. *Under assumptions of Theorem 2.3.2, we have*

$$\sqrt{\frac{n\phi(h_K)}{V_{JK}(x,y)}}Q_n(x,y)\xrightarrow{D} N(0,1).$$

Lemma 2.3.5. *Under assumptions (H1)–(H5) (i), (H6) and (H7), we have*

$$\widehat{F}_D(x)-1=o_p(1)$$

2.4 On simulated data

We now conduct a simulation study in which the finite sample performance of the local linear estimator given in formula (2.3) is compared to the following local constant estimator:

$$\widetilde{F}^x(y)=\frac{\sum_{j=1}^{n}K_j(x)J_j(y)}{\sum_{j=1}^{n}K_j(x)}. \tag{2.6}$$

Let us use the following regression model:

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

where the random variable ε is normally distributed with a variance equal to 0.075. The explanatory functional variables are constructed by:

$$X_i(t) = 2w_i t^2 + \frac{1}{2}\cos(\pi z_i t) \qquad i = 1, \ldots 200,\ t \in [0,1],$$

where w_i are n independent real random variables uniformly distributed over $[0,1]$ and $z_i = \frac{1}{3}z_{i-1} + \zeta_i$. Here ζ_i are i.i.d. realizations of $N(0,1)$ and are independent from w_i and z_i, which is generated independently by $z_0 \sim N(0,1)$. All the curves X_i's were discretized on the same grid generated from 200 equispaced measurements in (0,1) and are plotted in Figure 2.1.

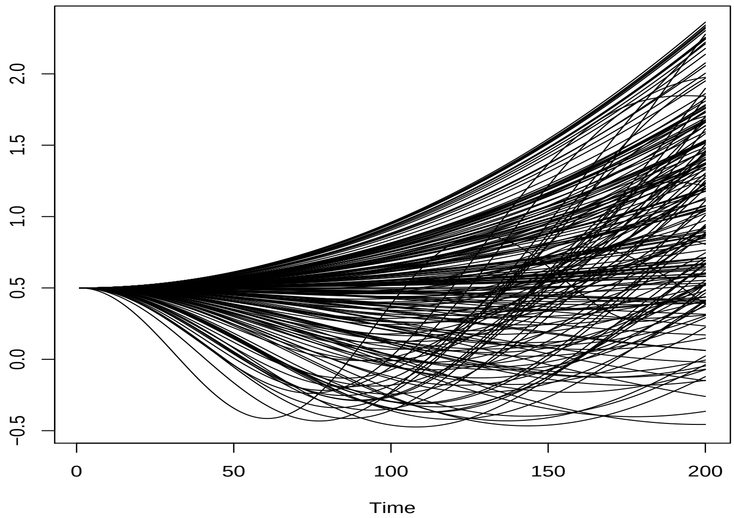

Figure 2.1: The curves X_i.

On the other hand, the building of the scalar response Y is obtained by considering the following regression operator:

$$\left(\int_0^1 X'(t)dt\right)^2$$

Recall that, the conditional distribution of Y given $X = x$ corresponding to this model is explicitly given by the law of ε_i shifted by $\left(\int_0^1 X'(t)dt\right)^2$.

When dealing with smooth curves such as those introduced herein, it is necessary to measure the proximity by means of a semi-metric based on the L_2 norm of some derivative of the curves. The smoothing parameters h_K and h_J are selected through automated cross validation, choos-

ing a value that minimizes the average error on the withheld data.
The behavior of our estimator is linked to the good choice of the functions δ and ρ. Because of the smoothness of the curves we take

$$\rho(x,x') = \int_0^1 \theta(t)(x^{(1)}(t) - x'^{(1)}(t))dt$$

and $$\delta(x,x') = \left(\int_0^1 (x^{(1)}(t) - x'^{(1)}(t))^2 dt\right)^{1/2}$$
with the functional index θ is selected among the eigenfunction of the empirical covariance operator $\frac{1}{n}\sum_{i=1}^{n}(X_i^{(1)} - \overline{X^{(1)}})^t((X_i^{(1)} - \overline{X^{(1)}}))$ corresponding to the biggest eigenvalues, where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.
For both competitors, the kernel $K(u) = (1-u^2)1\!\!1_{[0,1]}$ is used and the distribution function J is defined by:

$$J(u) = \frac{3u}{4}\left(1 - \frac{u^2}{3}\right)1\!\!1_{[-1,1]} + \frac{1}{2}.$$

In this illustration, we have followed the following steps:

- *Step 1:* We generate m replications of $(X_i, Y_i)_{i=1,\ldots,n}$.
- *Step 2:* We estimate the conditional local linear distribution (respectively the conditional kernel distribution).
- *Step 3:* we compare these estimators to the Gaussian distribution.

In order to eliminate the zero weight, we have removed the negative weighing. The obtained results are plotting in Figure 2.2.
It is clear that the local linear estimator of the CDF operator convincingly outperforms the local

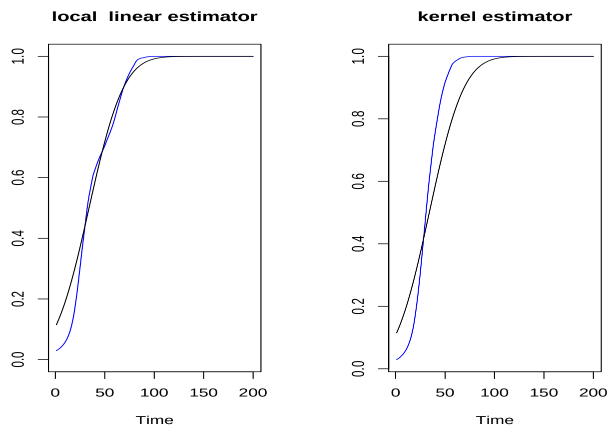

Figure 2.2: Comparison between the both estimators.

constant one.

2.5 Appendix

2.5.1 Preliminary technical lemmas

Firstly, we state the following lemmas which are needed to establish our asymptotic results.

Lemma. *A.1 Under assumptions (H1), (H2), (H4)(i), (H5) and (H7), we obtain*

i) $\sup_{x\in\mathcal{C}_\mathcal{F}} \bar{F}_D(x) = O(1)$.

ii) $\inf_{x\in\mathcal{C}_\mathcal{F}} \bar{F}_D(x) = O(1)$.

Lemma. *A.2 Under assumptions (B1), (H2), (B2)–(B5), we have*

(i) $h_K\mathbb{E}\left(\rho_j(x)K_j^a(x)|\mathfrak{F}_{j-1}\right) = o(h_K^2\phi_j(h_K))$ *for all* $a > 0$.

(ii) $\frac{1}{n\phi(h_K)}\sum_{j=1}^{n}\mathbb{E}\left(K_j^c(x)|\mathfrak{F}_{j-1}\right) = M_c + o(1)$ *for* $c = 1, 2$.

(iii) $\frac{1}{n\phi(h_K)}\sum_{j=1}^{n}\mathbb{E}\left(\Gamma_j^2K_j^2(x)|\mathfrak{F}_{j-1}\right) = (n-1)^2(N(1,2))^2h_K^4\phi^2(h_K)M_2 + o\left(h_K^4\phi^2(h_K)\right)$.

Proof of Lemma A.1

Before we start the proof of i), it is clear that by using Lemma A.1 of [3], we obtain

$$nCh_K^2\phi(h_K) \leq \mathbb{E}\left(\Gamma_1(x)K_1(x)\right) \leq nC'h_K^2\phi(h_K). \tag{2.7}$$

Then, by considering Lemma 5 of Ayad et *al.* [1] and (2.7) we get

$$\sup_{x\in\mathcal{C}_\mathcal{F}} \bar{F}_D(x) = O(1)\sup_{x\in\mathcal{C}_\mathcal{F}}\frac{1}{n\phi(h_K)}\sum_{j=1}^{n}\phi_j(h_K).$$

So, the claimed result i) of this lemma is a direct consequence of Assumption (H2)(ii).
The proof of (ii) is similar to that of (i), and is therefore omitted.

Proof of Lemma A.2

Firstly, the proof of (i) and (ii) are similar to the proof of (b) and (a) of Lemma A.1 in [27]. Secondly, in order to prove (iii), we use the definition of the conditional variance. Indeed,

$$\frac{1}{n\phi(h_K)}\sum_{j=1}^{n}\mathbb{E}\left(\Gamma_j^2K_j^2(x)|\mathfrak{F}_{j-1}\right) = \frac{1}{n\phi(h_K)}\sum_{j=1}^{n}\left(\mathrm{Var}\left(\Gamma_jK_j(x)|\mathfrak{F}_{j-1}\right) + \left(\mathbb{E}\left(\Gamma_jK_j(x)|\mathfrak{F}_{j-1}\right)\right)^2\right). \tag{2.8}$$

It remains to study each term of (2.8). For the first term on the right hand side of this equation, we have

$$\begin{aligned}
\mathrm{Var}\left(\Gamma_j K_j(x)|\mathfrak{F}_{j-1}\right) &= (n-1)\left(\mathrm{Var}\left(\rho_1^2(x)K_1(x)K_j(x)|\mathfrak{F}_{j-1}\right) + \mathrm{Var}\left(\rho_1(x)K_1(x)\rho_j(x)K_j(x)|\mathfrak{F}_{j-1}\right)\right)\\
&= (n-1)\Big(\underbrace{\mathbb{E}\left(\rho_1^4(x)K_1^2(x)\right)\mathbb{E}\left(K_j(x)^2|\mathfrak{F}_{j-1}\right)}_{T_1}\\
&\quad -\underbrace{\left(\mathbb{E}\left(\rho_1^2(x)K_1(x)\right)\mathbb{E}\left(K_j(x)|\mathfrak{F}_{j-1}\right)\right)^2}_{T_2}\\
&\quad +\underbrace{\mathbb{E}\left(\rho_1^2(x)K_1^2(x)\right)\mathbb{E}\left(\rho_j^2(x)K_j^2(x)|\mathfrak{F}_{j-1}\right)}_{T_3}\\
&\quad -\underbrace{\left(\mathbb{E}\left(\rho_1(x)K_1(x)\right)\mathbb{E}\left(\rho_j(x)K_j(x)|\mathfrak{F}_{j-1}\right)\right)^2}_{T_4}\Big).
\end{aligned}$$

Then, by using (i) of Lemma A.1 in [3] and (i) of Lemma 5 in [1], we find

$$\frac{n-1}{n\phi(h_K)}\sum_{j=1}^{n} T_i = O\left((n-1)h_K^4\phi(h_K)\right) \quad \text{for } i = 1,2,3,4.$$

It follows that

$$\frac{1}{n\phi(h_K)}\sum_{j=1}^{n}\mathrm{Var}\left(\Gamma_j K_j(x)|\mathfrak{F}_{j-1}\right) \longrightarrow 0, \qquad \text{as } n \to \infty.$$

On the other side, to complete the proof of (iii), we have to study the first term on the right hand side of (2.8). For that, we write:

$$\begin{aligned}
\frac{1}{n\phi(h_K)}\sum_{j=1}^{n}\left(\mathbb{E}\left(\Gamma_j K_j(x)|\mathfrak{F}_{j-1}\right)\right)^2 &= \frac{1}{n\phi(h_K)}\sum_{j=1}^{n}\Bigg(\mathbb{E}\Big(\sum_{i=1}^{n}\rho_i^2(x)K_i(x)K_j(x)\\
&\quad - \sum_{i=1}^{n}\rho_i(x)K_i(x)\rho_j(x)K_j(x)|\mathfrak{F}_{j-1}\Big)\Bigg)^2\\
&= \gamma_{n1} + \gamma_{n2} + \gamma_{n3},
\end{aligned}$$

where

$$\gamma_{n1} = \frac{(n-1)^2}{n\phi(h_K)}\left(\mathbb{E}\left(\rho_1^2(x)K_1(x)\right)\right)^2\sum_{j=1}^{n}\left(\mathbb{E}\left(K_j|\mathfrak{F}_{j-1}\right)\right)^2,$$

$$\gamma_{n2} = \frac{(n-1)^2}{n\phi(h_K)}\left(\mathbb{E}\left(\rho_1(x)K_1(x)\right)\right)^2\sum_{j=1}^{n}\left(\mathbb{E}\left(\rho_j(x)K_j|\mathfrak{F}_{j-1}\right)\right)^2,$$

and $\gamma_{n3} = -\frac{2(n-1)^2}{n\phi(h_K)}\mathbb{E}\left(\rho_1^2(x)K_1(x)\right)\mathbb{E}\left(\rho_1(x)K_1(x)\right)\sum_{j=1}^{n}\mathbb{E}\left(K_j|\mathfrak{F}_{j-1}\right)\mathbb{E}\left(\rho_j(x)K_j|\mathfrak{F}_{j-1}\right)$

Concerning the term γ_{n1}, by applying Jensen's Inequality, we have

$$\gamma_{n1} \leq \frac{(n-1)^2}{n\phi(h_K)}\left(\mathbb{E}\left(\rho_1^2(x)K_1(x)\right)\right)^2\sum_{j=1}^{n}\mathbb{E}\left(K_j^2|\mathfrak{F}_{j-1}\right),$$

then, we apply (c) of Lemma A.1 in [27] and (ii) of Lemma A.2 to obtain:

$$\gamma_{n1} = (n-1)^2 \left((N(1,2))^2 h_K^4 \phi^2(h_K) M_2 + o\left(h_K^4 \phi^2(h_K)\right)\right). \tag{2.9}$$

Concerning γ_{n2}, we apply (b) of Lemma A.1 in [27] and (i) of Lemma A.2 to get:

$$\gamma_{n2} = o\left((n-1)^2 h_K^4 \phi(h_K)\right). \tag{2.10}$$

For the last term, we apply (i) of Lemma A.1 in [3], (i) of Lemma 5 in [1] and (i) of Lemma A.2 to get:

$$\gamma_{n3} = o\left((n-1)^2 h_K^5 \phi(h_K)\right). \tag{2.11}$$

Combining (2.9), (2.10) and (2.11) permits to obtain the claimed result.

2.5.2 Proofs of Main results

Proof of Lemma 2.3.1

We start by writing

$$\sup_{x\in\mathcal{C}_\mathcal{F}} |B_n(x,y)| = \frac{\sup_{x\in\mathcal{C}_\mathcal{F}} |\widetilde{B}_n(x,y)|}{\inf_{x\in\mathcal{C}_\mathcal{F}} |\bar{F}_D(x)|},$$

where $\widetilde{B}_n(x,y) = \bar{F}_N^x(y) - F^x(y)\bar{F}_D(x)$.
First, observe that $\widetilde{B}_n(x,y)$ can be written as

$$\begin{aligned}
\widetilde{B}_n(x,y) &= \frac{1}{n\mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \left\{\mathbb{E}\left(\Gamma_j K_j(x) J_j | \mathfrak{F}_{j-1}\right) - F^x(y)\,\mathbb{E}\left(\Gamma_j K_j | \mathfrak{F}_{j-1}\right)\right\} \\
&= \frac{1}{n\mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \left\{\mathbb{E}\left(\Gamma_j K_j(x)\mathbb{E}\left(J_j|\mathcal{G}_{j-1}\right)|\mathfrak{F}_{j-1}\right) - F^x(y)\,\mathbb{E}\left(\Gamma_j K_j | \mathfrak{F}_{j-1}\right)\right\} \\
&\leq \frac{1}{n\mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \left\{\mathbb{E}\left(\Gamma_j \left|\mathbb{E}\left[J_j|X_j\right] - F^x(y)\right| | \mathfrak{F}_{j-1}\right)\right\}. \qquad (2.12)
\end{aligned}$$

The last inequality is obtained by using (H5) (iii).
Next, we have directly after integrating by parts, and changing of variables

$$\left|\mathbb{E}\left[J_j|X_j\right] - F^x(y)\right| \leq \int_{\mathbb{R}} J^{(1)}(t) \left|F^x(y - h_J t) - F^x(y)\right| dt.$$

Thus, from assumptions (H3) and (H5)(i) we get:

$$1\!\!1_{B(x,h_k)}(X_j) \left|\mathbb{E}\left[J_j|X_j\right] - F^x(y)\right| \leq \int_{\mathbb{R}} J^{(1)}(t) \left(h_K^{b_1} + |t|^{b_2} h_J^{b_2}\right) dt.$$

Since, $J^{(1)}$ is probability density function, and under the assumption (H5) (ii), we find that:

$$\mathbb{1}_{B(x,h_k)}(X_j)\,|\mathbb{E}[J_j|X_j] - F^x(y)| \le C\left(h_K^{b_1} + h_J^{b_2}\right). \tag{2.13}$$

Hence, By combining (2.12) together with (i) of Lemma A.1, we obtain

$$\sup_{x\in\mathcal{C}_{\mathcal{F}}} |\widetilde{B}_n(x,y)| = O\left(h_K^{b_1} + h_J^{b_2}\right) \sup_{x\in\mathcal{C}_{\mathcal{F}}} \bar{F}_D(x),$$

and the claimed result of this lemma is now checked.

Proof of Lemma 2.3.2

Firstly, by using (H4) (i) and because the kernel K is bounded on $[-1,1]$, it can be easily seen that

$$|\Gamma_j(x)| \le nCh_K^2 + nCh_K|\rho_j(x)|. \tag{2.14}$$

Secondary, for all $x \in \mathcal{C}_{\mathcal{F}}$, we denote: $k(x) = arg \min_{k\in\{1,2,\ldots,d_n\}} |\delta(x,x_k)|$

$$\begin{aligned}
\sup_{x\in\mathcal{C}_{\mathcal{F}}} |R_n(x,y)| \le\; & \underbrace{\sup_{x\in\mathcal{C}_{\mathcal{F}}} |\widehat{F}_N^x(y) - \widehat{F}_N^{x_{k(x)}}(y)|}_{\mathcal{Q}_1} + \underbrace{\sup_{x\in\mathcal{C}_{\mathcal{F}}} |\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)|}_{\mathcal{Q}_2} \\
+ \; & \underbrace{\sup_{x\in\mathcal{C}_{\mathcal{F}}} |\bar{F}_N^{x_{k(x)}}(y) - \bar{F}_N^x(y)|}_{\mathcal{Q}_3}
\end{aligned}$$

We will now treat each of the three terms involved in this decomposition. We start by the consistency of the term $\mathcal{Q}_1$. By using (2.7) and the boundeness on K and J, one can write:

$$\begin{aligned}
\mathcal{Q}_1 \le\; & \sup_{x\in\mathcal{C}_{\mathcal{F}}} \frac{1}{n}\sum_{j=1}^{n} |J_j(y)|\,\Big|\frac{1}{\mathbb{E}(\Gamma_1(x)K_1(x))}\Gamma_j(x)K_j(x)\mathbb{1}_{B(x,h_K)}(X_j) \\
- \; & \frac{1}{\mathbb{E}\left(\Gamma_1(x_{k(x)})K_1(x_{k(x)})\right)}\Gamma_j(x_{k(x)})K_j(x_{k(x)})\mathbb{1}_{B(x_{k(x)},h_K)}(X_j)\Big| \\
\le\; & \Big(\frac{C}{n^2h_K^2\phi(h_K)} \sup_{x\in\mathcal{C}_{\mathcal{F}}} \sum_{j=1}^{n} |\Gamma_j(x)\mathbb{1}_{B(x,h_K)}(X_j)| \\
\times\; & |K_j(x) - K_j(x_{k(x)})\mathbb{1}_{B(x_{k(x)},h_K)}(X_j)|\Big) \\
+\; & \Big(\frac{C}{n^2h_K^2\phi(h_K)} \sup_{x\in\mathcal{C}_{\mathcal{F}}} \sum_{j=1}^{n} K_j(x_{k(x)})\mathbb{1}_{B(x_{k(x)},h_K)}(X_j) \\
\times\; & |\Gamma_j(x)\mathbb{1}_{B(x,h_K)}(X_j) - \Gamma_j(x_{k(x)})|\Big) \\
:=\; & F_1 + F_2
\end{aligned}$$

Let us first deal with the term F_1. Because the kernel K satisfy the Lipschitz condition, and by using the inequality (2.14), we have

$$|\Gamma_j(x)|1\!\!1_{B(x,h_K)}(X_j)|K_j(x) - K_j(x_{k(x)})1\!\!1_{B(x_{k(x)},h_K)}(X_j)|$$

$$\leq nCh_K^2\left(\frac{r_n}{h_K}1\!\!1_{B(x,h_K)\cap B(x_{k(x)},h_K)}(X_j) + 1\!\!1_{B(x,h_K)\cap\overline{B(x_{k(x)},h_K)}}(X_j)\right)$$

Which implies that:

$$F_1 \leq \frac{Cr_n}{nh_K\phi(h_K)}\sup_{x\in\mathcal{C}_{\mathcal{F}}}\sum_{j=1}^{n}1\!\!1_{B(x,h_K)\cap B(x_{k(x)},h_K)}(X_j) + \frac{C}{n\phi(h_K)}\sup_{x\in\mathcal{C}_{\mathcal{F}}}\sum_{j=1}^{n}1\!\!1_{B(x,h_K)\cap\overline{B(x_{k(x)},h_K)}}(X_j)$$

Concerning the term F_2, we have that:

$$1\!\!1_{B(x_{k(x)},h_K)}(X_j)|\Gamma_j(x)1\!\!1_{B(x,h_K)}(X_j) - \Gamma_j(x_{k(x)})|$$

$$\leq \underbrace{1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\Gamma_j(x) - \Gamma_j(x_{k(x)})|}_{A} + \underbrace{nCh_K^2 1\!\!1_{B(x_{k(x)},h_K)\cap\overline{B(x,h_K)}}(X_j)}_{B}.$$

Now, we calculate the first part of the right side of this inequality

$$\begin{aligned} A &= 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\left(\sum_{i=1}^{n}\rho_i^2(x)K_i(x) - \rho_i^2(x_{k(x)})K_i(x_{k(x)})\right) \\ &-\left((\sum_{i=1}^{n}\rho_i(x)K_i(x))\rho_j(x)) - (\sum_{i=1}^{n}\rho_i(x_{k(x)})K_i(x_{k(x)})\rho_j(x_{k(x)})\right)| \\ &\leq A_1 + A_2, \end{aligned}$$

where

$$A_1 = 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\sum_{i=1}^{n}\rho_i^2(x)K_i(x) - \rho_i^2(x_{k(x)})K_i(x_{k(x)})|,$$

$$\begin{aligned} \text{and } A_2 &= 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\left((\sum_{i=1}^{n}\rho_i(x)K_i(x))\rho_j(x)\right) \\ &-\left((\sum_{i=1}^{n}\rho_i(x_{k(x)})K_i(x_{k(x)}))\rho_j(x_{k(x)})\right)| \end{aligned}$$

Let us now examine the terms A_1 and A_2 by putting

$$T^{k,l} = 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\left(\sum_{i=1}^{n}\rho_i^k(x)K_i(x)\right)\rho_j^l(x)$$
$$-\left(\sum_{i=1}^{n}\rho_i^k(x_{k(x)})K_i(x_{k(x)})\right)\rho_j^l(x_{k(x)})| \quad \text{with } k = 1,2 \text{ and } l = 0,1.$$

Therefore,

$$T^{k,l} \le T_1^{k,l} + T_2^{k,l},$$

with

$$T_1^{k,l} = 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\Big(\sum_{i=1}^{n}|\rho_i^k(x)|K_i(x)\times|\rho_j^l(x)-\rho_j^l(x_{k(x)})|\Big),$$

and

$$T_2^{k,l} = 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\Big(|\rho_j^l(x_{k(x)})|$$
$$\times|\sum_{i=1}^{n}(\rho_i^k(x)K_i(x)-\rho_i^k(x_{k(x)})K_i(x_{k(x)})|\Big),$$

By the assumption (H4)(ii) for $l = 1$, we can write:

$$1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\rho_j(x)-\rho_j(x_{k(x)})| \le Cr_n 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j).$$

So, for $l = 0, k = 2$

$$T_1^{k,l} = 0, \tag{2.15}$$

and for $l = 1, k = 1$

$$T_1^{k,l} \le nCr_nh_K 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j). \tag{2.16}$$

We now turn to the term $T_2^{k,l}$

$$T_2^{k,l} \le 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\left(\sum_{i=1}^{n}|\rho_j^l(x_{k(x)})|K_i(x)\times|\rho_i^k(x)-\rho_i^k(x_{k(x)})|\right)$$
$$+1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)\left(\sum_{i=1}^{n}|\rho_j^l(x_{k(x)})||\rho_i^k(x_{k(x)})|\times|K_i(x)-K_i(x_{k(x)})|\right).$$

Observe that:

$$1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\rho_i^2(x)-\rho_i^2(x_{k(x)})| \le Cr_nh_K 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j),$$

which implies that for $k = 1,2$

$$1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j)|\rho_i^k(x)-\rho_i^k(x_{k(x)})| \le Cr_nh_K^{k-1} 1\!\!1_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j).$$

Therefore, for $l = 0$, and $k = 2$

$$T_2^{k,l} \leq nCr_n h_K \mathbb{1}_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j), \tag{2.17}$$

and for $l = 1$, and $k = 1$

$$T_2^{k,l} \leq nCr_n h_K \mathbb{1}_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j). \tag{2.18}$$

Then, by Combining (2.15) with (2.17), we find that

$$A_1 \leq nCr_n h_K \mathbb{1}_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j).$$

In addition, by combining (2.16) with (2.18), allows us to find

$$A_2 \leq nCr_n h_K \mathbb{1}_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j).$$

Which implies that

$$A \leq nCr_n h_K \mathbb{1}_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j).$$

Thus

$$\begin{aligned} F_2 \leq & \frac{Cr_n}{nh_K\phi(h_K)} \sup_{x\in\mathcal{C}_\mathcal{F}} \sum_{j=1}^n \mathbb{1}_{B(x_{k(x)},h_K)\cap B(x,h_K)}(X_j) \\ & + \frac{C}{n\phi(h_K)} \sup_{x\in\mathcal{C}_\mathcal{F}} \sum_{j=1}^n \mathbb{1}_{B(x_{k(x)},h_K)\cap \overline{B(x,h_K)}}(X_j) \end{aligned}$$

Consequently, we obtain

$$\mathcal{Q}_1 \leq C \sup_{x\in\mathcal{C}_\mathcal{F}} \left(\mathcal{Q}_{1.1} + \mathcal{Q}_{1.2} + \mathcal{Q}_{1.3}\right),$$

where

$$\begin{aligned} \mathcal{Q}_{1.1} &= \frac{C}{n\phi(h_K)} \sum_{j=1}^n \mathbb{1}_{B(x_{k(x)},h_K)\cap \overline{B(x,h_K)}}(X_j), \\ \mathcal{Q}_{1.2} &= \frac{Cr_n}{nh_K\phi(h_K)} \sum_{j=1}^n \mathbb{1}_{B(x,h_K)\cap B(x_{k(x)},h_K)}(X_j), \\ \mathcal{Q}_{1.3} &= \frac{C}{n\phi(h_K)} \sum_{j=1}^n \mathbb{1}_{B(x,h_K)\cap \overline{B(x_{k(x)},h_K)}}(X_j). \end{aligned}$$

Now, we evaluate those last terms by an application of the standard inequality for sums of bounded random variables with Z_j identified such that:

$$Z_j = \begin{cases} \frac{1}{\phi(h_K)}[\mathbb{1}_{B(x_{k(x)},h_K)\cap\overline{B(x,h_K)}}(X_j)] & \text{for } \mathcal{Q}_{1.1} \\ \frac{r_n}{h_K\phi(h_K)}[\mathbb{1}_{B(x,h_K)\cap B(x_{k(x)},h_K)}(X_j)] & \text{for } \mathcal{Q}_{1.2} \\ \frac{1}{\phi(h_K)}[\mathbb{1}_{B(x,h_K)\cap\overline{B(x_{k(x)},h_K)}}(X_j)] & \text{for } \mathcal{Q}_{1.3} \end{cases}$$

It is clear that for $\mathcal{Q}_{1.1}$ and $\mathcal{Q}_{1.3}$, we have under the second part of (H1):

$$Z_j = O\left(\frac{1}{\phi(h_K)}\right), \qquad \mathbb{E}[Z_j] = O\left(\frac{r_n}{\phi(h_K)}\right) \text{ and } \quad \mathbb{E}\left(Z_j^2\right) = O\left(\frac{r_n}{\phi(h_K)^2}\right).$$

Therefore,

$$\mathcal{Q}_{1.1} = O\left(\frac{r_n}{\phi(h_K)}\right) + O_{a.co}\left(\sqrt{\frac{r_n \log n}{n\phi(h_K)^2}}\right).$$

With the same manner, assumption (H6) allows to get, for $\mathcal{Q}_{1.2}$:

$$Z_j = O\left(\frac{r_n}{h_K\phi(h_K)}\right), \qquad \mathbb{E}[Z_j] = O\left(\frac{r_n}{h_K}\right) \text{and} \quad \mathbb{E}\left(Z_j^2\right) = O\left(\frac{r_n^2}{h_K^2\phi(h_K}\right).$$

Which implies that:

$$\mathcal{Q}_{1.2} = O_{a.co}\left(\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right).$$

To finish the study of the term $\mathcal{Q}_1$, we need to put together all the intermediate result and to employ the second part of (H6) to obtain

$$\mathcal{Q}_1 = O_{a.co}\left(\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right).$$

Concerning the term $\mathcal{Q}_2$, we have for all $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P}\left(\mathcal{Q}_2 > \varepsilon\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right) &= \mathbb{P}\left(\max_{k\in 1,\dots,d_n} |\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)| > \varepsilon\right) \\ &\leq d_n \max_{k\in 1,\dots,d_n} \mathbb{P}\left(|\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)| > \varepsilon\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right). \end{aligned}$$

Let

$$\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y) = \frac{1}{\mathbb{E}(\Gamma_1 K_1)}\sum_{j=1}^{n} S_j$$

with

$$S_j = \Gamma_j(x_{k(x)})K_j(x_{k(x)})J_j(y) - \mathbb{E}\left(\Gamma_j(x_{k(x)})K_j(x_{k(x)})J_j(y)|\mathfrak{F}_{j-1}\right),$$

where S_j is a triangular array of bounded martingale differences with respect to the sequence of $\sigma-$fields $(\mathfrak{F}_{j-1})_{j\geq 1}$. So, we have

$$\begin{aligned}\mathbb{E}\left(S_j^2|\mathfrak{F}_{j-1}\right) &= \mathbb{E}\left((\Gamma_j K_j)^2 J_j^2|\mathfrak{F}_{j-1}\right) - \mathbb{E}\left((\Gamma_j K_j(x)J_j|\mathfrak{F}_{j-1})\right)^2 \\ &\leq \mathbb{E}\left((\Gamma_j K_j)^2 J_j^2|\mathfrak{F}_{j-1}\right).\end{aligned}$$

As $J_j \leq 1$, we deduce that

$$\mathbb{E}\left(S_j^2|\mathfrak{F}_{j-1}\right) \leq \mathbb{E}\left(\Gamma_j^2 K_j^2|\mathfrak{F}_{j-1}\right).$$

By using equation(2.14), (H4)(i) and (H5)(i), we obtain that:

$$\mathbb{E}\left(S_j^2|\mathfrak{F}_{j-1}\right) \leq 2C'n^2h_K^4\phi_j\left(h_K\right).$$

Now, we use the exponential inequality of Lemma 1 of [23] (with $d_j^2 = C'n^2h_K^4\phi_j(h_K)$ to obtain for all $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}\left(|\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)| > \varepsilon\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right) &= \mathbb{P}\left(|\frac{1}{n\mathbb{E}\left(\Gamma_1 K_1\right)}\sum_{j=1}^n S_j| > \varepsilon\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right) \\ &\leq 2\exp\left\{-C_0\varepsilon^2\log d_n\right\}.\end{aligned}$$

Thus, by choosing ε such that $C_0\varepsilon^2 = \varsigma$, we get

$$d_n \max_{k\in\{1,\ldots,d_n\}} \mathbb{P}\left(|\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)| > \varepsilon\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right) \leq C'd_n^{1-\varsigma}.$$

Since $\sum_{n=1}^{\infty} d_n^{1-\varsigma} < \infty$, we obtain that:

$$\mathcal{Q}_2 = O_{a.co}\left(\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right).$$

For the term $\mathcal{Q}_3$, clearly we have

$$\mathcal{Q}_3 \leq \mathbb{E}\left(\sup_{x\in\mathcal{C}_{\mathcal{F}}}|\widehat{F}_N^{x_{k(x)}}(y) - \bar{F}_N^{x_{k(x)}}(y)||\mathfrak{F}_{j-1}\right).$$

Subsequently, we follow the same steps used in studying the term $\mathcal{Q}_1$ to find

$$\mathcal{Q}_3 = O_{a.co}\left(\sqrt{\frac{\log d_n}{n\phi(h_K)}}\right).$$

This is enough to complete the proof of Lemma 2.3.2.

Proof of Lemma 2.3.3

i) This result can be deduced from Lemma 2.3.2 by taking $J_j = 1$. In this case, hypothesis (H5) (ii) and (iii) are not necessary.

ii) It is easy to see that
$\inf_{x\in\mathcal{C}_\mathcal{F}} |\widehat{F}_D(x)| \leq \frac{1}{2}$ implies that there exist $x \in \mathcal{C}_\mathcal{F}$ such that
$1 - \widehat{F}_D(x) \geq \frac{1}{2} \Longrightarrow \sup_{x\in\mathcal{C}_\mathcal{F}} |1 - \widehat{F}_D(x)| \geq \frac{1}{2}$.
According to (i) of this Lemma, we have:

$$\mathbb{P}\left(\inf_{x\in\mathcal{C}_\mathcal{F}} |\widehat{F}_D(x)| \leq \frac{1}{2}\right) \leq \mathbb{P}\left(\sup_{x\in\mathcal{C}_\mathcal{F}} |1 - \widehat{F}_D(x)| \geq \frac{1}{2}\right).$$

Consequently,

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{x\in\mathcal{C}_\mathcal{F}} |\widehat{F}_D(x)| \leq \frac{1}{2}\right) < \infty.$$

which ends the proof.

Proof of Lemma 2.3.4

For all $j = 1, \ldots, n$, let us denote

$$\eta_{n,j} = \frac{\sqrt{n\phi(h_K)}}{n\mathbb{E}(\Gamma_1 K_1)} (J_j - F^x(y)) \Gamma_j K_j,$$

and define $\xi_{n,j} = \eta_{n,j} - \mathbb{E}(\eta_{n,j}|\mathfrak{F}_{j-1})$. It is clear that

$$\sqrt{n\phi(h_K)} Q_n(x,y) = \sum_{j=1}^{n} \xi_{n,j}. \tag{2.19}$$

The summands in Equation (2.19) form a triangular array of stationary martingale differences with respect to the σ–fields $\mathfrak{F}_{j-1}$. Accordingly, the asymptotic normality of $Q_n(x,y)$ can be established by applying the central limit theorem for discrete time arrays of real–valued martingales. Therefore, to show Lemma 2.3.4, it suffices to prove the following two claims:

$$\sum_{j=1}^{n} \mathbb{E}\left(\xi_{n,j}^2 | \mathfrak{F}_{j-1}\right) \xrightarrow{\mathbb{P}} V_{JK}(x,y), \tag{2.20}$$

$$\text{and } \forall \varepsilon > 0 \; n\mathbb{E}\left(\xi_{n,j}^2 1\!\!1_{[|\xi_{n,j}|>\varepsilon]}\right) = o(1) \quad \text{(Lindeberg condition).} \tag{2.21}$$

Let us start the proof of (2.20) by remarking that

$$\mathbb{E}\left(\xi_{n,j}^2 | \mathfrak{F}_{j-1}\right) = \mathbb{E}\left(\eta_{n,j}^2 | \mathfrak{F}_{j-1}\right) - \left(\mathbb{E}\left(\eta_{n,j} | \mathfrak{F}_{j-1}\right)\right)^2.$$

Thus, it suffices to show that

$$\lim_{n\to\infty} \sum_{j=1}^{n} \Big(\mathbb{E}\left(\eta_{n,j} | \mathfrak{F}_{j-1}\right)\Big)^2 = 0 \quad \text{in probability,} \tag{2.22}$$

and

$$\lim_{n\to\infty} \sum_{j=1}^{n} \mathbb{E}\left(\eta_{n,j}^2 | \mathfrak{F}_{j-1}\right) = V_{JK}(x,y) \quad \text{in probability.} \tag{2.23}$$

Concerning the proof of (2.22), by applying Equation (2.13), Equation (2.7) and Lemma 5 of Ayad et *al.* [1], we get that

$$\begin{aligned} \left|\mathbb{E}\left(\eta_{n,j} | \mathfrak{F}_{j-1}\right)\right| &= \frac{\sqrt{n\phi(h_K)}}{n\mathbb{E}\left(\Gamma_1 K_1\right)} \left|\mathbb{E}\Big(\left(J_j - F^x(y)\right)\Gamma_j K_j(x) | \mathfrak{F}_{j-1}\Big)\right| \\ &\leq C\sqrt{n\phi(h_K)}\left(h_K^{b_1} + h_J^{b_2}\right)\frac{1}{n\phi(h_K)}\phi_j(h_K). \end{aligned}$$

Thus, by using (H2) (ii), we find

$$\sum_{j=1}^{n} \Big(\mathbb{E}\left(\eta_{n,j} | \mathfrak{F}_{j-1}\right)\Big)^2 = O_{a.co}\Big(n\phi(h_K)\left(h_K^{b_1} + h_J^{b_2}\right)^2\Big).$$

For the proof of Equation (2.23), we use (H5)(iii) to obtain

$$\begin{aligned} \sum_{j=1}^{n} \mathbb{E}\left(\eta_{n,j}^2 | \mathfrak{F}_{j-1}\right) &= \frac{\phi(h_K)}{n\left(\mathbb{E}\left(\Gamma_1 K_1\right)\right)^2} \sum_{j=1}^{n} \mathbb{E}\left(\Gamma_j^2 K_j^2 \left(J_j - F^x(y)\right)^2 | \mathfrak{F}_{j-1}\right) \\ &= \frac{\phi(h_K)}{n\left(\mathbb{E}\left(\Gamma_1 K_1\right)\right)^2} \sum_{j=1}^{n} \mathbb{E}\left(\Gamma_j^2 K_j^2 \mathbb{E}\left[\left(J_j - F^x(y)\right)^2 | X_j\right] | \mathfrak{F}_{j-1}\right). \end{aligned}$$

Next, by using the definition of the conditional variance, we find

$$\begin{aligned} \mathbb{E}\left[\left(J_j - F^x(y)\right)^2 | X_j\right] &= \text{Var}\left[J_j | X_j\right] + \left[\mathbb{E}\left(J_j | X_j\right) - F^x(y)\right]^2 \\ &:= \beta_{n1} + \beta_{n2} \end{aligned}$$

Concerning the term β_{n1}, we have

$$\text{Var}\left[J_j|X_j\right]=\mathbb{E}\left(J_j^2|X_j\right)-\left(\mathbb{E}\left(J_j|X_j\right)\right)^2 \tag{2.24}$$

An integration by part followed by a change of variable with Assumption (B2) permit us to deduce

$$\begin{aligned}\mathbb{E}\left[J_j|X_j\right] &= \int_{\mathbb{R}} J^{(1)}(t)\left[F^x(y-h_Jt)-F^x(y)\right]dt+F^x(y)\\ &= F^x(y). \end{aligned} \tag{2.25}$$

Similarly, the first term on the right hand side of Equation (2.24) is treated directly by using again (B2) combined with an integration by part and a change of variable. It follows that

$$\begin{aligned}\mathbb{E}\left[J_j^2|X_j\right] &= \int_{\mathbb{R}} J^2\left(\frac{y-z}{h_J}\right)f^x(z)dz\\ &= \int_{\mathbb{R}} 2J(t)J^{(1)}(t)\left[F^x(y-h_Jt)-F^x(y)\right]dt+\int_{\mathbb{R}} 2J(t)J^{(1)}(t)F^x(y)\,dt.\end{aligned}$$

Since $\int_{\mathbb{R}} 2J(t)J^{(1)}(t)F^x(y)\,dt=F^x(y)$, We infer that:

$$\mathbb{E}\left[J_j^2|X_j\right]\longrightarrow F^x(y), \qquad \text{as } n\to\infty. \tag{2.26}$$

Now, by combining the result (2.25) with (2.26), we arrive directly at the following result:

$$\text{Var}\left[J_j|X_j\right]=F^x(y)\left(1-F^x(y)\right) \tag{2.27}$$

Concerning the term β_{n2}, we deduce by (2.25) that

$$\beta_{n2}\longrightarrow 0, \qquad \text{as } n\to\infty.$$

Therefore,

$$\sum_{j=1}^n \mathbb{E}\left(\eta_{n,j}^2|\mathfrak{F}_{j-1}\right)=\frac{\phi(h_K)}{n\left(\mathbb{E}\left(\Gamma_1K_1\right)\right)^2}\sum_{j=1}^n \mathbb{E}\left(\Gamma_j^2K_j^2\beta_{n1}|\mathfrak{F}_{j-1}\right).$$

Combining (d) of Lemma A.1 [27], (iii) of Lemma A.2 and Equation (2.27) allow to obtain

$$\sum_{j=1}^n \mathbb{E}\left(\eta_{n,j}^2|\mathfrak{F}_{j-1}\right)\longrightarrow \frac{M_2}{M_1^2}F^x(y)\left(1-F^x(y)\right)=V_{JK}(x,y), \qquad \text{as } n\to\infty$$

which completes the proof of the claim (2.20).

Concerning the proof of (2.21), the Linderberg condition implies that $n\mathbb{E}\left(\xi_{n,j}^2\mathbb{1}_{[|\xi_{n,j}|>\varepsilon]}\right)\le 4n\mathbb{E}\left(\eta_{n,j}^2\mathbb{1}_{[|\eta_{n,j}|>\frac{\varepsilon}{2}]}\right)$. By using Markov's and Hölder's inequalities, we can write for all $\varepsilon > 0$,

$$\mathbb{E}\left(\eta_{n,j}^2 1\!\!1_{[|\eta_{n,j}|>\frac{\varepsilon}{2}]}\right) \leq \frac{\mathbb{E}\left(|\eta_{n,j}|\right)^{2a}}{(\varepsilon/2)^{2a/b}}.$$

Taking $a = 1+\frac{\delta}{2}$ for any $\delta > 0$, such that $\bar{G}_{2+\delta} = \mathbb{E}\left(|J_j - F^x(y)|^{2+\delta}|X_j\right)$ is a continuous function. It follow that

$$\begin{aligned}
4n\mathbb{E}\left(\eta_{n,j}^2 1\!\!1_{[|\eta_{n,j}|>\frac{\varepsilon}{2}]}\right) &\leq C\left(\frac{\phi(h_K)}{n}\right)^{\frac{2+\delta}{2}} \frac{n}{\left(\mathbb{E}\left(\Gamma_1 K_1\right)\right)^{2+\delta}} \\
&\times \mathbb{E}\left([|\left(J_j - F^x(y)\right)|\Gamma_j K_j]^{2+\delta}\right) \\
&\leq C\left(\frac{\phi(h_K)}{n}\right)^{\frac{2+\delta}{2}} \frac{n}{\left(\mathbb{E}\left(\Gamma_1 K_1\right)\right)^{2+\delta}} \\
&\times \mathbb{E}\left(|\Gamma_j K_j|^{2+\delta}\left[\mathbb{E}\left(|J_j - F^x(y))|^{2+\delta}|X_j\right]\right) \\
&\leq C\left(\frac{\phi(h_K}{n}\right)^{\frac{2+\delta}{2}} \frac{n}{\left(\mathbb{E}\left(\Gamma_1 K_1\right)\right)^{2+\delta}} \\
&\times \mathbb{E}\left(|\Gamma_j K_j|^{2+\delta}\right)\bar{G}_{2+\delta} \\
&= O\left(\left(n\phi(h_K)\right)^{\frac{-\delta}{2}}\right) \longrightarrow 0, \qquad \text{as } n \to \infty,
\end{aligned}$$

and the claimed result is checked.

Proof of Lemma 2.3.5

Observe that

$$\widehat{F}_D(x) - 1 = \underbrace{\widehat{F}_D(x) - \bar{F}_D(x)}_{I_1} + \underbrace{\bar{F}_D(x) - 1}_{I_2}.$$

Since $I_2 \longrightarrow 0$ almost completely as $n \longrightarrow \infty$ in view of (H2) (ii), it suffices to show that $I_1 = o(1)$ as $n \longrightarrow \infty$ Indeed, by using Lemma 2.3.3 (i) and hypothesis (H6), we obtain

$$\widehat{F}_D(x) - \bar{F}_D(x) = o(1) \text{ almost completely as } n \to \infty.$$

Which completes the proof of Lemma 2.3.5. ■

BIBLIOGRAPHY

[1] Ayad, S., Laksaci, A., Rahmani, S. and Rouane, R. On the local linear modelization of the conditional density for functional and ergodic data. *METRON.*, **78**, 237–254 (2020).

[2] Baìllo, A. and Grané, A. Local linear regression for functional predictor and scalar response. *J. of Multivariate Anal.*, **100**, 102–111(2009).

[3] Barrientos, J., Ferraty, F. and Vieu, P. Locally Modelled Regression and Functional Data. *J. Nonparametr. Statist.*, **22**, 617-632 (2010).

[4] Bouanani, O., Laksaci, A., Rachdi, M. and Rahmani, S. Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics, *Behaviormetrika*, **4**, 1–35 (2018)

[5] Bouanani, O., Rahmani, S. and Ait-Hennani, L. Local linear conditional cumulative distribution function with mixing data. *Arab. J. Math.* https://doi.org/10.1007/s40065-019-0247-7 (2019).

[6] Chaouch, M., Laïb, N. and Ould Saïd, E. Nonparametric M-estimation for right censored regression model with stationary ergodic data. *Statistical Methodology,***33**, 234–255 (2016).

[7] Cheng, M.-Y., J. Fan, and J. S. Marron. On automatic boundary corrections. Ann. Statist. **25**, 1691–1708 (1997).

[8] Delecroix, M. and Rosa, A.C. Nonparametric estimation of a regression function and its derivatives under an ergodic hypothesis, *J. Nonparametr. Statist.* **6**, 367–382 (1996).

[9] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. Functional data: local linear estimation of the conditional density and its application. *Statistics*, **47**, 26–44 (2013).

[10] Demongeot, J., Laksaci, A., Rachdi, M., and Rahmani, S. On the local linear modelization of the conditional distribution for functional data. S*ankhya A*, **76**, 328–355 (2014).

[11] Demongeot, J., Laksaci, A., Naceri, A. and Rachdi, M. Local linear regression modelization when all variables are curves. *Statist. Probab. Lett.,* **121**, 37–44 (2017).

[12] El Methni, M. and Rachdi, M. Local weighted average estimation of the regression operator for functional data. *Communications in Statistics-Theory and Methods*, **40**, 3141–3153 (2011).

[13] Fan, J. Design-adaptive nonparametric regression. *J. Amer. Statist. Assoc.*, **87**, 998–1004 (1992).

[14] Fan, J., Gijbels, I. *Local Polynomial Modelling and its Applications*. London, Chapman and Hall. (1996).

@Bookfannonlinear, title=Nonlinear time series: nonparametric and parametric methods, 2003, author=Fan, J. and Yao, Q., year=2003, publisher=Springer: New York

[15] Fan, J. and Yao, Q. *Nolinear Time Series: Nonparametric and Parametric Methods.* Springer-Verlag, New York.(2003).

[16] Ferraty, F. and Vieu, P. Nonparametric functional data analysis theory and practice. Springer-Verlag (2006).

[17] Ferraty, F. and Romain, Y. The Oxford Handbook of Fuctional Data Analysis, *Oxford University Press* (2010).

[18] Ferraty, F., Laksaci, A., Tadj. A. and Vieu, P. Rate of uniform consistency for nonparametric estimates with functional variables. *J. of Statistical Planning and Inference.*, **140**, 335–352 (2010).

[19] Gheriballah, A., Laksaci, A. and Sekkal, S. Nonparametric M-regression for functional ergodic data. *Statistics and Probability Letters.*, **83**, p. 902–908 (2013).

[20] Kara-Zaitri, L., Laksaci, A., Rachdi, M., and Vieu, P. Uniform in bandwidth consistency for various kernel estimators involving functional data. Journal of Nonparametric Statistics, **29**(1), 85-107 (2017).

[21] Kokoszka, P., and Reimherr, M. *Introduction to functional data analysis*. CRC Press. (2017).

[22] Laïb, N. and Ould-Saïd, E. A robust nonparametric estimation of the autoregression function under an ergodic hypothesis. Canad. J. Statist., **28**, 817–828, (2000).

[23] Laïb, N. and Louani D. Rates of strong consistencies of the regression function estimator for functional stationary ergodic data. *J. Stat. Plann. and Inf.*, **141**, p. 359–372 (2011).

[24] Laksaci, A., Rachdi, M. and Rahmani, S., Spatial modelization: local linear estimation of the conditional distribution for functional data. *Spat. Statist.*, **6**, 1–23 (2013).

[25] Ling, N., Meng, S., and Vieu, P. Uniform consistency rate of k NN regression estimation for functional time series data. Journal of Nonparametric Statistics, 31(2), 451-468 (2019).

[26] Ruppert, D. and Wand, M.P. Multivariate weighted least squares regression. *Ann. of Statist.*, **22**, 1346–1370 (1994).

[27] Zhou, Z. and Lin, Z.-Y. Asymptotic normality of locally modelled regression estimator for functional data. *J. Nonparametr. Statist.,* **28**, 116–131 (2016).

CHAPTER 3

ON THE LOCAL LINEAR MODELIZATION OF THE CONDITIONAL DENSITY FOR FUNCTIONAL AND ERGODIC DATA

This chapter is the subject of a publication in Journal of METRON.

On the local linear modelization of the conditional density for functional and ergodic data.

Somia Ayad [1], **Ali Laksaci** [2], **Saâdia Rahmani** [1] and **Rachida Rouane** [1].

[1] Laboratoire des Modèles Stochastiques, Statistique et Applications, Université Docteur Moulay Taher, 20000 Saïda, Algeria. E-mail: somiahadil@gmail.com. E-mail: saadia.rahmani@gmail.com. E-mail: rouane09@yahoo.fr

[2] Department of Mathematics, College of Science, King Khalid University, Abha, 61413, Saudi Arabia. E-mail:alilak@yahoo.fr.

Abstract : In this paper, we estimate the conditional density function using the local linear approach. We treat the case when the regressor is valued in a semi-metric space, the response is a scalar and the data are observed as ergodic functional times series. Under this dependence structure, we state the almost complete consistency (a.co.) with rates of the constructed estimator. Moreover, the usefulness of our results is illustrated through their application to the conditional mode estimation.

Keywords : Ergodic data, functional data, local linear estimator, conditional density, nonparametric estimation, conditional mode, ozone concentration.

Mathematics Subject Classification: 62G05, 62G08, 62G20, 62G35.
Secondary: 62H12.

3.1 Introduction and motivations

In recent years, several nonparametric estimators of the conditional models have been proposed in the literature when the explanatory random variables take their values in an infinite dimensional space (such as a Hilbert or Banach space). This field of research, known as Functional Data Analysis aims to analyze information on curves or surfaces or other complex objects. Besides, various literature deals with the limit properties of these estimators in both independent and dependent cases. For an overview, one may refer to the Oxford handbook of Ferraty and Romain [17] and to the pioneer work of Ferraty and Vieu [16] and the references therein.

It has been shown extensively in the literature that the prediction of a scalar response knowing an explaining functional variable is obtained by estimating the conditional expectation of Y given X, or alternatively, by the conditional mode or the conditional median which can be derived directly from the conditional density function. Moreover, it is well known that various nonparametric methods can be proposed for estimating these conditional models: kernel, spline or orthogonal series, among many others. Although, in general, these estimators have several important advantages, they do not necessarily yield more efficient estimates in practice. As discussed in Fan [14], the local polynomial estimators, and specially, the local linear ones, have some good features, such as the significative gain in the bias term with respect the classical kernel method, their appropriate boundary behaviour and their efficiency in an asymptotic minimax sense. For practical applications, local linear fitting is usually the most useful procedure (see for instance, Ruppert and Wand [24] and Fan and Gijbels [15] in the multivariate case, and more recently, see Barrientos et *al.* [3] and Laksaci et *al.* [22] in the FDA setup among others). Moreover it is worth mentioning that many recent works ([4],[5], [10] and [19]) followed the idea of Barrientos et *al.* [3] to construct the local linear estimates of several conditional models. In this work, we examine the extent to which the nice properties of the local linear estimator can be used to study the conditional density function under the only assumption that the process generating the functional data is stationary ergodic. We attempt to resolve this problem, which is not fully addressed in the literature yet.

On the other hand, the ergodic theory constitutes a recent and important research area in the study of stochastic processes. This study has a very wide range of applications, because most of the random phenomena we encounter around us are not independent. The ergodic processes are a class of stochastic processes that have the property that one sample of the process represents all the set. This theory represent now a very fashionable research area.

In the statistical literature, several papers have been devoted to the study of some properties of the nonparametric stationary ergodic processes estimators (see for instance, Didi and Louani [13] in the case of complete data and Chaouch et *al.* ([6] for right censored ones). This is due to the fact that despite that a large class of processes satisfies the condition of α-mixing (see for instance Laïb and Louani [21] and the references therein). However, there is still a great number of processes where such condition does not hold (as for example, the simple process $AR(1) : X_n = \rho X_{n-1} + \epsilon_n$, where $\rho \in]0, 1/2]$ and ϵ_n are i.i.d. with the binomial distribution). Furthermore, the ergodic assumption permits to avoid the complicated probabilistic calculations of the mixing condition. It is then essential to consider a general larger dependency framework, as is the ergodicity.

In this paper, we are concerned with the almost-complete convergence with rates of the local linear estimator of the conditional density function. For this purpose, it is assumed that the covariate takes its values in an infinite dimensional space and the data are sampled from a stationary ergodic process.

This paper is organized as follows. In Section 3.2, we construct the local linear estimator of the conditional density function. In section 3.3, we introduce and discuss in detail necessary conditions for establishing the almost-complete convergence of the constructed estimator. Section 3.4 states the main results achieved by the conditional density function estimator. In Section 3.5, we use the obtained results to derive some asymptotic properties for the local linear estimator of the conditional mode. In the same section, and to make our results as accessible as possible, we illustrate them with an application on real data, in which we demonstrate the practical performance of the proposed method. Finally, details of technical lemmas and all proofs are gathered in Appendix.

3.2 Local linear estimator construction

Let $Z_i = (X_i, Y_i)_{i=1,\ldots,n}$ be an $\mathcal{F}\times\mathbb{R}$-valued measurable strictly stationary process, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, where $\mathcal{F}$ is a semi-metric space, and d denotes the semi-metric. Furthermore, we assume that there exists a regular version of the conditional distribution of Y given X, which is absolutely continuous with respect to the *Lebesgue* measure on $\mathbb{R}$, and has a twice continuously differentiable probability density function denoted by $f^X(Y)$.
We focus on the estimation of the conditional density of Y given $X = x$ via the local linear method. For this purpose, it is well known that the main idea, in the local linear smoothing, is based on the fact that the function $f^x(y)$ admits a linear approximation in the neighborhood of the conditioning point. This consideration is motivated by the fact that the conditional density function can be expressed as a regression model with the response variable $\frac{1}{h_J}J(\frac{\cdot-Y}{h_J})$ instead Y, where J is a kernel function and $h_J = h_{J,n}$ is a sequence of positive real numbers under the condition $h_J \longrightarrow 0$. (See for instance, Fan [14] in the non-functional case, and Rachdi et *al.* [23] in the functional setting).
For this aim, we assume that the underlying process Z_i is functional stationary ergodic, and we propose to construct the estimator $\widehat{f^x}$ of f^x by $\widehat{f^x} = \widehat{a_0}$ which is obtained from the following minimization procedure:

$$\min_{(a_0,a_1)\in\mathbb{R}^2} \sum_{i=1}^{n} \left(\frac{1}{h_J} J\left(\frac{y-Y_i}{h_J}\right) - a_0 - a_1\rho(X_i, x)\right)^2 K\left(\frac{\delta(x, X_i)}{h_K}\right), \tag{3.1}$$

with $\rho(.,.)$ and $\delta(.,.)$ are known bi-functional operators defined from $\mathcal{F}^2$ into $\mathbb{R}$ such that $|\delta(x,z)| = d(x,z)$ and $\rho(z,z) = 0, \forall z \in \mathcal{F}$, (see Barrientos et *al.* [3] for some examples of these two locating functions). h_K is the smoothing parameter associated with the kernel K. Such fast version of functional local linear estimation has been proposed by Demongeot et *al.* [10] under the strong mixing condition usually assumed in functional time series analysis. They showed, by simple algebra, that:

$$\widehat{f}^x(y) = \frac{\sum_{i,j=1}^{n} W_{ij}(x) J\left(\frac{y-Y_j}{h_J}\right)}{h_J \sum_{i,j=1}^{n} W_{ij}(x)}, \tag{3.2}$$

where

$$W_{ij}(x) = \rho(X_i,x)\left(\rho(X_i,x)-\rho(X_j,x)\right) K\left(\frac{\delta(x,X_i)}{h_K}\right) K\left(\frac{\delta(x,X_j)}{h_K}\right).$$

For simplicity of notations, we write $\rho_i = \rho(X_i,x)$, $K_i = K\left(\frac{\delta(x,X_i)}{h_K}\right)$ $J_j = J\left(\frac{y-Y_j}{h_J}\right)$, and

$$\Gamma_j = K_j^{-1}\left(\sum_{i=1}^{n} W_{ij}\right) = \sum_{i=1}^{n} \rho_i^2 K_i - \left(\sum_{i=1}^{n} \rho_i K_i\right)\rho_j.$$

Then, it is obvious that (3.2) can be rewritten as

$$\widehat{f}^x(y) = \frac{\sum_{j=1}^{n} \Gamma_j K_j J_j}{h_J \sum_{j=1}^{n} \Gamma_j K_j}.$$

Notice that in their paper, Demongeot et *al.* [11] established the almost complete consistency (pointwise and uniform) of (3.2). While, the spatial version of this estimator was studied by *Laksaci* et *al.* [22]. Recently, the asymptotic normality of the same estimator was obtained by *Bouanani* et *al.* [4] in the i.i.d. case and by *Zhou* and *Lin* [25] in the strong mixing case.

3.3 Assumptions and notations

In order to state our results, we introduce some notations:
For $i = 1,\ldots,n$, let $\mathfrak{F}_i$ (resp. $\mathcal{G}_i$) be the σ-field generated by $((X_1,Y_1),\ldots,(X_i,Y_i))$, (resp. $((X_1,Y_1),\ldots,(X_i,Y_i),X_{i+1})$). For any fixed x in $\mathcal{F}$, $\mathcal{N}_x$ denotes a fixed neighborhood of x; and let $\mathscr{C}$ a fixed compact subset of $\mathbb{R}$. Moreover, let us denote by $\phi_x(r_1,r_2) = \mathbb{P}(r_2 \le \delta(X,x) \le r_1)$ the small ball probability function. Furthermore, when no confusion is possible, we will denote by C and C' some strictly positive generic constants.
Our consistency results are summarized in Theorems 3.4.1, 3.4.2 and relies on the following five assumptions:

(H.1) **On the ergodic functional variables:**
We suppose that the strictly stationary ergodic process $(X_i,Y_i)_{i\in\mathbb{N}^*}$ satisfies:

i) For all $r > 0$, $\phi_x(r) = \mathbb{P}(X \in B(x,r)) > 0$, where $B(x,r) = \{x' \in \mathcal{F}/|\delta(x',x)| < r\}$.

ii) For all $i = 1,\dots,n$ there exist a deterministic function $\phi_{i,x}(.)$ such that
$0 < \mathbb{P}(X_i \in B(x,r)|\mathfrak{F}_{i-1}) \le \phi_{i,x}(r),\ \forall r > 0,$

iii) For any $r > 0$, $\dfrac{1}{n\phi_x(r)}\sum_{i=1}^{n}\phi_{i,x}(r) \xrightarrow{P} 1$ and $n\phi_x(r) \to \infty$ as $r \to 0$.

The first part of this hypothesis is clearly unrestrictive, since it is the same as that classically used in the infinite-dimensional setting. The reader will be find in Gheriballah et *al.* [18] a deeper discussion on (ii). The ergodicity of functional data is exploited together with condition ((H.1) (iii)) which is less restrictive to the condition imposed by Laib and Louani [21].

(H.2) **On the regularity of the model:**
The conditional density function f^x will be supposed to verify one of the following constraints: $\forall x' \in \mathcal{N}_x$ and $\forall y' \in \mathscr{C}$:

i) $\begin{cases} f : \mathcal{F} \times \mathbb{R} \longrightarrow \mathbb{R}, \lim\limits_{|\delta(x,x')| \to 0} f^{x'}(y) = f^x(y) \\ \text{and } \lim\limits_{|y'-y| \to 0} f^x(y') = f^x(y). \end{cases}$

ii) There exist some positive constants b_1 and b_2 such that:

$$|f^x(y) - f^{x'}(y')| \le C\left(|\delta(x,x')|^{b_1} + |y - y'|^{b_2}\right),$$

where C is a positive constant depending on x.

The first part of this condition is a continuity-type constraint, while the second part is more restrictive and it is based on the Lipchitz-type condition ((H.2)(ii)). The part (i) is necessary to get pointwise convergence, while the second consideration (ii) is used to make precise the convergence rate of the estimate.

(H.3) **On the locating functions:**
The function ρ satisfies the following condition:
$\forall z \in \mathcal{F},\ C|\delta(x,z)| \le |\rho(x,z)| \le C'|\delta(x,z)|$.
This hypothesis is another regularity-type constraint in order to control the shape of the locating function ρ in relation with δ.

(H.4) **On the kernels:**

(i) K is a nonnegative bounded kernel supported on $[-1,1]$.

(ii) The kernel J is a positive, bounded and Lipschitzian continuous function, such that:

$$\int |t|^{b_2} J(t)dt < \infty \quad \text{and} \quad \int J^2(t)dt < \infty.$$

(iii) $\mathbb{E}\left(J(\frac{y-Y_j}{h_J})|\mathcal{G}_{j-1}\right) = \mathbb{E}\left(J(\frac{y-Y_j}{h_J})|X_j\right)$.

The boundedness of the kernel K in (H.4)(i) is standard and the reader will find in Ferraty and Vieu [16] some examples of kernels satisfying this hypothesis. (H.4) (ii) and (iii) are technical conditions for the brevity of proofs.

(H.5) **On the bandwidths h_K and h_J:**

$$\lim_{n\to\infty} h_K = 0,\ \lim_{n\to\infty} n^\alpha h_J^\gamma = \infty \text{ with } \gamma = 1, 2, \text{ and } \alpha > 1,$$

$$\lim_{n\to\infty} \frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)} = 0, \text{ where } \varphi_x(h_K) = \sum_{i=1}^{n} \phi_{i,x}(h_K)$$

and

$$h_K \int_{B(x,h_K)} \rho(u,x)\, dP(u) = o\left(\int_{B(x,h_K)} \rho^2(u,x)\, dP(u)\right),$$

where $dP(x)$ is cumulative distribution of X.

The local behaviour of ρ which models the local shape of our model is controlled by this last assumption and the rest of this assumption is standard.

3.4 Main results

Our first principal result is given in the following theorem which articulates the pointwise almost complete convergence.

Theorem 3.4.1. *Under the assumptions (H.1), (H.2) (i), (H.3)–(H.5) , we have*

$$\sup_{y\in\mathscr{C}} |\widehat{f^x}(y) - f^x(y)| = o(1), \qquad a.co.$$

So as to bestow a more definite asymptotic result, we replace (H.2)(i) by (H.2) (ii) and we obtain the following result:

Theorem 3.4.2. *Under the assumptions (H.1), (H.2) (ii), (H.3)–(H.5) , we have*

$$\sup_{y\in\mathscr{C}} |\widehat{f^x}(y) - f^x(y)| = O\left(h_K^{b_1}\right) + O\left(h_J^{b_2}\right) + O\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right), \quad a.co.$$

Remark 3.4.1. . *We point that the obtained convergence rate has the same structure of the almost complete convergence rate in nonparametric functional data analysis, in sense that the dimensionality of the model is explored in this bias term, whereas the functional and the correlation aspects are explored in the dispersion term, through the functions ϕ_x and φ_x, respectively.*

In particular the correlation is explored in the nominator of the dispersion term, while the concentration of the functional variable is explored in the denominator of the dispersion part. At this stage our convergence rate is similar to convergence rate obtained by Attouch et al. [1] for the M-regression of the functional mixing time series case.

Before starting the proof of our main result, it is convenient to denote that:

$$\widehat{f}_k^x(y) = \frac{1}{nh_J^k \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \Gamma_j K_j J_j^k,$$

and

$$\bar{f}_k^x(y) = \frac{1}{nh_J^k \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \mathbb{E}\left(\Gamma_j K_j J_j^k | \mathfrak{F}_{j-1}\right), \quad \text{with } k = 0, 1.$$

Next, the proof of Theorems 3.4.1 and 3.4.2 is based on the following decomposition:

$$\begin{aligned}\widehat{f}^x(y) - f^x(y) &= \left(\frac{\bar{f}_1^x(y)}{\bar{f}_0^x(y)} - f^x(y)\right) + \frac{1}{\widehat{f}_0^x}\left[\left(\frac{\bar{f}_1^x(y)}{\bar{f}_0^x(y)} - f^x(y)\right)\left(\bar{f}_0^x(y) - \widehat{f}_0^x(y)\right)\right. \\ &+ \left.\left(\left(\widehat{f}_1^x(y) - \bar{f}_1^x(y)\right) - f^x(y)\left(\widehat{f}_0^x(y) - \bar{f}_0^x(y)\right)\right)\right]\end{aligned} \tag{3.3}$$

and the following lemmas for which proofs are given in the appendix.

Lemma 3.4.1. *Under assumptions (H.1), (H.2) (i), (H.3) and (H.4), we have:*

$$\sup_{y\in\mathscr{C}}\left|\left(\frac{\bar{f}_1^x(y)}{\bar{f}_0^x(y)} - f^x(y)\right)\right| = o(1). \tag{3.4}$$

If we substitute (H.2) (i) by (H.2) (ii), we have:

$$\sup_{y\in\mathscr{C}}\left|\left(\frac{\bar{f}_1^x(y)}{\bar{f}_0^x(y)} - f^x(y)\right)\right| = O\left(h_K^{b_1}\right) + O\left(h_J^{b_2}\right). \tag{3.5}$$

Lemma 3.4.2. *Under the assumptions (H.1), (H.3)–(H.5), we have*

$$\widehat{f}_0^x(y) - \bar{f}_0^x(y) = O_{a.co}\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2\phi_x^2(h_K)}}\right).$$

Lemma 3.4.3. *Under the same assumptions of lemma 3.4.2, we have*

$$\exists\, C > 0 \text{ such that } \sum_{n=1}^{\infty} \mathbb{P}\left(\widehat{f}_0^x(y) < C\right) < \infty.$$

Lemma 3.4.4. *Under the assumptions (H.1),(H.2) (ii), (H.3)–(H.5) , we have*

$$\sup_{y\in\mathscr{C}} |\widehat{f_1^x}(y) - \bar{f}_1^x(y)| = O_{a.co}\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right).$$

3.5 Application to conditional mode

The purpose of this section is to study the nonparametric estimate $\widehat{\Theta}(x)$ of the conditional mode $\Theta(x)$ by the local linear approach. Recall that these questions in infinite dimension are particularly interesting, not only for the fundamental problemmas they formulate, but also for many applications, (see for instance Dabo and Laksaci [8] and Dabo et *al.* [7]).
We assume that the conditional density f^x has a unique mode $\Theta(x)$ on $\mathscr{C}$. A natural and usual estimator of $\Theta(x)$ is defined as the random variable $\widehat{\Theta}(x)$ which maximizes the local linear estimator $\widehat{f^x}(.)$ of $f^x(.)$ that is:

$$\widehat{\Theta}(x) = \arg\sup_{y\in\mathscr{C}} \widehat{f^x}(y).$$

So, in addition to the assumptions introduced along the previous section, we need the following one:

(H.6) (i) There exists some integer $j > 1$ such that $\forall x \in \mathcal{N}_x$, the function f^x is j-times continuously differentiable on the topological interior of $\mathscr{C}$ with respect to y.

(ii) $f^{x(l)}(\Theta(x)) = 0 \quad$ if $\quad 1 \leq l < j$.

(iii) $f^{x(j)}(.)$ is uniformly continuous on $\mathscr{C}$ such that $0 < |f^{x(j)}(\Theta(x))| < \infty$,

where $f^{x(j)}$ stands for the j^{th}-order derivative of the conditional density f^x.

Then, from Theorem 3.4.2, we deduce the consistency of $\widehat{\Theta}$ which is summarizing in the following corollary.

Corollary 3.5.1. *Suppose that assumptions of Theorem 3.4.2 are satisfied. In addition, if (H.6) holds, one gets:*

$$|\widehat{\Theta}(x) - \Theta(x)| = O\left(h_K^{\frac{b_1}{j}}\right) + O\left(h_J^{\frac{b_2}{j}}\right) + O\left(\left(\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}\right)^{\frac{1}{2j}}\right), \quad a.co.$$

3.5.1 A real data application

The main aim of this computational part is to examine the easy implementation of the conditional mode estimator $\widehat{\Theta}(x)$. Moreover, we highlight the superiority of the local linear approach compared to the kernel method for which the conditional mode is estimated by

$$\widetilde{\Theta}(x) = \sup_{y\in\mathscr{C}} \widetilde{f}^x(y)$$

where

$$\widetilde{f}^x(y) = \frac{h_J^{-1}\sum_{i=1}^{n} K(h_K^{-1}d(x,X_i))J(h_J^{-1}(y-Y_i))}{\sum_{i=1}^{n} K(h_K^{-1}d(x,X_i))}.$$

To do that we constructed an ergodic functional time series by cutting a continuous time process. Recall that one of the most usual example of functional time series prediction is the estimation of the future characteristic of continuous time process given its whole past observed in the continuous path. Formally, assume that we observe a continuous time process $(Z_t)_{t\in[0,b[}$ and we aim to estimate $Y = F(Z_{b+s})$ given future characteristic of the process $(Z_t)_{t\in[0,b[}$. To do that, we suppose that s is small enough to build $N < \frac{b}{s}$ functional random variables $(X_i)_{i=1,\dots,N}$ defined by:

$$\forall t \in [0,b[, \qquad X_i(t) = Z_{((i-1)b+t)/N}.$$

In this context the random variables $\widehat{\Theta}(X_N)$ and $\widetilde{\Theta}(X_N)$ are the best approximation of the quantity $F(Z_{b+s})$. The latter is obtained by using the $N-1$ pairs of random variables $(X_i, Y = F(Z_{(ib/N)+s}))_{i=1,\dots,N-1}$.
Now, we apply this idea to the problem of ozone concentration forecasting. Of course, there exist several interesting ozone characteristics which can be predicted one day ahead , such as the mean, the peak, the total or for fixed hour in the day. In this real data example, we focus in the prediction the total ozone in one day ahead using the conditional mode estimation. Precisely, we consider the the ozone data collected in Marylebone road monitoring site. The geographical locations of this site are 51.522530 (latitude) , -0.154611 (longitude). This site is an important thoroughfare within the Westminster city. This road is highly congested. Specifically, it is frequented by 90.000 vehicles per day. In this application study we focus on the hourly measurements of this polluting gas during the 2018-year. The data of this example is provided by the website https://www.airqualityengland.co.uk/.

Notice that the ozone gas is formed by a reaction between nitrogen oxides and the organic compounds. Undoubtedly this polluting gas is harmful for human health and its forecasting has been the subject of several studies in functional data analysis. The most of the previous works consider the ozone data as functional autoregressive model (see, Damon and Guillas [9]). Such common treatment motivates the use of this data as ergodic functional time series such as linear process. In this context, we apply the local linear mode estimation to predict the total ozone concentration in one day ahead the whole daily curves (one day before). Indeed, according to the previous algorithm , Z_t designs the ozone concentration for 8736 hours between 01/01/2018 and 31/12/2018. We cut this functional time series in $N+1=364$ pieces X_i of 24 hours (one day). These functionals variables X_i are presented by the following figure (Figure 3.1)

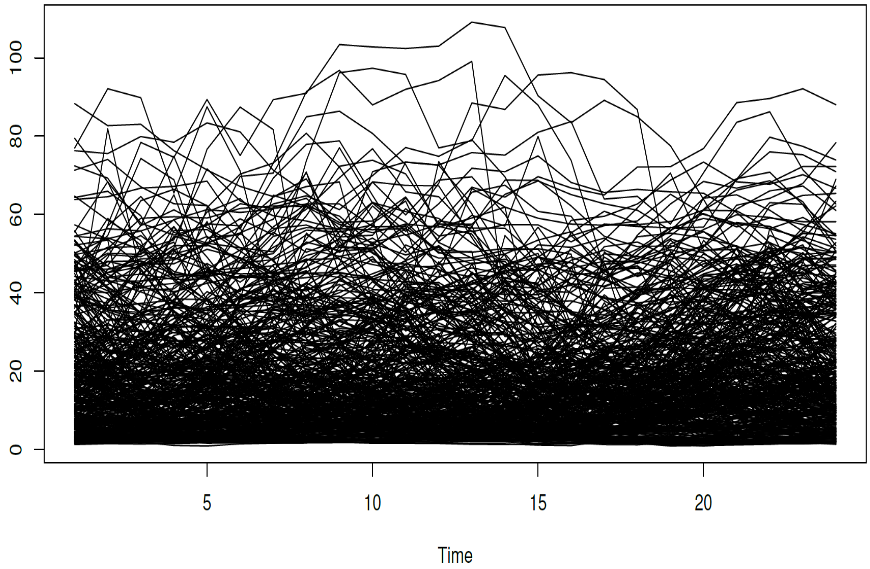

Figure 3.1: Hourly ozone concentration of the year 2018.

The scalar response variable Y is defined by $Y_i = \sum_{h=0}^{23} X_{i+1}(h)$. For this comparison study we compute both estimators in its optimal conditions. In particular, we choose the optimal bandwidths (h_K, h_J) locally by the cross-validation method on the k-nearest neighbors with respect the following MSE-criterion

$$MSE(Ker) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widetilde{\Theta}^{-i}(X_i))^2, \qquad \text{and} \qquad MSE(LL) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{\Theta}^{-i}(X_i))^2$$

where $\widetilde{\Theta}^{-i}$ (resp. $\widehat{\Theta}^{-i}$) designs the leave-one-out kernel (resp. local linear) estimator of the conditional mode. The behavior of the two estimators is also linked to the choice of the functions ρ, δ and d. Regarding the Figure 3.1, it appears the ACP metric is more adapted of these discontinuous curves. Thus, we have opted to take

$$\rho(x,z) = \delta(x,z) = d_{PCA}(x,z).$$

d_{PCA} is the semi-metric based on the $m=3$ first eigenfunctions of the empirical covariance operator associated to the $m=3$ greatest eigenvalues (see Ferraty and Vieu [16] for more discussion). The following quadratic kernel was selected:

$$J(x) = K(x) = \frac{3}{4}(1-x^2)\mathbb{1}_{[0,1]}.$$

Now, in order to compare both methods we split our data into two subsets I_1 and I_2. The 244 observations $(X_j, Y_j)_{j\in I_1}$ will be our statistical sample from which are calculated the estimators and the 120 remaining observations $(X_i, Y_i)_{i\in I_2}$ are considered as the testing sample. Next, we use the following algorithm:

- *Step 1.* For each curve X_j in the input sample we approximate the associated response variable Y_j by

$$\widehat{Y}_j = \widetilde{\Theta}(X_j)$$

  and

$$\widehat{Y}_j = \widehat{\Theta}(X_j).$$

- *Step 2.* For each X_{new} in the testing sample, we put

$$i^* = \arg\min_{j\in I_1} d(X_{new}, X_j).$$

- *Step 3.* For each X_{new} we put

$$h_K = \text{ the optimal bandwidth parameter associated to } X_i^*$$

  and

$$h_J = \text{ the optimal bandwidth parameter associated to } Y_i^*$$

- *Step 4.* We predict Y_{new} by

$$\widehat{Y}_{new} = \widetilde{\Theta}(X_{new})$$

  and

$$\widehat{Y}_{new} = \widehat{\Theta}(X_{new}).$$

- *Step 5.* We calculate the prediction error expressed by

$$\frac{1}{120}\sum_{i\in I_1}(Y_i - \widehat{T}(X_i))^2,$$

  where $\widehat{T}$ means either the kernel estimator or the local linear one.

- *Step 6.* We divide again our observations in the two subsets I_1 and I_2 and we repeat the step 1-5.

- *Step 7.* We repeat the Step 6 several times.

- *Step 8.* We end this analysis by plotting the box-plot of the mean square errors of each method.

The comparisons study is carried out by repeating the algorithm 60 times with random splitting of the observations between training and testing sample. In conclusion we point out that the scatter-plots indicates that the local linear method is significantly better than the kernel method. This statement confirms the superiority of the local linear approach over the kernel method. In the other hand, the errors results are really comparable with those obtained by Aneiros et *al.* (2004). Of course the small error differences is only due to the difference of the climatic conditions of the geographic areas under study.

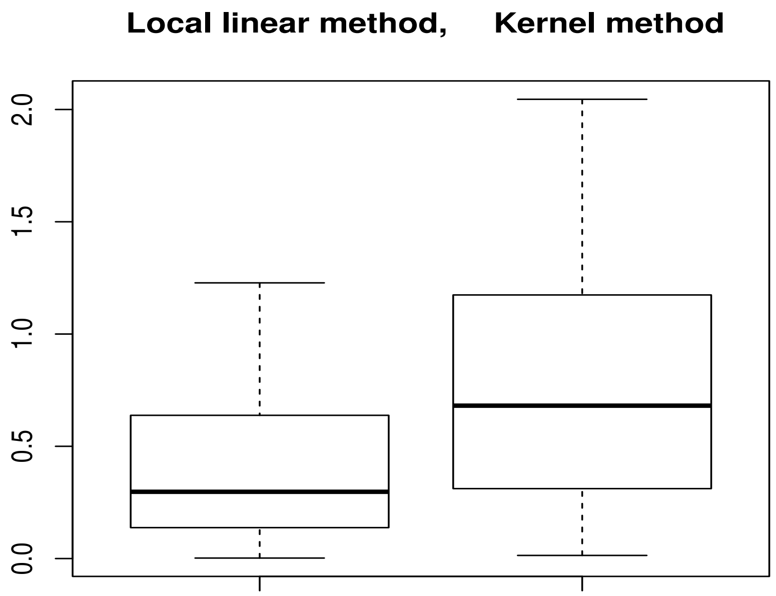

Figure 3.2: Comparison of the Ozone concentration prediction between the kernel method and the local linear approach

3.6 Appendix

3.6.1 Preliminary technical lemmas

Firstly, we state the following technical lemmas which are needed to establish our asymptotic results.

Lemma 3.6.1. *Under the assumptions (H.1),(H.3) and (H.4)(i), we have:*
$\forall\,(k,l)\in\mathbb{N}^*\times\mathbb{N}$,

(i) $\mathbb{E}\left(K_j^k|\rho_j|^l|\mathfrak{F}_{j-1}\right)\leq Ch_K^l\phi_{j,x}\left(h_K\right)$

(ii) $\mathbb{E}(\Gamma_j K_j|\mathfrak{F}_{j-1}) = O(nh_K^2\phi_{j,x}(h_K))$

(iii) $\mathbb{E}(\Gamma_1 K_1) = O(nh_K^2\phi_x(h_K))$

Proof.

(i) One starts by using (H.3) followed by using (H.4), we get

$$\begin{aligned} K_j^k|\rho_j|^l h_K^{-l} &\leq CK_j^k|\delta(X_j,x)|^l h_K^{-l} \\ &\leq C|\delta(X_j,x)|^l h_K^{-l}\mathbb{1}_{[-1,1]}(\delta(X_j,x)), \end{aligned}$$

and thereby, we have

$$\begin{aligned} \mathbb{E}\left(K_j^k|\rho_j|^l h_K^{-l}|\mathfrak{F}_{j-1}\right) &\leq C\mathbb{P}(X_j \in B(x,h_K)|\mathfrak{F}_{j-1}), \\ &\leq C\phi_{j,x}(h_K), \end{aligned}$$

which is the claimed result.

(ii) Recall that the fact that the kernel K is bounded on $[-1,1]$ and under (H.3), we have

$$|\Gamma_j| \leq nCh_K^2 + nCh_K|\rho_j|.$$

So, by using (i), we find

$$\begin{aligned} \mathbb{E}(\Gamma_j K_j|\mathfrak{F}_{j-1}) &\leq nCh_K^2\phi_{j,x}(h_K) + nCh_K^2\phi_{j,x}(h_K) \\ &\leq nC'h_K^2\phi_{j,x}(h_K). \end{aligned}$$

(iii) Combining (H.1)(iii) with part (ii) of the same lemma, and by considering $\mathfrak{F}_j$ as the trivial $\sigma-$ filed, part (iii) is directly verified. ■

Lemma 3.6.2. *Under the assumptions of lemma (3.6.1), we have*

$$\lim_{n\to\infty} \bar{f}_0^x(y) = O(1).$$

Proof.

We start by applying parts (ii) and (iii) of lemma 3.6.1 to get

$$\lim_{n\to\infty} \bar{f}_0^x(y) = O(1)\lim_{n\to\infty}\frac{1}{n\phi_x(h_K)}\sum_{j=1}^{n}\phi_{j,x}(h_K).$$

Finally, we just have to use part (iii) of assumption (H.1) to obtain the claimed result. ∎

3.6.2 Proofs of main results:

Proof of Lemma 3.4.1

Observe that

$$\begin{aligned}
\frac{\bar{f}_1^x(y)}{\bar{f}_0^x(y)} - f^x(y) &= \frac{1}{nh_J \mathbb{E}(\Gamma_1 K_1)\bar{f}_0^x(y)} \sum_{j=1}^{n} \left\{ \mathbb{E}(\Gamma_j K_j J_j | \mathfrak{F}_{j-1}) - h_J f^x(y) \mathbb{E}(\Gamma_j K_j | \mathfrak{F}_{j-1}) \right\} \\
&= \frac{1}{nh_J \mathbb{E}(\Gamma_1 K_1)\bar{f}_0^x(y)} \sum_{j=1}^{n} \left\{ \mathbb{E}(\Gamma_j K_j \mathbb{E}(J_j | \mathcal{G}_{j-1}) | \mathfrak{F}_{j-1}) - h_J f^x(y) \mathbb{E}(\Gamma_j K_j | \mathfrak{F}_{j-1}) \right\} \\
&\leq \frac{1}{nh_J \mathbb{E}(\Gamma_1 K_1)\bar{f}_0^x(y)} \sum_{j=1}^{n} \left\{ \mathbb{E}(\Gamma_j K_j |\mathbb{E}[J_j | X_j] - h_J f^x(y)| | \mathfrak{F}_{j-1}) \right\}.
\end{aligned}$$

The last inequality is obtained by (H.4) (iii). Next an integration par parts and the change of variable allow to get

$$\mathbb{E}(J_j | X_j) = h_J \int_{\mathbb{R}} J(t) f^x(y - h_J t)\, dt, \tag{3.6}$$

thus, we have

$$|\mathbb{E}[J_j | X_j] - h_J f^x(y)| \leq h_J \int_{\mathbb{R}} J(t) |f^x(y - h_J t) - f^x(y)|\, dt.$$

On one side, if we use the assumption (H.2)(i) followed by (H.4) (ii) and lemma 3.6.2, we obtain the part (3.4) of lemma 3.4.1.
And on the other side, if we replace (H.2) (i) by (H.2) (ii) we obtain

$$1\!\!1_{B(x,h_k)}(X_j) |\mathbb{E}[J_j | X_j] - h_J f^x(y)| \leq h_J \int_{\mathbb{R}} J(t) \left(h_K^{b_1} + |t|^{b_2} h_J^{b_2} \right) dt.$$

Hence, we get

$$\begin{aligned}
\bar{f}_1^x(y) - f^x(y)\bar{f}_0^x(y) &= \left(O\left(h_K^{b_1}\right) + O\left(h_J^{b_2}\right) \right) \times \frac{1}{n\mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \mathbb{E}(\Gamma_j K_j | \mathfrak{F}_{j-1}) \\
&= \left(O\left(h_K^{b_1}\right) + O\left(h_J^{b_2}\right) \right) \times \bar{f}_0^x(y).
\end{aligned}$$

Finally, making use lemma 3.6.2 allows us to obtain the part (3.5) of lemma 3.4.1. ∎

Proof of Lemma 3.4.2

Before proving this lemma let us start by writing that:

$$\begin{aligned}\widehat{f}_k^x(y) - \bar{f}_k^x(y) &= \frac{1}{nh_J^k \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^n \left(\Gamma_j K_j J_j^k - \mathbb{E}\left(\Gamma_j K_j J_j^k | \mathfrak{F}_{j-1}\right)\right) \\ &=: \frac{1}{nh_J^k \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^n T_j, \text{ with } k = 0, 1,\end{aligned}$$

and where T_j is a triangular array of martingale differences according to the σ- fields$(\mathfrak{F}_{j-1})_j$. In view that $\mathbb{E}\left(\Gamma_j K_j J_j^k | \mathfrak{F}_{j-1}\right)$ is $\mathfrak{F}_{j-1}$ measurable, it follows that

$$\begin{aligned}\mathbb{E}\left(T_j^2 | \mathfrak{F}_{j-1}\right) &= \mathbb{E}\left((\Gamma_j K_j)^2 J_j^{2k} | \mathfrak{F}_{j-1}\right) - \mathbb{E}\left((\Gamma_j K_j J_j^k | \mathfrak{F}_{j-1})\right)^2 \\ &\leq \mathbb{E}\left((\Gamma_j K_j)^2 \mathbb{E}\left(J_j^{2k} | \mathcal{G}_{j-1}\right) | \mathfrak{F}_{j-1}\right) \\ &\leq \mathbb{E}\left((\Gamma_j K_j)^2 \mathbb{E}\left(J_j^{2k} | X_j\right) | \mathfrak{F}_{j-1}\right).\end{aligned}$$

Now using (3.6) and by assumptions (H.2)(ii) and (H.4) (ii), we get

$$\mathbb{E}\left(J_j^{2k} | X_j\right) = O\left(h_J^k\right).$$

So,

$$\mathbb{E}\left(T_j^2 | \mathfrak{F}_{j-1}\right) \leq C h_J^k \mathbb{E}\left(\Gamma_j^2 K_j^2 | \mathfrak{F}_{j-1}\right).$$

Thus,

$$\begin{aligned}\mathbb{E}\left(T_j^2 | \mathfrak{F}_{j-1}\right) &\leq 2C h_J^k \left(\mathbb{E}\left(\left(\sum_{i=1}^n \rho_i^2 K_i\right)^2 K_j^2 | \mathfrak{F}_{j-1}\right) + \mathbb{E}\left(\left(\sum_{i=1}^n |\rho_i| K_i\right)^2 \rho_j^2 K_j^2 | \mathfrak{F}_{j-1}\right)\right). \\ &\leq 2C h_J^k \left(C n^2 h_K^4 \mathbb{E}\left(K_j^2 | \mathfrak{F}_{j-1}\right) + C n^2 h_K^2 \mathbb{E}\left(\rho_j^2 K_j^2 | \mathfrak{F}_{j-1}\right)\right).\end{aligned}$$

This last inequality is obtained under (H.3) and (H.4) (i).
Next, applying of lemma 3.6.1 (i) allows us to get

$$\mathbb{E}\left(T_j^2 | \mathfrak{F}_{j-1}\right) \leq 2C' n^2 h_J^k h_K^4 \phi_{j,x}(h_K).$$

Now, we use the exponential inequality of Lemma 1 in [21] (with $d_j^2 = C' n^2 h_J^k h_K^4 \phi_{j,x}(h_K)$ to obtain for all $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}\left(|\widehat{f}_k^x(y) - \bar{f}_k^x(y)| > \varepsilon\right) &= \mathbb{P}\left(|\frac{1}{nh_J^k \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^n T_j| > \varepsilon\right) \\ &\leq 2 \exp\left\{-\frac{\varepsilon^2 n^2 h_J^{2k} (\mathbb{E}(\Gamma_1 K_1))^2}{2\left(D_n + C\varepsilon n h_J^k \mathbb{E}(\Gamma_1 K_1)\right)}\right\}.\end{aligned}$$

Taking $\varepsilon = \varepsilon_0 \sqrt{\dfrac{\varphi_x(h_K)\log n}{n^2 h_J^k \phi_x^2(h_K)}}$, then

$$\mathbb{P}\left(|\widehat{f}_k^x(y) - \bar{f}_k^x(y)| > \varepsilon_0 \sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J^k \phi_x^2(h_K)}} \right)$$

$$\leq \quad 2\exp\left\{ -\frac{n^2 h_J^{2k} \left(\mathbb{E}(\Gamma_1 K_1)\right)^2 \varepsilon_0^2 \dfrac{\varphi_x(h_K)\log n}{n^2 h_J^k \phi_x^2(h_K)}}{2\left(D_n + C n h_J^k \mathbb{E}(\Gamma_1 K_1)\varepsilon_0 \sqrt{\dfrac{\varphi_x(h_K)\log n}{n^2 h_J^k \phi_x^2(h_K)}} \right)} \right\}.$$

Now using lemma 3.6.1 (iii), allows us to write

$$\mathbb{P}\left(|\widehat{f}_k^x(y) - \bar{f}_k^x(y)| > \varepsilon_0 \sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J^k \phi_x^2(h_K)}} \right)$$

$$\leq \quad 2\exp\left\{ -\frac{n^2 h_J^{2k} \left(O\left(n h_K^2 \phi_x(h_K)\right)\right)^2 \varepsilon_0^2 \dfrac{\varphi_x(h_K)\log n}{n^2 h_J^k \phi_x^2(h_K)}}{2 n h_J^k h_K^2 \varphi_x(h_K) \left(C' n h_K^2 + O\left(n \phi_x(h_K)\right)\varepsilon_0 \sqrt{\dfrac{\log n}{n^2 h_J^k \phi_x^2(h_K)\varphi_x(h_K)}} \right)} \right\}$$

$$\leq \quad 2\exp\left\{ -\frac{n^2 h_J^{2k} \left(O\left(n h_K^2 \phi_x(h_K)\right)\right)^2 \varepsilon_0^2 \dfrac{\varphi_x(h_K)\log n}{n^2 h_J^k \phi_x^2(h_K)}}{2 n h_J^k h_K^2 \varphi_x(h_K) \left(C' n h_K^2 + O(1)\varepsilon_0 \sqrt{\dfrac{\log n}{h_J^k \varphi_x(h_K)}} \right)} \right\}.$$

Now using the fact that, under (H.1) (ii) and (iii), for all n we have $\varphi_x(h_K) \geq C n \phi_x(h_K)$ which implies that

$$\frac{\log n}{h_J^k \varphi_x(h_K)} \leq C' \frac{\varphi_x(h_K)\log n}{n^2 h_J^k \phi_x^2(h_K)}.$$

Therefore, under (H.5) , we have:

$$\lim_{n\to\infty} \frac{\log n}{h_J^k \varphi_x(h_K)} = 0.$$

It follows that

$$\mathbb{P}\left(|\widehat{f}_k^x(y) - \bar{f}_k^x(y)| > \varepsilon_0 \sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J^k \phi_x^2(h_K)}} \right) \leq 2n^{-C_0 \varepsilon_0^2},$$

where C_0 is a positive constant.

Consequently, using Borel-Cantelli's lemma and by choosing ε_0 large enough, we can deduce that:

$$\widehat{f}_k^x(y) - \bar{f}_k^x(y) = O_{a.co}\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J^k \phi_x^2(h_K)}}\right). \tag{3.7}$$

Finally, taking $k = 0$, this last result finish the proof of lemma 3.4.2.

Proof of Lemma 3.4.3

Remarked that, under (H.1)(iii) and (H.4), we have

$$0 < \frac{C}{n\phi_x(h_K)}\sum_{j=1}^{n}\mathbb{P}(X_j \in B(x,h_K)|\mathfrak{F}_{j-1}) \leq \bar{f}_0^x(y) \leq |\widehat{f}_0^x(y) - \bar{f}_0^x(y)| + \widehat{f}_0^x(y).$$

Therefore,

$$\begin{aligned}\mathbb{P}\left(\widehat{f}_0^x(y) \leq \frac{C}{2}\right) &\leq \mathbb{P}\left(\frac{C}{n\phi_x(h_K)}\sum_{j=1}^{n}\mathbb{P}(X_j \in B(x,h_K)|\mathfrak{F}_{j-1}) < \frac{C}{2} + |\widehat{f}_0^x(y) - \bar{f}_0^x(y)|\right)\\ &\leq \mathbb{P}\left(|\frac{C}{n\phi_x(h_K)}\sum_{j=1}^{n}\mathbb{P}(X_j \in B(x,h_K)|\mathfrak{F}_{j-1}) - |\widehat{f}_0^x(y) - \bar{f}_0^x(y)| - C| > \frac{C}{2}\right).\end{aligned}$$

It is obvious that lemma 3.4.2 and (H.1) (iii) allow to obtain

$$\sum_n \mathbb{P}\left(|\frac{C}{n\phi_x(h_K)}\sum_{j=1}^{n}\mathbb{P}(X_j \in B(x,h_K)|\mathfrak{F}_{j-1}) - |\widehat{f}_0^x(y) - \bar{f}_0^x(y)| - C| > \frac{C}{2}\right) < \infty,$$

which gives the result. ■

Proof of Lemma 3.4.4

The compactness of $\mathscr{C}$ permits us to deduce that there exists a sequence of real numbers $(y_k)_{k=1,\ldots,d_n}$ such that:

$$\mathscr{C} \subset \bigcup_{k=1}^{d_n}\mathscr{C}_k, \text{ where } \mathscr{C}_k = (y_k - l_n, y_k + l_n),$$

with $l_n = n^{-1-\alpha}$ and $d_n = O(l_n^{-1})$.
We start our proof with the following decomposition:

$$\begin{aligned}\sup_{y\in\mathscr{C}}|\widehat{f}_1^x(y) - \bar{f}_1^x(y)| &\leq \underbrace{\sup_{y\in\mathscr{C}}|\widehat{f}_1^x(y) - \widehat{f}_1^x(z)|}_{S_1} + \underbrace{\sup_{y\in\mathscr{C}}|\widehat{f}_1^x(z) - \bar{f}_1^x(z)|}_{S_2}\\ &+ \underbrace{\sup_{y\in\mathscr{C}}|\bar{f}_1^x(z) - \bar{f}_1^x(y)|}_{S_3}.\end{aligned}$$

Now, we establish the three terms.
On the one side, for the term S_1, by using assumption (H.5), we obtain:

$$\begin{aligned} S_1 \le & \sup_{y\in\mathscr{C}} |\frac{1}{nh_J \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \Gamma_j K_j |J_j(y) - J_j(z)||, \\ \le & \sup_{y\in\mathscr{C}} \frac{C|y-z|}{h_J} \left(|\frac{1}{nh_J \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \Gamma_j K_j| \right), \\ \le & C\frac{l_n}{h_J^2} |\widehat{f_0^x}(y)|. \end{aligned}$$

Thus, using lemma 3.4.3, we get :

$$S_1 \le C\frac{l_n}{h_J^2}.$$

Since $l_n = n^{-1-\alpha}$, we obtain:

$$\frac{l_n}{h_J^2} = o\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}} \right).$$

So, for n large enough, we find a $\eta > 0$ such that

$$\mathbb{P}\left(S_1 > \eta\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}} \right) = 0. \tag{3.8}$$

Similarly, for the term S_3, we obtain

$$S_3 \le C\frac{l_n}{h_J^2} |\bar{f_0^x}(y)|.$$

Therefore, lemma 3.6.2 allows us to write:

$$S_3 \le C\frac{l_n}{h_J^2}.$$

Using analogous arguments as S_1, we can found for n large enough:

$$\mathbb{P}\left(S_3 > \eta\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}} \right) = 0. \tag{3.9}$$

On the other side, to complete the proof of this lemma, we need to prove that:

$$S_2 = O_{a.co}\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}} \right).$$

By using (3.7) for $k = 1$, we get for $\eta > 0$ and for all $z \in \mathscr{C}_k$:

$$\mathbb{P}\left(|\widehat{f_1^x}(z) - \bar{f}_1^x(z)| > \eta\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right) \leq C' n^{-C_0\eta^2}.$$

Thus, we have

$$\begin{aligned}
&\mathbb{P}\left(\sup_{y\in\mathscr{C}}|\widehat{f_1^x}(z) - \bar{f}_1^x(z)| > \eta\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right)\\
\leq\ & \mathbb{P}\left(\max_{z\in\mathscr{C}_k}|\widehat{f_1^x}(z) - \bar{f}_1^x(z)| > \eta\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right)\\
\leq\ & 2d_n \max_{z\in\mathscr{C}_k}\mathbb{P}\left(|\widehat{f_1^x}(z) - \bar{f}_1^x(z)| > \eta\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right)\\
\leq\ & C' n^{-C_0\eta^2+1+\alpha}.
\end{aligned}$$

Therefore, by choosing η such that $C_0\eta^2 = 2 + 2\alpha$, we find

$$\mathbb{P}\left(\sup_{y\in\mathscr{C}}|\widehat{f_1^x}(z) - \bar{f}_1^x(z)| > \eta\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right) \leq C' n^{-1-\alpha}. \tag{3.10}$$

Finally, lemma 3.4.4 can be deduced directly from (3.8), (3.9) and (3.10). ■

Proof of Corollary 3.5.1

The unimodality of f^x and assumption (H.6) (ii) permit us to write that $f^{x(l)}(\Theta(x)) = f^{x(l)}(\widehat{\Theta}(x))) = 0$. Furthermore, by a Taylor expansion of the function f^x at $\Theta(x)$, we have:

$$f^x(\widehat{\Theta}(x)) = f^x(\Theta(x)) + \frac{1}{j!} f^{x(j)}(\Theta^*(x))\left(\widehat{\Theta}(x) - \Theta(x)\right)^j, \tag{3.11}$$

where $\Theta^*(x)$ is between $\Theta(x)$ and $\widehat{\Theta}(x)$.
Next, by simple manipulation we show that

$$|f^x(\widehat{\Theta}(x)) - f^x(\Theta(x))| \leq 2\sup_{y\in\mathscr{C}}|\widehat{f^x}(y) - f^x(y)|. \tag{3.12}$$

To end the proof of Corollary 3.5.1, we only need to show the following claim.
Claim

$$\lim_{n\to\infty}|\widehat{\Theta}(x) - \Theta(x)| = 0. \qquad a.co.$$

Proof. By the continuity of the function f^x, it follows that:

$$\forall \varepsilon > 0, \exists \delta(\varepsilon) > 0, |f^x(y) - f^x(\Theta(x)| \leq \delta(\varepsilon) \Rightarrow |y - \Theta(x)| \leq \varepsilon.$$

Then, this last consideration implies that:

$$\forall \varepsilon > 0, \exists \delta(\varepsilon) > 0, \mathbb{P}\left(|\widehat{\Theta}(x) - \Theta(x)| > \varepsilon\right) \leq \mathbb{P}\left(|f^x(\widehat{\Theta}(x)) - f^x(\Theta(x))| > \delta(\varepsilon)\right). \tag{3.13}$$

Lastly, the claimed result can be deduced by combining (3.13) with the the statement (3.12) and Theorem 3.4.2. □

Now, we return to the proof of Corollary 3.5.1.
Since $f^{x(j)}(\Theta^*(x)) \to f^{x(j)}(\Theta(x))$ and by using (H.6)(iii) , we obtain

$$\exists c > 0, \sum_{n=1}^{\infty} \mathbb{P}\left(|f^{x(j)}(\Theta^*(x))| < c\right) < \infty. \tag{3.14}$$

Therefore, we have

$$|\widehat{\Theta}(x) - \Theta(x)|^j = O\left(\sup_{y \in \mathscr{C}} |\widehat{f^x}(y) - f^x(y)|\right), \quad a.co.$$

We find this last result by combining the statements (3.11) and (3.12) with (3.14).
Finally, the proof of Corollary 3.5.1 can be easily deduced from Theorem 3.4.2 ■

BIBLIOGRAPHY

[1] Attouch, M. K., Chouaf, B., and Laksaci, A. Nonparametric M-Estimation for Functional Spatial Data. *Communications for Statistical Applications and Methods*, **19**, 193–211 (2012).

[2] Baìllo, A. and Grané, A. Local linear regression for functional predictor and scalar response. *J. of Multivariate Anal.*, **100**, 102–111 (2009).

[3] Barrientos, J., Ferraty, F. and Vieu, P. Locally Modelled Regression and Functional Data. *J. Nonparametr. Statist.*, **22**, 617-632 (2010).

[4] Bouanani, O., Laksaci, A., Rachdi, M. and Rahmani, S. Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics, *Behaviormetrika*, **4**, 1–35 (2018)

[5] Bouanani, O., Rahmani, S. and Ait-Hennani, L. Local linear conditional cumulative distribution function with mixing data. *Arab. J. Math.* https://doi.org/10.1007/s40065-019-0247-7 (2019).

[6] Chaouch, M., Laïb, N. and Ould Saïd, E. Nonparametric M-estimation for right censored regression model with stationary ergodic data. *Statistical Methodology,***33**, 234–255 (2016).

[7] Dabo-Niang, S., Kaid, Z., et Laksaci, A. Asymptotic properties of the kernel estimate of spatial conditional mode when the regressor is functional. *Advances in Statistical Analysis*, **99**, 131–160 (2015).

[8] Dabo-Niang, S. and Laksaci, A. Note on conditional mode estimation for functional dependent data, *Statistica*, **70**, 83–94 (2010).

[9] Damon, J. and Guillas, S. The inclusion of exogenous variables in functional autoregressive ozone forecasting, *Environmetrics* **13** , 759–774 (2002).

[10] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. *A fast functional locally modeled conditional density and mode for functional time-series.* Recent Advances in Functional Data Analysis and Related Topics, Contributions to Statistics, Pages 85–90, DOI: 10.1007/978-3-7908-2736-1_13, Physica-Verlag/Springer (2011).

[11] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. Functional data: local linear estimation of the conditional density and its application. *Statistics*, **47**, 26–44 (2013).

[12] Demongeot, J., Laksaci, A., Rachdi, M., and Rahmani, S. On the local linear modelization of the conditional distribution for functional data. *Sankhya A*, **76**, 328–355 (2014).

[13] Didi, S. and Louani, D. Asymptotic results for the regression function estimate on continuous time stationary and ergodic data. *Statistics and Risk Modeling*, **31**, 129–150 (2014).

[14] Fan, J. Design-adaptive nonparametric regression. *J. Amer. Statist. Assoc.*, **87**, 998–1004 (1992).

[15] Fan, J., Gijbels, I. *Local Polynomial Modelling and its Applications*. London, Chapman and Hall. (1996).

[16] Ferraty, F. and Vieu, P. Nonparametric functional data analysis theory and practice. Springer-Verlag (2006).

[17] Ferraty, F. and Romain, Y. The Oxford Handbook of Fuctional Data Analysis, *Oxford University Press* (2010).

[18] Gheriballah, A., Laksaci, A. and Sekkal, S. Nonparametric M-regression for functional ergodic data. *Statistics and Probability Letters.*, **83**, 902–908 (2013).

[19] Kaid, Z., and Laksaci, A. Functional quantile regression: local linear modelisation, *In Functional Statistics and Related Fields*. Springer, Cham. 155–160 (2017).

[20] Laïb, N., and Ould-Saïd, E. Estimation non paramétrique robuste de la fonction de régression pour des observations ergodiques. C. R. Acad. Sci. Série 1, **322**, 271–276 (1996).

[21] Laïb, N. and Louani D. Rates of strong consistencies of the regression function estimator for functional stationary ergodic data. *J. Stat. Plann. and Inf.*, **141**, p. 359–372 (2011).

[22] Laksaci, A., Rachdi, M. and Rahmani, S., Spatial modelization: local linear estimation of the conditional distribution for functional data. *Spat. Statist.*, **6**, 1–23 (2013).

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

[24] Ruppert, D. and Wand, M.P. Multivariate weighted least squares regression. *Ann. of Statist.*, **22**, 1346–1370 (1994).

[25] Zhou, Z. and Lin, Z.-Y. Asymptotic normality of locally modelled regression estimator for functional data. *J. Nonparametr. Statist.*, **28**, 116–131 (2016).

CHAPTER 4

ON LOCAL LINEAR ESTIMATION OF THE CONDITIONAL DENSITY FUNCTION FOR FUNCTIONAL ERGODIC DATA UNDER RANDOM CENSORSHIP MODEL.

This chapter is the subject of a submitted paper .

On local linear estimation of the conditional density function for functional ergodic data under random censorship model.

Somia Ayad [1], **Ali Laksaci** [2], **Saâdia Rahmani** [1] and **Rachida Rouane** [1].

[1] Laboratory of Stochastic Models, Statistics and Applications,University of Saida–Dr. Moulay Tahar, 20000, Algeria. E-mail: somia.ayad@univ-saida.dz. E-mail: saadia.rahmani@gmail.com. E-mail: rouane09@yahoo.fr

[2] Department of Mathematics, College of Science, King Khalid University, Abha, 61413, Saudi Arabia. E-mail:alilak@yahoo.fr.

Abstract : In this paper, we estimate the conditional density function of a randomly censored scalar response variable given a functional random variable. Furthermore, we suppose that the data are sampled from stationary ergodic process. We introduce a local linear type estimator of the conditional density function, and, we state the almost complete convergence with explicit rates of the constructed estimator. Moreover, the usefulness of our results is illustrated through their application to the conditional mode estimation.

Keywords : Ergodic data, functional data, censored data, local linear estimator, conditional density function, nonparametric estimation, Asymptotic properties.

Mathematics Subject Classification: 62G05, 62G08, 62G20, 62N01, 62N02.

4.1 Introduction and motivations

The nonparametric methods are the practical way to deal with the functional data. There are many fields where functional data are collected such as medicine, econometrics and environments.

In the last decade, many statistical researchers preferred to use the kernel method for studying nonparametric functional data due to its easiness of implementation. We can refer to the pioneer work of Ferraty and Vieu [18]. In this context, the authors of this monograph established the almost complete convergence of several estimators, they precise their convergence rates in the case when the observations are i.i.d. However, they have the feeling that in most conditional models some ideas of local polynomial smoothing could be extended to the functional context. Literature attention has focused on their famous question "How can the local polynomial ideas be adapted to infinite dimensional settings?". It has received special attention in the literature by several scenarios. Firstly and exactly in 2009, Baillo and Grané [2] constructed a local linear estimator for the regression operator in the case where the covariate takes its values in a Hilbert space. One year later, that question has been responded by the same authors who asked it in collaboration with Barrientos-Marin [3]. They have introduced an estimator of the regression function considering the case of a polynomial of order one called local linear approach which is flexible and more general than the kernel method. They studied the almost complete convergence with a rate of the proposed estimator. The case of the conditional distribution function was addressed by Demongeot et *al.* [11]. In this work, the authors studied the almost complete convergence as well as the mean square error with rates of the constructed estimator in the i.i.d case. Recently Bouanani et *al.* [5] and [6] established the asymptotic normality of several conditional models in the both cases: $\alpha-$mexing and independent. Because there are some differences from one method to another, Berlinet et *al.* [4] introduced another local linear estimator when the explanatory variable belongs in a Hilbert space. In all the papers which we mentioned, the results have been established in the case when the data are functional with complete observed response.

In the local constant approach when the data are functional but incomplete observed such as missing at random or censored, Khardani et *al.* [26] studied the almost sure and asymptotic normality of the kernel estimate of the conditional mode when the data are censored and independent. However, in the dependent case Horrigue and Ould Saïd [24] established the uniform strong consistency with rates of a kernel conditional quantile estimator. The asymptotic normality of the previous estimator was studied by the same authors in [24]. Ling et *al.* [31] constructed a kernel estimator of the regression operator for functional stationary ergodic data with the fact that the responses are missing at random and they established the convergence with rate in probability and the asymptotic normality of the constructed estimator. The case when the response variable is subject to left-truncation by another random variable was studied by Derrar et *al.* [13]. The asymptotic properties studied were the almost complete convergence and the asymptotic normality of their estimator. Recently, Fetitah et *al.* [21] established asymptotic properties of a kernel estimator of the relative error regression for randomly censored data. Concerning the local linear approach for incomplete data, there are few results such as Chahad et *al.* [8]. It should be noted that, to the best of our knowledge, no asymptotic results have been available in the literature for local linear conditional density fitting neither in the special framework considered in this paper nor for the general case of polynomial fitting for functional covariate when data are randomly censored and assumed to be sampled from a stationary and ergodic process. Moreover, as will become clear from the following sections, establishing the almost complete convergence for the functional local linear fitting estimators is technically more involved than for the classical kernel method.

In this work, we construct a new local linear estimator of the conditional density function which extends previous results established by Ayad et *al.* [1] to the censored case.

This paper is organized as follow. Section 4.2 introduces the construction of our local linear estimator. In Section 4.3, we introduce some notations and hypotheses which are needed to obtain our main results. The later is stated in the same section with an application of the conditional mode function. Finally, the detailed proofs of our theoretical results and all technical lemmas needed are gathered in the Appendix.

4.2 Local linear estimator construction

Let $(X_i, Y_i)_{i=1,\ldots,n}$ be a sequence of stationary and ergodic functional random samples identically distribution as (X, Y), where X takes its values in a semi metric space $\mathcal{F}$ equipped with a semi metric d and Y takes its values in $\mathbb{R}$. In the case of complete data, it is well known that the local linear estimator of the conditional density $f^X(Y)$ (See ayad et *al.* [1]) is explicitly given by:

$$\widehat{f}_n^x(y) = \frac{\sum_{j=1}^{n} \Gamma_j K_j J_j}{h_J \sum_{j=1}^{n} \Gamma_j K_j}, \tag{4.1}$$

where

$$\Gamma_j = \sum_{i=1}^{n} \rho_i^2 K_i - \left(\sum_{i=1}^{n} \rho_i K_i \right) \rho_j,$$

$$\text{where} \quad \rho_i = \rho(X_i, x), K_i = K\left(\frac{\delta(x, X_i)}{h_K}\right),\ J_j = J\left(\frac{y - Y_j}{h_J}\right)$$

with the convention $0/0 := 0$, $\rho(.,.)$ and $\delta(.,.)$ are known bi-functional operators defined from $\mathcal{F}^2$ into $\mathbb{R}$, where the bi-functional $\delta(.,.)$ is lied with the topological structure of the functional space $\mathcal{F}$, that means $|\delta(x,z)| = d(x,z)$. The bi-functional operator ρ controls the local shape of our model. The functions K and J are kernels, where the first one is a density function and the second one is a distribution function. $h_K = h_{K,n}$ (resp. $h_J = h_{J,n}$) is a sequence of positive real numbers called the smoothing parameter.

In the censoring case, we can only observe the triplets $(X_i, T_i, \delta_i)_{1 \le i \le n}$, where

$$T_i = Y_i \wedge C_i \quad \text{and} \quad \delta_i = \mathbb{I}_{\{Y_i \le C_i\}} \qquad 1 \le i \le n,$$

with $\mathbb{I}_A$ denoting the indicator function on a set A and C_i is the censoring random variable with unknown continuous distribution function G and a survival function $\bar{G}$ defined by

$\bar{G}(y) = 1 - G(y)$. In the same way, $(Y_i)_{1\le i\le n}$ is a sequence representing the survival time which has a common unknown continuous distribution function L, $(\forall y \in \mathbb{R}, L(y) = \mathbb{P}(Y \le y))$. $\bar{L}(y) = 1 - L(y)$ is its survival function. Furthermore, Let $\tau_L = \sup\{y,\ L(y) < 1\}$ (resp. $\tau_G = \sup\{y,\ G(y) < 1\}$) is the upper endpoint of $\bar{L}$ (resp. $\bar{G}$).
We assume that $(C_i)_{1\le i\le n}$ and $(X_i, Y_i)_{1\le i\le n}$ are independent. The "pseudo" estimator of $f^x(y)$ is defined by:

$$\widetilde{f^x}(y) = \frac{\sum_{j=1}^{n} \delta_j \bar{G}^{-1}(T_j)\Gamma_j K_j J_j}{h_J \sum_{j=1}^{n} \Gamma_j K_j} := \frac{\widehat{f_1^x}(y)}{\widehat{f_0^x}(y)}, \tag{4.2}$$

where

$$\widehat{f_l^x}(y) = \frac{1}{nh_J^l \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \delta_j^l \bar{G}^{-l}(T_j)\Gamma_j K_j J_j^l \quad \text{for} \quad l = 0, 1$$

Since G is unknown in practice, it is not possible to use the estimator (4.2). Thus, in order to obtain the following explicit formula of our local linear estimator, we use the *Kaplan Meier* [25] estimator of G given by:

$$\bar{G}_n(y) = \begin{cases} \prod_{j}^{n} \left(1 - \frac{1 - \delta_{(j)}}{n - j + 1}\right)^{\mathbf{1}_{\{T_{(j)} \le y\}}} & \text{if} \quad y < T_{(n)} \\ 0 & \text{Otherwise} \end{cases},$$

where $T_{(1)} < T_{(2)} < \ldots < T_{(n)}$ are order statistics of T_j and $\delta_{(j)}$ is concomitant with $T_{(j)}$.
Thus a feasible estimator of $f^x(y)$ is given by

$$\widehat{f^x}(y) = \frac{\sum_{j=1}^{n} \delta_j \bar{G}_n^{-1}(T_j)\Gamma_j K_j J_j}{h_J \sum_{j=1}^{n} \Gamma_j K_j} := \frac{\widehat{f_{1,n}^x}(y)}{\widehat{f_0^x}(y)}, \tag{4.3}$$

where

$$\widehat{f_{1,n}^x}(y) = \frac{1}{nh_J \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \delta_j \bar{G}_n^{-1}(T_j)\Gamma_j K_j J_j.$$

Then, We assume that there exists a certain compact set $\mathcal{C}_{\mathbb{R}} \subset \mathbb{R}$, such that $f^x(y)$ is unimodal and its conditional unique mode is denoted by $\Theta(x)$ on $\mathcal{C}_{\mathbb{R}}$. A natural and usual estimator of $\Theta(x)$ is defined by:

$$\widehat{\Theta}(x) = \arg\sup_{y \in \mathcal{C}_{\mathbb{R}}} \widehat{f^x}(y).$$

4.3 Assumptions and main results

4.3.1 Pointwise almost complete convergence:

The main aim of this section is to establish the pointwise almost complete convergence of the estimator $\widehat{f^x}(y)$ under some mild regularity conditions. Note that we only give the main results; detailed proofs can be found in the Appendix.

Throughout this paper, when no confusion is possible, we will denote by C and C' some strictly positive generic constants and we fix a point x in $\mathcal{F}$, (respectively, a compact $\mathcal{C}_{\mathbb{R}} \in \mathbb{R}$). Moreover, for $i = 1, \ldots, n$, let $\mathfrak{F}_i$ be the σ-field generated by $((X_1, Y_1), \ldots, (X_i, Y_i))$, and $\mathcal{G}_i$ the one generated by $((X_1, Y_1), \ldots, (X_i, Y_i), X_{i+1})$.

We assume that $\Theta(x) \in \mathcal{C}_{\mathbb{R}} \subset (-\infty, \tau]$ where $\tau < \tau_G \wedge \tau_L$ and $\phi_x(r_1, r_2) = \mathbb{P}(r_2 \leq \delta(X, x) \leq r_1)$ is the small ball probability function.

To state our results we need the following hypotheses:

(H.1) We suppose that the strictly stationary ergodic process $(X_i, Y_i)_{i\in\mathbb{N}^*}$ satisfies:

i) For all $r > 0$, $\phi_x(r) = \mathbb{P}(X \in B(x, r)) > 0$, where $B(x, r) = \{x' \in \mathcal{F}/|\delta(x', x)| < r\}$.

ii) For all $i = 1, \ldots, n$ there exist a deterministic function $\phi_{i,x}(.)$ such that $0 < \mathbb{P}(X_i \in B(x, r)|\mathfrak{F}_{i-1}) \leq \phi_{i,x}(r)$, $\forall r > 0$,

iii) For any $r > 0$, $\frac{1}{n\phi_x(r)} \sum_{i=1}^{n} \phi_{i,x}(r) \overset{P}{\to} 1$ and $n\phi_x(r) \to \infty$ as $r \to 0$.

(H.2) i) $\exists b_1 > 0$ and $b_2 > 0$, $\forall x' \in \mathcal{N}_x$ and $\forall y' \in \mathcal{C}_{\mathbb{R}}$:

$$|f^x(y) - f^{x'}(y')| \leq C\left(|\delta(x, x')|^{b_1} + |y - y'|^{b_2}\right).$$

ii) $f^x(.)$ is twice differentiable, its second derivative $f^{x(2)}(.)$ is continuous on a neighborhood of $\Theta(x)$, $f^{x(1)}(\Theta(x)) = 0$ and $f^{x(2)}(\Theta(x)) < 0$.

(H.3) The function ρ satisfies the following condition:
$\forall\, w \in \mathcal{F},\, C|\delta\,(x, w)\,| \leq |\rho\,(x, w)\,| \leq C'|\delta\,(x, w)\,|$.

(H.4) (i) K is a nonnegative bounded kernel supported on $[-1, 1]$.

(ii) J is a positive kernel, bounded and Lipschitzian continuous function, such that:

$$\int |v|^{b_2} J(v)dv < \infty \quad \text{and} \quad \int J^{(2)}(v)dv < \infty.$$

(iii) $\mathbb{E}\left(J(\frac{y-Y_j}{h_J})|\mathcal{G}_{j-1}\right) = \mathbb{E}\left(J(\frac{y-Y_j}{h_J})|X_j\right)$.

(H.5) The bandwidths h_K and h_J satisfy:

$$\lim_{n\to\infty} h_K = 0,\ \lim_{n\to\infty} n^{\alpha} h_J^{\gamma} = \infty \text{ with } \gamma = 1, 2, \text{ and } \alpha > 1,$$

$$\lim_{n\to\infty} \frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)} = 0, \text{ where } \varphi_x(h_K) = \sum_{i=1}^{n} \phi_{i,x}(h_K)$$

and

$$h_K \int_{B(x,h_K)} \rho(u,x)\, dP(u) = o\left(\int_{B(x,h_K)} \rho^2(u,x)\, dP(u)\right),$$

where $dP(x)$ is cumulative distribution of X.

Remarks on the hypotheses:

Firstly, hypothesis (H.1) is the same (H.1) in [1], then hypotheses (H.2) and (H.3) are mild regularity hypotheses on the conditional density function. Finally, conditions (H.4) and (H.5) are technical assumptions for the brevity of proofs.

Theorem 4.3.1. *Under the assumptions (H.1), (H.2) (i), (H.3)–(H.5), we have*

$$\sup_{y\in\mathcal{C}_{\mathbb{R}}} |\widehat{f}^x(y) - f^x(y)| = O\left(h_K^{b_1}\right) + O\left(h_J^{b_2}\right) + O\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right), \quad a.co.$$

Proof of Theorem 4.3.1

In order to prove Theorem 4.3.1, we introduce the following decomposition:

$$\widehat{f}^x(y) - f^x(y) = \widehat{f}^x(y) - \tilde{f}^x(y) + \tilde{f}^x(y) - f^x(y). \tag{4.4}$$

Next, for a sake of simplicity the following notation is needed:

$$\bar{\widehat{f}}_l^x(y) = \frac{1}{n h_J^l \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \mathbb{E}\left(\delta_j^l \bar{G}^{-l}(T_j)\Gamma_j K_j J_j^l | \mathfrak{F}_{j-1}\right), \qquad \text{with} \quad l = 0, 1.$$

Then, we can write:

$$\begin{aligned} \tilde{f}^x(y) - f^x(y) &= \left(\frac{\bar{\widehat{f}}_1^x(y)}{\bar{\widehat{f}}_0^x(y)} - f^x(y)\right) + \frac{1}{\widehat{f}_0^x(y)}\left[\left(\frac{\bar{\widehat{f}}_1^x(y)}{\bar{\widehat{f}}_0^x(y)} - f^x(y)\right)\left(\bar{\widehat{f}}_0^x(y) - \widehat{f}_0^x(y)\right)\right. \\ &+ \left.\left(\left(\widehat{f}_1^x(y) - \bar{\widehat{f}}_1^x(y)\right) - f^x(y)\left(\widehat{f}_0^x(y) - \bar{\widehat{f}}_0^x(y)\right)\right)\right]. \end{aligned} \tag{4.5}$$

Thus, the proof of Theorem 4.3.1 is a direct consequence of Lemma 1 of Louani and Laib [27], Lemmas 3 and 5 of Ayad et *al.* and the following auxiliary results which play a main role and for which proofs are given in the appendix.

Lemma 4.3.1. *Assume that (H.1), (H.2)(i), (H.3) and (H.4) are satisfied. then, we have*

$$\sup_{y\in\mathcal{C}_{\mathbb{R}}}\left|\left(\frac{\bar{\widehat{f}}_1^x(y)}{\bar{\widehat{f}}_0^x(y)} - f^x(y)\right)\right| = O\left(h_K^{b_1}\right) + O\left(h_J^{b_2}\right).$$

Lemma 4.3.2. *Under the assumptions (H.1), (H.3)–(H.5), we have*

$$\widehat{f}_0^x(y) - \bar{\widehat{f}}_0^x(y) = O_{a.co}\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2\phi_x^2(h_K)}}\right). \tag{4.6}$$

And

$$\sup_{y\in\mathcal{C}_{\mathbb{R}}}\left|\left(\widehat{f}_1^x(y) - \bar{\widehat{f}}_1^x(y)\right)\right| = O_{a.co}\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2h_J\phi_x^2(h_K)}}\right). \tag{4.7}$$

Lemma 4.3.3. *Under the conditions (H.1), (H.2)(i), (H.3)–(H.5), we have*

$$\sup_{y\in\mathcal{C}_{\mathbb{R}}}\left|\widehat{f}^x(y) - \tilde{f}^x(y)\right| = O_{a.co}\left(\sqrt{\frac{\log\log n}{n}}\right).$$

4.3.2 Application to conditional mode

In addition to the assumptions introduced along the previous section, we need the following conditions to establish the consistency of the conditional mode estimator:

(H.6) There exists $\Theta(x)\in\mathcal{C}_{\mathbb{R}}$, such that $f^x(y) < f^x(\Theta(x))$, for all $y\neq\Theta(x),\ y\in\mathcal{C}_{\mathbb{R}}$

Theorem 4.3.2. *Assume that (H.1)–(H.6) hold, we have*

$$|\widehat{\Theta}(x) - \Theta(x)| = O\left(h_K^{\frac{b_1}{2}}\right) + O\left(h_J^{\frac{b_2}{2}}\right) + O\left(\left(\frac{\varphi_x(h_K)\log n}{n^2h_J\phi_x^2(h_K)}\right)^{\frac{1}{4}}\right), \quad a.co.$$

Proof of Theorem 4.3.2

Before starting the proof of this last Theorem, the following lemma is necessary:

Lemma 4.3.4.

$$\lim_{n\to\infty}|\widehat{\Theta}(x) - \Theta(x)| = 0. \qquad a.co.$$

Proof. Since $f^x(.)$ is uniformly continuous on $\mathcal{C}_{\mathbb{R}}$, it is easy to see that (H.6) implies that:

$$\forall\varepsilon>0, \exists\eta(\varepsilon)>0, |f^x(y) - f^x(\Theta(x)| \leq \eta(\varepsilon) \Rightarrow |y-\Theta(x)|\leq\varepsilon.$$

Which implies that:

$$\forall\varepsilon>0, \exists\eta(\varepsilon)>0, \mathbb{P}\left(|\widehat{\Theta}(x) - \Theta(x)|>\varepsilon\right) \leq \mathbb{P}\left(|f^x(\widehat{\Theta}(x)) - f^x(\Theta(x))|>\eta(\varepsilon)\right). \tag{4.8}$$

Then, by a simple algebra, we get

$$|f^x(\widehat{\Theta}(x)) - f^x(\Theta(x))| \leq 2 \sup_{y \in \mathcal{C}_{\mathbb{R}}} |\widehat{f}^x(y) - f^x(y)|. \tag{4.9}$$

Finally, the almost complete convergence of $\widehat{\Theta}(x)$ to $\Theta(x)$ can be deduced from the latter together with (4.8) and Theorem 4.3.1. □

Next, since $f^{x(1)}(\Theta(x)) = f^{x(1)}(\widehat{\Theta}(x)) = 0$ and by a Taylor expansion of the function f^x, we have:

$$f^x(\widehat{\Theta}(x)) = f^x(\Theta(x)) + \frac{1}{2} f^{x(2)}(\Theta^*(x)) \left(\widehat{\Theta}(x) - \Theta(x)\right)^2, \tag{4.10}$$

where $\Theta^*(x)$ is between $\Theta(x)$ and $\widehat{\Theta}(x)$. So, by using (H.2)(ii), we obtain

$$\exists c > 0, \sum_{n=1}^{\infty} \mathbb{P}\left(|f^{x(2)}(\Theta^*(x))| < c\right) < \infty. \tag{4.11}$$

Therefore, by combining the statements (4.9), (4.10) and (4.11) we obtain:

$$|\widehat{\Theta}(x) - \Theta(x)|^2 = O\left(\sup_{t \in \mathcal{C}_{\mathbb{R}}} |\widehat{f^x}(y) - f^x(y)|\right), \quad a.co.$$

Thus, the proof of Theorem 4.3.2 can be deduced from Theorem 4.3.1.

4.4 Appendix

4.4.1 Proof of Lemma 4.3.1

Firstly, remind that

$$\frac{\bar{\widehat{f}}{}_1^x(y)}{\bar{\widehat{f}}{}_0^x(y)} - f^x(y) = \frac{\bar{\widehat{f}}{}_1^x(y) - \bar{\widehat{f}}{}_0^x(y) f^x(y)}{\bar{\widehat{f}}{}_0^x(y)},$$

(H.4)(iii) combined with the property of the conditional expectation with respect to the σ-fields $\mathcal{G}_{j-1}$ and Y_j with the fact that $\mathbb{1}_{Y_j < C_j}\psi(T_j) = \mathbb{1}_{\{Y_j < C_j\}}\psi(Y_j)$, we obtain

$$\begin{aligned}
\bar{\widehat{f}}{}_1^x(y) &= \frac{1}{nh_J \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \mathbb{E}\left(\delta_j \bar{G}^{-1}(T_j)\Gamma_j K_j J_j | \mathfrak{F}_{j-1}\right) \\
&= \frac{1}{nh_J \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \mathbb{E}\left(\Gamma_j K_j \mathbb{E}\left(\delta_j \bar{G}^{-1}(T_j) J_j | \mathcal{G}_{j-1}, Y_j\right) | \mathfrak{F}_{j-1}\right) \\
&= \frac{1}{nh_J \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \mathbb{E}\left(\Gamma_j K_j \bar{G}^{-1}(T_j) J_j \mathbb{E}\left(\mathbb{1}_{\{Y_j < C_j\}} | X_j, Y_j\right) | \mathfrak{F}_{j-1}\right) \\
&= \frac{1}{nh_J \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \mathbb{E}\left(\Gamma_j K_j J_j | \mathfrak{F}_{j-1}\right) \qquad (4.12)
\end{aligned}$$

Furthermore, a double conditioning with respect to $\mathcal{G}_{j-1}$ leads to

$$\bar{\hat{f}}_1^x(y) - \bar{\hat{f}}_0^x(y) f^x(y) = \frac{1}{nh_J \mathbb{E}(\Gamma_1 K_1) \bar{\hat{f}}_0^x(y)} \sum_{j=1}^{n} \mathbb{E}\left(\Gamma_j K_j \left|\mathbb{E}[J_j|X_j] - h_J f^x(y)\right| |\mathfrak{F}_{j-1}\right).$$

Then, using an integration by parts followed by a change of variable, permits to get

$$\mathbb{E}(J_j|X_j) = h_J \int_{\mathbb{R}} J(u) f^x(y - h_J u)\, du, \tag{4.13}$$

thus, we have

$$\left|\mathbb{E}[J_j|X_j] - h_J f^x(y)\right| \le h_J \int_{\mathbb{R}} J(u) \left|f^x(y - h_J u) - f^x(y)\right| du.$$

Using (H.2)(i) permits us to find:

$$1\!\!1_{B(x,h_K)}(X_j) \left|\mathbb{E}[J_j|X_j] - h_J f^x(y)\right| \le h_J \int_{\mathbb{R}} J(u) \left(h_K^{b_1} + |y|^{b_2} h_J^{b_2}\right) du.$$

Hence, by assumption (H.4) (ii) and Lemma 6 of [1], we can obtain

$$\begin{aligned} \bar{\hat{f}}_1^x(y) - \bar{\hat{f}}_0^x(y) f^x(y) &= \left(O\left(h_K^{b_1}\right) + O\left(h_J^{b_2}\right)\right) \times \frac{1}{n\mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \mathbb{E}(\Gamma_j K_j | \mathfrak{F}_{j-1}) \\ &= O\left(h_K^{b_1}\right) + O\left(h_J^{b_2}\right). \end{aligned}$$

■

4.4.2 Proof of Lemma 4.3.2

For all $l = 0, 1$, we have

$$\begin{aligned} \hat{f}_l^x(y) - \bar{\hat{f}}_l^x(y) &= \frac{1}{nh_J^l \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} \delta_j^l \bar{G}^{-l}(T_j)\Gamma_j K_j J_j^l(y) - \mathbb{E}\left(\delta_j^l \bar{G}^{-l}(T_j)\Gamma_j K_j J_j^l(y) | \mathfrak{F}_{j-1}\right) \\ &= \frac{1}{nh_J^l \mathbb{E}(\Gamma_1 K_1)} \sum_{j=1}^{n} L_j(x, y), \end{aligned}$$

where $L_j(x, y)$ is a triangular array of martingale differences according to the σ- fields$(\mathfrak{F}_{j-1})_j$. Similar to the proof of Lemma 2 of [1] and under (H.1), (H.3) and (H.4), we can write

$$\mathbb{E}\left(L_j^2(x, y) | \mathfrak{F}_{j-1}\right) \le 2C' n^2 h_J^l h_K^4 \phi_{j,x}(h_K).$$

Then, we apply the exponential inequality of Lemma 1 in [27] (with $d_j^2 = C' n^2 h_J^l h_K^4 \phi_{j,x}(h_K)$ to obtain for $\varepsilon > 0$:

$$\mathbb{P}\left(|\widehat{f_l^x}(y) - \bar{\widehat{f_l^x}}(y) > \varepsilon\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J^l \phi_x^2(h_K)}}\right)$$

$$\begin{aligned}
&\leq \mathbb{P}\left(|L_j(x,y)| > n\varepsilon h_J^l \mathbb{E}(\Gamma_1 K_1)\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J^l \phi_x^2(h_K)}}\right)\\
&\leq 2\exp\left\{-\frac{n^2 h_J^{2l}(\mathbb{E}(\Gamma_1 K_1))^2 \varepsilon^2 \frac{\varphi_x(h_K)\log n}{n^2 h_J^l \phi_x^2(h_K)}}{2\left(D_n + Cnh_J^l \mathbb{E}(\Gamma_1 K_1)\varepsilon\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J^l \phi_x^2(h_K))}}\right)}\right\}.
\end{aligned}$$

Then, using Lemma 5(iii) of [1] leads:

$$\mathbb{P}\left(|\widehat{f_l^x}(y) - \bar{\widehat{f_l^x}}(y)| > \varepsilon\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J^l \phi_x^2(h_K)}}\right) \leq 2n^{-C'\varepsilon^2}.$$

Therefore, by using Borel-Cantelli's Lemma and by choosing ε large enough, we find:

$$\widehat{f_l^x}(y) - \bar{\widehat{f_l^x}}(y) = O_{a.co}\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J^l \phi_x^2(h_K)}}\right). \tag{4.14}$$

Finally, the proof of the first part of Lemma 4.3.2 can be deduced by replacing l by 0 in Equation (4.14).

Now, if we use the compactness of $\mathcal{C}_{\mathbb{R}}$, we can write $\mathcal{C}_{\mathbb{R}} \subset \bigcup_{k=1}^{d_n} \mathcal{C}_k$, where

$$\mathcal{C}_k = (y_k - l_n, y_k + l_n), \text{ with } l_n = n^{-1-\alpha} \text{ and } l_n d_n = O(1).$$

Thus, we obtain

$$\begin{aligned}
\sup_{y\in\mathcal{C}_{\mathbb{R}}} |\widehat{f_1^x}(y) - \bar{\widehat{f_1^x}}(y)| &\leq \underbrace{\sup_{y\in\mathcal{C}_{\mathbb{R}}} |\widehat{f_1^x}(y) - \widehat{f_1^x}(z)|}_{\mathscr{C}_1} + \underbrace{\sup_{y\in\mathcal{C}_{\mathbb{R}}} |\widehat{f_1^x}(z) - \bar{\widehat{f_1^x}}(z)|}_{\mathscr{C}_2}\\
&+ \underbrace{\sup_{y\in\mathcal{C}_{\mathbb{R}}} |\bar{\widehat{f_1^x}}(z) - \bar{\widehat{f_1^x}}(y)|}_{\mathscr{C}_3}.
\end{aligned}$$

For the term $\mathscr{C}_1$, by using assumptions (H.4)(ii) and (H.5), we obtain:

$$\begin{aligned}
\mathscr{C}_1 \leq & \sup_{y\in\mathcal{C}_{\mathbb{R}}} \Big|\frac{1}{nh_J\mathbb{E}(\Gamma_1 K_1)}\sum_{j=1}^{n}\delta_j\bar{G}^{-1}(T_j)\Gamma_j K_j|J_j(y)-J_j(z)|\Big|,\\
\leq & \sup_{y\in\mathcal{C}_{\mathbb{R}}}\frac{C|y-z|}{h_J}\left(\Big|\frac{1}{nh_J\mathbb{E}(\Gamma_1 K_1)}\sum_{j=1}^{n}\bar{G}^{-1}(T_j)\Gamma_j K_j\Big|\right),\\
\leq & C\frac{l_n}{h_J^2}|\widehat{f}_0^x(t)|.
\end{aligned}$$

Next, using Lemma 3 of [1] allows to have:

$$\mathscr{C}_1 \leq C\frac{l_n}{h_J^2}.$$

Since $l_n = n^{-1-\alpha}$ and by using the first part of (H.5), we obtain:

$$\frac{l_n}{h_J^2} = o\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right).$$

So, for n large enough, we have

$$\mathscr{C}_1 = O_{a.co}\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right). \tag{4.15}$$

Similarly, for $\mathscr{C}_3$, by using Formula (4.12) we obtain:

$$\mathscr{C}_3 \leq C\frac{l_n}{h_J^2}|\bar{f}_0^x(t)|.$$

Therefore, by using Lemma 6 of [1], we get:

$$\mathscr{C}_3 \leq C\frac{l_n}{h_J^2}.$$

Using similar arguments as $\mathscr{C}_1$, we can obtain for n large enough:

$$\mathscr{C}_3 = O_{a.co}\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right). \tag{4.16}$$

Concerning $\mathscr{C}_2$, by using (4.14) for $l=1$, we get for $\varepsilon_0 > 0$ and for all $z \in \mathcal{C}_k$:

$$\mathbb{P}\left(|\widehat{f}_1^x(z) - \bar{\widehat{f}}_1^x(z)| > \varepsilon_0\sqrt{\frac{\varphi_x(h_K)\log n}{n^2 h_J \phi_x^2(h_K)}}\right) \leq C' n^{-C_1\varepsilon_0^2}$$

Therefore, we have

$$\mathbb{P}\left(\sup_{y\in\mathcal{C}_{\mathbb{R}}}|\widehat{f}_1^x(z)-\bar{\widehat{f}}_1^x(z)|>\varepsilon_0\sqrt{\frac{\varphi_x(h_K)\log n}{n^2h_J\phi_x^2(h_K)}}\right)$$

$$\begin{aligned}
&\leq \mathbb{P}\left(\max_{z\in\mathcal{C}_k}|\widehat{f}_1^x(z)-\bar{\widehat{f}}_1^x(z)|>\varepsilon_0\sqrt{\frac{\varphi_x(h_K)\log n}{n^2h_J\phi_x^2(h_K)}}\right)\\
&\leq 2d_n\max_{z\in\mathcal{C}_k}\mathbb{P}\left(|\widehat{f}_1^x(z)-\bar{\widehat{f}}_1^x(z)|>\varepsilon_0\sqrt{\frac{\varphi_x(h_K)\log n}{n^2h_J\phi_x^2(h_K)}}\right)\\
&\leq C'n^{-C_1\varepsilon_0^2+1+\alpha}.
\end{aligned}$$

By choosing ε_0 such that $C_0\varepsilon_0^2=2+2\alpha$, we find

$$\mathbb{P}\left(\sup_{y\in\mathcal{C}_{\mathbb{R}}}|\widehat{f}_1^x(z)-\bar{\widehat{f}}_1^x(z)|>\eta\sqrt{\frac{\varphi_x(h_K)\log n}{n^2h_J\phi_x^2(h_K)}}\right)\leq C'n^{-1-\alpha}.$$

Then, by Borel-Cantelli's Lemma, we get

$$\mathscr{C}_2=O_{a.co}\left(\sqrt{\frac{\varphi_x(h_K)\log n}{n^2h_J\phi_x^2(h_K)}}\right). \tag{4.17}$$

Finally, The second part of Lemma 4.3.2 can be deduced directly from the results(4.15), (4.16) and (4.17). ∎

4.4.3 Proof of Lemma 4.3.3

From the explicit formulas (4.2) and (4.3) and by using Lemma 4.3.2, we have

$$\begin{aligned}
\sup_{y\in\mathcal{C}_{\mathbb{R}}}\left|\widehat{f}^x(y)-\tilde{f}^x(y)\right| &\leq \sup_{y\in\mathcal{C}_{\mathbb{R}}}\left|\frac{1}{nh_J\mathbb{E}(\Gamma_1K_1)}\sum_{j=1}^n\delta_j\Gamma_jK_jJ_j\left(\frac{1}{\bar{G}_n(T_j)}-\frac{1}{\bar{G}(T_j)}\right)\right|\\
&\leq \frac{\sup_{y\in\mathcal{C}_{\mathbb{R}}}\left|\bar{G}_n(y)-\bar{G}(y)\right|}{\bar{G}_n(\tau)}\left|\frac{1}{nh_J\mathbb{E}(\Gamma_1K_1)}\sum_{j=1}^n\delta_j\Gamma_jK_jJ_j\bar{G}^{-1}(T_j)\right|\\
&= \frac{\sup_{y\in\mathcal{C}_{\mathbb{R}}}\left|\bar{G}_n(y)-\bar{G}(y)\right|}{\bar{G}_n(\tau)}\left|\tilde{f}(x,y)\right|
\end{aligned}$$

Since $\bar{G}(\tau)>0$, in conjunction with the strong law of large numbers and the law of the iterated logarithm on the censoring law (see formula (4.28) in Deheuvels and Einmahl [12], 2000), the result is an immediate consequence of Lemmas 4.3.1 and 4.3.2. ■

BIBLIOGRAPHY

[1] Ayad, S., Laksaci, A., Rahmani, S. and Rouane, R. On the local linear modelization of the conditional density for functional and ergodic data. *METRON.*, **78**, 237–254 (2020).

[2] Baìllo, A. and Grané, A. Local linear regression for functional predictor and scalar response. *J. of Multivariate Anal.*, **100**, 102–111(2009).

[3] Barrientos, J., Ferraty, F. and Vieu, P. Locally Modelled Regression and Functional Data. *J. Nonparametr. Statist.*, **22**, 617-632 (2010).

[4] Berlinet, A., Elamine, A. and Mas, A. Local linear regression for functional data, *Annals of the Institute of Statistical Mathematics*, **63**, 1047–1075 (2011)

[5] Bouanani, O., Laksaci, A., Rachdi, M. and Rahmani, S. Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics, *Behaviormetrika*, **4**, 1–35 (2018)

[6] Bouanani, O., Rahmani, S. and Ait-Hennani, L. Local linear conditional cumulative distribution function with mixing data. *Arab. J. Math.* https://doi.org/10.1007/s40065-019-0247-7 (2019).

[7] Carbonez, A., Györfi, L., and vander Meulen, EC. Partitioning-estimates of a regression function under random censoring. *Statistics & Risk Modeling* **13**, 21–38 (1995).

[8] Chahad, A., Laksaci, A. and Ait-Hennani, L. Functional local linear relative regression. *2020 2nd International Conference on Mathematics and Information Technology (ICMIT),* 102–108 (2020).

[9] Chaouch, M., Laïb, N. and Ould Saïd, E. Nonparametric M-estimation for right censored regression model with stationary ergodic data. *Statistical Methodology,***33**, 234–255 (2016).

[10] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. Functional data: local linear estimation of the conditional density and its application. *Statistics*, **47**, 26–44 (2013).

[11] Demongeot, J., Laksaci, A., Rachdi, M., and Rahmani, S. On the local linear modelization of the conditional distribution for functional data. *Sankhya A*, **76**, 328–355 (2014).

[12] Deheuvels, P. and Einmahl, H. Functional limit laws for the increments of Kaplan-Meier product-limit processes and applications, 1301–1335 (2000).

[13] Derrar, S., Laksaci, A., and Ould Saïd, E. On the nonparametric estimation of the functional ψ-regression for a random left-truncation model. *Journal of statistical Theory and Practice*, **4**, 3823–849 (2015).

[14] El Methni, M. and Rachdi, M. Local weighted average estimation of the regression operator for functional data. *Communications in Statistics-Theory and Methods*, **40**, 3141–3153 (2011).

[15] Fan, J. Design-adaptive nonparametric regression. *J. Amer. Statist. Assoc.*, **87**, 998–1004 (1992).

[16] Fan, J., Gijbels, I. *Local Polynomial Modelling and its Applications*. London, Chapman and Hall. (1996).

[17] Fan, J. and Yao, Q. *Nolinear Time Series: Nonparametric and Parametric Methods.* Springer-Verlag, New York.(2003).

[18] Ferraty, F. and Vieu, P. Nonparametric functional data analysis theory and practice. Springer-Verlag (2006).

[19] Ferraty, F. and Romain, Y. The Oxford Handbook of Fuctional Data Analysis, *Oxford University Press* (2010).

[20] Ferraty, F., Laksaci, A., Tadj. A. and Vieu, P. Rate of uniform consistency for nonparametric estimates with functional variables. *J. of Statistical Planning and Inference.*, **140**, 335–352 (2010).

[21] Fetitah, O., Almanjahie, I. M and Attouch, M.K. and Righi, A.Strong convergence of the functional nonparametric relative error regression estimator under right censoring. *Mathematica Slovaca.*, **70**, 1469–1490 (2020).

[22] Gheriballah, A., Laksaci, A. and Sekkal, S. Nonparametric M-regression for functional ergodic data. *Statistics and Probability Letters.*, **83**, p. 902–908 (2013).

[23] Horrigue, W. and Ould Saïd, E. Strong uniform consistency of a nonparametric estimator of a conditional quantile for censored dependent data and functional regressors. *Random Operators & Stochastic Equations.*, **19**, p. 131–156 (2011).

[24] Horrigue, W. and Ould Saïd, E. Non Parametric Regression Quantile Estimation for Dependent Functional Data under Random Censorship: Asymptotic Normality. *Communications in Statistics-Theory and Methods.*, **44**, p. 4307–4332 (2015).

[25] Kaplan, EL., and Meier, P. Nonparametric estimation from incomplete observations. *Journal of the American statistical association* **53**, 457–481 (1958).

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

[27] Laïb, N. and Louani D. Rates of strong consistencies of the regression function estimator for functional stationary ergodic data. *J. Stat. Plann. and Inf.*, **141**, p. 359–372 (2011).

[28] Laksaci, A., Rachdi, M. and Rahmani, S., Spatial modelization: local linear estimation of the conditional distribution for functional data. *Spat. Statist.*, **6**, 1–23 (2013).

[29] Leulmi, S., Local linear estimation of the conditional quantile for censored data and functional regressors. *Communications in Statistics-Theory and Methods.*, , 1–15 (2019).

[30] Leulmi, S., Nonparametric local linear regression estimation for censored data and functional regressors. *Journal of the Korean Statistical Society.*, , 1–22 (2020).

[31] Ling, N. X., Liang, L. L. and Vieu, P..Nonparametric regression estimation for functional stationary ergodic data with missing at random. *Journal of Statistical Planning and Inference.*, **162**, 75–87 (2015).

[32] Ruppert, D. and Wand, M.P. Multivariate weighted least squares regression. *Ann. of Statist.*, **22**, 1346–1370 (1994).

[33] Zhou, Z. and Lin, Z.-Y. Asymptotic normality of locally modelled regression estimator for functional data. *J. Nonparametr. Statist.*, **28**, 116–131 (2016).

GENERAL CONCLUSION AND PROSPECTS

Conclusion

This thesis completes the recent advances existing in the local linear estimation for functional data, by giving not only the rates of convergence of the estimates but also the exact expressions of the constant terms involved in these rates which are expressed by means of some functional probability function such as the small ball probability function and its variation function. These theoretical results deal with the pointwise and uniform almost complete consistency as well as the asymptotic normality of several estimators related to the conditional cumulative distribution when the response is a scalar (censored or not) and the data are observed as ergodic functional times series.

The treatment of a dependence structure is of particular interest in the nonparametric estimation because of the possible applications in time series analysis and prediction problems. Although the widely use of the strong mixing condition and its variants to measure the dependency, it involves complicated probabilistic calculations. Moreover, several models which are given in the literature where mixing properties are still to be verified or even fail to hold for the process they induce. Therefore, we consider in the present thesis, the ergodic property to allow the maximum possible generality with regard to the dependence setting.

The main source of difficulty when dealing with local linear estimators consists in the fact that they are not written directly as one or more sums of independent and identically distributed variables. It is worth noting that the technical proofs for studying this kind of estimators (in the finite dimensional setup) are different; so we have opted for a "direct" way of giving relatively complex decompositions (for the consistence and the normality) of our estimators as sums of independent and identically distributed variables. These decompositions allow us to prove our results by a very traditional way. Moreover, the fast and the practical method introduced in this work is given by defining an optimization problem over real space which can be seen as the easiest estimation method. This is due to the fact that the other functional approaches are based

on an optimization problem in a functional space. However, in our work, we established our asymptotic results by optimization problems over a real space. This highlights another advantage of our method. As already mentioned above, this has been achievedbecause of the good choice of the decomposition, the assumptions that are detailed in Chapters 2, 3, 4; and the use of some basic probabilistic tools and the central limit theorems cited in the Appendix A.
It is well known that local polynomial fitting is an attractive method both from theoretical and practical point of view. This is confirmed by the different results obtained in the present thesis. Moreover, our practical studies confirm the superiority of our method on the kernel approach.

As it can be seen through the proofs presented in this thesis that the almost complete convergence has double advantages; it is in some sense easier to state than the almost sure one. Moreover, this mode of convergence implies other standard modes of convergence. This is why it became quite usual for many statisticians to express their asymptotic results in terms of complete convergence (pointwise or uniform).

On the other hand, our results validate the use of the asymptotic normality in the implementation of hypotheses tests and confidence interval estimation.

Prospects

To conclude the work of this thesis, we present in the following, some possible future developments in order to improve and extend our results.

- *Hight order of the polynom:* Our results confirm the superiority of the local linear method over the classical kernel method. Indeed, the kernel method provides a bias of order $O(h)$, whereas it is of order $O(h^2)$ for the local linear method. A naturel prospect of the present work is the functional local polynomial estimation of the conditional distribution function and its derivatives. Such a generalization can be obtained by reenforcing the regularity conditions of the model under study in order to approximate locally this function by a polynomial operator. Therefore, this estimate increases the gain in bias term over the classical kernel method. The explicit determination of the estimator can be obtained by a straightforward modification of the present approach.
- *Other models and/or other methods:*
  - *Estimation of the conditional quantile:* Unlike to the classical functional Nadaraya-Watson estimator, the local linear estimator of the conditional cumulative distribution function is not, necessarily, an increasing function. So, its inverse is very difficult to achieve in practice. In order to overcome the problem raised above we think

that it is possible to built an alternative estimator of the conditional quantile which is based on the L_1 approach. The constructed estimator will keep the robustness of the quantile regression function and advantages of the local linear method.

– *The robust regression function and the relative error:* The methodology is general, so, (except the quantile), it can be applied to any model, where a predictor X is functional, a response Y is scalar, and the data are observed as ergodic functional times series. So, it would be very important, in the future, to investigate the local linear estimation, particularly, of the robust regression function and the relative error.

– It is possible to generalize our results using other models such as the additive model, the semi-functional partial linear model, or the simple functional index model.

• *Data-driven automatic bandwidth selection:* The selection of the bandwidth parameter h_K and then h_J plays a crucial role. Indeed, a too large smoothing under-parametrizes the model (the CDF, the conditional density...), causing a large modelling bias, while a too small smoothing parameter over-parametrizes the unknown function and results in noisy estimates. Optimal theoretical choice of this parameter is obtained by minimizing the conditional Mean Squared Error (MSE). However, this theoretical choice is not directly practically usable since it depends on unknown quantities. Finding a practical technique for selecting the bandwidth parameter is one of the most important tasks. Furthermore, in our practical studies, the optimal bandwidths were chosen by the cross-validation procedure. One possible approach is to substitute the unknown quantities by pilot estimators, leading to so-called "plug-in" type bandwidth selectors. Another alternative approach for selecting the smoothing parameter is the functional version of wild bootstrapping ideas (see for instance Ferraty et al. [57]).

• *Missing data:* In a missing-data setting, we estimate the mean of a scalar outcome, based on a sample in which an explanatory variable is observed for every subject while responses are missing by happenstance for some of them. This situation is pervasive in most data samples, there is a need to extend our results to this kind of data.
It is quite possible to generalize our results in the case when the data are uncomplete (truncated).

APPENDIX A

SOME PROBABILISTIC TOOLS

A.1 Martingale differences sequence

Definition A.1.1. *[72]*
A sequence of random variables $(Z_n)_{n\geq 1}$ *is said to be a sequence of martingale differences (MDS) with respect to the sequence of* $\sigma-$ *fields* $(\mathfrak{F})_{n\geq 1}$ *whenever* Z_n *is* $\mathfrak{F}_n$ *measurable and* $\mathbb{E}\left(Z_n|\mathfrak{F}_{n-1}\right) = 0, \qquad a.co.$

The MDS is an extremely useful construct in modern probability theory because it implies much milder restrictions on the memory of the sequence than independence, yet most limit theorems that hold for an independent sequence will also hold for an MDS.

All along this thesis, we need an exponential inequality for partial sums of unbounded martingale differences that we use to derive asymptotic results for the local linear estimate built upon functional ergodic data. This inequality is given in the following Lemma.

Lemma A.1.1. *[72]*
Let $(Z_n)_{n\geq 1}$ *be a sequence of real martingale differences with respect to the sequence of* $\sigma-$*fields* $(\mathfrak{F}_n = \sigma(Z_1, \ldots, Z_n))_{n\geq 1}$, *where* $\sigma(Z_1, \ldots, Z_n)$ *is the* $\sigma-$*field generated by the random variables* $Z_1, \ldots, Z_n$. *Set* $S_n = \sum_{i=1}^{n} Z_i$. *For any* $p \geq 2$ *and any* $n \geq 1$, *assume that there exist some nonnegative constants* C *and* d_n *such that*

$$\mathbb{E}\left(Z_n^p|\mathfrak{F}_{n-1}\right) \leq C^{p-2}p!d_n^2 \qquad \textit{almost surely.} \tag{A.1}$$

Then for any $\varepsilon > 0$, *we have*

$$\mathbb{P}\left(|S_n| > \varepsilon\right) \leq 2\exp\{-\frac{\varepsilon^2}{2\left(D_n + C\varepsilon\right)}\},$$

where $D_n = \sum_{i=1}^{n} d_i^2$.

Theorem A.1.1. ***(Martingal Central Limit Theorem)*** *[67]*
For each n*, let* $\{M_{n,m}\}_{m\geq 0}$ *be a MG in* $\mathcal{L}^2$ *with respect to filtration* $\{\mathfrak{F}_{n,m}\}_{m\geq 0}$ *with corresponding MG differences* $Z_{n,m} = M_{n,m} - M_{n,m-1}$ *and conditional variance* $\sigma^2_{n,m} = \mathbb{E}\left(Z^2_{n,m}|\mathfrak{F}_{n,m-1}\right)$.
Assume that, for each $n, M_{n,m}$ *and* $\Gamma_{n,m} \equiv \sum_{r=1}^{m} \sigma^2_{n,r}$ *converge* $a.s$ *to a finite limit when* $n \longrightarrow \infty$.
suppose that

1) $\Gamma_{n,m} \equiv \sum_{m=1}^{\infty} \sigma^2_{n,m} \longrightarrow 1$ *in probability* $as\, n \longrightarrow \infty$.

2) $\forall \varepsilon > 0, \lim_{n\longrightarrow\infty} \sum_{m=1}^{\infty} \mathbb{E}\left(Z^2_{n,m}; |Z_{n,m}| > \varepsilon\right) = 0.$

Then

$$Z_{m,\infty} \equiv \sum_{m=1}^{\infty} Z_{n,m} \to N(0,1), \qquad as\, n \longrightarrow \infty.$$

A.2 Useful inequalities

Theorem A.2.1. ***(Jensen's inequality)****[84]*
Let X *be a real random variable and* φ *a convex function. Then*

$$\varphi(\mathbb{E}(X)) \leq \mathbb{E}(\varphi(X))$$

Theorem A.2.2. ***(Hölder's inequality)****[86]*
Let X *and* Y *be two random variables such that* $X \in L^p(\omega; A; \mathbb{P})$ *and* $Y \in L^q(\omega; A; \mathbb{P})$ *with* $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$ *and* $p \geq 1, q \geq 1$, *then*

$$\mathbb{E}\left(|XY|^{\frac{1}{r}}\right) \leq \left(\mathbb{E}\Big(|X|\Big)^p\right)^{\frac{1}{p}} \left(\mathbb{E}\Big(|Y|\Big)^q\right)^{\frac{1}{q}}.$$

Theorem A.2.3. ***(Markov's inequality)****[61]*
Let X *be a real random variable. Then, for all* $a > 0$,

$$\mathbb{P}(|X| > a) < \frac{\mathbb{E}(|X|)}{a}.$$

Now, we state the celebrated Borel Cantelli Lemma which is a useful technique in probability theory.

Lemma A.2.1. ***(Borel Canteli's Lemma)****[89]*
Let $B_1, B_2, \ldots$ *be a sequence of events in some probability space. The Borel-Cantelli Lemma states:*

i) If, for any sequence $B_1, B_2, \ldots$ of events, $\sum_{i=1}^{n} \mathbb{P}(B_n) < \infty$, then $\mathbb{P}(B_n\, i.o.) = 0$, where $i.o.$ is an abbreviation for "infinitively often".

ii) If $\sum_{i=1}^{n} \mathbb{P}(B_n) = \infty$, and $B_1, B_2, \ldots$ are independent events, then $\mathbb{P}(B_n\, i.o.) = 1$.

A.3 Ergodic Theorem and related concepts

Theorem A.3.1. ***(Birkhoff's ergodic Theorem)***
Let $\{Z_i\}_{-\infty}^{\infty}$ be a stationary and ergodic process with $\mathbb{E}(|Z_1|) < \infty$. Then

$$\lim_{n\longrightarrow\infty} \frac{1}{n}\sum_{i=1}^{n} Z_i = \mathbb{E}(Z_1), \qquad \text{almost surely.} \tag{A.2}$$

Definition A.3.1. *Let $\{X_n, n \in \mathbb{Z}\}$ be a stationary sequence. Consider the backward field $\mathcal{B}_n = \sigma(X_k; k \geq n)$ and the forward field $\mathcal{H}_m = \sigma(X_k; k \geq m)$. The sequence is ergodic if*

$$\lim_{n\longrightarrow\infty} \frac{1}{n}\sum_{k=0}^{n-1} \left|\mathbb{P}\left(A \cap \tau^{-k} B\right) - \mathbb{P}(A)\,\mathbb{P}(B)\right| = 0, \tag{A.3}$$

where τ is the time-evolution or shift transformation.

Remark A.3.1. *A stationary and strongly mixing process is a stationary and ergodic process. And there are some examples of stationary and ergodic processes which are not strongly mixing.*

A.4 Almost complete convergence

Definition A.4.1. *[55]*
One says that $(X_n)_{n\in\mathbb{N}}$ converges almost completely $(a.co)$ to some $r.r.v.X$, if and only if

$$\forall \varepsilon > 0, \qquad \sum_{n\in\mathbb{N}} \mathbb{P}(|X_n - X| > \varepsilon) < \infty,$$

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

$$\lim_{n\longrightarrow\infty} X_n = X, \qquad a.co.$$

Definition A.4.2. *[55] One says that the rate of almost complete convergence of $(X_n)_{n\in\mathbb{N}}$ to X is of order u_n if and only if*

$$\exists \varepsilon_0 > 0, \qquad \sum_{n\in\mathbb{N}} \mathbb{P}(|X_n - X| > \varepsilon_0 u_n) < \infty,$$

and we write

$$X_n - X = O_{a.co}(u_n).$$

Proposition A.4.1. *[55]*
Assume that $X_n - X = O_{a.co}(u_n)$, we have:

i) $X_n - X = O_p(u_n)$,

ii) $X_n - X = O_{a.s}(u_n)$.

Corollary A.4.1. *[55]*

i) If $\exists M < \infty, |Z_1| \leq M$, and denoting $\sigma^2 = \mathbb{E}Z_1^2$, we have

$$\forall \varepsilon \geq 0, \mathbb{P}\left(|\sum_{i=1}^{n} Z_i| > \varepsilon n\right) \leq \exp\left\{-\frac{\varepsilon^2 n}{2\sigma^2\left(1+\varepsilon\frac{M}{\sigma^2}\right)}\right\}.$$

ii) Assume that the variables depend on n (that is, $Z_i = Zi, n)$ and are such that $\exists M = M_n < \infty, |Z_1| \leq M$ and define $\sigma_n^2 = \mathbb{E}Z_1^2$. If $u_n = n^{-1}\sigma_n^2 \log n$ verifies $\lim_{n\longrightarrow\infty} u_n = 0$, and if $M/\sigma_n^2 < C < \infty$, then we have

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

Definition A.4.3. ***(Kolmogorov's entropy)****[59]*
Let $\mathcal{S}$ be a subset of a semi-metric space $\mathcal{F}$, and let $\varepsilon > 0$. A finite set of points $x_1, x_2, \ldots, x_n$ in $\mathcal{F}$ is called an $\varepsilon-$net for $\mathcal{S}$ if $\mathcal{S} \subset \bigcup_{k=1}^{N} B\left(x_k, \varepsilon\right)$. The quantity $\log\left(N_\varepsilon\left(\mathcal{S}\right)\right)$, where $N_\varepsilon\left(\mathcal{S}\right)$ is the minimal number of open balls in $\mathcal{F}$ of a radius ε which is necessary to cover $\mathcal{S}$, is called Kolmogorov's $\varepsilon-$entropy of the set $\mathcal{S}$.

BIBLIOGRAPHY

[1] Ait-Saïdi, A., Kassa, R. and Ferraty, F. Choix optimal du paramètre fonctionnel dans le modèle à indice fonctionnel simple. *Comptes Rendus Mathematique*, **346**, 217–220 (2008).

[2] Andrews, D. WK Non-strong mixing autoregressive processes. *Journal of Applied Probability*, 930–934(1984).

[3] Attouch, M., Laksaci, A., and Ould-Said, E. Asymptotic distribution of robust estimator for functional nonparametric models. *Communications in Statistics Theory and Methods*, **38**, 1317–1335(2009).

[4] Ardjoun, F. Z., and Hennani, L. A. and Laksaci, A. A recursive kernel estimate of the functional modal regression under ergodic dependence condition. *J. of Statistical Theory and Practice*, **10**, 475–496(2016).

[5] Aurzada, F. and Simon, T. Small ball probabilities for stable convolutions. *ESAIM: Probability and Statistics*, **11**,327–343(2007).

[6] Ayad, S., Laksaci, A., Rahmani, S. and Rouane, R. On the local linear modelization of the conditional density for functional and ergodic data. *METRON,* **78**, 237–254 (2020).

[7] Baìllo, A. and Grané, A. Local linear regression for functional predictor and scalar response. *J. of Multivariate Anal.*, **100**, 102–111(2009).

[8] Banon, G. Nonparametric identification for diffusion processes. *SIAM Journal on Control and Optimization.*, **16**, 380–395 (1978).

[9] Barrientos, J., Ferraty, F. and Vieu, P. Locally Modelled Regression and Functional Data. *J. Nonparametr. Statist.*, **22**, 617-632 (2010).

[10] Belarbi, F.,Chemikh, S. and Laksaci, A. Local linear estimate of the nonparametric robust regression in functional data. *Statistics & Probability Letters.*, **134**, 128–133 (2018).

[11] Benziadi, F., Laksaci, A. and Tebboune, F. Recursive kernel estimate of the conditional quantile for functional ergodic data. *Communications in Statistics-Theory and Methods,* **45**, 3097–3113 (2016).

[12] Benziadi, F., Gheriballah, A. and Laksaci, A. Asymptotic normality of kernel estimator of ψ-regression function for functional ergodic data. *New Trends in Mathematical Sciences,* **4**, 268–282 (2016). .

[13] Berlinet, A., Elamine, A. and Mas, A. Local linear regression for functional data, *Annals of the Institute of Statistical Mathematics*, **63**, 1047–1075 (2011)

[14] Besse, P., and Ramsay, J. O. Principal components analysis of sampled functions. *J. of Psychometrika* **51**, 285–311 (1986).

[15] Besse, P. Approximation spline de l'analyse en composantes principales d'une variable aléatoire hilbertienne. *Annales de la Faculté des sciences de Toulouse: Mathématiques*, **12**, 329–349 (1991).

[16] Besse, P., Cardot, H. and Ferraty, F. Simultaneous non-parametric regressions of unbalanced longitudinal data. *Computational Statistics & Data Analysis*, **24**, 255–270 (1997).

[17] Besse, P., and Cardot, H. Modélisation statistique de données fonctionnelles. *Analyse de données. Hermes*(2003).

[18] Bogachev, VI. Gaussian Measures (Mathematical Surveys and Monographs. *Providence, RI: American Mathematical Society*, **62**(1998).

[19] Bosq, D. (2012). Nonparametric statistics for stochastic processes: estimation and prediction (Vol. 110). Springer Science & Business Media.

[20] Bouanani, O., Laksaci, A., Rachdi, M. and Rahmani, S. Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics, *Behaviormetrika*, **46**, 1–35 (2018)

[21] Bouanani, O., Rahmani, S. and Ait-Hennani, L. Local linear conditional cumulative distribution function with mixing data. *Arabian Journal of Mathematics,* , 1–19 (2019).

[22] Bouchentouf, A. A, Djebbouri, T., Rabhi, A. and Sabri, K. Strong uniform consistency rates of some characteristics of the conditional distribution estimator in the functional single-index model. *Applicationes Mathematicae* **4**, 301–322 (2014).

[23] Carbonez, A., Györfi, L., and vander Meulen, EC. Partitioning-estimates of a regression function under random censoring. *Statistics & Risk Modeling* **13**, 21–38 (1995).

[24] Chahad, A., Ait-Hennani, L. and Laksaci, A. Functional local linear estimate for functional relative-error regression. *Journal of statistical theory and practice,* **11**, 771–789 (2017).

[25] Chaouch, M., & Khardani, S. (2015). Randomly censored quantile regression estimation using functional stationary ergodic data. Journal of nonparametric statistics, **27**(1), 65-87.

[26] Chaouch, M., Laïb, N. and Ould Saïd, E. Nonparametric M-estimation for right censored regression model with stationary ergodic data. *Statistical Methodology,* **33**, 234–255 (2016).

[27] Chernick, M. R. A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. *The Annals of Probability,* **9**, 145–149 (1981).

[28] Chikr-Elmezouar, Z., Almanjahie, IM., Laksaci, A and Rachdi, M. FDA: strong consistency of the k NN local linear estimation of the functional conditional density and mode. *Journal of Nonparametric Statistics,* **31**, 175–195 (2019).

[29] Chouaf, A. and Laksaci, A. On the functional local linear estimate for spatial regression. *Statistics & Risk Modeling,* **29**, 189–214 (2012).

[30] Dabo-Niang, S.and Rhomari, N. Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *Comptes Rendus Mathematique,* **336**, 75–80 (2003).

[31] Dabo-Niang, S. Sur l'estimation fonctionnelle en dimension infinie: application aux diffusions. thesis, Paris 6, France, (2002).

[32] Dabo-Niang, S.and Rhomari, N. Estimation non paramétrique de la régression avec variable explicative dans un espace métrique. *Comptes Rendus Mathematique,* **336**, 75–80 (2003).

[33] Dabo-Niang, S. Kernel density estimator in an infinite-dimensional space with a rate of convergence in the case of diffusion process. *Applied mathematics letters,* **17**, 381–386 (2004).

[34] Dabo-Niang, S., Ferraty, F. and Vieu, P. Mode estimation for functional random variable and its application for curves classification. *Far East Journal of Theoretical Statistics,* **18**, 93–119 (2006).

[35] Dabo-Niang, S. and Laksaci, A. Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle. *Comptes Rendus Mathematique,* **344**, 49–52 (2007).

[36] Davidson, J. *Stochastic limit theory: An introduction for econometricians.* OUP Oxford. (1994).

[37] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. *A fast functional locally modeled conditional density and mode for functional time-series.* Recent Advances in Functional Data Analysis and Related Topics, Contributions to Statistics, Pages 85–90, DOI: 10.1007/978-3-7908-2736-1_13, Physica-Verlag/Springer (2011).

[38] Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. Functional data: local linear estimation of the conditional density and its application. *Statistics*, **47**, 26–44 (2013).

[39] Demongeot, J., Laksaci, A., Rachdi, M., and Rahmani, S. On the local linear modelization of the conditional distribution for functional data. *Sankhya A*, **76**, 328–355 (2014).

[40] Demongeot, J., Naceri, A., Laksaci, A., and Rachdi, M. Local linear regression modelization when all variables are curves. *Statistics & Probability Letters*, **121**, 37–44 (2017).

[41] Deville, J. C. Méthodes statistiques et numériques de l'analyse harmonique. *Ann. Insee,*, **15**, 3–101 (1974).

[42] Didi, S. and Louani, D. Asymptotic results for the regression function estimate on continuous time stationary and ergodic data. *Statistics and Risk Modeling*, **31**, 129–150 (2014).

[43] Ezzahrioui, M. and Ould-Saïd, E. Asymptotic normality of a nonparametric estimator of the conditional quantile in the normed space. *Far East J. Theoretical Statist*, **25**, 15–38 (2008a).

[44] Ezzahrioui, M. and Ould-Saïd, E. Asymptotic results of a nonparametric conditional quantile estimator for functional time series. *Communications in Statistics Theory and Methods*, **37**, 2735–2759 (2008b).

[45] Ezzahrioui, M. and Ould-Saïd, E. Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *Journal of Nonparametric Statistics*, **20**, 3–18 (2008).

[46] Fan, J. Design-adaptive nonparametric regression. *J. Amer. Statist. Assoc.*, **87**, 998–1004 (1992).

[47] Fan, J., Gijbels, I. *Local Polynomial Modelling and its Applications*. London, Chapman and Hall. (1996).

[48] Fan, J. and Yao, Q. *Nolinear Time Series: Nonparametric and Parametric Methods.* Springer-Verlag, New York.(2003).

[49] Ferraty, F. and Vieu, P. Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics* **330**, 139–142 (2000).

[50] Ferraty, F. and Vieu, P. The functional nonparametric model and application to spectrometric data. *Computational Statistics* **17**, 545–564 (2002).

[51] Ferraty, F. and Vieu, P. Curves discrimination: a nonparametric functional approach. *Computational Statistics & Data Analysis* **44**, 161–173 (2003).

[52] Ferraty, F. and Vieu, P. Nonparametric models for functional data, with application in regression, time series prediction and curve discrimination. *Nonparametric Statistics* **16**, :111-125 (2004).

[53] Ferraty, F., Laksaci, A. and Vieu, P. Functional time series prediction via conditional mode estimation. *Comptes Rendus Mathematique* **340**, 389–392 (2005).

[54] Ferraty, F., Rabhi, A. and Vieu, P. Conditional quantiles for dependent functional data with application to the climatic El Niño phenomenon. *Mathematical Reviews (MathSciNet): MR2208895* **67**, 378–398 (2005).

[55] Ferraty, F. and Vieu, P. Nonparametric functional data analysis theory and practice. Springer-Verlag (2006).

[56] Ferraty, F., Laksaci, A. and Vieu, P. Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statistical Inference for Stochastic Processes.*,**9** , 47–76 (2006).

[57] Ferraty, F.,Mas, A. and Vieu, P. Nonparametric regression on functional data: inference and practical aspects. *Australian & New Zealand Journal of Statistics.*,**43** , 267–286 (2007).

[58] Ferraty, F., Rahbi, A. and Vieu, P. Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle. *Statistical Inference for Stochastic Processes.*,**53** , 1–18 (2008).

[59] Ferraty, F., Laksaci, A., Tadj. A. and Vieu, P. Rate of uniform consistency for nonparametric estimates with functional variables. *J. of Statistical Planning and Inference.*, **140**, 335–352 (2010).

[60] Ferraty, F. and Romain, Y. The Oxford Handbook of Fuctional Data Analysis, *Oxford University Press* (2011).

[61] Foata, D. and Fuchs, A. Cours et exercices corrigés. Enseignement des Mathématiques. Paris: Masson. x, 1996

[62] Gao, F. and Li, W. Small ball probabilities for the Slepian Gaussian fields. *Transactions of the American Mathematical Society.*, **359**, 1339–1350 (2007).

[63] Geenens, G. (2011). Curse of dimensionality and related issues in nonparametric functional regression. Statistics Surveys, **5**, 30-43.

[64] Gheriballah, A., Laksaci, A. and Sekkal, S. Nonparametric M-regression for functional ergodic data. *Statistics and Probability Letters.*, **83**, p. 902–908 (2013).

[65] Jolliffe, I. Principal component analysis: Wiley Online Library. *Kaiser, HF (1960). The application of electronic computers to factor analysis.*, (2002).

[66] Györfi, L., Härdle, W., Sarda, P., & Vieu, P. Nonparametric curve estimation from time series (Vol. 60) Springer. (2013).

[67] Heyde, CC and Hall, P. Martingale limit theory and its application. (1980).

[68] Horváth, L., & Kokoszka, P. (2012). Inference for functional data with applications (Vol. 200). Springer Science & Business Media.

[69] Kaplan, EL., and Meier, P. Nonparametric estimation from incomplete observations. *Journal of the American statistical association* **53**, 457–481 (1958).

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

[71] Laïb, N. and Louani D. Nonparametric kernel regression estimation for functional stationary ergodic data: asymptotic properties. *J. of Multivariate analysis*, **101**, 2266–2281(2010).

[72] Laïb, N. and Louani D. Rates of strong consistencies of the regression function estimator for functional stationary ergodic data. *J. Stat. Plann. and Inf.*, **141**, p. 359–372 (2011).

[73] Laksaci, A. Convergence en moyenne quadratique de l'estimateur à noyau de la densité conditionnelle avec variable explicative fonctionnelle. *Annales de l'ISUP*, **51**, 69–80 (2007).

[74] Laksaci, A. and Mechab, B Estimation non paramétrique de la fonction de hasard avec variable explicative fonctionnelle: cas des données spatiales. *Revue Roumaine de Mathématiques Pures et Appliquées*, **55**, 35–51 (2010).

[75] Li, W. V and Shao, Q.M. Gaussian processes: inequalities, small ball probabilities and applications. *Handbook of Statistics.*, **19**, 533–597 (2001).

[76] Lifshits, M. A., Linde, W. and Shi, Z. Small Deviations of Riemann-Liouville Processes In lq Spaces With Respect To Fractal Measures. *Proceedings of the London Mathematical Society.*, **92**, 224–250, (2006).

[77] Lifshits, M. A., and Simon, T. Small deviations for fractional stable processes. *Annales de l'IHP Probabilités et statistiques.*, **41**, 725–752, (2005).

[78] Mahiddine, A and Bouchentouf, A. A and Rabhi, A. Nonparametric estimation of some characteristics of the conditional distribution in single functional index model. *Malaya Journal of Matematik (MJM)*, **2**, 392–410 (2014).

[79] Mas, A. (2012). Lower bound in regression for functional data by representation of small ball probabilities. Electronic Journal of Statistics, 6, 1745-1778.

[80] Masry, E., Nonparametric regression estimation for dependent functional data: asymptotic normality. *Stochastic Processes and their Applications.*, **115**, 155–177 (2005).

[81] Mayer-Wolf, E and Zeitouni, O. The probability of small Gaussian ellipsoids and associated conditional moments. *The Annals of Probability.*, 14–24 (1993).

[82] Müller HG. Functional modeling of longitudinal data. In Longitudinal data analysis. Chapman and Hall/CRC, New York. (pp. 237-266). (2008).

[83] Nadaraya, E., On estimating regression. *Theory of Probability & Its Applications.*, **9**, 141–142 (1964).

[84] Nonnenmacher, DJF and Zagst, R. A new form of Jensen's inequality and its application to statistical experiments. *The ANZIAM Journal*, **36**,389–398 (1995).

[85] Parzen, E. On estimation of a probability density function and mode, *The annals of mathematical statistics* **33**, 1065–1076 (1962).

[86] Puntanen, S. Inequalities: Theory of Majorization and Its Applications, by Albert W. Marshall, Ingram Olkin, Barry C. Arnold, *Wiley Online Library* (2011).

[87] Rabhi, A., Hammou, Y. and Djebbouri, T. Nonparametric estimation of the maximum of conditional hazard function under dependence conditions for functional data. *Afrika Statistika,* **10**, 777–794 (2015).

[88] Rabhi, A., Soltani, S. and Traore, A. Conditional risk estimate for functional data under strong mixing conditions. *Journal of Statistical Theory and Applications,* **14**, 301–323 (2015).

[89] Serfling, R J. Approximation theorems of mathematical statistics. *John Wiley & Sons*, **192**, (2009).

[90] Ramsay, J. O and Silverman, B.W. Functional Data Analysis Springer-Verlag, *New York,* (1997).

[91] Ramsay, J., Silverman, B.W. Applied functional data analysis; Methods and case studies. *Springer-Verlag, New York* (2002).

[92] Rosenblatt, M. Remarks on some nonparametric estimates of a density function, *The Annals of Mathematical Statistics,* 832–837 (1956).

[93] Ramsay, J. O. and Silverman, B. W. (2005). Functional Data Analysis. 2nd edition Springer-Verlag, New York.

[94] Rice, J. and Rosenblatt, M. (1976). Estimation of the log survivor function and hazard function. Sankhya Ser. A 38, 60–78.

[95] Shmileva, E. Small ball probabilities for jump Lévy processes from the Wiener domain of attraction. *Statistics & probability letters.*, **76**, 1873–1881 (2006).

[96] Stone, C. J. Optimal global rates of convergence for nonparametric regression. *The annals of statistics.*, 1040–1053 (1982).

[97] Watson, G.S. Smooth regression analysis. Sankhya A, 26, pp. 359–372 (1964).

[98] Xiong, X., Zhou, P,. and Ailian, C. Asymptotic normality of the local linear estimation of the conditional density for functional time-series data. *Communications in Statistics-Theory and Methods*, **47**, 3418–3440 (2018).

[99] Zhou, Z. and Lin, Z.-Y. Asymptotic normality of locally modelled regression estimator for functional data. *J. Nonparametr. Statist.*, **28**, 116–131 (2016).

ملخص

في هذه الأطروحة، نأخذ في الاعتبار مشكل التقدير المحلي الخطي لوظيفة التوزيع الشرطي ومشتقاتها عندما يكون الانحدار ذو قيمة في فضاء ذو بعد غير محدود، الاستجابة عددية (تمت ملاحظتها بالكامل أو خضعت للرقابة) ويتم ملاحظة البيانات على أنها سلسلة أوقات وظيفية أرجوديك.

أولا، نبني تحت هيكل التبعية مقدرا محليا خطيا لوظيفة التوزيع الشرطي، وندرس في ظل افتراضات عامة معينة خصائصه المقاربة، مثل التقارب النقطي شبه الكامل (مع السرعة) والتقارب الطبيعي. يتم التحقق من ملاءمة المقدر المقترح من خلال دراسة المحاكاة.

ثانيا، في ظل نفس الشروط، نبني مقدرا محليا خطيا للكثافة الشرطية. ثم ندرس التقارب شبه الكامل، مع السرعة، لهذا المقدر، ونستنتج من ذلك خصائص مقاربة مشابهة للمقدر المحلي الخطي للمنوال الشرطي. يتم توضيح فائدة نتائجنا على بيانات حقيقية.

أخيرًا، نقوم بتعميم النتائج التي تم الحصول عليها مسبقًا في سياق خاضع للرقابة. نبني مرة أخرى مقدرًا للكثافة الشرطية بالطريقة المحلية الخطية وندرس سرعة التقارب شبه الكامل للمقدر المبني.

Résumé

Dans cette thèse, nous considérons le problème de l'estimation locale linéaire de la fonction de répartition conditionnelle et de ses dérivées lorsque le régresseur est évalué dans un espace de dimension infinie, la réponse est un scalaire (complètement observé ou censuré) et les données sont observées comme séries temporelles fonctionnelles ergodiques.

Tout d'abord, nous construisons sous cette structure de dépendance un estimateur local linéaire de la fonction de répartition conditionnelle, et nous établissons sous certaines hypothèses générales ses propriétés asymptotiques, telles que la convergence uniforme presque complète (avec taux) et la normalité asymptotique. La pertinence de l'estimateur proposé est vérifiée par une étude de simulation.

Deuxièmement, et sous les mêmes conditions, nous construisons un estimateur local linéaire de la densité conditionnelle. Ensuite, on établit la convergence presque complète, avec des taux, de cet estimateur, et on en déduit des propriétés asymptotiques similaires de l'estimateur linéaire local du mode conditionnel. L'utilité de nos résultats est illustrée sur des données réelles.

Enfin, nous généralisons les résultats précédemment obtenus dans un contexte censuré. On construit à nouveau un estimateur de la densité conditionnelle par la méthode locale linéaire et on établit la vitesse de convergence presque complète de l'estimateur construit.

Abstract

In this thesis, we consider the problem of the local linear estimation of the cumulative distribution function and its derivatives when the regressor is valued in an infinite dimensional space, the response is scalar (completely observed or censored) and the data are observed as ergodic functional times series.

Firstly, we build under this dependence structure a local linear estimator of the distribution function, and we establish under general assumptions its asymptotic properties, such as the uniform almost complete convergence (with rate) and the asymptotic distribution. The relevance of the proposed estimator is verified through a simulation study.

Secondly, under the same conditions, we construct a local linear estimator of the conditional density function. Afterward, we establish the almost complete convergence, with rates, of this estimator, and we deduce similar asymptotic properties of the local linear estimator of the conditional mode. The usefulness of our results is illustrated on some real data.

Finally, we generalize the results previously obtained in a censored context. We build again an estimator of the conditional density by the local method and we establish the strong consistency rate of the constructed estimator.