N°d’ordre:

Université de Saida– Dr. Moulay Tahar
Faculté des Sciences

Thèse

Présentée pour obtenir le diplôme de

Doctorat 3ième Cycle

Spécialité: Probabilités et Statistique

Filière: Mathématiques

Par:

Mlle. BOUAZZA Imane

Thème:

Estimation récursive de modèles non paramétriques pour des données ergodiques fonctionnelles

Thèse soutenue le 30/ 03/ 2022 devant le jury composé de:

N°	Nom et prénom	Grade	Etablissement	Qualité
01	Mr. Kandouci Abdeldjebbar	Prof.	Université de Saida – Dr. Moulay Tahar	Président
02	Mr. Guendouzi Toufik	Prof.	Université de Saida – Dr. Moulay Tahar	Rapporteur
03	Mr. Benchikh Tawfik	Prof.	Université de Sidi Bel Abbès	Examinateur
04	Mlle. Rahmani Saadia	Prof.	Université de Saida – Dr. Moulay Tahar	Examinatrice
05	Mr. Attouch Mohammed Kadi	Prof.	Université de Sidi Bel Abbès	Examinateur
06	Mme. Benziadi Fatima	MCB	Université de Saida – Dr. Moulay Tahar	Invitée

Dedication

To those who have great credit for my arrival and achievement of my goal and purpose. To those who helped and supported me in every step I took to the light of dawn, the symbol of tenderness and safety, the source of happiness and giving, which focuses on the secret of my struggle and my success, to the one who bears every moment of pain in my life and turns it into moments of joy. To those with whom I lived the most beautiful moments of my life, to candles my way. To the inhabitants of my heart, I dedicate my graduation to you, my little family.

♡

Acknowledgements

The time has come for me to express my great gratitude to all those who helped and encouraged me in the difficult moments of obtaining my PhD degree from the University of Saida- Dr. Moulay Tahar.

First and foremost, I would like to praise and thank **Allah**, the merciful and the passionate for giving me countless blessing, knowledge, and opportunity to step into the excellent world of sciences. All my praise is for you my creator for giving me the strength, wisdom, guidance and the physical ability to be able to accomplish this contribution.

This thesis would not see the light without the encouragement and guidelines of many people. So, I would like to warmly express my sincere gratitude to my thesis supervisors, *Pr. Toufik Guendouzi* and *Dr. Mme. Fatima Benziadi*, without whom this work would not have been possible, who accepted to supervise my first steps in research. Being their student was an honor and a pleasure, and I am very grateful for the unfailing availability, patience, advice and the confidence they showed in spite of their busy schedules. My professors, I learned a lot from you and your interesting lessons that helped me to progress in the thesis will always remain in my memory. I have acquired a lot of knowledge from them thanks to their unlimited generosity, both pedagogical and social.

I would also like to express my sincere thanks to *Pr.Abdeldjebbar Kandouci* for his advice and encouragement and for having done me the honor of chairing the thesis jury; may he find here the expression of my deepest gratitude. **I** would like to thank my thesis examiners, *Pr. Saadia Rahmani*, *Pr.Mohammed Kadi Attouch* and *Pr.Tawfik Benchikh* for accepting to examine the thesis, for being part of the jury and for carefully rereading this work, your remarks and for the interest you have shown in the current work. Furthermore, I am deeply grateful once again to *Pr.Abdeldjebbar Kandouci*, *Pr.Fethi Madani*, *Pr.Rekia Bourroumi*, *Dr.Soumia Idrissi* and *Dr.Khadidja Hadj Djelloul* for helpful discussions and valuable comments and advice. I will never forget the sense of unlimited generosity.

Special mention to all the members and staff of the department of mathematics at Saïda university and especially the laboratory of stochastic models, statistics and application where I did my first research internship, is gratefully acknowledged for providing us with the full equipment, capabilities whether material or moral and the pleasant working atmosphere.

Finally, I wish to thank all the members of my family especially my father, my mother and my brothers for their caring support, endurance, thoughtfulness and love that helped me to reach this point. And the last thank is for my friends who are my brothers. You have been my supporters and motivator for my success.

Imane **B**ouazza

Abstract

The main inspirational theme for this work is based on building recursive estimators for nonparametric conditional models, extending the works done previously to the issues recently discussed in nonparametric statistics. Hence, it is articulated around three main axes: Recursive Estimation, Functional Ergodic context and the survival data analysis. The Recursive Kernel method of statistics, which the work in hands focuses on, is presented in detail to assess its efficiency in nonparametric estimation, from which better conclusions can be drawn. Therefore, the main model considered here is the conditional distribution function and its derivatives such as: conditional quantile, density and mode functions, of the scalar response variable Y for a given random variable X taking its values in semi-metric space, by introducing then their recursive adaptations for ergodic random variables. The thesis uses appropriate statistical methodologies and theories to manage basic issues related to the possible prevalence of outliers and incomplete observations in the sample. Further, given the power of the recursive method, we continue to ask reasonable question of whether this method, proposed for evaluating previous models in complete case, can be considered a useful one and remains a viable alternative if data are incomplete. To this end, the estimate is extended to the case that observations can be right-censored for discrete variables.

Keywords: Recursive Estimate, Functional data, Conditional models, Almost sure convergence, Asymptotic distribution, Ergodic data, Incomplete data.

Résumé

Le principal thème d'inspiration de ce travail est basé sur la construction d'estimateurs récursifs pour les modèles conditionnels non paramétriques, en étendant les travaux réalisés précédemment aux questions récemment discutées en statistiques non paramétriques. Il s'articule donc autour de trois axes principaux: L'estimation récursive, le contexte ergodique fonctionnel et l'analyse des données de survie. La méthode statistique du noyau récursif, sur laquelle porte le travail en cours, est présentée en détail pour évaluer son efficacité dans l'estimation non paramétrique, ce qui permet de tirer de meilleures conclusions. Le modèle principal considéré ici est donc la fonction de distribution conditionnelle et ses dérivés tels que: les fonctions conditionnelles de quantile, de densité et de mode, de la variable scalaire de réponse Y pour une variable aléatoire X donnée prenant ses valeurs dans un espace semi-métrique, en introduisant ensuite leurs adaptations récursives pour des variables aléatoires ergodiques. La thèse utilise des méthodologies et des théories statistiques appropriées pour gérer les problèmes de base liés à la prévalence possible de valeurs aberrantes et d'observations incomplètes dans l'échantillon. En outre, étant donné la puissance de la méthode récursive, nous continuons à nous poser la question raisonnable de savoir si cette méthode, proposée pour évaluer les modèles précédents dans le cas complet, peut être considérée comme une méthode utile et reste une alternative viable si les données sont incomplètes. À cette fin, l'estimation est étendue au cas où les observations peuvent être censurées à droite pour des variables discrètes.
Mots clés: Estimation Récursive, Données fonctionnelles, Modèles conditionnels, Convergence presque sûre, Distribution asymptotique, Données ergodiques, Données incomplètes.

Publications and Conferences

Published Articles:

- I. Bouazza, F. Benziadi, F. Madani and T. Guendouzi. *Asymptotic results of a recursive double kernel estimator of the conditional quantile for functional ergodic data*. Bulletin of the Institute of Mathematics Academia Sinica (New Series), Vol. 16 (2021), 3, 217-239. https://doi.org/10.21915/BIMAS.2021302.

- I. Bouazza, F. Benziadi, T. Guendouzi. *Nonparametric recursive estimate for right-censored conditional mode function with ergodic functional data*. Journal of the Indian Society for Probability and Statistics, 22, 389-415 (2021). https://doi.org/10.1007/s41096-021-00110-5.

Submitted Articles:

- I. Bouazza, F. Benziadi, T. Guendouzi. *Consistency result for the recursive kernel estimate of the conditional quantile on continuous time stationary and ergodic data.*

Communications in conferences:

- **November 21-24, 2019.** *Consistency results for the conditional mode estimate on continuous time stationary and ergodic data.* Rencontre d'Analyse Mathématique et Applications RAMA11, Sidi Bel Abbès, Algérie. https://www.univ-sba.dz/rama11.

- **May 26-27, 2021.** *Nonparametric conditional density function estimation for randomly censored data.* The First Online International Conference on

Pure and Applied Mathematics, IC-PAM'21, Ouargla, ALGERIA. https://ic-pam.sciencesconf.org/.

- **June 26, 2021.** *Uniform strong convergence of recursive conditional distribution function estimator for truncated ergodic observations.* National E-Conference on Applied Mathematics and Didactics NCAMD2021, Constantine-Algeria. http://www.ensc.dz/ncamad/index.html.

- **October 27-28, 2021.** *The L^1 kernel recursive estimate of conditional distribution function for functional ergodic data in a random left-truncation model.* Proc. of Algerian Mathematicians Mini Congress AMMC2021, M'sila-Algeria. http://www.smath.dz/act/mcma2021.

Other training:

- **March 01-09, 2019.** Participation in the research CIMPA'ASA 2019 school on "Stochastic Analysis and Applications", University of Saida Dr. MOULAY Tahar.

- **February 24, 2020.** Participation in the day on "Applied Statistics" in honor of Pr. Mourid Tahar, University of Saida Dr. MOULAY Tahar.

- **March 16, 2022.** Participation in the first international day on probabilities, University of Saida Dr. MOULAY Tahar. https://www.univ-saida.dz/jiprob1/.

Contents

Chapter 1

General Introduction and Literature Review

1.1 Problem and Motivation

Conventionally, in usual statistical analysis, the identification of the link between a response variable Y and an explanatory one X is modeled with the so-called nonparametric regression function (the link function) that provides techniques which can help to achieve this aim. Thus, the regression model takes usually the form

$$Y = R(X) + \epsilon \tag{1.1}$$

where $R : x \mapsto \mathbb{E}(Y|X = x)$ is unknown in practical applications, known as the function of minimal L_2 risk with respect to Y and must be estimated from the possible observed sample $(X_k, Y_k)_{k\geq 1}$. The error terms ϵ are random and indicate that there is no exact relationship between it and the explanatory variable X. It is assumed also, that the disturbance term satisfies $\mathbb{E}(\epsilon/X = x) = 0$ and $Var(\epsilon/X = x) = \sigma^2(x)$. Hence, this function is one of the most widely used tools in statistics to predict the value of the random response variable, based on known values of one or more covariates (explanatory variables). The application of regression covers most fields such as biostatistics, economics or environmental sciences, and it is considered as one of the main solutions for those who try to study this link.

In a statistical population, the collected data contain what is called *outliers* [1]. This type of data is generally observed in medical follow-up studies, financial data (stock market indices) but also in sociological studies. However, in many statistical applications, a big drawback of the nonparametric regression model is that it is not efficient in some pathological situations (for instance, the existence of *outliers*). Notably, in all these cases, the conditional models such as: conditional density, distribution, mode and quantile functions are the pertinent ones to explore this relationship and then predicting certain events or behaviors. In the vast variety of papers, the authors have used the Nadaraya-Watson techniques as estimation method which is a particular case of the recursive kernel estimate considered in this thesis.

However, the existence of *outliers* in the estimation problem of the nonparametric regression function of the data available in a certain state, is not the only difficulty encountered. We are often confronted with the presence of *incomplete data* in the field of survival analysis. In fact, in observational studies, the data may also be subject to right-censoring generally and left-truncation sometimes. So that, the specificity of survival data is that they contain incomplete observations which lead normally to a loss of information.

1.2 Contribution of the thesis

In this thesis, we describe efforts in our laboratory to estimate models for non parametric statistics in which we have focused our attention on the conditional models (the mentioned models above) in the context of functional data analysis (i.e. $X \in \mathcal{H}$ where $\mathcal{H}$ is an infinite dimensional space) and $y \in \mathbb{R}$. Our principal objective is to study these nonparametric problems with recursive estimation method which extends the classical one and we present an overview of it. This thesis can be considered in its entirety as a contribution to the recursive method and its statistical applications. Whereas, the theory described in this work is based on several assumptions about functional and ergodic framework, censoring mechanism and continuous time process.

[1] An observation is said to be an outlier if it is "abnormally" distant from other observations made on a phenomenon

The outline of the thesis is briefly presented as follows: After a brief reminder on the basic mathematical concepts of functional, ergodic and recursive approach in addition to a general introduction to the area of survival data analysis in the first chapter, the work of the three thematic research axes mentioned above will be described in the next three chapters: Chapter 2 deals with the recursive double kernel estimator of the conditional quantile for functional ergodic data, in Chapter 3 we introduce a nonparametric recursive model for right-censored conditional mode function in the same context of ergodic functional data and in Chapter 4 we discuss on continuous time ergodic data, the recursive kernel estimate of the conditional quantile model. Finally, a general conclusion on the research perspectives closes this thesis.

In Chapter 2, based on an inverse of the conditional distribution function which is used as an important tool for modeling and then forecasting time series problems, we first present the functional non parametric conditional quantile estimation for discrete variables according to an approach based on the recursive method. Our approach is motivated by the data-generating functional process. The aim of this chapter is to confirm the prospect results achieved in Benziadi et al. [9], we study subsequently its asymptotic normality that dealing with ergodic data, construct the confidence interval and present a simulation study to illustrate the strong performance of the method and to validate our theoretical results. Hence, in our comparisons, a basic requirement is that the errors, i.e., the absolute differences between the function and its estimate, are small so that we can make the right decision. This chapter is accepted as a co-authored article for the Journal of Bulletin.

The growing success of this method has prompted us to examine its contribution to different models estimation, we consider in a next step, working on the estimation when we have incomplete information grouped in a variable of interest Y. Wherefore, Chapter 3 discusses the problem of modeling and then estimating the randomly right-censorship conditional mode function $\Theta(x)$. On the one hand, we extend to the basic concepts of randomly right censored data framework that are generally different from classical existing theoretical statistics consideration and are useful in the analysis of the asymptotic properties of our estimates. On the other hand, we focus on establishing under additional assumptions on the censoring and lifetime variables, the pointwise and uniform almost sure consistency (with rates) on a compact set of the involved nonparametric model (the conditional density as well as the conditional mode functions) for stationary and ergodic observations with respect to the new kind of data (censoring). In addition to a simple illustrative simulation to demonstrate the usefulness of the recursive approach, so that we analyze the effect of the censoring rates by comparing the performance of the estimator without/ and with censored data. This illustration is provided just to show in a simplified setting how this approach can be apply in prevalence of such observations.

As an extension of the previous described results with respect to functional context, for dependency structure of ergodic type, a slightly modified estimate of Chapter 2 is forthcoming as a submitted article. The conceptual idea here aims to consider two continuous time processes $(X_t, Y_t)_{t\geq 0}$. We describe the implementation of the estimation procedures and subsequently we establish the almost sure convergence with rates of the proposed estimator according to an approach based on the recursive method. Chapter 4 is completed by a discussion of some particular cases.

1.3 Functional data framework

As the advancement of the computer instruments, an ever increasing number of information are high dimensional vectors produced by estimating a continuous process on a discrete sampling network. Thus, numerous instances of this kind of information can be found in several fields such as, time series analysis, medicine, biology and biomechanics (for instance, growth curves), economic and financial applications and rural trial data of this sort ought to be broke down as functional information (i.e. every observation is a function coming from an infinite dimensional space) rather than a standard multivariate as functional. Indeed, in earlier, it was difficult to access, to reconstruct the form and then to register these random data for statistics (due to technological limitations), while, we have witnessed today a wide development and the ease in monitoring, storing this data and processing large amounts of it.

This thesis focuses significantly on the use of a specific fundamental branch in

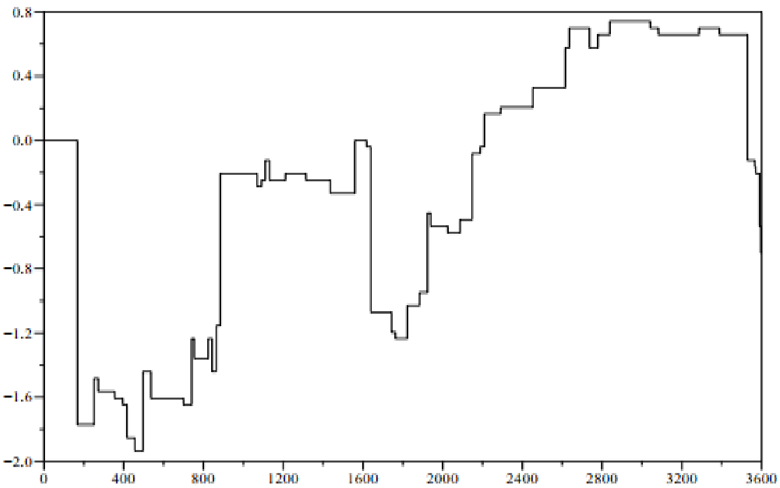

Figure 1.1: Example of functional data (Share index evolution during one hour)

nonparametric statistics that called "functional data analysis". Whereas most of the works done in the past was always concentrate on the analysis of random scalars ($X \in \mathbb{R}$) and vectors ($X \in \mathbb{R}^p, p \geq 2$). As a simple definition of functional data analysis, (FDA) is the statistical analysis of data represented by curves (random functions) or stochastic processes $(X_t)_{t \in \mathbb{I}}$ taking values in a semi metric space $\mathcal{H}$ of functions defined on some set $\mathbb{I}$ (this latter represents generally an interval of time), that extends the classical multivariate ones, which necessitates in turn to start with the development of suitable statistical methods.

Mathematicians have generally some main difficulties when dealing with functional data analysis (FDA) because of the infinite dimensional space that data belong to. For example, the non available of a definition for the distribution of a functional random variable or the definition of distances, whereas in practice one only has sampled curves observed into a finite set. In this regard, due to its prominence, several solutions have been proposed, and perhaps the most common of them requires of reducing in advance the infinite dimensional problem to a finite one by approximating data with elements belong to some finite dimensional space. Others preferred as a method to put forward, for any statistical model, a qualitative assumption on the underlying process. On the other hand, it consists to define specific distances or differences for functional data.

The study of statistical models in FDA has gained a great importance nowadays, such as, the non parametric conditional models, functional linear regression, functional autoregressive models, and so on. In which, the interested reader would find his goal in many excellent references, the most prominent of which are: the work of Bosq (2000)[12] for modeling dependent functional random variables, the popular monograph on functional data of Ramsay and Silverman (2002)[82] developing theory and applications of FDA. Also, the book of Ferraty and Vieu (2006)[39] on nonparametric models for functional data containing a review of the most recent contributions on this topic. In addition to a research group working on functional statistics in Toulouse (STAPH [2]) who have contributed greatly on regression models for functional data. Thus, according to Ferraty and Vieu (2006)[39], a functional random variable X is a random process with values in an infinite dimensional functional space $\mathcal{H}$. It is common to use $\mathcal{H}$ with its Borel $\sigma-$algebra generated by its open sets. Then, a functional data represents a set of random variables $(X_1, X_2, \ldots, X_n)$ of X and drawn from the same distribution as it.

Recently, several researchers have concentrated their efforts to solve non parametric problems when the dimension p in his absolute value, is very large or with respect to the size of some sample drawn from the distribution of X. Therefore, the study of the models from the functional data i.e. the modeling and the forecasting of several received problems such as the electricity demand; see, e.g., Ferraty and Vieu (2006)[39] and Attouch et al. (2010)[5].

1.4 Conditional models estimation in functional statistics

In the sequel, assume that all random variables are defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Thus, for measuring the dependence between variables, numerous probabilistic tools have been developed, on the one hand mixing assumptions, introduced by Rosenblatt (1956)[84], on the other hand martingales approximations; that are used

[2] http://univ-tlse3.fr/STAPH

to convey different ideas of asymptotic independence. For prediction, the statisticians have usually considered time series which are dependent by using regression approach. So kernel estimation was widely investigated under different notions of dependence to provide on a variety of results that in turn cover several ideas.

For independent samples, in the literature, several outcomes have been recorded that study conditional models estimate. Then again, researches for dependent samples should be highlighted, so that it is a question of responding to workable situations in which the data are not i.i.d. Within a sample, there are several types of dependency modeling: weak and strong. In fact, there are various popular models of weak dependency; $\alpha-$mixing[3](Rosenblatt (1956b)[84]), $\beta-$mixing[4](Volkonskii and Rozanov (1959)[91]) or $\phi-$mixing[5](Ibragimov (1962)[50]) and so on. For a wide view on the different sorts of mixing and examples we allude to Doukhan (1994)[26]. On the same path, Doukhan and Louhichi (1999)[27] have introduced a new concept of weak dependence condition that makes explicit the asymptotic independence between past and future. The authors have derived almost sure convergence of kernel density (Doukhan and Louhichi (2001)[28]). Whereas, consistency of kernel regression estimate has been studied by Ango Nzé et al. (2002)[3].

In our thesis, we focus on one type: ergodic case. The main advantage is that such a kind of dependence is very easy to verify, is the weakest and therefore the least restrictive that covers a large class of time series models. Thus, the main contribution

[3] Let $(X_k)_{k>0}$ be a sequence of real random variables. Denote by $\wp_1^j$ the $\sigma-$algebra generated by the $X_k, 1 \leq k \leq j$ and $\wp_{n+j}^{\infty}$ the ones generated by the $X_k, n+j \leq k < \infty$. We define the associated mixing coefficients between two $\sigma-$fields $\wp_{n+j}^{\infty}$ and $\wp_1^j$ to the sequence $(X_k)_{k>0}$ by:

$$\alpha(n) = \sup_{j\geq 1} \sup \left\{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|;\ A \in \wp_{n+j}^{\infty}, B \in \wp_1^j \right\}.$$

Then, we say that this sequence is $\alpha-$mixing if $\alpha(n) \to 0$ when $n \to \infty$.

[4]

$$\beta(n) = \sup_{j\geq 1} \sup \left\{ \frac{1}{2} \sum_{i=1}^{I} \sum_{s=1}^{S} |\mathbb{P}(A_i \cap B_s) - \mathbb{P}(A_i)\mathbb{P}(B_s)|;\ A_i \in \wp_{n+j}^{\infty}, B_s \in \wp_1^j \right\}.$$

Then, we say that this sequence is $\beta-$mixing if $\beta(n) \to 0$ when $n \to \infty$.

[5]

$$\phi(n) = \sup_{j\geq 1} \sup \left\{ |\mathbb{P}(B \backslash A) - \mathbb{P}(B)|;\ A \in \wp_{n+j}^{\infty}, B \in \wp_1^j \ \ and\ \ \mathbb{P}(A) \neq 0 \right\}.$$

Then, we say that this sequence is $\phi-$mixing if $\phi(n) \to 0$ when $n \to \infty$.

provides a pointwise and uniform almost sure convergence as well as the asymptotic distribution of recursive kernel estimates under weak dependence in the sense of Laib and Louani (2010)[58]. The results are in fact based on the use of the version of exponential type inequality for partial sums of unbounded martingale differences (see Laib and Louani (2011)[59]) to get our results in the framework of ergodic variables.

1.4.1 Reminder of ergodic theory

Before entering the heart of our thesis, we recall some basic definitions of this theory. So many definitions of what ergodic term is would be, have been suggested in literature that it is difficult to determine a specific definition due to the use of many techniques and examples from several areas of mathematics such as probability and number theory, statistical mechanics as well as dynamical systems and functional analysis.

The birth and abstract setup of ergodic theory

Origins of Ergodic theory go back to statistical mechanics, particularly, in Maxwell's and Gibbs's theories. Thus, the word *ergodic* was introduced first by Boltzmann taking into consideration his hypothesis; which is a mixture of two Greek words "ergon odos" meaning "energy path". From another perspective, establishing a connection between the sets typically studied in statistical mechanics and the properties of single systems evolving in time. More specifically, the study and solve of problems for demonstrating the equality of infinite time averages and phase averages. Another definition says that; it is a part of the theory of the long-term statistical behavior of dynamical systems (at its simplest form, a dynamical system is a function T defined on a set Ω). In other words, ergodic theory deals with measure preserving actions of measurable maps on a measure space, usually assumed to be finite.

In an attempt by several scientists to understand the long-term statistical or probabilistic behavior of dynamical systems such as the motions of a billiard ball or the motions of the Earth's atmosphere, this theory arises. It mainly focuses on certain mathematical objects called abstract dynamical systems or measure-preserving flows so that the main idea is to ignore the related properties of the dynamical system and to focus only on summarizing the statistical properties. From mathematical point of view, abstract dynamical systems are natural objects that arise in many different contexts (even in areas as far a field as number theory). In fact, the first rigorous result in mathematics is the famous *Poincaré's Recurrence Theorem* appearing in 1890. However, the development of such theory is considered to have taken place in 1931 when Neumann and G.D. Birkhoff have proved the pointwise ergodic theorems. It is at this point that ergodic theory became a legitimate mathematical discipline and it thus entered the framework of functional analysis for sure. Whereas, in practice, the ergodic framework is more convenient because it does not need to verify any condition as in the $\alpha-$mixing case for example.

Characterization of Ergodicity

Definition 1.4.1. *(Measure preserving transformations) Let* $(\Omega, \mathcal{B}, \mu)$ *be a probability space, and* $T : \Omega \to \Omega$ *measurable. The map* T *is said to be measure preserving with respect to* μ*, if* $\mu(T^{-1}A) = \mu(A)$ *for any event* $A \in \mathcal{B}$*.*

This definition implies that the processes f, $f \circ T$, $f \circ T^2$,... are stationary, for any measurable function $f : \Omega \to \mathbb{R}$. Such that, for all Borel sets $B_1, \dots, B_n$ and all integers $r_1 < r_2 < \dots < r_n$, one have for any $k \geq 1$,

$$\mu(\{x : f(T^{r_1}x) \in B_1, \dots, f(T^{r_n}x) \in B_n\}) = \mu(\{x : f(T^{r_1+k}x) \in B_1, \dots, f(T^{r_n+k}x) \in B_n\}).$$

Definition 1.4.2. *Let* T *be a measure preserving transformation on a probability space* $(\Omega, \mathcal{B}, \mu)$*. The map* T *is said to be ergodic if every measurable set* A *satisfying* $T^{-1}A = A$ *such that* $\mu(A) = 0$ *or 1.*

Example 1.4.1. *The identity application of* $(\Omega, \mathcal{B}, \mu)$ *is ergodic* $\Leftrightarrow \forall A \in \mathcal{B}$, $\mu(A) = 0\, or\, 1$.

Definition 1.4.3. *(Martingale difference) Let* $(\eta_t)_{t\geq 0}$ *be a sequence of real random variables and* $(\wp_t)_{t\geq 0}$ *be a sequence of* $\sigma-$*fields. The sequence* $(\eta_t, \wp_t)_{t\geq 0}$ *is said to be a martingale difference (or a sequence of martingale increments) if and only if:*

1. $\wp_t \subset \wp_{t+1}$;

2. η_t *is* $\wp_t-$*measurable;*

3. $\mathbb{E}|\eta_t| < \infty$;

4. $\mathbb{E}[\eta_{t+1}/\wp_t] = 0$.

Remark 1.4.1. *Obviously, there is a strong relation between the both terms; martingale and martingale difference, such that:*

1. *If* $(Y_t, \wp_t)_{t\geq 0}$ *is a martingale and* $\eta_0 = 0, \eta_t = Y_t - Y_{t-1}$*, then* $(\eta_t, \wp_t)_{t\geq 0}$ *is a martingale difference and* $\mathbb{E}[\eta_{t+1}/\wp_t] = \mathbb{E}[Y_{t+1}/\wp_t] - \mathbb{E}[Y_t/\wp_t] = 0$.

2. *If* $(\eta_t, \wp_t)_{t\geq 0}$ *is a martingale difference and* $Y_t = \eta_0 + \eta_1 + \ldots + \eta_t$*, then* $(Y_t, \wp_t)_{t\geq 0}$ *is a martingale and* $\mathbb{E}[Y_{t+1}/\wp_t] = \mathbb{E}[Y_t + \eta_{t+1}/\wp_t] = Y_t$.

In addition, for continuous time $t \geq 0$*, the random variable* $(\eta_t, \wp_t)_{t\geq 0}$ *is a martingale difference from filtration* $\wp_t$ *if* η_t *is* $\wp_t-$*measurable and* $\mathbb{E}[\eta_t/\wp_{t-s}] = 0$, $t \geq 0$, $s \geq 0$.

Ergodic Theorem

One of the fundamental results (theorems) of the ergodic theory is the following theorem known as Birkhoff's Ergodic Theorem or the Individual Ergodic Theorem appeared in 1931 and which applies to both stationary and ergodic processes (for more basic information and proofs on this theorem, the reader is directed to Peskir (2000)[79] for *Birkhoff's (Pointwise) Ergodic Theorem*, *Von Neumann's (Mean) Ergodic Theorem* on the one hand, and *Kingman's (Subadditive) Ergodic Theorem* on the other hand). Indeed, this theorem is a generalization of the *Kolmogorov's (Strong) Law of Large Numbers (SLLN)* which states that for a sequence $X = \{X_k;\ k \geq 1\}$ of i.i.d. random variables on the probability space $(\Omega, \mathcal{B}, \mu)$, with $\mathbb{E}|X_k| < \infty$, one have

$$\frac{1}{n}\sum_{k=1}^{n} X_k \longrightarrow \mathbb{E}(X_1) \ as \ n \to \infty \ \ a.s. \tag{1.2}$$

The Ergodic Theorem is originally proved by G.D. Birkhoff in 1931 and presented as follows

Theorem 1.4.1. *(Birkhoff's (Pointwise) Ergodic Theorem) Let $(\Omega, \mathcal{B}, \mu)$ be a probability space and $T : \Omega \to \Omega$ a measure-preserving transformation. Then, for any f in $L^1(\mu)$, there exists*

$$\frac{1}{n}\sum_{k=0}^{n-1} f(T^k x) \longrightarrow f^*(x)\ \mu - a.s.,\ as\ n \to \infty$$

is $T-$invariant and $\int_\Omega f d\mu = \int_\Omega f^* d\mu$. If moreover T is ergodic, then f^* is a constant a.s. and $f^* = \int_\Omega f d\mu$.

Under these assumptions, the temporal and spatial means of the observations are almost surely equal.

Definition 1.4.4. *(Stationary sequence) A process $(X_k)_{k\in\mathbb{U}}, \mathbb{U} = \{\mathbb{R}^+, \mathbb{Z}\}$ is said to be strictly stationary or stationary in the strict sense if the joint laws of $(X_{k_1}, \ldots, X_{k_s})$ and of $(X_{k_{1+h}}, \ldots, X_{k_{s+h}})$ are identical for any positive integer s and for all $k_1, \ldots, k_s$, $h \in \mathbb{Z}$.*
In other words, for $n \in \mathbb{N}$, we say that X_n is stationary process if for all $h \geq 1$ we have

$$\{X_{n+h};\, n \geq 0\} \stackrel{\mathcal{L}}{=} \{X_n;\, n \geq 0\}.$$

Theorem 1.4.2. *If $(X_k)_{k\geq 0}$ is a stationary real process and X_0 is integrable, then*

$$\frac{1}{n}\sum_{k=1}^{n} X_k \longrightarrow \mathbb{E}[X_0/\wp_k]\quad as\quad n \longrightarrow \infty \quad a.s.$$

where $\wp_k$ is the $\sigma-$algebra of invariant sets. If moreover, the process is ergodic, the limit coincides with the expectation of the variable X_0.

Remark 1.4.2. *There is an imperative link between stationarity and ergodicity such that: ergodicity leads to stationarity but the converse is false.*

Definition 1.4.5. *(Continuous time process) Let $(X_t)_{t\geq 0}$ be a process defined on $(\Omega, \wp, (\wp_t)_{t\geq 0}, \mathbb{P})$. The process $(X_t)_{t\geq 0}$ is said to be measurable if the following application is measurable with respect to $\mathcal{B}((0,\infty)) \otimes \wp_t$,*

$$\begin{aligned} X \quad : \quad & (0,\infty) \times \Omega \to (E, \varepsilon) \\ & (t, w) \longrightarrow X_t(w) \end{aligned}$$

In addition, the process $(X_t)_{t\geq 0}$ is said to be adapted if $\forall t \geq 0$, $(X_t)_{t\geq 0}$ is $\wp_t-$measurable.

Then, the previous pointwise ergodic theorem is given in the following result

Theorem 1.4.3. *(Birkhoff's ergodic theorem in continuous time). We say that a stationary continuous random process $(X_t)_{t\geq 0}$ is ergodic if*

$$\frac{1}{T}\int_{t=0}^{T} X(u)du \longrightarrow \mathbb{E}[X_0] \quad as \quad T \longrightarrow \infty \quad a.s.$$

1.4.2 Bibliographic context in i.i.d and $\alpha-$mixing cases

For an infinite dimensional covariate X, the first result refers to the work of Farraty et al. (2005)[33]. They have studied the almost complete convergence of a conditional density estimator and have examined an application of forecasting via the conditional mode defined by the random variable maximizing the conditional density in the i.i.d. case. On recent developments, by applying the small-ball probability theory, the authors have generalized this result to the α-mixing case in (2005). In the same framework of mixing functional observations, Masry (2005)[66] has systematically showed the asymptotic normality of Ferraty and Vieu's (2004)[38] estimator for the regression function.

Subsequently, Ferraty et al. (2006)[34] have constructed a double kernel estimator for a conditional distribution function and have specified the almost complete convergence with rates of this estimator on the one hand. The contribution of Ferraty et al. (2006)[34] also extends to functional statistics the estimation of the conditional density function and its derivatives of real random variable. These latter have assumed a probabilistic distribution on functional data and have obtained the almost complete convergence in the i.i.d. case. On the other hand, Ferraty in collaboration with Vieu (2006)[39] have established the almost complete convergence's rates of the conditional mode kernel estimator. Therefore, the estimation of the conditional distribution function has been also treated as a preliminary study of conditional quantile estimate. For example, Ezzahrioui and Ould-Saïd (2005[30], 2006[31]) have studied the asymptotic normality of this estimator in both cases (i.i.d and α-mixing). The convergence in L^p norm of the conditional density estimator has been interested obviously by Dabo-Niang and Laksaci (2007)[21] in both cases of finite and infinite dimensional regressors where the observations are i.i.d., whose asymptotic results are closely related to the property of small balls probability measure known as concentration property. While, Ezzahrioui and Ould-Saïd (2008)[32] gave the asymptotic normality of the conditional mode kernel estimator in the i.i.d. case. More recently, Ling and Xu (2012)[62] have discussed the asymptotic properties of semi-parametric conditional density estimate for functional time series data. They have mainly interested in the asymptotic normality of their proposed estimator as well as of the conditional mode in case of mixing processes.

Certainly, some work aimed at using the recursive method cannot be ignored. Given that this method is new, the work in it was rather few. We mention the most prominent ones: Given $\alpha-$mixing observations, Amiri (2010)[1] has established the consistency of the estimators defined above. Bouadjemi (2014)[13] has introduced a new nonparametric estimator of the conditional cumulative distribution function of a scalar response variable Y given a functional random variable X. To prove his result (asymptotic normality), he has hypothesized specific regularity conditions generated by the functional model. In the context of incomplete information, we mention among others, Wang and Liang (2004)[93] who have showed the almost sure convergence of truncated version of recursive estimator under $\phi-$mixing.

1.4.3 Bibliographic context in ergodic case

This part is devoted to a brief presentation of the results already available in the literature for ergodic data setting. We argue that these contributions are a direct continuation of the works done in the functional framework of the previous cases.

Therefore, based on the introduced Collomb and Härdle (1986)[20] estimate of the auto-regression model, Laib and Ould-Saïd in (2000)[57] have raised the question whether this estimate is still uniformly consistent for stationary ergodic process and it is answered positively. Subsequently, in functional statistics, for the same case and by using the well-known Nadaraya-Watson estimator, Laïb and Louani in (2010)[58] have considered the regression estimator of a real random variable Y on a functional one X. They have studied in fact the asymptotic properties of the estimate including the convergence in probability with rate, as well as the asymptotic normality which induces a confidence interval for this function, when the considered functional data are stationary and ergodic.

Sekkal et al. (2013)[88] have interested in extending the results introduced by Azzeddine et al. (2008)[6] for the estimation of the robust regression, in order to use them in the derivation of their results for handling different types of models. They have presented a $\psi-$regression function estimator using the robust method belongs to the class of non-parametric $M-$estimations introduced firstly by Huber in 1964 in the case where the explanatory process is functional. Then, the almost complete convergence is established in the case when the observations are ergodic. Whereas, the asymptotic normality of such model is discussed later by Benziadi et al. (2016)[8].

For recursive nonparametric kernel estimation of the conditional quantile of a scalar response variable Y with ergodic hilbertian explanatory variable X, Benziadi et al. (2016)[9] have considered two type of estimators: the first one is given by inverting the double kernels estimate of the conditional distribution function whereas the second is obtained by using the robust approach. They have achieved the almost complete consistency as well as the asymptotic normality under a stationary ergodic process assumption. Also, based on the same kernel method, the estimator of the conditional mode function is studied simultaneously by Ardjoun et al. (2016)[4]. These authors have considered an alternative estimator of this function when the explanatory variable is supposed to be functional and then, under the ergodicity hypothesis, they have quantified the asymptotic properties of their proposed estimate.

In recent decades, there has been a renewed interest in incomplete data modeling and statistical analysis in the scientific world. In fact, several studies take into account the existence of such data, as it cannot be ignored. It certainly has a clear impact on the approach we use. Moreover, to be more logical, in statistical studies that include many disciplines, such as lifetime studies and others that include the presence of samples that can often be incomplete. Statisticians are interested then in finding ways, methods and solutions to circumvent this difficulty by modeling it and then to see the extent of its impact. For these reasons, it is interesting to consider another theory, the theory of this kind of observations. In addition, due to numerous applications in different fields such as medical, social or economic, great importance is given to the case where the survival time may present some form of dependency. For example, in clinical trials, it frequently happens that patients in the same hospital have correlated survival times due to unmeasured variables such as the quality of the hospital's equipment. For more details about dependence in data, we can quote Lipshitz and Ibrahim (2000)[63].

1.5 Foundations of survival analysis theory: some generalities and main applications

It is well known that survival data analysis or generally speaking event history analysis is a group of statistical methods, that has been extended to deal with several fields of application to develop over the last few years in a variety of areas, prepared to analyze interval data and proposed as a rule to assess time to event data in one or more groups of individuals, such as: in randomized clinical trials and cohort studies. Survival theory is used then to evaluate incomplete data and the changes associated with their occurrence. The theoretical results on the statistical behavior of such observations allow us to propose a rigorous mathematical framework to perform such extrapolations. Thus, a survival time is a positive random variable (r.v) that measures the time from a particular starting point to a particular endpoint of interest i.e. the elapsed time since a specific *event under study* occurred and/ or until it occurs. Indeed, in most prospective studies, individuals are followed for an observation period fixed in advance.

For more explanation, the event under study is the irreversible transition between two states, for instance:

- Start of treatment⟶Time of death, Development of functional ability…
- Time of marriage⟶Birth of the first child, Divorce…

Thus, in such a field, the purpose we are interested in is the delay of occurrence of any well-defined event (i.e. one may want to know how long it takes before a certain event occurs). In other words, to examine the behavior of a population that experiences 'failures' over time. For example, this failure could be a part in an automobile wearing out or a subject in a clinical trial dying. In the biomedical domain, these durations can also contribute in the context of longitudinal studies such as cohort surveys (following patients over time: the time to remission or recrudescence of a patient) or therapeutic trials and the response to a given treatment (evaluating the effectiveness of a drug), the relapse of a disease or death. In industry, it identifies the time between two successive breakdowns, while in finance it refers to the inflation time of a stock index…and so on. In demography, it is used to construct life tables. These are used by actuaries to determine the amount of life insurance and annuities so that to understand many social and economic problems.

Equally important, the terminal event is not necessarily death: it can be the onset of a disease, the cure or the breakdown of a machine…

Survival function

In survival analysis, one seeks to draw conclusions for the non-negative random variable of interest Z referred as a lifetime of n objects under study from such incomplete data. Ordinarily, one of the most important and used function in such field of application and which best characterizes the distribution of this latter (i.e. explicit the probability that an individual survives to a given time point z) is the unknown *unconditional survival function* $K(z)$ denoted as

$$\begin{aligned} K(z) &:= \mathbb{P}(Z > z) \\ &= 1 - M(z). \end{aligned} \tag{1.3}$$

Equivalently, the probability distribution function $M(z) := \mathbb{P}(Z \leq z)$. In the literature, several authors have interested in estimating by a parametric or non-parametric approaches this function or functions that can be deduced from it, such as *the probability density function* $\xi(z) \geq 0$ defined $\forall t \geq 0$ by

$$\xi(z) = -\frac{dK(z)}{dz}$$

or *the hazard rate function* $h(z)$ expressed as

$$h(z) = \frac{\xi(z)}{K(z)} = -\frac{d\ln[K(z)]}{dz}$$

and *the cumulative hazard function* $H(z)$ that represents a measure of the risk of the occurrence of an event and given by

$$H(z) = \int_0^z h(u)du = -\ln[K(z)].$$

The incompleteness of data is fundamentally due to two principal phenomena that generate this type of processes: *Censoring and Truncation.* So that, the right-censored and left-truncated problems are often the most described in the setup of survival observations. Whereas, the modeling of incomplete data is an active field of research, especially because of the importance of their impacts on several fields. In particular, in recent years, there has been a growing interest in applying survival data theory to the modeling of such observations. Thus, these two models can be used to address this issue and also to get explicitly the estimators of the distribution of survival times (i.e the survival function) in a way that allow us to differentiate between them respectively or to analyze the way in which explanatory variables modify the survival functions. Also, as an additional information, traditional statistical tools (developed in a classical universe) are not adapted to the incomplete behavior: the classical empirical estimators do not exist in the new theory. So that, it is necessary to shed light on the problem of estimating this unbiased function in both cases that requires another approaches.

Another case is possible, when both left-truncation and right-censoring occur in the same sample, and this is called the LTRC (Left-Truncated and Right-Censored) model. It is also a common type in recent years, especially. Thus, we observe the triplet (Y, T, Δ) if and only if

$$Y \geq T, \ \ with \ Y = Z \wedge W \ \ and \ \Delta = \mathbb{I}_{[Z \leq W]}$$

where Z is the survival time, T the truncation time and W the time of censorship.

In what follows, we discuss these two concepts, provide a brief structured overview and recall some basic vocabularies used to designate a positive random variable in order to highlight and understand the difference between them, to derive under monotone constraints the nonparametric maximum likelihood estimator for the both cases and to fix notations.

1.5.1 Right-censorship

In survival studies, many scholars have encountered with censored data during their research. Unfortunately, censored data make analysis more complicated, because of exact event times are not observed and then the subject is censored. Therefore, the questions that are constantly and repeatedly asked in these studies are:

1. How to deal with censored data ?
2. What are the approved methods and how are they included in the statistical framework ?

As is clear and known to most of us, the analysis of incomplete data necessitates an acclimatized methodology to take into consideration the information contained in the censoring delay. In fact, when collecting survival data, the censoring is the most commonly encountered phenomenon which essentially represents censored data, i.e. observations for which the exact value of an event is not always recognized due to several facts or if a participant drop out, loss to follow-up or die. Despite this, we still have partial information that allow us to set a lower limit (right censoring) or an upper limit (left censoring), i.e. an information of the type $Z \leq M$ or $Z \geq m$ when we know the boundaries of an event. And since in most studies only one of the two bounds is known, this means that the available information is very pauper, pauper than to say that $Z \in [m, M]$. In particular, right-censoring is the most prevalent example of incomplete observations in survival analysis and here we can define it by the fact that the individual does not experience the event of interest on his last visit.

Example In clinical trials, there are different reasons for censoring, arise when studying "time to event" data (the event here is "death". It can be any event of interest), include:

- Some individuals are still alive at the end of the study or analysis so the event of interest has not occurred. Therefore we have right censored data.
- Loss to follow-up after time W: patients stop coming to clinic or move away (care in another hospital) or change of treatment (side effects or ineffectiveness). Here, W censors Z to the right, since, for them, the survival time Z is unknown but greater than $W : Z > W$.
- Deaths from other causes: competing risks.

The development of the proceedings used in such a framework requires several changes to construct for instance the likelihood function for this type of data. First of all, we shall present more standard methodologies and re-parametrization of the observed data as in [18] or [22]. Whereas, the contribution of [48] discusses statistical methods for the analysis of lifetime data and provides many interesting changes. The standard and some add-on survival packages in $\mathcal{R}$ routine handled with censored data are developed by [45], [46] and [47] by introducing deferent functions and demonstrating how it can be used in estimation. Thus, the results produced by these packages are very satisfactory.

We consider now the case of an observed population of n individuals. For individual number $k = 1, \ldots, n$; let us consider the random variable Z_k its strictly stationary, non-negative lifetime with the unknown distribution function $F(z)$ and density $\xi(z)$. Thus, due to the traditional right-censorship problem, we do not observe $(Z_k)_{k\geq 1}$ but only observe a censored version consist of n realizations; denoted by

$\{(Y_k, \Delta_k), k = 1, \ldots, n\}$ with $Y_k = Z_k \wedge W_k, k = 1, \ldots, n$ is the actual observed time, Δ_k is the nonnegative Bernoulli random variable (the censoring indicator stores the information) which allows us to know the nature of the observed data Y (i.e. if it is a true duration Z or if it is a censoring W) and such that

$$\Delta_k = \begin{cases} 1 & if \quad Z_k \leq W_k \\ 0 & if \quad Z_k > W_k \end{cases} \quad k = 1, \ldots, n$$

i.e. we observe the actual survival time if and only if it is less than W_k. Therefore, in such case, the observations are uncensored and $\Delta_k = 1$. Otherwise, if $\Delta_k = 0$, the observations are said to be right-censored by $(W_k)_{k \geq 1}$ which denote the censor points or detection thresholds (censoring times), with $z \wedge w$ is the minimum of z and w. Thus, the positive random variables Y_k have distribution function H defined by:

$$H(z) = 1 - (1 - F(z))(1 - G(z)) = 1 - \overline{F}(z)\overline{G}(z), \quad z \in \mathbb{R}.$$

This model assumes also that the positive time W which is caused by the censoring scenario has an unknown survival function $\overline{G}(z) := \mathbb{P}[W > z], \; z \in \mathbb{R}$. We put now for any distribution function L, the smallest upper bound of the support denoted by

$$T_L := \sup\{s, L(s) < 1\} \leq \infty. \tag{1.4}$$

Through which we can define each boundary of the two functions F and G respectively, by

$$T_F := \sup\{s, F(s) < 1\}, \qquad T_G := \sup\{s, G(s) < 1\}.$$

In general, note that these latter are unknowns too since F and G are unknowns. In the context of estimation problems when right-censoring mechanism is considered, such as in the diabetes study, usually, the technical procedures involve the regularity assumption of the independence between the variables: the possibly observable survival time Z and the right-censoring time W, (or (X, Z) and W) which is considered as a strong and important supposed hypothesis in the formulation of the likelihood. Note that, if this assumption does not hold, standard survival theory does not apply. Some prominent examples in the literature such as the cohort studies do not meet this necessary condition. Thus, this is what makes it impossible to cover the currently available status and by all accounts more troublesome to improve on the numerical issue in the rate of convergence's analysis and therefore, a proper analysis should be performed.

Back to the main point now, in case that no censoring arises, by using the optimal empirical distribution function, one can estimate the distribution function $F(z) := \mathbb{P}(Z \leq z)$ by:

$$\widehat{F}_n^Z(z) := \frac{1}{n} \sum_{1 \leq k \leq n} \mathbb{I}_{[Z_k \leq z]}. \tag{1.5}$$

While that this latter is not allowed in practice as one does not observe all the Z_ks in the presence of censoring but it is considered as a reference and a strongly uniform consistent estimate.

Then, we can conclude certainly $\widehat{S}_n^Z(z) = \widehat{\overline{F}}_n^Z(z)$ by the relationship $1 - \widehat{F}_n^Z(z)$ which is the main aim of researchers who are not familiar with the topic of survival analysis before and are asked to estimate it using only the observed data $(Y_k, \Delta_k)_k$. Indeed, they have proposed a first estimator obtained directly from the observations $(Y_k)_k$, by:

$$\widehat{S}_1^Y(z) = \frac{1}{n} \sum_{1 \leq k \leq n} \mathbb{I}_{[Y_k > z]}.$$

Whereas, a second estimator is obtained from only the uncensored observations $(Y_k, \Delta_k = 1)_k$, by:

$$\widehat{S}_2^Y(z) = \frac{1}{n} \sum_{1 \leq k \leq n} \mathbb{I}_{[Y_k > z, \Delta_k = 1]}.$$

However, in both scenarios, this is not the case where censorship is present, because it involves non observed quantities (the Y_k) and then $\widehat{S}_n^Z(z)$ is not calculable such that some simulation experiments have showed that these two naive estimators are behaved poorly in presence of censored data and are always failed to provide accurate estimations. In such case, we stress that the first proposed estimate of the survival function of a positive random lifetime; that generalizes the empirical one for right censorship, is the famous estimator suggested by Kaplan and Meier in 1958[51], based on the effective nonparametric maximum likelihood [6] estimate (NPMLE) and written as down:

$$\overline{F}_n^Z(z) := \begin{cases} 0 & ; z \geq Y_{(n)} \\ \displaystyle\prod_{k=1}^{n} \left(1 - \frac{d_k}{n_k}\right)^{\mathbb{I}\left(Y_{(k)} \leq z\right)} & ; z < Y_{(n)} \end{cases} \tag{1.6}$$

where $Y_{(1)} < Y_{(2)} < \cdots < Y_{(n)}$ represent the n distinct survival times among $(Y_1, \ldots, Y_n)$, d_k is the number of deaths at $Y_{(k)}$ and n_k is the number of individuals still at risk at time $Y_{(k)}$ defined by $n_k = \sum_{j=1}^{n} \mathbb{I}_{(Y_j \geq Y_{(k)})}$.

Because that the previous estimator is of very limited interest. Thus, another equivalent form which will be adopted in the rest of this thesis is presented as follows:

$$\begin{aligned} \overline{F}_n^Z(z) &= \overline{F}_n^Z(z, \{(Y_1, \Delta_1), \ldots, (Y_n, \Delta_n)\}) \\ &:= \begin{cases} 0 & ; z \geq Y_{(n)} \\ \displaystyle\prod_{k=1}^{n} \left(\frac{n-k}{n-k+1}\right)^{\left(\Delta_{(k)}\mathbb{I}\left(Y_{(k)} \leq z\right)\right)} & ; z < Y_{(n)} \end{cases} \end{aligned} \tag{1.7}$$

in addition to the survival function of the censoring time W,

$$\begin{aligned} \overline{G}_n^W(z) &= \overline{G}_n^W(z, \{(Y_1, \Delta_1), \ldots, (Y_n, \Delta_n)\}) \\ &:= \begin{cases} 0 & ; z \geq Y_{(n)} \\ \displaystyle\prod_{k=1}^{n} \left(\frac{n-k}{n-k+1}\right)^{\left((1-\Delta_{(k)})\mathbb{I}\left(Y_{(k)} \leq z\right)\right)} & ; z < Y_{(n)} \end{cases} \end{aligned} \tag{1.8}$$

Denoting here $Y_{(1)} < Y_{(2)} < \cdots < Y_{(n)}$ the order statistics of $(Y_k)_{k \in \{1,\ldots,n\}}$, along with their corresponding concomitant $\Delta_{(k)}$. Then, the previous estimators satisfy the following important properties.

[6]The maximum likelihood is the most popular and under certain conditions is the most effective method

Theorem 1.5.1. *(Peterson (1977)[77]) If F and G have no jumps in common, for all $z < T_K$. One have*

$$F_n(z) - F(z) \to 0 \ \ as \ \ n \to \infty \ \ a.s. \tag{1.9}$$

and

$$G_n(z) - G(z) \to 0 \ \ as \ \ n \to \infty \ \ a.s. \tag{1.10}$$

On the same path, Stute in collaboration with Wang in 1993 [89] have obtained the strong consistency for G under certain additional assumptions on the functions F and G.

Theorem 1.5.2. *(Stute and Wang (1993)[89]) Assume that F and G do not have jumps in common. Then, for $n \to \infty$*

$$\sup_{0 \le z \le T_K} |G_n(z) - G(z)| \to 0 \quad a.s. \tag{1.11}$$

if and only if $\mathbb{P}[W = T_K] \geq 0$ and $\mathbb{P}[Z \geq T_K] > 0$ with $T_K = T_F \wedge T_G$.

Whereas, the nonparametric estimate of the conditional survival function which generalizes the kaplan-meier one has been pioneered by Beran (1981)[10] and under suitable hypotheses on its structure and that Z and W are conditionally independent given X, Beran (1981)[10] and Dabrowska 1989[23] have showed that this estimate is weakly and strongly consistent. In such context, Lecoutre and Ould Saïd (1995)[60] have also established the almost complete uniform convergence of the conditional Kaplan-Meier estimator under mixing condition on the underlying distribution.

Remark 1.5.1. *After researching of the theoretical properties of the nonparametric Kaplan-Meier's estimator (NPKME). It is important to stress that:*

- *It is likewise called a product-limit estimator that is the handiest tailored approach in non-parametric statistics of estimating survival distribution of a right-censored r.v Z in practice and it is a very accurate estimator.*
- *It is not recursive but convenient and using it can slightly reduces the efficiency of the estimation in terms of computation time.*

- *It is not based on the assumption about the underlying probability distribution of failure times, which makes sense since survival data has a skewed distribution.*
- *It does not account for confounding or effect modification by other covariates.*

1.5.2 Left-truncation

Besides the first type, there is another interesting type related to incomplete data that is no less important than it; so-called "Left-Truncation". Statistically, the survival time in some studies is subject to left truncation, where the truncated survival time is the duration from disease onset to recording time. To be more precise, the experimental studies that consider lifelong events that must exceed a certain threshold. Or rather, the variable of interest Y must be greater than a given truncation variable T in order to be observable, are modeled by the left-truncation model (i.e. a survival time can be observed if and only if $Y \geq T$ and then the sampling weight depends on the underlying truncation time distribution).

Newly, there has been a growing interest in researching ways to make better use of the information about the truncation time. So that, compared with the type described in the previous paragraph, it is well known that there is some specific form that the left-truncation distribution follows. In other word, there is a striking similarity of the invented definition of the Kaplan-Meier's product limit estimator related to the case of right-censored observations to that of Lynden-Bell and Woodroofe's product limit estimator related to the left-truncated case. This supports the need for a new theory and it is worth noticing that there is a strong relationship between the both cases.

Under left truncation model, we observe (Y, T) if and only if $Y \geq T$, so among the total number in the pooled sample N, we observe the couple $\{(Y_k, T_k), k = 1, \ldots, n\}$ with $n \leq N$ (n is known compared with N) which has the same joint distribution as (Y, T). From the same subject, we impose the usual independent truncation assumption by assuming that the left truncation variable T is independent with the failure time Y. This latter condition is identical to that made with right-censoring data and credible in this case. Thus, for an event time subject to left-truncation, if the truncation variable has a continuous distribution with support $[0, a_F]$ where $a_F \in (0, \infty)$, then, the well known Kaplan-Meier estimator of the law of a positive random variable for right-censored case is consistently converted after some important modifications as the approach suggested first by Lynden-Bell (1971)[64], studied then by Woodroofe (1985)[95], and finally completed by another modification by Tsai, Jewell, and Wang in the 1980's [92], [90] and Gu and Lai (1990)[42], Lai and Ying (1991)[56], Gijbels and Wang (1993)[41] also to become the modified version as follow:

$$1 - F_n(y) = \prod_{s \leq y} \left[1 - \frac{F_n^*(s)}{K_n(s)} \right] \quad and \quad 1 - G_n(t) = \prod_{s > t} \left[1 - \frac{G_n^*(s)}{K_n(s)} \right].$$

These latter are then considered as the most widely accepted for estimating the marginal survival function of the nonterminal phenomenon and may be biased also. Afterward, Woodroofe (1985)[95] has further established the consistency of these product-limit estimates and has investigated the cases of decreasing hazard and discrete versions of the problem.

1.6 Modeling conditional models in the presence of censoring

In nonparametric regression analysis for such kind of observations, one wants to construct estimates of Y (the lifetime, from a study of the incomplete data of a series) after having observed X. Indeed, these estimates all depend on the unknown survival function G of the censoring time. Thus, the foundation stone is the application of the theory of these data. In the very recent statistical literature, there are so far published results that have been highlighted a special attention on analyzing and studying the convergence of several nonparametric estimates of the conditional models for the problem intensively posed by these incomplete observations of a series (particularly, the case where these latter are incomplete by right-censored data), both in a theoretical framework and application, which show the weak and strong consistency of various estimates with respect to the censoring mechanism and dependency of the lifetime and the censoring time. Therefore, the purpose of this part is to display the scope of the models investigated.

As far as the author of the present thesis know, Khardani et al. (2010)[52] have addressed the first nonparametric estimation of conditional mode function when the observations are subject to right-censoring, given by $\Theta(x) := \arg\max_{t\in\mathbb{R}} \zeta(t/x)$ for a sample of n i.i.d. observations $\{(X_k, Y_k, \delta_k), 1 \leq k \leq n\}$, $Y_k := min(T_k, C_k)$ and $\delta_k = \mathbb{I}_{[T_k \leq C_k]}$. Since the definition of the estimate is based on the function G (not F), they have replaced G with its Kaplan-Meier estimator G_n to create their estimate and such that

$$\widehat{\zeta}_n(t/x) = \frac{\sum_{k=1}^{n} \delta_k \overline{G}_n^{-1}(Y_k) L_1\left(a^{-1}(x - X_k)\right) L_2^{(1)}\left(a^{-1}(t - Y_k)\right)}{a \sum_{k=1}^{n} L_1\left(a^{-1}(x - X_k)\right)}. \tag{1.12}$$

Then, under the assumption that Y and C are conditionally independent given X, they have showed that this estimate is strongly consistent and asymptotically normal.

Inspired by the previous work, in the context of functional stationary ergodic data, Chaouch and Khardani (2014)[16] have investigated the conditional quantile estimator of a randomly censored scalar response variable by considering a kernel-based estimator of $F^x(t)$ given by

$$\widehat{F}_n^x(t) = \frac{\sum_{k=1}^{n} \delta_k \overline{G}_n^{-1}(Y_k) L_1\left(a^{-1} d(x, X_k)\right) L_2\left(b^{-1}(t - Y_k)\right)}{\sum_{k=1}^{n} L_1\left(a^{-1} d(x, X_k)\right)}. \tag{1.13}$$

They have stated the strong consistency rates as well as the asymptotic distribution of the latter under the following important assumptions.

(**A1**) L_1 is a nonnegative bounded kernel of class $\mathcal{C}^1$ over its support $[0, 1]$ such that $L_1(1) > 0$ and $L_2^{(j)}$ satisfies the Lipschitz condition and $\int |u|^\nu L_2^{(1)}(u)du < \infty$.

(**A2**) For $x, x' \in E$, $F_x(a) = \phi(a) f_1(x) o(\phi(a))$ as $a \to 0$.

(**A3**) The conditional df $F^x(t)$ and its derivative $f^x(t)$ satisfy: $\int_{\mathbb{R}} |t| f^x(t)dt < \infty$ and $|F^{(j)}(t_1/x_1) - F^{(j)}(t_2/x_2)| \leq C_x \left(d^\beta(x_1, x_2) + |t_1 - t_2|^\nu\right)$.

(**A4**) For any $m \geq 1$ and $j = 0, 1$, $\mathbb{E}[(L_2^{(j)}(b^{-1}(t - T_k)))^m/\mathcal{G}_{k-1}] = \mathbb{E}[(L_2^{(j)}(b^{-1}(t - T_k)))^m/X_k]$ and $\sup_{t\in\mathcal{S}} |\mathbb{E}[L_2^m(b^{-1}(t - T_1))/X_1 = x']| < \infty$.

(**A5**) $(C_n)_{n\geq 1}$ and $(X_n, T_n)_{n\geq 1}$ are independent.

(**A6**) $\mathbb{E}[(\delta_k \overline{G}^{-1}(Y_k) L_2 (b^{-1}(t - Y_k)) - F(t/X_k))^2/\mathcal{G}_{k-1}] = W_2(t/X_k)$ almost surely.

(**A7**) The df of the censored random variable G has a bounded first derivative $G^{(1)}$.

Theorem 1.6.1. *(Chaouch and Khardani (2014)[16]) Assume that conditions (A1)-(A5) hold true and* $\dfrac{\log n}{n\phi(a)} \to 0$ *as* $n \to \infty$, *then*

$$|\widehat{q}_{n,\alpha}(x) - q_\alpha(x)| = O_{a.s.}(a^\beta + b^\nu) + O_{a.s.}\left(\sqrt{\frac{\log n}{n\phi(a)}}\right).$$

and

Theorem 1.6.2. *(Chaouch and Khardani (2014)[16]) Assume that assumptions (A1)-(A7) hold true, then we have*

$$\sqrt{n\phi(a)}\,(\widehat{q}_{n,\alpha}(x) - q_\alpha(x)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \gamma^2(x, q_\alpha(x))),$$

where $\gamma^2(x, q_\alpha(x)) = \dfrac{M_2}{M_1^2 f_1(x)} \dfrac{\alpha[\overline{G}^{-1}(q_\alpha(x)) - \alpha]}{f^2(q_\alpha(x)/x)}$ *and* $M_j = L_1^j(1) - \int_0^1 (L_1^j)'\tau_0(u)du$.

For this type of data, that takes values in infinite dimensional space and is $\alpha-$mixing, Khardani and Thiam (2016)[54] have interested in the estimation of the classical conditional mode function defined as

$$\widehat{\Theta}_n(x) = \arg\sup_{t\in\Omega} \widehat{\zeta}_n^x(t)$$

where

$$\widehat{\zeta}_n^x(t) = \frac{\dfrac{1}{b}\displaystyle\sum_{k=1}^n \delta_k \overline{G}^{-1}(Y_k) L_1\left(a^{-1}(\|x - X_k\|)\right) L_2\left(b^{-1}(t - Y_k)\right)}{\displaystyle\sum_{k=1}^n L_1\left(a^{-1}(\|x - X_k\|)\right)}.$$

Such that, under certain conditions that we specify now

(**A1**) L_1 is a function with support $(0,1)$ such that $c_1 \mathbb{I}_{(0,1)} \leq L_1 \leq c_2 \mathbb{I}_{(0,1)}$, $c_1 > 0, c_2 > 0$. In addition, L_2 is a bounded Lipschitz function such that $\int_{\mathbb{R}} L_2(u)du = 1$ and $\int_{\mathbb{R}} |u|^{b_2} L_2(u)du < \infty$.

(**A2**) There exist $b_1 > 0, b_2 > 0$ such that $\forall (t_1, t_2) \in \Omega^2, \forall (x_1, x_2) \in \mathcal{F} \times \mathcal{F}$,

$$|\zeta(t_1/x_1) - \zeta(t_2/x_2)| \leq c \left(\|x_1 - x_2\|^{b_1} + |t_1 - t_2|^{b_2} \right).$$

(**A3**) There exists $\epsilon_1 \in (1/2, 1)$ such that $\sup_{i \neq j} \mathbb{P}[(X_i, X_j) \in B^2(x,a)] \leq (\phi_x(a))^{1+2\epsilon_1}$ and $b^2 \phi_x(a)^{\epsilon_1 - 1} \to \infty$.

(**A4**) There exists $\epsilon_2 \in (0,1)$ such that $\nu > \dfrac{1+\epsilon_2}{\epsilon_1 \epsilon_2}$ and $b\phi_x(a) = O(n^{-\epsilon_2})$.

(**A5**) $(X_n, T_n)_{n \geq 1}$ is a sequence of stationary $\alpha-$mixing rvs with coefficient $\alpha(n) = O(n^{-\nu})$, for some $\nu > 1$. Also, $(C_n)_{n \geq 1}$ and $(X_n, T_n)_{n \geq 1}$ are independent.

They have found that their estimate is almost completely convergence with rates.

Theorem 1.6.3. *(Khardani and Thiam (2016)[54]) Assume that (A1)-(A5) hold. In addition to the condition* $\dfrac{\log n}{nb\phi_x(a)} \to 0$ *as* $n \to \infty$*. Then, we have*

$$\left|\widehat{\Theta}_n(x) - \Theta(x)\right| = O\left(a^{\frac{b_1}{2}}\right) + O\left(b^{\frac{b_2}{2}}\right) + O\left(\left(\frac{\log n}{nb\phi_x(a)}\right)^{1/4}\right) \quad a.c. \quad as \ n \to \infty.$$

As a quantum leap, Khardani and Semmar (2014)[53] are the first authors to consider *recursive estimate* in *censored data context.* They have derived the almost sure uniform strong consistency with rates of convergence and the asymptotic normality of the kernel's recursive conditional density function estimator defined for $d-$dimensional co-variate by

$$\widehat{\phi}_n(t/x) = \frac{\sum_{k=1}^{n} a_k^{-(d+1)} \delta_k \overline{G}_n^{-1}(Y_k) L_1\left(a_k^{-1}(x - X_k)\right) L_2\left(a_k^{-1}(t - Y_k)\right)}{\sum_{k=1}^{n} a_k^{-d} L_1\left(a_k^{-1}(x - X_k)\right)} := \frac{\widehat{g}_n(x,t)}{l_n(x)}$$

where $g(\cdot,\cdot)$ is bounded function twice differentiable, $l(\cdot)$ is twice differentiable and satisfies a Lipschitz condition with $l(x) > \Gamma$ for all $x \in \mathcal{C}$ and $\Gamma > 0$ and let $\mathcal{C}$ and Ω be two compact sets of $\mathbb{R}^d$ and $\mathbb{R}$ respectively. Additionally,

(**A**) The kernels L_1 and L_2 are Lipschitz continuous functions and compactly supported, satisfy $\int_{\mathbb{R}^d} u_l L_1(u)du = 0,$ for $l = 1, \dots, d$ with $u = (u_1, \dots, u_d)^T$ and $\int_{\mathbb{R}} \nu L_2(\nu)d\nu = 0.$

(**C1**) The bandwidths satisfy $\lim_{n\to\infty} a_n^+ + \dfrac{\log n}{n a_n^{-d+1}} = 0$ and $\lim_{n\to\infty} n^\beta a_n^- = \infty, \forall \beta > 0.$

(**C2**) The variables $\{(X_k, T_k), k \geq 1\}$ and $\{C_k, k \geq 1\}$ are independent.

(**N**) $\lim_{n\to\infty} \dfrac{1}{n} \sum_{k=1}^{n} \left(\dfrac{a_n}{a_k}\right)^j = \theta_j$, $a_n^{(d+1)} \log\log n = o(1)$, $\lim_{n\to\infty} n a_n^{(d+1)} a_n^{+4} = 0$ and $\lim_{n\to\infty} n a_n^{(d+1)} = \infty.$

Theorem 1.6.4. *(Khardani and Semmar (2014)[53]) Under Assumptions (A), (C1), (C2) and let* $a_n^- = \inf_{k=1,\dots,n} a_k$ *and* $a_n^+ = \sup_{k=1,\dots,n} a_k$*. We have*

$$\sup_{x\in\mathcal{C}} \sup_{t\in\Omega} \left| \widehat{\phi}_n(t/x) - \phi(t/x) \right| = O\left\{ \max\left(\left(\sqrt{\frac{\log n}{n a_n^{-(d+1)}}} \right), a_n^{+2} \right) \right\}, \; a.s. \quad as \quad n \to \infty.$$

Theorem 1.6.5. *(Khardani and Semmar (2014)[53]) Under Assumptions (A), (C2) and (N), we have, for any* $(x,t) \in \mathcal{A}$,

$$\sqrt{n a_n^{(d+1)}} \left(\widehat{\phi}_n(t/x) - \phi(t/x) \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(x,t))$$

where $\xrightarrow{\mathcal{D}}$ *denotes the convergence in distribution and* $\mathcal{A} = \{(x,t),\, \sigma^2(x,t) \neq 0\}$ *with* $\sigma^2(x,t) = \theta_{d+1} \dfrac{\phi(t/x)}{l(x)\overline{G}(t)} \int_{\mathbb{R}^d} \int_{\mathbb{R}} L_1^2(z) L_2^2(y) dz dy.$

These models play a crucial role in nonparametric prediction setup. This is what makes the method of studying estimators expanded and many scholars tend particularly to use the recursive method because of its many benefits and practical usefulness.

1.7 Nonparametric recursive method

With the development of present day registering and information securing strategies, voluminous information are gathered in different applied areas including stargazing, computer networks, remote detecting, climate observing. . . . The study of this kind of phenomena requires some specific non-parametric procedures to model them. In fact, popular nonparametric techniques such as the kernel density of Rosenblatt and the kernel regression of Nadaraya and Watson have been proposed. However, statisticians suffer from serious computational drawbacks that effectively limit the relevance of these strategies in applications that contain a lot of information and need to be updated frequently.

Basically, in statistical frameworks, there are several types of non-parametric methods for estimating functional relationship between co-variates and a real response. Of course, these types of techniques have their specific advantages and shortcomings. Referring to the literature, we note among the most prominent of these methods that have been proposed; the *recursive* (termed also the *on-line kernel or real-time updating method*) is of undoubted significance in time series investigation, the most popular and simple one, which allows to achieve under certain conditions this aim. Therefore, according to Amiri (2010)[1] and Mezhoud et al. (2014)[68], the recursive approach is more appropriate and simpler type of estimation method in terms of simulation time and storage space for continuously updated data, i.e. when the sample size is large and un-prefixed. Here, the statistician should not recalculate the estimate (re-read the process again) whenever additional observations are obtained, he should just update the initial estimation via the iterative aspect and compute instantly without resorting to past data. Thus, this approach, from both practical and theoretical level aims at reducing the time of calculations, in addition to the good results compared with that obtained by the classical ones due to its recursive relationship which in turn allows to have a good approximation of the reality. Whether or not it is of such, more broad, importance, the undeniable value of recursive estimation lies not only in its undeniable elegance and adaptability, but also in its demonstrable practical utility.

Recursive estimation has become common place, taking its position as an essential component in most degree courses concerned with control and systems theory, signal and image processing, statistical estimation and econometrics; and it is becoming increasingly important in other applied science courses, such as the earth and atmospheric sciences (e.g. hydrology, oceanography, atmospheric science), as well as some courses in the social sciences (e.g. psychology, sociology). The algorithm for recursive estimation is being used increasingly in many applied fields, however, with a continuous review of the literature, we have seen that the other specialists have been slower than statisticians to exploit them. An extensive bibliography contains references to the work of statisticians on recursive estimation and the sources on which they have relied.

There are as of now colossal measure of distributed papers on asymptotic results of online kernel density and regression estimators for both i.i.d. and mixing observations. For an overall view, we carefully present the recursive kernel density and regression estimators and their asymptotic properties. In this setting, there are plenty of works using this approach. The first recursive kernel version of the *Parzen-Rosenblatt* estimator is developed by Wolverton and Wagner (1969)[94], defined, for all $n \geq 1$ and for all $x \in \mathbb{R}$, by the form

$$\widehat{f}_n(x) = n^{-1} \sum_{k=1}^{n} a_k^{-1} L_1 \left(a_k^{-1}(x - X_k) \right) \tag{1.14}$$

where a_k is the bandwidth sequence and $L_1(\cdot)$ is a kernel function. Such that several scholars have studied this estimator among whom Yamato (1971)[96], Masry (1987)[65], Gÿorfi and Masry (1990)[43]. Subsequently, the almost sure convergence and the asymptotic normality are established by Deheuvels (1974)[25] and given in the following result

Theorem 1.7.1. *(Deheuvels (1974)[25]) Suppose that f is derivable with bounded derivative. If the bandwidth a_n such that $a_n = n^{-\alpha}$ with $0 < \alpha < 1$. Then, for every*

$x \in \mathbb{R}$

$$\lim_{n\to\infty} \widehat{f}_n(x) = f(x) \quad a.s.$$

In addition, for $1/5 < \alpha < 1$, *then, the asymptotic normality is*

$$\sqrt{na_n}\left(\widehat{f}_n(x) - f(x)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\mu^2 f(x)}{(1+\alpha)}\right).$$

Another recursive kernel estimator exists in the literature to estimate the density function given for $n \geq 1$ and $x \in \mathbb{R}$ by Deheuvels (1973)[24] as follows

$$\widetilde{f}_n(x) = B_n^{-1} \sum_{k=1}^{n} L_1\left(a_k^{-1}(x - X_k)\right) \tag{1.15}$$

where $B_n = \sum_{k=1}^{n} a_k$. The latter author have established the asymptotic properties for his estimator as given

Theorem 1.7.2. *(Deheuvels (1973)[24]) Suppose that* f *is derivable with bounded derivative. If the bandwidth* a_n *such that* $a_n = n^{-\alpha}$ *with* $0 < \alpha < 1$. *Then, for every* $x \in \mathbb{R}$

$$\lim_{n\to\infty} \widetilde{f}_n(x) = f(x) \quad a.s.$$

In addition, for $1/5 < \alpha < 1$, *then, the asymptotic normality is*

$$\sqrt{na_n}\left(\widetilde{f}_n(x) - f(x)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \mu^2 f(x)\right).$$

For the second problem concerning the regression function $R(x) = \mathbb{E}[Y_k/X_k = x]$, for a couple (X_n, Y_n) arrives sequentially, a preferable estimate is then defined by extending the classical Nadaraya-Watson estimator (so-called the off-line or batch approach) to the recursive version (also called *the recursive Nadaraya-Watson estimator*) and given for $x \in \mathbb{R}$ by

$$\widehat{R}_n(x) = \frac{\sum_{k=1}^{n} Y_k L_1\left(a_k^{-1}(x - X_k)\right) / \sum_{k=1}^{n} a_k}{\sum_{k=1}^{n} L_1\left(a_k^{-1}(x - X_k)\right) / \sum_{k=1}^{n} a_k} = \frac{J_{n(x)}}{\widehat{g}_n(x)} \tag{1.16}$$

where $\widehat{g}_n(x)$ is the online estimator for the marginal density of X_k with $B_n = \sum_{k=1}^{n} a_k$.

Thus, under certain regularity assumptions on the kernel L_1 such that $\int_{-\infty}^{+\infty} L_1^2(x)dx = \mu^2 < \infty$ and $\int_{-\infty}^{+\infty} x^2 L_1(x)dx = \nu^2 < \infty$, Duflo (1997)[29] in his contribution has showed some asymptotic properties under mixing assumptions.

Theorem 1.7.3. *(Duflo (1997)[29]) We assume that f is bounded, twice derivable with bounded derivatives. Moreover, we suppose that X admits a density g which is bounded, twice derivable with bounded derivatives and let $0 < \alpha < 1$ with $a_n = n^{-\alpha}$. If the noise is integrable square and of variance $\sigma^2 > 0$, then for all $x \in \mathbb{R}$,*

$$\lim_{n\to\infty} \widehat{R}_n(x) = R(x) \quad a.s.$$

Moreover, if $1/3 < \alpha < 1$ and if ϵ_n admits a moment of order greater than 2, then, for any $x \in \mathbb{R}$ such that $g(x) \neq 0$,

$$\sqrt{na_n}\left(\widehat{R}_n(x) - R(x)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\mu^2\sigma^2}{(1+\alpha)g(x)}\right).$$

The growing success of this method has encouraged many scholars to explore its contribution to the estimation of some conditional models in several cases. Currently, work in the field of *functional* variables has turned into an interesting topic lately, thanks to its applications; so that we can mention the seminal work of Amiri et al. (2014)[2] who have addressed this problem first and have studied the asymptotic properties of the recursive nonparametric kernel estimator of the regression function. While a recursive estimator of the conditional geometric median is studied by Hervé Cardot et al. (2012)[15] in Hilbert spaces and they have proved the almost sure convergence together with L_2 rates of convergence. More recently, for independent and identically distributed observations, Bouadjemi (2014)[13] has introduced a new estimator of the conditional cumulative distribution function based on a recursive approach and he has elaborated under certain terms and general conditions a result on the asymptotic normality of built estimate.

However, what we are able to clearly notice by reading and looking at some of the few previous studies in addition to our simple personal study is that this method is not at all desirable in light of the presence of incomplete observations (censored in particular) in the studied sample. This is conclusive evidence that this method, despite its proven efficacy in the analytical framework compared to the classical method in terms of simulation problems and not consuming longer time, but it is a method that has its inconvenient as well, and this is evident in most studies so that we do not say all of them to maintain some honesty.

Bibliography

[1] A. Amiri (2010). Sur une famille paramétrique d'estimateurs séquentiels de la densité pour un processus fortement mélangeant, C. R. Acad. Sci. Paris, Ser. I, 347, 309-314.

[2] A. Amiri, Ch. Crambes, and B. Thiam (2014). Recursive estimation of nonparametric regression with functional covariate, Computational Statistics and Data Analysis, 69, 154-172.

[3] Y. P. Ango Nzé, P. Bühlmann and P. Doukhan. Weak Dependence Beyond Mixing and Asymptotics for Nonparametric Regression. The Annals of Statistics 2002, 30(2), 397-430.

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

[5] M. Attouch, A. Laksaci, and E. Ould Saïd (2010). Asymptotic normality of a robust estimator of the regression function for functional time series data. J. of the Korean Statist. Soc., 39, 489-500.

[6] N. Azzeddine, A. Laksaci, E. Ould-Saïd. (2008). On the robust nonparametric regression estimation for functional regressor. Statist. Probab. Lett., in press.

[7] G. Bassett, R. Koenker. (1978). Regression quantiles. Econometrica, pages 33-50. 54.

[8] F. Benziadi, A. Gheriballah, A. Laksaci, (2016). Asymptotic normality of kernel estimator of $\psi-$regression function for functional ergodic data. New Trends in Mathematical Sciences, 4, 268-282.

[9] F. Benziadi, A. Laksaci, F. Tebboune. (2016). Recursive kernel estimate of the conditional quantile for functional ergodic data. Comm in Statist-Theory and Methods 3097-3113.

[10] Beran. Nonparametric regression with randomly censored survival data. Technical report, Department of Statistics, University of California, Berkeley., 1981.

[11] A. Berlinet, A. Gannoun et E. Matzner-Lober. (2001). Asymptotic normality of convergent estimates of conditional quantiles. Statistics, 18:1400-1415. 50, 52, 95, 97

[12] D. Bosq (2000). Linear processs in function spaces. Théorie and Application. Lectures Notes in Statistics. Springer Verlag, New York. 149.

[13] A. Bouadjemi (2014). Asymptotic Normality of the Recursive Kernel Estimate of Conditional Cumulative Distribution Function. Journal of Probability and Statistical Science 12(2), 117-126.

[14] H. Cardot, C. Crambes and P. Sarda. (2005). Quantiles regression when the covariates are functions. Journal of Nonparametric Statistics, 17, 841-856.

[15] H. Cardot, P. Cénac and P-A. Zitt. Recursive estimation of the conditional geometric median in Hilbert spaces. Electronic Journal of Statistics, V. 6 (2012), 2535-2562.

[16] M. Chaouch, S. Khardani. (2014). Kernel-smoothed conditional quantiles of randomly censored functional stationary ergodic data. Statist, Theory.

[17] P. Chauduri (1991). Nonparametric estimates of regression quantiles and their local Bahadur representation. Annals of Statistics, 19(2):760-777.

[18] A. C. Cohen. Truncated and Censored Samples: Theory and Applications. Dekker, 1991.

[19] G. Collomb. (1980). Estimation non paramétrique de probabilités conditionnelles. Comptes Rendus de l'Académie des Sciences, 291: 427-430.

[20] G. Collomb, W. Härdle. (1986). Strong uniform convergence rates in robust nonparametric time series analysis and prediction: Kernel regression estimation from dependent observations. Stochastic Processes and Their Applications, 23, 77-89.

[21] S. Dabo-Niang, et A. Laksaci. (2007). Estimation non paramtrique de mode conditionnel pour variable explicative fonctionnelle. C. R. Math. Acad. Sci. Paris, 344(1), 49-52.

[22] D. M. Dabrowska (1987). Nonparametric regression with censored lifetime data. Scand. J. Statist. 14, 181-197.

[23] D. M. Dabrowska. Uniform consistency of the kernel conditional kaplan-meier estimate. Ann. of Statist., 17: 1157-1167, 1989.

[24] P. Deheuvels. Sur une famille d'estimateurs de la densité d'une variable aléatoire. C. R. Acad. Sci. Paris Sér. A-B, 276: A1013-A1015, 1973.

[25] P. Deheuvels. Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre et uniforme presque sûre des estimateurs de la densité. C. R. Acad. Sci. Paris Sér. A, 278: 1217-1220, 1974.

[26] P. Doukhan. Mixing: Properties and examples. Lecture Notes in Statistics, 85, SpringerVerlag, New York, 1994.

[27] P. Doukhan, S. Louhichi. A new weak dependence condition and applications to moment inequalities. Stochastic Processes and their Applications 84 (1999) 313-342.

[28] P. Doukhan, S. Louhichi. Functional Estimation of a Density Under a New Weak Dependence Condition. Board of the Foundation of the Scandinavian Journal of Statistics 28: 325-341, 2001.

[29] M. Duflo. Random iterative models, volume 34 de Applications of Mathematics. Springer-Verlag, Berlin, 1997.

[30] M. Ezzahrioui, and E. Ould-Saïd. Asymptotic normality of the kernel estimator of the conditional quantile in the normed space. Technical report, (2005b), 258, LMPA, Univ. Littoral Côte d'Opale. (A paraître dans Far East J. theoretical Statist.)

[31] M. Ezzahrioui, and E. Ould-Saïd. Asymptotic results of the kernel estimator of the conditional quantile in the normed space under $\alpha-$mixing hypothesis. Technical report, (2006b), 309, LMPA, Univ. Littoral Côte d'Opale.

[32] M. Ezzahrioui, E. and Ould-Saïd. (2008(a)). Asymptotic normality of the kernel estimators of the conditional mode for functional data. J. Nonparametric Statist, 20, 3-18.

[33] F. Ferraty, A. Laksaci, et P. Vieu. (2005). Functional time series prediction via conditional mode estimation. C. R. Math. Acad. Sci. Paris, 340(5), 389-392.

[34] F. Ferraty, A. Laksaci, P. Vieu. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. Statist. Infer. Stoch. Process 9:47-76.

[35] F. Ferraty, A. Rabhi, P. Vieu. (2005). Conditional quantiles for dependent functional data with application to the climatic El Nio phenomenon. Indian Journal of Statistics, 67(2): 378-398. 50, 126.

[36] F. Ferraty, et P. Vieu. (2000). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. Comptes Rendus de l'Académie des Sciences, 330: 403-406. 48.

[37] F. Ferraty, P. Vieu. (2003). Curves discrimination :a nonparametric functional approach. Comput. Statist. and Data Anal., 44, 161-173.

[38] F. Ferraty, P. Vieu. (2004). Nonparametric models for functional data, with application in regression times series prediction and curves discrimination. J. Nonparametric Statist. 16, 111-125.

[39] F. Ferraty, P. Vieu. (2006). Nonparametric Functional Data Analysis: Theory and Practice. Springer Series in Statistics, Springer. 48, 50, 54, 56, 68.

[40] A. Gannoun. (1989). Estimation de la médiane conditionnelle. Thèse de doctorat, Université de Paris VI. 48, 51.

[41] I. Gijbels, and J. L. Wang. (1993). Strong representations of the survival function estimator for truncated and censored data with applications. Journal of Multivariate Analysis 47, 210-229.

[42] M. G. Gu, and T. L. Lai. (1990). Functional laws of the iterated logarithm for the product-limit estimator of a distribution function under random censorship or truncation. Annals of Probability 18, 160-189.

[43] L. Gyorfi and E. Masry. The L_1 and L_2 strong consistency of recursive kernel density estimation from dependent samples, IEEE Trans. Inform. Theory, vol. 36, 531-539, 1990.

[44] X. He, and P. Shi. (1994). Convergence rate of B-spline estimators of non parametric conditional quantiles functions. Journal of Nonparametric Statistics, 3(4): 299-308. 51.

[45] D. R. Helsel. Statistics for Censored Environmental Data Using Minitab and $\mathcal{R}$. Wiley, 2nd edition, 2011

[46] A. Henningsen. censReg: Censored Regression (Tobit) Models, 2012. URL http://CRAN. $\mathcal{R}$-project.org/package=censReg. $\mathcal{R}$ package version 0.5-20. Date: 2013/08/20. Access Date May 20, 2014.

[47] A. Henningsen and O. Toomet. maxlik: A package for maximum likelihood estimation in $\mathcal{R}$. Computational Statistics, 26(3):443-458, 2011.

[48] P. Hougaard. (2000). Analysis of Multivariate Survival Data. New York: Springer-Verlag.

[49] Y. Huang, X. Chen, and W. B. Wu. Recursive Nonparametric Estimation For Time Series. IEEE TRANSACTIONS ON INFORMATION THEORY, 2013

[50] L. A. Ibragimov. Some limit theorems for stationary processes. Theory. Probab. Application, 7: 349-382, 1962.

[51] E. L. Kaplan, and P. Meier. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53, 457-481.

[52] S. Khardani, M. Lemdani, and E. Ould-Said. (2010). Some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship. J. Korean Statist. Soc., 39(4), 455-469.

[53] S. Khardani, S. Semmar. (2014) Nonparametric conditional density estimation for censored data based on a recursive kernel. Electron J Stat 8: 2541-2556.

[54] S. Khardani, B. Thiam. (2016). Strong consistency result of a non parametric conditional mode estimator under random censorship for functional regressors. Com In Stat-Theory and Methods, VOL. 45, 1863-1875.

[55] J. P. Klein, and M. L. Moeschberger. (1997). Survival Analysis: Techniques for Censored and Truncated Data. New York: Springer-Verlag.

[56] T. L. Lai and Z. Ying. (1991). Estimating a distribution function with truncated and censored data. Annals of Statistics 19, 417-442.

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

[58] N. Laïb, and D. Louani. (2010). Nonparametric kernel regression estimation for functional stationary ergodic data: asymptotic properties. J. Multivariate Anal., 101(10), 2266-2281.

[59] N. Laïb, and D. Louani. (2011). Rates of strong consistencies of the regression function estimator for functional stationary ergodic data J. Statist. Plan. and Inference , 141, 359-372.

[60] J.P. Lecoutre, E. Ould-Saïd. Convergence of the conditional Kaplan-Meier estimate under strong mixing. Journal of Statistical Planning and Inference 44 (1995), 359-369.

[61] J. van. Leth. Ergodic Theory, Interpretations of Probability and the Foundations of Statistical Mechanics, Studies in the History of the Philosophy of Modern Physics, 32, 581-594, (2001).

[62] N. Ling and Q. Xu. Asymptotic normality of conditional density estimation in the single index model for functional time series data. Statistics and Probability Letters 82 (2012), 2235-2243.

[63] S. R. Lipsitz and J. G. Ibrahim. Estimation with correlated censored survival data with missing covariates. Biostatistics, 1: 315-327, 2000.

[64] D. Lynden-Bell (1971). A method of allowing for known observational selection in small samples applied to 3CR quasars. Monthly Notices of the Royal Astronomical Society 155, 95-188.

[65] E. Masry. Almost sure convergence of recursive density estimators for stationary mixing processes. Statist. Probab. Lett., vol. 5, 249-254, 1987.

[66] E. Masry. (2005). Nonparametric regression estimation for dependent functional data: asymptotic normality. Stoch. Proc. and their Appl., 115, 155-177.

[67] A. I. McLeod and N. M. Mohammad. Censored time series analysis, 2012. URL http://CRAN.R-project.org/package=censReg. R package version 0.5-20. Date: 2013/08/20. Access Date May 20, 2014.

[68] K. Mezhoud, Z. Mohdeba, S. Louhichi. Recursive kernel estimation of the density under η-weak dependence. Journal of the Korean Statistical Society 43 (2014), 403-414.

[69] E. A. Nadaraja. On a regression estimate. Teor. Verojatnost. i Primenen., 9: 157-159, 1964.

[70] T. H. Nguyen. Asymptotic normality of recursive density estimators in Markov processes. Publ. Inst. Statist. Univ. Paris, vol. 26, pp. 73-93, 1981.

[71] T. M. Nguyen. Estimation récursive pour les modèles semi-paramétriques. Statistics [math.ST]. Université Sciences et Technologies-Bordeaux I, 2010. French.

[72] F.H. Nieto and J. Martinex. A recursive approach for estimating missing observations in univariate time series. Communications in Statistics (Theory and Methods), 25(9):2101-2116, 1996.

[73] D. Oakes. (1986). Semiparametric inference in bivariate survival data. Biometrika 73, 353-361.

[74] C.R. de Olivera and T. Werlang, Ergodic hypothesis in classical statistical mechanics, Rev. Brasileira de Ensino de Fisica, 29, 189-201, (2007).

[75] J. W. Park, Marc G. Genton, and Sujit K. Ghosh. Censored time series analysis with autoregressive moving average models. The Canadian Journal of Statistics, 35(1): pp. 151-168, 2007. URL http://www.jstor.org/stable/ 2044-5244.

[76] A. Patrascioiu. The ergodic hypothesis: a complicated problem in Mathematics and Physics, Los Alamos Science, (Special issue), 263-279, (1987).

[77] A. V. Jr. Peterson. (1977). Expressing the Kaplan-Meier estimator as a function of empirical subsurvival functions. J. Amer. Statist. Assoc. 72, 854-858

[78] J. Penzer and B. Shea. The exact likelihood of an autoregressive-moving average model with incomplete data. Biometrika, 84(4): pp. 919-928, 1997.

[79] G. Peskir. From Uniform Laws of Large Numbers to Uniform Ergodic Theorems. University of Aarhus, Department of Mathematics, 2000.

[80] K. Petersen, Ergodic Theory, Cambridge University Press 1983.

[81] H. Poincaré, Les méthodes nouvelles de la mécanique céleste, tome III, chapitre XXVI "Stabilité à la Poisson", Gauthier-Villars 1899, retirage librairie scientifique et technique Albert Blanchard 1987.

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

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

[84] M. Rosenblatt. A central limit theorem and a strong mixing condition. Proc. Nat. Acad. Sci. U. S. A., 42: 43-47, 1956b.

[85] G. G. Roussas. (1969). Nonparametric estimation of the transition distribution function of Markov process. Annals of Mathematical Statistics, 40(4):1386-1400. 46, 48.

[86] G.R. Roussas, L.T. Tran. (1992). Asymptotic normality of the recursive kernel regression estimate under dependence conditions. Ann. Statist., 20, 98-120.

[87] H. Schneider. Truncated and Censored Samples from Normal Populations. Dekker, 2nd edition, 1986.

[88] S. Sekkal, A. Gheriballah and A. Laksaci. Nonparametric $M-$regression for functional ergodic data. Statistics and Probability letters 83(2013), 902-908.

[89] W. Stute, and J.L. Wang. (1993). The strong law under random censorship. Ann. Statist. 21, 1591-1607.

[90] W. Y. Tsai, N. P. Jewell, and M. C. Wang. (1987). A note on the product-limit estimator under right censoring and left truncation. Biometrika 74, 883-886.

[91] V. A. Volkonskii and Y. A. Rozanov. Some limit theorem for random functions. I. Theory Probab. Appl., 4: 178-197, 1959.

[92] M. C. Wang, N. P. Jewell, and W. Y. Tsai. (1986). Asymptotic properties of the product limit estimate under random truncation. Annals of Statistics 14, 1597-1605.

[93] L. Wang and H. Y. Liang. (2004). Strong uniform convergence of the recursive regression estimator under $\phi-$mixing conditions. Metrika, 59(3), 245-261.

[94] C. T. Wolverton et T. J. Wagner. Asymptotically optimal discriminant functions for pattern classification. IEEE Trans. Information Theory, IT-15: 258-265, 1969.

[95] M. Woodroofe. (1985). Estimating a distribution function with truncated data. Annals of Statistics 13, 163-177.

[96] H. Yamato. Sequential estimation of a continous probability density function and mode. Bull. Math. Statist., vol. 14, 1-12, 1971.

Chapter 2

Asymptotic Results of a Recursive Double Kernel Estimator of the Conditional Quantile for Functional Ergodic Data

I. Bouazza, F. Benziadi, F. Madani and T. Guendouzi

Published in *Bulletin of the Institute of Mathematics Academia Sinica*

Abstract *The aim of this chapter is to investigate the estimation of conditional quantile of a scalar response variable Y given a random variable (rv) $X = x$ taking values in a semi-metric space. Hence, the asymptotic normality of the proposed estimator is obtained when the observations are sampled from a functional ergodic process. The result confirms the prospect proposed in Benziadi et al [3] and as applications, a comparison study based on a finite-sample behavior of the estimator is investigated by simulations as well.*

Key words and phrases *Asymptotic normality, Conditional quantile, Recursive estimate, Ergodic data, Functional data, Small ball probability.*

2.1 Introduction

In nonparametric statistics, the estimation of conditional quantiles is becoming increasingly an important problem which has been widely studied because of their importance in several applications such as agronomy, medicine, economic ... and so on. Historically, countless works have been documented on this problem: By a direct method, the estimator of conditional quantile $q_\alpha(x)$ is proposed and then studied firstly by Koenker and Bassett [15] and defined as follow

$$\widehat{q}_\alpha(x) = \arg\min_{q_\alpha} \sum_{i=1}^{n} \rho_\alpha(Y_i - q_\alpha), \quad for\ (x, y) \in \mathbb{R}^2$$

which made many scholars invest in this topic, among whom Cardot et al. [5], Koenker [14]. Subsequently, in the event that the observations are functional, Gannoun et al. [13] have proposed a nonparametric conditional median predictor based on the double kernel method. The asymptotic properties of no-parametric conditional quantile estimator are established by Ezzahrioui and Ould-Saïd [10]. Laksaci et al. [17] have studied the almost complete consistency and the asymptotic normality of a generalized L^1-approach for a kernel estimator of conditional quantile with functional regressor.

However, studies of conditional quantile estimation are a significant subject that has given rise to a large number of contributions, and their applications are very wide and cover various fields, often involve both prediction setting and in estimation of regression function. Several works on the regression quantile exist in the literature, the first idea for this subject, is proposed by Stone [20]. Cardot et al. [5] have proved the L^2-convergence rate of the conditional quantile as a linear regression model for functional data. Their results have been extended to the kernel case by Ferraty et al. [11] who have proposed a nonparametric estimator of this model and they have established the almost complete convergence for the i.i.d case. On the other hand, the convergence in L^p-norm is stated by Dabo-Niang and Laksaci [8]. The interested reader can refer also to some of the following additional references [18], [22] and [7] to expand further on this topic and take an overview.

Otherwise, considering the recursive conditional models estimate, literary, the first result on this topic is developed by Wolverton and Wagner [21], they have established the asymptotically optimal discriminant functions for pattern classification. Masry and Györfi [19] have also treated for weakly dependence stationary process the recursive estimator of probability density. Recently, Amiri [1] has investigated the asymptotic properties of the recursive regression estimator with application in the non-parametric prediction. In the same aim, the almost sure convergence rates of the conditional geometric median estimator have been proved by Cardot et al. [6] in Hilbert space. Amiri et al. [2] have obtained the asymptotic properties of the recursive estimator of the regression function with functional covariate.

In this present work, in regard to the dependence setting, our focus is to use ergodic variables to allow the maximum possible generality and to estimate the conditional quantile function in this case. Note here that, the nonparametric kernel regression estimation for functional stationary ergodic data is considered by Laib and Louani [16], they have studied the consistency in probability, with a rate, as well as the asymptotic normality of this estimator. Also, for i.i.d functional data, Bouadjemi [4] has established the asymptotic normality for the conditional cumulative distribution function. More recently, Benziadi et al. [3] have studied the almost complete (a.co) convergence with rates of the functional recursive kernel estimate of the conditional quantile.

The outline is described as follows: At first, we define the double-kernel recursive estimator when the covariate X is functional in Section 2.2. We establish then the asymptotic normality of this model as well as the confidence interval in Sections 2.4 and 2.5 under the assumptions given in 2.3. A computational study is carried out to evaluate and understand how effective this resulting model is, in Section 2.6. Finally, in Section 2.7 devoted to appendix, we present the detailed proofs of the auxiliary results. Indeed, in the setting of ergodic processes, to prove our results, our methodology is based mainly on the martingale approximation which allows to launch a systematic study for dependent data.

2.2 Definition of the model

To fix notation, let $(X_k, Y_k)_{k=1,\dots,n}$ be a sequence of strictly stationary dependent random variables valued in $\mathcal{H} \times \mathbb{R}$ and observable from the same subject, where $(\mathcal{H}, d_{\mathcal{H}})$ is a semi-metric space. For $\alpha \in [0,1]$, the conditional quantile of order α defined by the inverse of the conditional distribution function $F^x(y)$ of Y given $X = x$ is denoted $q_\alpha(x)$ and such that

$$q_\alpha(x) := F^{-1}(\alpha/x) = \inf\left\{y \in \mathbb{R} : F^x(y) \geq \alpha\right\}.$$

To estimate it, we use the estimator of the nonparametric function given usually by:

$$F^x(y) = \mathbb{P}\left[Y \leq y/X = x\right]. \tag{2.1}$$

Thus, based on the finding proof of Laksaci et al. [17], we note that this estimator can be given by:

$$\widehat{q}_\alpha(x) = \inf\left\{y \in \mathbb{R} : \widehat{F}^x(y) \geq \alpha\right\}.$$

In the sequel, let $\widehat{F}^x(y)$ be the recursive estimate of $F^x(y)$ defined as:

$$\widehat{F}^x(y) = \frac{\displaystyle\sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) L_2\left(b_k^{-1}(y - Y_k)\right)}{\displaystyle\sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)}, \quad \forall y \in \mathbb{R} \tag{2.2}$$

where L_1 is a kernel distribution defined for $x \in \mathcal{H}$, L_2 is a strictly increasing function defined for $y \in \mathbb{R}$ and (a_k) (resp (b_k)) is a sequence of positive real numbers such that $\lim\limits_{n\longrightarrow\infty} a_n = 0$ ($\lim\limits_{n\longrightarrow\infty} b_n = 0$). The main advantage of this estimation method is to update the estimate for each additional observation without resorting to past data.

2.3 General framework and assumptions

The general framework of our contribution is the nonparametric modeling in functional ergodic data. To this aim, in such study, we formulate these data by the following notations, for all $k = 1, \cdots, n$, we introduce $\wp_k$ the σ-field generated by $((X_1, Y_1), \cdots, (X_k, Y_k))$ and $\mathcal{B}_k$ the σ-field generated by $((X_1, Y_1), \cdots, (X_k, Y_k), X_{k+1})$. In addition, x will stand from now on for a fixed point in $\mathcal{H}$, C and C' denote some generic constant in $\mathbb{R}^{*+}$, $\mathcal{N}_x$ denotes the fixed neighborhood of x and $B(x,h) = \{x' \in \mathcal{H}/d_{\mathcal{H}}(x', x) < h\}$.

Therefore, our asymptotic results are stated under the following assumptions of our model that we gathered hereafter for easy reference.

(A.1) The strictly stationary ergodic process $(X_k, Y_k)_{k\in\mathbb{N}^*}$ satisfies:

$$\begin{cases} (i) \text{ The function } \phi(x,h) := \mathbb{P}(X \in B(x,h)) \text{ is such that } \phi(x,h) > 0, \quad \forall\, h > 0. \\ (ii) \text{ For all } k = 1, \dots, n, \text{ there exists a deterministic function, } \phi_k(x,\cdot) \text{ such that} \\ \text{ almost surely } 0 < \mathbb{P}\left(X_k \in B(x,h)/\wp_{k-1}\right) \le \phi_k(x,h), \forall\, h > 0, \\ \text{and } \phi_k(x,h) \to 0 \text{ as } h \to 0. \\ (iii) \text{ For all sequence } (h_k)_{k=1,\dots n} > 0, \ \dfrac{\sum_{k=1}^{n} \mathbb{P}\left(X_k \in B(x,h_k)/\wp_{k-1}\right)}{\sum_{k=1}^{n} \phi(x,h_k)} \to 1 \qquad a.s. \end{cases}$$

For a fixed neighborhood $\mathcal{N}_x$ of x, we assume that the regular version $F^{x'}$ of the conditional distribution function of Y given $X = x'$ exists for all $x' \in \mathcal{N}_x$ and we suppose that F^x has a continuous density ξ^x with respect to Lebesgue's measure over $\mathbb{R}$ and is such that:

(A.2) There exists $\delta > 0$, such that $\forall (y_1, y_2) \in [q_\alpha(x) - \delta, q_\alpha(x) + \delta]^2$, $\forall (x_1, x_2) \in \mathcal{N}_x^2$ and C', β_1 and $\beta_2 \in \mathbb{R}^{*+}$,

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \le C' \left(d_{\mathcal{H}}(x_1, x_2)^{\beta_1} + |y_1 - y_2|^{\beta_2}\right)$$

and for $y \in \mathbb{R}$ and $C > 0$, the conditional density function $\xi^x(y)$ of Y given $X = x$ verifies

$$\inf_{y \in [q_\alpha(x)-\delta,\, q_\alpha(x)+\delta]} \xi^x(y) > C.$$

(A.3) The positive bandwidths (a_k, b_k) satisfy: $\forall t \in [0,1]$,

$$\lim_{n\to\infty} \frac{\phi(x, ta_n)}{\phi(x, a_n)} = \beta_x(t), \ \lim_{n\to\infty} \frac{1}{n\psi_n(x)} \sum_{k=1}^{n} \Big((ca_k^{\beta_1} + c' b_k^{\beta_2})\phi(x, a_k) \Big)^2 = 0$$

and

$$n\psi_n(x) \longrightarrow \infty \text{ with } \psi_n(x) = n^{-1} \sum_{k=1}^{n} \phi(x, a_k).$$

(A.4) L_1 is a function and having a compact support on $(0,1)$, such that:

$$C\mathbb{I}_{(0,1)} < L_1(x) < C'\mathbb{I}_{(0,1)}.$$

(A.5) The function L_2 is of class $\mathcal{C}^1$ such that:

$$\begin{cases} \forall (y_1, y_2) \in \mathbb{R}^2, \ |L_2(y_1) - L_2(y_2)| \leq C|y_1 - y_2|; \\ \int |t|^{\beta_2} L_2^{(1)}(t)dt < \infty, \text{where } \beta_2 \text{ is given in (A.2)}. \end{cases}$$

2.3.1 Comments on the assumptions

For our model, these assumptions are very usual for estimating the conditional quantile; condition (A.1)(i) characterizes the property of concentration on small balls of the probability measure of the underlying explanatory variable. The ergodicity of functional data in assumption $(A.1)(ii)$ is the same as that classically imposed by Laib and Louani [16] for infinite-dimensional setting. The assumption $(A.2)$ ensures the regularity version $F^{x'}$ of the conditional distribution function of Y given $X = x'$ for a fixed neighborhood $\mathcal{N}_x$ of x, although there are several ways to define this nonparametric concept. For the semi-metric structure, this latter condition is more suitable. Condition (A.3) plays a crucial role in the asymptotic normality result: In other words, the function β_x is needed in the study of the variance term and thus $(A.4)$ is checked for kernel estimation. In order to explicit asymptotically the bias term, the condition $(A.5)$ is fulfilled.

2.4 Main results

In the following, we announce the asymptotic result of the double-recursive kernel estimator $\widehat{q}_\alpha(x)$ of $q_\alpha(x)$, where $\xrightarrow{D}$ denotes the convergence in distribution and $\mathcal{N}(\cdot,\cdot)$ denotes the gaussian distribution.

Proposition 2.4.1. *Suppose that the conditions $(A.1)-(A.5)$ hold true. In addition, if the following condition is verified*

$$\lim_{n\to\infty}\frac{\log n}{n\psi_n(x)}=0$$

then,

$$\left(\frac{n\psi_n(x)}{\sigma^2(x)}\right)^{1/2}\left(\widehat{F}^x(y)-F^x(y)\right)\xrightarrow{D}\mathcal{N}(0,1)\ as\ n\longrightarrow\infty.$$

Theorem 2.4.1. *Under the hypotheses of Proposition 2.4.1, we have for any $x\in\mathcal{A}$ and $\alpha\in[0,1]$*

$$\left(\frac{n\psi_n(x)}{\sigma^2(x)}\right)^{1/2}\left(\widehat{q}_\alpha(x)-q_\alpha(x)\right)\xrightarrow{D}\mathcal{N}(0,1)\ as\ n\longrightarrow\infty$$

where

$$\sigma^2(x)=\left(\frac{\alpha(1-\alpha)\gamma_1}{(\xi^x(q_\alpha))^2\gamma_2^2}\right);$$

with

$$\gamma_1=L_1^2(1)-\int_0^1(L_1^2(s))^{(1)}\beta_x(s)ds>0,\quad \gamma_2=L_1(1)-\int_0^1L_1^{(1)}(s)\beta_x(s)ds\neq 0$$

and $\mathcal{A}=\{x/\sigma^2(x)\neq 0\}$.

Proof of Proposition 2.4.1 We will rely on the following notations for the remainder of this paper: for any $(x,y)\in\ \mathcal{H}\times\mathbb{R}$ and $k=1,\cdots,n$

$$L_{1,k}=L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)\ \ and\ \ L_{2,k}=L_2\left(b_k^{-1}(y-Y_k)\right)$$

and then address the general decomposition used usually in this nonparametric case

$$\widehat{F}^x(y) - F^x(y) = \widehat{B}_n(x,y) + \frac{\widehat{R}_n(x,y)}{\widehat{F}_D(x)} + \frac{\widehat{Q}_n(x,y)}{\widehat{F}_D(x)}$$

where

$$\begin{aligned}\widehat{Q}_n(x,y) &:= \left(\widehat{F}_N^x(y) - \bar{F}_N^x(y)\right) - F^x(y)\left(\widehat{F}_D(x) - \bar{F}_D(x)\right),\\ \widehat{B}_n(x,y) &:= \frac{\bar{F}_N^x(y)}{\bar{F}_D(x)} \quad \text{and} \quad \widehat{R}_n(x,y) := -\widehat{B}_n(x,y)\left(\widehat{F}_D(x) - \bar{F}_D(x)\right)\end{aligned}$$

with

$$\begin{aligned}\widehat{F}_N^x(y) &:= \frac{1}{n\psi_n(x)}\sum_{k=1}^n L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}(y-Y_k)\right),\\ \bar{F}_N^x(y) &:= \frac{1}{n\psi_n(x)}\sum_{k=1}^n \mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}(y-Y_k)\right)/\wp_{k-1}\right],\\ \widehat{F}_D(x) &:= \frac{1}{n\psi_n(x)}\sum_{k=1}^n L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right),\\ \bar{F}_D(x) &:= \frac{1}{n\psi_n(x)}\sum_{k=1}^n \mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)/\wp_{k-1}\right].\end{aligned}$$

Thus, Proposition 2.4.1 is a consequence of the following intermediate results which proofs are given in the appendix.

Lemma 2.4.1. *Under hypotheses of Proposition 2.4.1, we have for any $x \in \mathcal{H}$*

$$\left(\frac{n\psi_n(x)\gamma_2^2}{\gamma_1\alpha(1-\alpha)}\right)^{1/2}\widehat{Q}_n(x,\tau_\alpha(u,x))\xrightarrow{D}\mathcal{N}(0,1) \text{ as } n \longrightarrow \infty.$$

Lemma 2.4.2. *Under hypotheses $(A.1)$ and $(A.4)$, we have:*

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

Lemma 2.4.3. *Under hypotheses* $(A.1)$*,* $(A.2)$ *and* $(A.3)$*, we have:*

$$\left(\frac{n\psi_n(x)\gamma_2^2}{\gamma_1\alpha(1-\alpha)}\right)^{1/2} \widehat{B}_n(x,\tau_\alpha(u,x)) = u + \circ_p(1) \text{ as } n \longrightarrow \infty.$$

Lemma 2.4.4. *Under hypotheses* $(A.1)$*,* $(A.2)$ *and* $(A.4)$*, we have*

$$\left(\frac{n\psi_n(x)\gamma_2^2}{\gamma_1\alpha(1-\alpha)}\right)^{1/2} \widehat{R}_n(x,\tau_\alpha(u,x)) = \circ(1) \text{ as } n \longrightarrow \infty.$$

Proof of Theorem 2.4.1 We define for all $u \in \mathbb{R}$, $\tau_\alpha(u,x) = q_\alpha(x) + u\left[n\psi_n(x)\right]^{-1/2}\sigma(x)$, such that

$$\begin{aligned}
\mathbb{P}\left\{\sqrt{n\psi_n(x)}\sigma^{-1}(x)(\widehat{q}_\alpha(x) - q_\alpha(x)) < u\right\} &= \mathbb{P}\left(\widehat{q}_\alpha(x) < \tau_\alpha(u,x)\right) \\
&= \mathbb{P}\left[\widehat{F}^x(\tau_\alpha(u,x)) > \alpha\right] \\
&= \mathbb{P}\left(\widehat{F}^x(\tau_\alpha(u,x)) > 0\right).
\end{aligned}$$

It follows that

$$\begin{aligned}
\mathbb{P}\left(\widehat{F}^x(\tau_\alpha(u,x)) > 0\right) &= \mathbb{P}\left(0 < \widehat{B}_n(x,y) + \frac{\widehat{R}_n(x,y)}{\widehat{F}_D(x)} + \frac{\widehat{Q}_n(x,y)}{\widehat{F}_D(x)}\right) \\
&= \mathbb{P}\left(-\widehat{F}_D(x)\widehat{B}_n(x,\tau_\alpha(u,x)) - \widehat{R}_n(x,\tau_\alpha(u,x)) - \widehat{Q}_n(x,\tau_\alpha(u,x)) < 0\right) \\
&= \mathbb{P}\left(-\widehat{F}_D(x)\widehat{B}_n(x,\tau_\alpha(u,x)) - \widehat{R}_n(x,\tau_\alpha(u,x)) < \widehat{Q}_n(x,\tau_\alpha(u,x))\right)
\end{aligned}$$

Thus, the proof of Theorem 2.4.1 is a direct consequence of Lemmas 2.4.1-2.4.4.

2.5 Application to predictive interval

The aim of this section is to construct the confidence band of asymptotic level $(1-\alpha)\%$ for the conditional distribution function estimator where $u_{\alpha/2}$ is the upper $\alpha/2$ quantile of standard normal $\mathcal{N}(0,1)$. So that, the following corollary gives us an asymptotic approximation

Corollary 2.5.1. *By following the above results, for any $x \in \mathcal{H}$ and every α, we get*

$$\left[\widehat{F}^x(y) - u_{\alpha/2}\sqrt{\frac{\sigma^2(x)}{n\psi_n(x)}}, \widehat{F}^x(y) + u_{\alpha/2}\sqrt{\frac{\sigma^2(x)}{n\psi_n(x)}}\right].$$

2.6 Computational studies

The highlighted result in this section is an important investigation of a small numerical study for evaluating the performance of the proposed estimator. More precisely, the main aim is to compare the efficiency of the double-kernel recursive estimation method, to the classical kernel one which has been extensively discussed in the previous many papers. To do that, we consider firstly the following non parametric model for all $k = 1, \dots, n$

$$Y_k = R(X_k) + \epsilon_k \tag{2.3}$$

where ϵ_k are random variables independent of X and follow a normal mixture distribution $(1-\lambda) * \mathcal{N}(0,1) + \lambda * \mathcal{N}(4,5)$ and we choose the contamination parameter λ to be respectively: 0.1, 0.2, 0.5, 0.7 and 0.9 with a sample size n= 100, 200 and 500.

In addition, in order to generate the functional variables $(X_k)_{k=1,\dots,n}$, we use the $\mathcal{R}$-routine *simul.far* of the *far* package in $\mathcal{R}$, This routine simulates a functional autoregressive process with a strong white noise. The simulation experiments used here considered a sinusoidal basis, with five functional axes, of the continuous functions from $[0,1]$ to $\mathbb{R}$. Also, we fix the diagonal matrix $(0.45, 0.9, 0.34, 0.45)$ to define the linear operator with a perturbation coefficient equal to 0.05. The X_k's curves are discretized in the same grid which is composed of 100 points and are plotted in Figure 2.1.

Moreover, the response variables Y_k are generated from the following operator:

$$R(x) = 5\int_0^1 \exp(x(t))dt.$$

Using this model permits the determination of the theoretical quantile $q_\alpha(x)$ such as the conditional distribution of Y given $X = x$ is explicitly given by the distribution of ϵ_k shifted by $R(x)$.

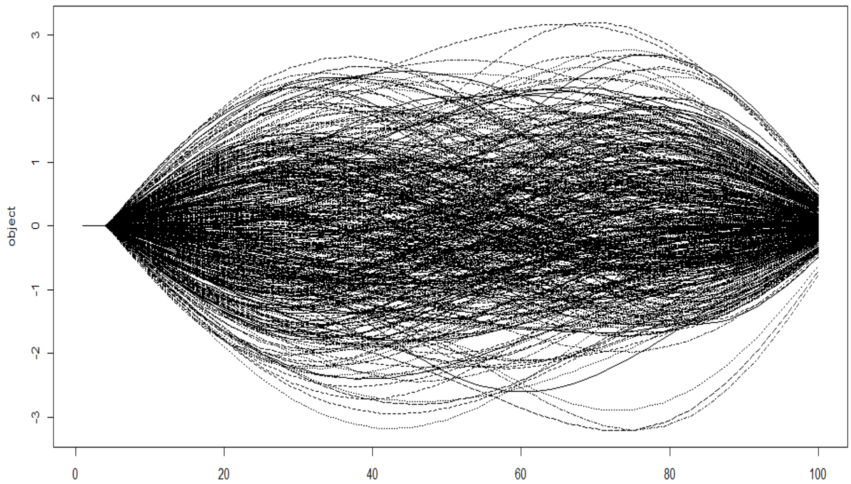

Figure 2.1: A sample of 100 curves

To give a fair comparison between the two methods we must treat each one under its optimal conditions and specify the different parameters of the both. Unfortunately, to the best of our knowledge, there is no automatic data-driven method available for selecting bandwidths when estimating a conditional quantile function with functional regressors. Thus, for our comparison study we consider a similar bandwidth selector to that used by Ferraty and Vieu (2006) [12]. Specifically, the bandwidths (a_k, b_k) in the recursive method are selected by the following leave-out-one-curve cross-validation procedure on the $k-$nearest neighbors

$$\arg \min_{(a_k,b_k)\in A_n\times B_n} \sum_{j=1}^{n} \left(Y_j - q_{0.5}^{[-j]}(X_j, a_k, b_k)\right)^2,$$

where $q_{0.5}^{[-j]}(X_j, a_k, b_k)$ denotes the double-kernel recursive estimator of the conditional median in the curve X_j and is computed by the bandwidths (a_k, b_k), $A_n \times B_n$ denotes a set of (a_k, b_k) such that, for a_k, the ball centered at X_k with radius a_k contains exactly n neighbors of X_k (resp. for b_k, the interval centered at Y_k with radius b_k contains exactly n neighbors of Y_k). Subsequently, for the kernel method we adapt the R-routine named *funopare.quantile.lcv*. We emphasize that, frequently, we consider a quadratic kernels defined as $L_1(u) = \dfrac{3}{2}(1-u^2)\mathbb{I}_{[0,1]}$ on $[0,1]$, $L_2(u) = \displaystyle\int_{-\infty}^{u} \frac{3}{4}(1-t^2)\mathbb{I}_{[-1,1]}(t)dt$ and the $\mathcal{L}_2$ semi-metric to measure the distance. Then, we compute the errors to evaluate the performance of these estimators as follows:

- The case of classical double-kernel method, the mean squared error (MSE) is

$$MSE(DKM) = \frac{1}{n}\sum_{k=1}^{n}\left(\widetilde{q}_{\alpha KM}(X_k) - q_\alpha(X_k)\right)^2 .$$

- The case of recursive double-kernel method, the mean squared error (MSE) is

$$MSE(RDKM) = \frac{1}{n}\sum_{k=1}^{n}\left(\widetilde{q}_{\alpha}(X_k) - q_\alpha(X_k)\right)^2 .$$

Therefore, the obtained results of mean squared error are summarized in Table 2.1, Table 2.2 and Table 2.3 for the different sample sizes $n = (100, 200, 500)$, while, Figure 2.2 simultaneously plots side by side, the estimated conditional quantiles by the RDKM and the ones estimated by the DKM.

e		$\lambda = 0.1$	$\lambda = 0.2$	$\lambda = 0.5$	$\lambda = 0.7$	$\lambda = 0.9$
	Q_1	4.9110	5.4123	6.9100	7.9121	8.9010
MSE(DKM)	Q_2	2.4420	2.9413	4.4401	5.4410	6.4321
	Q_3	4.2511	4.7515	6.2501	1.2552	2.2002
	Q_1	2.1302	2.3812	3.13402	3.6305	4.1328
MSE(RDKM)	Q_2	1.6920	1.9401	2.6955	3.1921	3.6921
	Q_3	2.2924	2.5443	3.2945	3.8421	4.3421

Table 2.1: *Mean Squared Error Results for* $n = 100$

		$\lambda = 0.1$	$\lambda = 0.2$	$\lambda = 0.5$	$\lambda = 0.7$	$\lambda = 0.9$
	Q_1	4.7801	5.2800	6.7843	7.7736	8.7718
MSE(DKM)	Q_2	2.3110	2.8109	4.3121	5.3102	6.3155
	Q_3	4.1220	4.6202	6.1201	1.1201	2.0761
	Q_1	2.0030	2.2500	3.0012	3.5050	4.0021
MSE(RDKM)	Q_2	1.5610	1.8191	2.5643	3.0667	3.5631
	Q_3	2.1609	2.4181	3.1661	3.7108	4.2157

Table 2.2: *Mean Squared Error Results for* $n = 200$

		$\lambda = 0.1$	$\lambda = 0.2$	$\lambda = 0.5$	$\lambda = 0.7$	$\lambda = 0.9$
	Q_1	4.3900	4.8901	5.7991	7.3865	8.3864
MSE(DKM)	Q_2	1.9210	2.4267	3.9267	4.9200	5.9101
	Q_3	3.7373	3.7201	3.7333	0.7343	1.6823
	Q_1	1.6101	1.8661	2.6113	2.6670	3.6753
MSE(RDKM)	Q_2	1.1702	0.9107	2.1787	2.6708	3.1702
	Q_3	1.7777	2.0266	2.7768	3.3252	3.8294

Table 2.3: *Mean Squared Error Results for* $n = 500$

Conclusion Table 2.1 (respectively Table 2.2 and Table 2.3) presents the MSE values of both estimates of the quartiles $Q_1(\alpha = 0.25)$, $Q_2(\alpha = 0.5)$ and $Q_3(\alpha = 0.75)$. This simulation involves two interpretations: The first is that Table 2.1 (respectively Table 2.2 and Table 2.3) is clearly showed that the proposed recursive double kernel method gives better results than the classical double kernel one in most of the studied situations. As can be seen from the Figure 2.2 as well. The second is that the MSE values increase more substantially (with respect to the value of λ) in the kernel method than in the recursive one and it decrease as the sample size n increases.

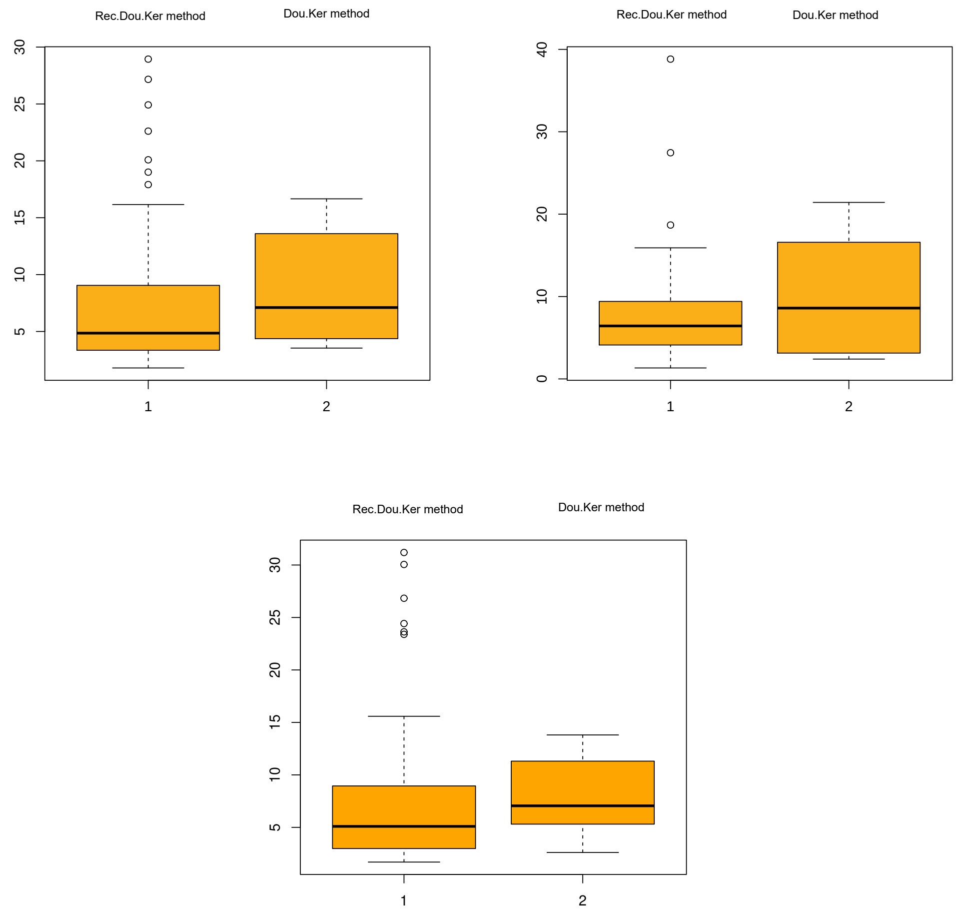

Figure 2.2: Conditional quantiles (Q1, Q2 and Q3) estimation by RDKM (on the left) versus DKM (on the right)

2.7 Appendix

Proof of Lemma 2.4.1 We use the same ideas as in Laib and Louani [16]. For all $k = 1, \ldots, n$, we start by defining

$$\eta_{n_k} = \left(\frac{\psi_n(x)\gamma_2^2}{\gamma_1 \alpha(1-\alpha)} \right)^{1/2} \left(L_{2,k}(\tau_\alpha(q_\alpha(x))) - F^x(\tau_\alpha(q_\alpha(x))) \right) \frac{L_{1,k}(x)}{n\psi_n(x)}$$

and then $\xi_{n_k} = \eta_{n_k} - \mathbb{E}(\eta_{n_k}/\wp_{k-1})$. Subsequently, it is easily seen that

$$\left(\frac{n\psi_n(x)\gamma_2^2}{\gamma_1\alpha(1-\alpha)}\right)^{1/2} \widehat{Q}_n(x,\tau_\alpha(t_\alpha(x))) = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\xi_{n_k}.$$

As ξ_{n_k} is a triangular array of martingale differences according the $\sigma-$fields $\wp_{k-1}$, we are in position to apply the central limit theorem based on unconditional lindeberg condition to prove the asymptotic normality of $\widehat{Q}_n(x)$. Thus, it suffices to establish the following two parts:

(a) $\displaystyle\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(\xi_{n_k}^2/\wp_{k-1}) \longrightarrow 1$ in probability;

(b) $\displaystyle\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(\xi_{n_k}^2\mathbb{I}_{\xi_{n_k}>\epsilon n}) \longrightarrow 0$ holds for any $\epsilon > 0$ (lindeberg condition).

Proof of part (a) The first part can be easily written as follows

$$\begin{aligned}\mathbb{E}(\xi_{n_k}^2/\wp_{k-1}) &= \mathbb{E}((\eta_{n_k} - \mathbb{E}(\eta_{n_k}/\wp_{k-1}))^2/\wp_{k-1})\\ &= \mathbb{E}(\eta_{n_k}^2/\wp_{k-1}) - \mathbb{E}^2(\eta_{n_k}/\wp_{k-1}).\end{aligned}$$

Thus, we need to prove the validity of the following two statements resulting from (a):

1. $\displaystyle\lim_{n\longrightarrow\infty}\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}^2(\eta_{n_k}/\wp_{k-1}) = 0_p$;

2. $\displaystyle\lim_{n\longrightarrow\infty}\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(\eta_{n_k}^2/\wp_{k-1}) = 1_p$.

For the first convergence, we have

$$\begin{aligned}
\mathbb{E}(\eta_{n_k}/\wp_{k-1}) &= \frac{1}{\psi_n(x)}\left(\frac{\psi_n(x)\gamma_2^2}{\gamma_1\alpha(1-\alpha)}\right)^{1/2}\mathbb{E}\Big[\Big[(L_{2,k}(\tau_\alpha(q_\alpha(x))) - F^x(\tau_\alpha(q_\alpha(x))))L_{1,k}(x)\Big]/\wp_{k-1}\Big] \\
|\mathbb{E}\left(\eta_{n_k}/\wp_{k-1}\right)| &= \frac{1}{\psi_n(x)}\left(\frac{\psi_n(x)\gamma_2^2}{\gamma_1\alpha(1-\alpha)}\right)^{1/2}\Big|\mathbb{E}\Big[\Big[(\mathbb{E}(L_{2,k}(\tau_\alpha(q_\alpha(x)))/\mathfrak{B}_{k-1})) \\
&\quad -F^x(\tau_\alpha(q_\alpha(x)))L_{1,k}(x)\Big]/\wp_{k-1}\Big]\Big| \\
&= \frac{1}{\psi_n(x)}\left(\frac{\psi_n(x)\gamma_2^2}{\gamma_1\alpha(1-\alpha)}\right)^{1/2}\Big|\mathbb{E}\Big[\Big[(\mathbb{E}(L_{2,k}(\tau_\alpha(q_\alpha(x)))/X_k)) \\
&\quad -F^x(\tau_\alpha(q_\alpha(x)))L_{1,k}(x)\Big]/\wp_{k-1}\Big]\Big|.
\end{aligned}$$

Then, under $(A.1)$ and $(A.4)$, we have:

$$C\phi_k(x,a_k) \le \mathbb{E}(L_{1,k}(x)/\wp_{k-1}) \le C'\phi_k(x,a_k).$$

After that, an integration by parts and a change of variable allow to get

$$\mathbb{E}(L_{2,k}((\tau_\alpha(q_\alpha(x)))/X_k)) = \int_{\mathbb{R}} L_2^{(1)}(t)F^{X_k}(y-b_kt)dt$$

and under $(A.2)$, we have

$$|\mathbb{E}(L_{2,k}(\tau_\alpha(q_\alpha(x)))/X_k) - F^x(\tau_\alpha(q_\alpha(x)))| \le ca_k^{\beta_1} + c'b_k^{\beta_2}.$$

Combining now these results, we have

$$|\mathbb{E}(\eta_{n_k}/\wp_{k-1})| \le C'\left(\frac{\psi_n(x)\gamma_2^2}{\gamma_1\alpha(1-\alpha)}\right)^{1/2}\left(ca_k^{\beta_1} + c'b_k^{\beta_2}\right)\frac{\phi_k(x,a_k)}{\psi_n(x)}.$$

Then, under $(A.3)$, we obtain

$$\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(\eta_{n_k}/\wp_{k-1})^2 \le C'^2\left(\frac{\gamma_2^2}{\gamma_1\alpha(1-\alpha)}\right)\sum_{k=1}^{n}\left(ca_k^{\beta_1} + c'b_k^{\beta_2}\right)^2\frac{\phi_k^2(x,a_k)}{n\psi_n(x)} = o_p(1)$$

Now, we treat the convergence (2). Indeed, we observe that

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n}\left(\mathbb{E}(\eta_{n_k}^2/\wp_{k-1})\right) &= \frac{1}{n\psi_n^2(x)}\left(\frac{\psi_n(x)\gamma_2^2}{\gamma_1\alpha(1-\alpha)}\right)\sum_{k=1}^{n}\mathbb{E}\Big[\Big(L_{2,k}(\tau_\alpha(q_\alpha(x)))\\ &-F^x(\tau_\alpha(q_\alpha(x)))\Big)^2 L_{1,k}^2(x)/\wp_{k-1}\Big]\\ &= \frac{1}{n\psi_n^2(x)}\left(\frac{\psi_n(x)\gamma_2^2}{\gamma_1\alpha(1-\alpha)}\right)\sum_{k=1}^{n}\Bigg[\mathbb{E}\left(L_{2,k}^2(\tau_\alpha(q_\alpha(x)))L_{1,k}^2(x)/\wp_{k-1}\right)\\ &-2F^x(\tau_\alpha(q_\alpha(x)))\mathbb{E}\left(L_{2,k}(\tau_\alpha(q_\alpha(x)))L_{1,k}^2(x)/\wp_{k-1}\right)\\ &+(F^x(\tau_\alpha(q_\alpha(x))))^2\mathbb{E}\left(L_{1,k}^2(x)/\wp_{k-1}\right)\Bigg].\end{aligned}$$

We put:

$$\begin{aligned}I_1 &= \sum_{k=1}^{n}\mathbb{E}\left(L_{2,k}^2(\tau_\alpha(q_\alpha(x)))L_{1,k}^2(x)/\wp_{k-1}\right)\\ I_2 &= \sum_{k=1}^{n}\mathbb{E}\left(L_{2,k}(\tau_\alpha(q_\alpha(x)))L_{1,k}^2(x)/\wp_{k-1}\right)\\ I_3 &= \sum_{k=1}^{n}\mathbb{E}\left(L_{1,k}^2(x)/\wp_{k-1}\right).\end{aligned}$$

We write:

$$\begin{aligned}I_1 &= F^x(\tau_\alpha(q_\alpha(x)))\sum_{k=1}^{n}\mathbb{E}\left[L_{1,k}^2(x)/\wp_{k-1}\right]+\sum_{k=1}^{n}\mathbb{E}\left[L_{2,k}^2(\tau_\alpha(q_\alpha(x)))L_{1,k}^2(x)/\wp_{k-1}\big)\right]\\ &- F^x(\tau_\alpha(q_\alpha(x)))\sum_{k=1}^{n}\mathbb{E}\left[L_{1,k}^2(x)/\wp_{k-1}\right]\\ &= F^x(\tau_\alpha(q_\alpha(x)))\sum_{k=1}^{n}\mathbb{E}\left[L_{1,k}^2(x)/\wp_{k-1}\right]+\sum_{k=1}^{n}\mathbb{E}\left[(\mathbb{E}(L_{2,k}^2(\tau_\alpha(q_\alpha(x)))/X_k)L_{1,k}^2(x))/\wp_{k-1}\big)\right]\\ &- F^x(\tau_\alpha(q_\alpha(x)))\sum_{k=1}^{n}\mathbb{E}\left[L_{1,k}^2(x)/\wp_{k-1}\right]\\ &\leq \sum_{k=1}^{n}\mathbb{E}\left[(\mathbb{E}(L_{2,k}^2(\tau_\alpha(q_\alpha(x)))/X_k)L_{1,k}^2(x))/\wp_{k-1}\big)\right]-F^x(\tau_\alpha(q_\alpha(x)))\sum_{k=1}^{n}\mathbb{E}\left[L_{1,k}^2(x)/\wp_{k-1}\right].\end{aligned}$$

Using the same argument as those used in proof of part (a), then we have:

$$\frac{1}{n\psi_n(x, a_k)} I_2 = \circ(1).$$

For I_3, we use the same ideas as in Ferraty et al (2009) to get:

$$\mathbb{E}\left[L_{1,k}^2(x)/\wp_{k-1}\right] = L_1^2(1)\phi_k(x, a_k) - \int_0^1 (L_1^2(u))'\phi_k(x, ua_k)du$$

and

$$\mathbb{E}\left[L_{1,1}(x)/\wp_{k-1}\right] = L_1(1)\phi_k(x, a_k) - \int_0^1 (L_1(u))'\phi_k(x, ua_k)du$$

so under (A.1), we have:

$$\begin{aligned}
\frac{1}{n\psi_n(x, a_k)} \sum_{k=1}^{n} \mathbb{E}\left[L_{1,k}^2(x)/\wp_{k-1}\right] &= \frac{L_1^2(1)}{n\psi_n(x, a_k)} \sum_{k=1}^{n} \phi_k(x, a_k) \\
&- \int_0^1 (L_1^2(u))' \frac{1}{n\psi_n(x, a_k)} \sum_{k=1}^{n} \phi_k(x, ua_k)du \\
&= L_1^2(1) - \int_0^1 (L_1^2(u))'\beta_x(u)du + o_p(1) = \gamma_2 + o_p(1)
\end{aligned}$$

and

$$\frac{1}{n\psi_n(x, a_k)} \mathbb{E}\left[L_{1,1}(x)/\wp_{k-1}\right] = \gamma_1 + o_p(1)$$

we deduce that $\lim_{n\longrightarrow\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}(\eta_{n_k}^2/\wp_{k-1}) = 1$ which complete the proof of part (a).

Proof of part (b) The lindeberg condition implies that

$$\mathbb{E}\left[\xi_{n_k}^2 \mathbb{I}_{\xi_{n_k} > n\epsilon}\right] \leq 4\mathbb{E}\left[\eta_{n_k}^2 \mathbb{I}_{\eta_{n_k} > n\epsilon/2}\right].$$

Let $a > 1$ and $b > 1$ such that $\frac{1}{a} + \frac{1}{b} = 1$. Making use now the hölder and markov inequalities, one can write for all $\epsilon > 0$

$$\mathbb{E}\left[\eta_{n_k}^2 \mathbb{I}_{\eta_{n_k} > n\epsilon/2}\right] \leq \frac{\mathbb{E}\left(\eta_{n_k}\right)^{2a}}{(n\epsilon/2)^{2a/b}}.$$

Taking $C_0 \in \mathbb{R}_+^*$ and $2a = 2+\delta$ for some $\delta > 0$, such that $\mathbb{E}\left(|Y_k|^{2+\delta}\right) < \infty$ and $\mathbb{E}\left(|L_{2,k} - F^x|^{2+\delta}/X_k = u\right) = \overline{W}_{2+\delta}(u)$ is a continuous function, to get the following

$$\begin{aligned}
4\mathbb{E}\left[\eta_{n_k}^2 \mathbb{I}_{\eta_{n_k} > n\epsilon/2}\right] &\leq C_0 \left(\frac{\psi_n(x,a_k)\gamma_1^2}{\gamma_2(\alpha(1-\alpha))}\right)^{(2+\delta)} \frac{1}{(\psi_n(x,a_k))^{2+\delta}} \left[(|L_{2,k} - F^x|^{2+\delta} L_{1,k}^{2+\delta}(x))^{2+\delta}\right] \\
&\leq C_0 \left(\frac{\psi_n(x,a_k)\gamma_1^2}{\gamma_2(\alpha(1-\alpha))}\right)^{(2+\delta)} \frac{1}{(\psi_n(x,a_k))^{2+\delta}} \mathbb{E}\left[\mathbb{E}\left[|L_{2,k} - F^x|^{2+\delta} L_{1,k}^{2+\delta}(x)/X_k\right]\right] \\
&\leq C_0 \left(\frac{\psi_n(x,a_k)\gamma_1^2}{\gamma_2(\alpha(1-\alpha))}\right)^{(2+\delta)} \frac{1}{(\psi_n(x,a_k))^{2+\delta}} \mathbb{E}\left[L_{1,k}^{2+\delta}(x)\overline{W}_{2+\delta}(X_k)\right] \\
&\leq C_0 \left(\frac{\psi_n(x,a_k)\gamma_1^2}{\gamma_2(\alpha(1-\alpha))}\right)^{(2+\delta)} \frac{1}{(\psi_n(x,a_k))^{2+\delta}} \mathbb{E}\Bigg[L_{1,k}^{2+\delta}(x)|\overline{W}_{2+\delta}(X_k) - \overline{W}_{2+\delta}(x)| \\
&\quad + |\overline{W}_{2+\delta}(x)\mathbb{E}(L_{1,k}^{2+\delta}(x))|\Bigg] \\
&\leq C_0 \left(\frac{\gamma_1^2}{\gamma_2(\alpha(1-\alpha))}\right)^{(2+\delta)} \left(\mathbb{E}(L_{1,k}^{2+\delta}(x))\left(|\overline{W}_{2+\delta}(x)| + \circ(1)\right)\right).
\end{aligned}$$

Consequently $\dfrac{1}{n}\displaystyle\sum_{k=1}^{n} \mathbb{E}\left[\xi_{n_k}^2 \mathbb{I}_{\xi_{n_k} > n\epsilon}\right] \longrightarrow 0$ *as* $n \longrightarrow \infty$ which completes the proof of this lemma.

Proof of Lemma 2.4.2 Observe that:

$$\begin{aligned}
\widehat{F}_D(x) - 1 &= \frac{1}{n\psi_n(x,a_k)} \sum_{k=1}^{n} \Big[[L_{1,k}(x) - \mathbb{E}(L_{1,k}(x)/\wp_{k-1})] + [\mathbb{E}(L_{1,k}(x)/\wp_{k-1})] - 1\Big] \\
&= \underbrace{\frac{1}{n\psi_n(x,a_k)} \sum_{k=1}^{n} [L_{1,k}(x) - \mathbb{E}(L_{1,k}(x)/\wp_{k-1})]}_{T_1} \\
&\quad + \underbrace{\frac{1}{n\psi_n(x,a_k)} \sum_{k=1}^{n} [\mathbb{E}(L_{1,k}(x)/\wp_{k-1})] - 1}_{T_2}.
\end{aligned}$$

The proof of this lemma follows then if we can show that:

1. $T_1 = \circ(1)$ as $n \longrightarrow \infty$;

2. $T_2 \longrightarrow 0$ in probability as $n \longrightarrow \infty$.

For T_2, under (A.1) and (A.3) we prove that:

$$\frac{1}{n\psi_n(x,a_k)}\sum_{k=1}^{n}\left[\mathbb{E}(L_{1,k}(x)/\wp_{k-1})\right]=\circ(1)\quad as\ n\longrightarrow\infty.$$

So, it is easily seen that

$$T_2\longrightarrow 0 \text{ in probability as } n\longrightarrow\infty.$$

For the first term T_1, observe that $T_1(x)=\sum_{k=1}^{n}L_{n_k}(x)$ where $\{L_{n_k}(x)\}$ is a triangular array of martingale differences with respect to the $\sigma-$field $\wp_{k-1}$. Combining next the Burkholder inequality (see P.Hall and C.Heyde 1980) and Jensen inequality (see Laib and Louani 2011), we obtain for any $\epsilon>0$, there exists a constant $C_0>0$ such that

$$\mathbb{P}\left(|T_1|>\epsilon\right)\leq C_0\frac{\mathbb{E}\left(L_{1,1}^2(x)\right)}{\epsilon^2 n(\psi_n(x,a_k))^2}=\circ\left(\frac{1}{\epsilon^2 n\psi_n(x)}+\circ(1)\right).$$

Since $n\psi_n(x)\longrightarrow\infty$, we conclude then that $T_1(x)=\circ(1)$ in probability as $n\longrightarrow\infty$ which completes the proof of Lemma 2.4.2.

Proof of Lemma 2.4.3 We have

$$\widehat{B}_n(x,\tau_\alpha(q_\alpha(x)))=\frac{\overline{F}_N(x,\tau_\alpha(q_\alpha(x)))}{\overline{F}_D(x,\tau_\alpha(q_\alpha(x)))}.$$

Thus, we write

$$\begin{aligned}|\widehat{B}_n(x,\tau_\alpha(q_\alpha(x)))| &= \frac{1}{\sum_{k=1}^{n}\mathbb{E}(L_{1,k}(x)/\wp_{k-1})}\sum_{k=1}^{n}\Big[\mathbb{E}\left[L_{1,k}(x)\mathbb{E}\left[L_{2,k}(\tau_\alpha(q_\alpha(x)))/\mathfrak{B}_{k-1}\right]/\wp_{k-1}\right]\\ &-F^x(\tau_\alpha(q_\alpha(x)))\mathbb{E}\left[L_{1,k}(x)/\wp_{k-1}\right]\Big]\\ &= \frac{1}{\sum_{k=1}^{n}\mathbb{E}(L_{1,k}(x)/\wp_{k-1})}\sum_{k=1}^{n}[\mathbb{E}\left[L_{1,k}(x)\mathbb{E}\left[L_{2,k}(\tau_\alpha(q_\alpha(x)))/X_k\right]/\wp_{k-1}\right]\\ &-F^x(\tau_\alpha(q_\alpha(x)))\mathbb{E}\left[L_{1,k}(x)/\wp_{k-1}\right]].\end{aligned}$$

Where

$$|\widehat{B}_n(x,\tau_\alpha(q_\alpha(x)))| \leq \frac{1}{\sum_{k=1}^n \mathbb{E}(L_{1,k}(x)/\wp_{k-1})} \sum_{k=1}^n \mathbb{E}\Big[L_{1,k}(x)|\mathbb{E}\left[L_{2,k}(\tau_\alpha(q_\alpha(x)))/X_k\right] - F^x(\tau_\alpha(q_\alpha(x)))|/\wp_{k-1}\Big].$$

Next, an integration by parts and a change of variable allow to get:

$$\mathbb{E}\left[L_{2,k}(\tau_\alpha(q_\alpha(x)))/X_k\right] = \int_{\mathbb{R}} L_2^{(1)}(t)F^{X_k}(\tau_\alpha(q_\alpha(x)) - b_k t)dt.$$

Thus, we have

$$|\mathbb{E}\left[L_{2,k}(\tau_\alpha(q_\alpha(x)))/X_k\right] - F^x(\tau_\alpha(q_\alpha(x)))| \leq \int_{\mathbb{R}} L_2^{(1)}(t)|F^{X_k}((\tau_\alpha(q_\alpha(x))) - b_k t) - F^x(\tau_\alpha(q_\alpha(x)))|dt \tag{2.4}$$

Under $(A.2)$, we obtain that

$$\mathbb{I}_{B(x,a_k)}(X_k)|\mathbb{E}\left[L_{2,k}(\tau_\alpha(q_\alpha(x)))/X_k\right] - F^x(\tau_\alpha(q_\alpha(x)))| \leq C\int_{\mathbb{R}} L_2^{(1)}(t)(a_k^{\beta_1} + |t|^{\beta_2} b_k^{\beta_2})dt \tag{2.5}$$

and under $(A.4)$, we prove easily that

$$\frac{1}{n}\sum_{k=1}^n \mathbb{E}\left[L_{1,k}(x)/\wp_{k-1}\right] = \circ(1). \tag{2.6}$$

As a final step, combining the statements (2.5) and (2.6) and we achieve the proof of our lemma.

Proof of Lemma 2.4.4 We put $t = \tau_\alpha(q_\alpha(x))$ and we write

$$\begin{aligned}\widehat{R}_n(x,t) &= -\left(\widehat{B}_n(x,t) - F^x(t)\right)\left(\widehat{F}_N(x,t) - \overline{F}_N(x,t)\right)\\ &= -\left(\frac{\overline{F}_N(x,t) - F^x(t)\overline{F}_D(x,t)}{\overline{F}_D(x,t)}\right)\left(\widehat{F}_N(x,t) - \overline{F}_N(x,t)\right).\end{aligned}$$

Clearly, it is suffices to show that:

1. $\left(\dfrac{\overline{F}_N(x,t) - F^x(t)\overline{F}_D(x,t)}{\overline{F}_D(x,t)}\right) = \circ(1);$

2. $\left(\widehat{F}_N(x,t) - \overline{F}_N(x,t)\right) = \circ(1).$

The proof of the first part uses arguments similar to those used in the proof of Lemma 2.4.3. While the second part will be established if these two following insertions are checked

(i) $\mathbb{E}\left(\widehat{F}_N(x,t) - \overline{F}_N(x,t)\right) = 0;$

(ii) $var\left(\widehat{F}_N(x,t) - \overline{F}_N(x,t)\right) \longrightarrow 0$ as $n \longrightarrow \infty$.

For all $k = 1, \cdots, n$, we put

$$\triangle_k(x,t) = \frac{1}{n\psi_n(x,a_k)}\left[L_{1,k}(x)L_{2,k}(t) - \mathbb{E}\left[L_{1,k}(x)L_{2,k}(t)/\wp_{k-1}\right]\right]$$

where $\triangle_k(x,t)$ is a triangular array of martingale differences according to the $\sigma-$fields $\wp_{k-1}$. Next, by $(A.1)(ii)$ and $(A.4)$ we obtain

$$\widehat{F}_N(x,t) - \overline{F}_N(x,t) = \sum_{k=1}^{n} \triangle_k(x,t).$$

By definition of $\triangle_k(x,t)$, $\mathbb{E}(\triangle_k(x,t)) = 0$. For **(ii)**, we have by Burkholder's inequality

$$\mathbb{E}\left[\sum_{k=1}^{n}(\triangle_k(x,t))\right]^2 \leq \sum_{k=1}^{n}\mathbb{E}\left[\triangle_k^2(x,t)\right].$$

Furthermore, by Jensen inequality we have:

$$\begin{aligned}
\mathbb{E}\left[\triangle_k^2(x,t)\right] &\leq \frac{1}{(n\psi_n(x,a_k))^2}\mathbb{E}\left[L_{1,k}^2(x)L_{2,k}^2(t)/\wp_{k-1}\right]\\
&\leq \frac{1}{(n\psi_n(x,a_k))^2}\mathbb{E}\left[L_{1,k}^2(x)/\wp_{k-1}\right]\\
&\leq \frac{1}{(n\psi_n(x,a_k))^2}\mathbb{P}\left(X_k \in B(x,a_k)/\wp_{k-1}\right)\\
&\leq \frac{1}{(n\psi_n(x,a_k))^2}\phi_k(x,a_k).
\end{aligned}$$

Likewise, we get

$$\sum_{k=1}^{n} \mathbb{E}\left[\triangle_k^2(x,t)\right] \leq \frac{\displaystyle\sum_{k=1}^{n} \phi_k(x,a_k)}{n^2\psi_n^2(x,a_k)}.$$

Thence, since $(A.1)(ii)$ is verified, we deduce that

$$var\left(\widehat{F}_N(x,t) - \overline{F}_N(x,t)\right) \longrightarrow 0 \quad as \quad n \longrightarrow \infty.$$

Bibliography

[1] A. Amiri, *Estimateurs fonctionnels récursifs et leurs applications à la prévision*, Doctoral Thesis, the University of Avignon and the Pays de Vaucluse, (2010).

[2] A. Amiri, C. Crambes and B. Thiam, Recursive estimation of nonparametric regression with functional covariate, *Computational statistics and data analysis*, **69** (2014), 154-172.

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

[4] A. Bouadjemi, Asymptotic normality of the recursive kernel estimate of conditional cumulative distribution function, *J Probab Stat Sci*, **12(2)** (2014), 117-126.

[5] H. Cardot, C. Crambes and P. Sarda, Quantile regression when the covariates are functions, *Nonparametric Statistics* **17(7)** (2005), 841-856.

[6] H. Cardot, P. Cénac, P. A. Zitt et al, Recursive estimation of the conditional geometric median in hilbert spaces, *Electronic journal of statistics* **6** (2012), 2535-2562.

[7] M. Chaouch and S. Khardani, Randomly censored quantile regression estimation using functional stationary ergodic data, *Journal of Nonparametric Statistics* **27:1** (2015), 65-87.

[8] S. Dabo-Niang and A. Laksaci, Nonparametric estimation of conditional quantiles when the regressor is valued in a semi-metric space, (2008).

[9] M. Ezzahrioui and E. Ould-Saïd, Asymptotic normality of the kernel estimator of conditional quantiles in a normed space, *Far East Journal of Theoretical Statistics* **25** (2008a), 15-38.

[10] M. Ezzahrioui and E. Ould-Saïd, Asymptotic results of a nonparametric conditional quantile estimator for functional time series, *Communications in Statistics, Theory and Methods* **37(17)**(2008b), 2735-2759.

[11] F. Ferraty, A. Laksaci and P. Vieu, Estimating some characteristics of the conditional distribution in nonparametric functional models, *Statistical Inference for Stochastic Processes* **9(1)** (2006), 47-76.

[12] F. Ferraty and P. Vieu, Nonparametric functional data analysis, *theory and practice. Springer Science and Business Media* (2006).

[13] A. Gannoun, J. Saracco and K. Yu, Nonparametric prediction by conditional median and quantiles, *Journal of statistical Planning and inference* **117(2)** (2003), 207-223.

[14] R. Koenker, Quantile regression, *econometric society monograph series* (2005).

[15] R. Koenker and G. Bassett, Regression quantiles, *econometrika*, **46** (1978), 33-50, Mathematical Reviews (MathSciNet): MR474644 Digital Object Identifier: doi, 10:1913643.

[16] N. Laib and D. Louani, Nonparametric kernel regression estimation for functional stationary ergodic data: asymptotic properties, *Journal of Multivariate analysis*, **101(10)** (2010), 2266-2281.

[17] A. Laksaci, M. Lemdani and E. Ould-Saïd, A generalized l1-approach for a kernel estimator of conditional quantile with functional regressors: Consistency and asymptotic normality, *Statistics and probability letters*, **79(8)** (2009), 1065-1073.

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

[19] E. Masry and L. Györfi, Strong consistency and rates for recursive probability density estimators of stationary processes, *Journal of Multivariate Analysis*, **22(1)** (1987), 79-93.

[20] C. J. Stone, Consistent nonparametric regression, *The annals of statistics*, (1977), 595-620.

[21] C. T. Wolverton and T. J. Wagner, Recursive estimates of probability densities, *IEEE Transactions on Systems Science and Cybernetics*, **5(3)** (1969), 246-247.

[22] L. Zhengyan and L. Degui, Asymptotic normality for L_1-norm kernel estimator of conditional median under association dependence, *Journal of Multivariate Analysis* **98** (2007), 1214-1230.

Chapter 3

Nonparametric Recursive Estimate for Right-Censored Conditional Mode Function with Ergodic Functional Data

I. Bouazza, F. Benziadi and T. Guendouzi

Published in *Journal of the Indian Society for Probability and Statistics*

Abstract *In this chapter, as an extension of some recent works that are essential references for this current contribution, we provide a recursive nonparametric approach to estimate the conditional mode function* $\Theta(x)$ *of the conditional density of a scalar positive random variable* Z *given an Hilbertian explanatory process* $X = x$*, denoted by* $\xi^x(z)$*, based on the randomly right-censorship model which is the new and main factor here. Our nonparametric model takes into account the fact that the response variable* Z *referred as a survival time is right-censored by another variable* W *independent of* (Z, X)*. Afterwards, we establish, under stationary and ergodic conditions, by using an adaptive exponential inequality to this context, some theoretical properties of the resulting estimator including the uniform almost sure convergence (with rates), after establishing the pointwise ones. Finally, an application based on simulated data is* *conducted to illustrate our results.*

Key words and phrases *Recursive nonparametric estimate, Conditional mode, Right-censoring, Functional data, Ergodic process, Almost sure convergence, Martingale differences.*

3.1 Introduction

The problem of modeling and then estimating the relationship between two variables, a covariate X and a real survival response variable Z which is assumed to be affected by a right-censorship phenomenon, has taken great attention in statistics and this has required the development of some statistical methods in this context. A quick look at the nonparametric literature shows that studying conditional models in the prevalence of incomplete information occupies an important space in major fields of application such as medicine, industrial sciences, sociology and others; and it has recorded a significant footprint in the last recent decade. The emergence of such incomplete samples is imperative in most experiments of a statistical nature. The right censorship is a key analytical problem, is the most popular type of incomplete data and is the result of competing failure modes. This is the case in survival studies, for example, when keeping track of the influence of a treatment on chronic diseases (alcoholic disease, diabetic, cancer,...). Two cases are then possible: when the patient is lost from the follow-up or withdraws or rather the study ends before the event has occurred. The use of such a kind of data is advisable, as there are, in reality, companies and institutions that have the task of processing the data before making it available to the public.

The core focus in this type of problems has always been to obtain results that are less sensitive to outliers. The conditional mode $\Theta(x) = \arg \max_{-\infty<y<\infty} \xi^x(y)$ is one of the appropriate and widely studied models in the right-censoring case which provides a robust statistical modeling. Literally, a great deal of recent works provides a comprehensive look on this topic, within the use of uncensored observations. This has been recorded for the estimation of the conditional mode function for both i.i.d. and

$\alpha-$mixing cases, including Dabo-Niang et al. (2014)[9], Didi and Louani (2014)[12] for the finite dimension and Dabo-Niang and Laksaci (2007)[8], Ezzahrioui and O.Saïd (2008(a) [13], 2010[14]), Ferraty et al. (2010)[17], Attaoui et al. (2011)[2] and Ling et al. (2016)[27] for the infinite dimension.

For the other case, when the response variable is right-censored, a specific methodology for this type of data is essential. This problem has been developed by O.Saïd and Cai (2005)[30] who have used a new smooth kernel estimator of the conditional mode function and have established the uniform strong consistency with the rates. A similar approach is discussed by Khardani et al. (2010)[20] who have investigated some asymptotic properties including the almost sure convergence as well as the asymptotic normality. On the other hand and quite recently, under the α-mixing condition, Khardani and Thiam (2016)[23] have proved the almost complete convergence when the process takes values in some functional space by using the Fuk-Nagaev inequality to get such results. In such a stationary mixing context, Baek in collaboration with Li-Niu (2016)[3] has proposed an estimator of this conditional model for the random left truncated and right censored (LTRC) type consisting of three components Y, T and W. He has obtained over a compact set its uniform strong consistency with rates and its asymptotic normality, and this expresses several ideas close to our approach. According to this framework, and despite the already obtained outcomes, most of the existing studies treat the finite case where the explanatory process is a scalar or a vector $(\mathbb{R}^p, p \geq 1)$.

In the functional data framework, no results have been recorded for the recursive nonparametric estimation combining incomplete (i.e censored) and ergodic data. The main idea is inspired from the previous study of Khardani and Semmar (2014)[22], which is considered as the first and most recent work that treats a recursive estimate for censored observations. Our principal goal here is to prove, under the condition of ergodicity and from such an incomplete data, that the proposed estimate is uniformly consistent by adopting the recursive kernel version. Indeed, it is essential to note that our investigation is an extension of the recent result developed in Ardjoun et al. (2016)[1] from a general setting (the case where the response of interest is complete) to the case where such a variable is subject to random right-censoring. According to

Ardjoun et al. (2016)[1], the nonparametric recursive kernel estimate of the conditional density is given by

$$\overline{\xi}_n^x(y) = \frac{\sum_{k=1}^{n} b_k^{-1} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) L_2\left(b_k^{-1}(y - Y_k)\right)}{\sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)} \tag{3.1}$$

and therefore

$$\overline{\Theta}_n(x) = \arg\max_{y \in \mathbb{R}} \overline{\xi}_n^x(y) \tag{3.2}$$

where L_1 and L_2 are two kernels, a_k (resp. b_k) is a sequence of positive real numbers tending to 0 as $n \to \infty$.

In an effort to arrange the remaining paper's ideas, Section 3.2 sets a description of the proposed model as well as its estimator in the context of right-censored data. The necessary assumptions with a brief discussion are also given in Section 3.3. Section 3.4 states the pointwise and uniform consistencies of this estimator with some particular cases which provide a better coverage of our study. In Section 3.6, a small numerical study highlights this factor's effects on the behavior of the estimator, and gives more efficiency to our results from a practical point of view. In terms of calculation, (under some additional conditions) there is no much difference between the modified estimator even by taking into account the "statistically right-censored" data and the above estimate, and this is detailed in Section 3.5 to make reading easier.

3.2 Construction of the estimator under random-censoring scheme

We consider on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ a nonnegative stationary and ergodic random sequence $\Gamma_k = \{(X_k, Z_k), k = 1, \ldots, n\}$ with a common joint probability distribution function $\varphi_{(X,Z)}$. We also assume that this observed sequence (X_k, Z_k) distributed as the couple (X, Z) such that X_k takes values in an infinite dimensional space $\mathcal{H}$ (whose distance is defined by $d_{\mathcal{H}}(x, X_k) = \|x - X_k\|$), whereas Z_k represents a real-valued random variable.

3.2.1 Randomly censored framework

First of all, under the randomly right-censored (RRC) model, there are some basic concepts and formal vocabularies which are fundamentally different from those usually encountered in statistics that we must highlight to better understand the structure of these data. For a random sample of n subjects, we don't observe the unknown variable Z_k which interests us, but we observe the triplet of data incorporating $\{(X_k, Y_k, \Delta_k), k = 1, \ldots, n\}$. Therefore, the time actually observed is defined as a minimum of two deadlines, such that

$$\begin{cases} Y_k & := Z_k \wedge W_k \\ \Delta_k & = \mathbb{I}_{[Z_k \leq W_k]} \end{cases} \quad k \in \{1, \ldots, n\}$$

where $(Z_k)_{1 \leq k \leq n}$ represents a sample of lifetime random variables under study, $(W_k)_{1 \leq k \leq n}$ a sample of right censoring times, with the associated unknown continuous distribution functions (f.d.r.) F^Z and G^W respectively, Δ_k the censoring status marking either occurrence or censorship (Bernoulli r.v, in other references) and which is worth 0 if the survival time is right-censored ($Z_k > W_k$) and 1 otherwise ($Z_k \leq W_k$). Here, we can add that in case of an uncensored framework, we have $Y_k = Z_k$ and $\Delta_k = 1$.

However, in most cases, the problem is that the used survival functions $\overline{F}^Z$ and $\overline{G}^W$ remain generally unknown. This requires a modification of the latter two taking into account the presence of censoring, by replacing them with their non-parametric consistent Kaplan-Meier estimates (KME also called product-limit estimates) introduced in [19] and which generalize the empirical ones to the censored case, respectively

$$\overline{F}_n^Z(z) = 1 - F_n^Z(z) = \begin{cases} 0 & ; z \geq Y_{(n)} \\ \prod_{k=1}^{n} \left(1 - \frac{\Delta_{(k)}}{n-k+1}\right)^{\mathbb{I}\left(Y_{(k)} \leq z\right)} & ; z < Y_{(n)} \end{cases}$$

and

$$\overline{G}_n^W(z) = 1 - G_n^W(z) = \begin{cases} 0 & ; z \geq Y_{(n)} \\ \prod_{k=1}^{n} \left(1 - \frac{1-\Delta_{(k)}}{n-k+1}\right)^{\mathbb{I}\left(Y_{(k)} \leq z\right)} & ; z < Y_{(n)} \end{cases}$$

Denoting $Y_{(1)} < Y_{(2)} < \cdots < Y_{(n)}$ the order statistics of $(Y_k)_{k\in\{1,\dots,n\}}$, along with their corresponding concomitant $\Delta_{(k)}$. Under certain regularity conditions, and as a basic result, this latter estimator converges almost surely and uniformly to $\overline{G}^W$ (see for instance Kohler et al (2002)[24]). In what follows, the generating process $(X_k, Z_k)_{k\in\{1,\dots,n\}}$ assumed for the right-censored structure satisfies the ergodic property. The data which achieve this property are abundant.

3.2.2 Estimating conditional mode function from ergodic censored data using recursive method

Given the covariate $X = x$, we assume that the conditional distribution function of Z exists and is often expressed as

$$F_{Z/X}(z/x) = \mathbb{E}\left[\mathbb{I}_{(Z\leq z)}/X = x\right], \ \forall z \in \mathbb{R}.$$

In this situation, if the purpose is the conditional mode which estimator is defined as a random variable $\widehat{\Theta}_n(x)$ maximizing $\widehat{\xi}_n^x(z)$, then, one wants to set for any fixed x, a continuously differentiable real function $\xi^x(z)$

$$\xi^x(z) = \frac{\partial F_{Z/X}}{\partial z}(z/x).$$

Thus, we suppose in the framework of ergodic rightly-censored data the uniqueness of $\widehat{\Theta}_n(x)$ on a compact of $\mathbb{R}$, such that

$$\widehat{\Theta}_n(x) = \arg\max_{z\in\mathbb{R}} \widehat{\xi}_n^x(z) \tag{3.3}$$

Similar to the construction of the above estimator (3.1), and on the same way of Carbonez et al. (1995)[6] and more recently Khardani and Semmar (2014)[22], one can define, given $X = x$, the new estimator $\widehat{\xi}_n^x(z)$ based on the recursive kernel technique of the conditional density function of Z for censored data as follow

$$\widehat{\xi}_n^x(z) = \frac{\widehat{\varphi}_n(x,z)}{\widehat{\gamma}_n(x)} = \frac{\sum_{k=1}^{n} \frac{\Delta_k}{b_k \overline{G}_n^W(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) L_2\left(b_k^{-1}(z - Y_k)\right)}{\sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)}, \tag{3.4}$$

where

$$\widehat{\varphi}_n(x,z) = (\chi_n(x, a_k))^{-1} \sum_{k=1}^{n} \frac{\Delta_k}{b_k \overline{G}_n^W(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) L_2\left(b_k^{-1}(z - Y_k)\right).$$

Here, $\varphi_{(X,Z)}(\cdot,\cdot)$ is assumed to be continuous and bounded and $\gamma_X(\cdot)$ is the marginal density of the explicative variable X which verifies $\gamma_X(x) > \mu$ for all $x \in \mathcal{H}$ and $\mu > 0$. According to our knowledge, this kind of estimate has been extensively studied over the past few years in many papers, combining censored and finite-dimensional processes.

3.3 Main assumptions

To prove that our estimate is strongly consistent, we have chosen to use the key notations often introduced in the literature, $\wp_k$ the $\sigma-$field generated by $\{(X_s, Z_s);\ 1 \leq s < k\}$ and $\mathcal{B}_k$ the one generated by $\{(X_s, Z_s), (X_r), 1 \leq s < k;\ k \leq r \leq k+1\}$. On the other hand, let $\mathcal{S}$ and $\mathcal{I}$ be respectively two compact sets of $\mathcal{H}$ and $\mathbb{R}$. We will set also $T_{F^Z} := \sup\{z \in \mathbb{R} : \overline{F}^Z(z) > 0\} < \infty$ ($T_{F^Z} := \sup\{z \in \mathbb{R} : \overline{F}^Z(z) < 1\}$) and $T_{G^W} := \sup\{z \in \mathbb{R} : \overline{G}^W(z) > 0\}$ the support right endpoints with clarification $\overline{F}^Z(z) = \mathbb{P}(Z > z)$ and $\overline{G}^W(z) = \mathbb{P}(W > z)$. Moreover, to simplify the presentation of our main results and their proofs, we need to set the next regularity assumptions.

(**A.1**) On the hilbertian variable: there is a ball B of radius $h > 0$ centered at x such that

(i) $\forall x \in \mathcal{S},\ 0 < \phi_x(h) \leq \mathbb{P}[X \in B(x,h)]$ *and* $\phi_x(h) \to 0$ *as* $h \to 0$;

(ii) *A deterministic function* $\phi_k(x,.)$ *exists for all* $k \in \{1,\ldots,n\}$ *such that* $\mathbb{P}[d_{\mathcal{H}}(x,X_k) \leq h/\wp_{k-1}] \leq \phi_k(x,h)$ *where* $\lim_{h\to 0}\phi_k(x,h) = 0$ *a.s*;

(iii) $(\chi_n(x,h_k))^{-1}\sum_{k=1}^{n}\mathbb{P}[X_k \in B(x,h_k)/\wp_{k-1}] \to 1,$ *almost surely as* $n \to \infty$.

(**A.2**) On the nonparametric model: $\forall (z_1,z_2) \in \mathcal{I}^2$, $\forall (x_1,x_2) \in \mathcal{N}_x^2$, $\xi^x(z)$ satisfies the lipschitz condition

$$|\xi^{x_1}(z_1) - \xi^{x_2}(z_2)| \leq C_1\left(d_{\mathcal{H}}^{\nu_1}(x_1,x_2) + |z_1 - z_2|^{\nu_2}\right),\ \textit{with}\ C_1 > 0, \nu_1 > 0, \nu_2 > 0.$$

(**A.3**) L_1 is a measurable non-negative continuous bounded function on its compact support $(0,1)$. Also, it is supposed to be Hölderian of order β_1 for $\beta_1 > 0$.

(**A.4**) L_2 is an increasing, continuous and bounded function satisfying:

(i) $\forall (z_1,z_2) \in \mathcal{I}^2,\ \ |L_2(z_1) - L_2(z_2)| \leq C_2|z_1 - z_2|\,, \int L_2(h)dh = 1$ *and* $\int_{\mathbb{R}} |h|^{\nu_2} L_2(h)dh < \infty$;

(ii) *For any* $p \geq 2$ *and* $j = 0,1$; *there exists a continuous bounded function* $l_p(\cdot)$ *in the neighborhood of* x *such that* $\mathbb{E}\left[\left(L_2^{(j)}\left(b_k^{-1}(z - Z_k)\right)\right)^p / X_k\right] < l_p(x) < \infty$.

(**A.5**) On the bandwidths: a_k and b_k satisfy the following conditions:

(i) $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0$ *and* $\lim_{n\to\infty} n^r b_n = \infty$ *for any* $r > 0$;

(ii) $\lim_{n\to\infty} \frac{\log n}{b_n \chi_n(x, a_n)} = 0$ *and* $\lim_{n\to\infty} \frac{\log n}{\chi_n(x, a_n)} = 0$;

(iii) $\chi_n(x, a_n) \to \infty$ *such that* $\chi_n(x, a_n) = n\psi_n(x, a_n) = \sum_{k=1}^{n} \phi_k(x, a_k)$ *with* $\psi_n(x, a_n) = \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_1)\right)\right]$.

(**A.6**)

(i) $(W_n)_{n\geq 1}$ *and the sequence* $(Z_n, X_n)_{n\geq 1}$ *are independent and* $\mathbb{I}_{[Z_k \leq W_k]}\Phi(Y_k) = \mathbb{I}_{[Z_k \leq W_k]}\Phi(Z_k)$;

(ii) $T_{F^Z} \leq T_{G^W}$ *with* $\overline{G}^W(T_{F^Z}) > 0$.

Comments on the assumptions The above conditions are not included in vain; they are important to prove the theorems below and are arranged along the same framework of the complete recursive estimation. From a theoretical point of view, the ergodicity context in the functional framework requires us to set the unchanged assumption (A.1) to guarantee the flexibility of our studied model (see Laïb and Louani (2011)[26] or Benziadi et al. (2016)[4]) and to show that the distribution of the process X is expressed in terms of small ball probabilities.

The condition (A.2) is formally correct and implies that the function $\xi^x(z)$, with respect to the random elements x and z respectively, satisfies the Lipschitz condition. In most cases, the use of such a condition marks an important advantage to prove the almost sure consistency of the estimator.

Usually, in the non censoring recursive kernel estimation, the regularity-type hypotheses (A.3) and (A.4) on the kernels L_1 and L_2 are used extensively to specify the convergence rates of the estimate. These two conditions are essentially used for the boundedness and they often follow the condition (A.2). The smoothness condition

(A.5) concerning the bandwidths a_n and b_n is stated in order to balance bias and variance terms and it characterizes the functional space of our model.

Let us finally point out, as it is known in this type of variables, that a strong additional condition which plays a crucial role to derive the results given for censored variables is stated in (A.6)(i) (see for instance Kohler et al. (2002)[24] for the non-functional case). It is considered as a key hypothesis and the primary engine for classical survival analysis models. Nearly all usual studies that take censorship into consideration assume almost always the independence of (X, Z) and the censorship variable W so as to manage the problem of the non-identifiability of the model that can appear (see also Guessoum and O. Saïd (2008)[18]). Otherwise, if this assumption is not verified, biased results will be obtained. In other hand, the fact that the observations are incomplete leads to a loss of information. Thus, it is not clear if we could deduce the law of Z and W by knowing the law of the elementary data. So as a solution to this problem, the independence in (A.6)(ii) of the latter random variables between them prove that the law of Z is identifiable from the couple (Y, Δ).

3.4 Some asymptotic results

3.4.1 Pointwise convergence with rates

To reduce the complication in analyzing such data, we assume that W is independent of the process Γ and we start by investigating first the pointwise almost sure convergence with rates of the conditional density; this will be useful to derive the asymptotic behavior of the conditional mode function estimate $\widehat{\Theta}_n(x)$ for censored data as stated in the following

Proposition 3.4.1. *Suppose that the assumptions (A.1)-(A.6) hold true. For n large enough, if* $\chi_n(x, a_n) \longrightarrow \infty$ *as* $n \to \infty$*, then, we have*

$$\sup_{z\in\mathcal{I}} \left| \widehat{\xi}_n^x(z) - \xi^x(z) \right| \overset{a.s.}{=} O\left\{ (a_n^{\nu_1} + b_n^{\nu_2}) + \left(\frac{\log n}{b_n \chi_n(x, a_n)} \right)^{1/2} \right\}.$$

Theorem 3.4.1. *Maintaining the same assumptions used in Proposition 3.4.1, we have*

$$\left|\widehat{\Theta}_n(x) - \Theta(x)\right| \overset{a.s.}{=} O\left\{ \left(a_n^{\nu_1/2} + b_n^{\nu_2/2}\right) + \left(\frac{\log n}{b_n \chi_n(x, a_n)}\right)^{1/4} \right\}.$$

Remark 3.4.1. *If censoring does not occur i.e* $\mathbb{P}(W = +\infty) = 1$*, our results can be seen as extensions of what already exists in the literature*

1. Classical estimation

- *Keeping in mind the functional data and by introducing another topological structure upon Kolmogorov's* ϵ*-entropy, the uniform almost complete convergence with rates of the conditional mode estimator is proved by Ferraty et al. (2010)[17] and presented for some integer* $j > 1$ *as*

$$\sup_{x \in \mathcal{S}_{\mathcal{F}}} |\widehat{\Theta}(x) - \Theta(x)|^j = O(h_K^{b_1}) + O(h_H^{b_2}) + O_{a.co.} \left\{ \sqrt{\frac{\psi_{\mathcal{S}_{\mathcal{F}}}\left(\frac{\log n}{n}\right)}{n^{1-\gamma}\phi(h_K)}} \right\}.$$

- *In the stationary and ergodic setting, for a real continuous time process* $(X_t, Y_t)_{t \in \mathbb{R}^+}$*, Didi and Louani (2014)[12] establish consistency results of the kernel estimator of the conditional mode using an extension of Bernstein inequality to the case of ergodic variables. The rate of convergence is given by the following*

$$\sup_{x \in \mathcal{C}} |\Theta_T(x) - \Theta(x)| \overset{a.s.}{=} O(a_T^{\beta/2} + a_T^{\nu/2}) + O\left\{ \left(\frac{\log T}{T a_T^d}\right)^{1/4} \right\} + O\left\{ \left(\frac{\log T}{T a_T^{d+1}}\right)^{1/4} \right\}$$

2. Recursive estimation
This is the case realized in Ardjoun et al's paper (2016)[1] in order to demonstrate that the functional modal regression defined in (3.2) is consistent under ergodic dependence. Roughly the same conditions are used in this complete data case to obtain rates identical to that of Theorem 3.4.1.

3.4.2 Uniform convergence with rates

Under the previous conditions, we present in this subsection the almost sure convergence with rates of the resulting model uniformly on a fixed compact sets of $\mathcal{H} \times \mathbb{R}$ dealing with stationary ergodic random variables. Complementary assumptions are therefore needed which imply the uniform uniqueness of the conditional mode as mentioned in Ould Saïd and Cai (2005)[30].

(A.7) For any $\epsilon > 0$ and any function $r(x)$, there exists a $\varsigma > 0$ such that

$$\sup_{x\in\mathcal{S}} |\Theta(x) - r(x)| \geq \epsilon \Rightarrow \sup_{x\in\mathcal{S}} |\xi^x(\Theta(x)) - \xi^x(r(x))| \geq \varsigma.$$

(A.8) $\inf_{x\in\mathcal{S}} \gamma_X(x) > 0$.

Note that the other conditions used for the uniform consistency are the same as for the pointwise case. As a preliminary result, we have

Proposition 3.4.2. *Assume that the assumptions (A.1)-(A.6) are verified, then*

$$\sup_{x\in\mathcal{S}} \sup_{z\in\mathcal{I}} \left|\widehat{\xi}_n^x(z) - \xi^x(z)\right| \overset{a.s.}{=} O\left\{(a_n^{\nu_1} + b_n^{\nu_2}) + \left(\frac{\log n}{b_n \chi_n(x, a_n)}\right)^{1/2}\right\}, \ as\, n \to \infty.$$

Then, the uniform almost sure convergence without/ with rates of $\widehat{\Theta}_n(x)$ is provided in the following theorem

Theorem 3.4.2. *Under the conditions of Proposition 3.4.2 and if (A.7), (A.8) hold, we have for all fixed x of $\mathcal{S}$*

(i)

$$\sup_{x\in\mathcal{S}} |\widehat{\Theta}_n(x) - \Theta(x)| \to 0 \ \textit{almost surely as } n \to \infty.$$

(ii)

$$\sup_{x\in\mathcal{S}} \left|\widehat{\Theta}_n(x) - \Theta(x)\right| \overset{a.s.}{=} O\left\{\left(a_n^{\nu_1/2} + b_n^{\nu_2/2}\right) + \left(\frac{\log n}{b_n \chi_n(x, a_n)}\right)^{1/4}\right\}.$$

Remark 3.4.2. *To remind the reader and to provide a clearer picture of our study, we present the theoretical results linked with the un-functional case (when $\mathcal{H} = \mathbb{R}$ and more generally $\mathbb{R}^p$)*

1. Classical case
Using some different assumptions, Khardani et al. (2010)[20] establish the uniform almost sure consistency of the classical kernel estimator of the conditional mode function under the i.i.d. random model. What differs from the studied functional case in this paper is that we do not need to impose hypotheses on the probability of small balls and we summarize the condition (A.1) for p ≥ 1 to $\phi_x(h) = h^p f(x) + O(h^p)$ as $h \to 0$ and $\mathbb{P}[d_{\mathcal{H}}(x, X_k) \leq h/\wp_{k-1}] = f_k^{\wp_{k-1}}(x)h^p + O(h^p)$, such that
$$\sup_{x\in I}\left|\widehat{\Theta}_n(x) - \Theta(x)\right| \overset{a.s.}{=} O\left\{\max\left(\left(\frac{\log n}{nh_n^{p+1}}\right)^{1/4}, h_n\right)\right\}, \text{ as } n \to \infty.$$

2. Recursive case
$\widehat{\xi}_n^x(z)$ reduces to the recursive kernel estimator of the conditional density previously studied by Khardani and Semmar (2014)[22] for the i.i.d. observations. Noting that they achieved under the independence condition between $(W_k)_k$ and $(X_k, Z_k)_k$ the rate of uniform strong consistency, using the same bandwidth parameter h_k for both kernels, and thinner arguments for simplification. These results were extended to the strong mixing data case by the same authors.

3.5 Proofs of the theoretical results

Right censoring being one novelty change related to the estimator, its influence appears in the first part of these proofs and for the sake of shortness we set the following notation throughout the rest of this section

$$\widetilde{Z}_k = \frac{\Delta_k L_2\left(b_k^{-1}(z - Y_k)\right)}{\overline{G}^W(Y_k)}.$$

It is worth being noted that the tools used to demonstrate the strong consistency of this estimate are analogous to those relating to the recursive estimate without the term of censorship. For that, we start by introducing the following decomposition so that our elements are well defined

$$\widehat{\xi}_n^x(z) - \xi^x(z) = \widehat{\xi}_n^x(z) - \widetilde{\xi}_n^x(z) + \widetilde{\xi}_n^x(z) - \xi^x(z),$$

and as is common in most of the reference papers in terms of censored processes, we have to set a "pseudo-estimator" of $\xi^x(z)$ for a continuous function $G^W(.)$ presented as follow

$$\widetilde{\xi}_n^x(z) := \frac{\widetilde{\varphi}_n(x,z)}{\widetilde{\gamma}_n(x)} = \frac{\sum_{k=1}^{n} \frac{\widetilde{Z}_k}{b_k} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)}{\sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)}, \tag{3.5}$$

with

$$\widetilde{\varphi}_n(x,z) := (\chi_n(x,a_k))^{-1} \sum_{k=1}^{n} \frac{\widetilde{Z}_k}{b_k} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)$$

and

$$\widetilde{\gamma}_n(x) := (\chi_n(x,a_k))^{-1} \sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)$$

respectively.

In this proof, we will clearly rely on the following decomposition

$$\widetilde{\xi}_n^x(z) - \xi^x(z) - B_n(x,z) = \frac{1}{\widetilde{\gamma}_n(x)} \left\{Q_n(x,z) - B_n(x,z)[\widetilde{\gamma}_n(x) - \ddot{\gamma}_n(x)]\right\} \tag{3.6}$$

where the two main terms are expressed for all couple $(x,z) \in \mathcal{S} \times \mathcal{I}$ by

$$Q_n(x,z) = [\widetilde{\varphi}_n(x,z) - \ddot{\varphi}_n(x,z)] - \xi^x(z)[\widetilde{\gamma}_n(x) - \ddot{\gamma}_n(x)], \tag{3.7}$$

$$B_n(x,z) = \frac{\ddot{\varphi}_n(x,z)}{\ddot{\gamma}_n(x)} - \xi^x(z). \tag{3.8}$$

3.5.1 Proof of Theorem 3.4.1

Obviously, the proof of Theorem 3.4.1 is a consequence of Proposition 3.4.1 and that is through

$$\sum_{n>1} \mathbb{P}\left[\left|\widehat{\Theta}_n(x) - \Theta(x)\right| > \epsilon\right] \le \sum_{n>1} \mathbb{P}\left[\sup_{z\in\mathcal{I}}\left|\widehat{\xi}_n^x(z) - \xi^x(z)\right| > \beta\right],$$

which is based primarily on the decomposition (3.6), valid for any $z \in \mathcal{I}$

$$\begin{aligned}\sup_{z\in\mathcal{I}} |\widetilde{\xi}_n^x(z) - \xi^x(z)| \quad &\le \sup_{z\in\mathcal{I}} |B_n(x,z)| + \frac{1}{\widetilde{\gamma}_n(x)}\Bigg\{\sup_{z\in\mathcal{I}} |\widetilde{\varphi}_n(x,z) - \ddot{\varphi}_n(x,z)| \\ &+ \left(\mu^{-1}\lambda + \sup_{z\in\mathcal{I}} |B_n(x,z)|\right) |\widetilde{\gamma}_n(x) - \ddot{\gamma}_n(x)|\Bigg\}.\end{aligned}$$

According to that, other functions must be defined

$$\ddot{\varphi}_n(x,z) := (\chi_n(x,a_k))^{-1} \sum_{k=1}^{n} \mathbb{E}\left[\frac{\widetilde{Z}_k}{b_k} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) / \wp_{k-1}\right] \tag{3.9}$$

and

$$\ddot{\gamma}_n(x) := (\chi_n(x,a_k))^{-1} \sum_{k=1}^{n} \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) / \wp_{k-1}\right]. \tag{3.10}$$

Then, the proof of Theorem 3.4.1 is a direct result of Lemmas 3.5.1-3.5.3 below extending Ardjoun et al's (2016)[1] results to the censored setting

Lemma 3.5.1. *Under Assumptions (A.1)-(A.3), (A.5) and (A.6), for $n \to \infty$ we have*

$$\sup_{z\in\mathcal{I}} |B_n(x,z)| = O\left\{a_n^{\nu_1} + b_n^{\nu_2}\right\}.$$

Lemma 3.5.2. *Under Assumptions (A.1) and (A.3)-(A.6), for $n \to \infty$ we have*

$$\sup_{z\in\mathcal{I}} |\widetilde{\varphi}_n(x,z) - \ddot{\varphi}_n(x,z)| = O\left\{\left(\frac{\log n}{b_n \chi_n(x,a_n)}\right)^{1/2}\right\} \quad a.s.$$

Lemma 3.5.3. *Under Assumptions (A.1) and (A.3), for $n \to \infty$ we have*

$$\widetilde{\gamma}_n(x) - 1 = O\left\{\left(\frac{\log n}{\chi_n(x, a_n)}\right)^{1/2}\right\} \quad a.s.$$

3.5.2 Proof of Theorem 3.4.2

(i) The key argument in this proof comes from the definition (3.3) of the conditional mode function $\widehat{\Theta}_n(x)$. In addition, by using the inequality expressed for n large enough as follows

$$\sup_{x\in\mathcal{S}}\left|\xi^x(\widehat{\Theta}_n(x)) - \xi^x(\Theta(x))\right| \leq 2\sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}\left|\xi^x(z) - \widehat{\xi}_n^x(z)\right| \tag{3.11}$$

which details given in the article by Didi and Louani (2014)[12]. Next, by the fact that $\sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}\left|\xi^x(z) - \widehat{\xi}_n^x(z)\right|$ converges almost surely to 0 as $n \to \infty$ (we can go back to Ezzahrioui and O.Saïd (2010)[14] or others) the proof of this part has therefore been verified.

Fortunately, to prove (ii) of Theorem 3.4.2, it is enough to study the uniform almost sure convergence of $\widehat{\xi}_n^x(z)$ shown in Proposition 3.4.2, from which we deduce the asymptotic results available for the estimator $\widehat{\Theta}_n(x)$, such that

$$\sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}\left|\widehat{\xi}_n^x(z) - \xi^x(z)\right| = \sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}\left|\widehat{\xi}_n^x(z) - \widetilde{\xi}_n^x(z)\right| + \sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}\left|\widetilde{\xi}_n^x(z) - \xi^x(z)\right|$$

Then, for the second term of the equality, we use the boundedness of the joint density φ described above and the fact that $\gamma(x) > \mu$ to present in a simpler form the decomposition below for any $(x, z) \in \mathcal{S} \times \mathcal{I}$

$$\begin{aligned}\sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}|\widetilde{\xi}_n^x(z) - \xi^x(z)| \quad &\leq \sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}|B_n(x,z)| + \frac{1}{\inf_{x\in\mathcal{S}}\widetilde{\gamma}_n(x)}\left\{\sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}|\widetilde{\varphi}_n(x,z) - \ddot{\varphi}_n(x,z)|\right. \\ &\left. + \left(\mu^{-1}\lambda + \sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}|B_n(x,z)|\right)\sup_{x\in\mathcal{S}}|\widetilde{\gamma}_n(x) - \ddot{\gamma}_n(x)|\right\}.\end{aligned}$$

To accomplish the proof of Theorem 3.4.2, we put the following auxiliary lemmas which correspond to the uniform versions of Lemmas 3.5.1-3.5.3

Lemma 3.5.4. *Under the same assumptions as those of Lemma 3.5.1, we have*

$$\sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}|B_n(x,z)| = O\left\{a_n^{\nu_1}+b_n^{\nu_2}\right\}.$$

Lemma 3.5.5. *Assume that (A.1)-(A.6) and (A.7) hold true, for any $x \in S$, we have*

$$\sum_{n\geq 1}\mathbb{P}\left\{\sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}|\widetilde{\varphi}_n(x,z)-\ddot{\varphi}_n(x,z)|>\eta\right\}<\infty \quad a.s.$$

Lemma 3.5.6. *Assume that (A.1) and (A.3)-(A.4) hold, for any $x \in \mathcal{S}$, we have*

$$\sum_{n\geq 1}\mathbb{P}\left\{\sup_{x\in\mathcal{S}}|\widetilde{\gamma}_n(x)-\ddot{\gamma}_n(x)|>\varsigma\right\}<\infty \quad a.s.$$

Lemma 3.5.7. *Under Assumptions (A.1)-(A.7), then for n large enough, we are able to get*

$$\mathbb{P}\left\{\sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}\left|\widehat{\xi}_n^x(z)-\widetilde{\xi}_n^x(z)\right| = O_{a.s.}\left(\left(\frac{\log_2 n}{n}\right)^{1/2}\right)\right\}=1.$$

3.6 Applications on simulated data

A simple computational study is realized in this short section in order to assess and compare the performance of the new proposed recursive kernel conditional mode function estimator $\widehat{\Theta}_n(x)$ when the data are right-censored and ergodic as defined in (3.3) with that given in (3.2). However, compared to the classical estimations, the effectiveness of recursive methods on simulated data is already confirmed in a number of recent numerical studies (Khardani and Semmar (2014)[22], Ardjoun et al. (2016)[1], Benziadi et al. (2016)[4] and others) comparing several conditional models.

We implement here two separate algorithms: as a first step, let us consider for $k = 1, \ldots, n$ the classical regression model $Y_k = R(X_k) + \epsilon_k$ in absence of censored data and using the $\mathcal{R}$-routine simul.far to generate an n-sample of functional autoregressive covariate $\{(X_k(t)),\ t \in [0,1],\ k = 1, \ldots, n\}$ (this specimen is used for the two methods). Figure 3.1 shows the curves $X_k(t)$, the time is discretized into 50 points.

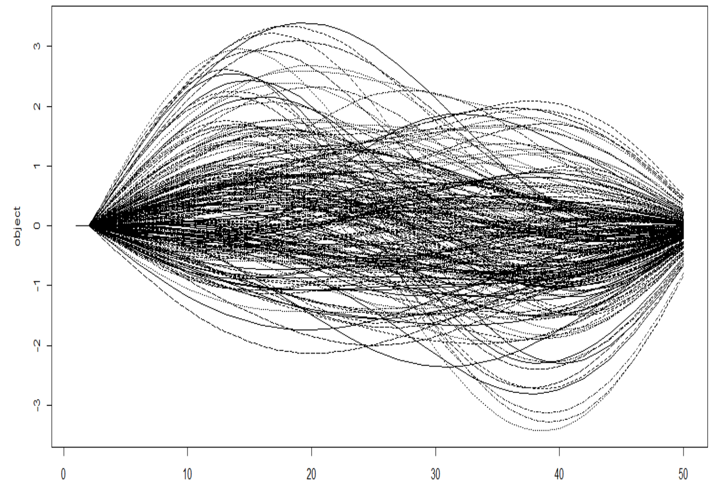

Figure 3.1: A sample of simulated curves of $\{X_k(t), t \in [0,1]\}$.

On the other hand, we try to test the effect of the (randomly-right) censored mechanism $(X_k, Y_k, \Delta_k)_{k\geq 1}$ on the consistency of our estimator when the data are simultaneously functional and ergodic, through variable parameters that include both various observed sample sizes n=200, 400 and 600 and censoring percentages CR=0.2, 0.4 and 0.6. For $k = 1, \ldots, n$, we also need to simulate the censoring variables $(W_k)_k$ through an exponential distribution of parameter $\lambda = 0.5$, which values are chosen according to the desired percentages of censored observations. Moreover, we simulate the censored regression model $Z_k = R(X_k) + \epsilon_k$ in order to deduce $Y_k = \min(Z_k, W_k)$, the censoring indicator $\Delta_k = \mathbb{I}_{[Z_k \leq W_k]}$ and the Kaplan-Meier estimator of the survival function which curve is plotted in Figure 3.2. For both simulations, we take the same operator $R(X_k) = \displaystyle\int_0^1 \exp(X_k(t))dt$ with the errors ϵ_k which are independent of X_k and are generated according to a Gaussian distribution $\mathcal{N}(0,1)$, without forgetting to highlight that our processes verify the required condition (ergodicity). Throughout the experiences, we choose to use the standard quadratic kernels and the semi-metric "deriv" in $\mathcal{H}$. Our original data of sizes n are divided into two subsets: a training

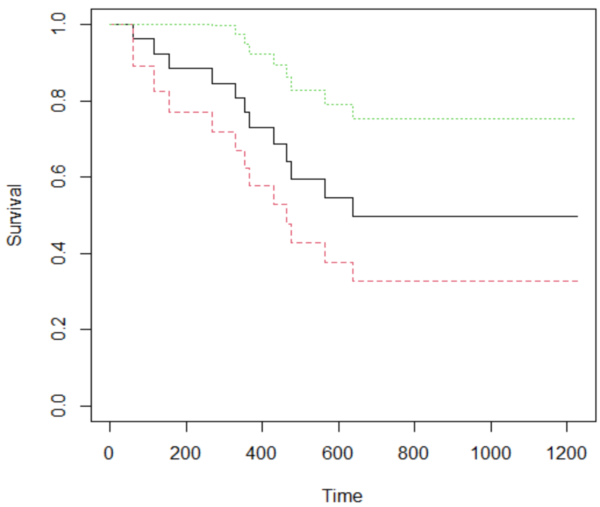

Figure 3.2: Kaplan-Meier curve.

sample I to build the functional recursive estimator of $\Theta(x)$ for both complete and censored cases and a testing sample J containing the last 50 observations from the total sample size n. Thence, the accuracy of the both estimators of the conditional mode is compared according to the MSE (Mean Squared Error) criterion:

The case of complete data

$$MSE(COMPLETE) = \frac{1}{n}\sum_{k=1}^{n}\left(\overline{\Theta}_n(X_k) - \Theta(X_k)\right)^2$$

The case of censored data

$$MSE(CENSORED) = \frac{1}{n}\sum_{k=1}^{n}\left(\widehat{\Theta}_n(X_k) - \Theta(X_k)\right)^2$$

The box-plot of the mean squared errors obtained by both cases considering various CR is then illustrated in Figure 3.3. Whereas, the obtained results under the randomly right-censorship for various sample sizes n are summarized in Table 3.1 below

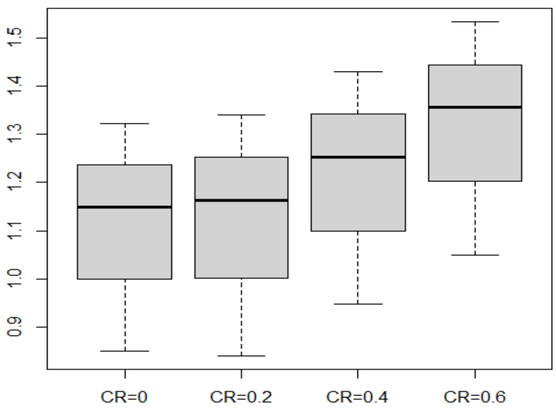

Figure 3.3: MSE comparison for different censoring rates considered: CR=0 (non censoring), CR=0.2, 0.4 and 0.6.

	MSE(Complete Data)	MSE(Censored Data)		
n	CR=0	CR=0.2	CR=0.4	CR=0.6
$n = 200$	1.3222	1.3402	1.4297	1.5327
$n = 400$	1.1493	1.1625	1.2524	1.3549
$n = 600$	0.8501	0.8410	0.9473	1.0485

Table 3.1: the MSE-Results

Conclusion and comments: Whoever reads the results obtained in Table 3.1 can observe that the new estimator $\widehat{\Theta}$ that includes censored data performs slightly less than $\overline{\Theta}$, when the percentage of censoring rates (CR) increases gradually (from 0.2 to 0.6); the quality of the proposed estimator is affected by the prevalence of censoring. There is an obvious influence of this factor on the performance of the estimate and this can also be shown through the graph in Figure 3.3. Moreover, this comparison shows that even in the presence of this data, when the sample size n gets larger (from 200 to 600), the decrease in MSE is noticeable and thus, as expected, the prediction accuracy of the conditional mode increases. This is what we call the sample size effect.

3.7 Appendix

Some necessary technical lemmas which have been used in the proofs of the asymptotic results (pointwise and uniform convergence) are mentioned in this section. To control the estimation based on functional ergodic data, we should use the following version of exponential inequality which is quoted in Lemma 3.7.1 (see Laib and Louani (2011)[26]) for partial sums of unbounded martingale differences.

Lemma 3.7.1. *Let $(Z_n)_{n\geq 1}$ be a sequence of real martingale differences with respect to the sequence of $\sigma-$ fields $\wp_n = \sigma(Z_1, Z_2, \ldots, Z_n)_{n\geq 1}$ generated by the random variables $Z_1, Z_2, \ldots, Z_n$. Set $S_n = \sum_{k=1}^{n} Z_k$. For any $p \geq 2$ and for any $n \geq 1$, assume that there exist some nonnegative constants C and d_n such that*

$$\mathbb{E}(Z_n^p/\wp_{n-1}) \leq C^{p-2} p! d_n^2 \quad \textit{almost surely.} \tag{3.12}$$

Then, for any $\epsilon > 0$, we have

$$\mathbb{P}(|S_n| > \epsilon) \leq 2\exp\left(-\frac{\epsilon^2}{2(D_n + C\epsilon)}\right) \textit{ where } D_n = \sum_{k=1}^{n} d_k^2.$$

Lemma 3.7.2. *Assume that conditions (A.1), (A.3) and (A.5)(ii)-(iii) are verified. Then, for any $x \in \mathcal{H}$, we have*

(i) $\widetilde{\gamma}_n(x) - \ddot{\gamma}_n(x) = O_{a.s.}\left\{\left(\frac{\log n}{(\chi_n(x, a_n))}\right)^{1/2}\right\}$,

(ii) $\ddot{\gamma}_n(x) \to 1$ *a.s. as* $n \to \infty$,

(iii) $\exists \delta > 0, \sum_{n\geq 1} \mathbb{P}\{\widetilde{\gamma}_n(x) < \delta\} < \infty.$

The proof of these insertions may be found in Laib and Louani (2010)[25].

Proof of Lemma 3.5.1

We first remark that for fixed (x, z), n, and from the expression (3.8), the bias term can be expressed as

$$B_n(x,z) = \frac{\ddot{\varphi}_n(x,z)}{\ddot{\gamma}_n(x)} - \xi^x(z) = \frac{\widetilde{B}_n(x,z)}{\ddot{\gamma}_n(x)}. \tag{3.13}$$

Using for all $1 \leq k \leq n$, the condition (A.6) and a double conditioning successively with respect to the δ-field $\mathcal{B}_{k-1}$ and Z_k. Thus, under the ergodicity of the random variables and the conditional expectation properties, we can write the following

$$\begin{aligned}
\mathbb{E}[\widetilde{Z}_k/\wp_{k-1}] &= \mathbb{E}\left[\mathbb{E}\left(\frac{\Delta_k L_2\left(b_k^{-1}(z-Y_k)\right)}{\overline{G}^W(Y_k)}/\mathcal{B}_{k-1}, Z_k\right)/\wp_{k-1}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left(\frac{\Delta_k L_2\left(b_k^{-1}(z-Z_k)\right)}{\overline{G}^W(Z_k)}/X_k, Z_k\right)/\wp_{k-1}\right] \\
&= \mathbb{E}\left[\frac{L_2\left(b_k^{-1}(z-Z_k)\right)}{\overline{G}^W(Z_k)}\mathbb{E}\left[\mathbb{I}_{(Z_k\leq W_k)}/X_k, Z_k\right]/\wp_{k-1}\right] \\
&= \mathbb{E}\left[L_2\left(b_k^{-1}(z-Z_k)\right)/\wp_{k-1}\right] \qquad (3.14)
\end{aligned}$$

which leads to

$$\begin{aligned}
|\ddot{\varphi}_n(x,z) - \xi^x(z)\ddot{\gamma}_n(x)| &\leq \frac{1}{\chi_n(x,a_k)}\sum_{k=1}^{n}\mathbb{E}\Big[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right) \\
&= \left|\mathbb{E}\left(b_k^{-1}L_2\left(b_k^{-1}(z-Z_k)\right)/X_k\right) - \xi^x(z)\right|/\wp_{k-1}\Big]
\end{aligned}$$

by a simple change of variable $h = \left(b_k^{-1}(z-u)\right)$, we obtain

$$\begin{aligned}
b_k^{-1}\mathbb{E}\left[L_2\left(b_k^{-1}(z-Z_k)\right)/X_k\right] &= b_k^{-1}\int_{\mathbb{R}} L_2(b_k^{-1}(z-u))\xi^{(X_k)}(u)du \\
&= \int_{\mathbb{R}} L_2(h)\xi^{(X_k)}(z-b_kh)dh
\end{aligned}$$

Meanwhile, it is clear that the condition (A.2) combined with the boundedness of the kernel L_2 allow us to have

$$\begin{aligned}\sup_{z\in\mathcal{I}}\left|\mathbb{E}\left[b_k^{-1}L_2\left(b_k^{-1}(z-Z_k)\right)/X_k\right]-\xi^x(z)\right| &= \sup_{z\in\mathcal{I}}\left|\int_{\mathbb{R}}L_2(h)\left[\xi^{(X_k)}(z-b_kh)-\xi^x(z)\right]dh\right|\\ &\leq C_1\int_{\mathbb{R}}L_2(h)\left(a_k^{\nu_1}+b_k^{\nu_2}|h|^{\nu_2}\right)dh.\end{aligned}$$

The assumptions (A.1)(iii) and (A.4) (i) permit to show that

$$\sup_{z\in\mathcal{I}}\left|\widetilde{B}_n(x,z)\right|\leq C_1\left(a_n^{\nu_1}+b_n^{\nu_2}\right)$$

from which we conclude the proof of Lemma 3.5.1.

Proof of Lemma 3.5.2

For all $x\in\mathcal{H}$, observe that we can write

$$\begin{aligned}\widetilde{\varphi}_n(x,z)-\ddot{\varphi}_n(x,z) &= \left(\chi_n(x,a_k)\right)^{-1}\sum_{k=1}^{n}\left\{\frac{\widetilde{Z}_k}{b_k}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)\right.\\ &\left.-\mathbb{E}\left[\frac{\widetilde{Z}_k}{b_k}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)/\wp_{k-1}\right]\right\}\\ &= \left(\chi_n(x,a_k)\right)^{-1}\sum_{k=1}^{n}U_{k,n}(x,z)\end{aligned}$$

where $\{U_{k,n}(x,z)\}_{k\geq1}$ forms a triangular array of martingale differences sequence with respect to the σ-fields $(\wp_k)_{1\leq k\leq n}$ and is defined for all $x\in\mathcal{H}$ by

$$\frac{\widetilde{Z}_k}{b_k}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)-\mathbb{E}\left[\frac{\widetilde{Z}_k}{b_k}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)/\wp_{k-1}\right]$$

with $S_n=\sum_{k=1}^{n}U_{k,n}(x,z)$. It is also essential to examine whether the condition of Lemma 3.7.1 is verified. By applying the Minkowski and Jensen inequalities for

$U_{k,n}(x,z)$ with $p \geq 2$, we can easily get

$$\begin{aligned}\left|\mathbb{E}\left[U^p_{k,n}(x,z)/\wp_{k-1}\right]\right| &= \left|\mathbb{E}\left[\left(\frac{\widetilde{Z}_k}{b_k}L_1(a_k^{-1}d_{\mathcal{H}}(x,X_k)) - \mathbb{E}\left[\frac{\widetilde{Z}_k}{b_k}L_1(a_k^{-1}d_{\mathcal{H}}(x,X_k))/\wp_{k-1}\right]\right)^p/\wp_{k-1}\right]\right| \\ &\leq \left(\mathbb{E}\left[\left|\frac{\widetilde{Z}_k}{b_k}L_1(a_k^{-1}d_{\mathcal{H}}(x,X_k))\right|^p/\wp_{k-1}\right]^{1/p}\right. \\ &\left.+\mathbb{E}\left[\mathbb{E}\left[\left|\frac{\widetilde{Z}_k}{b_k}L_1(a_k^{-1}d_{\mathcal{H}}(x,X_k))\right|^p/\wp_{k-1}\right]/\wp_{k-1}\right]^{1/p}\right)^p \\ &= 2^p\mathbb{E}\left[\left|\frac{\widetilde{Z}_k}{b_k}L_1(a_k^{-1}d_{\mathcal{H}}(x,X_k))\right|^p/\wp_{k-1}\right]\end{aligned}$$

by a simple manipulation and under the assumption (A.3), we have

$$\begin{aligned}\mathbb{E}\left[\left|\frac{\widetilde{Z}_k}{b_k}L_1(a_k^{-1}d_{\mathcal{H}}(x,X_k))\right|^p/\wp_{k-1}\right] &= \mathbb{E}\left[\mathbb{E}\left[\left|\frac{\widetilde{Z}_k}{b_k}L_1(a_k^{-1}d_{\mathcal{H}}(x,X_k))\right|^p/\mathcal{B}_{k-1},Z_k\right]/\wp_{k-1}\right] \\ &= \mathbb{E}\left[b_k^{-p}L_1^p(a_k^{-1}d_{\mathcal{H}}(x,X_k))L_2^p(b_k^{-1}(z-Z_k))/\wp_{k-1}\right]\end{aligned}$$

Next, by using the conditions (A.1)(ii) and (A.4)(ii), we deal with

$$\left|\mathbb{E}\left[U^p_{k,n}(x,z)/\wp_{k-1}\right]\right| \leq 2^pCb_n^{-p}\phi_k(x,a_n) \tag{3.15}$$

with $D_n = \sum_{k=1}^{n} d_k^2 = b_n^{-p}\sum_{k=1}^{n}\phi_k(x,a_n)$. From now on, replacing the difference $\widetilde{\varphi}_n(x,z) - \ddot{\varphi}_n(x,z)$ in (3.7) by $H_{k,n}(x,z)$ to avoid repetition of expressions

$$\begin{aligned}H_{k,n}(x,z) &= [\widetilde{\varphi}_n(x,z) - \widetilde{\varphi}_n(x,z_j)] + [\widetilde{\varphi}_n(x,z_j) - \ddot{\varphi}_n(x,z_j)] \\ &- [\ddot{\varphi}_n(x,z) - \ddot{\varphi}_n(x,z_j)] \\ &:= \widetilde{H}_{k,n}(x,z) + H_{k,n}(x,z_j).\end{aligned}$$

We use the compactness property of $\mathcal{I}$ by writing for any $z_1, z_2, \ldots, z_{r_n}$, $\mathcal{I} \subset \bigcup_{j=1}^{r_n} G_j$ where $G_j = (z_j - s_n, z_j + s_n)$. And since $\mathcal{I}$ is bounded we can take for a constant M, $s_n \leq Mn^{-\gamma}$ with $(\gamma = (3/2)\beta + 1/2)$ for $\beta > 0$ and $j(z) = \arg\min_{j\in\{1,\ldots,r_n\}}|z - z_j|$.

Thereafter, we have

$$\begin{aligned}\sup_{z\in\mathcal{I}}|H_{k,n}(x,z)| &\leq \max_{j\in\{1,\dots,r_n\}}\sup_{z\in\mathcal{I}}\left|\widetilde{H}_{k,n}(x,z)\right| + \max_{j\in\{1,\dots,r_n\}}|H_{k,n}(x,z_j)| \\ &= A_{1n}+A_{2n}. \qquad (3.16)\end{aligned}$$

Firstly, the term A_{1n} can be expressed as follow

$$\begin{aligned}A_{1n} &\leq \max_{j\in\{1,\dots,r_n\}}\sup_{z\in\mathcal{I}}|\widetilde{\varphi}_n(x,z)-\widetilde{\varphi}_n(x,z_j)| + \max_{j\in\{1,\dots,r_n\}}\sup_{z\in\mathcal{I}}|\ddot{\varphi}_n(x,z)-\ddot{\varphi}_n(x,z_j)| \\ &= \Sigma_1+\Sigma_2.\end{aligned}$$

Obviously, this latter can be deduced directly from the following two main statements.

- Σ_1 *tends to* 0 *almost surely as* $n\to\infty$;
- $\Sigma_2 = O_{a.s.}(1)$.

Concerning Σ_1, let us consider $\widetilde{Z}_{k,j} = \dfrac{\Delta_k L_2\left(b_k^{-1}(z_j-Y_k)\right)}{\overline{G}^W(Y_k)}$. Because L_1 is continuous and bounded, then by the condition (A.4)(i), we can write for almost all z and a fixed number C_2 the following

$$\begin{aligned}\Sigma_1 &\leq \max_{j\in\{1,\dots,r_n\}}\sup_{z\in\mathcal{I}}\frac{1}{\chi_n(x,a_k)}\sum_{k=1}^{n}\frac{L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)}{b_k}\left|\widetilde{Z}_k-\widetilde{Z}_{k,j}\right| \\ &\leq \max_{j\in\{1,\dots,r_n\}}\sup_{z\in\mathcal{I}}C_2\frac{|z-z_j|}{\chi_n(x,a_k)}\sum_{k=1}^{n}\frac{\Delta_k}{b_k^2\overline{G}^W(Z_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right) \\ &\leq (\chi_n(x,a_k))^{-1}M\sum_{k=1}^{n}s_n b_k^{-2}\end{aligned}$$

it follows from (3.12) and (A.5)(i) that

$$\Sigma_1 \to 0 \quad a.s. \quad as\ n\to\infty. \qquad (3.17)$$

Similarly, for Σ_2, using the same arguments to the foregoing such that

$$\Sigma_2 \leq \max_{j\in\{1,\dots,r_n\}}\sup_{z\in\mathcal{I}}\frac{1}{\chi_n(x,a_k)}\sum_{k=1}^{n}\mathbb{E}\left[\frac{L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)}{b_k}\left|\widetilde{Z}_k-\widetilde{Z}_{k,j}\right|/\wp_{k-1}\right]$$

with the fact that

$$\begin{aligned}\mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)/\wp_{k-1}\right] &\leq \mathbb{P}[d_{\mathcal{H}}(x,X_k)\leq a_k/\wp_{k-1}]\\ &\leq \phi_k(x,a_k)\end{aligned}$$

we deduce by the assumption (A.1)(ii) that

$$\Sigma_2 = O_{a.s.}(1) \ as \ n\to\infty. \tag{3.18}$$

Hence, the definition of A_{1n} and the statements (3.17) and (3.18) give directly

$$A_{1n} = O_{a.s.}\left\{\left(\frac{\log n}{b_n\chi_n(x,a_n)}\right)^{1/2}\right\}. \tag{3.19}$$

Next, let's highlight the second term A_{2n} of our decomposition (3.16), one can write

$$\begin{aligned}A_{2n} &= \max_{j\in\{1,\dots,r_n\}}|H_{k,n}(x,z_j)|\\ &\leq (\chi_n(x,a_k))^{-1}\max_{j\in\{1,\dots,r_n\}}\left|\sum_{k=1}^{n}U_{k,n}(x,z_j)\right|.\end{aligned}$$

Making use the principe of the technical Lemma 3.7.1 on $\sum_{k=1}^{n}U_{k,n}(x,z_j)$ with $r_n = O\left(n^{(3/2)\beta+1/2}\right)$, we obtain

$$\mathbb{P}\left\{(\chi_n(x,a_k))^{-1}\max_{j\in\{1,\dots,r_n\}}\left|\sum_{k=1}^{n}U_{k,n}(x,z_j)\right|>\theta\right\}\leq r_n\max_{j\in\{1,\dots,r_n\}}\mathbb{P}\left\{\left|\sum_{k=1}^{n}U_{k,n}(x,z_j)\right|>\theta\left(\chi_n(x,a_k)\right)\right\}$$

We have from (3.15), the following quantity is verified for all $(x,z)\in\mathcal{S}\times\mathcal{I}$ and $j=1,\dots,r_n$, $\left|\mathbb{E}\left[U_{k,n}^p(x,z_j)/\wp_{k-1}\right]\right|\leq 2^pCb_n^{-p}\phi_k(x,a_n)$ with $D_n = O_{a.s.}\left(b_n\chi_n(x,a_n)\right)$ and

$$\mathbb{P}\left\{\left|\sum_{k=1}^{n}U_{k,n}(x,z_j)\right|>\theta_0\left(\chi_n(x,a_k)\right)\left(\frac{\log n}{b_n\chi_n(x,a_k)}\right)^{1/2}\right\}\leq 2\exp\left\{-C\theta_0^2\log n\right\},$$

and therefore for the two positive constants C' and M_1 given, we have

$$\begin{aligned}
\mathbb{P}\left\{\max_{j\in\{1,\ldots,r_n\}}\left|\sum_{k=1}^{n}U_{k,n}(x,z_j)\right| > \theta_0\left(\chi_n(x,a_k)\right)\left(\frac{\log n}{b_n\chi_n(x,a_k)}\right)^{1/2}\right\} &\leq 2C'r_n\exp\left\{-C\theta_0^2\log n\right\}\\
&\leq M_1 r_n n^{-C\theta_0^2}\\
&\leq \frac{M_1}{n^{(C\theta_0^2-(3/2)\beta-1/2)}}
\end{aligned}$$

which implies by applying the Borel-Cantelli Lemma and taking θ_0 sufficiently big that

$$\sum_{n\geq 1}\mathbb{P}\left\{\max_{j\in\{1,\ldots,r_n\}}\left|\sum_{k=1}^{n}U_{k,n}(x,z_j)\right| > \theta\left(\chi_n(x,a_k)\right)\right\} < \infty$$

from which the same result holds

$$A_{2n} = O_{a.s.}\left\{\left(\frac{\log n}{b_n\chi_n(x,a_n)}\right)^{1/2}\right\} = O_{a.s.}(1),\ \ as\ n\to\infty. \tag{3.20}$$

Finally, combining (3.16) together with (3.19) and (3.20) lead to the required result.

$$\sum_{n\geq 1}\mathbb{P}\left\{\sup_{z\in\mathcal{I}}\left|\widetilde{\varphi}_n(x,z)-\ddot{\varphi}_n(x,z)\right| > \theta\right\} < \infty.$$

and the proof is thus valid.

Proof of Lemma 3.5.3

First of all, for $x\in\mathcal{H}$, note that the function $\widetilde{\gamma}_n(x)$ can be decomposed as

$$\begin{aligned}
\widetilde{\gamma}_n(x) &= \left\{(\chi_n(x,a_k))^{-1}\left(\sum_{k=1}^{n}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)-\mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)/\wp_{k-1}\right]\right]\right)\right\}\\
&+ \left\{(\chi_n(x,a_k))^{-1}\sum_{k=1}^{n}\mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)/\wp_{k-1}\right]-1\right\}+1\\
&= B_{1n}(x)+B_{2n}(x)+1.
\end{aligned}$$

Indeed, there is no effect here of the censoring factor on this latter decomposition, the proof is thus standard. We start directly by the first term

$$B_{1n}(x) = \widetilde{\gamma}_n(x) - \ddot{\gamma}_n(x) = (\chi_n(x, a_k))^{-1} \sum_{k=1}^{n} M_{k,n}(x).$$

We draw the readers' attention to the fact that the proof of this term uses a concept similar to that of Lemma 3.5.2 with a slight difference. Let $M_{k,n}(x) = L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) - \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) / \wp_{k-1}\right]$ be a sequence of martingale differences with respect to the σ-fields $(\wp_k)_{1 \leq k \leq n}$ and which satisfies the condition of Lemma 3.7.1 such as

$$\begin{aligned} \left|\mathbb{E}\left[M_{k,n}^p(x)/\wp_{k-1}\right]\right| &\leq 2^p \mathbb{E}\left[L_1^p\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) / \wp_{k-1}\right] \\ &\leq C 2^p \phi_k(x, a_n) \\ &:= d_k^2. \end{aligned}$$

Again, Lemma 3.7.1 with $D_n = \sum_{k=1}^{n} d_k^2$ and the result (i) of Lemma 3.7.2, which proof is given in Laib and Louani (2010)[25] allow us to conclude $B_{1n}(x)$. In view of the second term B_{2n}, we thus can write for any $x \in S$

$$B_{2n}(x) = \ddot{\gamma}_n(x) - 1,$$

a direct application of the insertion (ii) of Lemma 3.7.2 completes the proof and

$$B_{2n}(x) \to 0 \quad as \quad n \to \infty, \ almost\ surely.$$

Proof of Lemma 3.5.4

In this proof, we are satisfied with studying the quantity $\sup_{x \in \mathcal{S}} \sup_{z \in \mathcal{I}} |B_n(x, z)|$ uniformly on x and z. For that, using again the condition (A.6) for any measurable function Φ.

The conditions (A.1) and (A.3) allow us to conclude that

$$\chi_n(x, a_k) \ddot{\gamma}_n(x) \geq C \chi_n(x, a_k).$$

Moreover, we want to stress that despite the presence of censorship, the function $\ddot{\gamma}_n(\cdot)$ is not affected according to its definition (3.10). Furthermore, by taking into consideration the last inequality in establishing the proof of Lemma 3.5.1, in addition to the condition (A.4) (i), we terminate the proof of Lemma 3.5.4.

Proof of Lemma 3.5.5

By maintaining the same notations as the ones we used for Lemma 3.5.2, we add to that the topological condition $\mathcal{S} \subset \bigcup_{i=1}^{h_n} \mathcal{B}(x_i, s_n)$ realized by the compactness property of $\mathcal{S}$ for all sequence of positive real numbers h_n. In the following, we will note for all pairs (x, z),

$$\sup_{x\in\mathcal{S}} \sup_{z\in\mathcal{I}} |\widetilde{\varphi}_n(x, z) - \ddot{\varphi}_n(x, z)| = K_{k,n}(x, z)$$

and we set $h(x) = \arg \min_{i\in\{1,\dots,h_n\}} d_{\mathcal{H}}(x, x_i)$ where $\mathcal{B}(x_i, s_n) = \{x \in \mathcal{S};\, d_{\mathcal{H}}(x, x_i) \leq s_n\}$. Then, one can write the following decomposition

$$\begin{aligned}
K_{k,n}(x, z) &\leq \sup_{x\in\mathcal{S}} \sup_{z\in\mathcal{I}} |\widetilde{\varphi}_n(x, z) - \widetilde{\varphi}_n(x_i, z)| \\
&+ \max_{j\in\{1,\dots,r_n\}} \sup_{x\in\mathcal{S}} |\widetilde{\varphi}_n(x_i, z) - \widetilde{\varphi}_n(x_i, z_j)| \\
&+ \max_{i\in\{1,\dots,h_n\}} \max_{j\in\{1,\dots,r_n\}} |\widetilde{\varphi}_n(x_i, z_j) - \ddot{\varphi}_n(x_i, z_j)| \\
&+ \max_{j\in\{1,\dots,r_n\}} \sup_{x\in\mathcal{S}} |\ddot{\varphi}_n(x_i, z_j) - \ddot{\varphi}_n(x_i, z)| \\
&+ \sup_{x\in\mathcal{S}} \sup_{z\in\mathcal{I}} |\ddot{\varphi}_n(x_i, z) - \ddot{\varphi}_n(x, z)| \\
&= D_{1n} + D_{2n} + D_{3n} + D_{4n} + D_{5n}. \qquad (3.21)
\end{aligned}$$

Hence, the techniques are the same to treat the convergence of D_{1n} and that of D_{5n}. We will just treat one of them, let it be the first one. In this case, we use the fact that

L_1 is Lipschitzian on $(0,1)$

$$\begin{aligned} D_{1n} &\leq \sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}} \frac{1}{\chi_n(x,a_k)}\left|\sum_{k=1}^{n}\frac{\widetilde{Z}_k}{b_k}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)-L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)\right]\right| \\ &\leq \frac{C}{\chi_n(x,a_k)}\sup_{x\in\mathcal{S}}\sup_{z\in\mathcal{I}}\left|\sum_{k=1}^{n}R_k\right| \end{aligned}$$

with $R_k = \dfrac{\widetilde{Z}_k}{a_k b_k} d_{\mathcal{H}}(x,x_i)$, it comes for n large enough

$$D_{1n} = D_{5n} = O_{a.s.}\left\{\left(\frac{\log n}{b_n\chi_n(x,a_n)}\right)^{1/2}\right\}. \tag{3.22}$$

Turning to D_{2n}, by maintaining the condition of Lipschitz on the function L_2, the condition (A.6) and following the same passages of Lemma 3.5.2, one has

$$\begin{aligned} |\widetilde{\varphi}_n(x_i,z)-\widetilde{\varphi}_n(x_i,z_j)| &\leq \frac{1}{\chi_n(x,a_k)}\sum_{k=1}^{n}\frac{L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)}{b_k}|\widetilde{Z}_k-\widetilde{Z}_{k,j}| \\ &\leq C_2\frac{s_n}{\chi_n(x,a_k)}\sum_{k=1}^{n}T_k\frac{\Delta_k}{\overline{G}^W(Z_k)} \end{aligned}$$

with $T_k = \dfrac{1}{b_k^2}L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)$. The fact that $s_n = n^{-(3/2)\beta-1/2}$, allows us to obtain

$$D_{2n} = D_{4n} = O_{a.s.}\left\{\left(\frac{\log n}{b_n\chi_n(x,a_n)}\right)^{1/2}\right\} \tag{3.23}$$

In order to end the proof, it remains to study the third term on the right hand side of (3.21). We follow the same arguments as for proving the second term (A_{2n}) of Lemma 3.5.2, one may show that for all $\eta_0 > 0$

$$\begin{aligned} \mathbb{P}\left\{D_{3n}>\eta\right\} &= \mathbb{P}\left\{\max_{j\in\{1,\dots,r_n\}}\max_{i\in\{1,\dots,h_n\}}|\widetilde{\varphi}_n(x_i,z_j)-\ddot{\varphi}_n(x_i,z_j)|>\eta\right\} \\ &\leq r_n h_n \max_{j\in\{1,\dots,r_n\}}\max_{i\in\{1,\dots,h_n\}}\mathbb{P}\left\{|\widetilde{\varphi}_n(x_i,z_j)-\ddot{\varphi}_n(x_i,z_j)|>\eta\right\} \end{aligned}$$

where for $1 \leq k \leq n$, we have

$$\widetilde{\varphi}_n(x_i, z_j) - \ddot{\varphi}_n(x_i, z_j) = (\chi_n(x, a_k))^{-1} \sum_{k=1}^{n} U_{k,n}(x_i, z_j).$$

Applying the exponential inequality for bounded martingale differences sequence $\{U_{k,n}(x_i, z_j)\}_{k\geq 1}$ and without repeating the same calculations, we refer just to the first part of Lemma 3.5.2, the following inequality is still true

$$\left|\mathbb{E}\left[U^p_{k,n}(x_i, z_j)/\wp_{k-1}\right]\right| \leq 2^p C b_n^{-p} \phi_k(x_i, a_n).$$

Since $r_n = O(s_n^{-1})$, we have

$$r_n h_n \max_{j\in\{1,\dots,r_n\}} \max_{i\in\{1,\dots,h_n\}} \mathbb{P}\left\{|\widetilde{\varphi}_n(x_i, z_j) - \ddot{\varphi}_n(x_i, z_j)| > \eta\right\} \leq M' r_n h_n n^{-C\eta_0^2}.$$

Taking into account the hypothesis (A.5)(i), the Borel-Cantelli's Lemma with an appropriate choice of η_0 give directly

$$D_{3n} = O_{a.s.}\left\{\left(\frac{\log n}{b_n \chi_n(x, a_n)}\right)^{1/2}\right\}. \qquad (3.24)$$

From (3.22)-(3.24), the proof is then achieved.

Proof of Lemma 3.5.6

The positive real sequence s_n given previously checks the condition $h_n s_n \leq M_2$ for a positive finite constant M_2. Then, we have for all $x \in \mathcal{S}$ and $i = 1, \dots, h_n$ the following

$$\begin{aligned} \sup_{x\in\mathcal{S}} |\widetilde{\gamma}_n(x) - \ddot{\gamma}_n(x)| &= \sup_{x\in\mathcal{S}} |V_{k,n}(x)| \\ &\leq \sup_{x\in\mathcal{S}} |\widetilde{V}_{k,n}(x)| + \sup_{x\in\mathcal{S}} |V_{k,n}(x_i)|. \end{aligned} \qquad (3.25)$$

Roughly the same procedures are followed as in the second part of the proof of Lemma 3.5.2 with respect to x, and which give us

$$\begin{aligned}\sup_{x\in\mathcal{S}}|\widetilde{V}_{k,n}(x)| &\leq \sup_{x\in\mathcal{S}}\left\{(\chi_n(x,a_k))^{-1}\sum_{k=1}^{n}\left|L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)-L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)\right|\right\}\\ &+ \sup_{x\in\mathcal{S}}\left\{(\chi_n(x,a_k))^{-1}\sum_{k=1}^{n}\mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)-L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)/\wp_{k-1}\right]\right\}\\ &= \Lambda_1+\Lambda_2\end{aligned}$$

Making use now of the Hölderian property of the kernel L_1 on $(0,1)$ for $\beta_1=1$, the first term of the right hand-side can be rewritten as follows

$$\Lambda_1 \leq (\chi_n(x,a_k))^{-1}\sum_{k=1}^{n}\frac{d_{\mathcal{H}}(x,x_i)}{a_k}$$

and the same is true for the second term by using $s_n=n^{-(3/2)\beta-1/2}$. From this we conclude that

$$\sup_{x\in\mathcal{S}}|\widetilde{V}_{k,n}(x)|=O_{a.s.}(1).$$

It remains for us now to evaluate the term $\sup_{x\in\mathcal{S}}|V_{k,n}(x_i)|$, for n large enough and an optimal choice of ς_0 we obtain

$$\begin{aligned}\mathbb{P}\left\{\sup_{x\in\mathcal{S}}|V_{k,n}(x_i)|>\varsigma\right\} &\leq \mathbb{P}\left\{\max_{i\in\{1,\ldots,h_n\}}|V_{k,n}(x_i)|>\varsigma\right\}\\ &\leq h_n\max_{i\in\{1,\ldots,h_n\}}\mathbb{P}\left\{|\widetilde{\gamma}_n(x_i)-\ddot{\gamma}_n(x_i)|>\varsigma\right\}\end{aligned}$$

In order to apply Lemma 3.7.1, let $M_{k,n}(x_i) = L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right) - \mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)/\wp_{k-1}\right]$ as mentioned in proof of Lemma 3.5.3. A direct application of exponential inequality and the result of Lemma 3.7.2(i) achieve the proof.

Proof of Lemma 3.5.7

From the definition of the estimator $\widehat{\xi}_n^x(z)$ and (3.5), by the assumed condition on the function $\gamma_X(\cdot)$ and using the fact that $\mathbb{I}_{[Z_k\leq W_k]}\Phi(Y_k)=\mathbb{I}_{[Z_k\leq W_k]}\Phi(Z_k)$, it is enough to study this quantity

$$\begin{aligned}
|\widehat{\varphi}_n(x,z)-\widetilde{\varphi}_n(x,z)| &= (\chi_n(x,a_k))^{-1}\left|\sum_{k=1}^{n}\frac{\Delta_k}{b_k}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}(z-Y_k)\right)\right. \\
&\left.\left[\frac{1}{\overline{G}^W(Y_k)}-\frac{1}{\overline{G}_n^W(Y_k)}\right]\right| \\
&\leq (\chi_n(x,a_k))^{-1}\sum_{k=1}^{n}\frac{\Delta_k L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)}{b_k} \\
&\left|\left[\frac{\overline{G}_n^W(Z_k)-\overline{G}^W(Z_k)}{\overline{G}_n^W(Z_k)\overline{G}^W(Z_k)}\right]L_2\left(b_k^{-1}(z-Z_k)\right)\right|.
\end{aligned}$$

it follows

$$\begin{aligned}
\sup_{x\in\mathcal{S}}\sup_{0\leq z\leq T}|\widehat{\varphi}_n(x,z)-\widetilde{\varphi}_n(x,z)| &\leq \left\{\left[\overline{G}_n^W(T)\right]^{-1}\sup_{0\leq z\leq T}\left|\overline{G}_n^W(z)-\overline{G}^W(z)\right|\right\} \\
&\times \; (\chi_n(x,a_k))^{-1}\sum_{k=1}^{n}\frac{L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)}{b_k}\left|L_2\left(b_k^{-1}(z-Z_k)\right)\right|.
\end{aligned}$$

Then, from the very last inequality obtained and using the fact that $T < T_{F^Z}\wedge T_{G^W}$ and for any positive constant M_3, such that

$$\sup_{0\leq z\leq T}\left|\overline{G}_n^W(z)-\overline{G}^W(z)\right| \leq M_3\left\{n^{-1/2}\left(\log_2 n\right)^{1/2}\right\}$$

(the interested reader can be referred to Deheuvels and Einmahl (2000) [11]), we fall directly on the desired result

$$\sup_{x\in\mathcal{S}}\sup_{0\leq z\leq T}\left|\widehat{\xi}_n^x(z)-\widetilde{\xi}_n^x(z)\right| \longrightarrow 0 \; as \; n\to\infty.$$

This lemma remains valid for the punctual case and follows the same calculation structures.

Bibliography

[1] F. Ardjoun, L. Ait Hennani, A. Laksaci. (2016). A recursive kernel estimate of the functional modal regression under ergodic dependence condition. J. Statist Theory and Practice., 10, 475-496.

[2] S. Attaoui, A. Laksaci, E. Ould Saïd. (2011). A note on the conditional density estimate in the single functional index model. Stat Probab Lett 81:45-53.

[3] J. Baek, S. Li Niu. (2016). Estimation of conditional mode with truncated, censored and dependent data, Communications in Stats-Theory and Methods.

[4] F. Benziadi, A. Laksaci, F. Tebboune. (2016). Recursive kernel estimate of the conditional quantile for functional ergodic data. Comm in Statist-Theory and Methods 3097-3113.

[5] A. Berlinet, A. Gannoun, et E. Matzner-LÃ¸ber. (1998). Normalité asymptotique d'estimateurs convergents du mode conditionnel. Canad. J. Statist., 26(2), 365-380.

[6] A. Carbonez, L. Györfi, E. C. van der Meulen. (1995). Partitioning-estimates of a regression function under random censoring. Statist. Decisions, 13(1), 21-37.

[7] M. Chaouch, S. Khardani. (2019). Kernel-smoothed conditional quantiles of randomly censored functional stationary ergodic data. Statist, Theory.

[8] S. Dabo-Niang, A. Laksaci. (2007). Estimation non paramétrique de mode conditionnel pour variable explicative fonctionnelle. C. R. Math. Acad. Sci. Paris, 344(1), 49-52.

[9] S. Dabo-Niang, S. Ould-Abdi, A. Ould-Abdi, A. Diop. Consistency of a nonparametric conditional mode estimator for random fields. Stat Methods Appl (2014).

[10] D. M. Dabrowska. (1989). Uniform consistency of the kernel conditional Kaplan Meier estimate. The Annals of Statistics 17 (3):1157-67.

[11] P. Deheuvels, and J. H. J. Einmahl. (2000). Functional limit laws for the increments of Kaplan-Meier product-limit processes and applications. Ann. Probab., 28(3), 1301-1335.

[12] S. Didi, D. Louani. Consistency results for the conditional mode estimate on continuous time stationary and ergodic data. Stat. Risk Model. (2014); 31(2): 129-150.

[13] M. Ezzahrioui, and E. Ould-Saïd. (2008(a)). Asymptotic normality of the kernel estimators of the conditional mode for functional data. J. Nonparametric Statist, 20, 3-18.

[14] M. Ezzahrioui, and E. Ould-Saïd. (2010). Some asymptotic results of a nonparametric conditional mode estimator for functional time-series data. Statistica Neerlandica Vol. 64, nr. 2, pp. 171-201.

[15] F. Ferraty, A. Goia, P. Vieu. (2002). Functional nonparametric model for time series: a fractal approach for dimension reduction. TEST 11, 317-344.

[16] F. Ferraty, A. Laksaci, et P. Vieu. (2005). Functional time series prediction via conditional mode estimation. C. R. Math. Acad. Sci. Paris, 340(5), 389-392.

[17] F. Ferraty, A. Laksaci, A. Tadj and P. Vieu. Rate of uniform consistency for nonparametric estimates with functional variables. J. Statist. Plan and Inf, 140 (2010) 335-352.

[18] Z. Guessoum, and E. Ould Saïd. (2008). On nonparametric estimation of the regression function under random censorship model. Statist. Decisions, 26, 159-177.

[19] E. L. Kaplan, P. Meier. (1958). Nonparametric estimation from incomplete observations, J. Amer. Statist. Assoc., 53, 457-481.

[20] S. Khardani, M. Lemdani, and E. Ould-Saïd. (2010). Some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship. J. Korean Statist. Soc., 39(4), 455-469.

[21] S. Khardani, M. Lemdani, and E. Ould-Saïd. (2011). Uniform rate of strong consistency for a smooth kernel estimator of the conditional mode for censored time series. J. Statist. Plan and Inf, 141, 3426-3436.

[22] S. Khardani, S. Semmar. Nonparametric conditional density estimation for censored data based on a recursive kernel. Electronic Journal of Statist Vol. 8 (2014), 2541-2556.

[23] S. Khardani, B. Thiam. (2016). Strong consistency result of a non parametric conditional mode estimator under random censorship for functional regressors. Com In Stat-Theory and Methods, VOL. 45, 1863-1875.

[24] M. Kohler, K. Mathé, and M. Pintér. (2002). Prediction from randomly right censored data. J. Multivariate Anal., 80(1), 73-100.

[25] N. Laïb, D. Louani, (2010). Nonparametric kernel regression estimation for functional stationary ergodic data: asymptotic properties. J. Multivariate Anal., 101(10), 2266-2281.

[26] N. Laïb, D. Louani. (2011). Rates of strong consistencies of the regression function estimator for functional stationary ergodic data. J. Statist. Plan. and Inf, 141, 359-372.

[27] N.X. Ling, Y. Liu, P. Vieu. (2016). Conditional mode estimation for functional stationary ergodic data with missing at random, Statist, 50: 991-1013.

[28] D. Louani, and E. Ould-saïd. Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis, J. of Nonparametric Statist., 11 (2007), 413-442.

[29] B. Maillot. (2008). Conditional density estimation for continuous time processes with values in functional spaces, with applications to the conditional mode and the regression function estimators.

[30] E. Ould-Saïd, Z. Cai. (2005). Strong uniform consistency of nonparametric estimation of the censored conditional mode function. Nonparametric Statist. 17, 797-806.

Chapter 4

Consistency Result for the Recursive Kernel Estimate of the Conditional Quantile on Continuous Time Stationary and Ergodic Data

This chapter is the subject of an article submitted.

Abstract *The principal aim of this chapter is to consider the recursive kernel estimate of the conditional quantile function* q_α *when the response variable* Y *is real and the explanatory variable* X *takes values in some infinite-dimensional space. Then, by considering two continuous time processes* $(X_t, Y_t)_{t\geq 0}$*, we establish the almost sure convergence with rates of our proposed recursive estimator* $\widetilde{q}_\alpha(x)$ *under a stationary ergodic process assumption.*

Key words and phrases *Functional data, Recursive nonparametric estimation, Conditional quantile, Ergodicity, Almost sure convergence, Martingale difference, Continuous process.*

4.1 Introduction

As far as we know, for the last few years, there have been more and more works about the recursive kernel estimate, while that non-recursive method is very sensitive to the supplementary item of data added to the series, it should be re-read the process again. This approach, from a theoretical point of view, has many advantages because of its interest in reducing computation time, in addition to the results given that are closer to the correct one in prediction problems. In this setting, the first recursive estimators of the density and regression functions are introduced respectively by Wagner and Wolverton (1969)[27], Ahmed and Lin(1976)[1] and Devroye and Wagner (1980)[10]. Currently, work in the field of functional variables has turned into an interesting topic lately, thanks to its applications; so that we can mention the work of Amiri et al. (2014)[2] who have studied the asymptotic properties of the recursive nonparametric kernel estimator of the regression. When a recursive estimator of the conditional geometric median is studied by Hervé Cardot et al. (2012)[7] in Hilbert space and they have proved the almost sure convergence together with L^2 rates of convergence. For i.i.d observations, Bouadjemi (2014)[5] has introduced a new estimator of the conditional cumulative distribution function based on a recursive approach and he has elaborated under general conditions a result on the asymptotic normality of this estimate.

The conditional quantile $q_\alpha(x)$ is one of the most studied models in recursive estimation; this is interesting for a certain number of reasons, particularly since it provides a clearer view of the conditional distribution than the classical regression $R(x) = \mathbb{E}(Y/X = x),\ x \in \mathbb{R}^d$ that limits attention only to the conditional mean function influenced by outliers. It has been also known that it supplies a good solution to the prediction problem, thanks to its robustness and it is largely studied in finite dimensional spaces. The literature on the conditional quantile estimate has increased considerably in recent years, Roussas (1969)[24] is the first to be interested in and he has showed the convergence and the asymptotic normality for the kernel method, by using an estimator of the conditional distribution function for markov observations. We can cite the work of Samanta (1989)[26] (see also Chaudhuri (1991b)[8], Berlinet et al. (2001)[4] and Gardes et al. (2010)[18] as a general study and for additional details on the asymptotic properties).

Historically, several authors have discussed the nonparametric estimation of conditional quantiles as the inverse of the conditional distribution function. For example, the asymptotic normality in both cases (i.i.d. and $\alpha-$mixing) is studied by Ezzahrioui and Ould-Saïd ((2005b)[12], (2006b)[13]). The estimation of such a function for functional variables is introduced by Ferraty et al. (2006)[15]; they have constructed a dual kernel estimator for the conditional distribution function and have specified the almost complete convergence rate of this latter. Also, it has been studied in several articles among which we cite Ferraty et al. (2005)[14], Ferraty and Vieu (2006)[16] and Laksaci and Maaref (2009)[21].

Regarding the functional framework, to our knowledge, as a recent work, Benziadi in collaboration with Laksaci (2016)[3] have considered a recursive estimator of the conditional quantile by inverting the double-kernel estimate of the conditional distribution function, which gives efficient results in the case of a stationary ergodic process and they have got its almost complete convergence; this work is established with rates for n copies of a random vector $\{(X_k, Y_k), k = 1, \dots, n\}$. In the same field of stationary and ergodic hypotheses, notice that further studies are conducted in the case of continuous time processes; among whom Didi and Louani (2014)[11] who have established the almost sure convergence of the density, the regression and the conditional mode functions based on kernel estimators. Maillot (2008)[22] has also obtained the uniform convergence of the conditional density estimate and the almost sure convergence of the conditional mode estimator for continuous time processes when explanatory variable is functional. Thus, in this article, the main inspired objective is to extend the previously done work for the kernel recursive estimation of the conditional quantile given a functional covariate by achieving a result in the continuous time case, where the most physical phenomena are functional variables which are observable at continuous time, especially that there is no result investigated in this context before in the recent statistical literature.

Our scientific paper is arranged then according to the following way: In Section 4.2, we present the proposed recursive estimator. In the next section, some necessary hypotheses and notations are given in a very classical manner with comments. In addition to the main results and some particular cases to give a comprehensive look of our model. The last section is dedicated to prove the technical lemmas.

4.2 Presentation of the estimator

Let $Z_t = (X_t, Y_t)_{t\geq 0}$ be a stationary ergodic process defined on a probabilized space $(\Omega, \mathcal{A}, \mathbb{P})$. We also assume X_t take values in a semi-metric space $(\mathcal{H}, d_{\mathcal{H}})$ where $d_{\mathcal{H}}(.,.)$ is a semi-metric associated to the space (keeping in mind that $d_{\mathcal{H}}(x, X_k) = \|x - X_k\|$), whereas Y_t are real-valued random variables; distributed as the couple (X, Y) and both observed at continuous time.

For a fixed $x \in \mathcal{H}$, the conditional distribution function F^x of variable Y knowing $X = x$ given by

$$F(y/x) = \mathbb{P}(Y \leq y/X = x) \quad where\ y \in \mathbb{R}. \tag{4.1}$$

The study focuses on the conditional quantile $q_y(\alpha/x)$ of order α (denoted simply $q_\alpha(x)$) defined for all $x \in \mathcal{H}$ and $\alpha \in]0, 1[$ by

$$q_\alpha(x) := F^{-1}(\alpha/x) = inf\{y \in \mathbb{R} : F(y/x) \geq \alpha\}. \tag{4.2}$$

Thanks to the definition (4.2), it is easily shown that, to estimate the conditional quantile, we have estimate beforehand the conditional distribution function defined in (4.1) and then inverting it. For ensuring the existence and uniqueness of the latter, assuming that for any fixed $x \in \mathcal{H}$, $F(.|x)$ is strictly increasing and continuous in a neighborhood of $q_\alpha(x)$. It is easy then to make an estimator of $q_\alpha(x)$ by

$$\widetilde{q}_\alpha(x) := \widetilde{F}^{-1}(\alpha/x).$$

So, in the following, we may first introduce the recursive double kernels type estimator $\widetilde{F}^x$ (we use the notation $F^x(\cdot)$ instead of $F(\cdot/x)$) of the conditional distribution function F^x for the continuous time process, defined for $t \in [0, T]$ as follows:

$$\widetilde{F}^x(y) = \frac{\displaystyle\int_0^T L_1\left(\frac{d_{\mathcal{H}}(x, X_t)}{a_t}\right) L_2\left(\frac{(y - Y_t)}{b_t}\right) dt}{\displaystyle\int_0^T L_1\left(\frac{d_{\mathcal{H}}(x, X_t)}{a_t}\right) dt}, \ \forall y \in \mathbb{R}. \tag{4.3}$$

Where $L_1(\cdot)$ is the kernel with support $[0,1]$, $L_2(\cdot)$ is a strictly increasing distribution function and a_t, b_t are sequences of positive real numbers with the convention $\frac{0}{0} = 0$ and the denominator is not null.

4.3 Assumptions and main study

4.3.1 General assumptions

To establish our results, some notations are necessary, in the following, on the one hand, we use the notation $\wp_t$ the $\sigma-$field generated by $((X_1,Y_1),\ldots,(X_t,Y_t))$ and $\mathcal{B}_t$ the one generated by $((X_1,Y_1),\ldots,(X_t,Y_t),X_{t+1})$. On the other hand, we use $B(x,h)$ as a ball centered at x with radius h.

In addition, we need to set the necessary following assumptions which will help us in the proofs.

(A.1) Our strictly stationary ergodic process $(Z_t)_{t\geq 0}$ satisfies:

$$\begin{cases} (i)\ The\ function\ \phi(x,h) := \mathbb{P}(X \in B(x,h)) > 0, \forall h > 0. \\ \\ (ii)\ For\ \ all\ \ t \in [0,T],\ there\ \ exists\ a\ deterministic\ \ function\ \phi_t(x,.)\ such\ that \\ \mathbb{P}(X_t \in B(x,h)/\wp_{t-1}) = \mathbb{P}(d_{\mathcal{H}}(x,X_t) \leq h/\wp_{t-1}) \leq \phi_t(x,h),\ \forall h > 0\ \ a.s. \\ and\ \ \lim_{h\to 0} \phi_t(x,h) = 0. \\ \\ (iii)\ For\ \ all\ h > 0,\ \dfrac{\displaystyle\int_0^T \mathbb{P}(X_t \in B(x,h_t)/\wp_{t-1})dt}{\displaystyle\int_0^T \phi_t(x,h_t)dt} \to 1 \quad a.s. \end{cases}$$

Where $B(x,h) := \{x^{'} \in \mathcal{H}/d_{\mathcal{H}}(x,x^{'}) < h\}$, $\phi_t(x,h)$ is the conditional small ball probability.

(A.2) – The conditional distribution function F^x of Y given $X = x$ exists for all $x \in \mathcal{N}_x$ and satisfies the lipschitz condition with respect to each variable, there exists $\delta > 0, b_1 > 0, b_2 > 0$ such that: $\forall (y_1, y_2) \in S^2 = [q_\alpha(x) - \delta, q_\alpha(x) + \delta]^2$, $\forall (x_1, x_2) \in \mathcal{N}_x^2$, one has

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C^{'}(d_{\mathcal{H}}(x_1, x_2)^{b_1} + |y_1 - y_2|^{b_2})$$

– The function F^x is strictly monotonous, bounded, continuously differentiable and admits a continuous conditional density ξ^x with respect to the lebesgue's measure over $\mathbb{R}$ where $\inf_{y \in [q_\alpha(x)-\delta, q_\alpha(x)+\delta]} \xi^x(y) \geq C > 0$.

With $\mathcal{N}_x$ is a fixed neighborhood of x in $\mathcal{H}$, S is a fixed compact subset of $\mathbb{R}$ and C, $C^{'}$ are two strictly positive constants.

(A.3) L_1 is a spherically symmetric bounded kernel on its compact support $[0, 1]$, such that:

$$0 < C\mathbb{I}_{[0,1]} < L_1(t) < C^{'}\mathbb{I}_{[0,1]} < \infty.$$

(A.4)

$$\begin{cases} (i)\ The\ bandwidths\ a_t\ and\ \ b_t\ \ satisfy\ \ the\ following\ conditions: \\ \lim_{t\to\infty} a_t = 0,\ \lim_{t\to\infty} b_t = 0; \\ \\ (ii)\ The\ \ concentration\ function\ \ verifies \\ \lim_{T\to\infty} T\psi_T(x) = \infty,\ \lim_{T\to\infty} \dfrac{\log T}{T\psi_T(x)} = 0\ \ where\ \ \varphi_T(x) = \displaystyle\int_0^T \phi_t(x, a_t)dt \\ and\ \ \ \psi_T(x) = \dfrac{1}{T}\displaystyle\int_0^T \phi_t(x, a_t)dt. \end{cases}$$

(A.5) The kernel L_2 is an increasing continuous function satisfies:

$\forall (y_1, y_2) \in \mathbb{R}^2, |L_2(y_1) - L_2(y_2)| \leq C|y_1 - y_2|$ *and* $\int_{\mathbb{R}} |z|^{b_2} L_2^{(1)}(z)dz < \infty$, $\int_{\mathbb{R}} L_2^{(1)}(z)dz = 1$ where L_2, $L_2^{(1)}$ are bounded.

(A.6) For any $k \geq 1$ and $j = 0, 1$;

$$\mathbb{E}\left[\left(L_2^{(j)}(b_t^{-1}(y - Y_t))\right)^k / \mathcal{B}_{t-1}\right] = \mathbb{E}\left[\left(L_2^{(j)}(b_t^{-1}(y - Y_t))\right)^k / X_t\right] = h_k(x),$$

where the function $h_k(x)$ is continuous in the neighborhood of x.

(A.7) There exists $\eta > 0$ such that $\lim_{T\to\infty} T^{-\eta} \int_0^T \frac{1}{b_t} dt = 0$.

Remark 4.3.1. *For the discretization, we consider for all t, $n = T/\delta \in \mathbb{N}$, $(\delta > 0)$ and $k \in \mathbb{N} \cap [1, n]$, the partition $(T_k)_{1\leq k\leq n}$ of the interval $[0, T]$, such that*

$$[0, T] = \bigcup_{k=1}^{n} [T_{k-1}, T_k].$$

Comments on the assumptions

Assumption (A.1) known as concentration property, it is the same as one mentioned in Ferraty et al. (2006)[15] and analogous to that used by Gheriballah et al. (2013)[19] which plays an essential role in the ergodic and functional context of this paper as well as in the case of finite dimension, it is less demanding compared to the condition (A2) discussed for regression function case in the presence of a functional explanatory variable (see Laïb and Louani (2011)[20]).

Nevertheless, notice that only condition (A.1)(iii) will change and it is also satisfied in our continuous time case. While the condition (A.2) ensures the regularity of the conditional law of Y knowing $X = x$, this type of condition is necessary to achieve and give a precision of the convergence rates for the bias term given below. Furthermore, it is identical to assumption (H.2) made by Ferraty, Laksaci and Vieu (2006)[15].

The assumption (A.4) is very frequent in non-parametric estimation literature dealing with functional data for finite or infinite dimension spaces. Then, it is important to highlight that assumptions (A.3) and (A.5) are fundamental conditions which are intended to ease the computational complexity of recursive kernels. Assumption (A.6) is of Markov's nature which is previously used by Ferraty and al. (2006)[15]. Moreover, condition (A.7) is imposed in order to simplify and obtain the convergence rates and to outline our results proofs.

4.3.2 Strong consistency

We study the almost sure convergence (with rates of convergence) of our proposed family of estimators. Such that, the following theorem gives us the main result

Theorem 4.3.1. *Under the assumptions (A.1)-(A.7) above and for any $x \in \mathcal{H}$. If $\dfrac{T\psi_T(x)}{\log T} \to \infty$ as $T \to \infty$, then, we are able to obtain*

$$\widetilde{q}_\alpha(x) - q_\alpha(x) = O\left(\frac{1}{T\psi_T(x)} \int_0^T \phi_t(x, a_t)\left(a_t^{b_1} + b_t^{b_2}\right) dt\right) + O_{a.s}\left(\left(\frac{\log T}{T\psi_T(x)}\right)^{1/2}\right).$$

4.3.3 Particular cases

We present in this brief section some special cases. So that we discuss the important results in this area and provide a general overview in different types of our study as well as the impact of each of these cases, which correspond to some of the previously achieved ones by the statisticians. Note also that all these results are realized whenever the data are associated to a stationary and ergodic conditions.

- *The simple recursive kernel estimator*

  We start by replacing the cumulative distribution function H by an indicator one in the same model used in (4.3), this makes it possible to write the resulting estimator as follows

$$\widehat{F}_T^x(y) = \frac{\displaystyle\int_0^T L_1\left(\frac{d_{\mathcal{H}}(x, X_t)}{a_t}\right) \mathbb{I}_{(Y_t \leq y)} dt}{\displaystyle\int_0^T L_1\left(\frac{d_{\mathcal{H}}(x, X_t)}{a_t}\right) dt}, \ \forall y \in \mathbb{R}.$$

  Thus, it is clear that the conditional quantile expression is closely related to the distribution function, by the same arguments as of the recursive double kernel estimate, we study in this case the convergence rate which is given in the next corollary

Corollary 4.3.1. *Under assumptions (A.1)-(A.4), we have*

$$\widehat{q}_\alpha(x) - q_\alpha(x) = O\left(\frac{1}{T\psi_T(x)}\int_0^T a_t^{b_1}\phi_t(x,a_t)dt\right) + O_{a.s}\left(\left(\frac{\log T}{T\psi_T(x)}\right)^{1/2}\right)$$

- *Case of discrete processes*

  We consider a sequence of strictly stationary ergodic processes $\{Z_t\}_{t=1,\dots,n}$, then, we need to impose the following additional smoothness assumptions

(A.1$'$)(iii) The conditional concentration function satisfies for all sequence

$$(h_t)_{t=1,\dots,n} > 0, \quad \frac{\sum_{t=1}^{n} \mathbb{P}\left(X_t \in B(x,h_t)/\wp_{t-1}\right)}{\sum_{t=1}^{n}\phi(x,h_t)} \to 1 \quad a.co,$$

(A.7$'$) There exists $\eta > 0$ such that $\lim_{n\to\infty} n^{-\eta}\sum_{t=1}^{n}\frac{1}{b_t} = 0.$

Which will simply allow us to move from a continuous time case to a discrete time case.

Remark 4.3.2. *These two conditions are common and which characterize and control our model in the setting of functional discrete processes (as mentioned in Benziadi et al. (2016)[3]). On the other hand, as part of our methodology; respectively, they are necessary to obtain the convergence rates of the estimator and for the continuity of the model studied here.*

Then, the obtained estimator of the conditional distribution function is similar to the one defined by ([3]):

$$\widetilde{F}^x(y) = \frac{\sum_{t=1}^{n} L_1\left(\frac{d_{\mathcal{H}}(x,X_t)}{a_t}\right) L_2\left(\frac{(y-Y_t)}{b_t}\right)}{\sum_{t=1}^{n} L_1\left(\frac{d_{\mathcal{H}}(x,X_t)}{a_t}\right)}, \ \forall y \in \mathbb{R}.$$

Our theorem leads thus to the following corollary which gives the almost sure convergence with rate of the new constructed recursive kernel estimate.

Corollary 4.3.2. *Under assumptions (A.2)-(A.6) and the additional hypotheses given below, we have*

$$\widetilde{q}_\alpha(x) - q_\alpha(x) = O\left(\frac{1}{T\psi_T(x)}\sum_{t=1}^{n}\phi_t(x,a_t)\left(a_t^{b_1}+b_t^{b_2}\right)\right) + O_{a.s}\left(\left(\frac{\log T}{T\psi_T(x)}\right)^{1/2}\right).$$

- <u>*The classical kernel estimator*</u>

Comparing the recursive kernel method with the classical one by replacing the bandwidths $a_t = a_T$ and $b_t = b_T$ for all $t \in [0,T]$. Thus, in that case, the recursive estimator of the conditional distribution function can be re-write by this way

$$\widetilde{F}^x(y) = \frac{\displaystyle\int_0^T L_1\left(\frac{d_{\mathcal{H}}(x,X_t)}{a_T}\right) L_2\left(\frac{(y-Y_t)}{b_T}\right) dt}{\displaystyle\int_0^T L_1\left(\frac{d_{\mathcal{H}}(x,X_t)}{a_T}\right) dt}, \ \forall y \in \mathbb{R}.$$

Corollary 4.3.3. *Under assumptions (A.1)-(A.6), we have*

$$\widetilde{q}_\alpha(x) - q_\alpha(x) = O\left(a_T^{b1}+b_T^{b_2}\right) + O_{a.s}\left(\left(\frac{\log T}{T\psi_T(x)}\right)^{1/2}\right).$$

Remark 4.3.3. *We show that the obtained rates are similar to that realized in Sultana Didi and Louani's paper (2014)[11] in order to demonstrate the almost sure convergence of the conditional mode in the case of classical kernel estimate for continuous variables in finite dimensional setting.*

4.4 Appendix

From now on, we denote for all $t \geq 0$, by the quantities $L_1\left(\frac{d_{\mathcal{H}}(x,X_t)}{a_t}\right) = L_{1,t}(x)$, $L_2\left(\frac{y-Y_t}{b_t}\right) = L_{2,t}(y)$ and $\psi_T(x) = \mathbb{E}\left(L_1\left(\frac{d_{\mathcal{H}}(x,X_1)}{a_t}\right)\right)$.

To establish our almost sure pointwise result, we need to introduce the following exponential inequality used in Lemma 1 in Laib and Louani (2011) [20] for partial sums of unbounded martingale differences, which permits to control the estimation based on stationary ergodic data.

Lemma 4.4.1. *(Laib and Louani (2011)[20]). Let $(Z_n)_{n\geq 1}$ be a sequence of real martingale differences with respect to the sequence of $\sigma-$ fields $\mathcal{F}_n = \sigma(Z_1, Z_2, \ldots, Z_n)_{n\geq 1}$ generated by the random variables $Z_1, Z_2, \ldots, Z_n$. Set $S_n = \sum_{k=1}^{n} Z_k$. For any $p \geq 2$ and for any $n \geq 1$, assume that there exist some nonnegative constants C and d_n such that*

$$\mathbb{E}(Z_n^p/\wp_{n-1}) \leq C^{p-2} p! d_n^2 \quad almost \quad surely. \tag{4.4}$$

Then, for any $\epsilon > 0$, we have

$$\mathbb{P}(|S_n| > \epsilon) \leq 2\exp\left(-\frac{\epsilon^2}{2(D_n + C\epsilon)}\right), \quad D_n = \sum_{k=1}^{n} d_k^2.$$

The key argument in the proof of this theorem comes from the definition (4.2) of the conditional quantile $q_\alpha(x)$ and since $F^x(y)$ is supposed to be strictly increasing, in addition to the proof of Theorem 2.3.2 given in the article by Benziadi et al. (2016)[3] (for more details and demonstration of this result imposed, refer to this paper). It is thus expressed as follows

$$\sum_{n\geq 1} \mathbb{P}(|\widetilde{q}_\alpha(x) - q_\alpha(x)| > \epsilon) \leq \sum_{n\geq 1} \mathbb{P}\left(\sup_{y\in S}\left|\widetilde{F}^x(y) - F^x(y)\right| \geq C\epsilon\right) < \infty. \tag{4.5}$$

So, the proof of our theorem is an easy consequence of (4.5). This means that it is enough to study the almost sure convergence of the conditional distribution function estimator

$$\sup_{y\in S}\left|\widetilde{F}^x(y) - F^x(y)\right| = O\left(\frac{1}{T\psi_T(x)}\int_0^T \phi_t(x, a_t)\left(a_t^{b_1} + b_t^{b_2}\right) dt\right) + O\left(\left(\frac{\log T}{T\psi_T(x)}\right)^{1/2}\right), \ a.s.$$

From which we deduce the almost sure convergence of the conditional quantile estimator.

Proof of Theorem 4.3.1 The proof of this result may be demonstrated easily by using the following decomposition

$$\widetilde{F}^x(y) - F^x(y) = \widetilde{B}_T(x,y) + \frac{\widetilde{R}_T(x,y) + \widetilde{Q}_T(x,y)}{\widehat{F}_D(x)}. \tag{4.6}$$

Where

$$\widetilde{Q}_T(x,y) = [\widetilde{F}_T^x(y) - \ddot{F}_T^x(y)] - F^x(y)[\widehat{F}_D(x) - \overline{F}_D(x)], \tag{4.7}$$

$$\widetilde{R}_T(x,y) = -\widetilde{B}_T(x,y)[\widehat{F}_D(x) - \overline{F}_D(x)], \tag{4.8}$$

$$\widetilde{B}_T(x,y) = \frac{\ddot{F}_T^x(y)}{\overline{F}_D(x)} - F^x(y). \tag{4.9}$$

And let's define the used functions as follow

$$\begin{aligned}
\widetilde{F}_T^x(y) &:= \frac{1}{T\psi_T(x)} \int_0^T L_1\left(\frac{d_{\mathcal{H}}(x,X_t)}{a_t}\right) L_2\left(\frac{y-Y_t}{b_t}\right) dt, \\
\ddot{F}_T^x(y) &:= \frac{1}{T\psi_T(x)} \int_0^T \mathbb{E}\left[L_1\left(\frac{d_{\mathcal{H}}(x,X_t)}{a_t}\right) L_2\left(\frac{y-Y_t}{b_t}\right) / \wp_{t-1}\right] dt, \\
\widehat{F}_D(x) &:= \frac{1}{T\psi_T(x)} \int_0^T L_1\left(\frac{d_{\mathcal{H}}(x,X_t)}{a_t}\right) dt, \\
\overline{F}_D(x) &:= \frac{1}{T\psi_T(x)} \int_0^T \mathbb{E}\left[L_1\left(\frac{d_{\mathcal{H}}(x,X_t)}{a_t}\right) / \wp_{t-1}\right] dt.
\end{aligned}$$

Thus, Theorem 4.3.1 can be deduced from a sequence of the following auxiliary lemmas and corollary

Lemma 4.4.2. *Under hypotheses (A.1)-(A.7), for any $x \in \mathcal{F}$, we are able to obtain*

$$\sup_{y\in S}\left|\widetilde{F}_T^x(y) - \ddot{F}_T^x(y)\right| = O_{a.s}\left(\left(\frac{\log T}{T\psi_T(x)}\right)^{1/2}\right).$$

Proof of Lemma 4.4.2

For the demonstration, let's start by observing that $[0,T]=\bigcup_{k=1}^{n}[T_{k-1},T_k]$. So, for all $x\in\mathcal{H}$, we can write

$$\begin{aligned}\widetilde{F}_T^x(y)-\ddot{F}_T^x(y) &= \frac{1}{T\psi_T(x)}\int_0^T\left(L_{1,t}(x)L_{2,t}(y)-\mathbb{E}\left[L_{1,t}(x)L_{2,t}(y)/\wp_{t-1}\right]\right)dt\\ &= \frac{1}{T\psi_T(x)}\sum_{k=1}^{n}\int_{T_{k-1}}^{T_k}\left(L_{1,t}(x)L_{2,t}(y)-\mathbb{E}\left[L_{1,t}(x)L_{2,t}(y)/\wp_{t-1}\right]\right)dt\\ &= \frac{1}{T\psi_T(x)}\sum_{k=1}^{n}Z_{T,k}(x).\end{aligned}$$

Where $(Z_{T,k}(x))_{k\geq 1}$ is a martingale difference sequence with respect to the sequence of σ-fields $(\wp_t)_{t\geq 0}$ and defined for all $x\in\mathcal{H}$ by

$$Z_{T,k}(x)=\int_{T_{k-1}}^{T_k}\left[L_{1,t}(x)L_{2,t}(y)-\mathbb{E}\left[L_{1,t}(x)L_{2,t}(y)/\wp_{t-1}\right]\right]dt.$$

Applying now Jensen and Minkowski inequalities for $Z_{T,k}$ with $p\geq 2$, to verify the conditions of the Lemma 4.4.1. Thus, we have

$$\begin{aligned}\left|\mathbb{E}\left[Z_{T,k}^p(x)/\wp_{t-1}\right]\right| &\leq \mathbb{E}\left[\left|\int_{T_{k-1}}^{T_k}\left[L_{1,t}(x)L_{2,t}(y)-\mathbb{E}\left[L_{1,t}(x)L_{2,t}(y)/\wp_{t-1}\right]\right]dt\right|^p/\wp_{t-1}\right]\\ &\leq \int_{T_{k-1}}^{T_k}\left(\mathbb{E}\left[|L_{1,t}(x)L_{2,t}(y)|^p/\wp_{t-1}\right]^{\frac{1}{p}}+\mathbb{E}\left[(\mathbb{E}[|L_{1,t}(x)L_{2,t}(y)|/\wp_{t-1}])^p/\wp_{t-1}\right]^{\frac{1}{p}}\right)^p dt\\ &\leq \int_{T_{k-1}}^{T_k}\left(\mathbb{E}\left[|L_{1,t}(x)L_{2,t}(y)|^p/\wp_{t-1}\right]^{\frac{1}{p}}+\mathbb{E}\left[\mathbb{E}[|L_{1,t}(x)L_{2,t}(y)|^p/\wp_{t-1}]/\wp_{t-1}\right]^{\frac{1}{p}}\right)^p dt\\ &\leq \int_{T_{k-1}}^{T_k}\left(\mathbb{E}\left[|L_{1,t}(x)L_{2,t}(y)|^p/\wp_{t-1}\right]^{\frac{1}{p}}+\mathbb{E}\left[|L_{1,t}(x)L_{2,t}(y)|^p/\wp_{t-1}\right]^{\frac{1}{p}}\right)^p dt\\ &= 2^p\int_{T_{k-1}}^{T_k}\mathbb{E}\left[|L_{1,t}(x)L_{2,t}(y)|^p/\wp_{t-1}\right]dt.\end{aligned}$$

By using the condition (A.1)(ii), (A.6) when $\mathbb{E}\left[|L_{2,t}(y)|^p/X_t\right] \leq C < \infty$ and since the kernel L_1 is bounded from hypothesis (A.3), then, we have for any $p \geq 2$

$$\begin{aligned}
\mathbb{E}\left[|L_{1,t}(x)L_{2,t}(y)|^p/\wp_{t-1}\right] &= \mathbb{E}\left[\mathbb{E}\left[|L_{1,t}(x)L_{2,t}(y)|^p/\mathcal{B}_{t-1}\right]/\wp_{t-1}\right] \\
&= \mathbb{E}\left[L_{1,t}^p(x)\mathbb{E}\left[|L_{2,t}(y)|^p/\mathcal{B}_{t-1}\right]/\wp_{t-1}\right] \\
&= \mathbb{E}\left[L_{1,t}^p(x)\mathbb{E}\left[|L_{2,t}(y)|^p/X_t\right]/\wp_{t-1}\right] \\
&\leq C\phi_t(x, a_t).
\end{aligned}$$

According to Lemma 4.4.1, this form of writing satisfies the conditions imposed, so, we can simply observe that $d_t^2 = \phi_t(x, a_t)$, and this will allow us to apply the exponential inequality in later proofs.

In the rest of this part, to establish the proof of our lemma, the idea is to introduce the compactness of $S = [q_\alpha(x) - \delta,\, q_\alpha(x) + \delta] \in \bigcup_{j=1}^{d_n} S_j$ where $S_j = (y_j - l_n, y_j + l_n)$ and since S is bounded we can take $l_T \leq T^{-1/2\alpha}$. Thus, it suffices to consider the following decomposition where y is in S

$$\begin{aligned}
\sup_{y\in[q_\alpha(x)-\delta,\, q_\alpha(x)+\delta]} \left|\widetilde{F}_T^x(y) - \ddot{F}_T^x(y)\right| &\leq \sup_{y\in S}\left|\widetilde{F}_T^x(y) - \widetilde{F}_T^x(z_y)\right| \\
&+ \sup_{y\in S}\left|\widetilde{F}_T^x(z_y) - \ddot{F}_T^x(z_y)\right| \\
&+ \sup_{y\in S}\left|\ddot{F}_T^x(z_y) - \ddot{F}_T^x(y)\right| \\
&= D_1 + D_2 + D_3. \qquad (4.10)
\end{aligned}$$

Where $z_y = \arg\min_{z\in\{y_1,\ldots,y_{d_n}\}} |y - z|$. Clearly, two kinds of proofs are considered for this decomposition, the first and third terms of these are depend on the same method of computation. When the second one is a dispersion term, controlled by Lemma 4.4.1.

We treat the three terms separately. Firstly, let us begin with the term (D_1), by using the fact that L_2 is a lipschitz function and for a large enough T, we get:

$$\begin{aligned}
\sup_{y\in S}\left|\widetilde{F}_T^x(y)-\widetilde{F}_T^x(z_y)\right| &= \sup_{y\in S}\left|\frac{1}{T\psi_T(x)}\int_0^T L_{1,t}(x)[L_{2,t}(y)-L_{2,t}(z_y)]\right| dt\\
&\leq \frac{1}{T\psi_T(x)}\sup_{y\in S}\int_0^T L_{1,t}(x)\left|L_{2,t}(y)-L_{2,t}(z_y)\right| dt\\
&\leq \frac{C}{T\psi_T(x)}\sup_{y\in S}|y-z_y|\int_0^T \frac{L_{1,t}(x)}{b_t}dt\\
&\leq C\frac{l_T}{T\psi_T(x)}\left(\int_0^T \frac{1}{b_t}dt\right).
\end{aligned}$$

By taking $l_T=T^{-\eta}$, it remains to show that

$$\frac{l_T}{T\psi_T(x)}\left(\int_0^T \frac{1}{b_t}dt\right)=O\left(\left(\frac{\log T}{T\psi_T(x)}\right)^{1/2}\right).$$

Which implies that, there exists a finite constant $\theta>0$ such that, for T great enough, we conclude then

$$\mathbb{P}\left(\sup_{y\in S}\left|\widetilde{F}_T^x(y)-\widetilde{F}_T^x(z_y)\right|>\theta\sqrt{\frac{\log T}{T\psi_T(x)}}\right)=0. \tag{4.11}$$

Turning now our interest to the second term (D_2) of the decomposition, such that

$$\begin{aligned}
\mathbb{P}\left(\sup_{y\in S}\left|\widetilde{F}_T^x(z_y)-\ddot{F}_T^x(z_y)\right|>\theta\sqrt{\frac{\log T}{T\psi_T(x)}}\right) &\leq \mathbb{P}\left(\max_{z\in\{y_1,y_2,\dots y_{d_n}\}}\left|\widetilde{F}_T^x(z)-\ddot{F}_T^x(z)\right|>\theta\sqrt{\frac{\log T}{T\psi_T(x)}}\right)\\
&\leq \sum_{z\in S_j}\mathbb{P}\left(\left|\widetilde{F}_T^x(z)-\ddot{F}_T^x(z)\right|>\theta\sqrt{\frac{\log T}{T\psi_T(x)}}\right)\\
&\leq 2C^{"}\, l_T^{-1}\exp(-C\theta^2\log T)\\
&\leq \frac{2C^{"}}{T^{(C\theta^2-\eta)}}.
\end{aligned}$$

Where $C^"$ is a positive constant. Thus, we are in position to apply Borel-Cantelli Lemma and by taking θ big enough to have $T^{(C\theta^2-\eta)} \to \infty$ as $T \to \infty$. Therefore,

$$\sum_{n\geq 1} \mathbb{P}\left(\sup_{y\in S}\left|\widetilde{F}_T^x(z_y) - \ddot{F}_T^x(z_y)\right| > \theta\sqrt{\frac{\log T}{T\psi_T(x)}}\right) \quad \textit{converges almost surely.}$$

It follows that

$$D_2 = O_{a.s}\left(\left(\frac{\log T}{T\psi_T(x)}\right)^{1/2}\right). \tag{4.12}$$

Similarly, it remains to study the term (D_3) in order to end the proof. So, to demonstrate it, we follow the same steps as for proving (D_1) and condition (A.1)(ii), which allow us to also obtain after a classical calculation the same rate and we write

$$\sup_{y\in S}\left|\ddot{F}_T^x(z_y) - \ddot{F}_T^x(y)\right| \to 0 \ a.s. \ \ as \ \ T \to \infty. \tag{4.13}$$

Finally, by combining the decomposition (4.10) and the three last statements (4.11), (4.12) and (4.13), we get the result when T tending to infinity.

Lemma 4.4.3. *Under hypotheses (A.1)-(A.3) and (A.5), for any $x \in \mathcal{H}$, we have*

$$\sup_{y\in S}\left|\widetilde{B}_T(x,y)\right| = O\left(\frac{1}{T\psi_T(x)}\int_0^T \phi_t(x,a_t)\left(a_t^{b_1} + b_t^{b_2}\right)dt\right).$$

Proof of Lemma 4.4.3

Making use the markov property of the process $(X_t)_{t\geq 0}$, it follows that

$$\begin{aligned}
|\widetilde{B}_T(x,y)| &= \left|\frac{\ddot{F}_T^x(y) - F^x(y)\overline{F}_D(x)}{\overline{F}_D(x)}\right| \\
&= \left|\frac{1}{T\psi_T(x)\overline{F}_D(x)}\int_0^T \left(\mathbb{E}[L_{1,t}(x)L_{2,t}(y)/\wp_{t-1}] - F^x(y)\mathbb{E}[L_{1,t}(x)/\wp_{t-1}]\right)dt\right| \\
&= \left|\frac{1}{T\psi_T(x)\overline{F}_D(x)}\int_0^T \left(\mathbb{E}[L_{1,t}(x)\mathbb{E}\left(L_{2,t}(y)/\mathcal{B}_{t-1}\right)/\wp_{t-1}] - F^x(y)\mathbb{E}[L_{1,t}(x)/\wp_{t-1}]\right)dt\right| \\
&\leq \frac{1}{T\psi_T(x)\overline{F}_D(x)}\int_0^T \mathbb{E}\left(L_{1,t}(x)\left[|\mathbb{E}(L_{2,t}(y)/X_t) - F^x(y)|\right]/\wp_{t-1}\right)dt.
\end{aligned}$$

The kernels are bounded, so we can easily see, after using successively an integration by parts and a change of variable $z = (y - Y_t)/b_t$, that we get

$$\mathbb{E}(L_{2,t}(y)/X_t) = \int_{\mathbb{R}} L_2^{(1)}(z) F^{X_t}(y - b_t z) dz.$$

Moreover, under hypotheses $(A.1)$ and $(A.3)$, we have

$$T\psi_T(x)\overline{F}_D(x) \geq C\varphi_T(x) \geq CT\psi_T(x).$$

This inequality is useful in the sequel. Then, condition (A.2) allows us to write

$$\begin{aligned} |\widetilde{B}_T(x,y)| &\leq \frac{C'}{T\psi_T(x)} \int_0^T \mathbb{E}\left[L_{1,t}(x)\left(\int_{\mathbb{R}} L_2^{(1)}(z)|F^{X_t}(y - b_t z) - F^x(y)|dz\right)/\wp_{t-1}\right] dt \\ &\leq \frac{C'}{T\psi_T(x)} \int_0^T \phi_t(x, a_t)\left(\int_{\mathbb{R}} L_2^{(1)}(z)\left(a_t^{b_1} + |z|^{b_2} b_t^{b_2}\right) dz\right) dt. \end{aligned}$$

Now, by the assumption (A.5) and the fact that $L_2^{(1)}(\cdot)$ is a density of probability, we immediately have

$$\begin{aligned} \int_{\mathbb{R}} L_2^{(1)}(z)\left(a_t^{b_1} + |z|^{b_2} b_t^{b_2}\right) dz &\leq \int_{\mathbb{R}} L_2^{(1)}(z) a_t^{b_1} dz + \int_{\mathbb{R}} L_2^{(1)}(z)|z|^{b_2} b_t^{b_2} dz \\ &\leq a_t^{b_1} + M b_t^{b_2}. \end{aligned}$$

Therefore, we get the desired result with C_1 is a constant supposed does not depend on x and y

$$|\widetilde{B}_T(x,y)| \leq C_1 \left[\frac{1}{T\psi_T(x)} \int_0^T \phi_t(x, a_t)\left(a_t^{b_1} + b_t^{b_2}\right) dt\right]. \tag{4.14}$$

And here we have come to complete the proof of the Lemma 4.4.3.

Remark 4.4.1. *The term $\widetilde{R}_T(x,y)$ can be easily deduced from the decomposition (4.8) together with the statement (4.14) and the result of the next lemma.*

Lemma 4.4.4. *Under the hypotheses (A.1), (A.3) and (A.4)(ii), we have*

$$\widehat{F}_D(x) - \overline{F}_D(x) = O_{a.s}\left(\left(\frac{\log T}{T\psi_T(x)}\right)^{1/2}\right).$$

Proof of Lemma 4.4.4

It is very simple to show that for all $x \in \mathcal{H}$, we can write without loss of generality

$$\widehat{F}_D(x) - \overline{F}_D(x) = \frac{1}{T\psi_T(x)} \sum_{k=1}^{n} \Delta_{T,k}(x),$$

with

$$\Delta_{T,k}(x) = \int_{T_{k-1}}^{T_k} [L_{1,t}(x) - \mathbb{E}(L_{1,t}(x)/\wp_{t-1})]\, dt.$$

Now, applying the exponential inequality since $\Delta_{T,k}$ fulfills the conditions of Lemma 4.4.1, for all $\epsilon_0 > 0$, it follows that

$$\begin{aligned}
\mathbb{P}\left(\left|\widehat{F}_D(x) - \overline{F}_D(x)\right| > \epsilon_0 \sqrt{\frac{\log T}{T\psi_T(x)}}\right) &= \mathbb{P}\left(\left|\sum_{k=1}^{n} \Delta_{T,k}(x)\right| > \epsilon_0 T\psi_T(x)\sqrt{\frac{\log T}{T\psi_T(x)}}\right) \\
&\leq 2\exp\left(-\frac{T^2\psi_T^2(x)\epsilon_0^2 \frac{\log T}{T\psi_T(x)}}{2\left(D_n + CT\psi_T(x)\epsilon_0\sqrt{\frac{\log T}{T\psi_T(x)}}\right)}\right) \\
&\leq 2\exp\left(-\frac{T\psi_T(x)\epsilon_0^2 \log T}{2\left(\varphi_T(x) + CT\psi_T(x)\epsilon_0\sqrt{\frac{\log T}{T\psi_T(x)}}\right)}\right) \\
&\leq 2\exp\left(-\frac{\epsilon_0^2 \log T}{2\left(1 + C\epsilon_0\sqrt{\frac{\log T}{T\psi_T(x)}}\right)}\right) \\
&\leq \frac{2}{T^{C'\epsilon_0^2}}.
\end{aligned}$$

While, according to the condition (A.4)(ii), we have

$$\lim_{T\to\infty} \frac{\log T}{T\psi_T(x)} = 0$$

with $C^{'}$ is a positive constant. So, it is enough then to take ϵ_0 sufficiently large and to use once again the Borel-Cantelli Lemma to complete the proof and obtain almost surely

$$\sum_{n\geq 1} \mathbb{P}\left(\left|\sum_{k=1}^{n} \Delta_{T,k}(x)\right| > T\psi_T(x)\epsilon_0\sqrt{\frac{\log T}{T\psi_T(x)}}\right) < \infty$$

Corollary 4.4.1. *Under assumptions* $(A.1)(iii)$ *and* $(A.3)$*, we have*

$$\exists C > 0, \quad \sum_{n\geq 1} \mathbb{P}\left(\widehat{F}_D(x) \leq C\right) < \infty.$$

Proof of Corollary 4.4.1

Assume that assumptions $(A.1)(iii)$ and $(A.3)$ are satisfied, there exists $0 < C < C^{'} < \infty$. Then, we have for any $x \in \mathcal{H}$

$$0 < C\frac{\displaystyle\int_0^T \mathbb{P}(X_t \in B(x,a_t)/\wp_{t-1})dt}{\displaystyle\int_0^T \phi_t(x,a_t)dt} \leq \overline{F}_D(x) \leq |\widehat{F}_D(x) - \overline{F}_D(x)| + \widehat{F}_D(x).$$

Recall that we already consider in the assumptions the writing

$$\varphi_T(x) = \int_0^T \phi_t(x,a_t)dt,$$

so, the previous inequality can be reformulated as shown below

$$\frac{C}{\varphi_T(x)}\int_0^T \mathbb{P}(X_t \in B(x,a_t)/\wp_{t-1})dt - |\widehat{F}_D(x) - \overline{F}_D(x)| \leq \widehat{F}_D(x).$$

Hence, we can reasonably assume

$$\begin{aligned}\mathbb{P}\left(\widehat{F}_D(x) \leq \frac{C}{2}\right) &\leq \mathbb{P}\left(\left|\frac{C}{\varphi_T(x)}\int_0^T \mathbb{P}(X_t \in B(x,a_t)/\wp_{t-1})dt - |\widehat{F}_D(x) - \overline{F}_D(x)| - C\right| > \frac{C}{2}\right)\\ &\leq \mathbb{P}\left(|\widehat{F}_D(x) - \overline{F}_D(x)| > \frac{C}{2}\right).\end{aligned}$$

Consequently, it follows that

$$\sum_{n\geq 1}\mathbb{P}\left(|\widehat{F}_D(x)-\overline{F}_D(x)|>\frac{C}{2}\right)<\infty$$

from Lemma 4.4.4's results, which completes the proof.

Lemma 4.4.5. *(Laib and Louani, 2010) Assume that conditions* $(A.1)$ *and* $(A.3)$ *hold true. Then, we have*

$$\widehat{F}_D(x)-1=O_{a.s}\left(\left(\frac{\log T}{T\psi_T(x)}\right)^{1/2}\right)+O_{a.s}(1).$$

Proof of Lemma 4.4.5

Observe that we can write for each $x\in\mathcal{H}$ the following decomposition

$$\begin{aligned}\widehat{F}_D(x)-1 &= (\widehat{F}_D(x)-\overline{F}_D(x))+(\overline{F}_D(x)-1)\\ &= F_{D,1}(x)+F_{D,2}(x).\end{aligned}\tag{4.15}$$

Where

$$\begin{aligned}F_{D,1}(x) &= \frac{1}{T\psi_T(x)}\int_0^T\left(L_{1,t}\left(x\right)-\mathbb{E}\left[L_{1,t}\left(x\right)/\wp_{t-1}\right]\right)dt,\\ F_{D,2}(x) &= \frac{1}{T\psi_T(x)}\int_0^T\mathbb{E}\left[L_{1,t}\left(x\right)/\wp_{t-1}\right]dt-1.\end{aligned}$$

So, to prove this lemma, we will argue these two terms as follows:

- To handle the first term $F_{D,1}$, we start by taking into consideration that $F_{D,1}(x)=\dfrac{1}{T\psi_T(x)}\sum_{k=1}^{n}\Delta_{T,k}(x)$ where $(\Delta_{T,k}(x))$ is an $\wp_{t-1}$ triangular array of martingale differences. Considering now the definition of $F_{D,1}$ and by keeping in mind the ratings of Lemma 4.4.4 and using the same reasoning, it's obvious to conclude that this term converges almost surely to zero as T goes to infinity and we write

$$F_{D,1}(x)=O_{a.s}\left(\left(\frac{\log T}{T\psi_T(x)}\right)^{1/2}\right)=O_{a.s}\left(1\right).\tag{4.16}$$

- Next, for the second one $F_{D,2}$, assuming that the hypotheses (A.1)(ii) and (A.3) are verified, when it is easy to see below that we have

$$\begin{aligned}\mathbb{E}\left[L_{1,t}(x)/\wp_{t-1}\right] &\leq C\mathbb{P}(X_t \in B(x,a_t)/\wp_{t-1}) \\ &\leq C\phi_t(x,a_t).\end{aligned}$$

Then, in view of the last inequality and by a simple manipulation, we thus get

$$\begin{aligned}F_{D,2}(x)+1 &= \frac{1}{T\psi_T(x)}\int_0^T \mathbb{E}\left[L_{1,t}\left(x\right)/\wp_{t-1}\right]dt \\ &\leq \frac{C}{T\psi_T(x)}\int_0^T \phi_t(x,a_t)dt.\end{aligned}$$

So, by assumption (A.1)(iii), we can prove very simply that:

$$F_{D,2}(x) = O_{a.s}(1) \quad as \quad T \to \infty. \tag{4.17}$$

Consequently, Lemma 4.4.5 can be deduced from the decomposition (4.15) and the statements (4.16), (4.17). Therefore, the proof is readily achieved.

Bibliography

[1] I. Ahmad. (1976). Uniform strong convergence of a generalized failure rate estimate. Bull. Math. Statist. 17, 77-84.

[2] A. Amiri, Ch. Crambes, and B. Thiam. (2014). Recursive estimation of nonparametric regression with functional covariate, Computational Statistics and Data Analysis, 69, 154- 172.

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

[4] A. Berlinet, A. Gannoun et E. Matzner-Lober. (2001). Asymptotic normality of convergent estimates of conditional quantiles. Statistics, 18 :1400-1415. 50, 52, 95, 97.

[5] A. Bouadjemi. (2014). Asymptotic Normality of the Recursive Kernel Estimate of Conditional Cumulative Distribution Function. Journal of Probability and Statistical Science 12(2), 117-126.

[6] H. Cardot, C. Crambes and P. Sarda. (2004). Estimation spline de quantiles conditionnels pour variables explicatives fonctionnelles, C. R. Math. Paris,339, 141-144.

[7] H. Cardot, P. Cénac and P. Zitt. (2012). Recursive estimation of the conditional geometric median in Hilbert spaces, Electronic Journal of Statistics.

[8] P. Chaudhuri. Global nonparametric estimation of conditional quantile functions and their derivatives, J. Multivariate Anal., (1991b), 39, 246-269.

[9] S. Dabo-Niang, A. Laksaci. (2011). Nonparametric quantile regression estimation for functional dependent data. Comm in Statist. Theory amp; Methods, 41, 1254-1268.

[10] L. Devroye, T.J. Wagner. (1980). Distribution-free consistency results in nonparametric discrimination and regression function estimation. Ann. Statist. 8, 231-239.

[11] S. Didi, N. Louani. Consistency results for the conditional mode estimate on continuous time stationary and ergodic data. Stat. Risk Model. (2014); 31(2): 129-150.

[12] M. Ezzahrioui, and E. Ould-Saïd. Asymptotic normality of the kernel estimator of the conditional quantile in the normed space. Technical report, (2005b), 258, LMPA, Univ. Littoral Côte d'Opale. (A paraître dans Far East J. theoretical Statist.)

[13] M. Ezzahrioui, and E. Ould-Saïd. Asymptotic results of the kernel estimator of the conditional quantile in the normed space under α-mixing hypothesis. Technical report, (2006b), 309, LMPA, Univ. Littoral Côte d'Opale.

[14] F. Ferraty, A. Rabhi, et P. Vieu. (2005). Conditional quantiles for dependent functional data with application to the climatic El Nio phenomenon. Indian Journal of Statistics, 67(2) :378-398. 50, 126.

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

[16] F. Ferraty, P. Vieu. Nonparametric functional data analysis. Springer Series in Statistics, Theory and practice, Springer, New York, (2006a).

[17] A. Gannoun, J. Saracco, K. Yu. Nonparametric prediction by conditional median and quantiles. J. Stat. Plann. Inference. (2003), 117, No.2 207-223.

[18] L. Gardes, S. Girard, A. Lekina. (2010). Functional nonparametric estimation of conditional extreme quantiles. Journal of Multivariate Analysis, 101 :419-433. (Cité en pages xix, 23, 81 et 134.)

[19] A. Gheriballah, A. Laksaci, and S. Sekkal. (2013). Nonparametric M-regression for functional ergodic data. Statist. Probab. Lett. 83; 902-908.

[20] N. Laïb, and D. Louani. (2011). Rates of strong consistencies of the regression function estimator for functional stationary ergodic data J. Statist. Plan. and Inference , 141, 359-372.

[21] M. Laksaci, and F. Maref. (2009). Conditional cumulative distribution estimation and its applications Journal of probability and statistical sciences, 13, Pages 47-56.

[22] B. Maillot. (2008). Conditional density estimation for continuous time processes with values in functional spaces, with applications to the conditional mode and the regression function estimators.

[23] E. Masry, L. Györfi. (1987). Strong consistency and rates for recursive density estimators of stationary processes. Journal of Multivariate Analysis, 22, 79-93.

[24] G.G. Roussas. (1969). Nonparametric estimation of the transition distribution function of a Markov process, Ann. Math. Statist., 40, 1386-1400.

[25] G.R. Roussas, L.T. Tran. (1992). Asymptotic normality of the recursive kernel regression estimate under dependence conditions. Ann. Statist., 20, 98-120.

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

[27] C. Wolverton, and T. J. Wagner. (1969). Recursive estimates of probability densities, IEEE Trans. Syst. Cybern. 5, 307-308.

Conclusion and Prospects

Conclusion

During this work, researchers are interested in studying the asymptotic properties of conditional nonparametric models related to recursive kernel approach, and thereafter assume that the sample we study consists of functional ergodic observations.

- To make the work easier to read, Chapter 1 has recalled the necessary key concepts and notions related to statistics on incomplete data and other properties needed for the study. In Chapter 2, 3 and 4, the interest has been in the first time, the presentation of new estimator of the conditional distribution function based on the recursive method, in addition to a generalization of this result in continuous time case. We have examined in the second time the problem of estimating the conditional mode function by the recursive kernel method for a sample consisting of variables and that the variable of interest is randomly censored on the right. Fortunately, we have conducted asymptotic normality results as well as almost sure convergence with rates of the estimators.

- Complementary to this work, we have applied these theories to simulated data to test the numerical behavior of the estimators and confirm our theoretical result for an infinite sample size and different censoring rates that measure the degree to which the data are randomly censored on the right. Thus, the estimators are shown to perform poorly in terms of mean square error (MSE) as we increase the percentage of censoring. This indicates that censorship has a significant impact on estimates even in case of increasing the sample size, this will not improve the performance of the methodology.

Prospects

Work in collaboration with Pr. T. Guendouzi and Dr. F. Benziadi has already done on the recursive estimation of some conditional models for right-censored data. At the end of this satisfying work, here are some research perspectives and future challenges on the continuity of this thesis.

Extension of chapters 3 to a truncated response variable and other types

- Following the results concerning incomplete and dependent data, we can think in our future short-term research to determine the asymptotic properties of some estimators in the case where the variable of interest Y is randomly left-truncated. To our knowledge, this is the least discussed case in the literature and we are currently working on this topic.

- It would also be interesting to think that it is possible to extend our results in the presence of other types of incomplete data (left-censored, interval, LTRC, the mixed model,...). Even though this class of data is of high relevance for practical problems, it is still rarely applied in the empirical literature.

- Subsequently, we would like to combine our strengths to extend our work to the case where both variables X and Y are assumed to be incomplete.

Semi-parametric study of certain model in functional context

Another line of research that I am particularly interested in and would like to invest myself in future research, is to study the estimation of these models in a semi-parametric context (the single index domain, in particular). The mentioned models generally depend on one or more parameters as well as a link function to be estimated. Then, as an extension of some existent contributions, the idea would be to propose a recursive estimator of the desired function next to a one for the parameter θ. Indeed, one of the popular methods today that serves the required topic is the recursive SIR (Sliced Inverse Regression).

Reminder of some tools

We recall here without demonstration some basic results, used in the proofs of the achieved results in this thesis.

Theorem 4.4.1. *(Hölder's Inequality for Integrals) Let $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. If f and g are real functions defined on $[a, b]$ and if $|f|^p$, $|g|^q$ are integrable functions on $[a, b]$, then*

$$\int_a^b |f(t)g(t)|dt \leq \left\{ \int_a^b |f(t)|^p dt \right\}^{1/p} \left\{ \int_a^b |g(t)|^q dt \right\}^{1/q},$$

with equality holding if and only if $C_1|f(t)|^p = C_2|g(t)|^q$ almost everywhere, where C_1 and C_2 are constants.

Theorem 4.4.2. *(Minkowski's Inequality) Let $a, b \in \mathbb{T}$ ($\mathbb{T}$ is a subset of the reals) and $p > 1$. For rd-continuous $f, g : [a, b] \to \mathbb{R}$, we have*

$$\left\{ \int_a^b |(f+g)(t)|^p dt \right\}^{1/p} \leq \left\{ \int_a^b |f(t)|^p dt \right\}^{1/p} + \left\{ \int_a^b |g(t)|^p dt \right\}^{1/p}.$$

Theorem 4.4.3. *(Jensen's Inequality) Let $a, b \in \mathbb{T}$ and $c, d \in \mathbb{R}$. Suppose $g : [a, b] \to (c, d)$ is rd-continuous and $F : (c, d) \to \mathbb{R}$ is convex. Then*

$$F\left(\frac{\int_a^b g(t)dt}{b-a} \right) \leq \frac{\int_a^b F(g(t))dt}{b-a}.$$

Theorem 4.4.4. *(Markov's Inequality) Let X be a non-negative random variable. Then, for any $a > 0$*

$$\mathbb{P}\left[X \geq a\right] \leq \frac{\mathbb{E}(X)}{a}.$$

Theorem 4.4.5. *(Lindeberg Condition) Suppose* (X_{nk}) *is a triangular array with* $Z_n = \sum_{k=1}^{n} X_{nk}$ *and* $s_n^2 = var(Z_n)$*. If the Lindeberg condition holds for every* $\epsilon > 0$,

$$s_n^{-2} \sum_{k=1}^{n} \mathbb{E}\left[X_{nk}^2 \mathbb{I}_{(|X_{nk}| \geq \epsilon s_n)}\right] \to 0, \quad as \;\; n \to \infty.$$

Then, $Z_n/s_n \xrightarrow{d} \mathcal{N}(0,1)$.

Lemma 4.4.6. *(Borel-Cantelli Lemma) Let* $\{A_n\}$ *be a sequence of events such that* $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$*. Then, almost surely, only finitely many* A_n*'s will occur.*

Characterization of an Ergodic Application

Corollary 4.4.2. *Let* $(\Omega, \mathcal{B}, \mu)$ *be a probability space and* $T : \Omega \longrightarrow \Omega$ *be a measure preserving transformation, then we have for* $n \to \infty$,

$$T \, is \, ergodic \Leftrightarrow \forall A, B \in \mathcal{B}, \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k} A \cap B) \to \mu(A)\mu(B).$$

Recursive estimation of nonparametric models for functional ergodic data

The main inspirational theme for this work is based on building recursive estimators for nonparametric conditional models, extending the works done previously to the issues recently discussed in nonparametric statistics, such as: functional ergodic context and the survival data analysis. Therefore, the main model considered here is the conditional distribution function and its derivatives of a scalar response variable Y given a random variable X taking its values in semi-metric space. The thesis uses appropriate statistical methodologies and theories to manage basic issues related to the possible prevalence of outliers and incomplete observations in the sample. Given the power of the recursive method, we continue to ask reasonable question of whether this method, proposed for evaluating previous models in complete case, can be considered a useful one and remains a viable alternative if data are incomplete.

Estimation récursive de modèles non paramétriques pour des données ergodiques fonctionnelles

Le thème principal d'inspiration de ce travail est basé sur la construction des estimateurs récursifs pour des modèles conditionnels non paramétriques, en étendant les travaux réalisés précédemment aux questions récemment discutées en statistique non paramétrique, telles que: le contexte ergodique fonctionnel et l'analyse des données de survie. Par conséquent, le modèle principal considéré ici est la fonction de distribution conditionnelle et ses dérivées d'une variable réponse scalaire Y étant donné une variable aléatoire X prenant ses valeurs dans un espace semi-métrique. La thèse utilise des méthodologies et des théories statistiques appropriées pour gérer les problèmes de base liés à la prévalence possible de valeurs aberrantes et d'observations incomplètes dans l'échantillon. Compte tenu de la puissance de la méthode récursive, nous continuons à poser la question raisonnable de savoir si cette méthode, proposée pour évaluer les modèles précédents dans le cas complet, peut être considérée comme utile et reste une alternative viable si les données sont incomplètes.

التقدير التكراري للنماذج اللامعلمية للبيانات الترابطية الوظيفية

يعتمد الموضوع الرئيسي الملهم لهذا العمل على بناء المقدرات العودية للنماذج الشرطية اللامعلمية، وتوسيع الأعمال المنجزة سابقًا إلى القضايا التي تمت مناقشتها مؤخرًا في الإحصائيات اللامعلمية، مثل: السياق الوظيفي الترابطي وتحليل بيانات البقاء. لذلك، فإن النموذج الرئيسي الذي تم النظر فيه هنا هو دالة التوزيع الشرطي ومشتقاتها من متغير الاستجابة العددية نظرًا لمتغير عشوائي يأخذ قيمه في مساحة شبه مترية. تستخدم الأطروحة المنهجيات والنظريات الإحصائية المناسبة لادارة القضايا الأساسية المتعلقة بالانتشار المحتمل للقيم المتطرفة والملاحظات غير المكتملة في العينة. فنظرًا لقوة الطريقة العودية، واصلنا طرح سؤال معقول عما إذا كانت هذه الطريقة، المقترحة لتقييم النماذج السابقة في الحالة الكاملة، يمكن اعتبارها مفيدة وتظل بديلاً قابلاً للتطبيق إذا كانت البيانات غير كاملة.