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

N° Attribué par la bibliothèque

|_|_|_|_|_|_|_|_|_|_|

Année univ.: 2024/2025

Contribution to the nonparametric estimation in relation with the truncated and dependent data

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université de Saïda - Dr Moulay Tahar
Discipline : MATHEMATIQUES

Spécialité: Analyse stochastique, statistique des processus et applications (ASSPA)

par
Mrs. Nedjeh Menasria[1]

sous la direction de
Dr. Imane Bouazza

Soutenu le 16/06/2025 devant le jury composé de

Dr. F. Benziadi	Université de Saïda - Dr Moulay Tahar	Présidente
Dr. I. Bouazza	Université de Saïda - Dr Moulay Tahar	Encadreur
Dr. R. Rouane	Université de Saïda - Dr Moulay Tahar	Examinatrice

[1] menasrianedjeh@gmail.com

Dedication

I dedicate this work to my beloved husband the source of my happiness and success, my little queens shining my life, and my entire family, whose unwavering support and love have been my guiding light throughout this work.
Thank you for being my supports of strength.

I dedicate my graduation to all those who helped and supported me especially the academic and administrative oversight of the university without forgetting all my schoolmate.

Acknowledgements

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

This master-thesis would not have been possible without the invaluable support and guidance of so many people. To whom i am greatly grateful:

I want to express my great appreciation and deepest gratitude to my supervisor, **Dr. Imane BOUAZZA**, without whom this work would not have emerged. It was a great honor and an immense pleasure to be able to accomplish this work under his direction.

I could benefit from her sound advice and careful guidance, availability, kindness and her constant support to carry out this work.

I would like to warmly thank **Mme Benziadi Fatima** for chairing my committee members.

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

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

I want to thank more generally all members of the Laboratory of Stochastic Models, Statistics and Applications University of Moulay Tahar Saïda.

Finally, my thanks goes to all those who participated directly or indirectly in the successful completion of the defense of this work.

Contents

1 General Introduction **5**

1.1 Survival data analysis 7

1.1.1 The origin date 7

1.1.2 Life-time 8

1.1.3 Survival functions 8

1.2 Incomplete data 11

1.2.1 Censored data 11

1.2.2 Truncated data 12

1.2.3 The randomly truncated framework 14

1.3 Dependent data 16

1.4 Nonparametric recursive method 19

2 Conditional Distribution Function Estimation under α-Mixing in Complete Case **21**

2.1 Definition of the estimator 21

2.2 Assumptions and main results 22

2.2.1 Assumptions 22

2.2.2 Discussion on the assumptions 24

2.2.3 Almost sure convergence rate of the conditional distribution function 25

2.2.4 Almost sure convergence rate of the conditional quantile function . 27

2.3 Technical proofs 28

3 Conditional Distribution Function Estimation in case of Truncated and Dependent Data **32**

3.1 The model and its estimate 32

3.2 Assumptions and main results 33
3.2.1 Assumptions 33
3.2.2 Discussion on the assumptions 34
3.2.3 Uniform almost sure consistency of the conditional distribution function 35
3.2.4 Uniform almost sure consistency of the conditional quantile function 35
3.3 Particular cases 38
3.3.1 The real case (p=1) 38
3.3.2 The L^1 recursive estimate 39
3.4 Technical proofs 40

4 Computational Study and Conclusion **53**
4.1 Computational study 53
4.1.1 Simulated data 54
4.1.2 Literature studies 58
4.2 Conclusion 58

Bibliography **59**

Chapter 1

General Introduction

In today's data-driven world, non-parametric statistics has gained significant prominence among researchers and practitioners across various fields. As data sets grow in complexity and size, traditional parametric methods, which rely on strict assumptions about data distributions, often fall short in accurately capturing the underlying patterns. Non-parametric approaches offer a flexible solution, allowing for the analysis of a wide array of data types without the limitations imposed by parametric constraints. This flexibility makes it particularly valuable in modern research such as economics, healthy, and environmental science. Among the various non-parametric methods gaining traction is the double kernel method. This innovative approach is especially relevant for estimating dependent and truncated data. 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. In fact, there are various popular models of weak dependency, α- mixing Rosenblatt (1956)[31]. For a wide view on the different sorts of mixing and examples the reader can refer to Doukhan (1994) [12]. Bisids, on recent works by applying the small-ball probability theory, Farraty et al. (2005) [17] have studied the almost complete convergence of a conditional density estimator and generalized this result to the α-mixing case. In the same framework of mixing functional observations, Masry

(2005) [28] has systematically showed the asymptotic normality of Ferraty and Vieu's (2004) [14] estimator for the regression function. After that Farraty et al. (2006) [16] have constructed a double kernel estimator for a conditional distribution function and have specified the almost complete convergence with rates of this estimator. Later Farraty et al. (2006) [15] have treated the estimation of the conditional distribution function as a preliminary study of conditional quantile estimate. More recently Bouadjemi (2014) [6] has introduced a new nonparametric estimator of the conditional cumulative distribution function of a scalar response variable Y given a functional random variable X. For recursive nonparametric kernel estimation of the conditional quantile of a scalar response variable Y but with ergodic hilbertian explanatory variable X, Benziadi (2016) [2] has used two type of estimators. In the same framework of mixing functional observations, Bouazza (2021) [8] has focused on building recursive estimators for nonparametric conditional models, extending the works done before to deal with the issues recently discussed in nonparametric statistics.

This work has been elaborated as a contribution to the recursive method and its statistical applications. Especially in the case of incomplete and dependent data under α-mixing. The outlines of this work are briefly presented as follows: After a brief reminder on the basic concepts of survival data analysis. Censored and truncated observations are also discussed as part of the issues about incomplete data. Then we delve into the notion of dependence or α-mixing. After that a brief overview to the nonparametric estimation was reviewed in the Chapter 1.

Chapter 2 introduce a new work about the recursive double kernel estimator of the conditional distribution function (*cdf*) in the context of functional data analysis (i.e. $X \in \mathcal{H}$ where $\mathcal{H}$ is an infinite dimensional space) and $y \in R$ in the case of complete data. Our objective is to study the nonparametric issues with recursive estimation method which extends the classical one. In fact we explore the definition of the estimator, its properties, present the main theoretical results on the estimation of the *cdf* such as almost sure convergence under divers assumptions.

In Chapter 3 we also deal with the nonparametric recursive method but now in case of the left-truncated and dependent data of the *cdf* in the same context functional data, working on the estimation when we have incomplete information with a vectorial random variable of interest Y (with $Y \in \mathbb{R}^p$). Furthermore we present the main theoretical results such as the uniform almost sure consistency of cdf under adequate assumptions.

The last Chapter 4 evaluate the performance of the estimators by using simulated data. The computational experiments demonstrate the applicability and effectiveness of the methods. Also, we present some literature studies that compare the performance of both classical and recursive double kernel estimators when faced with incomplete data. Finally, a general conclusion on this study conclude the work.

1.1 Survival data analysis

The analysis of survival data is the study of the arisen, in time, of one precise event for one or several groups of given individuals. This event, often called death (deaths), can be as well the death of an individual as the arisen of a disease, the answer to a treatment or the breakdown of a machine (generally it is a change of state), every observation is defined by :

1.1.1 The origin date

It is the birth date of the subject, if we study the age of the subject when arises the event or date of putting in touch with an infectious agent, if we study the duration of incubation of an infectious disease. Every individual has a date of origin The measure different on the calendar, but which interests us is the extension since this date. The date of origin defined for every individual the time 0.

To allow for comparison of survival durations between the individuals, one precise definition of the event of interest is necessary. If it is the death caused by a disease, it should be made sure that each death is indeed due to the disease studied, and not with other causes [20].

1.1.2 Life-time

We call life-time a positive random variable T, in general, is the time between two events. This variable is observed in various fields, including reliability (duration of hardware life, between two breakdowns of repairable), demography and healthy care (duration of human life, between disease outbreak and recovery, between two births) or economics and insurance (duration of an unemployment episode, life-time of a company, between two claims, moment of default, etc.).

From this point we can say that the data are completely observed. But in life-time studies, it may occur that we will not be able to observe the variable of interest, this is commonly referred to as incomplete data.

1.1.3 Survival functions

Let T be a positive random variable corresponding to the duration of survival. The probability law of T can be characterized by several functions, as shown by the following definitions that can be founded in **Ferraty** and **Vieu** (2006) [15].

Definition 1.1.1. *The probability density function, noted $f(t)$:*

$$f(t) = \lim_{\Delta t \to 0^+} \frac{\mathbb{P}(t \leq T \leq t + \Delta t)}{\Delta t}$$

$f(t)\Delta t + o(\Delta t)$ is thus the probability of knowing the event of interest between t and $t + \Delta t$. The distribution function, noted $F(t)$, satisfy :

$$F(t) = \mathbb{P}(T \leq t) = \int_0^t f(u)du$$

$F(t)$ define the probability of knowing the event of interest between $[0, t]$, this function is monotonous, and we have:

$$F(0) = 0 \quad \text{and} \quad \lim_{t\to\infty} F(t) = 1$$

Definition 1.1.2. *[13] Given a sample $T_i, 1 < i < n$ of the variable T of the repartition function F where the data are actually observed, a natural estimator of F is then the empirical estimator given by :*

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{T_i \leq t\}}$$

Definition 1.1.3. *The survival function, denoted $S(t)$, is defined as:*

$$S(t) = \mathbb{P}(T > t) = 1 - F(t)$$

The survival function is the probability that the time of death is later than some specified time t. survival function $S(t)$ is monotonically decreasing, such that:

$$S(0) = 1 \quad \text{and} \quad \lim_{t\to\infty} S(t) = 0$$

It also characterized the law of T.

Definition 1.1.4. *The risk function, or fate function, or the immediate risk of change of state noted $h(t)$, is defined as being the immediate probability that a duration T of "stay" in a state ends at the moment $t + \Delta t$ knowing that we were at the moment t there, i.e. :*

$$h(t) = \lim_{\Delta t \to 0^+} \frac{\mathbb{P}(t \leq T \leq t + \Delta t / T \geq t)}{\Delta t}$$

We easily show that :

$$\begin{aligned} h(t) &= \frac{f(t)}{S(t)} \\ &= \frac{-d\log(s(t))}{d(t)} \end{aligned}$$

Thus a $h(t)\Delta t$ represent, when Δt is small, the probability "approached" for an individual to reach the event of interest before $t + \Delta t$, conditionally in the fact that it is still in the previous state just before t. This function is also called immediate risk at the moment t. We also notice that the function of risk characterizes the law of T (or $S(t)$).

Definition 1.1.5. *The function of accumulated risk, noted $H(t)$ defined by :*

$$H(t) = \int_0^t h(u)du$$

Remark 1.1.1. *By manipulation of the previous definitions, we find easily the following relations :*

$$\begin{aligned} f(t) &= -\frac{dS(t)}{dt} \\ S(t) &= \exp\left(-\int_0^t h(u)du\right) \\ S(t) &= \exp(-H(t)) \\ f(t) &= h(t)\exp\left(-\int_0^t h(u)du\right) \end{aligned}$$

Thus, the accumulated risk function characterizes the law of T (or $S(t)$).

The distribution of the duration of survival T can be described by one of the functions defined above. However, one of the most interesting is the risk function $h(t)$ because it is a probabilistic description of the immediate future of the subject "still with risk" and reflects differences between the models, often less visible through the distribution functions or survival functions. In epidemiology, in certain cases, it can be interpreted in terms of incidence.

Note that if $h(t)$ is constant (as is noted by λ), then,

$$S(t) = \exp\left(-\int_0^t h(u)du\right) = e^{-\lambda t}$$

becomes the tail of a distribution of exponential law. That supposes that one can adopt the Markovian model in two states to estimate survival and the problem becomes purely parametric. However, in general, $h(t)$ is not constant, which leaves one to deal with the problem using functional statistics.

1.2 Incomplete data

Incomplete data in statistical analysis refer to the case where data or measurements are unavailable for one or more variables. Possibly caused by participant dropout, data collection and sampling errors, or survey non-response. This situation leads to bias, power loss, and misleading inferences if left untreated. for this purpose, several analytical methods are available to handle this.

1.2.1 Censored data

This phenomenon is commonly encountered in survival analysis. The variable of interest T is not observed (the individual has not suffered the event), and is increased or decreased by a variable or a censored value denoted C which, for its part, has been observed. We consider a variable of interest T (a life-time, for example). Instead of observing the variables $T_1, T_2, ...T_n$, which interest us, we observe T_i only when $(T_i < C)$, otherwise we only know that $(T_i > C)$. We are then talking about right censoring, the most frequent (see Figure 1.1).

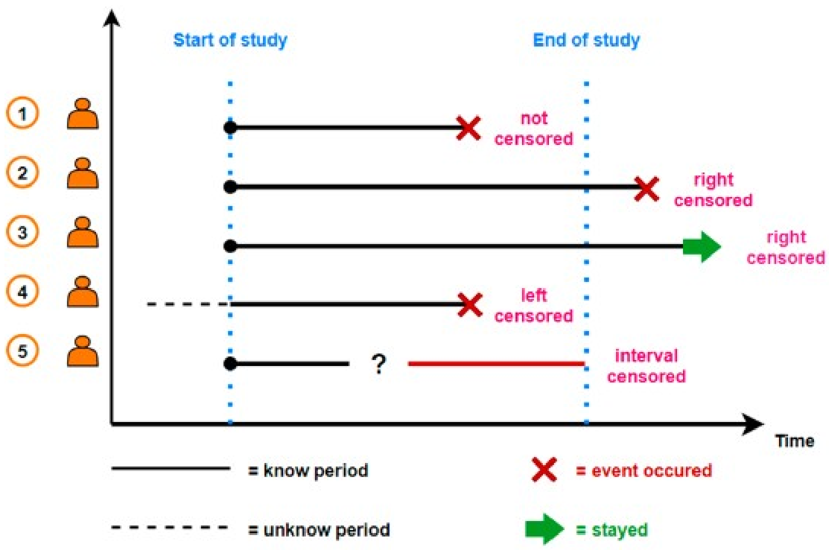

Figure 1.1: Types of censored data [23]

.

Remark 1.2.1. *[13] If the subject has already experienced the event before being observed (we have C instead of T_i), and we know that $(T_i < C)$, it is left censoring. If an event is only observed between two dates $(C_1 < T_i < C_2)$, it is considered censored by interval. If Y is the observed variable, use the notation $Y_i = T_i \wedge C = \min(T_i; C)$ This form of right-censoring is frequently seen in dependability (for the life-time of components created within a given period), medicine (to assess the efficacy of a treatment), biology, etc.*

1.2.2 Truncated data

Truncated data is another form of incomplete data. Truncations are distinct from censures in that they concern the sampling itself (see Figure 1.2). An observation is considered to be truncated if it is dependent on another occurrence. We say that the lifetime variable T is truncated if T is observable only under a certain condition dependent on the value of T. They are classified into three types, according to **Hellal** [20], as folows :

① **Left truncation**

Definition 1.2.1. *Let Y be an independent random variable; we say that there is a left truncation when T is only observable if $T > Y$.*

Example 1.2.1. *If the lifetime of a population is studied using a random cohort in this population, only the survival of subjects living at inclusion can be studied.*

② **Right truncation**

Definition 1.2.2. *Similarly, we have a right truncation where T is only observable if $T < Y$.*

③ Interval truncation

Definition 1.2.3. *When a duration is truncated to the right and left, we say it is truncated by interval.*

Example 1.2.2. *while analyzing patients from a registry, we meet this form of truncation: patients diagnosed before the register's inception or listed after consulting the registry are excluded from the study.*

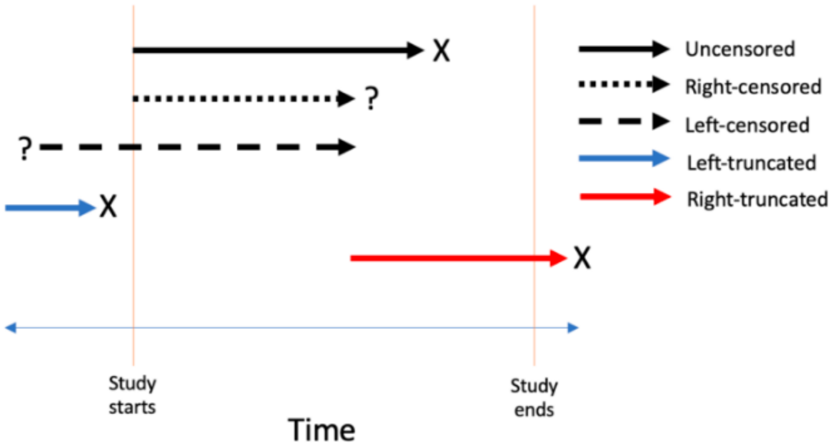

Figure 1.2: Truncation vs. Censoring [26]

.

It should be noted that the literature is much larger with regard to censoring than the truncation which is more recent. Indeed, the estimator of the repartition function of Y in this case, appears for the first time in the work of **Lynden-Bell** (1971)[27]. Since then several articles have appeared interested on it including us, which is the framework of our study.

1.2.3 The randomly truncated framework

Since the truncation does not allow the application of ordinary statistical techniques, we present a reminder of some basic structures based on bibliography corresponding to this context.
We consider $Y_1, ..., Y_N$ a sequence of real random variables of interest that has a distribution function F. And $T_1, ..., T_N$ a sequence of random variables of truncation, with $T \in [0, a_F]$ and has an unknown distribution function G. From that, let $(Y_1, T_1), ..., (Y_N, T_N)$ be the sample truly observed, with T_k are assumed to be independent of Y_k.
We point out that for annotation, let $P_n(\cdot) = P(\cdot \mid n)$ be the conditional probability. Since independence is preserved we can write $P(\cdot) = \mathbb{P}(\cdot \mid Y \geq T)$. Basically this study will be based mainly on the truncation probability defined for the two pairs of observable variables Y and T, by :

$$\tau := \mathbb{P}[Y \geq T] = \int G(v)F(dv) > 0 \tag{1.1}$$

To be able to continue the study, we assume, among the total number in the pooled sample N, that the pairs $(Y_k, T_k), k = 1, ..., n$ can be observed, with the conventions $n \leq N$ (n is known) and $P\left[n/N \to \tau\right] = 1$. So since that, the construction of the distribution F (resp. G) of Y (resp. T) can be reformulated in terms of the size n, thus one must be aware that their joint distribution is also changeful (see **Stute** (1993) [33]), such that :

$$\begin{aligned} H^*(y,t) &= \mathbb{P}(Y_1 \leq y, T_1 \leq t \mid Y_1 \geq T_1) \\ &= P(Y_1 \leq y, T_1 \leq t) \\ &= \tau^{-1} \int_{-\infty}^{y} G(t \wedge v)F(dv) \end{aligned}$$

where $t \wedge u = min(t, u)$.
with the marginal ones which depend on this latter, that generate the distribution of the positive data Y and T respectively :

$$F^*(y) := \tau^{-1} \int_{-\infty}^{y} G(v)F(dv) \text{ and } G^*(t) := \tau^{-1} \int_{-\infty}^{\infty} G(t \wedge v)F(dv)$$

and thus :

$$\begin{aligned}K(y) &= G^*(y) - F^*(y)\\ &= \mathbb{P}(T_1 \le y \le Y_1 \mid Y_1 \ge T_1)\\ &= \tau^{-1} G(y)\bar{F}(y)\end{aligned}$$

with their empirical estimators defined by :

$$F_n^*(y) = n^{-1}\sum_{k=1}^{n} \mathbb{I}_{(Y_k \le y)} \quad \text{and} \quad G_n^*(t) = n^{-1}\sum_{k=1}^{n} \mathbb{I}_{(T_k \le t)}$$

and the consistent estimator of $K(y)$ for $a_F \le y < +\infty$ given by:

$$K_n(y) = n^{-1}\sum_{k=1}^{n} \mathbb{I}_{(T_k \le y \le Y_k)}$$

where $\mathbb{I}_A$ denotes the indicator function of the event A. Note that, the star notation $(*)$ relates to any characteristic of the actually observed data.
For the random left-truncation model, similar to the nonparametric Kaplan-Meier estimator (NPKME) for censored data, the astrophysicist **Lynden-Bell** (1971) [27] has proposed the unique nonparametric estimator (NPLBE) based on maximum likelihood (ML) of the continuous functions F and G expressed as :

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

Note that the KME and LBE always give a valid redistribution of the upper limits, though the result may not be applicable in wider context. In addition, we will set the identifiability conditions on the support of F and G:

$$a_G \le a_F; \quad b_G \le b_F \quad \text{and} \quad \int_{a_F}^{\infty} \frac{1}{G} dF < \infty$$

where a_G, b_G and a_F, b_F denote the extreme points of the supports of G and F respectively. In which, the main asymptotic properties of the later estimates, including the weak and strong uniform convergence with rates of convergence, have been provided from the paper of the statistician **Woodroofe** (1985) [35], such that :

$$\sup_{y \geq a_F} |F_n(y) - F(y)| \xrightarrow{\text{P.a.s.}} 0 \text{ and } \sup_{t \geq a_G} |G_n(t) - G(t)| \xrightarrow{\text{P.a.s.}} 0$$

with a simpler form for the estimator of τ :

$$\widehat{\tau}_n := \frac{G_n(y)\bar{F}_n(y)}{K_n(y)} \tag{1.2}$$

For this, in some references, $F_n(y)$ and $G_n(t)$ called the Lynden-Bell-Woodroofe estimators (NPLBWE). **Honda** [22] proved that $\widehat{\tau}_n$ does not depend on y and its value can then be obtained for any y such that $K_n(y) \neq 0$. Furthermore, she showed its $P - a.s.$ consistency.

Remark 1.2.2. *To be more precise, we note that the strong uniform consistency for the improved product limit estimator of the distribution function F over $[a_F, \infty)$ was proved under the only condition $a_F > a_G$. However, in complementary case $(a_F \leq a_G)$, the desired asymptotic property does not achieved as described by **Chen** ad **al** (1995) [9].*

1.3 Dependent data

In the context of nonparametric estimation, it is appropriate to model the dependence between the random variables. The type of dependency "mixture" is largely used in the literature.

Dependent observations are more adjusted to reality. There are many notions of dependence. We are interested here in those that are expressed in terms of mixing coefficients between tribes generated by the past and the future of the sequence of random variables $(X_n)_n \leq 1$. **Rosen-blatt** (1956) [31] defined the $\alpha - mixing$ coefficients α_n, measuring the difference between the indicators of the events which belong respectively to the tribe generated by the future of the variables after the instant n and that generated by the past of the variables before the instant zero.

The concept of mixing is a set of conditions that are usual structures that model the dependence of a sequence of random variables, we will be interested in the notion of the alpha-mixing sequences witch have a lot of interest [28], the linear processes are under certain mixing conditions, and their mixing coefficients have an explicit order of magnitude, however it is not easy to evaluate them. It is defined as follows.

Definition 1.3.1. *[8] We consider a sequence of random variables $(X_k)_{k>0}$ defined on probabilistic space $(\Omega, \mathcal{F}, \mathbb{P})$. Let us denote by $\mathcal{F}_1^j$ the σ-algebra generated by the $X_k, 1 \leq k \leq j$ and $\mathcal{F}_{n+j}^{\infty}$ the ones generated by the $X_k, n+j \leq k < \infty$. We define the associated mixing coefficient between two σ-fields $\mathcal{F}_{n+j}^{\infty}$ and $\mathcal{F}_1^j$ to the sequence $(X_k)_{k>0}$ by :*

$$\alpha(n) = \sup_{j \geq 1} \sup_{A,B} \left\{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| ; A \in \mathcal{F}_{n+j}^{\infty}, B \in \mathcal{F}_1^j \right\} \tag{1.3}$$

we say that this sequence is α-mixing if :

$$\alpha(n) \to 0 \quad \textit{when} \quad n \to \infty$$

There are two types of strong mixing, such that :

Definition 1.3.2. *[24] The sequence $(X_k)_{k>0}$ is said arithmetically equivalently algebraically α-mixing with rate $\alpha > 0$ if :*

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

it is called geometrically α-mixing if :

$$\exists C > 0, \quad \exists \rho \in]0,1[\quad \alpha(n) \leq C \rho^n.$$

There are various popular coefficients of mixing other than α-mixing [11], as quoted below :

1. $\beta(n) = \sup_{j \geq 1} \sup_{A,B} \left\{ \frac{1}{2} \sum_{i=1}^{I} \sum_{s=1}^{S} |\mathbb{P}(A_i \cap B_s) - \mathbb{P}(A_i)\mathbb{P}(B_s)| ; A_i \in \mathcal{F}_{n+j}^{\infty}, B_s \in \mathcal{F}_1^j \right\}$

2. $\phi(n) = \sup_{j\geq 1}\sup_{A,B}\left\{|\mathbb{P}(B\backslash A) - \mathbb{P}(B)|; A \in \mathcal{F}_{n+j}^{\infty}, B \in \mathcal{F}_1^{j} \text{ and } \mathbb{P}(A) \neq 0\right\}$
3. $\rho(n) = \sup_{j\geq 1}\sup_{A,B}\left\{|corr(X,Y)|; X \in L^2(\mathcal{F}_{n+j}^{\infty}), Y \in L^2(\mathcal{F}_1^{j})\right\}$

These coefficients satisfy the following inequalities :

$$2\alpha \leq \beta \leq \phi,$$

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

Then,

$$\phi - mixing \Rightarrow \beta - mixing \Rightarrow \alpha - mixing$$

$$\phi - mixing \Rightarrow \rho - mixing \Rightarrow \alpha - mixing$$

The tool that will be used in a decisive way in almost sure convergence problems is the Fuk-Nagaev exponential inequality [10].

Lemma 1.3.1. *"Fuk-Nagaev type Inequality under algebraic mixing" Let $\{\Delta_i, i \in \mathbb{N}\}$ be a family of random variables valued in $\mathbb{R}$, of algebraically mixing decreasing coefficient. One pose :*

$$s_n^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} |\mathrm{cov}\,(\Delta_i, \Delta_j)|\,,$$

$\forall i, \quad \|\Delta_i\|_\infty < \infty$, *then for all $\theta > 0$ and any $q > 1$, we have:*

$$\mathbb{P}\left(\left|\sum_{i=1}^{n}\Delta_i\right| > 4\theta\right) \leq 4\left(1 + \frac{\theta^2}{qs_n^2}\right)^{\frac{-q}{2}} + 2ncq^{-1}\left(\frac{2q}{\theta}\right)^{a+1} \tag{1.4}$$

We also use the following Lemma as a necessary tool [24] :

Lemma 1.3.2. *We consider a family of random variables $\{\Delta_i, i \in \mathbb{N}\}$ valued in $\mathbb{R}$. If the condition of strongly mixing is verified and if $\|\Delta_i\| < \infty$ there are for all $i \neq j$:*

$$|\mathrm{cov}\,(\Delta_i, \Delta_j)| \leq 4\alpha(|i - j|)$$

1.4 Nonparametric recursive method

Nonparametric analysis aims at data examination procedures that do not rely on parameterized distribution models, to achieve more flexibility with various types of data. The procedure is most useful when traditional parametric assumptions, such as normality, might not be valid. Recursive models within this context enable analysts to refresh models with emerging data, strengthening their result validity. This approach enhances our understanding of models by uncovering trends and correlations in high-dimensional datasets while minimizing bias from the peculiarities of the data.

The idea of recursive methods is to use the estimates calculated on the basis of the initial data and to update them with only new observations arriving in the database. A major advantage of these methods is that it is not necessary to restart all the calculations of the model parameters whenever the data base is completed by new observations. In general, the advantage of these methods is the ability to take into account the successive collection of the data and to refine the implemented estimation algorithms, in addition to reducing the computing time.

Historically, recursive estimation with rate was introduced by **Wolverten and Wagner** (1969) [34]. Later, **Roussas** (1992) [32] has presented in his work the objective of recursive kernel estimator of the cdf in case of functional data. After that **Baltagi and Li** (1994) [1] proposed a simple recursive estimation method for linear regression models with AR(p) disturbances.

At the beginning of the third millennium, the estimate of the *cdf* in a functional setting has been introduced by **Ferraty** et **al.** (2006) [15]. The authors built a double kernel estimator for the *cdf* and they established the almost complete convergence rate of the estimator when observations are independent and identically distributed ($i.i.d$). The case of $\alpha-mixing$ observations has been studied earlier by **Ferraty** et **al.** (2006) [16]. The first uniforms results available in the literature on the estimation of the distribution function conditionally to a functional variable were established in **Ferraty** et **al.** (2006) [18].

More recently, the asymptotic normality of the kernel estimator of the *cdf* was studied by

Bouadjemi (2014) [6], the author introduced a new nonparametric estimator of the *cdf* of a scalar response variable Y given a functional random variable X. **Benziadi** (2016) [2] built a new estimator for the *cdf*, under certain terms and conditions, she proved the asymptotic normality of the built model. In 2018 **Keddani** et **al** [25] proposed an estimator of the *cdf* when the explanatory variable takes its values in a functional space. **Benziadi** and **Bouazza** (2022) [3] studied the nonparametric recursive estimation of *cdf* of a vectorial response valued variable Y explained by a Hilbertian random variable $X = x$, based on the double-kernel approach.

Chapter 2

Conditional Distribution Function Estimation under α-Mixing in Complete Case

First of all, this chapter presents a new work that has been based on previous study, which is interested to discuss the theoretical results obtained about the almost sure convergence rate of the conditional distribution function under α-mixing in complete case when using the recursive double kernal estimator. To support this work at the end we present a simulation study.

2.1 Definition of the estimator

In order to simplify and give more flexibility for our framework and to focus on the main interest of our study, let us consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\{(X_k, Y_k), k = 1, \ldots, n\}$ be a sample of n random pairs, each one distributed as (X, Y), where X is the random covariate taking its values in a distanced functional space $(\mathcal{H}, d)$, Y is the interest random real variable. The conditional distribution function of Y given

X, denote F, is defined for any $y \in \mathbb{R}$ and any $x \in \mathcal{H}$ by :

$$F^x(y) = \mathbb{P}(Y \leq y \mid X = x)$$

To achieve the desired objective of this study, we based on the modification of **Ferraty** et **al.** estimator (2006) [16] in the case of non-truncated data, witch has been introduced by **Benziadi** et **al.** (2016) [4] to recursively estimate the nonparametric conditional distribution function, they proposed the following estimator:

$$\ddot{F}_n^x(y) = \frac{\sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x, X_k\right)\right) L_2\left(b_k^{-1}\left(y - Y_k\right)\right)}{\sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x, X_k\right)\right)} = \frac{\ddot{\Psi}_n(x,y)}{\ddot{\Upsilon}_n(x)} \tag{2.1}$$

where :

$$\ddot{\Psi}_n(x,y) = \frac{1}{n\psi_n\left(x, a_n\right)} \sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x, X_k\right)\right) L_2\left(b_k^{-1}\left(y - Y_k\right)\right)$$

and :

$$\ddot{\Upsilon}_n(x) = \frac{1}{n\psi_n\left(x, a_n\right)} \sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x, X_k\right)\right)$$

with $\Psi(\cdot,\cdot)$ is the joint probability function assumed to be bounded, $\Upsilon(\cdot)$ is the marginal one and the functions L_1 and L_2 are kernels and a_k, b_k are two positive real numbers tending to 0 as n goes to infinity. In addition to,

$$\psi_n\left(x, a_n\right) = \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x, X_1\right)\right)\right]$$

2.2 Assumptions and main results

2.2.1 Assumptions

In order to achieve the desired results, let us start by proving that our estimate leads to obtain asymptotic properties, for that we first use the notation often introduced in many studies, $\mathcal{B}_k$ σ-field generated by $\{(X_s, Y_s), (X_r), 1 \leq s < k; \quad k \leq r \leq k+1\}$. Thus, let $\mathcal{I}$ be a compact set of $\mathbb{R}$.

Then, to simplify the demonstration of our main results and their proofs, from now on we assume that certain important assumptions are assumed to hold.

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

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

(ii) The joint distribution exists, is bounded and satisfies :

$$0 < \sup_{k \neq l} \mathbb{P}[X_k \in B(x, a_k), X_l \in B(x, a_l)] = O\left\{ \frac{(\phi(x, a_k))^{\frac{(1+a)}{a}}}{n^{\frac{1}{a}}} \right\}$$

(**H.2**) $(X_k, Y_k)_{k \in \mathbb{N}}$ is a stationary sequence of $\alpha-$ dependent real-valued random variables whose coefficients of mixture $\alpha(n)$ satisfy the condition :

$$\exists a, c \in \mathbb{R}_+^* : \forall n \in \mathbb{N}, \alpha(n) = O\left(n^{-a}\right)$$

(**H.3**) On the nonparametric model: $\forall (y_1, y_2) \in \mathcal{I}^2, \forall (x_1, x_2) \in \mathcal{N}_x^2, F^x(y)$ satisfies the *Lipschitz* condition :

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C_1 \left(d_{\mathcal{H}}^{\nu_1}(x_1, x_2) + |y_1 - y_2|^{\nu_2}\right),$$

with,

$$C_1 > 0, \nu_1 > 0, \nu_2 > 0$$

.

(**H.4**) L_1 is a function with support $[0, 1]$, such that :

$$0 < C_1 < L_1(t) \leq C_2 < \infty$$

(**H.5**) L_2 is an increasing, continuous and bounded distribution function satisfying:

$$\forall (y_1, y_2) \in \mathcal{I}^2, |L_2(y_1) - L_2(y_2)| \leq C_3 |y_1 - y_2|$$

where,

$$\int L_2^{(1)}(|t|)\, dt = 1, \quad \text{and} \quad \int |t|^{\nu_2} L_2^{(1)}(|t|)\, dt < \infty$$

(**H.6**) 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) $\sum_{k=1}^{n} \phi_k(x, a_k) = n\psi_n(x, a_n) \to \infty$ and $\lim_{n\to\infty} \frac{\log n}{n\psi_n(x, a_n)} = 0$

2.2.2 Discussion on the assumptions

Commonly, in nonparametric classical and/ or recursive estimation for conditional distribution functions in α-mixing context with dependent processes, which have been adopted by **Doukhan** [12], all the assumptions used in this work are necessary.

The assumption (H.1)(i) is a standard condition for functional estimate. While, (H.1)(ii) is the same as used in **Ferraty** et **al.** [17] among which the small-ball probability satisfies :

$$\sup_{k\neq l} \frac{\mathbb{P}[X_k \in B(x, a_k), X_l \in B(x, a_l)]}{\mathbb{P}[X \in B(x, a_k)]} = O\left\{\left(\frac{\phi(x, a_k)}{n}\right)^{\frac{1}{a}}\right\} \tag{2.2}$$

Compared with *Theorem 4.1* in **Hellal** and **O. Said** [21] for the independent framework in which they used the classical Bernstein exponential inequality for the classical kernel estimate. In the case of dependent observations, when the process (X_k, Y_k) has algebraically decreasing mixing coefficients $\alpha(n)$, we should to set the condition (H.2) in order to use the adapted inequality (1.3.2), then, to study the consistency of the estimator. While, statisticians see that the dependency structure is more complex than the previous one and has many practical applications.

Now, we introduce the regularity condition (H.3), defining the Holderian property of the continuous conditional distribution which makes the proof's steps easier and enables us to obtain the rates of convergence.

Moreover, the hypotheses (H.4), (H.5) are considered as classical assumptions of kernel estimation which are necessary, sufficient and always keep track of the above condition (H.3) in terms of function's class, as well as on the conditional distribution.

Furthermore, (H.6) is an important technical condition on the sequences a_n and b_n, however, rather classic in recursive kernel estimation.

At this stage, we are finally in a position to state our main theoretical results.

2.2.3 Almost sure convergence rate of the conditional distribution function

In this section we recall some results since our estimator of the conditional distribution function has been defined. Recall that, in the case of complete data, a well-known double-kernel estimator is more appropriate in functional analysis. Then, we establish its almost sure convergence rate [1], which is the object of the following theorem.

Theorem 2.2.1. *Under the above hypotheses, and if the below condition posed by* ***Ferraty*** *et* ***al*** *(2005) [18] :*

$$\exists \eta \in \mathbb{R}^*, C n^{\frac{3-a}{1+a}+\eta} \leq \phi\left(x, a_k\right) \leq C' n^{\frac{1}{1-a}}$$

holds with $a > (5+\sqrt{17})/2$, C and C' denote some generic constant in $\mathbb{R}^{*+}$, we have:

$$\sup_{y \in I} \left|\ddot{F}_n^x(y) - F^x(y)\right| = O\left(\sum_{k=1}^{n} a_k^{\nu_1} + \sum_{k=1}^{n} b_k^{\nu_2}\right) + O_{a.s}\left(\frac{\log n}{n \psi_n\left(x, a_n\right)}\right)^{1/2} \tag{2.3}$$

Proof of Theorem 2.2.1 The proof techniques are based mainly on the following stan-

[1] Recall that the sequence $(X_n)_{n\in\mathbb{N}}$ of random variables is said to converge almost surely to some variable X, if [11]:

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1 \quad \text{in short} \quad X_n \xrightarrow{\text{a.s.}} X.$$

dard decomposition :

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

with,

$$\ddot{B}_n(x,y) = \frac{\mathbb{E}\left[\ddot{\Psi}_n(x,y)\right]}{\mathbb{E}\left[\ddot{\Upsilon}_n(x)\right]} - F^x(y)$$

and,

$$\ddot{R}_n(x,y) = -\ddot{B}_n(x,y)\left(\ddot{\Upsilon}_n(x) - \mathbb{E}\left[\ddot{\Upsilon}_n(x)\right]\right)$$

also,

$$\ddot{Q}_n(x,y) = \left(\ddot{\Psi}_n(x,y) - \mathbb{E}\left[\ddot{\Psi}_n(x,y)\right]\right) - F^x(y)\left(\ddot{\Upsilon}_n(x) - \mathbb{E}\left[\ddot{\Upsilon}_n(x)\right]\right)$$

where,

$$\ddot{\Psi}_n(x,y) = \frac{1}{n\psi_n(x,a_n)} \sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}|y - Y_k|\right)$$

and,

$$\ddot{\Upsilon}_n(x) = \frac{1}{n\psi_n(x,a_n)} \sum_{k=1}^{n} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right)$$

in addition to,

$$\mathbb{E}\left[\ddot{\Psi}_n(x,y)\right] = \frac{1}{n\psi_n(x,a_n)} \sum_{k=1}^{n} \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}|y - Y_k|\right)\right]$$

and,

$$\mathbb{E}\left[\ddot{\Upsilon}_n(x)\right] = \frac{1}{n\psi_n(x,a_n)} \sum_{k=1}^{n} \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right)\right]$$

Then, the proof of Theorem 2.2.1 is a direct consequence of the below Lemmas.

Lemma 2.2.1. *Under Hypotheses (H1)-(H3) and (H5), we have :*

$$\sup_{y\in\mathcal{I}}\left|\ddot{B}_n(x,y)\right| = O\left(\sum_{k=1}^{n} a_k^{\nu_1}\right) + O\left(\sum_{k=1}^{n} b_k^{\nu_1}\right) \tag{2.4}$$

Lemma 2.2.2. *Under Hypotheses of Theorem 2.2.1, we have :*

$$\ddot{\Upsilon}_n(x) - \mathbb{E}\left[\ddot{\Upsilon}_n(x)\right] = O_{a.s.}\left(\left(\frac{\log n}{n\psi_n\left(x, a_n\right)}\right)^{1/2}\right) \tag{2.5}$$

Lemma 2.2.3. *Under Hypotheses of Theorem 2.2.1, we have :*

$$\sup_{y\in\mathcal{I}}\left|\ddot{\Psi}_n(x,y) - \mathbb{E}\left[\ddot{\Psi}_n(x,y)\right]\right| = O_{a.s.}\left(\left(\frac{\log n}{n\psi_n\left(x, a_n\right)}\right)^{1/2}\right) \tag{2.6}$$

2.2.4 Almost sure convergence rate of the conditional quantile function

We contend that the accuracy of the conditional quantile estimator fundamentally depends on the properties of the conditional distribution function estimator used in its construction. In fact, we obtain the quantile estimate by inverting the estimated *cdf*.

Consequently, the uniform consistency of the conditional quantile estimator is intrinsically related to *cdf* estimator. To ensure the existence and uniqueness of the conditional quantile function, we explicitly assume that $F^x(\cdot)$ is strictly increasing and continuous over its domain. This assumption is crucial, as it guarantees a well-defined inverse and stabilizes the estimation process, allowing for reliable inference of the conditional quantiles.

From that, we can easily estimate the conditional quantile $q_\alpha(x)$ by :

$$\ddot{q}_{\alpha,n}(x) = \ddot{F}_n^{-1}(\alpha/x) = \inf\left\{y : \ddot{F}_n(y/x) \geq \alpha\right\} \tag{2.7}$$

In that case, we have to introduce another hypothesis (H.7) as follow :

(H.7) For each fixed $\alpha \in (0,1)$, the function $q_\alpha(x)$ satisfies that, for any $\epsilon > 0$ and $\eta_\alpha(x)$, there exists a $\beta > 0$ such that $q_\alpha(x) - \eta_\alpha(x) \geq \epsilon$ implies that $F^x\left(q_\alpha(x)\right) - F^x\left(\eta_\alpha(x)\right) \geq \beta$.
The following Corollary gives the almost sure convergence rate of the estimate $\ddot{q}_\alpha(x)$.

Corollary 2.2.1. *Let the assumptions of Theorem 2.2.1 hold. In addition to (H.7), then,* *we have :*

$$\ddot{q}_{\alpha,n}(x) - q_\alpha(x) = O\left(\sum_{k=1}^{n} a_k^{\nu_1} + \sum_{k=1}^{n} b_k^{\nu_2}\right) + O_{a.s}\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2} \tag{2.8}$$

Proof of Corollary 2.2.1 By following the same steps as for Theorem 2.2.1 we can proof this result.

2.3 Technical proofs

Proof of Lemma 2.2.1 To prove this Lemma we go through the following steps, consider now for all $k = 1, ..., n$ the following notations:

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

Than we write :

$$\begin{aligned}\left|\ddot{B}_n(x,y)\right| &= \left|\frac{1}{n\psi_n(x,a_n)\,\mathbb{E}(\ddot{\Upsilon}_n(x))} \sum_{k=1}^{n} \left\{\mathbb{E}\left(L_{1.k}(x)\mathbb{E}\left[L_{2.k}(y) \mid X\right]\right) - F^x(y)\mathbb{E}\left[L_{1.k}(x)\right]\right\}\right| \\ &\le \frac{1}{n\psi_n(x,a_n)\,\mathbb{E}(\ddot{\Upsilon}_n(x))} \sum_{k=1}^{n=1} \left\{\mathbb{E}\left(L_{1.k}(x)\left|\mathbb{E}\left[L_{2.k}(y) \mid X\right] - F^x(y)\right|\right)\right\}.\end{aligned}$$

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

$$E\left(L_{2.k}(y) \mid X\right) = \int_{\mathbb{R}} L_2^{(1)}(t) F^X(y - b_k t)\, dt.$$

Thus, we have :

$$\left|\mathbb{E}\left[L_{2.k}(y) \mid X\right] - F^x(y)\right| \le \int_{\mathbb{R}} L_2^{(1)}(t) \left|F^X(y - b_k t) - F^x(y)\right| dt$$

Moreover, it follows by (H5) that :

$$\left|\mathbb{E}\left[L_{2.k}(y) \mid X\right] - F^x(y)\right| \le C \int_{\mathbb{R}} L_2^{(1)}(t)\left(a_k^{\nu_1} + |t|^{\nu_2} b_k^{\nu_2}\right) dt$$

This last inequality is uniform on y which achieves the proof of lemma.

Proof of Lemma 2.2.2 : To deal with the proof requirements, we follow the same ideas used by **Ferraty** ad **al** (2004) [14], the main point consists in using a pseudo-exponential inequality taking into account the α-mixing structure. Witch allow as to start by writing :

$$\ddot{\Upsilon}_n(x) - \mathbb{E}\left[\ddot{\Upsilon}_n(x)\right] = \frac{1}{n\mathbb{E}(L_{1.1}(x))}\sum_{k=1}^{n}\Delta_k(x)$$

where,

$$\Delta_k(x) = L_1\left(a_k^{-1}d_{\mathcal{H}}\left(x, X_k\right)\right) - \mathbb{E}\left(L_1\left(a_k^{-1}d_{\mathcal{H}}\left(x, X_k\right)\right)\right)$$

.

The Fuk-Nagaev's inequality allows one to get, for all $\theta > 0$ and $q > 1$:

$$P\left(\left|\ddot{\Upsilon}_n(x) - \mathbb{E}\left[\ddot{\Upsilon}_n(x)\right]\right| > 4\theta\right)$$
$$\leq C\{\underbrace{\frac{n}{q}\left(\frac{q}{\theta n\mathbb{E}(L_{1.1}(x))}\right)^{a+1}}_{Q_1} + \underbrace{\left(1 + \frac{\theta^2 n^2\left(\mathbb{E}(L_{1.1}(x))\right)^2}{q s_n}\right)^{q/2}}_{Q_2}\}$$

where,

$$s_n = \sum_{k=1}^{n}\sum_{l=1}^{n}\mathrm{Cov}\left(\Delta_k(x), \Delta_l(x)\right)$$

By taking,

$$q = C(\log n)^2 \text{ and } \qquad \theta = \theta_0\frac{\sqrt{n\psi_n\left(x, a_n\right)\log n}}{n\mathbb{E}(L_{1.1}(x))}, \tag{2.9}$$

and by using the left part of (H.7), it follows that,

$$Q_1 \leq Cn^{-1-\nu} \tag{2.10}$$

Before we focus on Q_2, we have to study the asymptotic behaviour of :

$$s_n = \underbrace{\sum_{k\neq l}\mathrm{Cov}\left(\Delta_k(x), \Delta_l(x)\right)}_{s_n^{\mathrm{cov}}} + \underbrace{\sum_{k=1}^{n}\mathrm{Var}\left(\Delta_k(x)\right)}_{s_n^{\mathrm{var}}} \tag{2.11}$$

On one hand, by using successively (H1), (H2), (H4) and the right part of (H.7), we have :

$$|\mathrm{Cov}\left(\Delta_k(x), \Delta_l(x)\right)| = O\left(\left(\frac{\psi_n\left(x, a_n\right)}{n}\right)^{1/a} \psi_n\left(x, a_n\right)\right) \tag{2.12}$$

On the other hand, these covariances can be controlled by means of the usual Davydov's covariance inequality for mixing processes (see **Rio**, 2000) [30]. Together with (H.2), this inequality leads to :

$$\forall k \neq l, |\mathrm{Cov}\left(\Delta_k(x), \Delta_l(x)\right)| \leq C|k-l|^{-a} \tag{2.13}$$

Thus, by using the following classical technique (see **Bosq**, 1998) [5], we can write :

$$s_n^{\mathrm{cov}} = \sum_{0<|k-l|\leq u_n} |\mathrm{Cov}\left(\Delta_k(x), \Delta_l(x)\right)| + \sum_{|k-l|>u_n} |\mathrm{Cov}\left(\Delta_k(x), \Delta_l(x)\right)|$$

By putting $u_n = \left(\frac{\psi_n(x,a_n)}{n}\right)^{-1/a}$, and using (2.12) (resp. (2.13)) to treat the first (resp. second) covariance term, we get :

$$s_n^{cov} = O\left(n\psi_n\left(x, a_n\right)\right) \tag{2.14}$$

The variance terms can be calculated by following the same arguments as those invoked to obtain to get (2.12), and we arrive at :

$$\mathrm{Var}\left(\Delta_k(x)\right) = O\left(\psi_n\left(x, a_n\right)\right) \tag{2.15}$$

Now, (2.14) and (2.15) lead directly to :

$$s_n = O\left(n\psi_n\left(x, a_n\right)\right) \tag{2.16}$$

This is ample to study the quantity Q_2, since (2.9) and (2.16) allow us to write that, for n and θ_0 large enough :

$$\exists \nu' > 0, \quad Q_2 \leq Cn^{-1-\nu'} \tag{2.17}$$

Finally, we put together (2.10) and (2.17) in addition to use (H.4) we achieve the proof of Lemma.

Proof of Lemma 2.2.3. It proceeds along the same steps and by invoking the same arguments, just changing the variables $\Delta_k(x)$ into the following ones:

$$\Gamma_k(x) = L_{2.k}(y)L_{1.k}(x) - \mathbb{E}\left[L_{2.k}(y)L_{1.k}(x)\right]$$

Due to the fact that L_2 is a cumulative kernel, so $L_{2.k}(y) \leq 1$. Using this fact systematically to bound the variables $L_{2.k}$, all the calculus made previously with the variables $\Delta_k(x)$ remain valid with the variables $\Gamma_k(x)$. Which completes the proof of Lemmas and, therefore, the Theorem 2.2.1.

Chapter 3

Conditional Distribution Function Estimation in case of Truncated and Dependent Data

The main objective of this chapter is to discuss the results obtained related to this study. Indeed, the sensitivity of the double kernal estimator in the case of truncated data under α-mixing condition, which is a research topic encountered several times in the literature for many conditional models, in particular the distribution and the quantile functions.

3.1 The model and its estimate

In regard to treat our framework in this chapter, let us take the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and consider an infinite stationary dependent random vectors $(X_k, Y_k), k = 1, ..., n$ $(n \leq N)$ taken from the pair (X, Y), where X is functional and Y is the random vectorial variable of interest with a continuous distribution function F. In addition, we consider for this case, the scenario in which the response variable Y is assumed to be subject to the truncation time T.

Therefore, the conditional distribution function of Y given the covariate $X = x$ under the truncation condition exists and is frequently defined by :

$$F_{Y/X}(y/x) = \mathbb{E}\left[\mathbb{I}_{(Y\leq y)}/X = x\right], \forall y \in \mathbb{R}^p$$

With respect to the main objective of this research case, truncated data, our purpose is to adapt the version of a semi recursive double kernel estimator of the model given above denoted $\widehat{F}_n^x(.)$ proposed by **Benziadi** et **al.** (2022) [3] witch is defined as follow, taking into account that G_n as defined in section 1.2.3 :

$$\widehat{F}_n^x(y) = \frac{\sum_{k=1}^{n} G_n^{-1}(Y_k) L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right)}{\sum_{k=1}^{n} G_n^{-1}(Y_k) L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)} = \frac{\widehat{\Psi}_n(x, y)}{\widehat{\Upsilon}_n(x)} \tag{3.2}$$

where,

$$\widehat{\Psi}_n(x, y) = \frac{\widehat{\tau}_n}{n\psi_n(x, a_n)} \sum_{k=1}^{n} \frac{1}{G_n(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right)$$

and,

$$\widehat{\Upsilon}_n(x) = \frac{\widehat{\tau}_n}{n\psi_n(x, a_n)} \sum_{k=1}^{n} \frac{1}{G_n(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right)$$

3.2 Assumptions and main results

3.2.1 Assumptions

Firstly, to ensure that our estimate achieves the asymptotic properties, in this case study we choose the notation often introduced in many studies, $\mathcal{B}_k$ the σ-field generated by $\{(X_s, Y_s), (X_r), 1 \leq s < k; k \leq r \leq k+1\}$. Thus, let $\mathcal{S}$ and $\mathcal{I}$ be, respectively, two compact sets of $\mathcal{H}$ and $\mathbb{R}^p$. Secondly, following **Woodroofe** (1985) [35] in addition to

He and **Yang** (1998) [19], let us note for the distribution function L of Z, the lower and upper limits of the support by :

$$a_L = \inf\{z : L(z) > 0\} \quad \text{and} \quad b_L = \sup\{z : L(z) < 1\}$$

Now, to make the steps of the main results and their proofs clearer, we point out that the previous assumptions of the Theorem 2.2.1 from (H.1) to (H.6) remain valid taking into consideration that $Y \in \mathbb{R}^p$, in addition, we need to add the following assumptions as :

(**H.5**) (ii) There exists a continuous bounded function $l_\infty(\cdot)$ in the neighborhood of x such that the conditional distribution of the couple (Y_k, Y_l) knowing (X_k, X_l) exists and verifies :

$$\max\left[F\left(y_k/x_k\right), F_{k,l}\left(y_k, y_l/x_k, x_l\right)\right] \leq l_\infty(x) < \infty$$

(**H.6**) (iii) $\exists\gamma > 0;\quad \frac{1}{n^\gamma \log n}\sum_{k=1}^{n} b_k^{-1} \to 0$ as $n \to \infty$.

(**H.7**) The variables $(T_k)_{k=1,\ldots,n}$ are independent of $(Y_k)_{k=1,\ldots,n}$.

3.2.2 Discussion on the assumptions

Since, we are still concerned with nonparametric recursive estimation for conditional distribution functions in functional α-mixing context, the previous assumptions from (H.1) to (H.6) remain valid taking into consideration that $Y \in \mathbb{R}^p$. We point out that (H.5)(ii) introduce a continuous bounded function to drive uniform convergence rate, and to avoid instabilities due to explosive behavior of conditional distributions near the boundaries. Furthermore, we add (H.6)(iii) to ensure that we have enough smoothing of the estimator, and consequently converge properly. Moreover, to deal with the difficulty of the problem and to treat properties of the estimator when the sample contains truncated data, we point out that the truncation mechanism would be examined by assumption (H.7) which is considered as a powerful tool in nonparametric truncation estimation in the sense it gives a valid solution.

Now, we state our main theoretical results.

3.2.3 Uniform almost sure consistency of the conditional distribution function

We first establish the rate of the uniform almost sure consistency[2], which is the object of the following theorem. Throughout the rest of this work, K_i (resp. M_i) for $i = 1, \ldots, 6$ will be used to denote the positive constants whose values may vary, in addition to the previous constants mentioned above.

Theorem 3.2.1. *(**Benziadi** and **Bouazza** (2022)) [3]. Suppose that the assumptions (H.1)-(H.7) hold true. For n large enough, we have :*

$$\lim_{n\to\infty}\sup \sup_{x\in\mathcal{S}} \sup_{y\in\mathcal{I}} \frac{\left|\widehat{F}_n^x(y) - F^x(y)\right|}{\left(\left(\sum_{k=1}^{n} a_k^{\nu_1} + \sum_{k=1}^{n} b_k^{\nu_2}\right) + \left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}\right)} \leq K_1 \qquad a.s.$$

The application of Theorem 3.2.1 is needed to obtain the following result.

3.2.4 Uniform almost sure consistency of the conditional quantile function

Considering that the conditional quantile estimator depends on the construction of the conditional distribution function estimator. Thus, its uniform consistency depends basically on that of the previous ones. Normally, it is necessary to assume that $F^x(\cdot)$ is

[2] A sequence of estimators $\hat{X}_n$ is said to be *uniformly almost surely consistent* if:

$$\sup_{x\in\mathcal{S}} |\hat{X}_n(x) - X(x)| \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad n \to \infty$$

strictly increasing and continuous, in order to ensure the existence and the uniqueness of the conditional quantile function. Now, we interest to the conditional quantile $q_\alpha(x)$ that naturally estimated by :

$$\widehat{q}_{\alpha,n}(x) = \widehat{F}_n^{-1}(\alpha/x) = \inf\left\{y : \widehat{F}_n(y/x) \geq \alpha\right\}$$

then, we have to introduce additional condition

(H.8) For each fixed $\alpha \in (0,1)$, the function $q_\alpha(x)$ satisfies that, for any $\epsilon > 0$ and $\eta_\alpha(x)$, there exists a $\beta > 0$ such that $\sup_{x\in\mathcal{S}} |q_\alpha(x) - \eta_\alpha(x)| \geq \epsilon$ implies that $\sup_{x\in\mathcal{S}} |F^x(q_\alpha(x)) - F^x(\eta_\alpha(x))| \geq \beta$.

Corollary 3.2.1. *(**Benziadi** and **Bouazza** (2022)).Let the assumptions of Theorem 3.2.1 hold. In addition to (H.8), one gets :*

$$\limsup_{n\to\infty} \sup_{x\in\mathcal{S}} \frac{|\widehat{q}_{\alpha,n}(x) - q_\alpha(x)|}{\left(\left(\sum_{k=1}^{n} a_k^{\nu_1} + \sum_{k=1}^{n} b_k^{\nu_2}\right) + \left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}\right)} \leq K_2 \qquad a.s.$$

Proof of Theorem 3.2.1 The proof techniques based mainly on the following standard decomposition :

$$\begin{aligned}\widehat{F}_n^x(y) - F^x(y) - \widehat{B}_n(x,y) =& \frac{1}{\widehat{h}_n(x)}\left\{\widehat{Q}_n(x,y) - \widehat{B}_n(x,y)\left[\left(\widehat{\Upsilon}_n(x) - \widetilde{\Upsilon}_n(x)\right)\right.\right.\\ &\left.\left. + \left(\widetilde{\Upsilon}_n(x) - \mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]\right)\right]\right\}\end{aligned}$$

with :

$$\widehat{B}_n(x,y) := \frac{\mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right] - F^x(y)\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]}{\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]}$$

and :

$$\begin{aligned}\widehat{Q}_n(x,y) :=& \left[\left(\widehat{\Psi}_n(x,y) - \widetilde{\Psi}_n(x,y)\right) + \left(\widetilde{\Psi}_n(x,y) - \mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right]\right)\right]\\ &- F^x(y)\left[\left(\widehat{\Upsilon}_n(x) - \widetilde{\Upsilon}_n(x)\right) + \left(\widetilde{\Upsilon}_n(x) - \mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]\right)\right]\end{aligned}$$

where :

$$\widetilde{\Psi}_n(x,y) = \frac{\tau}{n\psi_n(x,a_n)} \sum_{k=1}^{n} \frac{1}{G(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1} \|y - Y_k\|_{\mathbb{R}^p}\right)$$

and :

$$\widetilde{\Upsilon}_n(x) = \frac{\tau}{n\psi_n(x,a_n)} \sum_{k=1}^{n} \frac{1}{G(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right)$$

in addition to :

$$\mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right] = \frac{\tau}{n\psi_n(x,a_n)} \sum_{k=1}^{n} \mathbb{E}\left[\frac{1}{G(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1} \|y - Y_k\|_{\mathbb{R}^p}\right)\right]$$

and :

$$\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right] = \frac{\tau}{n\psi_n(x,a_n)} \sum_{k=1}^{n} \mathbb{E}\left[\frac{1}{G(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right)\right]$$

Then, the proof of Theorem 3.2.1 is a direct consequence of the fowling Lemmas extending several results to the left-truncation setting.

Lemma 3.2.1. *Under the assumptions (H.1), (H.2), (H.3) and (H.6), we have :*

$$\lim_{n\to\infty} \sup_{x\in\mathcal{S}} \sup_{y\in\mathcal{I}} \frac{\left|\widetilde{\Psi}_n(x,y) - \mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right]\right|}{\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}} \leq K_3 \qquad a.s.$$

Lemma 3.2.2. *Under the assumptions (H.2), (H.4), (H.5) and (H.7) one get :*

$$\lim_{n\to\infty} \sup_{x\in\mathcal{S}} \sup_{y\in\mathcal{I}} \frac{\left|\widehat{\Psi}_n(x,y) - \widetilde{\Psi}_n(x,y)\right|}{(n^{-1/2})} \leq K_4$$

Lemma 3.2.3. *Assume that (H.1) and (H.4) hold true, for any $x \in S$, we have :*

$$\lim_{n\to\infty} \sup_{x\in\mathcal{S}} \frac{\left|\widetilde{\Upsilon}_n(x) - \mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]\right|}{\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}} \leq K_5 \qquad a.s.$$

Lemma 3.2.4. *Assume that (H.2), (H.4) and (H.7) hold true, for any $x \in S$, one get :*

$$\limsup_{n\to\infty} \sup_{x\in\mathcal{S}} \frac{\left|\widehat{\Upsilon}_n(x) - \widetilde{\Upsilon}_n(x)\right|}{(n^{-1/2})} \leq K_6$$

Lemma 3.2.5. *Under the assumptions (H.1), (H.3), (H.4) and (H.6), we have :*

$$\sup_{x\in\mathcal{S}} \sup_{y\in\mathcal{I}} \left|\widehat{B}_n(x,y)\right| = O\left(\sum_{k=1}^{n} a_k^{\nu_1}\right) + O\left(\sum_{k=1}^{n} b_k^{\nu_1}\right)$$

Proof of Corollary 3.2.1 It is easy to see that this Corollary can be deduced from the relation :

$$\sup_{x\in\mathcal{S}} \left|F^x\left(\widehat{q}_{\alpha,n}(x)\right) - F^x\left(q_\alpha(x)\right)\right| \leq 2 \sup_{x\in\mathcal{S}} \sup_{y\in\mathcal{I}} \left|\widehat{F}_n^x(y) - F^x(y)\right|$$

which is based primarily on the decomposition (3.4).

3.3 Particular cases

3.3.1 The real case (p=1)

We have previously studied the strong consistency of our estimator when the random variable of interest Y is of vector nature. It remains to treat the particular case where this variable is real (i.e. $p = 1$). In this case, some current assumptions will be modified to fit the situation considered, such as (H.3) and (H.5) becomes respectively (H.3) and (H.5) of the previous Chapter 2.

In fact, we will not repeat here the proofs which are the same as for the previously studied case and the result remains the same too, such that :

Corollary 3.3.1. *(**Benziadi** and **Bouazza** (2022)). Based on the same assumptions* *used in Theorem 3.2.1, in addition to (H.3) and (H.5) used in Theorem 2.2.1, one have :*

$$\limsup_{n\to\infty}\sup_{x\in\mathcal{S}}\sup_{y\in\mathcal{I}}\frac{\left|\widetilde{F}_n^x(y)-F^x(y)\right|}{\left(\left(\sum_{k=1}^{n}a_k^{\nu_1}+\sum_{k=1}^{n}b_k^{\nu_2}\right)+\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}\right)}\leq M_1 \qquad a.s.$$

3.3.2 The L^1 recursive estimate

For $x\in\mathcal{H}$, the L^1 estimator of the conditional probability distribution of Y given $X=x$ is given as follows :

$$\bar{F}_n^x(y)=\frac{\sum_{k=1}^{n}G_n^{-1}\left(Y_k\right)L_1\left(a_k^{-1}d_{\mathcal{H}}\left(x,X_k\right)\right)\mathbb{I}_{(-\infty,y)}\left(Y_k\right)}{\sum_{k=1}^{n}G_n^{-1}\left(Y_k\right)L_1\left(a_k^{-1}d_{\mathcal{H}}\left(x,X_k\right)\right)}:=\frac{\bar{\Psi}_n(x,y)}{\bar{\Upsilon}_n(x)} \tag{3.4}$$

where $\mathbb{I}_A$ denotes the indicator function of the set A.

Theorem 3.3.1. *(**Benziadi** and **Bouazza** (2022)) [3] Under the assumptions (H.1), (H.4), (H.6) and (H.3 hold in Theorem 2.2.1), one have :*

$$\limsup_{n\to\infty}\sup_{x\in\mathcal{S}}\sup_{y\in\mathcal{I}}\frac{\left|\bar{F}_n^x(y)-F^x(y)\right|}{\left(\left(\sum_{k=1}^{n}a_k^{\nu_1}\right)+\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}\right)}\leq M_2 \qquad a.s.$$

The proof of this theorem is based on the main following results.

Lemma 3.3.1. *Let assumptions of Theorem 3.2.1 hold true, then :*

(i) $$\limsup_{n\to\infty}\sup_{x\in\mathcal{S}}\sup_{y\in\mathcal{I}}\frac{\left|\ddot{\Psi}_n(x,y)-\mathbb{E}\left[\ddot{\Psi}_n(x,y)\right]\right|}{\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}}\leq M_3 \qquad a.s.$$

(ii) $$\limsup_{n\to\infty}\sup_{x\in\mathcal{S}}\sup_{y\in\mathcal{I}}\frac{\left|\bar{\Psi}_n(x,y)-\ddot{\Psi}_n(x,y)\right|}{(n^{-1/2})}\leq M_4$$

Lemma 3.3.2. *Let the assumptions (H.1), (H.4) and (H.6) hold. Then, one have :*

$$(i)\ \limsup_{n\to\infty}\sup_{x\in\mathcal{S}}\frac{\left|\ddot{\Upsilon}_n(x)-\mathbb{E}\left[\ddot{\Upsilon}_n(x)\right]\right|}{\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}}\leq M_5 \qquad a.s.$$

$$(ii)\ \limsup_{n\to\infty}\sup_{x\in\mathcal{S}}\frac{\left|\bar{\Upsilon}_n(x)-\ddot{\Upsilon}_n(x)\right|}{(n^{-1/2})}\leq M_6$$

Lemma 3.3.3. *Under the same assumptions as those of Lemma 3.2.5. then, we have :*

$$\sup_{x\in\mathcal{S}}\sup_{y\in\mathcal{I}}\left|\widetilde{B}_n(x,y)\right|=O\left(\sum_{k=1}^{n}a_k^{\nu_1}\right)$$

3.4 Technical proofs

In following, the proof of Theorem 3.2.1 is essentially based on the Fuk-Nagaev Inequality showed in Lemma 1.3.1 and the following lemma adapted to the α-mixing context.

Lemma 3.4.1. *(O. Said and Tatachak (2009)[29]). Under assumption (H.2) of mixing random variables, we have :*

$$|\widehat{\tau}_n-\tau|=O\left\{n^{-1/2}\left(\log_2 n\right)^{1/2}\right\}$$

Proof of Lemma 3.2.1 We start by noting for all couple $(x,y)\in\mathcal{S}\times\mathcal{I}$ that :

$$\begin{aligned}
&\widetilde{\Psi}_n(x,y)-\mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right]\\
&=\frac{1}{n}\sum_{k=1}^{n}\frac{\tau}{\psi_n(x,a_n)}\left\{\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right.\\
&\left.-\mathbb{E}\left[\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right]\right\}\\
&=\frac{1}{n}\sum_{k=1}^{n}Z_{k,n}(x,y).
\end{aligned}$$

with :

$$\begin{aligned} Z_{k,n}(x,y) =& \frac{\tau}{\psi_n(x,a_n)} \left\{ \frac{1}{G(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right) \right. \\ & \left. -\mathbb{E}\left[\frac{1}{G(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right]\right\} \end{aligned}$$

The use of the boundedness of L_1 from the assumption (H.4) ensures that for two existent constants $(c_1, c_2) \in \mathbb{R}_+^2$, we would have :

$$0 < c_1\phi(x,a_k) \leq \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right)\right] \leq c_2\phi(x,a_k) \tag{3.5}$$

In view of the following quantity, since the condition $\mathbb{I}_{(T_k \leq Y_k)} = 1$ is always validated in the left truncated model by definition of the probability τ. Then, by condition (3.5) and applying the assumption (H.5)(ii), we can write it in its simplest form, thus :

$$\begin{aligned} & \left|\mathbb{E}\left[\frac{1}{G(Y_k)} L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right) L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right]\right| \\ &= \left|\mathbb{E}\left[\mathbb{E}\left[\frac{1}{G(Y_k)} L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\mathbb{I}_{(T_k\leq Y_k)} \mid X\right] L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right)\right]\right| \\ &\leq \mathbb{E}\left[\frac{1}{G(Y_k)}\mathbb{E}\left[L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right) \mid X\right]\mathbb{P}\left[T_k \leq Y_k\right] L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right)\right] \\ &\leq l_\infty(x)\mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}(x,X_k)\right)\right] \\ &\leq l_\infty(x) c_2 \phi(x,a_k). \end{aligned}$$

and so, we employ the decomposition :

$$\begin{aligned} \sup_{x\in\mathcal{S}}\sup_{y\in\mathcal{I}}\left|\sum_{k=1}^n Z_{k,n}(x,y)\right| \leq & \underbrace{\sup_{x\in\mathcal{S}}\sup_{y\in\mathcal{I}}\left|\sum_{k=1}^n Z_{k,n}^*(x_i,y)\right|}_{Q_1} \underbrace{\max_{j\in\{1,\ldots,r_n\}}\sup_{x\in\mathcal{S}}\left|\sum_{k=1}^n \widetilde{Z}_{k,n}(x,y)\right|}_{Q_2} \\ & + \underbrace{\max_{i\in\{1,\ldots,h_n\}}\max_{j\in\{1,\ldots,r_n\}}\left|\sum_{k=1}^n Z_{k,n}(x_i,y_j)\right|}_{Q_3} \end{aligned} \tag{3.6}$$

Therefore, the compactness property of the two subsets $\mathcal{I}$ and $\mathcal{S}$ help us to :

write for any $y_1, y_2, \ldots, y_{r_n}$ and $x_1, x_2, \ldots, x_{h_n}$,

$$\mathcal{I} \subset \bigcup_{j=1}^{r_n} \mathcal{B}(y_j, s_n) \text{ and } \mathcal{S} \subset \bigcup_{i=1}^{h_n} \mathcal{B}(x_i, s_n)$$

Thus, we can take for a constant $M, s_n \leq Mn^{-\beta}$ with $(\beta > 0)$ and $j(y) = \arg\min_{j\in\{1,\ldots,r_n\}} \|y - y_j\|_{\mathbb{R}^p}$ and $h(x) = \arg\min_{i\in\{1,\ldots,h_n\}} d_{\mathcal{H}}(x, x_i)$.

For the first term of the decomposition (3.6), we have for any $(x, y) \in \mathcal{S} \times \mathcal{I}$:

$$\begin{aligned}
\left|\sum_{k=1}^{n} Z^*_{k,n}(x_i, y)\right| &\leq \left|\widetilde{\Psi}_n(x,y) - \widetilde{\Psi}_n(x_i,y)\right| + \left|\mathbb{E}\left[\widetilde{\Psi}_n(x_i,y)\right] - \mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right]\right| \\
&\leq \frac{\tau}{\psi_n(x,a_n)} \sum_{k=1}^{n} \frac{L_2\left(b_k^{-1}\|y_-Y_k\|_{\mathbb{R}^p}\right)}{G(Y_k)} \Bigg| L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) \\
&- L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) | \\
&+ \frac{\tau}{\psi_n(x,a_n)} \sum_{k=1}^{n} \mathbb{E}\left[\frac{L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right)}{G(Y_k)} \Bigg| L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right)\right. \\
&\left. -L_1\left(a_k^{-1} d_{\mathcal{H}}(x, X_k)\right) |\right]
\end{aligned}$$

Using the fact that the kernel L_1 is of *Lipschitz* class. Then, one get :

$$\begin{aligned}
Q_1 \leq & C_2 \frac{\tau}{G(a_F)\psi_n(x,a_n)} \sup_{x\in\mathcal{S}} \sup_{y\in\mathcal{I}} \sum_{k=1}^{n} \frac{d_{\mathcal{H}}(x,x_i)}{a_k} \left|L_2\left(b_k^{-1}\|y_-Y_k\|_{\mathbb{R}^p}\right)\right| \\
&+ C_2 \frac{\tau}{G(a_F)\psi_n(x,a_n)} \sup_{x\in\mathcal{S}} \sup_{y\in\mathcal{I}} \sum_{k=1}^{n} \frac{d_{\mathcal{H}}(x_i,x)}{a_k} \mathbb{E}\left[L_2\left(b_k^{-1}\|y_-Y_k\|_{\mathbb{R}^p}\right)\right]
\end{aligned}$$

and since for all $s_n = n^{-\beta}$, it follows that $Q_1 \longrightarrow 0$ as $n \to \infty$.

Then, for the study of Q_2, we first write the following decomposition which leads to :

$$\begin{aligned}
Q_2 = & \max_{j\in\{1,\ldots,r_n\}} \sup_{x\in\mathcal{S}} \left|\widetilde{\Psi}_n(x_i,y) - \widetilde{\Psi}_n(x_i,y_j)\right| \\
&+ \max_{j\in\{1,\ldots,r_n\}} \sup_{x\in\mathcal{S}} \left|\mathbb{E}\left[\widetilde{\Psi}_n(x_i,y_j)\right] - \mathbb{E}\left[\widetilde{\Psi}_n(x_i,y)\right]\right| \\
\leq & \max_{j\in\{1,\ldots,r_n\}} \sup_{x\in\mathcal{S}} \sum_{k=1}^{n} \frac{\tau}{\psi_n(x,a_n)} \left\{\frac{1}{G(Y_k)}\right| L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y - Y_k\|_{\mathbb{R}^p}\right)
\end{aligned}$$

$$
\begin{aligned}
&-L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x_i, X_k\right)\right) L_2\left(b_k^{-1}\left\|y_j-Y_k\right\|_{\mathbb{R}^p}\right) \mid\Big\} \\
&+\max_{j \in\{1, \ldots, r_n\}} \sup_{x \in \mathcal{S}} \sum_{k=1}^n \frac{\tau}{\psi_n\left(x, a_n\right)}\left\{\mathbb{E}\left[\left.\frac{1}{G\left(Y_k\right)}\right| L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x_i, X_k\right)\right) L_2\left(b_k^{-1}\left\|y_j-Y_k\right\|_{\mathbb{R}^p}\right)\right.\right. \\
&\left.\left.-L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x_i, X_k\right)\right) L_2\left(b_k^{-1}\left\|y-Y_k\right\|_{\mathbb{R}^p}\right) \mid\right]\right\} \\
&=Q_2^{(1)}+Q_2^{(2)}
\end{aligned}
$$

So that under the assumptions (H.5) and (H.6)(iii), we have :

$$
\begin{aligned}
Q_2^{(1)} & \leq C_3 \frac{\left\|y-y_j\right\|_{\mathbb{R}^p}}{\psi_n\left(x, a_n\right)} \sum_{k=1}^n \frac{\tau}{b_k G\left(Y_k\right)} L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x_i, X_k\right)\right) \\
& \leq C_3 \frac{\tau s_n}{G\left(a_F\right) \psi_n\left(x, a_n\right)} \sum_{k=1}^n \frac{L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x_i, X_k\right)\right)}{b_k} \\
& \leq \frac{M_1 n^{-\gamma}}{\psi_n\left(x, a_n\right)} \sum_{k=1}^n b_k^{-1} \\
& \leq \frac{M_1 \log n}{\psi_n\left(x, a_n\right)} \frac{1}{n^\gamma \log n} \sum_{k=1}^n b_k^{-1} \longrightarrow 0 \text { as } n \rightarrow \infty
\end{aligned} \tag{3.7}
$$

For the second term of the decomposition, the same arguments as for $Q_2^{(1)}$ with condition (3.5) lead as n goes to infinity to :

$$
Q_2^{(2)} \leq C_3\left[\psi_n\left(x, a_n\right)\right]^{-1} \sum_{k=1}^n \frac{\left\|y_j-y\right\|_{\mathbb{R}^p}}{b_k} \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x_i, X_k\right)\right)\right] \longrightarrow 0
$$

We move now to the last term :

$$
Q_3=\max_{i \in\{1, \ldots, h_n\}} \max_{j \in\{1, \ldots, r_n\}}\left|\sum_{k=1}^n Z_{k, n}\left(x_i, y_j\right)\right|
$$

At first we calculate :

$$
S_n^2=\underbrace{\sum_{k \neq l}\left|\operatorname{Cov}\left(Z_{k, n}\left(x_i, y_j\right), Z_{l, n}\left(x_i, y_j\right)\right)\right|}_{S_n^{Cov}}+\underbrace{\sum_{k=l}\left|\operatorname{Cov}\left(Z_{k, n}\left(x_i, y_j\right), Z_{k, n}\left(x_i, y_j\right)\right)\right|}_{S_n^{Var}} \tag{3.8}
$$

The definition of the probability τ and because of the boundedness of the kernels L_1 and L_2, one can show that $Z_{k,n}$ really satisfies the condition $\left|Z_{k,n}\left(x_i, y_j\right)\right| < \infty$. In particular, it is bounded $\forall k \in \mathbb{N}$, such that :

$$\begin{aligned}
|Z_{k,n}(x_i, y_j)| \leq & \frac{\tau}{G(a_F)\psi_n(x,a_n)} \Big| L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right) \\
& - \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right)\right] | \\
\leq & C \frac{\tau}{G(a_F)\psi_n(x,a_n)} \\
= & O\left(\frac{1}{\phi(x,a_k)}\right).
\end{aligned}$$

Then, the linearity of the expectation with the standard Jensen inequality lead directly to :

$$\begin{aligned}
& |\mathbb{E}[Z_{k,n}(x_i, y_j)]| \\
& \leq 2\mathbb{E}\left[\frac{\tau}{G(Y_k)\psi_n(x,a_n)}\left|L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right)\right|\right] \\
& \leq 2\frac{\tau}{G(a_F)}\frac{1}{\psi_n(x,a_n)}\mathbb{E}\left[\left|L_1\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right)\right|\right] \\
& = O(1)
\end{aligned} \tag{3.9}$$

and,

$$\begin{aligned}
& \mathbb{E}\left[Z_{k,n}^2(x_i, y_j)\right] \\
& \leq C\frac{1}{G(a_F)\psi_n^2(x,a_n)}\mathbb{E}\left[\frac{\tau}{G(Y_k)}L_1^2\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right) L_2^2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right)\right] \\
& \leq C\frac{1}{G(a_F)\psi_n^2(x,a_n)}\mathbb{E}\left[L_1^2\left(a_k^{-1} d_{\mathcal{H}}(x_i, X_k)\right)\mathbb{E}\left(L_2^2\left(b_k^{-1}\|y_j - Y_k\|_{\mathbb{R}^p}\right) \mid X\right)\right] \\
& \leq l_\infty(x)\frac{C}{G(a_F)}\frac{1}{\phi(x,a_k)} \leq O\left(\frac{1}{\phi(x,a_k)}\right).
\end{aligned} \tag{3.10}$$

furthermore,

$$\begin{aligned}
&\left|\mathbb{E}\left[Z_{k,n}\left(x_i,y_j\right)\cdot Z_{l,n}\left(x_i,y_j\right)\right]\right| \\
&=\left|\mathbb{E}\left[\frac{\tau}{G\left(Y_k\right)G\left(Y_l\right)\psi_n^2\left(x,a_n\right)}L_1\left(a_k^{-1}d_{\mathcal{H}}\left(x_i,X_k\right)\right)L_2\left(b_k^{-1}\left\|y_j-Y_k\right\|_{\mathbb{R}^p}\right)\right.\right. \\
&\quad\left.\times L_1\left(a_l^{-1}d_{\mathcal{H}}\left(x_i,X_l\right)\right)L_2\left(b_l^{-1}\left\|y_j-Y_l\right\|_{\mathbb{R}^p}\right)\right]\Big| \\
&\leq C\frac{\tau}{G^2\left(a_F\right)\psi_n^2\left(x,a_n\right)}\mathbb{E}\left[\left(L_1\left(a_k^{-1}d_{\mathcal{H}}\left(x_i,X_k\right)\right)L_1\left(a_l^{-1}d_{\mathcal{H}}\left(x_i,X_l\right)\right)\right)\right. \\
&\qquad\qquad\left.\times\left|\mathbb{E}\left[L_2\left(b_k^{-1}\left\|y_j-Y_k\right\|_{\mathbb{R}^p}\right)L_2\left(b_l^{-1}\left\|y_j-Y_l\right\|_{\mathbb{R}^p}\right)\mid X_k,X_l\right]\right|\right] \\
&\leq C\frac{\tau}{G^2\left(a_F\right)\psi_n^2\left(x,a_n\right)}l_\infty(x)\mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}\left(x_i,X_k\right)\right)L_1\left(a_l^{-1}d_{\mathcal{H}}\left(x_i,X_l\right)\right)\right]
\end{aligned}$$

then, by assumptions (H.1)(ii), (H.4) and condition (3.3), one get :

$$\begin{aligned}
\left|\mathbb{E}\left[Z_{k,n}\left(x_i,y_j\right)\cdot Z_{l,n}\left(x_i,y_j\right)\right]\right| &\leq C\frac{\tau}{G^2\left(a_F\right)\psi_n^2\left(x,a_n\right)}l_\infty(x)\left(\frac{\left(\phi\left(x,a_k\right)\right)^{1+1/a}}{n^{1/a}}\right) \\
&\leq C\frac{\tau}{G^2\left(a_F\right)}l_\infty(x)\left[\frac{\phi\left(x,a_k\right)}{n}\right]^{1/a}\cdot\frac{1}{\phi\left(x,a_k\right)} \qquad (3.11)
\end{aligned}$$

in which we deduce on the one hand from (3.9) and (3.11)

$$\begin{aligned}
&\left|\operatorname{Cov}\left(Z_{k,n}\left(x_i,y_j\right),Z_{l,n}\left(x_i,y_j\right)\right)\right| \\
&\leq\left|\mathbb{E}\left[Z_{k,n}\left(x_i,y_j\right)\cdot Z_{l,n}\left(x_i,y_j\right)\right]\right|+\left(\mathbb{E}\left(Z_{k,n}\left(x_i,y_j\right)\right)\right)^2 \\
&=C\left\{\left(\frac{\phi\left(x,a_k\right)}{n}\right)^{1/a}\cdot\frac{1}{\phi\left(x,a_k\right)}\right\}+1. \qquad (3.12)
\end{aligned}$$

In the other hand, applying Lemma 1.3.2 the usual modified Davydov-Rio's covariance inequality for the mixing processes. $\forall k\neq l$, we have :

$$\begin{aligned}
\left|\operatorname{Cov}\left(Z_{k,n}\left(x_i,y_j\right),Z_{l,n}\left(x_i,y_j\right)\right)\right| &\leq 4\left\|Z_{k,n}\left(x_i,y_j\right)\right\|\left\|Z_{l,n}\left(x_i,y_j\right)\right\| \\
&\leq C|k-l|^{-a} \qquad (3.13)
\end{aligned}$$

It is therefore useful to set from now on the two subsets :

$$J_1 = \{(k,l); 0 < |k-l| \leq \mu_n\} \quad \text{and} \quad J_2 = \{(k,l); \mu_n < |k-l| \leq n-1\}$$

A simply combination between (3.12) and (3.13) allows us to have the following :

$$\begin{aligned} S_n^{Cov} &= \sum_{J_1}\sum |\mathrm{Cov}\left(Z_{k,n}\left(x_i, y_j\right), Z_{l,n}\left(x_i, y_j\right)\right)| + \sum_{J_2}\sum |\mathrm{Cov}\left(Z_{k,n}\left(x_i, y_j\right), Z_{l,n}\left(x_i, y_j\right)\right)| \\ &\leq \mathrm{Cn}\, \mu_n \left\{ \left(\frac{\phi\left(x, a_k\right)}{n}\right)^{1/a} \cdot \frac{1}{\phi\left(x, a_k\right)} + 1 \right\} + Cn^2 \mu_n^{-a} \end{aligned} \tag{3.14}$$

For the variance term, we apply the general definition and we get :

$$\begin{aligned} |\mathrm{Var}\left[Z_{k,n}\left(x_i, y_j\right)\right]| &\leq \mathbb{E}\left[Z_{k,n}^2\left(x_i, y_j\right)\right] + \mathbb{E}\left[Z_{k,n}\left(x_i, y_j\right)\right]^2 \\ &= O\left\{\frac{1}{\phi\left(x, a_k\right)}\right\} \end{aligned}$$

thus :

$$S_n^{Var} = O\left\{\frac{n}{\phi\left(x, a_k\right)}\right\} \tag{3.15}$$

It follows from (3.14) and (3.15) that :

$$S_n^2 = O\left\{ n\mu_n \left(\left(\frac{\phi\left(x, a_k\right)}{n}\right)^{1/a} \cdot \frac{1}{\phi\left(x, a_k\right)} + 1 \right) + n^2 \mu_n^{-a} \right\} + O\left\{\frac{n}{\phi\left(x, a_k\right)}\right\} \tag{3.16}$$

The complementarity of this proof depends primarily on the choice of the sequence μ_n. We put $\mu_n = \left(\frac{\phi(x,a_k)}{n}\right)^{-1/a}$, thus, we will have $S_n^2 = O\left\{\frac{n}{\phi(x,a_k)}\right\}$.
At this stage of the proof, we use the Fuk-Nagaev inequality adapted to the $\alpha-$ mixing

context for $\theta = \theta_0 \left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}$ yields :

$$\begin{aligned}
&\mathbb{P}\left\{\left|\sum_{k=1}^{n} Z_{k,n}\left(x_i, y_j\right)\right| > 4\left(\frac{n\theta}{4}\right)\right\} \\
&\leq 4\left(1+\frac{\theta^2}{qS_n^2}\right)^{-q/2} + \frac{2nc}{q}\left(\frac{2q}{\theta}\right)^{a+1} \\
&\leq 4\left(1+\frac{\frac{n^2\theta_0^2\log n}{16n\phi(x,a_k)}}{q\frac{n}{\phi(x,a_k)}}\right)^{-q/2} + \frac{2nc}{q}\left(\frac{8q\left(\frac{\log n}{n\phi(x,a_k)}\right)^{-1/2}\Big)^{a+1}}{n\theta_0}\right)^{-q/2} \\
&\leq 4\left(1+\frac{\theta_0^2\log n}{16q}\right)^{-q}\Bigg) + \frac{2nc}{q}\left(\frac{8q\left(\phi\left(x,a_k\right)\right)^{1/2}}{\theta_0\sqrt{n\log n}}\right)^{a+1} \\
&= Q_4 + Q_5
\end{aligned}$$

We see here that the preferred choice of q is $\log^2 n$, such that the first term in the right hand side is thus increased by :

$$Q_4 \leq cn^{-\frac{\theta_0^2}{32}} \longrightarrow 0$$

moreover, for the second term and by the same choice of q we also find as $n \to \infty$ the following :

$$Q_5 \leq cn^{\frac{-1}{2(1-a)}\left(-4a^2+a+1\right)} \longrightarrow 0$$

Hence, both (Q_4) and (Q_5) fall on the following result :

$$\mathbb{P}\left[A_3 > 4\left(\frac{n\theta}{4}\right)\right] \leq r_n h_n \max_{i\in\{1,\ldots,h_n\}} \max_{j\in\{1,\ldots,r_n\}} \mathbb{P}\left[\left|\sum_{k=1}^{n} Z_{k,n}\left(x_i, y_j\right)\right| > 4\left(\frac{n\theta}{4}\right)\right]$$

which give us for an appropriate choice of θ_0 :

$$\sum_{n\geq 1} \mathbb{P}\left\{\max_{i\in\{1,\ldots,h_n\}} \max_{j\in\{1,\ldots,r_n\}} \left|\sum_{k=1}^{n} Z_{k,n}\left(x_i, y_j\right)\right| > n\theta\right\} < \infty \tag{3.17}$$

Proof of Lemma 3.2.2 We have form the definition of the estimators $\widehat{\Psi}$ and $\widetilde{\Psi}$:

$$\begin{aligned}
&\widehat{\Psi}_n(x,y)-\widetilde{\Psi}_n(x,y)\\
&=\frac{\widehat{\tau}_n}{n\psi_n(x,a_n)}\sum_{k=1}^n\frac{1}{G_n(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\\
&-\frac{\tau}{n\psi_n(x,a_n)}\sum_{k=1}^n\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right).
\end{aligned}$$

such that :

$$\begin{aligned}
&\left|\widehat{\Psi}_n(x,y)-\widetilde{\Psi}_n(x,y)\right|\\
&\leq\frac{1}{n\psi_n(x,a_n)}\left\{|\widehat{\tau}_n-\tau|\sum_{k=1}^n\frac{1}{G_n(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right.\\
&\left.+\tau\sum_{k=1}^n\left|\frac{G_n(Y_k)-G(Y_k)}{G(Y_k)G_n(Y_k)}\right|L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right\}\\
&\leq\frac{1}{n\psi_n(x,a_n)}\left\{\frac{|\widehat{\tau}_n-\tau|}{G_n(a_F)}\sum_{k=1}^n L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right.\\
&\left.+\frac{\tau}{G_n(a_F)}\left|\frac{G_n(y)-G(y)}{G(a_F)}\right|\sum_{k=1}^n L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right\}
\end{aligned}$$

Thus :

$$\sup_{x\in\mathcal{S}}\sup_{y\in\mathcal{I}}\left|\widehat{\Psi}_n(x,y)-\widetilde{\Psi}_n(x,y)\right|\leq\frac{1}{G_n(a_F)}\left\{|\widehat{\tau}_n-\tau|+\tau\frac{\sup_{y\geq a_F}|G_n(y)-G(y)|}{G(a_F)}\right\}$$
$$\sup_{x\in\mathcal{S}}\sup_{y\in\mathcal{I}}|\Psi_n^*(x,y)|$$

with :

$$\Psi_n^*(x,y)=\frac{1}{n\psi_n(x,a_n)}\sum_{k=1}^n L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)$$

Recall the Lemma 3.4.1, when the process (X_k,Y_k) has a decreasing mixing coefficients $\alpha(n)$, such that :

$$|\widehat{\tau}_n-\tau|=O_{\text{a.s.}}\left\{n^{-1/2}(\log_2 n)^{1/2}\right\}$$

and the direct application of *Remark 6* in **Woodroofe** (1985)[35] which gives :

$$|G_n(a_F)-G(a_F)|=O\left\{n^{-1/2}\right\}$$

Thus, the result is an immediate consequence of what has already been mentioned.

Proof of Lemma 3.2.3 Keeping the same conditions concerning the compactness of $\mathcal{S}$, almost certainly identical as in Lemma 3.2.1 and we decompose the studied quantity as follows :

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{x\in\mathcal{S}}\left|\widetilde{\Upsilon}_n(x)-\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]\right|>3\eta\right\} \\
&\leq \underbrace{\mathbb{P}\left\{\sup_{x\in\mathcal{S}}\left|\widetilde{\Upsilon}_n(x)-\widetilde{\Upsilon}_n\left(x_i\right)\right|>\eta\right\}}_{I_1}+\underbrace{\mathbb{P}\left\{\sup_{x\in\mathcal{S}}\left|\widetilde{\Upsilon}_n\left(x_i\right)-\mathbb{E}\left[\widetilde{\Upsilon}_n\left(x_i\right)\right]\right|>\eta\right\}}_{I_2} \\
&\quad+\underbrace{\mathbb{P}\left\{\sup_{x\in\mathcal{S}}\left|\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]-\mathbb{E}\left[\widetilde{\Upsilon}_n\left(x_i\right)\right]\right|>\eta\right\}}_{I_3} \qquad (3.18)
\end{aligned}$$

For the first term of the decomposition (3.18), L_1 being a Lipschitzian kernel. In addition, $s_n=O\left\{n^{-\beta}\right\}$; implies that :

$$\begin{aligned}
&\sup_{x\in\mathcal{S}}\left|\widetilde{\Upsilon}_n(x)-\widetilde{\Upsilon}_n\left(x_i\right)\right| \\
&\leq \sup_{x\in\mathcal{S}}\sum_{k=1}^{n}\frac{\tau}{\psi_n\left(x,a_n\right)}\frac{1}{G\left(Y_k\right)}\left|L_1\left(a_k^{-1}d_{\mathcal{H}}\left(x,X_k\right)\right)-L_1\left(a_k^{-1}d_{\mathcal{H}}\left(x_i,X_k\right)\right)\right| \\
&\leq \frac{\tau}{G\left(a_F\right)\psi_n\left(x,a_n\right)}\sum_{k=1}^{n}\frac{d_{\mathcal{H}}\left(x,x_i\right)}{a_k} \\
&\leq M\frac{s_n}{\psi_n\left(x,a_n\right)}\sum_{k=1}^{n}\frac{1}{a_k}\longrightarrow 0 \quad \text{as} \quad n\longrightarrow\infty
\end{aligned}$$

and the same for I_3 : $I_3\to 0$ as $n\to\infty$ and therefore we deal with :

$$I_1 \overset{\text{a.s}}{=} O\left\{\left(\frac{\log n}{n\psi_n\left(x,a_n\right)}\right)^{1/2}\right\} \quad \text{and} \quad I_3 \overset{\text{a.s}}{=} O\left\{\left(\frac{\log n}{n\psi_n\left(x,a_n\right)}\right)^{1/2}\right\}$$

We now move on to the second term I_2, we appeal again to the Bernstein's type inequality adapted to this context of dependence by taking $\eta=\eta_0\left(\frac{\log n}{n\psi_n(x,a_n)}\right)^{1/2}>0$

$$\begin{aligned}
I_2 &= \mathbb{P}\left\{\sup_{x\in\mathcal{S}}\left|\widetilde{\Upsilon}_n(x_i) - \mathbb{E}\left[\widetilde{\Upsilon}_n(x_i)\right]\right| > \eta\right\} \\
&\leq \sum_{i=1}^{h_n}\mathbb{P}\left\{\left|\widetilde{\Upsilon}_n(x_i) - \mathbb{E}\left[\widetilde{\Upsilon}_n(x_i)\right]\right| > \eta\right\} \\
&\leq h_n \max_{1\leq i\leq h_n}\mathbb{P}\left\{\left|\widetilde{\Upsilon}_n(x_i) - \mathbb{E}\left[\widetilde{\Upsilon}_n(x_i)\right]\right| > \eta\right\}
\end{aligned}$$

where, we have for all $k \in \mathbb{N}$:

$$\begin{aligned}
\widetilde{\Upsilon}_n(x_i) - \mathbb{E}\left[\widetilde{\Upsilon}_n(x_i)\right] =& \frac{1}{n}\sum_{k=1}^{n}\frac{\tau}{\psi_n(x,a_n)}\left\{\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)\right. \\
&\left. -\mathbb{E}\left[\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)\right]\right\} \\
=& \frac{1}{n}\sum_{k=1}^{n}\Lambda_{k,n}(x_i)
\end{aligned}$$

with :

$$\begin{aligned}
\Lambda_{k,n}(x_i) =& \frac{\tau}{\psi_n(x,a_n)}\left\{\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)\right. \\
&\left. -\mathbb{E}\left[\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x_i,X_k)\right)\right]\right\}.
\end{aligned}$$

For the reminder of this proof, the same steps as term (Q_3) in the proof of Lemma 3.2.1 are followed, in which under assumption (H.4), one can check that $\Lambda_{k,n}$ satisfies the condition of Lemma 1.3.1, such that

$$|\Lambda_{k,n}(x_i)| \leq M\frac{\tau}{G(a_F)\,\phi(x,a_k)} = O\left\{\frac{1}{\phi(x,a_k)}\right\}$$

Furthermore, we deduce that :

$$|\mathrm{Cov}(\Lambda_{k,n}(x_i),\Lambda_{l,n}(x_i))| \leq |\mathbb{E}[\Lambda_{k,n}(x_i)\Lambda_{l,n}(x_i)]| = O\left\{\frac{(\phi(x,a_k))^{-1+1/a}}{n^{1/a}}\right\} \tag{3.19}$$

and that :

$$|\mathrm{Var}[\Lambda_{k,n}(x_i)]| \leq \left|\mathbb{E}\left[\Lambda_{k,n}^2(x_i)\right]\right| = O\left\{\frac{1}{\phi(x,a_k)}\right\} \tag{3.20}$$

Finally, combining (3.19) with (3.20) and following some additional classical calculations, we get $S_n^2 = O\left\{\frac{n}{\phi(x,a_k)}\right\}$. Therefore, a direct application of Fuk-Nagaev exponential inequality makes it possible to deduce the proof.

Proof of Lemma 3.2.4 Similarly to the proof of Lemma 3.2.2, one may follows the same lines and arguments, such that :

$$
\begin{aligned}
\sup_{x\in\mathcal{S}}\left|\widehat{\Upsilon}_n(x)-\widetilde{\Upsilon}_n(x)\right| &\leq \frac{1}{n\psi_n(x,a_n)}\sup_{x\in\mathcal{S}}\left\{\widehat{\tau}_n\sum_{k=1}^{n}\frac{1}{G_n(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)\right.\\
&\left.-\tau\sum_{k=1}^{n}\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)\right\}\\
&\leq \frac{1}{G_n(a_F)}\left\{|\widehat{\tau}_n-\tau|+\tau\frac{\sup_{y\geq a_F}|G_n(y)-G(y)|}{G(a_F)}\right\}\sup_{x\in\mathcal{S}}|\Upsilon_n^*(x)|
\end{aligned}
$$

with,

$$
\Upsilon_n^*(x)=\frac{1}{n\psi_n(x,a_n)}\sum_{k=1}^{n}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right).
$$

Again, a direct application of Lemma 3.4.1 with *Remark 6* in **Woodroofe** (1985)[35] complete the proof.

Proof of Lemma 3.2.5 According to the definition of the bias term $\widehat{B}_n(x,y)$ above, which is not affected by the dependence condition. We use the fact that $\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]$ is bounded. Then, we can rewrite it in the following form :

$$
\widehat{B}_n(x,y)=\frac{\mathbb{E}\left[\mathbb{E}\left(\widetilde{\Psi}_n(x,y)-F^x(y)\mid X\right)\widetilde{\Upsilon}_n(x)\right]}{\mathbb{E}\left[\widetilde{\Upsilon}_n(x)\right]}
$$

As we have already proved in previous steps the fact that this writing is satisfied by definition of the probability of trancature :

$$
\begin{aligned}
\mathbb{E}\left[\widetilde{\Psi}_n(x,y)\right] &= \frac{\tau}{n\psi_n(x,a_n)}\sum_{k=1}^{n}\mathbb{E}\left[\frac{1}{G(Y_k)}L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right]\\
&= \frac{\tau}{n\psi_n(x,a_n)}\sum_{k=1}^{n}\mathbb{E}\left[L_1\left(a_k^{-1}d_{\mathcal{H}}(x,X_k)\right)L_2\left(b_k^{-1}\|y-Y_k\|_{\mathbb{R}^p}\right)\right]
\end{aligned}
$$

Therefore, by a conditioning to X_k we have :

$$\left|\mathbb{E}\left(\widetilde{\Psi}_n(x,y) - F^x(y) \mid X=u\right)\right| \leq \frac{\tau}{n\psi_n\left(x,a_n\right)} \sum_{k=1}^{n} \mathbb{E}\left[L_1\left(a_k^{-1} d_{\mathcal{H}}\left(x,X_k\right)\right)\right.$$
$$\left.\times\left|\mathbb{E}\left[L_2\left(b_k^{-1}\left\|y-Y_k\right\|_{\mathbb{R}^p}\right) - F^x(y) \mid X=u\right]\right|\right].$$

Next, an integration by parts, a change of variable and because of condition (H.3), for any $u \in \mathcal{B}\left(x,a_k\right)$, we get :

$$\begin{aligned}
&\mid \mathbb{E}\Big[L_2\left(b_k^{-1}\left\|y-Y_k\right\|_{\mathbb{R}^p}\right) - F^x(y) \mid X=u\Big] \mid \\
&\leq \int_{\mathbb{R}^p} L_2^{(1)}\left(\|t\|_{\mathbb{R}^p}\right)\left|F^{(u)}\left(y-b_k t\right) - F^{(x)}(y)\right| dt \\
&\leq \int_{\mathbb{R}^p} L_2^{(1)}\left(\|t\|_{\mathbb{R}^p}\right)\left|F^{(u)}\left(y-b_k t\right) - F^{(u)}(y)\right| dt + \left|F^{(u)}(y) - F^{(x)}(y)\right|
\end{aligned}$$

$$\leq C_1 \int_{\mathbb{R}^p} L_2^{(1)}\left(\|t\|_{\mathbb{R}^p}\right)\left(a_k^{\nu_1} + \left|b_k\right|^{\nu_2}\|t\|_{\mathbb{R}^p}^{\nu_2}\right) dt$$

which completes the proof of this Lemma.

Chapter 4

Computational Study and Conclusion

4.1 Computational study

This section presents a brief numerical study assessing the performance of the proposed estimator. Firstly, we deal with simulated data to show how recursive estimator of the conditional quantile function behaves in functional and dependent contexts, comparing to the classical one. Secondly, we discuss about the sensitivity of a kernel estimator to the presence of incomplete data especially when dealing with the truncated and censored data.

4.1.1 Simulated data

We point out that the primary goal is to compare the efficiency of the double-kernel recursive method (RDKM) with the double-kernel approach (DKM) [3], based on the work of **Bouazza** (2021) [8], we first introduce the following nonparametric model for all $k = 1, \ldots, n$ when n = 100 and 500.

$$Y_k = R\left(X_k\right) + \epsilon_k \tag{4.1}$$

Where ϵ_k are r.v. independent of X and follow a normal mixture distribution, with λ takes $0.1, 0.5$ and 0.9 :

$$\epsilon_k \sim (1-\lambda) * \mathcal{N}(0,1) + \lambda * \mathcal{N}(4,5) \tag{4.2}$$

$(X_k)_{k=1,\ldots,n}$ are the functional data which are considered as a sinusoidal basis with five functional axes of the continuous functions from the interval $[0,1]$ to $\mathbb{R}$, generated by simul.far on **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 4.1

[3]Recall that the double kernel estimator $\hat{F}_n^x(.)$ of $F^x(.)$ is defined as follows [18]:

$$\hat{F}_n^x(y) = \frac{\sum_{k=1}^{n} L_1\left(a_n^{-1} d_{\mathcal{H}}\left(x, X_k\right)\right) L_2\left(b_n^{-1}\left(y - Y_k\right)\right)}{\sum_{k=1}^{n} L_1\left(a_n^{-1} d_{\mathcal{H}}\left(x, X_k\right)\right)}$$

where L_1 is a kernel function, L_2 is a cumulative distribution function, and a_n (resp. b_n) a sequence of positive real numbers.

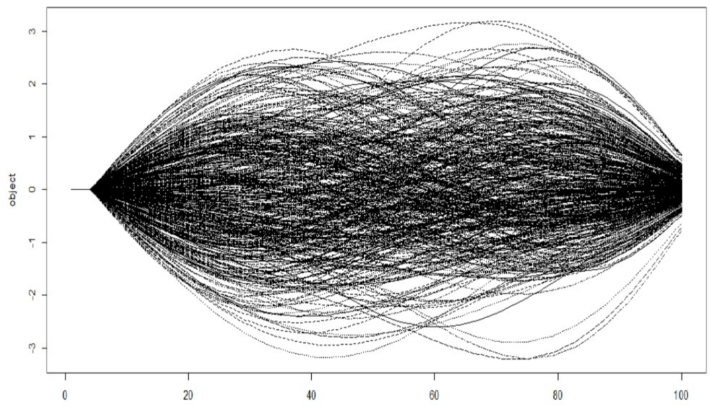

Figure 4.1: A sample of 100 curves

Furthermore, the response variables Y_k are obtained through the following operator:

$$R(x) = 5 \int_0^1 \exp(x(t))dt \tag{4.3}$$

This model enables to have 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 $R(x)$.

Besides, to a reasonable extent and compare the two methods, each is evaluated under optimal conditions with specified parameters. As no automatic, data-driven bandwidth selection method exists for estimating conditional quantiles with functional regressors, we adopt a bandwidth selector similar to that used by **Ferraty** and **Vieu** (2006) [15]. 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 :

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

where $q_{0.5}^{[-j]}\left(X_j, a_k, b_k\right)$ denotes the double-kernel recursive estimator of the conditional median in the curve X_j.

As an aside, for the kernel method we adapt the R-routine named funopare.quantile.lcv.

So the quadratic kernel chosen on $[0,1]$, is given by :

$$L_1(u) = \frac{3}{2}\left(1-u^2\right)\mathbb{I}_{[0,1]}$$

and the distribution function $L_2(.)$ is defined by

$$L_2(u) = \int_{-\infty}^{u}\frac{3}{4}\left(1-t^2\right)\mathbb{I}_{[-1,1]}(t)dt$$

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

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

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

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

Therefore, the obtained results of mean squared error are summarized in Table 4.1 and Table 4.2 for the two sample sizes $n = (100, 500)$, while, Figure 4.2 simultaneously plots side by side, the estimated conditional quantiles by the *RDKM* and the ones estimated by the *DKM*.

		$\lambda = 0.1$	$\lambda = 0.5$	$\lambda = 0.9$
MSE(DKM)	Q_1	4.9110	6.9100	8.9010
	Q_2	2.4420	4.4401	6.4321
	Q_3	4.2511	6.2501	2.2002
MSE(RDKM)	Q_1	2.1302	3.1340	4.1328
	Q_2	1.6920	2.6955	3.6921
	Q_3	2.2924	3.2945	4.3421

Table 4.1: Mean Squared Error Results for $n = 100$

		$\lambda = 0.1$	$\lambda = 0.5$	$\lambda = 0.9$
MSE(DKM)	Q_1	4.3900	5.7991	8.3864
	Q_2	1.9210	3.9267	5.9101
	Q_3	3.7373	3.7333	1.6823
MSE(RDKM)	Q_1	1.6101	2.6113	3.6753
	Q_2	1.1702	2.1787	3.1702
	Q_3	1.7777	2.7768	3.8294

Table 4.2: Mean Squared Error Results for $n = 500$

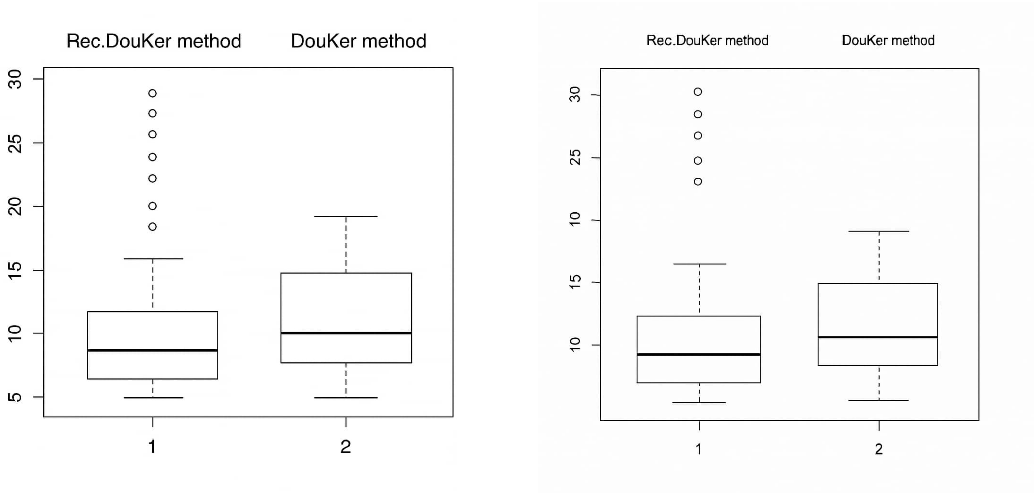

Figure 4.2: Conditional quantiles (Q1, Q2 and Q3) estimation by RDKM (on the left) versus DKM (on the right)

Results: Tables 4.1 and 4.2 present the Mean Squared Error (MSE) values for the estimated quartiles $Q_1(\alpha = 0.25), Q_2(\alpha = 0.5)$ and $Q_3(\alpha = 0.75)$. The simulation results support two key interpretations. First, both tables clearly demonstrate that the proposed recursive double kernel method outperforms the classical one in most of the examined scenarios, as also illustrated by Figure 4.2. Second, the MSE values tend to increase more noticeably (with respect to the parameter λ), in the classical kernel method compared to the recursive method. Additionally, the MSE decreases as the sample size n increases, further favoring the recursive approach.

4.1.2 Literature studies

Recognizing the prevalence of incomplete data and the need for reliable inference, a significant research explores adaptive statistical techniques. To contribute to this understanding, we present several studies that compare the performance of both classical and recursive double kernel estimators when faced with incomplete data, such as :

- ✓ **Helal** and **Ould Said** (2016)[21] provided significant simulation evidence for their kernel conditional quantile estimator in functional spaces with varying truncation levels. Their results clearly demonstrated that the mean squared error (MSE) decreased progressively with increasing sample size ($n = 100, 300, 500$), while the accuracy of the estimate worsened as the truncation percentage (TR) increased ($0\%, 12\%, 32\%, 66\%$).

- ✓ **Bouazza** et **al** (2021)[7] investigated the recursive estimator's effectiveness with incomplete (censored) dependent data in a semi-metric space for conditional mode estimation. Their simulations showed that the MSE decreased with larger sample sizes ($n = 200, 400, 600$), indicating improved estimator quality. Furthermore, they observed a slight decline in performance when censoring rates (CR) increased ($20\%, 40\%, 60\%$) compared to the complete data scenario.

4.2 Conclusion

In conducting this work, we acknowledge that the asymptotic properties of nonparametric conditional models associated with the recursive kernel approach have consistently been a focus of researchers; consequently, they occasionally presume that the sample under investigation comprises functional α mixing observations.

In order to achieve this work, we started with Chapter 1 to review the key principles and concepts related to statistics such as recursive method, dependent and incomplete data, as well as other features required for the research, namely truncated data.

The focus of Chapter 2 and 3 was then mostly on highlighting the challenges that arise when we attempt to deal with issues in particular truncated and dependent data, employing an estimator of the conditional distribution function and a quantile that is based on the recursive technique. Firstly, we use the recursive double-kernel approach to estimate the conditional distribution function for complete sample of random variables when the variable of interest is real. Secondly, by using the same approach we tried to estimate the conditional distribution function for an incomplete sample (truncated data) of variables when the variable of interest is a vector. In the meantime of the both cases, we performed the almost sure convergence with the estimators' rates.

Finally, Chapter 4 came to verify the theoretical findings for an infinite sample size and various truncation rates, we attempted to simulate at the end of this study in the purpose to examining the numerical behavior of the estimators.

Bibliography

[1] Baltagi, B. H. and Li, Q. “A simple recursive estimation method for linear regression models with AR(p) disturbances”. In: *Stat. Papers* (1994), "35, n 1, 93U100".

[2] Benziadi, F. “Recursive kernel estimate of the conditional quantile for functional ergodic data”. PhD thesis. University of Sidi Bel Abbes - Dr. Djillali Liabes, Algeria, 2016, "57, 13–92".

[3] Benziadi, F. and Bouazza, I. “Consistency result of recursive conditional distribution estimate for dependent data under left truncation, with applications to the conditional quantile”. In: *Bulletin of the Institute of Mathematics Academia Sinica (New Series)* (2022), "55–59, 53–81".

[4] Benziadi F., Laksaci A. and Tebboune, F. “Recursive kernel estimate of the conditional quantile for functional ergodic data”. In: *Comm in Statist-Theory and Methods* (2016), "45, No.11, 3097–3113".

[5] Bosq, D. “Nonparametric Statistics for Stochastic Processes: Estimation and Prediction (Second edition)”. In: *Lecture Notes in Statistics, Springer-Verlag,* (1998), "110".

[6] Bouadjemi, A. “Asymptotic Normality of the Recursive Kernel Estimate of Conditional Cumulative Distribution Function”. In: *Journal of Probability and Statistical Science* (2014), "12, 117–126".

[7] Bouazza I., Benziadi F. and Guendouzi, T. “Nonparametric recursive estimate for right-censored conditional mode function with ergodic functional data”. In: *J Indian Soc Probab Stat* (2021), "22, 389–415".

[8] Bouazza, I. “Estimation récursive de modèles non paramétriques pour des données ergodiques fonctionnelles”. PhD thesis. Université de saida- Dr. Mouly Taher, 2022, "7, 1–147".

[9] Chen K., Chao-M. T. and Lo, S. H. “On strong uniform consistency of the Lynden-Bell estimator for truncated data”. In: *The Annals of Statistics,* (1995), "23, 440–449".

[10] Djabbouri, T. “Statistique fonctionnelle de certaines caracteristiques sur les modeles de survie”. PhD thesis. Université de saida - Dr. Moulay Taher, 2016, "21, 10–90".

[11] Djillali, S. “On the consistency and the asymptotic normality of several conditional models, dependent case”. MA thesis. University Of Saida - Dr. Moulay Taher, 2018, "10–11, 5–63".

[12] Doukhan, P. “Mixing: Properties and examples”. In: *Lecture Notes in Statistics, Springer, Berlin* (1994), "85".

[13] Ferrany, Y. “Sur l’estimation non paramétrique de la densité et du mode dans les modèles de données incomplètes et associées”. PhD thesis. USTHB, Algeria, 2014, "7–10, 2–76".

[14] Ferraty, F. and Vieu, P. “Nonparametric methods for functional data, with applications in regression, time-series prediction and curves discrimination”. In: *J. Nonparametr. Statist.* (2004), "16, 11–126".

[15] Ferraty, F. and Vieu, P. *Nonparametric functional data analysis, Theory and practice.* Springer Science with Business Media, Inc., 2006, "30, 6–261".

[16] Ferraty F., Laksaci A. and Vieu, P. “Estimating Some Characteristics of the Conditional Distribution in Nonparametric Functional Models”. In: *Statistical Inference for Stochastic Processes* (2006), "9, 47–76".

[17] Ferraty F., Mas A. and Vieu, P. “Nonparametric regression on functional data: inference and practical aspects”. In: *Aust. N. Z. J. Stat.* (2007), "49, 267–286".

[18] Ferraty F., Rabhi A. and Vieu, P. “Conditional Quantiles for Functionally Dependent Data with Application to the Climatic El Niño Phenomeno”. In: *The Indian Journal of Statistics* (2005a), "6–16, Volume 67, Part 2, 378–398".

[19] He, S. and Yang, G. “NEstimation of the truncation probability in the random truncation model”. In: *Annals of Statistics* (1998), "26, 1011–1027".

[20] Helal, N. “Etude de la fonction des quantiles conditionnels pour des variables tronquees aleatoirement a gauche et a covariables fonctionnelles”. PhD thesis. University Of Sidi Bel Abbes - Dr. Djillali Liabes, Algeria, 2016, "14–16, 12–123".

[21] Helal, N. and O. Said, E. “Kernel conditional quantile estimator under left truncation for functional regressors”. In: *Opuscula Math* (2016), "36, 25–48".

[22] HONDA, T. “Nonparametric estimation of a conditional quantile for alpha-mixing processes”. In: *Ann. Inst. Statist. Math* (2000), "52, 459–470".

[23] Ian, L. “Introduction to Survival Analysis”. Medium, Applied Statistician. 2023.

[24] Keddani, L. “L’étude asymptotique d’estimation récursive pour des données fonctionnelles”. PhD thesis. Université de sada- Dr. Mouly Taher, 2020, "22, 1–60".

[25] Keddani L., Meghnafi-M. Rabhi A. and Sabri, K. “Recursive estimation for nonparametric estimators with functional data”. In: *International Journal of Mathematics and Computation* (2018), "2, 95–112".

[26] Lathan, L. “How well do you “really” know survival analysis?” Data Science. 2021.

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

[28] MASRY, E. and GYORFI, L. “The L1, and L2, Strong Consistency Kernel Density Estimation of Recursive from Dependent Samples”. In: *IELE Transactions on Information Theory* (1990), "3, 531–539".

[29] Oueld Said, E. and Tatachak, A. “Strong Consistency Rate for the Kernel Mode Estimator Under Strong Mixing Hypothesis and Left Truncation”. In: *Communications in Statistics - Theory and Methods* (2009), "38, No. 8, 1154–1169".

[30] Rio, E. “Théorie Asymptotique des Processus Aléatoires Dépendants”. In: *Mathématiques et Applications, Springer, Berlin* (2000), "31".

[31] Rosenbaltt, M. “A central limit theorem and a strong mixing condition”. In: *Proc. Nat. Acad. Sci. U. S. A* (1956b), "2, 43–47".

[32] Roussas, G. G. “Exact rates of almost sure convergence of a recursive kernel estimate of a probability densiy function: Application to regression and hazard rate estimation”. In: *Journal of Nonpara- metric Statistics* (1992), "1, 171–195".

[33] Stute, W. and Wang, J. L. “The strong law under random censorship”. In: *Ann. Statist.* (1993), "21, 1591–1607".

[34] Wolverton, C. and Wagner, T. J. “Recursive estimates of probability densities”. In: *IEEE Trans* (1969), "5, 307–308".

[35] Woodroofe, M. “Estimating a distribution function with truncated data”. In: *Annals of Statistics* (1985), "13, 163–177".