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

Université Frères Mentouri Constantine
Faculté des sciences exactes
Département de mathématiques

N° d'ordre: ....
N° de série: ....

THÈSE

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

"Estimation non paramètrique des données fonctionnelles et censurées"
"Nonparametric estimation of functional and censored data"

Par Farid Leulmi

OPTION : Probabilités et Statistique

Devant le jury :

Président	M.	T. Hamaizia	Prof.	Université frères Mentouri, Constantine 1
Directrice de thèse	Mme	S. Leulmi	M.C.A.	Université frères Mentouri, Constantine 1
Examinatrice	Mme	Kharfouchi	Prof.	Université Salah Boubnider, Constantine 3
Examinatreur	M.	M. Boukeloua	M.C.A.	Ecole Polytechnique de Constantine

Soutenue le: 19/01/2023.

Acknowledgment

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

I wish to sincerely thank Madam S. Leulmi as my supervisor and sister. I know it is a cliche to say " I wouldn't be here without my supervisor" but it is 100 % true in this situation. I am extremely grateful to she for it invaluable advice, continous support and patience during my doctorat study. Its immense knowledge and plantiful experience have encouraged me in all the time of my scientific researc..... Frankly I do not think these little words grant her all what she deserves.

Apart from my supervisor, I won't forget to express the gratitude to Madam S. Kharfouchi, my former supervisor and a member of the jury, for giving the encoragement and sharing insightful suggestions. She has played a major role in polishing my research writing skills. Her endless guidance is hard to forget throughout my life.

I extend my sincere thanks to Mr. T. Hamaizia for having accepted to present the jury, and for all the good times we shared during my thesis's years.

I am olso pleased to say thank you to Mr. M. Boukeloua, for doing me the honor of being part of the jury, and for the time given to the attentive reading of my thesis.

Contents

General Introduction

Functional data analysis

Functional data analysis is a branch of statistics that has been the subject of many studies and developments in the recent years. It deals with the analysis and theory of data that are in the form of functions, images and shapes, or more general objects. While the term functional data analysis" was coined by Ramsay (1982) and Ramsay and Dalzell (1991), the history of this area is much older and dates back to Grenander (1950) and Rao (1958).

Functional data are intrinsically infinite dimensional. The high intrinsic dimensionality of these data poses challenges both for theory and computation, where these challenges vary with how the functional data were sampled. On the other hand, the high or infinite dimensional structure of the data is a rich source of information, which brings many opportunities for research and data analysis.

Functional variables can be only observed on a finite grid of discretization points, the estimation can then be viewed as a multidimensional problem. This technic fails because of the great number of discretization points which leads to the well-known problem of curse of dimensionality, linked to the sparseness of the data. This motivates the extension of the finite dimensional statistical technics to the infinite dimensional data setting. The nonparametric methods are then reasonable ways to deal with this type of data sets.

There is nowadays a large number of fields where functional data are collected such as environmetrics, medicine, finance and pattern recognition. To illustrate, here are two real data examples :
The first one, deals with spectrometric data, which can be found at *http ://lib.stat. cmu.edu/datasets/tecator*, and concerns the food industry. More precisely, a food sample has been considered, which contains finely chopped pure meat with a different fat content. The aim is to predict the fat content of a meat sample based on the near-infrared absorbance spectrum, each spectrum being recorded on a Tecator Infratec Food and Feed Analyzer. The data consist of a 100-channel spectrum of absorbances in the wavelength range $850 - 1050$ nm. Here, the response is the corresponding fat content obtained by an analytical chemical process.
The second example deals with the US monthly electricity consumption observed during 338 months (from January 1973 up to February 2001) which can be found at *http://www.economagic.com.* As pointed out in Ferraty and Vieu (2006), this time series can be viewed as dependent functional data. The consumption of a year is the explanatory variable and the consumption of each month of the following year is the response one. We eliminate the 337 and 338 months and we retain the remaining 28 years. Fix $s \in \{1, 2, \dots, 12\}$, in order to predict the electricity consumption of the s^{th} month of the last year (the 28^{th}) by each cited method, we use the 27 first years to define the training sample $(X_i, Y_i^s)_{(i=1,\dots,26)}$ used to build the estimators under investigation, where X_i stands for the consumption of the whole i^{th} year and Y_i^s is the consumption of the s^{th} month of the $(i+1)^{th}$ year. Then, for all $s \in \{1, 2, \dots, 12\}$, we predict Y_{27}^s, which is the consumption of the s^{th} month of the 28^{th} year, given X_{27}.
Studying the link between a scalar response variable Y given a new value for the explanatory variable X is an important subject in nonparametric statistics, and there are several ways to explain this link. For example, the

conditional expectation, the conditional distribution, the conditional densityand the conditional hazard function.

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

Nonparametric Estimation of the regression function

This thesis focuses on the nonparametric estimation of the regression operator defined by:

$$Y = m(x) + \epsilon,$$

where the explanatory variable x is valued in some infinite dimensional space $\mathcal{F}$ equipped with a semi-metric d and Y is a scalar response.
Let $(X_1;Y_1), ..., (X_n;Y_n)$ be a random sample of bivariate data that we have available to estimate the unknown regression function $m(x) = E(Y/X = x)$. In finite dimension, and more precisely when X_i is a real-valued random variable, we can approximate $m(x)$ by using the Taylor Series, around x_0, as follows :

$$m(x) \approx m(x_0)+m^{(1)}(x_0)\,(x-x_0)+\frac{m^{(2)}(x_0)}{2!}\,(x-x_0)^2+...+\frac{m^{(p)}(x_0)}{p!}\,(x-x_0)^p$$

provided that all the required derivatives exist. This is a polynomial of degree p. We can then use this in a minimization problem with the data on x and Y. This is the local polynomial regression problem in which we use the data to estimate that polynomial of degree p which best approximates $m(x)$ in a small neighborhood around the point x_0. ie. we minimize with respect to a, b_1, ..., b_p the function

$$\sum_{i=1}^{n}\left[Y_i - a - b_1(X_i - x_0) - ... - b_p(X_i - x_0)^p\right]^2 K\left(h^{-1}(X_i - x_0)\right).$$

This is a weighted least squares problem where the weights are given by the kernel functions. Here K is a kernel and $h = h_n$ the bandwidth indexed by the sample size. Then $\hat{a}$, one of the two solutions of the display above is the estimate of $m(x_0)$.
It is convenient to define the following vectors and matrices :

$$C^t = \begin{bmatrix} 1 & 1 & ... & 1 \\ (X_1-x_0) & (X_2-x_0) & ... & (X_n-x_0) \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ (X_1-x_0)^p & (X_2-x_0)^p & ... & (X_n-x_0)^p \end{bmatrix}, \quad \mathrm{Y}=\begin{bmatrix} Y_1 \\ ... \\ Y_n \end{bmatrix}, \quad \mathrm{B}=\begin{bmatrix} a \\ b_1 \\ ... \\ b_p \end{bmatrix}$$

and

$$
\text{W}= \begin{bmatrix} K(h^{-1}|(X_1-x_0)) & 0 & ... & 0 \\ 0 & K(h^{-1}|(X_2-x_0)) & ... & 0 \\ . & . & ... & . \\ . & . & ... & . \\ . & . & ... & . \\ 0 & 0 & ... & K(h^{-1}|(X_n-x_0)) \end{bmatrix}
$$

However, the notation used here emphasizes the fact that the local polynomial regression is a weighted regression using data centered around x_0. The least squares problem is then to minimize the weighted sum-of-squares function

$$
(Y-CB)^t\, W\, (Y-CB) \tag{1}
$$

with respect to the parameters b. The solution is

$$
\widehat{m} = e_1^t (C^t W C)^{-1} C^t W Y, \tag{2}
$$

where $e_1 = (1, 0, ..., 0) \in \mathbb{R}^{p+1}$.

When x belongs to a functional space, the principle remains the same. The function m is now approximated by the solution for a of the following minimization problem with respect to a, b_1, ..., b_p

$$
\sum_{i=1}^{n} [Y_i - a - b_1\beta(X_i,x_0) - ... - b_p\beta(X_i,x_0)^p]^2\, K\left(h^{-1}d(X_i,x_0)\right),
$$

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

When $p = 0$ (local constant), the estimator of $m(x)$ is equivalent to the Nadaraya-Watson estimator

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

When $p=1$ (local linear), we can express the estimator as

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

where

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

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

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

so, if the denominator of the estimator $\widehat{m}(x)$ is zero, it is the same for its numerator.

Remark that the expression of $\overline{m}$ allows fast computational issue and that the choices of β and d will be crucial.

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

Kolmogorov's entropy

In practice, the uniform consistency has great importance because it is used to improve the efficiency of the estimation and to solve some problems such as data-driven bandwidth choise (see Benhenni et al. (2007)), or bootstrapping (see Ferraty et al. (2008)).
Noting that, unlike in the multivariate case, the uniform consistency is not a standard extension of the pointwise one. So, suitable additional tools and topological conditions are needed.
For the uniform consistency, where the main tool is to cover a subset $S_{\mathcal{F}}$ with a finite number of balls,one introduces a topological concept defined asfollows

Definition 0.1. *Let S be a subset of a semi-metric space $\mathcal{F}$, and let $\varepsilon > 0$ be given. A finite set of points $x_1, x_2, ..., x_N$ in $\mathcal{F}$ is called an ε-net for S if $S \subset \bigcup_{k=1}^{N} B(x_k, \varepsilon)$. The quantity $\psi_S(\varepsilon) = \ln(N_\varepsilon(S))$, where $N_\varepsilon(S)$ is the minimal number of open balls in $\mathcal{F}$ of radius ε which is necessary to cover S, is called Kolmogorov's ε-entropy of the set S.*

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

details and examples on this topic.
Uniform convergence of other local linear nonparametric estimators has been investigated in some papers as Demongeot et al. (2010) and Demongeot et al. (2011) for the conditional density and Messaci et al. (2015) for the conditional quantile. Leulmi and Messaci (2019) established the uniform almost complete convergence of the local linear estimator of a generalized regression function which generalizes the regression estimator studied in Barrientos et al. (2010) and to focus on a robust tool of prediction (a conditional quantile estimator).

The strong mixing condition

The field of mixing conditions is of great interest in statistics. This comes mainly from the fact that it opens the door for application involving time series. Notice that, there are many ways of modelling the dependence of a sequence of random variables in the case of mixing. But, In this section we focus on the α-mixing (or strong mixing) notion, which is one of the most general among the different mixing structures introduced in the literature (see for instance Roussas and Ioannides (1987) or Chapter 1 in Yoshihara (1994) for definitions of various other mixing structures and links between them). For the strong mixing in the functional context, we refer to Ferraty and Vieu (2006), especially sections 10.3 and 10.4.
All that can be done here is to give a narrow snapshot of part of the strong mixing in the functional context which applied in the theoretical advances in Chapters 1 and 3.
To start with, some notations are introduced. Let $(Z_n)_{n\in\mathbb{Z}}$ be a sequence of random variables on the probability space (Ω, A, P), which takes values in the measurable space (Ω', A'). Denote σ_j^k, $-\infty \leqslant j \leqslant k \leqslant +\infty$, the σ-algebra, which is generated by the random variables $\{Z_j, ..., Z_k\}$.

Definition 0.2. *The strong mixing coefficient of a sequence* $(Z_n)_{n\in\mathbb{Z}}$ *of random variables is defined as*

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

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

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

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

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

Definition 0.3. *The sequence* $(Z_n)_{n\in\mathbb{Z}}$ *is said to be arithmetically* α*-mixing with rate* $a > 0$ *if*

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

It is called geometrically α*-mixing if*

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

To study the nonparametric kernel functional statistical methods (see our chapters 1 and 3), we need the following proposition

Proposition 0.1. *Assume that* Ω' *is a semi-metric space with semi-metric* d*, and that* A *is the* σ*-algebra spanned by the open balls for this semi-metric. Let* x *be a fixed element of* Ω'*. Then we have*
i) $(Z_n)_{n\in\mathbb{R}}$ *is* α*-mixing then* $(d(Z_n, x))_{n\in\mathbb{Z}}$ *is* α*-mixing.*
ii) In addition, if the coefficients of $(Z_n)_{n\in\mathbb{Z}}$ *are geometric (resp. arithmetic) then those of* $(d(Z_n, x))_{n\in\mathbb{Z}}$ *are also geometric (resp. arithmetic with the same order).*

Many works etablished the dependence condition, we cite Laksaci et al. (2011) and Attaoui et al. (2014) for papers dealing with such functional dependent data. In the last works, the pointwise almost complete convergence has been studied, while Laib and Louani (2010), and Ling et al. (2015) obtained the asymptotic properties of the nonparametric kernel estimator for functional stationary ergodic data, Benhenni et al. (2008) for the long memory dependent case. In 2005, Masry (2005) investigated the asymptotic normality of the nonparametric kernel estimator for α-mixing functional data. Demongeot et al. (2013) established the pointwise almost-complete consistency of a fast functional local linear estimator of the conditional density when the explanatory variable is functional and the observations are dependent and Ferraty et al. (2012b) treated the case when the response variable is also functional for the β-mixing observations. Furthermore, Leulmi and Messaci (2018) used the local linear approach to estimate the regression function and established its pointwise and uniforme almost-complete convergences, in the functional α-mixing case.

Censored Data

In survival and reliability analysis, survival or failure time is the duration that an event of intrest takes to occur. It is a positive random variable and often assumed to be bounded. It can be the lifespan of a patient after treatment, the duration of unemployment, the downtime of a device, the age at which a child learns to accomplish a given task, etc. It often happens, for various reasons, that the duration of interest cannot be observed. This may be due to the loss of sight of some subjects, at the beginning or at the end of the study period, or may occur when some subjects have not experienced the event of interest at the end of the study or time of analysis. For example, some patients may still be alive or disease-free at the end of the study period. The exact survival times of these subjects are unknown. These are called censored observations or censored times. We should take them in consideration to obtain correct estimations and presise conclusions. There are three types of censoring :

Right censoring occurs when a subject leaves the study before an event occurs, or the study ends before the event has occurred. In other words, there is right censoring when we observe the censoring variable C instead of the lifetime of interest Y and that we know that $Y > C$. This model is the most frequent in practice, it is for example adapted to the case where the event of interest is the time of survival to a disease and where the duration of the study is previously fixed; patients alive at the end of the study provide right-censored data. The observations are replicas of the pair $(Z = min(Y, C); \delta = 1_{Y \leq C})$ where δ is equal to 1 when the observation is complete, which means that it corresponds to a true value of the variable of interest and is equal to 0 otherwise (censored data).

Left censoring is when the event of interest has already occurred before enrolment. This is very rarely encountered. In other words, there is left censoring when we observe C censoring instead of the lifetime Y and that we know that $Y < C$. The observations are replicas of the pair $(Z = max(Y, C); \delta = 1_{Y \geq C})$. A symmetrical phenomenon at the previous (right censor) occurs when an epidemiologist wishes to know the age at diagnosis in a follow-up study of diabetic retinopathy. At the time of the examination, a 50-year-old participant was found to have already developed retinopathy, but there is no record of the exact time at which initial evidence was found. Thus the age at examination (i.e., 50) is a left-censored observation. It means that the age of diagnosis for this patient is at most 50 years. Notice that the left censorship is generally accompanied by right censoring, this is the case in the twice, double and interval censoring models.

Many researchers consider survival data analysis to be merely the application of two conventional statistical methods to a special type of problem: parametric and nonparametric. This assumption would be true if the survival times of all the subjects were exact and known; however, some survival times are not. Whereas, in the nonfunctional case, the nonparametric estimation based on censored data has been considered by Ould Said, E (2006). They obtained the rate of strong uniform convergence of a kernel estimator of the conditional quantile under an i.i.d. case and El Ghouch and Van Keilegom (2009) studied the asymptotic properties of a local linear estimator of the regression function under the α-mixing assumption. Gannouni et al. (2018) constructed a local linear estimator of the quantile regression and obtained its rate of the almost sure consistency as well as its asymptotic normality in the i.i.d. case. For nonparametric estimation of the regression function under random censorship model we cite Guessoum and Ould Said (2008).
Recently, many topics concerning the analysis of functional and censored data have been developed. We refer to Ait Hennania et al. (2018), where the authors gave a family of robust nonparametric estimators for which consistency and asymptotic normality results are established under independent data. For the same data, Leulmi (2019) and Leulmi (2020) investigated the rates of the pointwise and the uniform almost-complete convergence of a local linear estimator of the conditional quantile and the regression function. She improved that the local linear method outperforms the kernel method even for censored data.
Concerning the dependence data case, we can cite the results on strong consistency and asymptotic normality of the kernel estimator of the conditional quantile function, when the response random variable is subject to random censorship by Horrigue and Ould Said (2011) and Horrigue and Ould Said (2015) for the α-mixing data and Chaouch and Khardani (2014) for the stationary ergodic data. For the kernel estimator of nonparametric regres-

sion function, Ling (2016) obtained its asymptotic properties for the same data. Mechab et al. (2019) examined the almost-complete consistency and the asymptotic normality of the estimator of the relative error regression for the strictly stationary data. Furthermore, Benkhaled et al. (2020) used the local linear approach to estimate the conditional density and established its pointwise almost sure convergence, in the censored and functional α-mixing case.

Whereas, In the single functional index model for α-mixing functional data under random censorship, the pointwise and the uniform almost-complete convergence (with rates), are investigated by Kadiri et al. (2018) for the conditional quantile estimator an citeBouchentouf for the conditional hazard estimator.

Some probabilistic tools

The almost Complete Convergence

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

Definition 0.4. *Let $(Z_n)_{n\in\mathbb{N}^*}$ be a sequence of real random variables (r.r.v.). We say that $(Z_n)_{n\in\mathbb{N}^*}$ converges almost completely to some r.r.v. Z, and we note $Z_n \overset{a.co.}{\longmapsto} Z$, if and only if*

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

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

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

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

Proposition 0.2. *Let* l_x *and* l_y *be two deterministic real numbers and let* $(u_n)_{n\in\mathbb{N}^*}$ *be a sequence of real numbers going to zero.*
i). If $lim_{n\to+\infty} X_n = l_x$, *a.co. and* $lim_{n\to+\infty} Y_n = l_y$, *a.co., we have*
a) $lim_{n\to+\infty}(X_n + Y_n) = l_x + l_y$ *a.co.,*
b) $lim_{n\to+\infty}(X_n \times Y_n) = l_x \times l_y$ *a.co.,*
c) $lim_{n\to+\infty} \frac{1}{Y_n} = \frac{1}{l_y}$ *a.co. as long* $l_y \neq 0$.

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

Exponential Inequalities

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

Independent case

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

Proposition 0.3. *(See Corollary A.8 in Ferraty and Vieu (2006))*
i). if $\forall m \geq 2$, $\exists C_m > 0$; $E|Z_1^m| \leq C_m a^{2(m-1)}$, *we have* $\forall \epsilon > 0$

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

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

$$\frac{1}{n}\sum_{i=1}^{n} Z_i = O_{a.co.}\left((u_n)^{1/2}\right).$$

Mixing case

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

Proposition 0.4. *(See Proposition A.10 in Ferraty and Vieu (2006)) Assume that* $(T_n)_{n\in\mathbb{Z}}$ *is* α*-mixing. Let us, for some* $k \in \mathbb{Z}$, *consider a real variable* τ *(resp.* τ'*) which is* $\sigma_{-\infty}^{k}$*-measurable (resp.* $\sigma_{n+k}^{+\infty}$*-measurable).*
i). If τ *and* τ' *are bounded, then*

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

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

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

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

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

Proposition 0.5. *(See Proposition A.11 in Ferraty and Vieu (2006))*
i). If $\exists p > 2$ and $M > 0$ such that $\forall t > M$; $P(|Z_1| > t) \leq t^{-p}$, then we have for any $r \geqslant 1$, $\epsilon > 0$ and for some $C < +\infty$

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

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

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

Organization of the thesis

After a general introduction which we briefly recall, in it, some basic definitions and probabilistic tools needed in this thesis, our thesis is organized as follows.

Chapter 1 : In this chapter, we study a local modelling approach when one regresses a scalar response on an explanatory functional variable via a regression estimator nder α-mixing condition. This dependence complicates considerably the theoretical study. Then, we establish, the pointwise and the uniform almost complete convergences (see Sections 1.2 and 1.3) of the local linear estimator of the regression function.

Chapter 2 : It consists to introduce a local linear nonparametric estimation of the regression function for a censored scalar response random variable, given a functional random covariate, when the data are independent and identically distributed. Under standard conditions, we establish the pointwise and the uniform almost-complete convergences, with rates, of the proposed estimator (see Sections 2.2 and 2.3).

Chapter 3 : In this chapter, we are concerned with local linear nonparametric estimation of the regression function in the censorship model when the covariates take values in a semimetric space, when the sample is a strong mixing sequence. Then, we establish the pointwise almost-complete convergences, with rate, of the proposed estimator (see sections 3.2 and 3.3).

Chapter 4 : To lend further support to our theoretical results, a simulation study is carried out to illustrate the good accuracy of the studied method. More precisely, we conducted a comparison between kernel and local linear estimators, in the tree cases : Functional and complete case under dependant condition (see 4.1), Functional and censored case under independant condition (see 4.2) and Functional and censored case under dependant condition (see 4.3).

Chapter 1

Nonparametric local linear estimation of the functional regression based on complete data under strong mixing condition

The nonparametric methods are practical ways to deal with the functional data sets. There are nowadays a large number of fields where functional data are collected such as environmetrics, medicine and finance. A classical statistical problem is that of the regression which consists in the study of the relationship between two observed variables with the aim of predicting the value of the response variable when a new value of the explanatory one is observed. The estimation of the regression function arouses a growing interest on both theoretical and practical terms and has been extensively studied for independent data. However, in practice, observed data can exhibit a dependence form. The α-mixing dependence, which is under investigation in this chapter, is not only reasonably weak but it is also fulfilled by many stochastic processes including some time series models. We refer to Leulmi and Messaci (2018) which established the both pontwise and uniforme almost complete convergences of the local linear estimator of generalized regression function $m_\varphi(x) = E(\varphi(Y)/X = x)$ in the case of a scalar response variable Y given a random variable X taking values in a semimetric space. In this chapter, we stady the partical case when $\varphi(t) = t$ (the regression function).

1.1 Model

Let us draw n pairs $(X_i, Y_i)_{i=1,\dots,n}$ of random variables identically distributed from the pair (X, Y) which is valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a infinite-dimensional space equipped with a semi-metric d.
In the complete functional case, Barrientos et al. (2010) proposed the local linear estimator of regression function $m(x) = E(Y|X = x)$ as the solution for a of the following minimization problem

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

where $\beta(.,.)$ is a known operator from $\mathcal{F} \times \mathcal{F}$ into $\mathbb{R}$ such that, $\forall x \in \mathcal{F}, \beta(x, x) = 0$, the function K is a kernel and $h := h_n$ is a sequence of strictly positive real numbers which plays a smoothing parameter role. Then Leulmi and Messaci (2018)) extended their results to α-mixing depented case.
This approach assumes that $a + b\beta(., x)$ is a good approximation of $m(.)$ around x. As $\beta(x, x) = 0$, a will be a suitable estimate for $m(x)$.
We can easilly derive the following explicit estimator

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

where

$$W_{ij}(x) = \beta(X_i, x)\left(\beta(X_i, x) - \beta(X_j, x)\right) K(h^{-1}d(X_i, x))K(h^{-1}d(X_j, x)). \tag{1.1}$$

1.2 Pointwise almost-complete convergence

Let x be a fixed point in $\mathcal{F}$, for any positive real h, $B(x,h) := \{y \in \mathcal{F}/\ d(x,y) \leq h\}$ denotes a closed ball in $\mathcal{F}$ of center x and radius h. We also define $\Phi_x(r_1, r_2) := P(r_1 \leq d(X,x) \leq r_2)$, where r_1 and r_2 are two real numbers.

Leulmi and Messaci (2018) studied the asymptotic behavior of the local linear estimator $\overline{m}$, under the following assumptions.

(H1) For any $h > 0$, $\Phi_x(h) := \Phi_x(0,h) > 0$.

(H2) $\exists\, b > 0$ such that: $\forall x_1, x_2 \in B(x,h),\ |\ m(x_1) - m(x_2)| \leq C_x d^b(x_1, x_2)$, where C_x is a positive constant depending on x.

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

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

(H4) The kernel K is a positive and differentiable function on its support $[0,1]$ and $\exists\, C, C'$ such that $0 < C 1_{[0,1]}(t) \leq K(t) \leq C' 1_{[0,1]}(t) < \infty$.

(H5) The sequence (X_i, Y_i) is an α-mixing sequence with coefficient $\alpha(n)$, moreover (H5a) and (H5b) are satisfied, where

(H5a) $\exists\, C > 0, \exists a > \sup(3, \frac{1+u}{ud}), \forall n \in \mathbb{N}; \alpha(n) \leq C n^{-a}$,

(H5b) $\exists\, 0 < d \leq 1$, $\exists\, C, C' > 0$ such that:

$$C' \left[\Phi_x(h)\right]^{1+d} < \psi_x(h) \leq C \left[\Phi_x(h)\right]^{1+d},$$

where

$$\psi_x(h) := \psi_x(0,h)$$

and

$$\psi_x(h_1, h_2) := P\left(h_1 \leq d(X_1, x) \leq h_2, 0 \leq d(X_2, x) \leq h_2\right).$$

(H6) $\forall m \geq 2 : \delta_m : x \mapsto E(|Y|^m/X = x)$ is a continuous operator at x and

$$\exists C > 0,\ \sup_{i \neq j} E\left(|Y_i Y_j|/(X_i, X_j)\right) \leq C < \infty.$$

(H7) The bandwidth h satisfies

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

and

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

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

(H8) $\lim_{n\to\infty} h = 0$ and $\exists\ 0 < \eta_0 < \frac{a-3}{a+1}$, $\exists\ 0 < u < 1$, $\exists\ C_1, C_2 > 0$ such that $C_1 n^{\frac{3-a}{a+1}+\eta_0} \leq \Phi_x(h) \leq C_2 n^{-u}$.

Hypotheses (H1)-(H3) are standard and have been assumed in the independent case (see Barrientos et al. (2010)). (H4) is a technical condition. (H5a) means that (X_i, Y_i) is arithmetically mixing which is a standard choice of the mixing coefficient in the time series as well as in the context of functional data. (H6) is obviously satisfied whenever m is the conditional distribution function and assumes the boundedness of the response variable (when φ is the identity) which is a reasonable hypothesis in several practical cases. (H7) is of the same kind as (H7) together with (H8) in Barrientos et al. (2010). The choice of bandwidth is given by (H8) which implies that $n\Phi_x(h)/\ln n \to \infty$ as $n \to \infty$.

Now, let us present the pointwise almost-complete convergence of $\overline{m}(x)$.

Theorem 1.1. *Under the assumptions (H1)–(H8), we have*

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

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

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

Proof. Firstly, to treat the pointwise almost-complete convergence of $\hat{m}(x)$, we need the following preliminary technical lemma 1.1.
Secondly, the proof of the theorem 1.1 is based on the following standard decomposition, for all $x \in \mathcal{F}$,

$$\overline{m}(x) - m(x) = \frac{1}{m_0(x)}\left[(m_1(x) - Em_1(x)) - (m(x) - Em_1(x))\right] - \frac{m(x)}{m_0(x)}(m_0(x) - 1) \quad (1.3)$$

where, for $l \in \{0, 1\}$

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

Then, the following lemmas are applied (see Lemmas 1.3–1.5). □

Lemma 1.1. *Under assumptions (H1), (H3), (H4), (H5b) and (H7), we obtain*

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

ii) $\forall (p_1, p_2, l_1, l_2) \in \mathbb{N}^\star \times \mathbb{N}^\star \times \mathbb{N} \times \mathbb{N}$,
$E\left[K_1^{p_1}(x)K_2^{p_2}(x)|\beta_1^{l_1}(x)||\beta_2^{l_2}(x)|\right] \leq Ch^{(l_1+l_2)}\left[\Phi_x(h)\right]^{1+d}.$

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

Proof. **i)**(see Lemma A.1-i in (Barrientos et al. (2010)).

ii) In view of hypotheses (H3) and (H4), we get

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

so, we derive the claimed result by using (H5b).

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

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

Then by (H4) and Fubini's theorem, we can write

$$\begin{aligned} E\left[K_1(x)d^2(X_1,x)K_2(x)\right] &= h^2\int_0^1\int_0^1 t^2K(t)K(u)dP_{(h^{-1}d(X_1,x),h^{-1}d(X_2,x))}(t,u) \\ &> Ch^2\int_0^1\int_0^1 t^2K(t)dP_{(h^{-1}d(X_1,x),h^{-1}d(X_2,x))}(t,u) \\ &> Ch^2\int_0^1\int_0^1\left(\int_0^t \frac{d}{dz}(z^2K(z))dz\right)dP_{(h^{-1}d(X_1,x),h^{-1}d(X_2,x))}(t,u) \\ &> Ch^2\int_0^1\left(\int_0^1\int_0^1 1_{[z,1]}(t)dP_{(h^{-1}d(X_1,x),h^{-1}d(X_2,x))}(t,u)\right)\frac{d}{dz}(z^2K(z))dz. \end{aligned}$$

Moreover, it is easy to check that

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

We end the proof by applying hypothesis (H7). □

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

$$S^2_{n,l,k}(x) = \sum_{i=1}^n\sum_{j=1}^n |Cov(\Lambda_i^{(k,l)}(x),\Lambda_j^{(k,l)}(x))|, \tag{1.4}$$

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

$$\Lambda_i^{(k,l)}(x) = \frac{1}{h^k}\left\{K_i(x)\beta_i^k(x)Y_i^l - E[K_i(x)\beta_i^k(x)Y_i^l]\right\}. \tag{1.5}$$

We deal with these covariance terms in the following result.

Lemma 1.2. *Under assumptions (H1)–(H7) Then, we have:*

$$S^2_{n,l,k}(x) = O(n\Phi_x(h)). \tag{1.6}$$

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

$$S^2_{n,l,k}(x) = \sum_{i=1}^n\sum_{j=1}^n |Cov(\Lambda_i^{(k,l)}(x),\Lambda_j^{(k,l)}(x))| = J_{1,n}(x)+J_{2,n}(x)+nVar(\Lambda_1^{(k,l)}(x)) \tag{1.7}$$

where

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

and

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

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

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

Under (H6) and because $E[Y|X] = m(X) = m(x) + o(1)$ in view of hypothesis (H2), together with the application of Lemma 1.1, we obtain

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

To apply a covariance inequality for no bounded mixing sequences , we must calculate the absolute moments of the r.r.v. $\Lambda_i^{(k,l)}(x)$.

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

the last inequality is obtained by conditionning on X_i. In addition,(H6) implies that $E(|Y|^j|X) = \delta_j(X) = \delta_j(x) + o(1)$ and using Lemma 1.1, we get

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

Now, we can use Rio inequality (see Proposition A.10.(ii) in Ferraty and Vieu (2006)) together with hypothesis (H5a) to obtain

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

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

$$J_{1,n}(x) = O(n\Phi_x(h)) \tag{1.8}$$

and

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

the last result comming from the condition (H8). Now, in view of (H5a) we can choose q such that $u\frac{(q-2)(ad-1)}{q} > 1$. So, we obtain

$$J_{2,n}(x) = O(n\Phi_x(h)) \tag{1.9}$$

For the variance term, Lemma 1.1 and hypothesis (H6) permit to write

$$\begin{aligned}
Var(\Lambda_1^{(k,l)}(x)) &\le C\left[\Phi_x(h) + (\Phi_x(h))^2\right] \\
&\le C\Phi_x(h).
\end{aligned} \tag{1.10}$$

We readily derive the claimed result from (1.7), (1.8), (1.9) and (1.10). □

Lemma 1.3. *Assume that hypotheses (H1)–(H5) and (H7) satisfied, then*

$$m(x) - Em_1(x) = O(h^b).$$

Proof. The proof works exactly as that of Lemma 4.3 in Barrientos et al. (2010), because the dependence condition does not affect the bias terms. Remark that $EW_{1,2}(x) > 0$ under the assumed hypotheses (see relation (1.11)). □

Lemma 1.4. *Under assumptions of Theorem 1.1, we obtain*

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

Proof. Using the decomposition given in the proof of Lemma 4.4 in Barrientos et al. (2010), we set

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

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

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

and

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

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

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

• It is easy to see that under (H1)–(H4), for $p \in \{2,3,4\}$ and $l \in \{0,1\}$,we have
$ES_{p,l}(x) = O(1)$.
• Treatment of the term $Q(x)$
We have

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

together with

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

and (H7) implies that

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

By applying (H3) and (H5b), we get

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

which implies that

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

Now, Lemma 1.1-(iii) and the last result allow to write

$$EW_{12}(x) > Ch^2\left[\Phi_x(h)\right]^{1+d}. \tag{1.11}$$

So, for n sufficiently large

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

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

We have

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

where

$$\Gamma_i^{(p,l)}(x) = \Lambda_i^{(p-2,l)}(x) = \frac{1}{h^{p-2}} \left\{K_i(x)\beta_i^{p-2}(x)Y_i^l - E\left[K_i(x)\beta_i^{p-2}(x)Y_i^l\right]\right\}, \tag{1.12}$$

with $\Lambda_i^{(k,l)}(x)$ is defined in (1.5).

Note that, because $E(\Gamma_1^{(k,l)}(x)) = 0$, $E|\Gamma_1^{(k,l)}(x)|^q = O(\Phi_x(h))$ for $q > 2$ and using Tchebychev's inequality, we can apply Proposition A.11-i in Ferraty and Vieu (2006), to get:

there exist $q > 2$, for any $\varepsilon > 0$, $r \geq 1$ and for some $0 < C < \infty$

$$\begin{aligned} P\left(|S_{p,l}(x) - E\left[S_{p,l}(x)\right]| > \varepsilon\right) &= P\left(|\sum_{i=1}^{n} \Gamma_i^{(p,l)}(x)| > n\varepsilon\Phi_x(h)\right) \\ &\leq C\left[A_1(x) + A_2(x)\right], \end{aligned} \tag{1.13}$$

where

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

Now, taking for $\eta > 0$

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

we obtain

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

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

$$A_2(x) = O(n^{-1-\nu}). \tag{1.14}$$

Moreover, in view of equation (1.6) and the fact that $\ln(x+1) = x - x^2/2 + o(x^2/2)$ where x tends to zero, we can write

$$A_1(x) \leq Cn^{-\eta^2/2}, \tag{1.15}$$

which shows that $A_1(x)$ is the general term of a convergent series for an appropriate choice of η.

Hence, by combining relations (1.13), (1.14) and (1.15), we derive

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

• Finally, by following similar arguments used to prove (1.6), we obtain

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

and

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

In view of (H8), this last rate is negligible with respect to $O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$.

The proof is then completed. □

Lemma 1.5. *If assumptions (H1),(H3), (H4), (H5a), (H5b), (H7) and (H8) hold, we have*

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

and

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

Proof. We can deduce the first part of the claimed results directly from the proof of Lemma 1.4 by taking for all i, $Y_i = 1$.

Therefore, $m_0(x)$ converges almost completely to 1 and this involves that

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

□

1.3 Uniform almost-complete convergence

This section is devoted to the uniform version of Theorem 1.1. In practice, the uniform consistency has great importance because it is used to improve the efficiency of the estimation and to solve some problems such as data-driven bandwidth choise (see Benhenni et al. (2007)), or bootstrapping (see Ferraty et al. (2008)). Noting that, unlike in the multivariate case, the uniform consistency is not a standard extension of the pointwise one. So, suitable additional tools and topological conditions are needed.
More precisely, we establish the uniform almost-complete convergence of $\overline{m}(x)$ on some subset $S_{\mathcal{F}} \subset \mathcal{F}$, such that $S_{\mathcal{F}} \subset \bigcup_{k=1}^{d_n} B(x_k, r_n)$, where for all $k \in \{1, \dots, d_n\}$, $x_k \in S_{\mathcal{F}}$ and (r_n) (resp. (d_n)) is a sequence of positive real (resp. integer) numbers.
The covering condition on the subset $S_{\mathcal{F}}$ is linked to the topological structure of the functional space $\mathcal{F}$. It controls the Kolmogorov's entropy of the set $S_{\mathcal{F}}$. Such considerations have been introduced in nonparametric functional data analysis by Ferraty et al. (2010). The latter contains several examples of subsets $S_{\mathcal{F}}$ and functional spaces $\mathcal{F}$ which satisfy this condition.
For this purpose,we set the following assumptions

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

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

and

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

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

(U2) The regression function m satisfies:

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

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

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

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

(U5) The sequence (X_i, Y_i) satisfies (H5a) and

(U5b) $\exists\, 0 < d \leq 1,\, \exists\, C_1 > 0,\; C_2 > 0$ such that $\forall\, x_1 \in S_{\mathcal{F}}, \forall\, x_2 \in S_{\mathcal{F}}$,
$0 < C_1 \left[\Phi(h)\right]^{1+d} \leq P\left[(X_1, X_2) \in B(x_1,h) \times \in B(x_2,h)\right] \leq C_2 \left[\Phi(h)\right]^{1+d}.$

(U6) $\forall m \geq 2, \exists C_1 > 0, E(|Y|^m/X) \leq C_1$ and $\exists C_2 > 0,\; \sup_{i \neq j} E\left(|Y_i Y_j|/(X_i, X_j)\right) \leq C_2 < \infty$.

(U7) The bandwidth h satisfies (H7) uniformly on $x \in S_{\mathcal{F}}$.

(U8) The bandwidth h satisfies (H8) and for $r_n = O\left(\frac{\ln n}{n}\right)$, the sequence d_n satisfies for n large enough $d_n \sim n$

Roughly speaking, most of these hypotheses are uniform version of the corresponding conditions in the pointwise case. (U3) and (U7) are introduced to deal with the local linear method and are respectively similar to (U3) and (U6) in Messaci et al. (2015). (U5) and (U8) allow to treat the dependence terms and were inspired by imposed conditions in Laksaci et al. (2011) and Attaoui et al. (2014). (U1) and (U2) are commonly used to get the uniformity (see Messaci et al. (2015)) and (U4) is a technical assumption. The last condition on entropy in (U8) is satisfied in some common cases (see example 4 on page 338 in Ferraty et al. (2010)) and leads to find again the same rate as in the pointwise case but uniformly on x.
Our result is as follows.

Theorem 1.2. *Under assumptions (U1)–(U8), we have*

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

To treat the uniform convergence of $\hat{m}(x)$, we need the following preliminary technical lemma. This is the uniform version of Lemma 1.1 and its proof works in the same manner.

Lemma 1.6. *Under assumptions (U1), (U3), (U4), (U5b) and (U7), we obtain*

i) $\forall (p,l) \in \mathbb{N}^\star \times \mathbb{N},\ \sup_{x\in S_{\mathcal{F}}} E\left(K_1^p(x)|\beta_1^l(x)|\right) \le Ch^l\Phi(h).$

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

$\sup_{x\in S_{\mathcal{F}}} E\left(K_1^{p_1}(x)K_2^{p_2}(x)|\beta_1^{l_1}(x)||\beta_2^{l_2}(x)|\right) \le Ch^{(l_1+l_2)}\left[\Phi(h)\right]^{1+d}.$

iii) $\exists n_0 \in \mathbb{N}, \forall n > n_0,\ \ \inf_{x\in S_{\mathcal{F}}} E\left[K_1(x)K_2(x)\beta_1^2(x)\right] > Ch^2\left[\Phi(h)\right]^{1+d}.$

Proof. The proof of Theorem 1.2 is a direct consequence of the decomposition (2.4) and the following lemmas. □

Lemma 1.7. *Assume that hypotheses (U1)–(U5) and (U7) are satisfied, then*

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

Proof. We have

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

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

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

So, we get under assumption (U4)

$|m(x) - Em_1(x)| = \frac{1}{|EW_{12}(x)|}\left|E\left\{W_{12}(x)\left[m(x) - m(X_2)\right]\right\}\right| \le \sup_{x'\in B(x,h)}|m(x) - m(x')|.$

We need to take into account hypothesis (U2) to obtain

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

□

Lemma 1.8. *Under assumptions of Theorem 1.2, we have*

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

Proof. Following the same steps as in the proof of Lemma 1.4, but using Lemma 1.6 instead of Lemma 1.1, we obtain under assumptions (U1) and (U3)–(U8), for $p \in \{2,3,4\}$ and $l \in \{0,1\}$

$$\sup_{x\in S_{\mathcal{F}}} Q(x) = O(1), \quad \sup_{x\in S_{\mathcal{F}}} ES_{p,l}(x) = O(1) \tag{1.16}$$

and

$$\sup_{x\in S_{\mathcal{F}}} Cov\left[S_{2,1}(x), S_{4,0}(x)\right] = O\left(\frac{1}{n\Phi(h)}\right), \quad \sup_{x\in S_{\mathcal{F}}} Cov\left[S_{3,1}(x), S_{3,0}(x)\right] = O\left(\frac{1}{n\Phi(h)}\right). \tag{1.17}$$

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

$$\sup_{x\in S_{\mathcal{F}}} |S_{p,l}(x) - ES_{p,l}(x)| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi(h)}}\right). \tag{1.18}$$

To this aim, let us set

$$j(x) = \arg\min_{j\in\{1,2,\ldots,d_n\}} d(x, x_j),$$

and consider the following decomposition

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

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

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

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

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

where $\Gamma_i^{p,l}(x)$ is defined in (1.12). By applying Proposition A.11-i in Ferraty and Vieu (2006) and since $E|\Gamma_1^{(k,l)}(x)|^q = O(\Phi_x(h))$ for $q > 2$, we have, there exist $q > 2$, for any $\varepsilon > 0$, $r \geq 1$ and for some $0 < C < \infty$

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

where

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

and $S_{n,l,p}^2 := \sup_{x\in S_{\mathcal{F}}} S_{n,l,p}^2(x) = O(n\Phi(h))$ in view of relation (1.6) together with hypothesis (U1).

Choosing for $\eta > 0$

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

we obtain

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

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

Hence, we get for η large enough

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

where $\xi > 0$.

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

First, let us analyze the term $F_1^{p,l}$. Since K is supported in $[0,1]$ and according to (U1), we can write

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

Analysis of the term $F_{1.1}^{p,l}$

According to assumption (U3), we get

$$\begin{aligned}
&1_{B(x,h)}(X_i)\left|\beta_i(x) - \beta_i(x_{j(x)})1_{B(x_{j(x)},h)}(X_i)\right| \\
&\leq \; Cr_n 1_{B(x,h)\bigcap B(x_{j(x)},h)}(X_i) + Ch1_{B(x,h)\bigcap \overline{B(x_{j(x)},h)}}(X_i)
\end{aligned}$$

and

$$\begin{aligned}
&1_{B(x,h)}(X_i)\left|\beta_i^2(x) - \beta_i^2(x_{j(x)})1_{B(x_{j(x)},h)}(X_i)\right| \\
&\leq \; Cr_n h1_{B(x_{j(x)},h)\bigcap B(x,h)}(X_i) + Ch^2 1_{B(x,h)\bigcap \overline{B(x_{j(x)},h)}}(X_i).
\end{aligned}$$

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

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

which yields to the following inequality

$$F_{1.1}^{p,l} \leq \frac{Cr_n}{nh\Phi(h)} \sup_{x\in\mathcal{S}_{\mathcal{F}}} \sum_{i=1}^{n} |Y_i^l| K_i(x) 1_{B(x,h)\bigcap B(x_{j(x)},h)}(X_i)$$

$$+\frac{C}{n\Phi(h)}\sup_{x\in\mathcal{S}_{\mathcal{F}}}\sum_{i=1}^{n}|Y_i^l|K_i(x)1_{B(x,h)\bigcap\overline{B(x_{j(x)},h)}}(X_i). \qquad (1.19)$$

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

Using the following inequality

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

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

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

which leads to

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

This last inequality combined with (1.19) allow us to write

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

Taking into account hypothesis (U4), we find

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

Let

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

In the same manner as for proving (1.6), we have under hypotheses (U1), (U5b), (U6) and (U8)

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

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

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

Second, since

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

we deduce that

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

Applying (1.16), (1.17) and (1.18) together with the last condition of hypothesis (U8), the result of Lemma 1.8 is immediately obtained. □

Lemma 1.9. *If assumptions (U1),(U3), (U4), (H5a), (U5b), (U7) and (U8) are satisfied, we obtain*

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

and

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

Proof. The first part of the claimed results can be directly deduced from the proof of Lemma 1.8 by taking for all i, $Y_i = 1$ and this yields easily to the second part. □

Chapter 2

Nonparametric local linear estimation of the functional regression based on censored data under independant condition

Since the pioneer works in Ferraty and Vieu (2006), various studies dealt with the nonparametric functional kernel estimation. This research field is motivated by the fact that several data collected in practice, are given in the form of curves and that the progress of the digital computing tools allows treatment of such observations. In the complete data case and when the regressors are of functional type, Leulmi and Messaci (2019)established the rates of the pointwise and the uniforme almost-complete convergences for the lacal linear estimator of the generalized regression function. Besides, there exists an extensive literature on the conditional quantile function estimation, when the data are independent and identically distributed. Unfortunately, in many practical applications such as reliability and survival time studies, the interest response variable may be incompletely observed, which make the study of censored data more useful in practice. In this chapter we give the works of Leulmi (2020), where the author investigated the rates of the pointwise and the uniform almost-complete convergences of a local linear nonparametric regression estimator for a censored scalar response random variable, given a functional random covariate, in the independent and identically distributed data case.

2.1 Definition of the estimator

Let us draw n pairs of random variables i.i.d. from the pair (X, Y) which is valued in $\mathcal{F} \times \mathbb{R}$.
We report that in the complete case, the local linear estimator of the regression function $m(x) = E(Y|X = x)$ is presented in Barrientos et al. (2010) as follows

$$\overline{m}(x) = \frac{\sum_{i,j=1}^{n} W_{ij}(x)Y_j}{\sum_{i,j=1}^{n} W_{ij}(x)},$$

with the convention $0/0 := 0$ and $W_{ij}(x)$ are defined by (1.1).
As Y_i is not disponible in practice, we can only observe a sample $(X_i, Z_i, \delta_i)_{1\leq i\leq n}$ of i.d. observations of $(X, Z = Y \wedge C, \delta)$ where R is nonnegative censoring random variable with unknown continuous survival function G ($\forall t$, $S(t) = P(C > t)$) and $\delta = 1_{\{Y\leq C\}}$ (where 1_A denotes the indicator function of the set A) and Y is a nonnegative random variable.
Let $T_Y = \sup\{t, F_Y(t) < 1\}$ denote the upper endpoint of the support of F_U, where $F_U(t) = P(U \leq t)$ denote the distribution of a real random variable (r.r.v.) U.
All over this chapter, we will assume that $T_C < \infty$ and let T be a positive real number such that $T < T_Y$.
Let (A1) be the following assumption.

- C and (X, Y) are independent and $T_Y < T_C < \infty$.

- $\exists T < T_Y$ such that $\forall i, 1 \leq i \leq n;\ Z_i \leq T$.

This assumption is a standard condition in nonparametric censoring estimation which permits us to obtain an unbiased estimator. Like so, the independence assumption between R and (X, Y) is plausible whenever the censoring is independent of the patients modality, $T_Y < T_C$ implies that $S(T) > 0$ because $T < T_Y$.
Inspiring by the idea of Barrientos et al. (2010) combined with that of Guessoum and Ould Said (2008) and by hypothesis (A1), we can construct a local linear estimator of $m(x)$ by

$$\widetilde{m}(x) = \frac{\sum_{i,j=1}^{n} W_{ij}(x) \frac{\delta_j Z_j}{S(Z_j)}}{\sum_{i,j=1}^{n} W_{ij}(x)}, \tag{2.1}$$

with the convention $0/0 := 0$ and $W_{ij}(x)$ is defined in (1.1).
unfortunately, we can not use the estimator (2.1) because S is unknown in practice. however we replace it by its Kaplan and Meier (1958) estimator S_n defined as

$$S_n(t) = \begin{cases} \prod_{i=1}^{n} \left(1 - \frac{1-\delta_{(i)}}{n-i+1}\right)^{1_{\{Z_{(i)} \leq t\}}} & \text{if } t < Z_{(n)} \\ 0 & \text{if } t \geqslant Z_{(n)}, \end{cases} \tag{2.2}$$

where $Z_{(1)} < Z_{(2)} < ... < Z_{(n)}$ are the order statistics of Z_i and $\delta_{(i)}$ the noncensoring indicator corresponding to $Z_{(i)}$.
So, the feasible estimator of $m(x)$ is given by

$$\widehat{m}(x) = \frac{\sum_{i,j=1}^{n} W_{ij}(x) \frac{\delta_j Z_j}{S_n(Z_j)}}{\sum_{i,j=1}^{n} W_{ij}(x)}. \tag{2.3}$$

with the convention $0/0 := 0$. Notice that for all $1 \leq j \leq n$, $S_n(Z_j) = 0$ implies that $\delta_j = 0$.

2.2 Pointwise almost-complete convergence

To establish the pointwise almost-complete convergence of $\widehat{m}(x)$, for a fixed point x in $\mathcal{F}$, we need the assumptions (H1)–(H3) in addition to the following assumptions.

(D1) The kernel K is a positive and differentiable function on its support $[0,1]$.

(D2) The bandwidth h satisfies

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

(D3) There exists an integer n_0, such that

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

and

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

where dP_X is the distribution of X.

(D4) For all $m \geq 2$, $\delta_m : x \mapsto E(|Y|^m|X=x)$ is a continuous operator at x.

Remark that the hypotheses (H1)–(H3) and (D1)–(D3) are the same conditions assumed in Barrientos et al. (2010) and Leulmi and Messaci (2019). The condition (D4) is the same condition (H6) in Leulmi and Messaci (2018) and (H7) in Leulmi and Messaci (2019) with $\varphi(t) = t$.
Now, we are able to give the rate of the pointwise almost-complete convergence of $\widehat{m}(x)$.

Theorem 2.1. *Under assumptions (A1), (H1)–(H3) and (D1)–(D4), we get*

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

To treat the pointwise almost-complete convergence of $\widehat{m}(x)$, we need Lemma A.1 introduced in Barrientos et al. (2010).

Proof. It is easy to see that the proof of Theorem 2.1 is a direct consequence of the following decomposition given by

$$\begin{aligned}\widehat{m}(x) - m(x) &= \frac{1}{m_0(x)}\left[(\widehat{m}_1(x) - \widetilde{m}_1(x)) + (\widetilde{m}_1(x) - E\widetilde{m}_1(x)) + (E\widetilde{m}_1(m) - m(x))\right] \\ &+ \frac{m(x)}{m_0(x)}\left(1 - m_0(x)\right), \end{aligned} \tag{2.4}$$

where

$$\widehat{m}_1(x) = \frac{1}{n(n-1)E\left[W_{12}(x)\right]}\sum_{i\neq j} W_{ij}(x)\frac{\delta_j Z_j}{S_n(Z_j)}, \quad m_0(x) = \frac{1}{n(n-1)E\left[W_{12}(x)\right]}\sum_{i\neq j} W_{ij}(x) \tag{2.5}$$

and

$$\widetilde{m}_1(x) = \frac{1}{n(n-1)E\left[W_{12}(x)\right]}\sum_{i\neq j} W_{ij}(x)\frac{\delta_j Z_j}{S(Z_j)}. \tag{2.6}$$

Then, we apply the following Lemmas. □

Lemma 2.1. *Assume that hypotheses (A1), (H1), (H2) and (D1) are satisfied, then*

$$m(x) - E(\widetilde{m}_1(x)) = O\left(h^b\right).$$

Proof. As (X_i, Z_i, δ_i) are i.i.d., we get

$$E\widetilde{m}_1(x) - m(x) = \frac{1}{E[W_{12}(x)]}E\left\{W_{12}(x)\left[E\left(Z_2 S^{-1}(Z_2)\delta_2|X_2\right) - m(x)\right]\right\}.$$

Hypothesis (D1), combining with the fact that $E(\delta_2|X_2, Y_2) = S(Y_2)$, give that

$$E\left[Z_2 S^{-1}(Z_2)\delta_2|X_2\right] = E\left[Y_2 S^{-1}(Y_2)E\left(\delta_2|X_2, Y_2\right)|X_2\right] = m(X_2).$$

Then, we get

$$E\widetilde{m}_1(x) - m(x) = \frac{1}{E[W_{12}(x)]} E\left[W_{12}(x)\left(m(X_2) - m(x)\right)\right]. \tag{2.7}$$

The claimed result is obtained by using the last relation and the condition (H2). □

Lemma 2.2. *If the assumptions of Theorem 2.1 hold, we obtain*

$$\widetilde{m}_1(x) - E(\widetilde{m}_1(x)) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\,\Phi_x(h)}}\right).$$

Proof. We need to show that

$$\sum_n P\left(|\widetilde{m}_1(x) - E(\widetilde{m}_1(x)| > \epsilon\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right) < \infty.$$

By following the same decomposition idea as in the proof of Lemma 4.4 in Barrientos et al. (2010), we can write

$$\begin{aligned}
\widetilde{m}_1(x) &= \frac{1}{n(n-1)E\left[W_{12}(x)\right]}\sum_{i,j=1}^{n} W_{ij}(x)\delta_j Z_j S^{-1}(Z_j) \\
&= \frac{n^2h^2\Phi_x^2(h)}{n(n-1)E\left[W_{12}(x)\right]}[\left(\frac{1}{n\Phi_x(h)}\sum_{j=1}^{n} K_j(x) Z_j\delta_j S^{-1}(Z_j)\right)\left(\frac{1}{n\Phi_x(h)}\sum_{j=1}^{n}\frac{K_j(x)\beta_j^2(x)}{h^2}\right) \\
&- \left(\frac{1}{n\Phi_x(h)}\sum_{j=1}^{n}\frac{K_j(x)\beta_j(x)Z_j\delta_j S^{-1}(Z_j)}{h}\right)\left(\frac{1}{n\Phi_x(h)}\sum_{j=1}^{n}\frac{K_j(x)\beta_j(x)}{h}\right)] \\
&= Q(x)[M_{2,1}(x)M_{4,0}(x) - M_{3,1}(x)M_{3,0}(x)],
\end{aligned} \tag{2.8}$$

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

$$M_{p,l}(x) = \frac{1}{n\Phi_x(h)}\sum_{j=1}^{n}\frac{K_j(x)\beta_j^{p-2}(x)Z_j^l\delta_j^l S^{-l}(Z_j)}{h^{p-2}} \;\; and \;\; Q(x) = \frac{n^2h^2\Phi_x^2(h)}{n(n-1)E\left[W_{12}(x)\right]}.$$

So, we have

$$\begin{aligned}
\widetilde{m}_1(x) - E\widetilde{m}_1(x) &= Q(x)\{[M_{2,1}(x)M_{4,0}(x) - E(M_{2,1}(x)M_{4,0}(x))] \\
&= [M_{3,1}(x)M_{3,0}(x) - E(M_{3,1}(x)M_{3,0}(x))]\}.
\end{aligned}$$

Notice that, $Q(x) = O(1)$ (see the proof of Lemma 4.4 in Barrientos et al. (2010)), so, we have to show that, for $p \in \{2, 3, 4\}$ and $l \in \{0, 1\}$

$$\sum_n P\left(|M_{p,l}(x) - E(M_{p,l}(x))| > \epsilon\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right) < \infty, \quad E[M_{p,l}(x)] = O(1),$$

and that almost-surely

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

and

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

- Firstly we have

$$\begin{aligned}
M_{p,l}(x) - E(M_{p,l}(x)) &= \frac{1}{nh^{p-2}\Phi_x(h)}\sum_{i=1}^{n}[K_i(x)\beta_i^{p-2}(x)Z_i^l\delta_i^l S^{-l}(Z_i) \\
&- E\left(K_i(x)\beta_i^{p-2}(x)Z_i^l\delta_i^l S^{-l}(Z_i)\right)] \\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h^{p-2}\Phi_x(h)}[K_i(x)\beta_i^{p-2}(x)Z_i^l\delta_i^l S^{-l}(Z_i) \\
&- E\left(K_i(x)\beta_i^{p-2}(x)Z_i^l\delta_i^l S^{-l}(Z_i)\right)] \\
&:= \frac{1}{n}\sum_{i=1}^{n}\eta_i^{(p,l)}(x),
\end{aligned}$$

where

$$\eta_i^{(p,l)}(x) := \frac{1}{h^{p-2}\Phi_x(h)}\left[K_i(x)\beta_i^{p-2}(x)Z_i^l\delta_i^l S^{-l}(Z_i) - E\left(K_i(x)\beta_i^{p-2}(x)Z_i^l\delta_i^l S^{-l}(Z_i)\right)\right] \quad (2.9)$$

In order to apply an exponential inequality, we focus on the absolute moments of the r.r.v. $\eta_i^{(p,l)}(x)$. By Lemma A.1(i) in Barrientos et al. (2010), we can write

$$E|\eta_i^{(p,l)}(x)|^m = O\left([\Phi_x(h)]^{-m+1}\right).$$

Finally, it suffices to apply Corollary A.8–(ii) in Ferraty and Vieu (2006) with $a_n^2 = [\Phi_x(h)]^{-1}$ to get, for $p \in \{2,3,4\}$ and $l \in \{0,1\}$

$$M_{p,l}(x) - EM_{p,l}(x) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right). \tag{2.10}$$

• It is easy to see that under (H1), (H3), (D1) and (A1), we get, for $p \in \{2,3,4\}$ and $l \in \{0,1\}$,

$$E[M_{p,l}(x)] = h^{2-p}\Phi_x(h)^{-1}E\left[K_1(x)\beta_1^{p-2}(x)Z_1^l\delta_1^l S^{-l}(Z_1)\right] \leq C, \tag{2.11}$$

the last inequality is obtained by using the Lemma A.1(i) in Barrientos et al. (2010).

• Treatment of the term $E(M_{2,1}(x))E(M_{4,0}(x)) - E(M_{2,1}(x)M_{4,0}(x))$

We can write

$$\begin{aligned} E(M_{2,1}(x))E(M_{4,0}(x)) - E(M_{2,1}(x)M_{4,0}(x)) &= \frac{1}{nh^2\Phi_x(h)^2}E[K_1(x)\beta_1^2(x)]E[K_1(x)m(X_1)] \\ &+ O\left((n\Phi_x(h))^{-1}\right). \end{aligned}$$

By using Lemma A.1(i) in Barrientos et al. (2010), it is easy to see that

$$E(M_{2,1}(x))E(M_{4,0}(x)) - E(M_{2,1}(x)M_{4,0}(x)) = O\left((n\Phi_x(h))^{-1}\right), \tag{2.12}$$

which is negligible with respect to $O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$, under (D2).

• By similar arguments, one can state

$$E(M_{3,1}(x))E(M_{3,0}(x)) - E(M_{3,1}(x)M_{3,0}(x)) = O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right). \tag{2.13}$$

□

Lemma 2.3. *(see Barrientos et al. (2010))*

Under the assumptions (H1), (H3), (D1) and (D3), we get

$$\widehat{m}_0(x) - 1 = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\,\Phi_x(h)}}\right) \textit{ and } \sum_{n=1}^{\infty} P\left(\widehat{m}_0(x) < \frac{1}{2}\right) < \infty.$$

Lemma 2.4. *Under assumptions (A1), (H1), (H3), (D1) and (D3), we obtain*

$$\widehat{m}_1(x) - \widetilde{m}_1(x) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\,\Phi_x(h)}}\right).$$

Proof. Because the assumption (A1) and the definitions of $\widehat{m}_1(x)$ and $\widetilde{m}_1(x)$ in (2.5) and (2.6), we can write

$$\begin{aligned} |\widehat{m}_1(x) - \widetilde{m}_1(x)| &= \left|\frac{1}{n(n-1)E\left[W_{12}(x)\right]}\sum_{i\neq j} W_{ij}(x)\delta_j Z_j\left(\frac{1}{S_n(Z_j)} - \frac{1}{S(Z_j)}\right)\right| \\ &\leq \frac{|T|\sup_{t\leq T}|S_n(t)-S(t)|}{S_n(T)S(T)}\left|\frac{1}{n(n-1)E\left[W_{12}(x)\right]}\sum_{i\neq j} W_{ij}(x)\right| \\ &\leq \frac{|T|\sup_{t\leq T}|S_n(t)-S(t)|}{S_n(T)S(T)}|m_0(x)|, \qquad (2.14) \end{aligned}$$

where $m_0(x)$ is defined in (2.5).
In order hands, by adapt Theorem 1 of Bitouzé et al. (1999), we get

$$\sup_{t\leq T}|S_n(t) - S(t)| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n}}\right), \qquad (2.15)$$

which is equals to $O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$. The proof is completed by using Lemma 2.3. □

2.3 Uniform almost-complete convergence

In this section, we give the uniform version of the Theorem 2.1 on some subset $S_{\mathcal{F}}$ of $\mathcal{F}$ which can be covered by a finite number of balls. This number has to be related to the radius of these balls.
We suppose that $x_1, \ldots, x_{d_n}$ is an r_n-net for $S_{\mathcal{F}}$ where for all $k \in \{1, \ldots, d_n\}$, $x_k \in S_{\mathcal{F}}$ and (r_n) is a sequence of positive real numbers.

To establish the uniform almost-complete convergence of $\widehat{m}(x)$, we need the assumptions (H1)–(H3) in addition to the following assumptions.

(E1) The kernel K fulfills (D1) and is Lipschitzian on $[0,1]$.

(E2) The bandwidth h satisfies (H5) and for $r_n = O\left(\frac{\ln n}{n}\right)$, the sequence d_n satisfies, for n large enough

$$\frac{(\ln n)^2}{n\Phi(h)} < \ln d_n < \frac{n\Phi(h)}{\ln n}$$

and

$$\sum_{n=1}^{\infty} d_n^{1-\alpha} < \infty, \quad for \quad some \quad \alpha > 1.$$

(E3) The hypothesis (D3) is satisfies uniformly on $x \in S_{\mathcal{F}}$.

(E4) $\exists\ C > 0$, $\forall m \geq 2$, $E(|Y|^m|X) < \delta_m(x) < C$, with $\delta_m(x)$ continous on $S_{\mathcal{F}}$.

Notice that most of these conditions are uniform version of the corresponding conditions in the pointwise case. We refer to Leulmi and Messaci (2019) for the conditions (U1)–(U3), (E1) and (E3). The condition (E4) is the same condition (H3) in Ferraty et al. (2010) and (U7) in Leulmi and Messaci (2019) with $\varphi(t) = t$. The assumption (E2) is linked with the topological structure of the functional variable (for more details and examples, see Kolmogorov and Tikhomirov (1959) and Ferraty et al. (2010)).
Our main result is as follows.

Theorem 2.2. *Under assumptions (A1), (U1)–(U3) and (E1)–(E4), we have*

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

This result shows that, contrary to the finite case, the rate of convergence obtained may differ from that of the pointwise consistency. Notice that this rate of convergence is the same as that of Leulmi and Messaci (2019) in the case of uncensored data.
To treat the uniform convergence of $\widehat{m}(x)$, we need to Lemma 4.1 introduced in Messaci et al. (2015).

Proof. It is clear that the proof of Theorem 2.2 is a direct consequence of the decomposition (2.4) and of the following Lemmas which correspond to the uniform versions of Lemmas 2.1–2.4. □

Lemma 2.5. *Assume that hypotheses (A1), (U1), (U2) and (E1) hold, then*

$$\sup_{x\in S_{\mathcal{F}}} |m(x) - E(\widetilde{m_1}(x))| = O\left(h^b\right).$$

Proof. It is a direct proof, by combining equation (2.7) and hypothesis (U2). □

Lemma 2.6. *Under assumptions of Theorem 3.2, we obtain*

$$\sup_{x\in S_{\mathcal{F}}} |\widetilde{m}_1(x) - E(\widetilde{m}_1(x))| = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\Phi(h)}}\right).$$

Proof. We use again the decomposition (2.8) and by following the same steps as in (2.11), (2.12) and (2.13), with using Lemma 4.1 in Messaci et al. (2015) instead of lemma A.1 in Barrientos et al. (2010), we obtain under the assumptions (U1), (U3), (E1), (E3) and (A1), for $p = 2, 3, 4$ and $l = 0, 1$,

$$\sup_{x\in S_{\mathcal{F}}} E(M_{p,l}(x)) = O(1), \quad \sup_{x\in S_{\mathcal{F}}} Q(x) = O(1), \tag{2.16}$$

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

and

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

which is, in view of hypothesis (E2), equals to $O\left(\sqrt{\frac{\ln d_n}{n\Phi(h)}}\right)$.
So, we need to check that for $p = 2, 3, 4$ and $l = 0, 1$,

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

Now, we consider the following decomposition

$$\begin{aligned}\sup_{x\in S_{\mathcal{F}}} |M_{p,l}(x) - E(M_{p,l}(x))| &\leq \sup_{x\in S_{\mathcal{F}}} \left|M_{p,l}(x) - M_{p,l}(x_{j(x)})\right| \\ &+ \sup_{x\in S_{\mathcal{F}}} \left|E(M_{p,l}(x_{j(x)})) - E(M_{p,l}(x_{j(x)}))\right| \\ &+ \sup_{x\in S_{\mathcal{F}}} \left|E(M_{p,l}(x_{j(x)})) - E(M_{p,l}(x))\right| \\ &:= \sum_{i=1}^{3} T_i^{p,l}.\end{aligned}$$

Study of the terms $T_1^{p,l}$ and $T_3^{p,l}$.
First, let us analyze the term $T_1^{p,l}$. Since K is supported in $[0, 1]$ and according to (U1), we can write

$$T_1^{p,l} \leq \frac{Cr_n}{nh\Phi(h)} \sup_{x\in S_{\mathcal{F}}} \sum_{i=1}^{n} |Z_i^l|\delta_i^l S^{-l}(Z_i) 1_{B(x,h)\cup B(x_{j(x)},h)}(X_i).$$

Let

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

The assumptions (A1) and (E4), implie that

$$E|\xi_1^m| \leq \frac{Cr_n^m}{h^m\Phi(h)^{m-1}}, \tag{2.17}$$

so, by applying corollary A.8-(ii) in Ferraty and Vieu (2006), with $a_n^2 = \frac{r_n}{h\Phi(h)}$,

$$\frac{1}{n}\sum_{i=1}^{n} \xi_i = E(\xi_1) + O_{a.co.}\left(\sqrt{\frac{r_n \ln n}{nh\Phi(h)}}\right).$$

Applying (2.17) again (for $m = 1$), one gets

$$T_1^{p,l} = O(\frac{r_n}{h}) + O_{a.co.}\left(\sqrt{\frac{r_n \ln n}{nh\Phi(h)}}\right).$$

Combining this equation with assumption (E2) and the second part of the assumption (U1), we obtain

$$T_1^{p,l} = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\Phi(h)}}\right). \tag{2.18}$$

Second, since

$$T_3^{p,l} \leq E\left(\sup_{x\in S_{\mathcal{F}}} \left|M_{p,l}(x) - M_{p,l}(x_{j(x)})\right|\right),$$

we deduce that

$$T_3^{p,l} = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\Phi(h)}}\right). \tag{2.19}$$

Study of the term $T_2^{p,l}$.

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

$$P\left(T_2^{p,l} > \zeta\sqrt{\frac{\ln d_n}{n\Phi(h)}}\right) = P\left(\sup_{x\in S_{\mathcal{F}}} \left|M_{p,l}(x_{j(x)}) - E(M_{p,l}(x_{j(x)})\right| > \zeta\sqrt{\frac{\ln d_n}{n\Phi(h)}}\right)$$
$$\leq \ d_n \max_{x_{j(x)}\in\{x_1,\dots,x_{d_n}\}} P\left(\frac{1}{n}\left|\sum_{i=1}^{n} \eta_i^{(p,l)}(x_{j(x)})\right| > \zeta\sqrt{\frac{\ln d_n}{n\Phi(h)}}\right),$$

where $\eta_i^{(p,l)}$ is defined in (2.9). By using again Corollary A.8-(ii) in Ferraty and Vieu (2006) and the assumption (E2), we obtain

$$T_2^{p,l} = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\Phi(h)}}\right). \tag{2.20}$$

Finally, the result of Lemma 2.6 follows from the relations (2.18), (2.20) and (2.19). □

Lemma 2.7. *(see Messaci et al. (2015)) If assumptions (U1), (U3), (E1)–(E3) are satisfied, we get*

$$\sup_{x\in S_{\mathcal{F}}} |\widehat{m}_0(x) - 1| = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\Phi(h)}}\right) \text{ and } \sum_{n=1}^{\infty} P\left(\inf_{x\in S_{\mathcal{F}}} \widehat{m}_0(x) < \frac{1}{2}\right) < \infty.$$

Lemma 2.8. *Under assumptions (A1), (U1), (U3) and (E1)–(E3), we obtain*

$$\sup_{x\in S_{\mathcal{F}}} |\widehat{m}_1(x) - \widetilde{m}_1(x)| = O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\Phi(h)}}\right).$$

Proof. By the relation (2.14), we can write

$$\sup_{x\in S_{\mathcal{F}}} |\widehat{m}_1(x) - \widetilde{m}_1(x)| \leq C \sup_{t\leq T} |S_n(t) - S(t)| \sup_{x\in S_{\mathcal{F}}} |\widehat{m}_0(x)|.$$

The proof is completed by the relation (2.15) and Lemma 2.7. □

Chapter 3

Nonparametric local linear estimation of the functional regression based on censored data under strong mixing condition

In Chapter 2, we study a nonparametric local linear regression estimator in the functional and censored case under independant condition. The present work gives its extension to the dependent case (α-mixing) and this fact complicates considerably the study. The interest comes mainly from the fact that some application fields, for functional methods, need to analyze time series. This work has been published in a international journal (see Leulmi et al. (2022) for the sections 3.1 and 3.2).

3.1 Definition of the estimator

Consider n pairs of random variables $(X_i, Y_i)_{i=1,\dots,n}$ identically distributed as the pair (X, Y) which is valued in $\mathcal{F} \times \mathbb{R}$.

Unfortunately, Y_i is not available in our setting. We can only observe a sample $(X_i, Z_i, \delta_i)_{1 \le i \le n}$ of i.i.d. observations of $(X, Z = Y \wedge C, \delta)$ where C is nonnegative censoring random variable with unknown continuous survival function S ($\forall t$, $S(t) = P(C > t)$) and $\delta = 1_{\{Y \le C\}}$ (where 1_A denotes the indicator function of the set A) and Y is a nonnegative random variable.
For the rest of the Chapter, we will assume that the sequences $(X_i)_{1 \le i \le n}$, $(Y_i)_{1 \le i \le n}$ and $(C_i)_{1 \le i \le n}$ are stationary and α-mixing with mixing coefficients $\alpha_1(n)$, $\alpha_2(n)$ and $\alpha_3(n)$ respectively. Notice that, in view of Lemma 2 in Cai (2001), we can show that, the sequences $(X_i, Y_i)_{1 \le i \le n}$, $(Z_i)_{1 \le i \le n}$ and then $(X_i, Z_i, \delta_i)_{1 \le i \le n}$ are α-mixing with coefficients $a(n)$, $b(n)$ and $\alpha(n)$ respectively, where

$$a(n) = 4 \max(\alpha_1(n), \alpha_2(n)),$$
$$b(n) = 4 \max(\alpha_2(n), \alpha_3(n))$$

and

$$\alpha(n) = 4 \max(\alpha_1(n), b(n)) = 4 \max\left(\alpha_1(n), 4 \max(\alpha_2(n), \alpha_3(n))\right).$$

Furthermore, the dependence assumption of $(X_i)_{1 \le i \le n}$, $(Y_i)_{1 \le i \le n}$ and $(C_i)_{1 \le i \le n}$, seems to be more general and one can think to replace it by a classical dependence assumption of $(X_i, Y_i)_{1 \le i \le n}$ and the sequence $(C_i)_{1 \le i \le n}$ is i.i.d. censoring random variable, see for example Benkhaled et al. (2020). Because, since $(X_i, Y_i)_{1 \le i \le n}$ is stationary and α-mixing, it is straightforward that the sequences $(X_i)_{1 \le i \le n}$ and $(Y_i)_{1 \le i \le n}$ are also stationary and α-mixing. This can be deduced from the fact that the later can be seen as a projection-image of the former. On other hand, the α-mixing condition of $(C_i)_{1 \le i \le n}$ is more comprehensive than the independence assumption, we put $\alpha_3 = 0$.
A feasible local linear nonparametric estimator of $m(x)$, constructed in Leulmi (2020), is defined by

$$\widehat{m}(x) = \frac{\sum_{i,j=1}^{n} W_{ij}(x) \frac{\delta_j Z_j}{S_n(Z_j)}}{\sum_{i,j=1}^{n} W_{ij}(x)}, \quad \left(\frac{0}{0} = 0\right), \tag{3.1}$$

where $W_{ij}(x)$ is defined in (1.1) and S_n is the well known Kaplan and Meier (1958) estimator of S, which given in (2.2). Notice that for all $1 \leq j \leq n$, $G_n(Z_j) = 0$ implies that $\delta_j = 0$.
From now on, we have that $(X_i, Z_i, \delta_i)_{1 \leq i \leq n}$ is strongly mixing with mixing's coefficient $\alpha(n)$.
Now we are in position to give our assumptions and main result.

3.2 Pointwise almost-complete convergence

The aim of this section is to establish the pointwise almost-complete convergence of $\widehat{m}$. For this purpose, we need the assumptions (H1)–(H8) and (A1).

Theorem 3.1. *Assume that assumptions (A1) and (H1)-(H8) are satisfied, then*

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

One of the main features of the present paper is studding the local linear estimation under the dependent and censored case, which is generalizes several usual situations. In particular, we consider the independent case (see Leulmi (2020)), the complete case(see Leulmi and Messaci (2018)) and the kernel method (see Ling (2016)).

Proof. To treat the pointwise almost-complete convergence of $\widehat{m}(x)$, we need Lemma A1 introduced in Leulmi and Messaci (2018) and the following preliminary tecnical Lemma 3.1. Then, the proof of the Theorem 3.1 is based on the decomposition (2.4) and the following Lemmas. □

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

$$G^2_{n,l,k}(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} |Cov(\Delta_i^{(k,l)}(x), \Delta_j^{(k,l)}(x))|, \tag{3.2}$$

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

$$\Delta_i^{(k,l)}(x) = \frac{1}{h^k}\left\{K_i(x)\beta_i^k(x)\delta_i^l Z_i^l S^{-l}(Z_i) - E[K_i(x)\beta_i^k(x)\delta_i^l Z_i^l S^{-l}(Z_i)]\right\}. \tag{3.3}$$

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

Lemma 3.1. *Under assumptions (A1) and (H1)-(H7), we have*

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

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

$$G^2_{n,l,k}(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} |Cov(\Delta_i^{(k,l)}(x), \Delta_j^{(k,l)}(x))| = J_{1,n}(x)+J_{2,n}(x)+nVar(\Delta_1^{(k,l)}(x)) \tag{3.5}$$

with

$$J'_{1,n}(x) = \sum_{S_1} |Cov(\Delta_i^{(k,l)}(x), \Delta_j^{(k,l)}(x))|; \;\; S_1 = \{(i,j) : 1 \leq |i-j| \leq m_n\}.$$

and

$$J'_{2,n}(x) = \sum_{S_2} |Cov(\Delta_i^{(k,l)}(x), \Delta_j^{(k,l)}(x))|; \;\; S_2 = \{(i,j) : m_n+1 \leq |i-j| \leq n-1\},$$

where the sequence (m_n) will be specified below.

For the $J'_{1,n}(x)$ term, since for all i in $\{1,...,n\}$, $E(\Delta_i^{(k,l)}(x)) = 0$, we have

$$
\begin{aligned}
J'_{1,n}(x) &= \sum_{S_1} |E[\Delta_i^{(k,l)}(x)\Delta_j^{(k,l)}(x)]| \\
&\leq \frac{1}{h^{2k}} \sum_{S_1} \{E[K_i(x)\beta_i^k(x)K_j(x)\beta_j^k(x)E(|\delta_i^l Z_i^l S^{-l}(Z_i)\delta_j^l Z_j^l S^{-l}(Z_j)||(X_i, X_j)] \\
&+ |E[K_i(x)\beta_i^k(x)E(\delta_i^l Z_i^l S^{-l}(Z_i)|X_i)]||E[K_j(x)\beta_j^k(x)E(\delta_j^l Z_j^l S^{-l}(Z_j)|X_j]|\}. \\
&\leq \frac{1}{h^{2k}S^{2l}(T)} \sum_{S_1} \{E[K_i(x)\beta_i^k(x)K_j(x)\beta_j^k(x)E(|\delta_i^l Y_i^l \delta_j^l Y_j^l||(X_i, X_j)] \\
&+ |E[K_i(x)\beta_i^k(x)E(\delta_i^l Y_i^l|X_i)]||E[K_j(x)\beta_j^k(x)E(\delta_j^l Y_j^l|X_j]|\},
\end{aligned}
$$

the last inequality is obtained by the condtion (A1). Under the conditions (H2) and (H6) and by using the Lemma A1 in Leulmi and Messaci (2018) , we get

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

For the $J'_{2,n}(x)$ term, we must calculate the absolute moments of the r.r.v. $\Delta_i^{(k,l)}(x)$.

$$
\begin{aligned}
E|\Delta_i^{(k,l)}(x)|^q &\leq h^{-qk} \sum_{j=0}^{q} C_{j,q} E|K_i^j(x)\beta_i^{kj}\delta_i^{lj} Z_i^{lj} S^{-lj}(Z_i)||EK_i(x)\beta_i^k(x)\delta_i^l Z_i^l S^{-l}(Z_i)|^{q-j} \\
&\leq h^{-qk} \sum_{j=0}^{q} C_{j,q} E[K_i^j(x)\beta_i^{kj}(x)E(|\delta_i^{lj} Y_i^{lj} S^{-lj}(Y_i)|X_i)] \\
&[E(K_i(x)\beta_i^k(x)E(|\delta_i^l Y_i^l S^{-l}(Y_i)||X_i))]^{q-j}, \\
&\leq G^{-lq}(T)h^{-qk} \sum_{j=0}^{q} C_{j,q} E[K_i^j(x)\beta_i^{kj}(x)E(|Y_i^{lj}|X_i)][E(K_i(x)\beta_i^k(x)E(|Y_i^l||X_i))]^{q-j},
\end{aligned}
$$

the last inequality is obtained by conditionning on X_i. Now using Lemma A1 Leulmi and Messaci (2018) and under (H6), we obtain

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

Now, we can use a covariance inequality for unbounded mixing sequences (see Proposition A.10.(ii) in Ferraty and Vieu (2006)) together with (H5a) to obtain

$$\begin{aligned} J'_{2,n}(x) &= \sum_{S_2} |Cov(\Delta_i^{(k,l)}(x), \Delta_j^{(k,l)}(x))| \\ &\leq C \left[E|\Delta_1^{(k,l)}(x)|^q \right]^{2/q} \sum_{S_2} [\alpha(|i-j|)]^{1-\frac{2}{q}} \\ &\leq C \left[\varPhi_x(h)\right]^{\frac{2}{q}} \sum_{S_2} |i-j|^{-a(1-\frac{2}{q})} \\ &\leq C \left[\varPhi_x(h)\right]^{\frac{2}{q}} n^2 m_n^{-a(1-\frac{2}{q})}. \end{aligned}$$

Taking $m_n = (\varPhi_x(h))^{-d}$, we get

$$J'_{1,n}(x) = O(n\varPhi_x(h)) \tag{3.6}$$

and

$$\begin{aligned} J'_{2,n}(x) &\leq C\left(n\varPhi_x(h)\right) \left[n \left(\varPhi_x(h)\right)^{\frac{(q-2)(ad-1)}{q}} \right] \\ &\leq C\left(n\varPhi_x(h)\right) n^{1-\frac{u(q-2)(ad-1)}{q}}, \end{aligned}$$

the last result coming from the condition (H8). Now, we can choose q such that $u\frac{(q-2)(ad-1)}{q} > 1$. So, we obtain

$$J'_{2,n}(x) = O(n\varPhi_x(h)) \tag{3.7}$$

For the variance term, Lemma A1 in Leulmi and Messaci (2018) and conditions (A1) and (H6) permit to write

$$\begin{aligned} Var(\Delta_1^{(k,l)}(x)) &\leq \frac{C}{S^2(T)} \left[\varPhi_x(h) + (\varPhi_x(h))^2\right] \\ &\leq C\varPhi_x(h). \end{aligned} \tag{3.8}$$

The poof is completed by (3.5), (3.6), (3.7) and (3.8). □

Lemma 3.2. *Assume that hypotheses (A1), (H1)-(H5) and (H7) hold, then*

$$m(x) - E(\widetilde{m}_1(x)) = O\left(h^b\right).$$

Proof. The bias term is not affected by the dependence condition. Therefore, by the equiprobability of the couples (X_i, Z_i, δ_i), we get

$$E\widetilde{m}_1(x) - m(x) = \frac{1}{E[W_{12}(x)]} E\left\{W_{12}(x)\left[E\left(Z_2 S^{-1}(Z_2)\delta_2 | X_2\right) - m(x)\right]\right\}.$$

Hypothesis (H4), combining with the facts that $E(\delta_2|X_2, Y_2) = S(Y_2)$ and $\delta_2 Z_2 = \delta_2 Y_2$, give that

$$E\left[Z_2 S^{-1}(Z_2)\delta_2|X_2\right] = E\left[Y_2 S^{-1}(Y_2) E\left(\delta_2|X_2, Y_2\right)|X_2\right] = m(X_2).$$

Then, we have

$$E\widetilde{m}_1(x) - m(x) = \frac{1}{E[W_{12}(x)]} E\left[W_{12}(x)\left(m(X_2) - m(x)\right)\right]. \tag{3.9}$$

The claimed result is obtained by using the last relation and the condition (H2). □

Lemma 3.3. *Under assumptions of Theorem 3.1, we get*

$$\widetilde{m}_1(x) - E(\widetilde{m}_1(x)) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\,\Phi_x(h)}}\right).$$

Proof. Inspiring by the proof of Lemma 4.4 in Barrientos et al. (2010), we consider the following decomposition

$$\begin{aligned}
\widetilde{m}_1(x) &= \frac{1}{n(n-1)E\left[W_{12}(x)\right]} \sum_{i,j=1}^{n} W_{ij}(x)\delta_j Z_j S^{-1}(Z_j) \\
&= \frac{n^2 h^2 \Phi_x^2(h)}{n(n-1)E\left[W_{12}(x)\right]} [\left(\frac{1}{n\Phi_x(h)} \sum_{j=1}^{n} K_j(x) Z_j \delta_j S^{-1}(Z_j)\right)\left(\frac{1}{n\Phi_x(h)} \sum_{j=1}^{n} \frac{K_j(x)\beta_j^2(x)}{h^2}\right) \\
&- \left(\frac{1}{n\Phi_x(h)} \sum_{j=1}^{n} \frac{K_j(x)\beta_j(x) Z_j \delta_j S^{-1}(Z_j)}{h}\right)\left(\frac{1}{n\Phi_x(h)} \sum_{j=1}^{n} \frac{K_j(x)\beta_j(x)}{h}\right)] \\
&= Q(x)[D_{2,1}(x)D_{4,0}(x) - D_{3,1}(x)D_{3,0}(x)],
\end{aligned} \tag{3.10}$$

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

$$D_{p,l}(x) = \frac{1}{n\Phi_x(h)} \sum_{j=1}^{n} \frac{K_j(x)\beta_j^{p-2}(x) Z_j^l \delta_j^l S^{-l}(Z_j)}{h^{p-2}} \quad and \quad Q(x) = \frac{n^2 h^2 \Phi_x^2(h)}{n(n-1)E\left[W_{12}(x)\right]}.$$

Notice that, $Q(x) = O(1)$ (see the proof of Lemma 2 in Leulmi and Messaci (2018)), so, we have to show that, for $p \in \{2,3,4\}$ and $l \in \{0,1\}$

$$\sum_n P\left(|D_{p,l}(x) - E(D_{p,l}(x))| > \epsilon\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right) < \infty, \quad E[D_{p,l}(x)] = O(1),$$

and that almost surely

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

and

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

- Firstly we have

$$D_{p,l}(x) - ED_{p,l}(x) = \frac{1}{n\Phi_x(h)} \sum_{i=1}^{n} \Delta_i^{(p-2,l)}(x),$$

with $\Delta_i^{(k,l)}(x)$ is defined in (3.3).
Note that, because $E(\Delta_1^{(k,l)}(x)) = 0$, $E|\Delta_1^{(k,l)}(x)|^q = O(\Phi_x(h))$ for $q > 2$ and using Tchebychev's inequality, we can apply Proposition A.11-i in Ferraty and Vieu (2006), to get for any $q > 2$, $\varepsilon > 0$, $r \geq 1$ and for some $0 < C < \infty$

$$\begin{aligned} P\left(|D_{p,l}(x) - E\left[D_{p,l}(x)\right]| > \varepsilon\right) &= P\left(|\sum_{i=1}^{n} \Delta_i^{(p,l)}(x)| > n\varepsilon\Phi_x(h)\right) \\ &\leq C\left[A_1(x) + A_2(x)\right], \end{aligned} \tag{3.11}$$

where

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

Now, choosing for $\eta > 0$

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

In view of Lemma 3.1, we have $G^2_{n,l,k}(x) = O(n\Phi_x(h))$. So, we obtain

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

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

$$A_2(x) = O(n^{-1-\nu}). \tag{3.12}$$

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

$$A_1(x) \leq Cn^{-\eta^2/2}, \tag{3.13}$$

which shows that $A_1(x)$ is the general term of a convergent series for an appropriate choice of η.
Hence, by combining relations (1.13), (3.12) and (3.13), we derive

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

• It is easy to see that under (H1), (H3), (H4) and (A1), we get, for $p \in \{2,3,4\}$ and $l \in \{0,1\}$,

$$E[D_{p,l}(x)] = h^{2-p}\Phi_x(h)^{-1}E\left[K_1(x)\beta_1^{p-2}(x)Z_1^l\delta_1^l S^{-l}(Z_1)\right] \leq C, \tag{3.14}$$

the last inequality is obtained by using the Lemma A1(i) in Leulmi and Messaci (2018) and the condition (A1).
• Finally, by following similar arguments used to prove (3.4), we obtain

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

and

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

In view of (H8), this last rate is negligible with respect to $O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$. The proof is then completed. □

Lemma 3.4. *(see Leulmi and Messaci (2018)) If assumptions (H1),(H3), (H4), (H5a), (H5b), (H7) and (H8) are satisfied, we obtain*

$$\widehat{m}_0(x) - 1 = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\,\Phi_x(h)}}\right) \text{ and } \sum_{n=1}^{\infty} P\left(\widehat{m}_0(x) < \frac{1}{2}\right) < \infty.$$

Lemma 3.5. *Under assumptions (A1), (H1),(H3), (H4), (H5a), (H5b) and (H7), we have*

$$\widehat{m}_1(x) - \widetilde{m}_1(x) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\,\Phi_x(h)}}\right).$$

Proof. Because the assumption (A1) and the definitions of $\widehat{m}_1(x)$ and $\widetilde{m}_1(x)$ in (2.5) and (2.6), we can write

$$\begin{aligned}
|\widehat{m}_1(x) - \widetilde{m}_1(x)| &= \left|\frac{1}{n(n-1)E\left[W_{12}(x)\right]}\sum_{i\neq j} W_{ij}(x)\delta_j Z_j\left(\frac{1}{S_n(Z_j)} - \frac{1}{S(Z_j)}\right)\right| \\
&\leq \frac{|T|\sup_{t\leq T}|S_n(t) - S(t)|}{S_n(T)S(T)}\left|\frac{1}{n(n-1)E\left[W_{12}(x)\right]}\sum_{i\neq j} W_{ij}(x)\right| \\
&\leq \frac{|T|\sup_{t\leq T}|S_n(t) - S(t)|}{S_n(T)S(T)}|m_0(x)|, \qquad (3.15)
\end{aligned}$$

where $m_0(x)$ is defined in (2.5).
In order hands, following Cai (2001) and Rouabah et al. (2018), we obtain

$$\sup_{t\leq T}|S_n(t) - S(t)| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n}}\right), \qquad (3.16)$$

which is equals to $O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$. The proof is completed by using Lemma 3.4.

□

3.3 Uniform almost-complete convergence

In this section, we give the uniform version of the Theorem 3.1 on some subset $S_{\mathcal{F}}$ of $\mathcal{F}$ which can be covered by a finite number of balls. This number has to be related to the radius of these balls.
We suppose that $x_1, \ldots, x_{d_n}$ is an r_n-net for $S_{\mathcal{F}}$ where for all $k \in \{1, \ldots, d_n\}$, $x_k \in S_{\mathcal{F}}$ and (r_n) is a sequence of positive real numbers.
To study the uniform almost-complete convergence of $\widehat{m}(x)$, we need the assumptions (U1)–(U8) in addition to the assumption (A1).

Theorem 3.2. *Under assumptions (A1) and (U1)–(U8), we have*

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

Proof. It is clear that the proof of Theorem 3.2 is a direct consequence of the decomposition (2.4) and of the following Lemmas. □

Lemma 3.6. *Assume that hypotheses (A1), (U1), (U2) and (U4) hold, then*

$$\sup_{x\in S_{\mathcal{F}}} |m(x) - E(\widetilde{m_1}(x))| = O\left(h^b\right).$$

Proof. It is a direct proof, by combining equation (3.9) and hypothesis (U2).

□

Lemma 3.7. *Under the assumptions of Theorem 3.1, we obtain*

$$\sup_{x\in S_{\mathcal{F}}} |\widetilde{m}_1(x) - E(\widetilde{m}_1(x))| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\,\Phi_x(h)}}\right).$$

Proof. Following the same steps as in the proof of Lemma 3.3, but using Lemma 1.6 instead of Lemma 1.1, we obtain under assumptions (A1), (U1) and (U3)–(U8), for $p \in \{2,3,4\}$ and $l \in \{0,1\}$

$$\sup_{x\in S_{\mathcal{F}}} E(D_{p,l}(x)) = O(1), \quad \sup_{x\in S_{\mathcal{F}}} Q(x) = O(1), \tag{3.17}$$

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

and

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

which is, in view of hypothesis (U5), equals to $O\left(\sqrt{\frac{\ln n}{n\Phi(h)}}\right)$.
So, we need to check that for $p = 2,3,4$ and $l = 0,1$,

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

Now, we consider the following decomposition

$$\begin{aligned}\sup_{x\in S_{\mathcal{F}}} |D_{p,l}(x) - E(D_{p,l}(x))| &\leq \sup_{x\in S_{\mathcal{F}}} \left|D_{p,l}(x) - D_{p,l}(x_{j(x)})\right| \\ &+ \sup_{x\in S_{\mathcal{F}}} \left|E(D_{p,l}(x_{j(x)})) - E(D_{p,l}(x_{j(x)}))\right| \\ &+ \sup_{x\in S_{\mathcal{F}}} \left|E(D_{p,l}(x_{j(x)})) - E(D_{p,l}(x))\right| \\ &:= \sum_{i=1}^{3} L_i^{p,l}.\end{aligned}$$

Study of the term $L_2^{p,l}$.

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

$$\begin{aligned}P\left(L_2^{p,l} > \varepsilon\right) &= P\left(\max_{j\in\{1,\ldots,d_n\}} |D_{p,l}(x_j) - ED_{p,l}(x_j| > \varepsilon\right) \\ &\leq d_n \max_{j\in\{1,\ldots,d_n\}} P\left(|D_{p,l}(x_j) - ED_{p,l}(x_j| > \varepsilon\right) \\ &\leq d_n \max_{j\in\{1,\ldots,d_n\}} P\left(\left|\sum_{i=1}^{n} \Delta_i^{p,l}(x_j)\right| > n\Phi(h)\varepsilon\right),\end{aligned}$$

where $\Delta_i^{p,l}(x)$ is defined in (3.3). By applying Proposition A.11-i in Ferraty and Vieu (2006) and since $E|\Delta_1^{(k,l)}(x)|^q = O(\Phi_x(h))$ for $q > 2$, we have, there exist $q > 2$, for any $\varepsilon > 0$, $r \geq 1$ and for some $0 < C < \infty$

$$P\left(L_2^{p,l} > \varepsilon\right) \leq C(A_1 + A_2),$$

where

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

and $G_{n,l,p}^2 := \sup_{x \in S_{\mathcal{F}}} G_{n,l,p}^2(x) = O(n\Phi(h))$ in view of relation (3.4) together with hypothesis (U1).

Choosing for $\eta > 0$

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

we obtain

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

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

Hence, we get for η large enough

$$P\left(L_2^{p,l} > \eta \sqrt{\frac{\ln n}{n\Phi(h)}}\right) \leq C n^{-1-\xi},$$

where $\xi > 0$.

Thus

$$L_2^{p,l} = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi(h)}}\right). \tag{3.18}$$

Study of the terms $L_1^{p,l}$ and $L_3^{p,l}$.

First, let us analyze the term $L_1^{p,l}$. Since K is supported in $[0,1]$ and according to (U1), we can write

$$L_1^{p,l} \leq \frac{C r_n}{n h \Phi(h)} \sup_{x \in S_{\mathcal{F}}} \sum_{i=1}^{n} |Z_i^l| \delta_i^l S^{-l}(Z_i) 1_{B(x,h) \cup B(x_{j(x)},h)}(X_i).$$

Let
$$\xi_i = \frac{Cr_n|Z_i^l|\delta_i^l S^{-l}(Z_i)}{h\Phi(h)} \sup_{x\in S_{\mathcal{F}}} 1_{B(x,h)\cup B(x_{j(x)},h)}(X_i).$$
In the same manner as for proving (3.4), we have under hypotheses (U1), (U5b), (U6) and (U8)
$$G_n^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} |Cov(\xi_i,\xi_j)| = O(n\Phi(h)).$$
It remains to use similar arguments as to treat $L_2^{p,l}$ to obtain
$$L_1^{p,l} = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi(h)}}\right). \tag{3.19}$$
Second, since
$$L_3^{p,l} \leq E\left(\sup_{x\in S_{\mathcal{F}}} \left|D_{p,l}(x) - D_{p,l}(x_{j(x)})\right|\right),$$
we deduce that
$$L_3^{p,l} = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi(h)}}\right). \tag{3.20}$$
Finally, the result of Lemma 3.7 follows from the relations (3.19), (3.18) and (3.20). □

Lemma 3.8. *(see Messaci et al. (2015)) Assume that assumptions (U1), (U6) are satisfied, we get*
$$\sup_{x\in S_{\mathcal{F}}} |\widehat{m}_0(x) - 1| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi(h)}}\right) \text{ and } \sum_{n=1}^{\infty} P\left(\inf_{x\in S_{\mathcal{F}}} \widehat{m}_0(x) < \frac{1}{2}\right) < \infty.$$
Lemma 3.9. *Under assumptions (A1), (U1), (U6), we obtain*
$$\sup_{x\in S_{\mathcal{F}}} |\widehat{m}_1(x) - \widetilde{m}_1(x)| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi(h)}}\right).$$
Proof. By the relation (3.15), we can write
$$\sup_{x\in S_{\mathcal{F}}} |\widehat{m}_1(x) - \widetilde{m}_1(x)| \leq C \sup_{t\leq T} |S_n(t) - S(t)| \sup_{x\in S_{\mathcal{F}}} |\widehat{m}_0(x)|.$$
The proof is completed by the relation (3.16) and Lemma 3.8. □

Chapter 4

Simulation study

In this chapter, we carry out a simulation study in order to compare the performances of the local linear methodology with those of the kernel method, in the tree cases : Functional and complete case under dependant condition (see 4.1), Functional and censored case under independant condition (see 4.2) and Functional and censored case under dependant condition (see 4.3).

4.1 Functional and complete case under dependant condition

In this section, two examples of simulation are presented to illustrate the performance of the proposed estimator (LLR), for functional and complet data under dependant condition. More precisely, we compare the LLR estimator to the kernel regression estimator (KR) studied in Ferraty and Vieu (2006). For the computation of the (LLR) and the (KR) estimators, we use the quadratic kernel $K(x) = \frac{3}{2}(1-x^2)1_{[0,1]}(x)$ and the bandwith h is chosen by the 2-fold cross-validation method. Take into account of the smoothness of the curves $X_i(t)$ (see Figure 4.6), we choose the semi-metric d based on the derivative described in Ferraty and Vieu (2006) (see routines "semimetric.deriv" and "semimetric.pca" in the website http://www.lsp.ups-tlse.fr/staph/npfda) and we take $\beta = d$ (for the LLR estimator).

4.1.1 Example

Let us consider the following nonparametric regression model

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

where

$$m(X) = \frac{1}{4} \exp \left\{ 2 - \frac{1}{\left(\int_0^1 X'(t)dt \right)^2} \right\}.$$

and ϵ is the error supposed to be generated by an autoregressive model defined by

$$\epsilon_i = \frac{1}{\sqrt{2}}(\epsilon_{i-1} + \xi_i), \quad i = 1, \cdots, n$$

with ξ_i are centered random variables normally distributed (i.i.d.) with a variance equal to 0.1 ($\xi_i \rightsquigarrow \mathcal{N}(0, 0.1)$). The functional covariate $X(t)$ is defined, for $t \in [0, \pi/3]$ by

$$X(t) = 2 - cos\left(W\left(t - \frac{2\pi}{3}\right)\right), \quad t \in [0, \frac{2\pi}{3}]$$

where W is an α-mixing process generated by $W_i = \frac{2}{9}W_{i-1} + \eta_i$ with η_i are i.i.d $\mathcal{N}(0,1)$ and are independent from W_i, which is generated independently by $W_0 \rightsquigarrow \mathcal{N}(0,1)$ (see Figure 4.1 below for a sample of these curves). Notice that the conditional mean function will coincide and will be equal to $m(x)$.

In this simulation, to illustrate the performance of our estimator, we proceed as follows:

• Step 1. For a different sample sizes $n = 100, 200, 300, 500$, we split our data

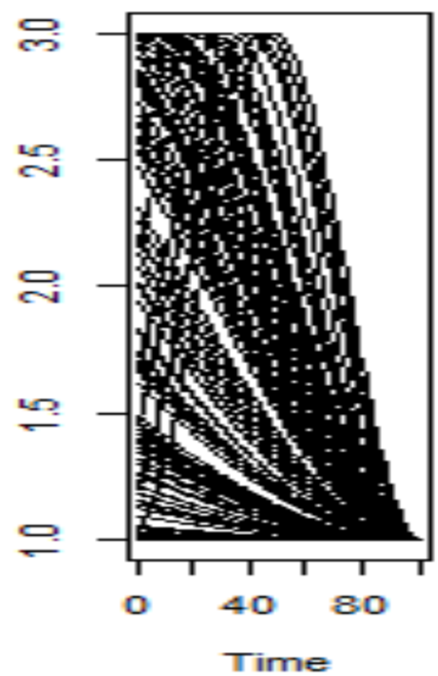

Figure 4.1: The curves X_i.

into two subsets:

- $(X_i, Y_i)_{1\le i\le n_1}$: The learning sample used to build the estimators, where $n_1 = n/2$.

- $(X_i, Y_i)_{n_2\le i\le n}$: The testing sample used to make a comparison, with $n_2 = n_1 + 1$.

• Step 2. We calculate the two estimators by using the learning sample and we find the LLR and KR estimators of the conditional expectation ($\widehat{m}$ and $\widehat{m}_{KR}$), for a different sample sizes $n = 100, 200, 300, 500$.

• Step 3. We plot the true values $m(X_i)$ for all i $(n_2 \le i \le n)$ against the predicted ones by means of the two estimators, one in each graph (for a fixed sample size $n = 100$, see Figure 4.2).

• Step 4. To be more precise, we measure the prediction accuracy, for

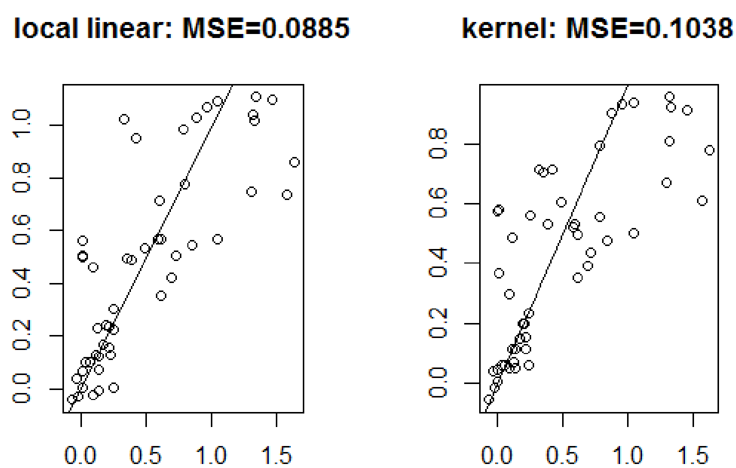

Figure 4.2: From left to right representation of the LLR and KR estimators $(n = 100)$.

different values of n, by using the mean absolute errors (MAE), given by

$$\begin{cases} MAE(LLR) := \frac{1}{n_2+1}\sum_{j=n_2}^{n} |\widehat{m}(X_j) - m(X_j)| \\ MAE(KR) := \frac{1}{n_2+1}\sum_{j=n_2}^{n} |\widehat{m}_{KR}(X_j) - m(X_j)| \end{cases}$$

and the prediction errors (MSE) such that

$$\begin{cases} MSE(LLR) := \frac{1}{n_2+1}\sum_{j=n_2}^{n} (\widehat{m}(X_j) - m(X_j))^2 \\ MSE(KR) := \frac{1}{n_2+1}\sum_{j=n_2}^{n} (\widehat{m}_{KR}(X_j) - m(X_j))^2 \end{cases}$$

The obtained results are in the table 4.1.

Table 4.1: MSE and MAE comparaison for LLR and KR methods according to sample sizes.

	$n = 100$		$n = 200$		$n = 300$		$n = 500$	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
LLR	0.0885	0.1998	0.0695	0.1686	0.0536	0.1471	0.0369	0.0962
KR	0.1038	0.2635	0.0764	0.1853	0.0674	0.1541	0.0562	0.1422

From Table 4.1 and Figure 4.2, we observe that the quality of the two estimators perform better when the sample size n increase. Also, we can be seen that our predictor has a good behavior than the kernel one for functional and complete data under dependant condition.

4.2 Functional and censored case under independant condition

In this section, we conduct two examples of simulation to illustrate the performance of the local linear regression estimator (LLR) studied in Leulmi

(2020) (see chapter 2), for functional and censored data under independant condition. for further illustration. More precisely, we compare the LLR estimator to the conditional expectation kernel estimator (KR) studied in Ling (2016).
For the computation of the LLR and the KR estimators, we use the quadratic kernel $K(x) = \frac{3}{2}(1 - x^2)1_{[0,1]}(x)$ and the bandwith h is chosen by the 2-fold cross-validation method. The semi-metric d is based on the derivative described in Ferraty and Vieu (2006)(see routines "semimetric.deriv" in the website http://www.lsp.ups-tlse.fr/staph/npfda) and we take $\beta = d$ (for the LLR estimator).

4.2.1 Example

We consider a functional covariate $X(t)$ on the interval $[0, 1]$

$$X(t) = A\,(2 - cos(\pi t W)) + (1 - A)\,cos(\pi t W),$$

where W is a centered random variable normally distributed with a variance equal to 1 and A is a random variable having a Bernoulli distribution with parameter $p = 0.5$. We carried out the simulation with a 400-sample of the curve X which is represented in the Fig. 4.3.
The scalar response variable is defined as

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

where X and ϵ are independent, the error ϵ is a centered random variable normally distributed with a variance equal to 0.1 and

$$r(X) = \frac{1}{4}\int_0^1 (X'(t))^2\,dt.$$

Given $X = x$, we can easily see that $Y \rightsquigarrow \mathcal{N}\,(r(x), 0.1)$, and therefore, the conditional mean functions will coincide and will be equal to $r(x)$.

Moreover, as a first model, the censoring variable C has an exponential distribution with parameter $\lambda = 1.5$. In the second model, C is distributed as Weibull distribution with parameters $k = 2$ and $\lambda = 0.5$.

Under these two models, we compute the kernel regression estimator (KR) studied in Ling (2016) and defined by

$$\widehat{m}_{KR}(x) = \frac{\sum_{i=1}^{n} K(h^{-1}d(X_i, x))Z_i\delta_i S_n^{-1}(Z_i)}{\sum_{i=1}^{n} K(h^{-1}d(X_i, x))} \qquad \left(\frac{0}{0} := 0\right) \tag{4.1}$$

and our LLR estimator $\widehat{m}(x)$ on the basis of the sample $(X_i, Z_i, \delta_i)_{1\le i\le n}$, where $Z_i = \min(Y_i, C_i)$ and $\delta_i = 1_{\{Y_i \le C_i\}}$.

In this simulation, to illustrate the performance of our estimator, we proceed as follows:

• Step 1. We split our data into two subsets:

- $(X_i, Y_i)_{1\le i\le 300}$: The learning sample used to build the estimators.
- $(X_i, Y_i)_{301\le i\le 400}$: The testing sample used to make a comparison.

• Step 2. We calculate the two estimators by using the learning sample and we find the LLR and KR estimators of the conditional expectation ($\widehat{m}$ and $\widehat{m}_{KR}$).

• Step 3. We plot the true values ($r(X_i)$) for all i ($301 \le i \le 400$) against the predicted ones by means of the two estimators (one in each graph), this is displayed in Fig. 4.3 for the first model and Fig. 4.4 for the second model.

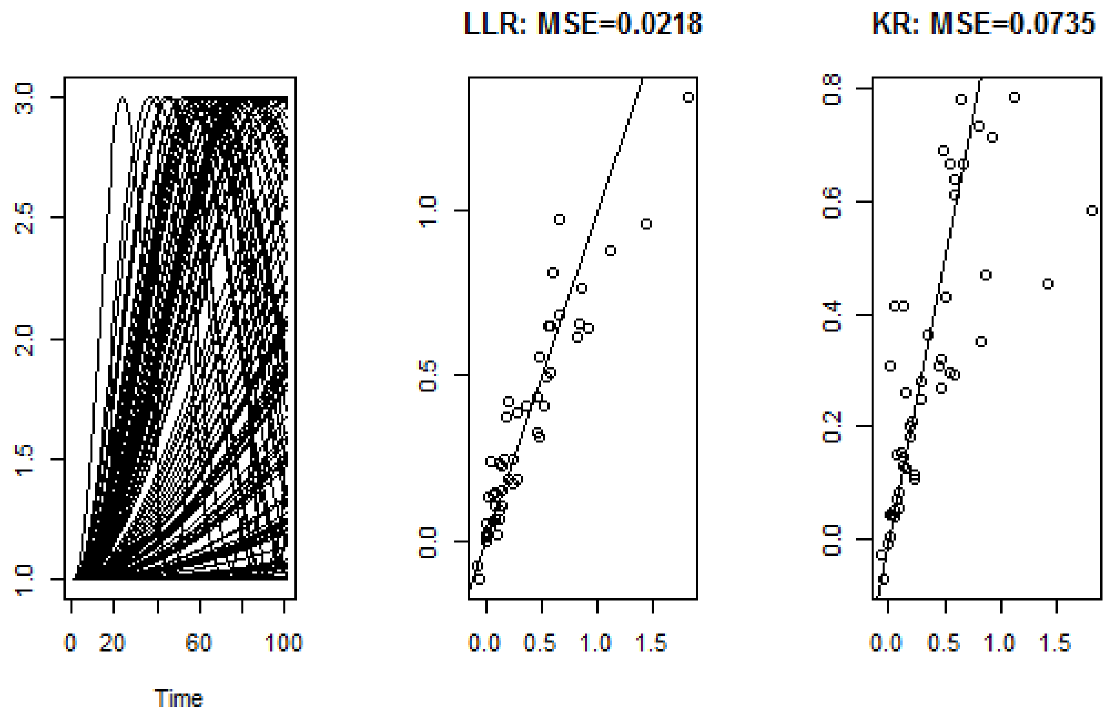

Figure 4.3: From left to right the curves X_i, the LLR estimator and the KR estimator for the first data model.

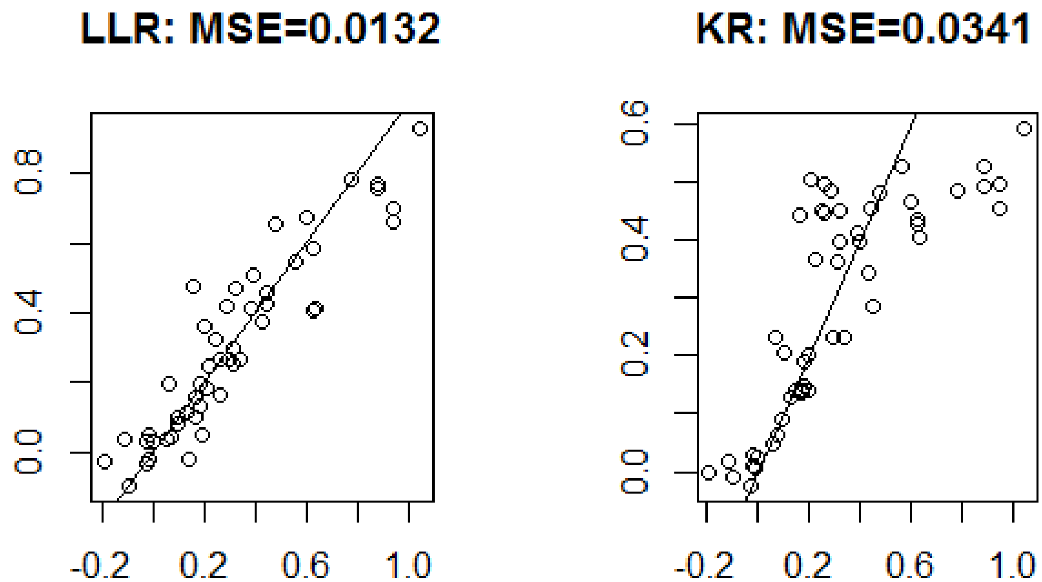

Figure 4.4: Representation of the LLR and the KR estimators for the second data model.

- Step 4. To be more precise we evaluate the prediction errors given by

$$MSE(LLR) := \frac{1}{100}\sum_{j=301}^{400}\left(\widehat{m}(X_j) - r(X_j)\right)^2 \quad and \quad MSE(KR) := \frac{1}{100}\sum_{j=301}^{400}\left(\widehat{m}_{KR}(X_j) - r(X_j)\right)^2$$

and the mean absolute errors (MAE) defined by

$$MAE(LLR) := \frac{1}{100}\sum_{j=301}^{400}|\widehat{m}(X_j) - r(X_j)| \quad and \quad MAE(KR) := \frac{1}{100}\sum_{j=301}^{400}|\widehat{m}_{KR}(X_j) - r(X_j)|.$$

The obtained results are in the table below (see talbe 4.2.

Table 4.2: MSE comparaison for KR and LLR methods for the two models.

	The first model		The second model	
	MSE	MAE	MSE	MAE
LLR	0.0218	0.1239	0.0132	0.0831
KR	0.0735	0.1652	0.0341	0.1296

To make a decision, we choose an other example.

4.2.2 Example

In this example, The functional covariate X is generated by the following equation:

$$X(t) = 2 - cos\left(W\left(t - \frac{2\pi}{3}\right)\right), \quad t \in [0, \frac{2\pi}{3}]$$

where W is a random variable having a standard normal distribution. The curves are discretized on the same grid which is composed of 215-equidistant values in $[0, \frac{2\pi}{3}]$ (see Fig. 4.5).

The scalar response is defined as

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

where X and ϵ are independent, the error ϵ is a centered random variable normally distributed with a variance equal to 0.1 and

$$r(X) = \frac{1}{4}\exp\left\{2 - \frac{1}{\left(\int_0^1 X'(t)dt\right)^2}\right\}.$$

For this model, we adopt the censored mechanism $(X_i, Z_i, \delta_i)_{1\leq i\leq n}$, where $Z_i = \min(Y_i, C_i)$, $\delta_i = 1_{\{Y_i\leq C_i\}}$ and the censoring random variable C has an exponential distribution with parameter λ which is adapted in order to get different censoring rates (CR).

In this simulation, to illustrate the efficiency of our estimator, we proceed as follows:

• Step 1. We divide our sample of size 215 into the learning sample $(X_i, Y_i)_{1\leq i\leq 108}$ and the testing $(X_i, Y_i)_{109\leq i\leq 215}$

• Step 2. We calculate the two estimators by using the learning sample and we find the LLR and KR estimators of the conditional expectation ($\widehat{m}$ and $\widehat{m}_{KR}$).

• Step 3. We plot the true values $(r(X_i))$ for all i $(109 \leq i \leq 215)$ against the predicted ones by means of the two estimators (one in each graph) ($CR = 28.04\%$, see Fig. 4.5).

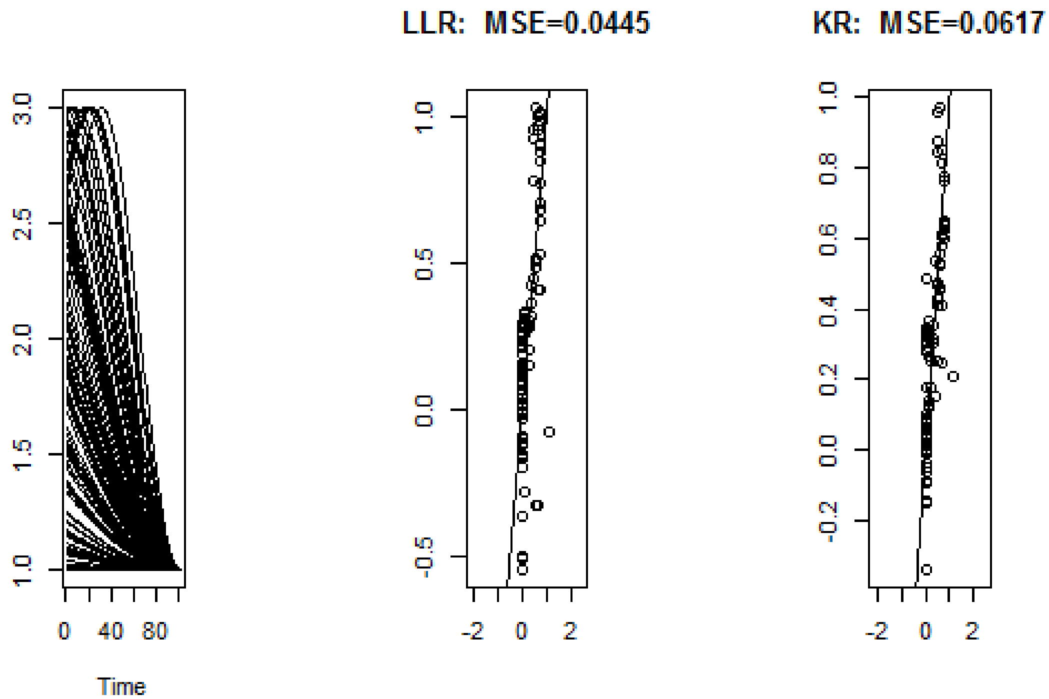

Figure 4.5: From left to right the curves X_i, the KR estimator and the LLR estimator ($CR = 28.04\%$).

- Step 4. To be more precise, we measure the prediction accuracy, for different values of CR, by using the mean absolute errors (MAE), given by

$$MAE(LLR) := \frac{1}{107}\sum_{j=109}^{215} |\widehat{m}(X_j) - r(X_j)| \quad and \quad MAE(KR) := \frac{1}{107}\sum_{j=109}^{215} |\widehat{m}_{KR}(X_j) - r(X_j)|,$$

as well as the prediction errors (MSE) such that

$$MSE(LLR) := \frac{1}{107}\sum_{j=109}^{215} (\widehat{m}(X_j) - r(X_j))^2 \quad and \quad MSE(KR) := \frac{1}{107}\sum_{j=109}^{215} (\widehat{m}_{KR}(X_j) - r(X_j))^2 .$$

The obtained results are in the table below.

Table 4.3: MSE and MAE comparaison for LLR and KR methods according to CR.

	$CR = 1.87\%$		$CR = 28.04\%$		$CR = 49.53\%$		$CR = 73.83\%$	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
LLR	0.0056	0.0419	0.0445	0.1234	0.1337	0.243	0.1594	0.2746
KR	0.0125	0.0652	0.0617	0.1462	0.1409	0.2448	0.1595	0.2755

Note from Table 4.3 that the MSE and the MAE values for both kernel and local linear methods become more important when the CR increases. In addition, we remark that the local linear estimator performs better than the kernel estimator. (see Figs. 4.3, 4.4 and 4.5 and Tables 4.2 and 4.3).

Conclusion and comments

In conclusion, We remark that our estimator has a good performance and seems to outperform the kernel estimator even for censored data. Moreover, we note that the prediction accuracy of the regression function decreases whenever the CR increases.

4.3 Functional and censored case under dependant condition

In this section, two examples of simulation are presented to illustrate the performance of the proposed estimator (LLR), for functional and censored data under dependant condition. More precisely, we compare the LLR estimator to the kernel regression estimator (KR) studied in Ling (2016).

For the computation of the (LLR) and the (KR) estimators, we use the quadratic kernel $K(x) = \frac{3}{2}(1-x^2)1_{[0,1]}(x)$ and the bandwith h is chosen by the 2-fold cross-validation method. Take into account of the smoothness of the curves $X_i(t)$ (see Figure 4.6 and 4.7), we choose the semi-metric d based on the derivative (for the first example) and the PCA (for the second example) described in Ferraty and Vieu (2006) (see routines "semimetric.deriv" and "semimetric.pca" in the website http://www.lsp.ups-tlse.fr/staph/npfda) and we take $\beta = d$ (for the LLR estimator).

4.3.1 Example

Let us consider the following nonparametric regression model

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

where

$$m(X) = \frac{1}{4}\exp\left\{2 - \frac{1}{\left(\int_0^1 X'(t)dt\right)^2}\right\}.$$

and ϵ is the error supposed to be generated by an autoregressive model defined by

$$\epsilon_i = \frac{1}{\sqrt{2}}(\epsilon_{i-1} + \xi_i), \quad i = 1, \cdots, n$$

with ξ_i are centered random variables normally distributed (i.i.d.) with a variance equal to 0.1 ($\xi_i \rightsquigarrow \mathcal{N}(0, 0.1)$). The functional covariate $X(t)$ is defined, for $t \in [0, \pi/3]$ by

$$X(t) = 2 - cos\left(W\left(t - \frac{2\pi}{3}\right)\right), \quad t \in [0, \frac{2\pi}{3}]$$

where W is an α-mixing process generated by $W_i = \frac{2}{9}W_{i-1} + \eta_i$ with η_i are i.i.d $\mathcal{N}(0,1)$ and are independent from W_i, which is generated independently by $W_0 \rightsquigarrow \mathcal{N}(0,1)$ (see Figure 4.6 for a sample of these curves). Notice that the conditional mean function will coincide and will be equal to $m(x)$.
For this model, we adopt the censored mechanism $(X_i, Z_i, \delta_i)_{1\le i\le n}$, where $Z_i = \min(Y_i, C_i)$, $\delta_i = 1_{\{Y_i \le C_i\}}$ and the censoring random variable $C_i =$

$a_i C_{i-1} + \zeta_i$ with $a_i \rightsquigarrow \mathcal{N}(0, 0.1)$ and ζ_i are i.i.d. exp(1.5) and are independent from C_i, which is generated independently by $C_0 \rightsquigarrow \exp(1.5)$.

In this simulation, to illustrate the performance of our estimator, we proceed as follows:

• Step 1. For a different sample sizes $n = 100, 200, 300, 500$, we split our data into two subsets:

- $(X_i, Y_i)_{1 \le i \le n_1}$: The learning sample used to build the estimators, where $n_1 = n/2$.

- $(X_i, Y_i)_{n_2 \le i \le n}$: The testing sample used to make a comparison, with $n_2 = n_1 + 1$.

• Step 2. We calculate the two estimators by using the learning sample and we find the LLR and KR estimators of the conditional expectation ($\widehat{m}$ and $\widehat{m}_{KR}$), for a different sample sizes $n = 100, 200, 300, 500$.

• Step 3. We plot the true values $m(X_i)$ for all i ($n_2 \le i \le n$) against the predicted ones by means of the two estimators, one in each graph (for a fixed sample size $n = 300$, see Figure 4.6).

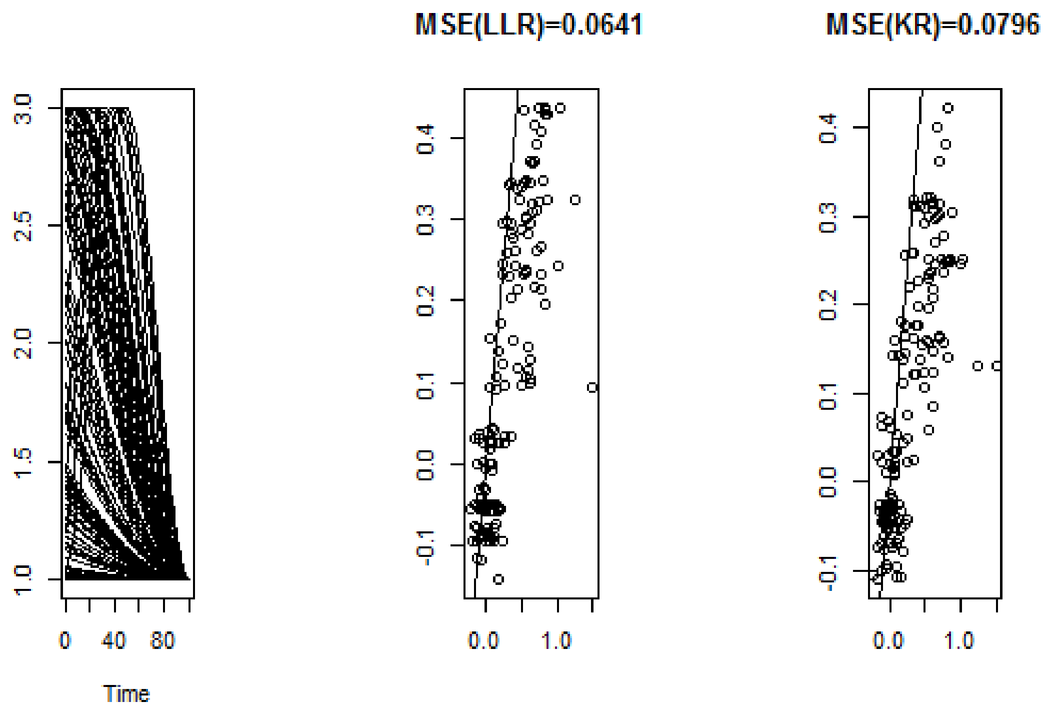

Figure 4.6: From left to right the curves X_i, the LLR and KR estimators ($n = 300$).

• Step 4. To be more precise, we measure the prediction accuracy, for different values of n, by using the mean absolute errors (MAE), given by

$$\begin{cases} MAE(LLR) := \frac{1}{n_2+1}\sum_{j=n_2}^{n} |\widehat{m}(X_j) - m(X_j)| \\ MAE(KR) := \frac{1}{n_2+1}\sum_{j=n_2}^{n} |\widehat{m}_{KR}(X_j) - m(X_j)| \end{cases}$$

and the prediction errors (MSE) such that

$$\begin{cases} MSE(LLR) := \frac{1}{n_2+1}\sum_{j=n_2}^{n} (\widehat{m}(X_j) - m(X_j))^2 \\ MSE(KR) := \frac{1}{n_2+1}\sum_{j=n_2}^{n} (\widehat{m}_{KR}(X_j) - m(X_j))^2 \end{cases}$$

The obtained results are in the table 4.4.

Table 4.4: MSE and MAE comparaison for LLR and KR methods according to sample sizes.

	$n = 100$		$n = 200$		$n = 300$		$n = 500$	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
LLR	0.0896	0.2008	0.0775	0.1796	0.0641	0.1529	0.0396	0.1062
KR	0.1190	0.2338	0.0867	0.1933	0.0796	0.1590	0.0471	0.1529

From Table 4.4 and Figure 4.6, we observe that the quality of the two estimators perform better when the sample size n increase. Also, we can be seen that our predictor has a good behavior than the kernel one.
We preffer to give a second example to make a better decision.

4.3.2 Example

We fixe $n = 200$ and we generated the functional explanatory variables $X(t)$ as follows

$$X_i(t) = a_i \sin\left(4(b_i - t)\right) + c_i, \quad i = 1, \cdots, 200$$

where $a_i \rightsquigarrow \mathcal{N}(4, 3)$, $c_i \rightsquigarrow \mathcal{N}(0, 0.01)$ and b_i is an α-mixing process generated by $b_i = \frac{1}{3}a_{i-1} + \eta_i$ with η_i are i.i.d. $\mathcal{N}(0, 1)$ and are independent from b_i, which is generated independently by $b_0 \rightsquigarrow \mathcal{N}(0, 3)$. We carried out the simulation with a 300-sample of the curve $X(t)$ (see 4.7).
The scalar response variable is defined as

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

where

$$m(X) = \int_0^1 \frac{1}{1 + |X(t)|} dt$$

and ϵ is the error generated by an autoregressive model defined by

$$\epsilon_i = \frac{1}{\sqrt{2}}\,\epsilon_{i-1} + \xi_i, \quad i = 1, \cdots, 200$$

with $\xi_i \rightsquigarrow \mathcal{N}(0, 0.1)$. Notice that the conditional median function will coincide and will be equal to $m(x)$.
We also simulate n i.i.d. rondom (C_i) exponentially distributed with parameter λ which is adapted in order to get different censoring rates (CR). We compute our estimator with the observed data $(X_i, Z_i, \delta_i)_{1\leq i\leq n}$, where $Z_i = \min(Y_i, C_i)$ and $\delta_i = 1_{\{Y_i \leq C_i\}}$. Next, we split our data into a learning sample with size 135 and a test sample with size 65. The true values are plotted against the predicted ones by means of our estimator $\widehat{m}(x)$ and the kernel estimator $\widehat{m}_{KR}(x)$ ($CR = 1.48\%$).

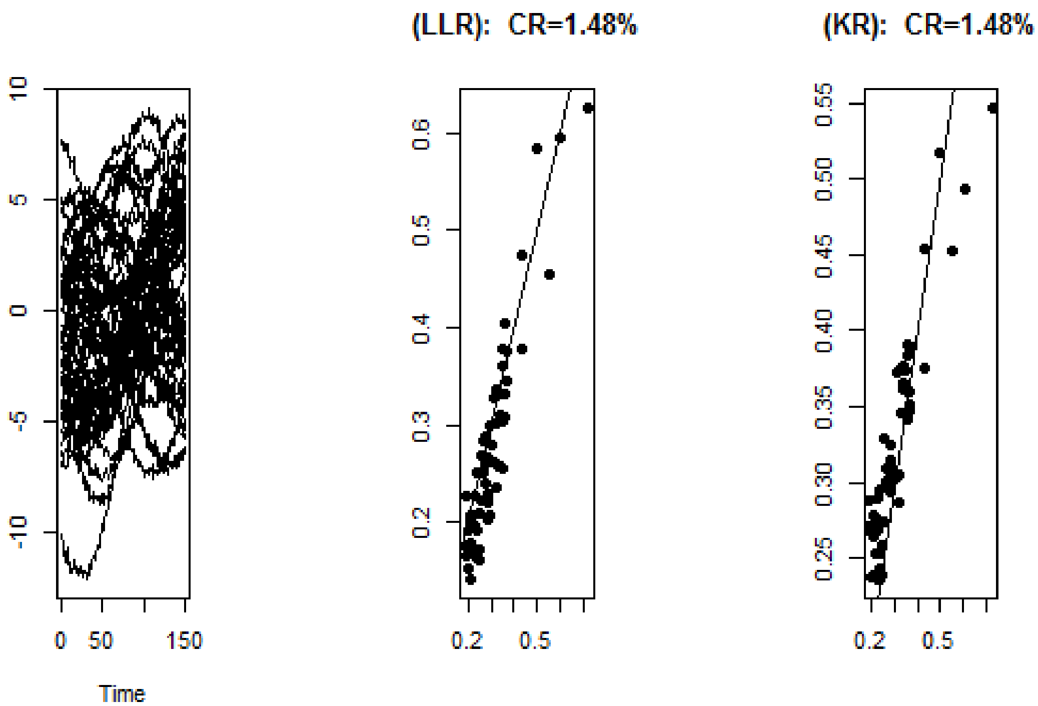

Figure 4.7: From left to right the curves X_i, the LLR and KR estimators ($CR = 1.48\%$).

To be more precise, we measure the prediction accuracy, for different values of CR, by using the mean absolute errors (MAE), given by

$$\begin{cases} MAE(LLR) := \frac{1}{65}\sum_{j=136}^{200} |\widehat{m}(X_j) - m(X_j)| \\ MAE(KR) := \frac{1}{65}\sum_{j=136}^{200} |\widehat{m}_{KR}(X_j) - m(X_j)| \end{cases}$$

and the prediction errors (MSE) such that

$$\begin{cases} MSE(LLR) := \frac{1}{65}\sum_{j=136}^{200} (\widehat{m}(X_j) - m(X_j))^2 \\ MSE(KR) := \frac{1}{65}\sum_{j=136}^{200} (\widehat{m}_{KR}(X_j) - m(X_j))^2 \end{cases}$$

The obtained results arein the table 4.5.

Figure 4.7 and Table 4.5 show that, our estimator performs better than

Table 4.5: MSE and MAE comparaison for LLR and KR methods according to CR.

	$CR = 1.48\%$		$CR = 28.67\%$		$CR = 48.15\%$		$CR = 73.33\%$	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
LLR	0.0019	0.0331	0.0182	0.1044	0.0260	0.1271	0.05458	0.2106
KR	0.0037	0.0353	0.0220	0.1098	0.0295	0.1474	0.0610	0.2314

the kernel estimator. It is also clear that, the quality of the both estimators become slightly worse when we have high percentage of censoring, however it remains acceptable.

Conclusion and comments

In conclusion, our Our theoretical and practical studies confirmed without surprise that the quality of the LLR and the KR estimators are better for a bigger sample size n and a weak rate of censoring CR. Furthermore, as for independent and censored data, the LLR estimator stay more accurate than the KR one in all cases.

Perspectives

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

• Show the almost complete convergence results (pointwise and uniform) for the condtional quantile and coditional mode for functional, censored and α-mixing data.

• Study the quadratic mean convergence and the asymptotic normality for functional, censored and α-mixing data.

• Establish the almost complete convergence results when the explanatory variable is valued in fuctional space in the setting of the association dependence condition.

Bibliography

Ait Hennania, L., Lemdania, M., and Ould Said, E. (2018). Robust regression analysis for a censored response and functional regressors. *Journal of Nonparametric Statistics*, https ://doi.org/10.1080/10485252.2018.15463 86.

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

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

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

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

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

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

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

Benkhaled, A., Madani, F., Khardani, S., (2020). Strong consistenand cy of local linear estimation of a conditional density function under random censorship, *Arabian Journal of Mathematics*, https://doi.org/10.1007s40065-020-00282-1.

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

Bouchentouf, A., Hamza, A., and Rabhi, A. (2017). Strong uniform consistency rates of conditional hazard estimation in the single functional index model for dependant functional data under random censorship. *International Journal of Statistics and Economics*, 18, 82–101.

Bitouzé, D., Laurent, B. & Massart, P. (1999). A dvoretzky-kiefer-wolfowitz type inequality for the Kaplan-Meier estimator. *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*. 35: 735-763.

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

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

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

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

Cai, Z. , (2001). *E*stimating a distribution function for censored time series data, *Journal of Multivariate Analysis*, 78, 299-318.

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

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

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

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

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

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

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

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

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

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

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

Chaouch, M. and Khardani, S. (2014). Randomly censored quantile regression estimation using functional stationary ergodic data. *Journal of Nonparametric Statistics*, https ://doi.org/10.1080/10485252.2014.982651.

El Ghouch, A., and Van Keilegom, I. (2009). Local linear quantile regression with dependent censored data. *Statistica Sinica*, 19, 1621–1640.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Gannoun, A., J. Saracco, A. Yuan, and G. Bonney. 2005. Non-parametric quantile regression with censored data. *Scandinavian Journal of Statistics*, 32 (4), 527–50. doi:10.1111/j.1467-9469.2005.00456.x.

Guessoum, Z. Ould Said, E. (2008). On nonparametric estimation of the regression function under random censorship model, *Statistics & Decisions*, 26, 159-177.

Grenander U. (1950). Stochastic processes and statistical inference. *Arkiv for Matematik*, 1, 195–277.

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

Horrigue, W., and Ould Said, E. (2011). Strong uniform consistency of a nonparametric estimator of a conditional quantile for censored dependent data and functional regressors conditional quantile for functionnal times series. *Random operators and stochastic equations*, 131–156. Basel: Birkhusser Verlag.

Horrigue, W., and Ould Said, E. (2015). Nonparametric regression quantile estimation for dependent functional data under random censorship: Asymptotic normality. *Communications in Statistics-Theory and Methods*, 44, 4307–4332.

Kadiri, N., Rabhi, A., and Bouchentouf, A. A. (2018). Strong uniform consistency rates of conditional quantile estimation in the single functional index model under random censorship. *Dependent Modeling*, 16, 197–227.

Kaplan, E. M., Meier, P., (1958). Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association.* 53, 457-481.

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

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

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

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

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

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

Leulmi, F. Leulmi, S. Kharfouchi S.(2022). On the nonparametric estimation of the functional regression based on censored data under strong mixing condition *Journal of Siberian Federal University*, 15(4), 523–536.

Leulmi, S., (2019). Local linear estimation of the conditional quantile for censored data and functional regressors. *Communications in Statistics-Theory and Methods*, 1-15.

Leulmi, S., (2020). Nonparametric local linear regression estimation for censored data and functional regressors. *Journal of the Korean Statistical Society* (2020), 1-22.

Leulmi, S., Messaci, F., (2018). Local linear estimation of a generalized regression function with functional dependent data. *Communications in Statistics - Theory and Methods* 47 no.23, 5795-5811.

Leulmi, S., Messaci, F., (2019). A class of local linear estimators with functional data *Journal of Siberian Federal University* 12, 379-391.

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

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

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

Ling, N., Liu, Y., (2016) The kernel regression estimation for randomly censored functional stationary ergodic data, *Communications in Statistics-Theory and Methods*, DOI10.1080/03610926.2016.1185117.

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

Mechab, B., Hamidi, N., Benaissa,S. (2019) Nonparametric estimation of the relative error in functional regression and censored data, *Chilean Journal of Statistics*, 10 , no.2, 177-195.

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

Rachdi, M., Laksaci, A., Demongeot, J., Abdali, A., and Madani, F. (2014). Theoretical and Practical Aspects of the Quadratic Error in the Local Linear Estimation of the Conditional Density for Functional Data. *Computational Statistics & Data Analysis,*, 73, 53–68.

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

Ramsay JO. (1982). When the data are functions. *Psychometrika*, 47, 379–396.

Ramsay JO and Dalzell C. (1991). Some tools for functional data analysis. Journal of the Royal Statistical Society. *Series B (Methodological)*, 539–572.

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

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

Rao CR. (1958). Some statistical methods for comparison of growth curves. *Biometrics*, 14, 1–17.

Rouabah, N., Nemouchi, N., Messaci, F., (2018). A rate of consistency for nonparametric estimators of the distribution function based on censored dependent data, *Statistical Methods & Applications*, https://doi.org/10.1007/s10260-018-00445-7.

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

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

Ould Said, E. (2006). A strong uniform convergence rate of kernel conditional quantile estimator under random censorship. *Statistics & Probability Letters*, 76, 579–586

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

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

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

Abstract

In this thesis, we consider the problem of the nonparametric estimation of the regression function when the response variable is real and the regressor is valued in a functional space (space of infinite dimension), by using the local linear method.

Firstly, we suppose that the observations are strongly mixing and we establish the pointwise and the uniform almost complete convergence, with rates, of the local linear regression estimator.

Secondly, we consider a sequence of independent and identically distributed observations and we introduce a local linear nonparametric estimation of the regression function for a censored scalar response random variable. Then, we establish their pointwise and the uniform almost-complete convergences, with rates.

Our main results is based on the functional and censored data under strong mixing condition and we study the rate of the pointwise almost-complete convergence of the local linear regression estimator.

Finally, a simulation study illustrates the performance of the local linear methodology with respect to other kernel method, in the tree cases: Functional and complete case under dependent condition, Functional and censored case under independent condition and Functional and censored case under dependent condition.

Keywords: Functional data, Censored data, Local linear Estimation, Almost-complete convergence, α-mixing.

Résumé

Dans cette thèse nous considérons le problème de l'estimation non paramétrique de la fonction de régression d'une variable réponse réelle conditionnée par une variable explicative fonctionnelle (á valeurs dans un espace de dimension infinie), par utilisation de la méthode locale linéaire.

Dans un premier temps, nous supposons que les observations sont fortement mélangeantes et nous étudions la convergence presque complète ponctuelle et uniforme, avec taux, de l'estimateur local linéair de la fonction de régression.

Puis, dans un second temps, nous considérons une suite d'observations indépendantes et identiquement distribuées et nous introduisons l'estimateur local linéaire de la fonction de régression dans le cas censuré. Ensuite, nous étudions leurs convergences presque complètes ponctuelle et uniforme.

Mes principaux résultats sont basés sur des données fonctionnelles et censurées sous la condition de mélange forte et nous étudions les vitesses de convergences presque complète ponctuelle et uniforme de l'estimateur local linéair de la fonction de régression.

Finalement, les études de simulations illustre la performance de la méthode locale linéaire par rapport Ã la méthode de noyau, dans les trois cas : le cas fonctionnel et complet sous la condition de dépendance, le cas fonctionnel et censuré sous la condition d'indépendance et le cas fonctionnel et censuré sous la condition de dépendance.

Mots clés: Données fonctionnelles, Données censurées, Estimation locale linéaire, Convergence presque compète, α-mélange.