N° d'ordre :

REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES EXACTES
SIDI BEL ABBES

THESE DE DOCTORAT

Présentée par :

Aicha HAMZA

Spécialité : Mathématiques
Option : Statistique

Intitulée

Étude asymptotique sur les modèles non-paramétriques conditionnels de durées de vie en statistique fonctionnelle.

Soutenue le :
Devant le jury composé de :

Président :
Mr. Boubaker MECHAB (Professeur à l'Université Djillali Liabès)
Examinateurs :
Mr. Toufik Guendouzi (Professeur à l'Université de Saida)
Mr. Fethi Madani (Professeur à l'Université de Saida)
Mr. Hakim OUADJED (M.C.A. à l'Université de Mascara)
Co-directeur de thèse :
Mr. Abbes RABHI (Professeur à l'Université Djillali Liabès)

Directeur de thèse :
Madame Amina Angelika BOUCHENTOUF (Professeur à l'Université Djillali Liabès)

Année universitaire : 2021/2022

Dedication

To my family.

Acknowledgments

All thanks to ALLAH, the most gracious, the most merciful for giving me strength to achieve this thesis. I would like to express my deep and sincere gratitude to my research supervisor, Professor Amina Angelika Bouchentouf who expertly guided me and shared the excitement of years of discovery enjoyable. My appreciation and acknowledgements also extends to my co-supervisor Professor Abbes Rabhi for giving me the opportunity to do this research, and for giving advice and invaluable help throughout this work. I want to thank very warmly, Professor Boubaker Mechab for chairing the committee members, Professor Toufik Guendouzi, Professor Fethi Madani and Doctor Hakim Ouadjed, who honoured me by agreeing to participate in my jury, thank you for the time spent and efforts in reviewing this work. My acknowledgements, to all those who directly or indirectly participated in the successful completion of this thesis. Finally, I owe debt gratitude to my family for their love, understanding, prayers and continuous support.

Contents

ملخص

المشكلة التي نتناولها في هذه الأطروحة تتعلق بالتقدير اللا معلمي لوظيفة الخطر الشرطي ذات متغير عشوائي دالي و متغير الاستجابة حقيقيًا و نعرض نتائجنا في حالة الملاحظات المرتبطة بنوع المعامل α تم النظر في حالتين ؛ البيانات الكاملة والبيانات الخاضعة للرقابة.

ندرس في ظل ظروف معينة ، نقطة التقارب شبه الكاملة والتقارب المنتظم شبه الكامل لمقدر النواة لهذا النموذج.

كلمات مفتاحية:

النماذج الشرطية؛ التقدير اللا معلمي؛ تحليل البيانات الوظيفية؛ احتمالية الكرات الصغيرة.

Abstract

The problem addressed in this thesis concerns nonparametric estimation of the conditional hazard function, when the explanatory variable is of functional nature and the response variable is real. Our results are presented in the case where the observations are strongly mixing ($\alpha - mixing$). Two cases are considered; complete data and censored data. We establish under certain conditions, the almost complete point convergence and the almost complete uniform convergence of the kernel estimator of this model.

Key words:

Conditional models, functional data analysis (FDA), nonparametric estimation, small ball probability.

Résumé

La problématique abordée dans cette thèse concerne l'estimation non paramétrique de la fonction de hasard conditionnelle, lorsque la variable explicative est fonctionnelle et la variable réponse est de type réel. Nos résultats sont présentés dans le cas où les observations sont fortement mélangeantes (α-mixing). Deux cas sont considérés; données complètes et données censurées.
Nous établissons sous certaines conditions, la convergence ponctuelle presque complète et la convergence uniforme presque complète de l'estimateur à noyau de ce modèle.

Mots clés:

Modèles conditionnels, estimation non paramétrique, analyse de données fonctionnelles (FDA), probabilité de petite boules petite.

Chapter 1

Introduction

In recent decades, functional statistics have become a very important field in statistical research. It is closely related to the study of data sets in which the observations are curves or surfaces. These are infinite-dimensional data which appear in many scientific fields such as meteorology, quantitative chemistry, biometrics, econometrics, medical imaging, and so on.

Estimation theory plays an important role in many fields such as finance, economics, medicine, weather forecasting, etc. In the statistical literature, two types of estimation are omnipresent; parametric estimation (the law of the random variable admits a general known form which depends on one or more unknown parameters to be estimated) and nonparametric estimation (information of the random variable law is so vague). Parametric parameters are introduced by Rosenblatt (1969) and Parzen (1962) to estimate a probability density, and by Nadaraya-Watson (1964) to estimate a regression function, which are called kernel estimators. The nonparametric study of functional data is much more recent than the parametric analysis. The first work in this subject was introduced by Ferraty and Vieu (2006).

This thesis concerns the problem of a conditional hazard function estimate in the single index under complete and censored data. Single-index models have a strong approximation capability in the way that any nonlinear relationship may be invariably detected by the model. So, instead of studying the relationship between a real response variable and a functional explanatory variable, we only have to observe the relation between two real variables: The response variable Y and the variable $<\theta, X>$. Further, it is known that the use of semi-metrics permits us to have very nice rates of convergence. The case considered in this investigation is when the functional space $\mathcal{F}$ is an Hilbert space with inner product $<,>$ and when d is the semi-metric constructed (for a fixed functional direction θ) as $d(x,y) =< x-y, \theta >$. Moreover, in addition to the dimension reduction, the functional index also plays a nice role in the interpretation of the data. In fact, when the covariate variances are small (close to zero), it is the same for the variance of the functional index, and vice versa. Further, nonparametric conditional models has become pertinent for the censored survival data analysis. It is known for its flexibility and ability to provide a more complete perspicacity into the stochastic relationship between variables. Unfortunately, in contexts with multivariate covariates, we are faced with the problem of the "curse of dimensionality". This makes it difficult to estimate conditional models. In the setting of survival analyses, the problem is further aggravated by the presence of censored observations. The most of the literature is dedicated to the case where the variable of interest is completely observed. This is not the case in many interesting applications, notably survival analysis, where censorship prevents the direct application of classical methods.

The remainder of this chapter is organized as follows: In Section 1.1, we provide a literature review on functional data. In Sections 1.2, 1.3, and 1.4, we give a fairly broad set of results on regression, conditional models and single index models, respectively. Finally, we present the contribution and outline of the thesis in Sections 1.5 and 1.6, respectively.

1.1 Functional data

The statistical issues concerned with modeling functional random variables study have recited a great advantage in statistics. The pioneer work is relying on the discretization of these functional observations to adapt traditional multivariate statistical techniques. While, due to the progress of the data-processing tool enabling the recovery of increasingly large data, an alternative has been developed consisting in treating this kind of data in its own dimension, that is to say by preserving the functional character.

In recent years, functional models have been very privileged topics. Within the linear framework, the contribution of Ramsay and Silverman (1997,2002) presented an important collection of statistical methods for the functional variables. In the same way, note that Bosq (2000) has significantly contributed to the development of statistical methods within the framework of process of auto-regression linear functional. By using functional principal components analysis, Cardot et al. (1999) built an estimator for the model of the Hilbertien linear regression similar to Bosq estimator (1991) in the case of Hilbertien process auto-regressive. This estimator is defined using the spectral properties of the empirical version of variance-covariance operator of the functional explanatory variable. They obtained convergence of probability for some cases and almost complete convergence of the built estimator for other cases. Norm convergence in L^2 for a regularized version (spline) of the preceding estimator was established by the same authors in 2000. A comprehensive overview on this field can be found in Hastie et al. (1995), Gasser et al. (1998), Hallet al. (1999), Ferraty and Vieu (2000), Besse et al. (2000), Dabo-Niang (2002), Hall and Heckman (2002), Dabo-Niang and Rhomari (2003), Cardot et al. (2003), Cardot et al. (2004), Ferraty et al. (2003), Cuevas et al. (2004), Ferraty and Vieu (2004), Aït-Saïd et al. (2005,2008), Ferré and Villa (2005), Ferraty Laksaci et al. (2013,2005), Gannoun et al. (2007), Geenens (2011), Ferraty et al. (2011), Laksaci et al. (2013), Ling and Vieu (2018), and Aneiros et al. (2019).

1.1.1 Application area

Increased interest in the application of statistical modeling to diverse domains including engineering, environmental science, biology, medicine, finance etc, has greatly been driven by the need for good data. It is important to note that these models will only be useful in the long term if they are accurate, based on good quality data and generated by the application of appropriate and robust statistical methods. Functional data analysis (FDA) is one such time series data modeling approach that has begun to gain attention in the literature, particularly in terms of public health and biomedical applications.

From Ullah and Finch (2013), we present in Table 1 a brief overview of some fields of application.

Table 1.1: Some fields of application of Functional data analysis (FDA)

Field of study	Outcome of interest	Reference
Engineering	Radar waveforms	Dabo-niang and Vieu (2007)
Biology	Temporal fertility trajectories of medfly	Muller et al. (2009)
	Time-course gene expression yeast cell cycle	Song et al. (2007)
	Protein expression profiles	Bensmail et al. (2005)
Demography	Age-specific mortality rates	Hyndman and Shang (2010)
	Mortality, fertility and migration rates	Hyndman and Booth (2008)
	Mortality and fertility rate	Hyndman and Ullah (2007)
Environment	Gas emissions	Torres et al.(2010)
	Diurnal ozone and NOx cycles for transportation emission control	Guo (2004)
	Stratospheric ozone levels	Meiring (2007)
Finance	Cash flow and transactions	Laukaitis (2008)
	Price formation and online auctions	Bapna et al.(2008)
	Cash flows in point of sale and ATM networks	Laukaitis (2005)
Linguistics	Speech production variability in fricatives of children and adults	Koenig et al. (2008)
	Tongue tip velocity	Lee et al. (2006)
	Speech movement records	Lucero (2005)

Biomedicine	Diffusion tensor imaging fiber images	Zhu et al. (2010)
	Gene expression microarray data	Wu and Müller (2010)
	Spinal cord dorsal horn neurons	Kim et al.2010
	3-Tesla magnetic resonance imaging data	Gouttard et al. (2009)
	Denaturing gradient gel electrophoresis data	Illian et al. (2009)
Medicine	Human growth	Hermanussen and Auxology (2010)
	Age–specific breast cancer mortality rates	Erbas et al. (2010)
	Age–specific fall injury incidence rates	Ullah and Finch (2010)
	Haemoglobin levels in renal anaemia	West et al. (2007)
	Women urinary hormone profiles at midlife	Meyer et al. (2007)
Neurology	Joint coordination data in motor development	Harrison et al. (2007)
Behavioural	Male medfly calling behaviour	Zhang et al. (2006)
Chemistry	Molecular weight distributions	Hutchinson et al. (2004)
Ecology	Plankton monitoring data	Ikeda et al. (2008)
Biomechanics	Kinematic gait data	Roislien et al. (2009)

1.2 Regression models

The first results in functional nonparametric statistics were developed by Ferraty and Vieu (2000). Authors established the almost complete convergence of a kernel estimator of the nonparametric model in the i.i.d case. Then, Ferraty and Vieu (2004) generalized these results to the α-mixing case and they exploited the importance of nonparametric modeling of functional data by applying their studies problems such as time series prediction and curves discrimination. In the context of functional observations α-mixing, Masry (2005) proved asymptotic normality of the estimator of Ferraty and Vieu (2004) for the regression function. The reader can find in the book of Ferraty and Vieu (2006), a wide range of applications of the regression function in functional statistics. The mean square convergence was investigated by Ferraty et al. (2007). Specifically, they explained the exact asymptotic term of the quadratic error. This result was used by Rachdi and Vieu (2007) to determine an criterion for automatic smoothing parameter selection based on cross validation. The local version of this criterion has been studied by Benhenni et al. (2007). That article presented a comparative study between the local and global approach. Different research works were interested in estimating the regression function using different approaches; the method of k nearest neighbors (Burba et al. (2008)); robust technical (Azzidine et al. (2008) and Crambes et al. (2008)), and the estimate via the simplified method of local polynomial (Barrientos-Marin et al. (2010)). For more literature, we refer the reader to Delsol (2007,2009), Delsol et al. (2011), Ferraty and Vieu (2011), Delsol (2011), Mechab and Laksaci (2016), and Akkal et al. (2018).

1.3 Conditional models

1.3.1 On conditional distribution

The estimation of the conditional distribution function in a functional framework was introduced by Ferraty et al. (2006). Authors constructed a double-kernel estimator for the conditional distribution function and they specified the almost complete convergence rate of this estimator when the observations are independent and identically distributed. The case of α-mixing observations was studied by Ferraty et al. (2005). Since then, several authors dealt with the estimation of the conditional distribution function (e.g. Ezzahrioui and Ould-Saïd (2005,2006), Ferraty et al. (2011), Mahiddine et al. (2014), Demongeot et. al (2014), Attaoui and Ling (2015), and Bouanani et al. (2019)).

1.3.2 On conditional density

The estimation of the conditional density function in functional statistics was at first presented by Ferraty et. al (2006). Authors obtained the almost complete convergence in the i.i.d case. Then, an abundant literature has been developed on the estimation of the conditional density, in particular in order to use it to estimate the conditional mode. Indeed, considering $\alpha-$mixing observations, Ferraty et al. (2005) established the almost complete convergence of a kernel estimator of the conditional mode defined by the random variable maximizing the conditional density. Alternatively. Ezzahrioui and Ould-Said (2005, 2006) estimated the conditional mode. The latter focused on the asymptotic normality of the proposed estimator in both cases (i.i.d. and $\alpha-$mixing). The precision of the dominant terms of the quadratic error of the kernel estimator of the conditional density was obtained by Laksaci (2007). After that, an extensive literature has been done on the subject (e.g. Laksaci et al. (2013), Demongeot et al. (2013), Izbichi and Lee (2016), Xianzhu et al. (2017), Daoudi et al. (2019), and Xiong and Meijuan (2020)).

1.3.3 On conditional hazard function

The literature on estimating the conditional hazard function is relatively restricted into functional statistics. The article by Ferraty et al. (2008) is precursor work on the subject, the authors introduced a nonparametric estimate of the conditional hazard function, when the covariate is functional. The α-mixing case was handled by Quintela-Del-Rio (2010). The latter established the asymptotic normality of the estimator proposed by Ferraty et al. (2008), the authors have illustrated these asymptotic results by an application on seismic data. Then, Laksaci and Mechab (2010) gave the estimation of conditional hazard function for functional data spatially dependent. After that, several research works have been given on the subject (e.g. Rabhi and Benaissa (2013), Laksaci and Mechab (2014), Benaissa and Mechab (2015), Rabhi et al. (2015), Massim and Mechab (2016), Hamel et al. (2017), Merouan and Mechab (2018), Tabti and Ait Saidi (2018), and Daoudi et al. (2020)).

1.4 On single index models

For several years, an increasing interest has been devoted to the study of semi-parametric models. This is mainly due to the problems associated with the poor specification of certain models. Tackle a problem of misspecification semiparametric way consists in not specifying the functional form of some model components. This approach completes those non-parametric models, which can not be useful in small samples, or with a large number of variables. In the classical regression case, the regression function of Y knowing the covariate X, is denoted by $r(x) = \mathbb{E}(Y \mid X = x),\ X, Y \in \mathbb{R}^d \times \mathbb{R}$. For this model, the non-parametric method considers only regularity assumptions on the function r. Obviously, this method has some drawbacks. One can cite the problem of curse of dimensionality.

This problem appears when the number of regressors d increases, the rate of convergence of the nonparametric estimator r which is supposed k times differentiable is $O(n^{-k/2k+d})$ deteriorates. The second drawback is the lack of means to quantify the effect of each explanatory variable. To alleviate in these drawbacks, an alternative approach is naturally provided by the semi-parametric model which supposes the introduction of a parameter on the regressors, by considering

$$\mathbb{E}_\theta(Y \mid X) = \mathbb{E}(Y \mid < X, \theta >= x).$$

The models are known in the literature as the single-index models. These models allow to obtain a compromise between parametric models, generally too restrictive and nonparametric models where the rate of convergence of the estimators deteriorate quickly in the presence of a large number of explanatory variables. In this area, different types of models have been studied in the literature: amongst the most famous, there may be mentioned additive models, partially linear models or single index models. The idea of these models, in the case of estimating the conditional density or regression consists in bringing to the covariate a dimension smaller than dimension of the variable space, thus allowing to overcome the problem of curse of dimensionality. For example, in the partially linear model, we decompose the quantity to be estimated, into a linear part and a functional part. This latter quantity does not pose estimation problem since it's expressed as a function of explanatory variables of finite dimension, thus avoiding the problems associated with curse of dimensionality. In order to treat the problem of curse of dimensionality in the case of chronological series, several semi-parametric approaches have been proposed. A general presentation of this type of model is given in Ichimura et al. (1993) where the convergence and asymptotic normality are obtained. In the case of M-estimators, Delecroix et al. (1999) proved the consistency and asymptotic normality of the estimator, and they study its effectiveness. The literature on these methods is abundant. Huber (1985) and Hall (1989) presented an estimation method which consists in projecting the density and the regression function on a space of dimension one, to bring a nonparametric estimation for dimensional covariate. Attaoui et al. (2011) established the pointwise and the uniform almost complete convergence (with rate) of the kernel estimate of this model. The interest of their study is to show how the estimate of the conditional density can be used to obtain an estimate of the simple functional index if the latter is unknown. More precisely, this parameter can be estimated by pseudo-maximum likelihood method which is based on the preliminary estimate of the conditional density. Later, Mahiddine et al. (2014) established the pointwise almost complete convergence and the uniform almost complete convergence (with the rate) of some characteristics of the conditional distribution and the successive derivatives of the conditional density when the observations are linked with a single-index structure and they are applied to the estimations of the conditional mode and conditional quantiles.

The single-index approach is widely applied in econometrics as a reasonable compromise between nonparametric and parametric models. Such kind of modelization is intensively studied in the multivariate case (e.g. Härdle et al. (1993) and Hristache et al. (2001)). Based on the regression function, Delecroix et al. (2003) studied the estimation of the single-index and established some asymptotic properties. The literature is strictly limited in the case where the explanatory variable is functional (that is a curve). The first asymptotic properties in the fixed functional single-model were obtained by Ferraty et al. (2003). They established the almost complete convergence, in the i.i.d. case, of the link regression function of this model. Their results were extended to dependent case by Aït-Saidi et al. (2005). Aït-Saidi Saidi et al. (2008) studied the case where the functional single-index is unknown. They proposed an estimator of this parameter, based on the cross-validation procedure. Attaoui (2014) and Attaoui and Ling (2016) studied, respectively, the estimation of the conditional density and the conditional cumulative distribution function based on a single functional index model under strong mixing condition. Bouchentouf et al.

(2014) investigated the semi-parametric estimation of the hazard function. Goia and Vieu (2015) introduced a semi-parametric methodology, which approximates the unknown regression operator through a single index approach, taking possible structure changes into account. For more recent review on functional single-index models, one refers the reader to Ling and Xu (2012), Shang (2018) and Sang and Cao (2020).

1.5 Brief presentation of the results established in this thesis

In this section, we give a brief presentation of different results obtained in this thesis.

1.5.1 Presentation of the considered model

Let X denote a random variable associated to a lifetime (ie, a random variable with values in $\mathbb{R}^+$).

When X has a density f with respect to the measure of Lebesgues, the hazard rate is written, for all x as follows:

$$h(x) = \frac{f(x)}{S(x)},$$

where, f is the density function, $S = 1 - F$ is survival function of X, and F denotes the distribution function of X such that $F(x) < 1$.

Let the conditional random rate for $x > 0$,

$$h^Z(x) = \frac{f^Z(x)}{S^Z(x)},$$

with $f^Z(\cdot)$ the conditional density, $S^Z = 1 - F^z$ the conditional survival function and, $F^Z(\cdot)$ the conditional distribution function of X knowing Z.

Let z be a fixed element of the functional space $\mathcal{H}$, $\mathcal{N}_z$ denotes a fixed neighborhood of z and $\mathcal{S}_{\mathbb{R}}$ is a fixed compact of $\mathbb{R}^+$.

Let $(X_i, Z_i)_{1\le i\le n}$ be random variables, each of them follows the same law of a couple (X, Z) where X is valued in $\mathbb{R}$ and Z has values in the Hilbert space $(\mathcal{H}, < \cdot;\cdot >)$. Here, we assume that X_i and Z_i are observed.

Let the following functional kernel estimators:

$$\hat{F}(\theta,x,z) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right) H\left(h_H^{-1}(x - X_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right)},$$

and

$$\hat{f}(\theta,x,z) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right) H'\left(h_H^{-1}(x - X_i)\right)}{h_H \sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right)},$$

with K a kernel, H a distribution function and, $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers.

A kernel estimator of the functional conditional hazard function $h(\theta,\cdot,Z)$ is therefore given as:

$$\hat{h}(\theta,x,Z) = \frac{\hat{f}(\theta,x,Z)}{1 - \hat{F}(\theta,x,Z)}.$$

1.5.2 The model under censorship

Let C be a positive variable, and variables (T_i, Δ_i, Z_i) the observed random, where $T_i = \min(X_i,\ C_i)$ and $\Delta_i = I_{X_i \leq C_i}$. $F_1(\theta,\cdot,Z)$ and $f_1(\theta,\cdot,Z)$ to describe the distribution function and conditional density of C knowing Z and we use the notation $S_1(\theta,\cdot,Z) = 1 - F_1(\theta,\cdot,Z)$.

Let $L(\theta,\cdot,Z) = 1 - S_1(\theta,\cdot,Z)S(\theta,\cdot,Z)$ and $\varphi(\theta,\cdot,Z) = f(\theta,\cdot,Z)S_1(\theta,\cdot,Z)$, we can reformulate the expression (4.1) as follow:

$$h(\theta,t,Z) = \frac{\varphi(\theta,t,Z)}{1 - L(\theta,t,Z)},\ \forall t,\ L(\theta,t,Z) < 1.$$

So, we can define $\varphi(\theta,\cdot,Z)$ and $L(\theta,\cdot,Z)$ by setting

$$\hat{L}(\theta,t,Z) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right) H\left(h_H^{-1}(t - T_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right)}$$

and

$$\widetilde{\varphi}(\theta,t,Z) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right) \Delta_i H'\left(h_H^{-1}(t - T_i)\right)}{h_H \sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right)}.$$

Finally, the hazard function estimator is given as:

$$\widetilde{h}(\theta,t,Z)=\frac{\widetilde{\varphi}(\theta,t,Z)}{1-\hat{L}(\theta,t,Z)}.$$

1.5.3 First result: Pointwise almost complete Convergence

The objective is to establish the pointwise almost complete convergence of the kernel estimator $\hat{h}(\theta,\cdot,Z)$ of the conditional hazard function $h(\theta,\cdot,Z)$ including censored and uncensored variables.

Case of complete data

Theorem 1.1. *Under some hypotheses, we have:*

$$\sup_{x\in\mathcal{S}_{\mathbb{R}}}|\hat{h}(\theta,x,z)-h(\theta,x,z)|=\mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_{\theta,z}(h_K)}}\right).$$

Case of censored data

Theorem 1.2. *Under some assumptions, we have:*

$$\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\widetilde{h}(\theta,t,z)-h(\theta,t,z)|=\mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_{\theta,z}(h_K)}}\right).$$

1.5.4 Second result: Uniform almost complete convergence

The following condition is necessary for our results. Consider

$$\mathcal{S}_{\mathcal{H}}\subset\bigcup_{k=1}^{d_n^{\mathcal{S}_{\mathcal{H}}}}B(z_k,r_n)\text{ and }\Theta_{\mathcal{H}}\subset\bigcup_{j=1}^{d_n^{\Theta_{\mathcal{H}}}}B(s_j,r_n)$$

with x_k (resp. t_j) $\in\mathcal{H}$ and $r_n,d_n^{\mathcal{S}_{\mathcal{H}}},d_n^{\Theta_{\mathcal{H}}}$ are sequences of positive real numbers which tend to infinity as n goes to infinity.

Case of complete data

Let

$$s_{n,0}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left|Cov\left(\psi_i(x,\theta),\psi_j(x,\theta)\right)\right|,$$
$$s_{n,1}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left|Cov\left(\psi_i(x,\theta)H_i(t),\psi_j(x,\theta)H_j(t)\right)\right|,$$
$$s_{n,3}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left|Cov\left(\Lambda_i,\Lambda_j\right)\right|,\ s_{n,4}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left|Cov\left(\Omega_i,\Omega_j\right)\right|,$$
$$s_{n,5}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left|Cov\left(\psi_i\left(x_{k(x)},\theta_{m(\theta)}\right),\psi_j\left(x_{k(x)},\theta_{m(\theta)}\right)\right)\right|,$$
$$s_{n,6}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left|Cov\left(\Gamma_i,\Gamma_j\right)\right|,\ s_{n,7}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left|Cov\left(\Gamma_i^{(l)},\Gamma_j^{(l)}\right)\right|,$$

where

$$\psi_i(x,\theta) = \frac{K(h_K^{-1}(<x-X_i,\theta>))}{\mathbb{E}K_1(\theta,x)},$$
$$\Lambda_i(x,\theta) = \frac{1}{h_K\phi(h_K)}\mathbf{1}_{B_\theta(x,h)\cup B_\theta(x_{k(x)},h)}(X_i),$$
$$\Omega_i(x,\theta) = \frac{1}{h_K\phi(h_K)}\mathbf{1}_{B_\theta(x_{k(x)},h)\cup B_{\theta_{m(\theta)}}(x_{k(x)},h)}(X_i),$$
$$\psi_i\left(x_{k(x)},\theta_{m(\theta)}\right) = \frac{K\left(h_K^{-1}<x_{k(x)}-X_i,\theta_{m(\theta)}>\right)}{\mathbb{E}K\left(h_K^{-1}<x_{k(x)}-X_i,\theta_{m(\theta)}>\right)},$$

$$\begin{aligned}\Gamma_i &= \frac{K\left(h_K^{-1}<x_{k(x)}-X_i,\theta_{m(\theta)}>\right)}{\mathbb{E}K\left(h_K^{-1}<x_{k(x)}-X_i,\theta_{m(\theta)}>\right)}H\left(h_H^{-1}(t_y-Y_i)\right)\\ &- \mathbb{E}\left(\frac{K\left(h_K^{-1}<x_{k(x)}-X_i,\theta_{m(\theta)}>\right)}{\mathbb{E}K\left(h_K^{-1}<x_{k(x)}-X_i,\theta_{m(\theta)}>\right)}H\left(h_H^{-1}(t_y-Y_i)\right)\right)\end{aligned}$$

and

$$\begin{aligned}\Gamma_i^{(l)} &= \frac{1}{h_H^l}\frac{K\left(h_K^{-1}<x_{k(x)}-X_i,\theta_{m(\theta)}>\right)}{\mathbb{E}K\left(h_K^{-1}<x_{k(x)}-X_i,\theta_{m(\theta)}>\right)}H^{(l)}\left(h_H^{-1}(t_y-Y_i)\right)\\ &- \frac{1}{h_H^l}\mathbb{E}\left(\frac{K\left(h_K^{-1}<x_{k(x)}-X_i,\theta_{m(\theta)}>\right)}{\mathbb{E}K\left(h_K^{-1}<x_{k(x)}-X_i,\theta_{m(\theta)}>\right)}H^{(l)}\left(h_H^{-1}(t_y-Y_i)\right)\right).\end{aligned}$$

Theorem 1.3. *Under some hypotheses, we have:*

$$\begin{aligned}\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}|\hat{h}(\theta,y,x)-h(\theta,y,x)| &= \mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}+\log d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right)\\ &+ \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n'^{*2}\log d_n^{\mathcal{S}_{\mathcal{F}}}d_n^{\Theta_{\mathcal{F}}}}}{n}\right).\end{aligned}$$

where

$s_n^{'*} = \max\{s_{n,0}; s_{n,1}; s_{n,2}; s_{n,3}; s_{n,4}; s_{n,5}; s_{n,7}\}$.

When functional single-index is fixed, we get

Corollary 1.4. *Under some assumptions, as n goes to infinity, we have*

$$\begin{aligned}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}|\hat{h}(\theta,y,x)-h(\theta,y,x)| &= \mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right)\\ &+ \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n^{'*2}\log d_n^{\mathcal{S}_{\mathcal{F}}}}}{n}\right).\end{aligned}$$

Case of censored data

Theorem 1.5. *Under some hypotheses, we get:*

$$\begin{aligned}\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\widetilde{h}(\theta,t,z)-h(\theta,t,z)| &= \mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}+\log d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right)\\ &+ \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n^{'*2}\log d_n^{\mathcal{S}_{\mathcal{F}}}d_n^{\Theta_{\mathcal{F}}}}}{n}\right).\end{aligned}$$

In the case of fixed functional single-index, we have:

Corollary 1.6. *Under some assumptions, as n goes to infinity, we have*

$$\begin{aligned}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\widetilde{h}(\theta,t,z)-h(\theta,t,z)| &= \mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right)\\ &+ \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n^{'*2}\log d_n^{\mathcal{S}_{\mathcal{F}}}}}{n}\right).\end{aligned}$$

1.6 Outline of the thesis

This thesis consists of three chapters including the introductory chapter.

In the first chapter, we give the basics notion of nonparametric statistics for functional data. Then, we give an overview on the conditional models. The single index models have been also introduced. We finish this chapter by including a synthesis of the results obtained in this thesis. In second chapter, we present the basic concepts in survival analysis. In chapter three, we give some asymptotic notations and definitions, where we provide importance tools useful for our results; almost complete convergence, properties of different kernels, and some results of strongly mixing conditions. In chapter four, we present some asymptotic properties related to the nonparametric estimation of the conditional hazard function with functional data. Firstly, we introduce the estimator of the of the conditional hazard function from the estimates of the conditional distribution and the conditional density, in two cases, namely, with complete and censored and censored data. Then, we study the both almost complete and uniform convergence of our estimator.

We finish the thesis by a conclusion, we summarize our results given in this thesis. We also give some points prospects.

Chapter 2

Basic concepts in survival analysis

Analysis of survival data dates back to 1693 with the famous English astronomer Edmond Halley who studied the birth and death records of Breslau city, which had been transmitted to the Royal Society by Caspar Neumann. He produced a life table showing the number of people surviving to any age from a cohort born the same year. He also employed his table to compute the price of life annuities. These analyzes were refined in the 19th century, with the appearance of the exogenous variables. In the 20th century, the analysis of survival data began to go beyond the strict framework of demography, especially after the second world war, where the analysis of survival data has been very significant for industrial applications using parametric models with exponential or Weibull laws. It is only recently, motivated by medical applications (pharmaceutical, biomedical), that the nonparametric methods (Kaplan-Meier (1958)) appeared for the estimation of the survival function.

2.1 Survival data Analysis

Survival data analysis is the study of the occurrence, over time, of a specific event for one or more groups of given individuals. This event is generally a change of a state, it is often referred to a death, which can be the death of an individual as well as the illness onset, response to treatment, or machine failure. Each observation is defined by:

The date of origin: each individual has an original date. In order to make comparison of survival times between individuals, a precise definition of the interest event is necessary. If it is a death caused by an illness, it must be ensured that each death is due to the illness being studied, and not to other causes.

Survival time: it is defined as the time between the starting date and the occurrence of the event of interest.

Individuals or groups of individuals are likely to differ in one or more factors. These factors, called explanatory variables, can explain a significant difference in the survival time of the subjects studied. Their effects are analyzed by regression models. These may be individual factors (sex, age, biological parameters relating to a disease, genetic parameters, etc.), or linked to a therapeutic test (belonging to the treatment group or to the placebo group, drug dosage, etc.). The analysis of survival data is attached to the description of survival times and to see how much they depend on these explanatory variables. Classical approaches of the survival data analysis are of stochastic type, the time of occurrence of an event is assumed to be the realization of a random process associated with a particular distribution.

Excellent research works devoted to the analysis of survival data can be found in Kalbeisch and Prentice (1980), Cox and Oakes (1984), and Klein and Moeschberger (1997).

2.2 Incomplete data

One of the characteristics of survival data is the existence of incomplete observations. In fact, data are often partially collected, in particular because of the censorship and truncation processes. Censored or truncated data results from incomplete access to all the information. Instead of observing i.i.d realizations of Y duration, we observe the realization of the variable Y subjected to various disturbances, whether or not independent of the studied event. The mechanisms of censorship and truncation can occur simultaneously.

2.2.1 Truncation

The censored data are not the unique type of incomplete data. The other classical case is the one of the so-called truncated data, that models the lifetime by a variable Y which must be big enough to be observed. Contrarily to the censored data, variables are not still observed being given that if $Y < T$, nor Y nor the truncation variable T can not be observed. It is a model that first appeared in astronomy where is composed of astral objects. The truncated data are frequently used on the lifetime study. At the end of 1980, some statistical studies were undertaken on the time of incubation of the virus of the $AIDS$, that is the time during which a person is seropositive without to develop the illness as much.

Definition 2.2.1. *Truncation is a variant of censoring but different which occurs when the incomplete nature of the observation is due to a systematic selection process inherent to the study design.*

Randomly truncated data frequently arise in medical studies, other application areas include economics, insurance and astronomy.... In a broad sense, random truncation corresponds to biased sampling, where only partial or incomplete data are available about the variable of interest. One has two type of truncation, as follows:

i) ***Right truncation***: only individuals with event time less than some threshold are included in the study. As example, if you ask a group of smoking school pupils at what age they started smoking, you necessarily have truncated data, as individuals who start smoking after leaving school are not included in the study.

ii) ***Left truncation***: due to structure of the study design, we can only observe those individuals whose event time is greater than some truncation threshold. As example, imagine you Wish to study how long people who have been hospitalized for a heart attack survive taking some treatment at home. The start time is taken to be the time of the heart attack. Only those individuals who survive their stay in hospital are able to be included in the study.

ii) ***Interval truncation***: this is due when Y is truncated on the right and left. This type of truncation is encountered when studying patients in a registry: patients diagnosed before the registry is set up or listed after consulting the registry will not be included in the study.

2.2.2 Censoring

Definition 2.2.2. *Censoring is when an observation is incomplete due to some random case. The cause of the censoring must be independent of the event of interest if we are to use standard methods of analysis.*

In what follow, we distinguish the different types of censorship:

(i) ***Type I censorship***: instead of looking at the variables Y_1, Y_2,...,Y_n which we are interested in, we observe Y_i when it is less than a fixed duration C, otherwise we only know that Y_i is greater than C. We therefore observe a variable T_i such that $T_i = \min(Y_i, C)$.

(ii) ***Type II censorship***: we observe the life time of n patients until m of them are died and we stop at that moment. If we order the Y_1, Y_2,...,Y_n, we obtain the statistics of order $Y_{(1)}$, $Y_{(2)}$, ...,$Y_{(n)}$. The date of censorship is then $Y_{(r)}$ and we observe $T_{(1)} = Y_{(1)}, T_{(2)} = Y_{(2)}, ..., T_{(m)} = Y_{(m)}, T_{(m+1)} = Y_{(m)}, ... T_{(n)} = Y_{(m)}$.

(iii) ***Type III censorship***: there are two cases:

- ***Right-censoring:*** we observe the pair (T, δ) where T is the observed duration and δ is a variable representing the nature of this duration which takes the value 1 if it this is a true life time and 0 if it is censorship.
- ***Left-censoring:*** instead of observing Y_1, Y_2,...,Y_n we observe (T_i, δ_i) where $T_i = \max(Y, C_i)$ and $\delta_i = 1_{\{Y_i \geq C\}}$ for $i = 1, ..., n$ and C_i is a random censorship.

(IV) ***Double censorship:*** there are double censorship in a data sample if there is both left and right censorship in that sample (data are censored both right and left).

(IV) ***Interval censoring:*** a date is interval censored if, instead of observing with certainty the time of the event, the only information available is that it took place between two known dates.

Remark 2.2.1. *The commonest form of censoring is right censoring. Subjects followed until some time, at which the event has yet to occur, but then talks no further part in the study. This may be because:*

- *the subject dies from another cause, independently of the cause of interest,*

- *the study ends while the subject survives, or*
- *the subject is lost to the study, by dropping out, moving to a different area, etc.*

Remark 2.2.2. *Left censoring is much rare. Examples of left censoring include: infection with a sexually transmitted disease such as HIV/AIDS and time at which teenagers begin to drink alcohol.*

Remark 2.2.3. *Examples of interval censoring include: infection with HIV/AIDS with regular testing and failure of a machine during the Chinese new Year.*

2.2.3 Kaplan-Meier estimator

A very popular estimator of the distribution function in the setting of right censored data is the Kaplan-Meier estimator (also known as product-limit estimator) introduced by Kaplan and Meier (1958).

Let $X_1,...,X_n$ be a sample representing the durations of interest (these variables are therefore assumed to be positive), with distribution function F, and let $C_1,...,C_n$ denote a sample representing the censoring times, which we suppose to be independent of the durations of interest, of distribution function G. In the right random censorship model, we do not observe the duration of interest X_i but rather the smaller of the two values $Z_i = \min(X_i; C_i)$, as well as the censorship indicator δ_i which is equal to 1 if the duration of interest is observed, and 0 if it is censored, ie $\delta_i = 1_{X_i \leq C_i}$. In this kind of data, the distribution function F is estimated by the estimator introduced by Kaplan and Meier (1958):

$$\hat{F}^{KM}(t) := 1 - \prod_{[0,t]} \left\{ 1 - \frac{\sum_j 1_{\{Z_j = s, \Delta_j = 0\}}}{\sum_j 1_{\{Z_j \geq s\}}} \right\}.$$

The Kaplan-Meier estimator can also be represented as

$$\hat{F}^{KM}(t) = 1 - \Pi_{[0,t]}\{1 - \Lambda^{-}(ds)\},$$

where $\hat{\Lambda}^{-}(ds)$ is the Nelson-Aalen estimator of the predictable hazard measure.

$$\Lambda_n^{-}(ds) := \frac{N(ds)}{\sum_j 1_{\{Z_j \geq s\}}},$$

and $N(t) := \sum_j 1_{\{Z_j \leq t, \Delta_j = 1\}}$.

Chapter 3

Some probability tools and definitions

In this chapter, based on what has been done in the book of Ferraty (2005), we present different definitions and properties that are useful for our research work.

Let $X_1, X_2, ..., X_n$ denote n functional random variables valued in E and let χ denote a fixed element of E. A functional extension of multivariate kernel local weighting ideas would be to transform the n functional random variables $X_1, X_2, ..., X_n$ into the n quantities

$$\frac{1}{V(h)} = K\left(\frac{d(\chi, \chi_i)}{h}\right),$$

where d is a semi-metric on E, K is an asymmetrical real kernel. In this expression $V(h)$ would be the volume of

$$B(\chi, h) = \{\chi^{'} \in E, d(\chi, \chi') \leq h\}$$

which is the ball, with respect to the topology induced by d, centered at χ and of radius h. Nevertheless, this approach requests to define $V(h)$. That is, this needs to have at hand a measure on E. This is the main difference with real and multivariate cases for which the Lebesgue measure is implicitly used whereas in the functional space E we do not have such a universally accepted reference measure. Thus, we build the normalization by utilizing directly the probability distribution of the functional random variables. The functional kernel local weighted variables are defined by:

$$\Delta_i = \frac{K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)}{\mathbb{E}\left(K(\frac{d(\chi,\mathcal{X}_i)}{h})\right)}.$$

Note that for the multivariate case, we have for some constant C depending on K and on the norm $\|.\|$ used in $\mathbb{R}^p$,

$$EK(\|x - X_i\|/h) \sim cf(x)h^p$$

as long as X_i has a density f with respect to Lebesgue measure which is continuous and such that $f(x) > 0$, this result is known in the literature as the Bochner's type Theorem.

3.1 Kernel types and proprieties

We present two kinds of kernels for weighting functional variables and consider their man properties from Ferraty (2005).

Definition 3.1.1. ***(Ferraty(2005)).***

i) *A function K from $\mathbb{R}$ into $\mathbb{R}^+$ such that $\int K = 1$ is called a kernel of type I if there exit two real constants $0 < C_1 < C_2 < \infty$ such that :*

$$C_1 1_{[0,1]} \leq K \leq c_2 1_{[0,1]}.$$

ii) *A function K from $\mathbb{R}$ into $\mathbb{R}^+$ such that $\int K = 1$ is called a kernel of type II if its support is $[0,1]$ and if its derivative K' exists on $[0,1]$ and satisfies for two real constants $-\infty < C_1 < C_2 < 0$:*

$$C_2 \leq K' \leq C_1$$

.

Definition 3.1.2. ***(Ferraty(2005)).*** *A function K from $\mathbb{R}$ into $\mathbb{R}^+$ such that $\int K = 1$ with compact support $[-1,1]$ and such that $\forall u \in [0,1]$, $K(0) > 0$ is called a kernel of type 0.*

Next, we can put:

$$\mathbb{E}\left(1_{[0,1]}\left(\frac{d(\chi,X)}{h}\right)\right) = \mathbb{E}(1_{B(\chi,h)(X)}) = \mathbb{P}(X \in B(\chi,h)).$$

It is quite clear that the probability of the ball $B(\chi, h)$ appears in the normalization. The smoothing parameter h (also said the bandwidth) decreases with the size of the sample of the functional variables (more precisely, h tends to zero when n tends to ∞). Therefore, when we take n very large, h is close to zero and then $B(\chi, h)$ is considered as a small ball and $P(\chi \in B(\chi, h))$ as a small ball probability. Thus, for all χ in E and for all positive real h, we will use the notation:

$$\varphi_\chi(h) = \mathbf{P}(\chi \in B(\chi, h)).$$

It should be emphasized that small ball probabilities play an important role both from a theoretical and practical point of view. The notion of ball being strongly linked with the semi-metric d, the choice of this semi-metric is very important since the convergence rates of our nonparametric functional estimates are systematically linked with d through the behaviour, around 0, of the small ball probability function φ_χ.

Next, we present two results, according to the fact that the kernel is of type I or II. Let X denote a functional random variable taking its values in the semi-metric space (E, d), let χ denote a fixed element of E, let h be a real positive number and let K be a kernel function.

Lemma 3.1.1. ***(Ferraty(2005)).*** *If K is a kernel of type I, then there exist non negative finite real constants C and C' such that:*

$$C_{\varphi_\chi(h)} \leq \mathbb{E}K(\frac{d(\chi, X)}{h}) \leq C'_{\varphi_\chi(h)}.$$

Lemma 3.1.2. ***(Ferraty(2005)).*** *If K is a kernel of type II and if $\varphi_\chi(.)$ satisfies*

$$\exists C_3 > 0, \exists \epsilon_0, \int_0^\epsilon \varphi_\chi(u) d(u) > C_3 \varphi_\chi(\epsilon),$$

then there exist non negative finite real constants C and C' such that ,for hsmall enough:

$$C_{\varphi_\chi(h)} \leq \mathbb{E}K(\frac{d(\chi, X)}{h}) \leq C'_{\varphi_\chi(h)}.$$

3.2 Kernel estimators

3.2.1 Estimating the conditional distribution function (c.d.f.)

Let $\hat{F}_Y^{\mathcal{X}}$ be the estimator of the conditional c.d.f. $F_Y^{\mathcal{X}}$. The $F_Y^{\mathcal{X}}(\chi, y) = \mathbb{P}(Y \leq y | \mathcal{X} = \chi)$ can be expressed in terms of conditional expectation:

$$F_Y^{\mathcal{X}}(\chi, y) = \mathbb{E}(1_{(-\infty, y]}(Y) | \mathcal{X} = \chi),$$

and by analogy with the functional regression context, a naive kernel conditional c.d.f. estimator could be defined as follows:

$$\tilde{F}_Y^{\mathcal{X}}(\chi, y) = \frac{\sum_{i=1}^{n} K(h^{-1} d(\chi, \mathcal{X}_i)) \mathbb{1}_{(-\infty, y]}(Y_i)}{\sum_{i=1}^{n} K(h^{-1} d(\chi, \mathcal{X}_i))}.$$

By following the ideas previously developed by Roussas (1969)and Samanta (1989)in the finite dimensional case, it is easy to construct a smooth version of this naive estimator. To do so, it suffices to change the basic indicator function into a smooth c.d.f. Let K_0 be an usual symmetrical kernel , let H be defined as:

$$\forall u \in \mathbb{R} \qquad H(u) = \int_{-\infty}^{u} K_0(v) dv,$$

and define the kernel conditional c.d.f. estimator as follows:

$$\hat{F}_Y^{\mathcal{X}}(\chi, y) = \frac{\sum_{i=1}^{n} K(h^{-1} d(\chi, \mathcal{X}_i)) H(g^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(h^{-1} d(\chi, \mathcal{X}_i))}.$$

It is clear that the parameter g acts as the bandwidth h. The smoothness of the function $\hat{F}_Y^{\mathcal{X}}(\chi, .)$ is controled both by the smoothing parameter g and by the regularity of the c.d.f. H. The idea to build such a smooth c.d.f. estimate was introduced by Azzalini (1981) and Reiss (1981). The roles of the other parameters involved in this functional kernel c.d.f. estimate [i.e., the roles of K and h] are the same as in the regression setting.

3.2.2 Estimating the conditional density

It is known that,under some differentiability assumption, the conditional density function can be obtained by derivating the conditional c.d.f. . Since we have the estimator $\hat{F}_Y^{\mathcal{X}}$ of $F_Y^{\mathcal{X}}$, we suggest the following estimate:

$$\hat{f}_Y^{\mathcal{X}}(\chi, y) = \frac{\partial}{\partial y}\hat{F}_Y^{\mathcal{X}}.$$

Assuming the differentiability of H, we have

$$\frac{\partial}{\partial y}\hat{F}_Y^{\mathcal{X}} = \frac{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i))\frac{\partial}{\partial y}H(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i))},$$

and this is motivating the following expression for the kernel functional conditional density estimate:

$$\hat{f}_Y^{\mathcal{X}}(\chi, y) = \frac{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i))\frac{1}{g}H'(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i))}.$$

More generally, we can state for any kernel K_0 the following definition:

$$\hat{f}_Y^{\mathcal{X}}(\chi, y) = \frac{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i))\frac{1}{g}HK_0(g^{-1}(y - Y_i))}{\sum_{i=1}^n K(h^{-1}d(\chi, \mathcal{X}_i))}.$$

Note that we can easily get the following kernel functional conditional mode estimator of $\theta(\chi)$:

$$\hat{\theta}(\chi) = \arg\sup_{y \in S} \hat{f}_Y^{\mathcal{X}}(\chi, y).$$

3.2.3 Estimating the regression

We propose for the nonlinear operator r the following functional kernel regression estimator:

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

where K is an asymmetrical kernel and h (depending on n) is a strictly positive real. It is a functional extension of the familiar Nadaraya- Watson estimate (see Nadaraya (1964) and Watson (1964) which was previously introduced for finite dimensional nonparametric regression (see Härdle (1990) for extensive discussion). The main change comes from the semi-metric d which measures the proximity between functional objects. To see how such an estimator works, let us consider the following quantities:

$$w_{i,h} = \frac{K(h^{-1}d(x,X_i))}{\sum_{i=1}^{n} K(h^{-1}d(x,X_i))}.$$

Thus, it is easy to rewrite estimator $\hat{r}$ as follows:

$$\hat{r} = \sum_{i=1}^{n} w_{i,h}(x)Y_i,$$

which is really a weighted average because:

$$\sum_{i=1}^{n} w_{i,h}(x) = 1.$$

The behavior of the $w_{i,h}(x)$ can be deduced from the shape of the asymmetrical kernel function K.

3.2.4 Estimation of the hazard function

The estimation of the hazard function is of great interest in statistics. Indeed, it is used in risk analysis or for the study of survival phenomena. The hazard rate $h(t)$ is defined by:

$$h(t) = \lim_{\Delta t \longrightarrow 0^+} \frac{P(t \leq \tau < t + \Delta t / \tau \geq t)}{\Delta t}, \; t > 0.$$

It is not difficult, to see that the hazard rate can be rewritten as the rate of the density $f(.)$ of which it is absolutely continuous with respect to the Lebesgue measure and the survival function $S(.) = 1 - F(.)$ of t. In other words:

$$h(t) = \frac{f(t)}{S(t)},$$

where the survival function $S(t)$ is none other than the complement of the distribution function. In fact it is the derivative of a probability that the duration is between t and Δ, t knowing that the period t is reached. More practically, this is an instantaneous rate of exit from the state at the date t. The survival curve takes a particular meaning given by:

$$S(t) = exp\left(-\int_0^t h(u)d(u)\right).$$

There is an extended literature, on the estimator of the nonparametric hazard rate, in an approximate way and for the nonparametric case, two methods have been proposed to estimate the hazard rate. The first approach replaces $f(t)$ and $S(t)$ in the expression of $h(t)$ by their estimators$f(t)$ and $S(t)$ respectively, which gives us the estimator of the hazard rate by:

$$\hat{h}(t) = \hat{f}(t)\hat{S}(t).$$

Nielsen and Linton (1995) named this type of estimator by (external estimator). The estimator with external kernel of the hazard rate of non censured data has been introduced by Watson and Leadbetter (1964).

The second method is based on the relation between cumulative hazard and the hazard rate, where the cumulative hazard is defined by:

$$\lambda(t) = -\int_0^t h(u)d(u).$$

Nielsen and Linton (1995) named this type of estimators by (internal estimator). The relation between cumulative hazard and the hazard rate suggests that h(t) can be obtained by smoothing $H(t)$ using a kernel, in other words:

$$h(t) = \int_0^t K_h(t-u)d\hat{\lambda}(u),$$

where h is a window width such that $h \longrightarrow 0$ when $n \longrightarrow 1$.

The internal hazard rate estimator for censored data has also been introduced by Watson and Leadbetter (1964), and Tanner and Wong (1983, 1984). Then, Tanner and Wang (1984), and Sarda and Vieu (1996) used selection in window width for this type of hazard rate estimators. In a more interesting work, Rice and Rosenblatt compared the asymptotic properties of the two classes of the kernel estimator of the hazard rate, they showed that the two have the same asymptotic variance, but their asymptotic biases are different. Until now, to take interest to the hazard rate will generally depend on certain covariances, for example, the survival time of a patient will be affected by many characteristics such as age and gender. The conditional hazard rate of t knowing $Z = z$ is defined by:

$$h(t/z) = \lim_{\Delta \longrightarrow} \frac{P(t \leq \tau < t + \Delta t / \tau \geq t, Z = z)}{\Delta t}, \ \ t > 0.$$

As well as the conditional hazard function τ knowing $Z = z$ is defined by:

$$h(t/z) = \frac{f(t/z)}{1 - F(t/z)},$$

such that $F(t/z)$ (resp: $f(t/z)$ is the conditional distribution (resp. the conditional density) of T knowing $Z = z$, which is supposed absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.

3.3 On mixing conditions

For many phenomena of the real world, observations in the past and present may have considerable influence on observations in the near future, but rather weak influence on observations in the far future. Random sequences that satisfy strong mixing conditions are used to model such phenomena.
In the reality, the treated data present a certain form of dependence or mixing, and there exist several form of mixing according to coefficients, noted: $\alpha, \beta, \rho, \psi$ and ϕ among those, the alpha-mixing is weakest and is therefore least restraining. Thus, all results statement for alpha mixing data will be valid for the submissive data to another type of mixing.

3.3.1 Strong mixing conditions

Suppose $X : (X_k; k \in \mathbb{Z})$ is a (not necessarily stationary) sequence of random variables. For $-\infty \le I \le J \le \infty$, define the $\sigma-$field:

$$\mathcal{F}_I^J := \sigma(X_k; I \le k \le J, (k \in \mathbb{Z})).$$

The notation $\sigma(\cdots)$ means the $\sigma-$ field $\subset \mathcal{F}$ generated by $(\cdots)$. For each $n \ge 1$, let us define the following dependence coefficients:

$$\alpha(n) := \sup_{i\in\mathbb{Z}} \alpha(\mathcal{F}_{-\infty}^j, \mathcal{F}_{j+n}^\infty), \qquad \phi(n) := \sup_{i\in\mathbb{Z}} \phi(\mathcal{F}_{-\infty}^j, \mathcal{F}_{j+n}^\infty)$$
$$\psi(n) := \sup_{i\in\mathbb{Z}} \psi(\mathcal{F}_{-\infty}^j, \mathcal{F}_{j+n}^\infty), \quad \rho(n) := \sup_{i\in\mathbb{Z}} \rho(\mathcal{F}_{-\infty}^j, \mathcal{F}_{j+n}^\infty), \text{ and}$$
$$\beta(n) := \sup_{i\in\mathbb{Z}} \beta(\mathcal{F}_{-\infty}^j, \mathcal{F}_{j+n}^\infty).$$

The random sequence X is said to be:

- $\alpha-$mixing (or strong mixing) if $\alpha(n) \to 0$ as $n \to \infty$,

- $\phi-$mixing if $\phi(n) \to 0$ as $n \to \infty$,
- $\psi-$mixing if $\psi(n) \to 0$ as $n \to \infty$,
- $\rho-$mixing if $\rho(n) \to 0$ as $n \to \infty$,
- $\beta-$mixing (or absolutely regular) if $\beta(n) \to 0$ as $n \to \infty$,

3.3.2 $\alpha-$Mixing conditions

We recall some definitions and fix some notations of the $\alpha-$mixing (or strong mixing) notion, which is one of the most general among the different mixing structures introduced in the literature. Let $(\xi_n)_{n\in\mathbb{Z}}$ be a sequence of random variables defined on some probabilistic space $(\Omega, \mathcal{A}, \mathrm{P})$ and taking values in some space $(\Omega^{'}, \mathcal{A}^{'})$. Let us denote, for $-\infty \le j \le k \le +\infty$, by $\mathcal{A}_j^k$ the $\sigma-$algebra generated by the random variables $(\xi_s, j \le s \le k)$. The strong mixing coefficients are defined to be the following quantities:

$$\alpha(n) = \sup_k \sup_{A\in\mathcal{A}_k^{-\infty}} \sup_{A\in\mathcal{A}_{n+k}^{+\infty}} |\,\mathrm{P}(A\cap B) - \mathrm{P}(A)\mathrm{P}(B)\,|\,.$$

Definition 3.3.1. ***(Ferraty(2005)).*** *The sequence $(\xi_n)_{n\in\mathbb{Z}}$ is said to be $\alpha-$mixing (or strongly mixing), if* $\lim_{n\to+\infty} \alpha(n) = 0$.

In order to simplify the presentation of the results and not to mask our main purpose, we will mainly consider both of the following subclasses of mixing sequences:

Definition 3.3.2. ***(Ferraty(2005)).*** *The sequence $(\xi_n)_{n\in\mathbb{Z}}$ is said to be arithmetically (or equivalently algebraically) $\alpha-$mixing with rate $a > 0$ if $\exists C > 0,\ \alpha(n) \le Cn^{-a}$. It is called geometrically $\alpha-$mixing if $\exists C > 0,\ \exists t \in (0,1),\ \alpha(n) \le Ct^n$.*

Next, we present some general results on mixing sequences of functional random variables by using probabilistic results for mixing sequences of real random variables.

Proposition 3.3.1. ***(Ferraty(2005)).***

Assume that $\Omega^{'}$ is a semi-normed space with semi-norm $\|.\|$ and that $\mathcal{A}^{'}$ is the $\sigma-$algebra spanned by the open balls for this semi-norm then we have:

- *$(\xi_n)_{n\in\mathbb{Z}}$ is $\alpha-$mixing $\Rightarrow (\|\xi_n\|_{n\in\mathbb{Z}})$ is $\alpha-$mixing.*

- *In addition, if the coefficients of $(\xi_n)_{n\in\mathbb{Z}}$ are geometric (resp. arithmetic) then those of $(\|\xi_n\|_{n\in\mathbb{Z}})$ are also geometric (resp. arithmetic with the same order).*

3.4 Almost complete convergence

Definition 3.4.1. ***(Ferraty(2006)).*** *One says that $(X_n)_{n\in\mathbb{N}}$ converges almost completely to some r.r.v. X, if and only if*

$$\forall\epsilon > 0, \sum_{n\in\mathbb{N}}(P \mid X_n - X \mid > \varepsilon) < \infty,$$

and the almost complete convergence of $(X_n)_{n\in\mathbb{N}}$ to X is denoted by $\lim_{n\longrightarrow\infty} X_n = X$, a.co.

Proposition 3.4.1. *If $\lim\limits_{n\longrightarrow+\infty} X_n = X$ a.co., then w have:*

i) $\lim\limits_{n\longrightarrow+\infty} X_n = X$, *p.*

ii) $\lim\limits_{n\longrightarrow+\infty} X_n = X$, *a.s.*

Definition 3.4.2. *One says that the rate of almost complete convergence of $(X_n)_{n\in\mathbb{N}}$ to X is of order one if and only if*

$$\forall\epsilon > 0, \sum_{n\in\mathbb{N}}(P \mid X_n - X \mid > \varepsilon u_n) < \infty,$$

and we write

$$X_n - X = O_{a.co.}(u_n).$$

Proposition 3.4.2. *Assume that $X_n - X = O_{a.co.}(u_n)$. We have:*

i) $X_n - X = O_p(u_n)$.

ii) $X_n - X = O_{a.s}(u_n)$.

The proof of the above propositions can also be found in Ferraty and Vieu (2006).

Proposition 3.4.3. *Assume that $\lim\limits_{n\longrightarrow+\infty} u_n = 0$, $\lim\limits_{n\longrightarrow+\infty} X_n = \iota_X$ a.co., and $\lim\limits_{n\longrightarrow+\infty} Y_n = \iota_Y$ a.co., where ι_X and ι_Y are two deterministic real numbers.*

i) *We have:*

a) $\lim\limits_{n\longrightarrow+\infty} X_n + Y_n = \iota_X + \iota_Y$, *a.co.*

b) $\lim_{n\to+\infty} X_n Y_n = \iota_X \iota_Y$, *a.co.*

c) $\lim_{n\to+\infty} \frac{1}{Y_n} = \frac{1}{\iota_Y}$, *a.co. as along as* $\iota_Y \neq 0$.

ii) *If* $X_n - \iota_X = O_{a.co}(u_n)$ *and* $Y_n - \iota_Y = O_{a.co}(u_n)$, *we have:*

a) $(X_n + Y_n) - (\iota_X + \iota_Y) = O_{a.co.}(u_n)$;

b) $X_n Y_n - \iota_X \iota_Y = O_{a.co.}(u_n)$,

c) $\frac{1}{Y_n} - \frac{1}{\iota_Y} = O_{a.co.}(u_n)$ *as along as* $\iota_Y \neq 0$.

3.5 Some useful inequalities

3.5.1 Hölder inequality

Theorem 3.1. ***(Baillo and Grane(2009)).*** *Let X and Y be two random variables such that* $X \in L^p(\Omega, A, \mathbb{P})$ *and* $Y \in L^q(\Omega, A, \mathbb{P})$ *with* $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$ *and* $p \geq 1, q \geq 1$. *Then*

$$\left(\mathbb{E} \mid XY \mid^{\frac{1}{2}}\right) \leq \mathbb{E}(\mid X \mid^p)^{\frac{1}{2}} \mathbb{E}(\mid X \mid^q)^{\frac{1}{2}}.$$

3.5.2 Bernstein's inequality

Let $Z_1, ..., Z_b$ be a independent real random variables with zero mean. It is worth being pointed out that the statement of almost complete convergence properties needs to find an upper bound for some probabilities involving sum of real random variables such as

$$\mathbb{P}\left(\Big|\sum_i^n Z_i\Big| > \varepsilon\right),$$

where eventually, the positive real decreases with n. In this context, there exist powerful probabilistic tools, generically called exponential inequalities. We focus here on the so-called Bernstein inequality.

Proposition 3.5.1. ***(Ferraty(2005)).*** *Assume that* $\forall m \geq 2$, $\mathbb{E}|Z_i^m| \leq (m/2)(a_i)^2 b^{m-2}$, *and let* $(A_n^2 = (a_1^2 + ... + a_1^2)$. *Then we have*

$$\forall \varepsilon \geq 0,\ \mathbb{P}\left(\Big|\sum_i^n Z_i\Big| > \varepsilon A_n\right) \leq exp\left\{-\frac{\varepsilon^2}{2(2+\frac{\varepsilon b}{\varepsilon A_n})}\right\}.$$

Note that this inequality is stated for non-identically distributed real random variables. Note also that each variable Z_i may depend on n. Indeed, for our statistical purpose, the next Corollary is used more often than the previous general proposition.

Corollary 3.2. *(Ferraty(2005)).*

- *If* $\forall m > 2,\ \exists C_m > 0,\ \mathbb{E} \mid Z_i^m \mid \leq C_m a^{2(m-1)}$, *we have*

$$\forall \varepsilon \geq 0,\ \mathbb{P}\left(\left| \sum_i^n Z_i \right| > \varepsilon n \right) \leq exp\left\{ - \frac{\varepsilon^2 n}{2a^2(1+\varepsilon)} \right\}.$$

- *Assume that the variables depend on n* ($Z_i = Z_{i,n}$).
  If $\forall m > 2,\ \exists C_m > 0,\ \mathbb{E} \mid Z_i^m \mid \leq C_m a^{2(m-1)}$, *and if* $u_n = n^{-1} a_n^2 \log n$ *verifies* $\lim_{n\to\infty} u_n = 0$, *we have:*

$$\frac{1}{n} \sum_i^n Z_i = o_{a.co.}(\sqrt{u_n}).$$

Note that all previous inequalities are given for unbounded random variables, which is useful for functional nonparametric regression. They apply directly for bounded variables, such as those appearing along functional conditional density or c.d.f. studies.

Corollary 3.3. *(Ferraty(2005)).*

- *if* $\exists M < \infty,\ \mid Z_1 \mid \leq M$, *and denoting* $\sigma^2 = \mathbb{E} Z_i^2$, *we have*

$$\forall \varepsilon \geq \mathbb{P}\left(\mid \sum_i^n Z_i \mid > \varepsilon n \right) \leq exp\left\{ - \frac{\varepsilon^2 n}{2\sigma^2(1 + \varepsilon \frac{M}{\sigma^2})} \right\}.$$

- *Assume that the variables depend on n (that is, assume that* $Z_i = Z_{i,n}$*).and are such that* $\exists M = M_n < \infty, \mid Z_1 \mid \leq M$ *and define* $\sigma^2 = \mathbb{E} Z_i^2$. *If* $u_n = n^{-1} \sigma_n^2 \log n$ *verifies* $\lim_{n\to\infty} u_n = 0$, *and if* $M/\sigma_n^2 < C < \infty$ *then we have :*

$$\frac{1}{n} \sum_i^n Z_i = O_{a.co.}(\sqrt{u_n}).$$

3.6 Topological considerations

3.6.1 Kolmogorov's entropy

As specified in Ferraty and Vieu (2006), all the asymptotic results in nonparametric statistics for functional variables are closely related to the concentration properties of the probability measure of the functional variable X. Here, we have to consider the uniformity aspect. To this end, let $\mathcal{S}_{\mathcal{F}}$ be a fixed subset of $\mathcal{H}$; we the following assumption:

$$\forall x \in \mathcal{S}, 0 < C\phi_x(h) \leq \mathbb{P}(X \in B(x,h)) \leq C'\phi_x(h) < \infty.$$

The firs contribution of the topological structure of the functional space can be seen through the function ϕ_x controlling the concentration of the measure of probability of the functional variable on as mall ball. Further, for the uniform consistency, where the main tool is to cover a subset $\mathcal{S}_{\mathcal{F}}$ with a finite number of balls, one introduces an other topological concept defined as follows:

Definition 3.6.1. *Let $\mathcal{S}_{\mathcal{F}}$ be a subset of a semi-metric space $\mathcal{H}$, and let $\epsilon > 0$ be given. A finite set of points $x_1, x_2, ..., x_N$ in $\mathcal{F}$ is called an $\epsilon-$ set for $\mathcal{S}_{\mathcal{F}}$ if $\mathcal{S}_{\mathcal{F}} \subset \bigcup_{k=1}^{N} B(x_k, \epsilon)$. The quantity $\psi_{\mathcal{S}_{\mathcal{F}}} = \log(N_\epsilon(\mathcal{S}_{\mathcal{F}}))$, where $N_\epsilon(\mathcal{S}_{\mathcal{F}})$ is the minimal number of open balls in $\mathcal{F}$ of radius ϵ which is necessary to cover $\mathcal{S}_{\mathcal{F}}$, is called the Kolmogorov's ϵ-entropy of the set $\mathcal{S}_{\mathcal{F}}$.*

This concept was presented for the first time by Kolmogorov in the mid-1950's (see Kolmogorov and Tikhomirov (1959)) and it represents a measure of the complexity of a set, in sense that, high entropy means that much information is needed to describe an element with an accuracy ϵ. Then, the choice of the topological structure will play a crucial role when one is looking at uniform (over some subset $\mathcal{S}_{\mathcal{F}}$) of $\mathcal{F}$ asymptotic results. In Ferraty et al. (2006), the phenomenon of concentration of the probability measure of the functional variable by computing the small ball probabilities in standard situations has been highlighted. For more details on entropy and small ball probabilities) or/and the use of the Kolmogorov's $\epsilon-$entropy in dimensionality reduction problems, see Kuelbs and Li (1993) or/and Theodoros and Yannis (1997).

Chapter 4

Strong uniform consistency rates of conditional hazard estimation in the single functional index model for dependant functional data under random censorship

4.1 Introduction

The estimation of the hazard function is a problem of considerable interest, especially to inventory theorists, medical researchers, logistics planners, reliability engineers and seismologists Rabhi et al. (2015). Nonparametric estimation of the hazard function has been extensively discussed in the literature Quintela (2008). The first who introduced the estimation of the hazard rate was Watson and Leadbetter (1964), after that many works were given on these topic; Ahmad (1976), Singpurwalla and Wong (1983) and many others. Recently, Massim and Mechab (2016) presented the local linear estimation of the conditional hazard function, Quintela (2007) can be sited for a survey.

Single-index models are becoming increasingly popular in many scientific fields including biostatistics, medicine, economics and financial econometrics Cui et al. (2011). This sort of kind modelization is excessively studied in the multivariate case, let's cite for instance Härdle et al. (1993) and Hristache et al. (2001). Based on the regression function, Delecroix et al. (2003) constructed an asymptotically efficient estimator for general conditional single-index response models; the estimation and some asymptotic properties of the single-index models were established. Let's note that when the explanatory variable is functional the literature is strictly limited. Ferraty et al. (2003) were the first who obtained the asymptotic properties in the fixed functional single-model, authors established the almost complete convergence, in the i.i.d. case of the link regression function of this model. Their results were extended to dependent case by Aït Saidi et al. (2005). The case where the functional single-index is unknown was studied by Aït Saidi et al. (2008). Later, many authors focused on the study of conditional single-index models; Mahiddine et al. (2014) studied the nonparametric estimation of some characteristics of the conditional distribution in single functional index model, Bouchentouf et al. (2014) presented a nonparametric estimation of hazard function with functional explicatory variable in single functional index and the variables are independent. The authors proved some consistency properties (with rates) in various situations, including censored and complete data. Bouchentouf et al. (2015) gave a strong uniform consistency rates of conditional quantiles for time series data in the single functional index model, they estimated nonparametrically the quantiles of a conditional distribution when the sample is considered as an α-mixing sequence.

The present work presents a study of a nonparametric estimation of the conditional hazard function, when the covariate is functional and when the sample is considered as an α-mixing sequence. We prove the consistency properties in various situations; censored and complete variables. The pointwise almost complete convergence and the uniform almost complete convergence (with rate) of the kernel estimator of this model are established.

The chapter is organized as follows, in Section 4.2 the model of the hazard rate predictor for functional single functional index is presented. In Section 4.3.1 we give results in the simple framework of dependent and complete data. Extensions to censored framework are presented in Section 4.3.2. To complete the range of our results, we extend them to the frame of the uniform almost complete convergence. Specifically, we give in sections 4.4.1 and 4.4.2 the uniform convergence results when the variables are derived from a dependent process.

The proof of the results presented in Section 4.3 will be given using up the existing literature. Then, technical details of the proofs of the results done in Sections 4.3.2 and 4.4.2 are shown at the end of chapter.

4.2 The model

Let X be a random variable associated to a lifetime (ie, a random variable with values in $\mathbb{R}^+$).

When X has a density f with respect to the measure of Lebesgues, the hazard rate is written, for all x as follows:

$$h(x) = \frac{f(x)}{S(x)},$$

where, f is the density function, $S = 1 - F$ is survival function of X, and F denotes the distribution function of X such that $F(x) < 1$.

Let the conditional random rate for $x > 0$,

$$h^Z(x) = \frac{f^Z(x)}{S^Z(x)}, \tag{4.1}$$

with $f^Z(\cdot)$ the conditional density, $S^Z = 1 - F^z$ the conditional survival function and $F^Z(\cdot)$

the conditional distribution function of X knowing Z.

Let z be a fixed element of the functional space $\mathcal{H}$, $\mathcal{N}_z$ denotes a fixed neighborhood of z and $\mathcal{S}_{\mathbb{R}}$ is a fixed compact of $\mathbb{R}^+$. Here, we give an assumption on the concentration function $\phi_{\theta,z}$:

(H0) $\forall h_K > 0,\ \mathbb{P}(Z \in B_\theta(z,h_K)) = \phi_{\theta,z}(h_K) > 0,\ \phi_{\theta,z}(h_K) \underset{h_K \to 0}{\longrightarrow} 0,$

where K is a kernel, H is a distribution function and $h_K = h_{K,n}$.

(H1a) The sequence $(X_i, Z_i)_{i\in\mathbb{N}}$ is α-mixing and its mixing coefficients $\alpha(n)$ are such that:

$$\exists a,\ c \in \mathbb{R}_+^* :\ \forall n \in \mathbb{N}\ \ \alpha(n) \le cn^{-a}.$$

(H1b) The joint density (Y_i, Y_j) knowing (Z_i, Z_j) exists and is bounded, and

$$\exists \gamma_1 \in]0,1],\ 0 < \sup_{i\neq j} \mathbb{P}\Big((Z_i, Z_j) \in B_\theta(z,h) \times B_\theta(z,h)\Big) = \mathcal{O}\Big(\phi_z(h))^{1+\gamma_1}\Big).$$

(H1c) $$\exists \gamma_2 \in]0,1[, a > \frac{1+\gamma_1}{\gamma_1\gamma_2} \text{ and } h_H \phi_{\theta,z}(h_K) = \mathcal{O}(n^{-\gamma_2}).$$

Let's note that these hypotheses are common in nonparametric estimation problems with dependent variables, functional or not (see Ferraty and Vieu (2006), Chapter 11).

The nonparametric model on the estimated function h^Z will be determined by the regularity conditions on the conditional distribution of X knowing Z. These conditions are the following:

(H2) $\exists A_{\theta,z} < \infty, \exists b_1, b_2 > 0, \forall (x_1,x_2) \in \mathcal{S}_{\mathbb{R}}^2,\ \forall (z_1,z_2) \in \mathcal{N}_z^2 :$

$$|F(\theta,x_1,z_1) - F(\theta,x_2,z_2)| \le A_{\theta,z}\Big(\|z_1,z_2\|^{b_1} + |x_1 - x_2|^{b_2}\Big),$$
$$|f(\theta,x_1,z_1) - f(\theta,x_2,z_2)| \le A_{\theta,z}\Big(\|z_1,z_2\|^{b_1} + |x_1 - x_2|^{b_2}\Big);$$

(H3) $\exists \nu < \infty, \forall (x,z') \in \mathcal{S}_{\mathbb{R}} \times \mathcal{N}_z,\ \ f(\theta,x,z') \le \nu;$

(H4) $\exists \beta > 0,\ \forall (x,z') \in \mathcal{S}_{\mathbb{R}} \times \mathcal{N}_z,\ \ F(\theta,x,z') \le 1 - \beta.$

Let $(X_i, Z_i)_{1\le i\le n}$ be random variables, each of them follows the same law of a couple (X, Z) where X is valued in $\mathbb{R}$ and Z has values in the Hilbert space $(\mathcal{H}, <\cdot;\cdot>)$. In this section, we suppose that X_i and Z_i are observed.

Now, it is natural to construct an estimator of the function $h(\theta,\cdot,Z)$. To estimate the conditional distribution function and the conditional density in the presence of variable Z, Mahiddine et al. (2014) proposed the following functional kernel estimators:

$$\hat{F}(\theta,x,z) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1}(< z-Z_i,\theta >)\right) H\left(h_H^{-1}(x-X_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}(< z-Z_i,\theta >)\right)}, \tag{4.2}$$

and

$$\hat{f}(\theta,x,z) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1}(< z-Z_i,\theta >)\right) H'\left(h_H^{-1}(x-X_i)\right)}{h_H \sum_{i=1}^{n} K\left(h_K^{-1}(< z-Z_i,\theta >)\right)}, \tag{4.3}$$

with K a kernel, H a distribution function and $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers.

A kernel estimator of the functional conditional hazard function $h(\theta,\cdot,Z)$ may therefore be constructed in the following way:

$$\hat{h}(\theta,x,Z) = \frac{\hat{f}(\theta,x,Z)}{1-\hat{F}(\theta,x,Z)}. \tag{4.4}$$

The assumptions we need later for the parameters of the estimator, i.e. on K, H, h_H and h_K are not restrictive.
Next, we introduce the following conditions which guarantee the good behavior of the estimators $\hat{F}(\theta,x,Z)$ and $\hat{f}(\theta,x,Z)$ (see Ferraty and Vieu (2006)):

(H5) The cumulative kernel H is derivable such that:

- i) $\exists A < \infty, \forall (x_1,x_2) \in \mathbb{R}^2, |H'(x_1)-H'(x_2)| \le A|x_1-x_2|$;
- ii) H' is of compact support with values in $[-1,1]$ and $H'(t) > 0, \forall t \in [-1,1]$.

(H6) The functional kernel K verifies the following conditions:

- i) K is of compact support with values in $(0,1)$;
- ii) $\exists A_1, A_2, \forall t \in (0,1),\ 0 < A_1 < K(t) < A_2 < \infty$.

(H7) The bandwidth h_K has to satisfy

$$\lim_{n\to\infty} h_K = 0 \text{ and } \lim_{n\to\infty} \frac{\log n}{n h_H \phi_{\theta,x}(h_K)} = 0,$$

(H8) The bandwidth h_H has to satisfy

$$\lim_{n\to\infty} h_H = 0 \text{ and } \exists a > 0, \lim_{n\to\infty} n^a h_H = \infty.$$

Under these general conditions, we establish in Section 1.4 the pointwise convergence of the kernel estimator $\hat{h}(\theta, x, z)$ of the functional conditional hazard function $h(\theta, x, z)$ when the observed sample is complete. In section 1.4, these results will be generalized to censored variables.

4.2.1 Censored data

Estimation of the hazard function when the data are censored is an important problem in medicine. This problem is usually modeled by considering a positive variable called C, and the observed random variables (T_i, Δ_i, Z_i), where $T_i = \min(X_i,\ C_i)$ and $\Delta_i = I_{X_i \leq C_i}$. In the following we use the notations $F_1(\theta, \cdot, Z)$ and $f_1(\theta, \cdot, Z)$ to describe the distribution function and conditional density of C knowing Z and we use the notation $S_1(\theta, \cdot, Z) = 1 - F_1(\theta, \cdot, Z)$.

If we introduce the notation $L(\theta, \cdot, Z) = 1 - S_1(\theta, \cdot, Z) S(\theta, \cdot, Z)$ and $\varphi(\theta, \cdot, Z) = f(\theta, \cdot, Z) S_1(\theta, \cdot, Z)$, we can reformulate the expression (4.1) as follow:

$$h(\theta, t, Z) = \frac{\varphi(\theta, t, Z)}{1 - L(\theta, t, Z)},\ \forall t,\ L(\theta, t, Z) < 1. \tag{4.5}$$

So, we can define $\varphi(\theta, \cdot, Z)$ and $L(\theta, \cdot, Z)$ by setting

$$\hat{L}(\theta, t, Z) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right) H\left(h_H^{-1}(t - T_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right)}, \tag{4.6}$$

and

$$\widetilde{\varphi}(\theta, t, Z) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right) \Delta_i H'\left(h_H^{-1}(t - T_i)\right)}{h_H \sum_{i=1}^{n} K\left(h_K^{-1}(< z - Z_i, \theta >)\right)}. \tag{4.7}$$

Finally, the hazard function estimator is given as:

$$\widetilde{h}(\theta,t,Z)=\frac{\widetilde{\varphi}(\theta,t,Z)}{1-\hat{L}(\theta,t,Z)}. \tag{4.8}$$

In addition to the assumptions introduced above, we need some additional conditions:

(H9) Conditionally to Z, the variables X and C are independent.

(H10) $\exists A_{\theta,z}<\infty, \exists b_1,b_2>0, \forall(t_1,t_2)\in\mathcal{S}_{\mathbb{R}}^2,\ \forall(z_1,z_2)\in\mathcal{N}_z^2:$

$$|L(\theta,t_1,z_1)-L(\theta,t_2,z_2)|\le A_{\theta,z}\left(\|z_1-z_2\|^{b_1}+|t_1-t_2|^{b_2}\right)$$
$$|\varphi(\theta,t_1,z_1)-\varphi(\theta,t_2,z_2)|\le A_{\theta,z}\left(\|z_1-z_2\|^{b_1}+|t_1-t_2|^{b_2}\right);$$

(H11) $\exists\mu<\infty,\ \varphi(\theta,t,z')<\mu,\ \forall(t,z')\in\mathbb{R}_+\times\mathcal{N}_z$.

(H12) $\exists\eta>0,\ L(\theta,t,z')\le 1-\eta,\ \ \forall(t,z')\in\mathbb{R}_+\times\mathcal{N}_z$.

We begin by studying a statistical samples satisfying a classical assumption of dependency, the couples (X_i,Z_i) are dependent.

(H13a) The sequence $(X_i,C_i,Z_i)_{i\in\mathbb{N}}$ is α-mixing and its mixing coefficients $\alpha(n)$ are as:

$$\exists a,\ c\in\mathbb{R}_+^*:\ \forall n\in\mathbb{N}\ \ \alpha(n)\le cn^{-a}.$$

(H13b) The join distribution of (Y_i,Y_j) knowing (Z_i,Z_j) exists and is bounded, and

$$\exists\gamma_1\in]0,1]:\ 0<\sup_{i\neq j}\mathbb{P}\left((Z_i,Z_j)\in B_\theta(z,h)\times B_\theta(z,h)\right)=\mathcal{O}\left(\phi_{\theta,z}(h_K)^{1+\gamma_1}\right).$$

(H13c) $\exists\gamma_2\in]0,1[,\ a>\frac{1+\gamma_1}{\gamma_1\gamma_2}$ and $h_H\phi_{\theta,z}(h_K)=\mathcal{O}(n^{-\gamma_2})$.

4.3 Pointwise almost complete Convergence

The objective of this part is to establish the pointwise almost complete convergence of the kernel estimator $\hat{h}(\theta,\cdot,Z)$ of the conditional hazard function $h(\theta,\cdot,Z)$ including censored and complete variables.

4.3.1 Case of complete data

Theorem 4.1. *Under hypotheses (H0)-(H8), we have:*

$$\sup_{x\in\mathcal{S}_{\mathbb{R}}}|\hat{h}(\theta,x,z)-h(\theta,x,z)|=\mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_{\theta,z}(h_K)}}\right). \tag{4.9}$$

Proof. The proof of the Theorem is based on the following inequality, valid for any $x \in \mathcal{S}_{\mathbb{R}}$:

$$\begin{aligned}
\hat{h}(\theta,x,z) - h(\theta,x,z) &= \frac{1}{(1-\hat{F}(\theta,x,z))(1-F(\theta,x,z))}\Big(\hat{f}(\theta,x,z) - f(\theta,x,z)\Big) \\
&\quad + \frac{f(\theta,x,z)}{(1-\hat{F}(\theta,x,z))(1-F(\theta,x,z))}\Big(\hat{F}(\theta,x,z) - F(\theta,x,z)\Big) \\
&\quad - \frac{F(\theta,x,z)}{(1-\hat{F}(\theta,x,z))(1-F(\theta,x,z))}\Big(\hat{f}(\theta,x,z) - f(\theta,x,z)\Big) \\
&= \frac{1}{1-\hat{F}(\theta,x,z)}\Big(\hat{f}(\theta,x,z) - f(\theta,x,z)\Big) \\
&\quad + \frac{h(\theta,x,z)}{1-\hat{F}(\theta,x,z)}\Big(\hat{F}(\theta,x,z) - F(\theta,x,z)\Big).
\end{aligned}$$

Thus,

$$\begin{aligned}
\sup_{x\in\mathcal{S}_{\mathbb{R}}}\Big|\hat{h}(\theta,x,z) - h(\theta,x,z)\Big| &\leq \frac{1}{\inf_{x\in\mathcal{S}_{\mathbb{R}}}\Big|1-\hat{F}(\theta,x,z)\Big|}\left(\sup_{x\in\mathcal{S}_{\mathbb{R}}}\Big|\hat{f}(\theta,x,z) - f(\theta,x,z)\Big|\right) \\
&\quad + \frac{\sup_{x\in\mathcal{S}_{\mathbb{R}}}|h(\theta,x,z)|}{\inf_{x\in\mathcal{S}_{\mathbb{R}}}\Big|1-\hat{F}(\theta,x,z)\Big|}\left(\sup_{x\in\mathcal{S}_{\mathbb{R}}}\Big|\hat{F}(\theta,x,z) - F(\theta,x,z)\Big|\right), \quad (4.10)
\end{aligned}$$

which leads to (a constant $C < \infty$):

$$\sup_{x\in\mathcal{S}_{\mathbb{R}}}\Big|\hat{h}(\theta,x,z) - h(\theta,x,z)\Big| \leq C\,\frac{\Big\{\sup_{x\in\mathcal{S}_{\mathbb{R}}}\Big(\Big|\hat{f}(\theta,x,z) - f(\theta,x,z)\Big| + \Big|\hat{F}(\theta,x,z) - F(\theta,x,z)\Big|\Big)\Big\}}{\inf_{x\in\mathcal{S}_{\mathbb{R}}}\Big|1-\hat{F}(\theta,x,z)\Big|}. \quad (4.11)$$

Then, the rest of the proof is based on the following properties:

$$\sup_{x\in\mathcal{S}_{\mathbb{R}}}|F(\theta,x,z) - \hat{F}(\theta,x,z)| = \mathcal{O}\Big(h_K^{b_1} + h_H^{b_2}\Big) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,\phi_{\theta,z}(h_K)}}\right), \quad (4.12)$$

and

$$\sup_{x\in\mathcal{S}_{\mathbb{R}}}|f(\theta,x,z) - \hat{f}(\theta,x,z)| = \mathcal{O}\Big(h_K^{b_1} + h_H^{b_2}\Big) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_{\theta,z}(h_K)}}\right), \quad (4.13)$$

and the next result which is a consequence of property (4.12).

Corollary 4.2. *Under the conditions of Theorem 4.1, we have*

$$\exists \delta > 0 \text{ such that } \sum_{n=1}^{\infty}\mathbb{P}\left\{\inf_{x\in\mathcal{S}_{\mathbb{R}}}|1-\hat{F}(\theta,x,z)| < \delta\right\} < \infty.$$

The proof of the properties (4.12) and (4.13) is similar as in Ferraty et al. (2008). These results are an be seen as a particular case of Propositions 11.22.ii et 11.23.ii given in Ferraty and Vieu (2006).

4.3.2 Case of censored data

The main goal of this part is to study the asymptotic properties in the broader context of a censored sample. We begin by studying statistical samples satisfying a standard assumption of dependency, ie. the triples (X_i, C_i, Z_i) are dependent such that condition (H13) is satisfied.

Theorem 4.3. *Under assumptions (H0)-(H12), and (H13a)-(H13c), we have:*

$$\sup_{t\in\mathcal{S}_{\mathbb{R}}} |\widetilde{h}(\theta,t,z) - h(\theta,t,z)| = \mathcal{O}\left(h_K^{b_1} + h_H^{b_2}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\, h_H\, \phi_{\theta,z}(h_K)}}\right). \tag{4.14}$$

Proof. The result is based on the bellow decomposition, wherein C is a real constant strictly positive:

$$\begin{aligned}\sup_{t\in\mathcal{S}_{\mathbb{R}}} \left|\widetilde{h}(\theta,t,z) - h(\theta,t,z)\right| \;\leq\; & \frac{1}{\inf_{t\in\mathcal{S}_{\mathbb{R}}} |1-\hat{L}(\theta,t,z)|}\left\{\sup_{t\in\mathcal{S}_{\mathbb{R}}} |\widetilde{\varphi}(\theta,t,z) - \varphi(\theta,t,z)|\right. \\ + & \left.\frac{\sup_{t\in\mathcal{S}_{\mathbb{R}}} |h(\theta,t,z)|}{\inf_{t\in\mathcal{S}_{\mathbb{R}}} |1-\hat{L}(\theta,t,z)|}\sup_{t\in\mathcal{S}_{\mathbb{R}}} |\hat{L}(\theta,t,z) - L(\theta,t,z)|\right\},\end{aligned} \tag{4.15}$$

which leads to (a constant $C<\infty$):

$$\sup_{t\in\mathcal{S}_{\mathbb{R}}} \left|\widetilde{h}(\theta,t,z) - h(\theta,t,z)\right| \leq \frac{\sup_{t\in\mathcal{S}_{\mathbb{R}}}\left\{|\widetilde{\varphi}(\theta,t,z) - \varphi(\theta,t,z)| + \left|L(\theta,t,z) - \hat{L}(\theta,t,z)\right|\right\}}{\inf_{t\in\mathcal{S}_{\mathbb{R}}}\left|1-\hat{L}(\theta,t,z)\right|}. \tag{4.16}$$

The announced result follows from the following property:

$$\sup_{t\in\mathcal{S}_{\mathbb{R}}} \left|\hat{L}(\theta,t,Z) - L(\theta,t,Z)\right| = \mathcal{O}\left(h_K^{b_1} + h_H^{b_2}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_{\theta,z}(h_K)}}\right), \tag{4.17}$$

Lemma 4.3.1. *Under the hypotheses of Theorem 4.1, we have*

$$\sup_{t\in\mathcal{S}_{\mathbb{R}}} |\widetilde{\varphi}(\theta,t,Z) - \varphi(\theta,t,Z)| = \mathcal{O}\left(h_K^{b_1} + h_H^{b_2}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\, h_H\, \phi_{\theta,z}(h_K)}}\right), \tag{4.18}$$

and

Corollary 4.4. *Under the conditions of Theorem 4.1, we have*

$$\exists\delta > 0 \text{ such that } \sum_{n=1}^{\infty} \mathbb{P}\left\{\inf_{x\in\mathcal{S}_{\mathbb{R}}} |1 - \hat{L}(\theta, x, z)| < \delta\right\} < \infty.$$

Following Ferraty and Vieu (2006), the property (4.17) remains valid in the case of single index. Therefore, the result (4.3) follows from (4.16) and Lemma 4.3.1.

4.4 Uniform almost complete convergence

In this part of chapter, we derive the uniform version of Theorem 4.1. To this end, some additional tools and topological conditions are required (see Ferraty et al. (2003) for more discussion on the uniform convergence in nonparametric functional statistics). Thus, in addition to the conditions introduced previously, we need the following ones. Firstly, consider

$$\mathcal{S}_{\mathcal{H}} \subset \bigcup_{k=1}^{d_n^{\mathcal{S}_{\mathcal{H}}}} B(z_k, r_n) \text{ and } \Theta_{\mathcal{H}} \subset \bigcup_{j=1}^{d_n^{\Theta_{\mathcal{H}}}} B(s_j, r_n) \tag{4.19}$$

with x_k (resp. t_j) $\in \mathcal{H}$ and $r_n, d_n^{\mathcal{S}_{\mathcal{H}}}, d_n^{\Theta_{\mathcal{H}}}$ are sequences of positive real numbers which tend to infinity as n goes to infinity.

4.4.1 Case of complete data

To establish the uniform almost complete convergence of our estimator defined in (4.4), we need the following assumptions:

(A1) There exists a differentiable function $\phi(\cdot)$ such that $\forall x \in \mathcal{S}_{\mathcal{H}}$ and $\forall\theta \in \Theta_{\mathcal{H}}$,

$$0 < C\phi(h) \le \phi_{\theta,x}(h) \le C'\phi(h) < \infty \text{ and } \exists\eta_0 > 0,\ \forall\eta < \eta_0,\ \phi'(\eta) < C,$$

(A2) $\exists b_1, b_2 > 0,\ \forall(x_1, x_2) \in \mathcal{S}_{\mathbb{R}} \times \mathcal{S}_{\mathbb{R}}, \forall(z_1, z_2) \in \mathcal{S}_{\mathcal{H}} \times \mathcal{S}_{\mathcal{H}}$ and $\forall\theta \in \Theta_{\mathcal{H}}$,

$$|F(\theta, x_1, z_1) - F(\theta, x_2, z_2)| \le A\left(\|z_1, z_2\|^{b_1} + |x_1 - x_2|^{b_2}\right),$$
$$|f(\theta, x_1, z_1) - f(\theta, x_2, z_2)| \le A\left(\|z_1, z_2\|^{b_1} + |x_1 - x_2|^{b_2}\right);$$

(A3) $\exists\nu < \infty, \forall(x, z') \in \mathcal{S}_{\mathbb{R}} \times \mathcal{N}_z,\ \ \forall\theta \in \Theta_{\mathcal{H}},\ \ f(\theta, x, z') \le \nu$;

(A4) $\exists\beta > 0,\ \forall (x,z') \in \mathcal{S}_{\mathbb{R}} \times \mathcal{N}_z, \quad \forall \theta \in \Theta_{\mathcal{H}}, \ \ F(\theta,x,z') \leq 1-\beta.$

(A5) The kernel K satisfies (H3) and Lipschitz's condition holds:

$$|K(x) - K(y)| \leq C\|x-y\|.$$

(A6) For $r_n = \mathcal{O}\left(\frac{\log n}{n}\right)$ the sequences $d_n^{\mathcal{S}_{\mathcal{H}}}$ and $d_n^{\Theta_{\mathcal{H}}}$ satisfy:

$$\frac{(\log n)^2}{n\phi(h_K)} < \log d_n^{\mathcal{S}_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}} < \frac{n\phi(h_K)}{\log n},$$

and $\sum_{n=1}^{\infty} n^{1/2b_2} (d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}})^{1-\beta} < \infty$ for some $\beta > 1$.

(A7) For some $\gamma \in (0,1)$, $\lim_{n\to\infty} n^{\gamma} h_H = \infty$, and for $r_n = \mathcal{O}\left(\frac{\log n}{n}\right)$ the sequences $d_n^{\mathcal{S}_{\mathcal{F}}}$ and $d_n^{\Theta_{\mathcal{F}}}$ satisfy:

$$\frac{(\log\ n)^2}{n h_H \phi(h_K)} < \log\ d_n^{\mathcal{S}_{\mathcal{F}}} + \log d_n^{\Theta_{\mathcal{F}}} < \frac{n h_H \phi(h_K)}{\log n},$$

and $\sum_{n=1}^{\infty} n^{(3\gamma+1)/2} (d_n^{\mathcal{S}_{\mathcal{F}}} d_n^{\Theta_{\mathcal{F}}})^{1-\beta} < \infty$, for some $\beta > 1$

Let

$$s_{n,0}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \left| Cov\left(\psi_i(x,\theta), \psi_j(x,\theta)\right)\right|$$

$$s_{n,1}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \left| Cov\left(\psi_i(x,\theta) H_i(t)), \psi_j(x,\theta) H_j(t))\right)\right|$$

$$s_{n,3}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \left| Cov\left(\Lambda_i, \Lambda_j\right)\right|, \quad s_{n,4}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \left| Cov\left(\Omega_i, \Omega_j\right)\right|$$

$$s_{n,5}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \left| Cov\left(\psi_i\left(x_{k(x)}, \theta_{m(\theta)}\right), \psi_j\left(x_{k(x)}, \theta_{m(\theta)}\right)\right)\right|, \quad s_{n,6}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \left| Cov\left(\Gamma_i, \Gamma_j\right)\right|$$

$$s_{n,7}^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \left| Cov\left(\Gamma_i^{(l)}, \Gamma_j^{(l)}\right)\right|,$$

where

$$\psi_i(x,\theta) = \frac{K(h_K^{-1}(< x - X_i, \theta >))}{\mathbb{E}K_1(\theta,x)}$$

$$\Lambda_i(x,\theta) = \frac{1}{h_K\phi(h_K)} \mathbf{1}_{B_\theta(x,h)\cup B_\theta(x_{k(x)},h)}(X_i),$$

$$\Omega_i(x,\theta) = \frac{1}{h_K\phi(h_K)} \mathbf{1}_{B_\theta(x_{k(x)},h)\cup B_{\theta_{m(\theta)}}(x_{k(x)},h)}(X_i),$$

$$\psi_i\left(x_{k(x)}, \theta_{m(\theta)}\right) = \frac{K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{m(\theta)} >\right)}{\mathbb{E}K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{m(\theta)} >\right)},$$

$$
\begin{aligned}
\Gamma_i &= \frac{K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{m(\theta)} >\right)}{\mathbb{E}K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{m(\theta)} >\right)} H\left(h_H^{-1}(t_y - Y_i)\right) \\
&- \mathbb{E}\left(\frac{K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{m(\theta)} >\right)}{\mathbb{E}K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{m(\theta)} >\right)} H\left(h_H^{-1}(t_y - Y_i)\right)\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\Gamma_i^{(l)} &= \frac{1}{h_H^l} \frac{K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{m(\theta)} >\right)}{\mathbb{E}K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{m(\theta)} >\right)} H^{(l)}\left(h_H^{-1}(t_y - Y_i)\right) \\
&- \frac{1}{h_H^l}\mathbb{E}\left(\frac{K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{m(\theta)} >\right)}{\mathbb{E}K\left(h_K^{-1} < x_{k(x)} - X_i, \theta_{m(\theta)} >\right)} H^{(l)}\left(h_H^{-1}(t_y - Y_i)\right)\right)
\end{aligned}
$$

Remark 4.4.1. *Note that assumptions (A1)-(A4) are respectively, the uniform version of (H1a)-(H4). Assumptions (A1) and (A6) are linked with the the topological structure of the functional variable, see Ferraty and Vieu (2006).*

Theorem 4.5. *Under hypotheses (H0)-(H1),(H13b)-(H13c) and (A1)-(A7) we have:*

$$
\begin{aligned}
\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in\mathcal{S}_{\mathcal{H}}} \sup_{y\in\mathcal{S}_{\mathbb{R}}} |\hat{h}(\theta,y,x) - h(\theta,y,x)| &= \mathcal{O}\left(h_K^{b_1} + h_H^{b_2}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n h_H \phi(h_K)}}\right) \\
&+ \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n^{'*2} \log d_n^{\mathcal{S}_{\mathcal{F}}} d_n^{\Theta_{\mathcal{F}}}}}{n}\right).
\end{aligned}
$$

where $s_n^{'*} = \max\{s_{n,0}; s_{n,1}; s_{n,2}; s_{n,3}; s_{n,4}; s_{n,5}; s_{n,7}\}$.

In the particular case, where the functional single-index is fixed we get the following result.

Corollary 4.6. *Under assumptions of Theorem 4.5, as n goes to infinity, we have*

$$
\begin{aligned}
\sup_{x\in\mathcal{S}_{\mathcal{H}}} \sup_{y\in\mathcal{S}_{\mathbb{R}}} |\hat{h}(\theta,y,x) - h(\theta,y,x)| &= \mathcal{O}\left(h_K^{b_1} + h_H^{b_2}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}}{n h_H \phi(h_K)}}\right) \\
&+ \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n^{'*2} \log d_n^{\mathcal{S}_{\mathcal{F}}}}}{n}\right).
\end{aligned}
$$

Proof of Theorem 4.5. Clearly, the proofs of these two results, namely Theorem 4.5 and Corollary 4.6 can be deduced easily from the following intermediate results given in Bouchentouf et al. (2015).

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}\left|\widehat{F}(\theta,y,x)-F(\theta,y,x)\right| = \mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right) +\mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n^{*2}\log d_n^{\mathcal{S}_{\mathcal{F}}}d_n^{\Theta_{\mathcal{F}}}}}{n}\right)$$

and

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}\left|\widehat{f}(\theta,y,x)-f(\theta,y,x)\right| = \mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right) +\mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n^{'*2}\log d_n^{\mathcal{S}_{\mathcal{F}}}d_n^{\Theta_{\mathcal{F}}}}}{nh_H}\right)$$

where $s_n^* = \max\{s_{n,0};s_{n,1};s_{n,2};s_{n,3};s_{n,4};s_{n,5};s_{n,6}\}$, $s_n^{'*} = \max\{s_{n,0};s_{n,1};s_{n,2};s_{n,3};s_{n,4};s_{n,5};s_{n,7}\}$.

4.4.2 Case of censored data

To study the uniform almost complete convergence of our estimator defined above (4.8), we need the following assumptions:

(A2a) $\forall(t_1,t_2)\in\mathcal{S}_{\mathbb{R}}\times\mathcal{S}_{\mathbb{R}}, \forall(z_1,z_2)\in\mathcal{S}_{\mathcal{H}}\times\mathcal{S}_{\mathcal{H}}$ and $\forall\theta\in\Theta_{\mathcal{H}}$,

$$|L(\theta,t_1,z_1)-L(\theta,t_2,z_2)| \le A\left(\|z_1,z_2\|^{b_1}+|t_1-t_2|^{b_2}\right),$$
$$|\varphi(\theta,t_1,z_1)-\varphi(\theta,x_2,z_2)| \le A\left(\|z_1,z_2\|^{b_1}+|t_1-t_2|^{b_2}\right),$$

(A3a) $\exists\nu<\infty, \forall(t,z')\in\mathcal{S}_{\mathbb{R}}\times\mathcal{N}_z,\ \ \forall\theta\in\Theta_{\mathcal{H}},\ \ \varphi(\theta,t,z')\le\nu;$

(A4a) $\exists\beta>0,\ \forall(t,z')\in\mathcal{S}_{\mathbb{R}}\times\mathcal{N}_z,\ \ \forall\theta\in\Theta_{\mathcal{H}},\ \ L(\theta,t,z')\le 1-\beta.$

Theorem 4.7. *Under hypotheses (A1), (A5)-(A7) and (A2a)-(A4a), we get:*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\widetilde{h}(\theta,t,z)-h(\theta,t,z)| = \mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}+\log d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right) + \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n^{'*2}\log d_n^{\mathcal{S}_{\mathcal{F}}}d_n^{\Theta_{\mathcal{F}}}}}{n}\right).$$

Now, when the functional single-index is fixed we have.

Corollary 4.8. *Under assumptions (A1), (A5)-(A7), (A2a)-(A4a) and (H4), as n goes to infinity, we have*

$$\begin{aligned}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\widetilde{h}(\theta,t,z)-h(\theta,t,z)| &= \mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right)\\ &+ \mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n^{'*2}\log d_n^{\mathcal{S}_{\mathcal{F}}}}}{n}\right).\end{aligned}$$

Proof of Theorem 4.7. The result is based on the decomposition (4.15). Clearly the proofs of these two results namely the Theorem 4.7 and Corollary 4.8 can be deduced from the following intermediate results which are only uniform version of properties (4.17) and (4.18).

The properties of the estimators $\hat{L}(\theta,\cdot,z)$ and $\widetilde{\varphi}(\theta,\cdot,z)$ are given in Lemma 4.4.5.

Finally, the desired result is obtained directly from (4.15), (4.20) and (4.21).

$$\begin{aligned}\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}\left|\widehat{L}(\theta,y,x)-L(\theta,y,x)\right| &= \mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right)\\ &+\mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n^{*2}\log d_n^{\mathcal{S}_{\mathcal{F}}}d_n^{\Theta_{\mathcal{F}}}}}{n}\right)\end{aligned}$$

and

$$\begin{aligned}\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}\left|\widehat{\varphi}(\theta,y,x)-\varphi(\theta,y,x)\right| &= \mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right)\\ &+\mathcal{O}_{a.co.}\left(\frac{\sqrt{s_n^{'*2}\log d_n^{\mathcal{S}_{\mathcal{F}}}d_n^{\Theta_{\mathcal{F}}}}}{nh_H}\right)\end{aligned}$$

Lemma 4.4.1. *Under assumptions (A1), (A2) and (H5), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|L(\theta,t,z)-\mathbb{E}(\hat{L}_N(\theta,t,z))|=\mathcal{O}(h_K^{b_1})+\mathcal{O}(h_H^{b_2}).$$

Lemma 4.4.2. *Under assumptions (A1), (A5)-(A7) and (A2a)-(A4a) we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\hat{L}_N(\theta,t,z)-\mathbb{E}\left[\hat{L}_N(\theta,t,z)\right]|=\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}+\log\, d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

Lemma 4.4.3. *Under assumptions (A1), (A2a) and (H5), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{F}}}\sup_{z\in\mathcal{S}_{\mathcal{F}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\varphi(\theta,t,z)-\mathbb{E}(\widetilde{\varphi}_N(\theta,t,z))|=\mathcal{O}(h_K^{b_1})+\mathcal{O}(h_H^{b_2}).$$

Lemma 4.4.4. *Under the assumptions (A1), (A5) ,(A2a), (A7) and (H5), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{F}}}\sup_{w\in\mathcal{S}_{\mathcal{F}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\widetilde{\varphi}_N(\theta,t,w)]-\mathbb{E}\,[\widetilde{\varphi}_N(\theta,t,w)]|=\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{nh_H\phi_{\theta,w}(h_K)}}\right).$$

Lemma 4.4.5. *Under hypotheses of Theorem 4.7, we have:*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\hat{L}(\theta,t,z)-L(\theta,t,z)|=\mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}+\log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right), \tag{4.20}$$

and

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\widetilde{\varphi}(\theta,t,z)-\varphi(\theta,t,z)|=\mathcal{O}\left(h_K^{b_1}+h_H^{b_2}\right)+\mathcal{O}_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}+\log d_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right). \tag{4.21}$$

Corollary 4.9. *Under assumptions (A1), (A5) and (A6), we have as $n\to\infty$*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}|\hat{\varphi}_D(\theta,z)-1|=\mathcal{O}_{a.co}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}+\log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right), \tag{4.22}$$

and

$$\sum_{n=1}^{\infty}\mathbb{P}\left(\inf_{\theta\in\Theta_{\mathcal{H}}}\inf_{z\in\mathcal{S}_{\mathcal{H}}}\hat{\varphi}_D(\theta,z)<\frac{1}{2}\right)<\infty. \tag{4.23}$$

4.5 Proofs of technical lemmas

In what follows C and c denote generic strictly positive real constants. Furthermore, the following notation are introduced: $K_i(\theta,z)=K(h_K^{-1}(<z-Z_i,\theta>)), H_i(t)=H'\left(h_H^{-1}(t-T_i)\right)$,

$$\widetilde{\varphi}_N(\theta,t,z) = \frac{1}{nh_H\mathbb{E}K_1(\theta,z)}\sum_{i=1}^{n} K_i(\theta,z)H_i(t)\Delta_i,\ \ \hat{\varphi}_D(\theta,z) = \frac{1}{n\mathbb{E}K_1(\theta,z)}\sum_{i=1}^{n} K_i(\theta,z),$$
$$V_i = \frac{1}{\mathbb{E}K_1(\theta,z)}K_i(\theta,z),$$
$$W_i = \frac{1}{h_H\mathbb{E}K_1(\theta,z)}K_i(\theta,z)H_i(t)\Delta_i,$$
$$s_n^2 = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} cov(V_{i_1},V_{i_2}),$$
$$S_n^2 = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} cov(W_{i_1},W_{i_2}).$$

Proof of Lemma 4.3.1. By using the following decomposition:

$$\widetilde{\varphi}(\theta,t,z) - \varphi(\theta,t,z) = \frac{(\widetilde{\varphi}_N(\theta,t,z) - \varphi_N(\theta,t,z))\,\varphi_D(\theta,z) - (\hat{\varphi}_D(\theta,z) - \varphi_D(\theta,z))\,\varphi_N(\theta,t,z)}{\hat{\varphi}_D(\theta,z)\varphi_D(\theta,z)}, \tag{4.24}$$

and because of the dependency of the variables and under the Proposition A6ii of Ferraty and Vieu (2006), the result of Lemma 4.3.1 will arise directly following the three properties:

$$|\hat{\varphi}_D(\theta,z) - 1| = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_{\theta,z}(h_K)}}\right), \tag{4.25}$$

$$\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\mathbb{E}\widetilde{\varphi}_N(\theta,t,z) - \varphi(\theta,t,z)| = \mathcal{O}(h_K^{b_1} + h_H^{b_2}), \tag{4.26}$$

and

$$\frac{1}{\hat{\varphi}_D(z)} sup_{t\in\mathcal{S}_{\mathbb{R}}}|\widetilde{\varphi}_N(\theta,t,z) - \mathbb{E}\widetilde{\varphi}_N(\theta,t,z)| = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_{\theta,z}(h_K)}}\right), \qquad a.co. \tag{4.27}$$

⇝ **Proof of (4.25).** It suffices to note that we can write

$$\hat{\varphi}_D(\theta,z) = \frac{1}{n}\sum_{i=1}^{n} V_i,$$

with

$$|V_i| = \mathcal{O}\left(\frac{1}{\phi_{\theta,z}(h)}\right), \tag{4.28}$$

and

$$\mathbb{E}V_i^2 = \mathcal{O}\left(\frac{1}{\phi_{\theta,z}(h)}\right). \tag{4.29}$$

The main step of the proof is to obtain the assessment of the sum of the covariances s_n^2. For $i_1 \neq i_2$, according to hypothesis (H13b) we have:

$$|EV_{i_1}V_{i_2}| \leq \frac{C}{(\mathbb{E}K_1(\theta,z))^2}\phi_{\theta,z}(h_K)^{1+\gamma_1} = \mathcal{O}(\phi_{\theta,z}(h_K)^{-1+\gamma_1}),$$

thus,

$$|cov(V_{i_1},V_{i_2})| \leq \frac{C}{(\mathbb{E}K_1(\theta,z))^2}\phi_{\theta,z}(h_K)^{1+\gamma_1} = \mathcal{O}(max\{\phi_{\theta,z}(h_K)^{-1+\gamma_1},1\}). \tag{4.30}$$

On other hand, by using the inequality of covariance for the mixing process (see Proposition A10i of Ferraty and Vieu (2006)) we can write:

$$cov(V_{i_1},V_{i_2}) \leq C\phi_{\theta,z}(h_K)^{-2}\alpha(|i_1 - i_2|). \tag{4.31}$$

Finally, for each positive sequence v_n we can write

$$s_n^2 = \sum_{i=1}^{n} var(V_i) + \sum_{0<|i_1-i_2|\leq v_n} cov(V_{i_1},V_{i_2}) + \sum_{|i_1-i_2|>v_n} cov(V_{i_1},V_{i_2}), \tag{4.32}$$

and by using respectively (4.42), (4.30) and (4.31) to treat each of the three terms of (4.32) we get:

$$\begin{aligned} s_n^2 = \mathcal{O}\left(\frac{n}{\phi_{\theta,z}(h_K)}\right) &+ \mathcal{O}(nv_n max\{\phi_{\theta,z}(h_K)^{-1+\gamma_1},1\}) \\ &+ \mathcal{O}(\phi_{\theta,z}(h_K)^{-2}\sum_{|i_1-i_2|>v_n}\alpha(|i_1-i_2|)). \end{aligned}$$

It suffices now to choose $v_n = \phi_{\theta,z}(h_K)^{-\gamma_1}$ to obtain

$$\begin{aligned} s_n^2 &= \mathcal{O}\left(\frac{n}{\phi_{\theta,z}(h_K)}\right) + \mathcal{O}(\phi_{\theta,z}(h_K)^{-2}n(n-v_n)\alpha(v_n)) \\ &= \mathcal{O}\left(\frac{n}{\phi_{\theta,z}(h_K)}\right) + \mathcal{O}(\phi_{\theta,z}(h_K)^{-2}n^2\phi_{\theta,z}(h_K)^{a\gamma_1}) \\ &= \mathcal{O}\left(\frac{n}{\phi_{\theta,z}(h_K)}\right), \end{aligned} \tag{4.33}$$

the last inequality flows directly from the condition (H1c).

By using the boundaries given by (4.42) and (4.42), and by applying the exponential inequality for mixing bounded variables (Corollary A13ii of Ferraty and Vieu (2006)), we obtain

$$\hat{\varphi}_D(\theta,z) - \mathbb{E}\hat{\varphi}_D(\theta,z) = \mathcal{O}(n^{-1}\sqrt{\log ns_n^2}),\ a.co. \tag{4.34}$$

We get (4.25) directly from (4.33) and (4.34).

⇝ **Proof of (4.26.)** We have, for any $t \in \mathcal{S}_{\mathbb{R}}$:

$$\begin{aligned} \mathbb{E}\widetilde{\varphi}_N(\theta,t,z) &= \frac{1}{h_H \mathbb{E}K_1(\theta,z)}\mathbb{E}\left(K_1(\theta,z)H_1(t)\Delta_1\right) \\ &= \frac{1}{h_H \mathbb{E}K_1(\theta,z)}\mathbb{E}\Big(K_1(\theta,z)1_{B_\theta(z;h_K)}\mathbb{E}\Big(H_1(t)I_{X_1\le C_1}\Big|Z_1\Big)\Big) \\ &= \frac{1}{h_H \mathbb{E}K_1(\theta,z)}\mathbb{E}\left(K_1(\theta,z)E(H_1(t)S_1(\theta,X_1,Z_1)|Z_1)\right). \qquad (4.35) \end{aligned}$$

Furthermore, we have

$$\begin{aligned} \mathbb{E}(H_1(t)S_1(\theta,X_1,z)|Z_1) &= \int H'\left(\frac{t-u}{h_H}\right)S_1(\theta,u,z))f^{Z_1}(u)du \\ &= h_H\int H'(v)\varphi(\theta,t-vh_H,Z_1)dv \\ &= h_H\left(\varphi(\theta,t,z)+o(h_H^{b_2}+h_K^{b_1})\right), \qquad (4.36) \end{aligned}$$

the last equality is arising from the property of Lipschitz function $\varphi(\theta,.,z)$ introduced in (H10) and the fact that H' is a probability density. It should be noted again that because of the condition (H10), the therm $o()$ involved in the result (4.36) is uniform for $t \in \mathcal{S}_{\mathbb{R}}$. Thus, the result (4.26) is an immediate consequence of (4.35) and (4.36).

⇝ **Proof of (4.27)**. The compactness of the set $\mathcal{S}_{\mathbb{R}}$ can be covered by a u_n disjoint intervals as follows:

$$\mathcal{S}_{\mathbb{R}} \subset \cup_{k=1}^{u_n}[\tau_k - l_n, \tau_k + l_n[,$$

where $\tau_1,\ldots,\tau_{u_n}$ are points of $\mathcal{S}_{\mathbb{R}}$ and where l_n and u_n are chosen such that

$$\exists C>0, \exists c>0, l_n = Cu_n^{-1} = n^{-c}. \qquad (4.37)$$

For each $t \in \mathcal{S}_{\mathbb{R}}$, let note τ_t the single τ_k such as $t \in [\tau_k - l_n, \tau_k + l_n[$. Finally, (4.27) can be easily deduced from the following results:

$$\frac{1}{\hat{\varphi}_D(\theta,z)}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\widetilde{\varphi}_N(\theta,t,z)-\widetilde{\varphi}_N(\theta,\tau_x,z)| = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_{\theta,z}(h_K)}}\right), \qquad (4.38)$$

$$\frac{1}{\hat{\varphi}_D(\theta,z)}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\mathbb{E}\widetilde{\varphi}_N(\theta,t,z)-\mathbb{E}\widetilde{\varphi}_N(\theta,\tau_x,z)| = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_{\theta,z}(h_K)}}\right), \qquad (4.39)$$

and

$$\frac{1}{\hat{\varphi}_D(\theta,z)}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\widetilde{\varphi}_N(\theta,\tau_x,z)-\mathbb{E}\widetilde{\varphi}_N(\theta,\tau_x,z)| = \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_{\theta,z}(h_K)}}\right). \qquad (4.40)$$

⇝ **Proof of (4.38**.) Because of the condition (H5), there is exists a finite constant C such that for all $t \in \mathcal{S}_{\mathbb{R}}$:

$$\begin{aligned}|\widetilde{\varphi}_N(\theta,t,z)-\widetilde{\varphi}_N(\theta,\tau_t,z)| &= \frac{1}{nh_H\mathbb{E}K_1(\theta,z)}\sum_{i=1}^{n}\Delta_iK_i(\theta,z)(H_i(t)-H_i(\tau_t))\\ &\leq \frac{C}{nh_H\mathbb{E}K_1(\theta,z)}\sum_{i=1}^{n}K_i(\theta,z)\frac{|t-\tau_t|}{h_H}\\ &\leq C\hat{\varphi}_D(\theta,z)l_nh_H^{-2}. \qquad (4.41)\end{aligned}$$

By using (4.37) and choosing c large enough, we obtain directly (4.38).

⇝ **Proof of (4.39**). This result is obtained directly from (4.25) and (4.41) using Proposition A6ii of Ferraty and Vieu (2006).

⇝ **Proof of (4.40**). Note that we can have:

$$\hat{\varphi}_D(\theta,z)=\frac{1}{n}\sum_{i=1}^{n}W_i,$$

with

$$|W_i| = \mathcal{O}\left(\frac{1}{h_H\phi_{\theta,z}(h)}\right),$$

and

$$\mathbb{E}W_i^2 = \mathcal{O}\left(\frac{1}{h_H\phi_{\theta,z}(h)}\right).$$

Then, by using condition (4.37), we get

$$\begin{aligned}&\mathbb{P}\left(\sup_{x\in\mathcal{S}_{\mathbb{R}}}|\widetilde{\varphi}_N(\theta,\tau_t,z)-\mathbb{E}(\widetilde{\phi}_N(\theta,\tau_t),z)|>\varepsilon\sqrt{\frac{\log n}{nh_H\phi_{\theta,z}(h)}}\right)\\ &\leq n^c\max_{j=1,\dots,u_n}\mathbb{P}\left(|\widetilde{\varphi}_N(\theta,\tau_t,z)-\mathbb{E}(\widetilde{\phi}_N(\theta,\tau_t),z)|>\varepsilon\sqrt{\frac{\log n}{nh_H\phi_{\theta,z}(h)}}\right). \qquad (4.42)\end{aligned}$$

The main step of the demonstration is to get the evaluation of the sum of covariances S_n^2. For $i_1 \neq i_2$, we have:

$$|\mathbb{E}W_{i_1}W_{i_2}|\leq\frac{C}{(h_H^2\mathbb{E}K_1(\theta,z))^2}\phi_{\theta,z}(h_K)^{1+\gamma_1}=\mathcal{O}(h_H^{-2}\phi_{\theta,z}(h_K)^{-1+\gamma_1}),$$

and thus,

$$|cov(W_{i_1}, W_{i_2})| \leq \frac{C}{(h_H^2 \mathbb{E}K_1(\theta, z))^2} \phi_{\theta,z}(h_K)^{1+\gamma_1} = \mathcal{O}(h_H^{-2} max\{\phi_{\theta,z}(h_K)^{-1+\gamma_1}, 1\}). \quad (4.43)$$

On another side, by using the covariance inequality for the mixing process (Proposition A10i of Ferraty and Vieu (2006)) we can write:

$$cov(W_{i_1}, W_{i_2}) \leq C h_H^{-2} \phi_{\theta,z}(h_K)^{-2} \alpha(|i_1 - i_2|). \quad (4.44)$$

Finally, for any positive sequence v_n we can write

$$S_n^2 = \sum_{i=1}^{n} var(W_i) + \sum_{0<|i_1-i_2|\leq v_n} cov(W_{i_1}, W_{i_2}) + \sum_{|i_1-i_2|>v_n} cov(W_{i_1}, W_{i_2}), \quad (4.45)$$

and using respectively (4.42), (4.43) and (4.44) to treat the three terms (4.45) we get:

$$\begin{aligned} S_n^2 = \mathcal{O}\left(\frac{n}{h_H \phi_{\theta,z}(h_K)}\right) &+ \mathcal{O}(n v_n h_H^{-2} max\{\phi_{\theta,z}(h_K)^{-1+\gamma_1}, 1\}) \\ &+ \mathcal{O}(h_H^{-2} \phi_{\theta,z}(h_K)^{-2} \sum_{|i_1-i_2|>v_n} \alpha(|i_1 - i_2|)). \end{aligned}$$

Now, we have just to choose $v_n = \phi_{\theta,z}(h_K)^{-\gamma_1}$ to get

$$S_n^2 = \mathcal{O}\left(\frac{n}{h_H \phi_{\theta,z}(h_K)}\right), \quad (4.46)$$

the last inequality flows directly from the condition (H1c).

By using (4.42), (4.42), and (4.42), and applying the exponential inequality for mixing bounded variables (for instance the Corollary A13ii of Ferraty and Vieu (2006)), we obtain

$$\widetilde{\varphi}_N(\theta, \tau_j, z) - \mathbb{E}\widetilde{\varphi}_N(\theta, \tau_j, z) = \mathcal{O}(n^{-1}\sqrt{\log n S_n^2}), \ a.co. \quad (4.47)$$

the result (4.25) flows directly from (4.46) and (4.47).

Proof of Lemma 4.4.2. The proof can be completed along the same line as that of Lemma 4.2(ii)of Bouchentouf et al. (2015).

Proof of Lemma 4.4.3. Let $H_i^{(1)}(t) = H^{(1)}\left(h_H^{-1}(t - T_i)\right)$, note that

$$\mathbb{E}\widetilde{\varphi}_N(\theta, t, z) - \varphi(\theta, t, z) = \frac{1}{h_H \, \mathbb{E}K_1(z, \theta)} \mathbb{E}\left(K_1(z, \theta)\left[\mathbb{E}\left(H_1^{(1)}(t)| < Z, \theta >\right) - h_H \varphi(\theta, t, z)\right]\right). \quad (4.48)$$

Condition (H8) allows us to write:

$$\left|\mathbb{E}\Big(H_1^{(1)}(t)| < Z,\theta >\Big) - h_H\varphi(\theta,t,z)\right| \le h_H \int_{\mathbb{R}} H^{(1)}(t)\,|\varphi(\theta,t-h_H t,Z) - \varphi(\theta,t,z)|\,dt.$$

Finally, (A5) allows us to write

$$\left|\mathbb{E}\Big(H_1^{(1)}(t)| < Z,\theta >\Big) - h_H\varphi(\theta,t,z)\right| \le C_{\theta,z} h_H \int_{\mathbb{R}} H^{(1)}(t)\Big(h_K^{b_1} + |t|^{b_2} h_H^{b_2}\Big)\,dt. \tag{4.49}$$

This inequality is uniform on $(\theta,t,z) \in \Theta_{\mathcal{F}} \times \mathcal{S}_{\mathcal{F}} \times \mathcal{S}_{\mathbb{R}}$, now to finish the proof it is sufficient to use (H5). □

Proof of Lemma 4.4.4. Let $\mathcal{S}_{\mathcal{F}} \in \bigcup_{k=1}^{w_n} (t_j - l_n, t_j + l_n)$ with $l_n = n^{-\frac{3}{2}\gamma - \frac{1}{2}}$ and $w_n \le C n^{-\frac{3}{2}\gamma - \frac{1}{2}}$. Taking $j(t) = \arg\min_{\{1\ldots w_n\}} |t - s_j|$. Consider the following decomposition

$$\begin{aligned}
|\widetilde{\varphi}_N(\theta,t,z) - \mathbb{E}(\widetilde{\varphi}_N(\theta,t,z))| = \;& \underbrace{\left|\widetilde{\varphi}_N(\theta,t,z) - \widetilde{\varphi}_N(\theta,t,z_{k(z)})\right|}_{T_1} \\
& + \underbrace{\left|\widetilde{\varphi}_N(\theta,t,z_{k(z)}) - \mathbb{E}\Big(\widetilde{\varphi}_N(\theta,t,z_{k(z)})\Big)\right|}_{T_2} \\
& + 2\underbrace{\left|\widetilde{\varphi}_N(s_{j(\theta)},t,z_{k(z)}) - \widetilde{\varphi}_N(s_{j(\theta)},t_{j(t)},z_{k(z)})\right|}_{T_3} \\
& + 2\underbrace{\left|\mathbb{E}\Big(\widetilde{\varphi}_N(s_{j(\theta)},t,z_{k(z)})\Big) - \mathbb{E}\Big(\widetilde{\varphi}_N(s_{j(\theta)},t_{j(t)},z_{k(z)})\Big)\right|}_{T_4} \\
& + \underbrace{\left|\mathbb{E}\Big(\widetilde{\varphi}_N(\theta,t,z_{k(z)})\Big) - \mathbb{E}(\widetilde{\varphi}_N(\theta,t,z))\right|}_{T_5}
\end{aligned}$$

⇝ Concerning T_1. We use the Hölder continuity condition on K, the Cauchy-Schwartz's inequality and the Bernstein's inequality. With theses arguments we get

$$T_1 = \mathcal{O}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n h_H \phi(h_K)}}\right).$$

Then using the fact that $T_5 \le T_1$, we obtain

$$T_5 = \mathcal{O}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n h_H \phi(h_K)}}\right). \tag{4.50}$$

⇝ For T_2, we follow the same idea given for Γ_2, we get

$$T_2 = \mathcal{O}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{H}} + \log d_n^{\Theta_\mathcal{H}}}{nh_H\phi(h_K)}}\right)$$

⇝ Concerning T_3 and T_4. Using Lipschitz's condition on the kernel H,

$$\left|\widetilde{\varphi}_N(s_{j(\theta)}, t, z_{k(z)}) - \widetilde{\varphi}_N(s_{j(\theta)}, t_{j(t)}, z_{k(z)})\right| \leq \frac{l_n}{h_H^2\phi(h_k)},$$

using the fact that $\lim_{n\to\infty} n^\gamma h_H = \infty$ and choosing $l_n = n^{-\frac{3}{2}\gamma-\frac{1}{2}}$ implies

$$\frac{l_n}{h_H^2\phi(h_k)} = o\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{H}} + \log d_n^{\Theta_\mathcal{H}}}{nh_H\phi(h_K)}}\right).$$

So, for n large enough, we have

$$T_3 = \mathcal{O}_{a.co}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{H}} + \log d_n^{T_\mathcal{H}}}{nh_H\phi(h_K)}}\right).$$

And as $T_4 \leq T_3$, we obtain

$$T_4 = \mathcal{O}_{a.co}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{H}} + \log d_n^{\Theta_\mathcal{H}}}{nh_H\phi(h_K)}}\right). \tag{4.51}$$

Finally, the lemma can be easily deduced from (4.50) and (4.51).

Proof of Lemma4.4.5.

The proof of (4.20) is based on some results depending on the following decomposition;

$$\begin{aligned}\hat{L}(\theta,t,z) - L(\theta,t,z) &= \frac{1}{\hat{\varphi}_D(\theta,z)}\left\{\left(\hat{L}_N(\theta,t,z) - \mathbb{E}\hat{L}_N(\theta,t,z)\right) - \left(L(\theta,t,z) - \mathbb{E}\hat{L}_N(\theta,t,z)\right)\right\} \\ &\quad + \frac{L(\theta,t,z)}{\hat{\varphi}_D(\theta,z)}\{1 - \hat{\varphi}_D(\theta,z)\}.\end{aligned} \tag{4.52}$$

Then, the rest of the proof is deduced directly from Corollary 4.9 and Lemma 4.4.2.

The proof of these points are similar to ones given in Bouchentouf et al. (2015), so it is sufficient to replace $\hat{F}_D(\theta,z)$, $F(\theta,t,z)$ and $\mathbb{E}(\hat{F}_N(\theta,t,z))$ (Lemma 6, Corollary 3 and Lemma 7) by $\hat{L}_D(\theta,z)$, $L(\theta,t,z)$, and $\mathbb{E}(\hat{L}_N(\theta,t,z))$ respectively.

Concerning(4.21), we consider the following decomposition:

$$\begin{aligned}\widetilde{\varphi}(\theta,t,z)-\varphi(\theta,t,z) &= \frac{1}{\hat{\varphi}_D(\theta,z)}(\widetilde{\varphi}_N(\theta,t,z)-\mathbb{E}(\widetilde{\varphi}_N(\theta,t,z)) \\ &\quad -\frac{1}{\hat{\varphi}_D(\theta,z)}(\varphi(\theta,t,z)-\mathbb{E}\widetilde{\varphi}_N(\theta,t,z)) \\ &\quad +\frac{\varphi(\theta,t,z)}{\hat{\varphi}_D(\theta,z)}(1-\hat{\varphi}_D(\theta,z))\end{aligned}$$

The rest of the proof is deduced directly from Lemma 4.4.2, Lemma 4.4.3, Lemma 4.4.4 and Corollary 4.9.

Conclusion and future work

In this thesis, we consider the problem of conditional hazard function estimation. Although extensive research has been done on this topic in the recent decades, we believe that combining different features including functional, censorship, and single index can find its has not been addressed, yet. In what follows, we first reiterate the main conclusions of this thesis, then we propose some possible extensions that can contribute to the literature on subject.

In Chapter 1, a comprehensive review of the recent literature on functional data analysis, conditional models, survival models, and single index models was given.

In Chapter 2, we provided some definitions and tools utilized for our research.

In chapter 3, we established the consistency properties, with rates, of the conditional hazard function in the single functional index model for dependant functional data under random censorship; the pointwise almost complete and the uniform almost complete convergence (with rates) of the kernel estimate of this model are obtained.

The study established in this thesis offers different perspectives, let us cite for instance:

– Asymptotic normality of the model studied.

– It is quite possible to generalize our results to the case of ergodic data as well as spatial data.

– Recursive estimation of the considered model.

– Recursive estimation of the considered model for truncated data.

– Recursive estimation of the considered model for ergodic observations.

Bibliography

Ahmad, I. A. (1976). Uniform strong convergence of the generalized failure rate estimate, *Bull. Math. Statist.* **17**, 77-84.

Aït Saidi, A., Ferraty, F., Kassa, R. (2005). Single functional index model for a time series. *R. Roumaine Math. Pures et Appl.* **50**, 321-330.

Aït Saidi, A., Ferraty, F., Kassa, R., Vieu, P. (2008). Cross-validated estimation in the single functional index model. *Statistics*. **42**, 475-494.

Aneiros, G., Cao, R., Fraiman, R., Genest, C., and Vieu, V. (2019). Recent advances in Functional Data Analysis and High-Dimensional Statistics,*Journal of Multivariate Analysis*. **170**, 3-9.

Attaoui, S. (2014). On the Nonparametric Conditional Density and Mode estimates in the single functional index model with strongly mixing data, *Sankhya A*. **76**, 356-378.

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

Attaoui,S, Nengxiang Ling (2015). Asymptotic results of a nonparametric conditional cumulative distribution estimator in the single functional index modeling for time series data with applications, *Metrika* .00184-015-0564-6.

Attaoui, S., and Ling, N. (2016), Asymptotic Results of a Nonparametric Conditional Cumulative Distribution Estimator in the Single Functional Index Modeling for Time Series Data with Applications,*Metrika*. **79**, 485-511.

Azzeddine, N., Laksaci, A., Ould-Saïd, E. (2008). On the robust nonparametric regression estimation for functional regressor, *Statist. Probab.Lett.* **78**, 3216-3221.

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

Bapna R, Jank W, Shmueli G(2008). Price formation and its dynamics in online auctions, *Decis Support Syst*, **44**,641-656.

Benaissa et Mechab (2015).recursive estimate of the conditional hazard function

Benaissa and al (2013) mean square error of the estimator of the conditional hazard function.*Applicationes Mathematicae*.4,405-420.

Benhenni, K., Ferraty, F., Rachdi, M., Vieu, P. (2007). Locally smoothing regression with functional data. *Computat. Statist.* **22**, 353-370.

Bensmail H, Aruna B, SemmesOJ, Haoudi A(2005). Functional clustering algorithm for high-dimensional proteomics data. *J Biomed Biotechno*.**2**, 80-86.

Besse, P., Cardot, H., Stephenson, D. (2000). Autoregressive forecasting of some functional climatic variations, *Scand. J. Stat.* **27**, 673-687.

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

Bosq, D., (1991). Modelization, non-parametric estimation and prediction for continuous time processes. In nonparametric functional estimation and Related topics (1990), 509-529, NATO, Adv. Sci. Inst. Ser. C Math. Phys. Sci. **335**, Kluwer Acad. Publ., Dordrecht.

Bosq, D. Delecroix, M. (1985). Noparametric prediction of a Hilbert space valued random variable. *Stochastic Processes Appl*.**19**, 271-280.

Bosq, D., Lecoutre, J. P.(1987). Théorie de l'estimation fonctionnelle.*Economica, Paris* .

Bouchentouf, A. A., Djebbouri, T., Rabhi, A. and Sabri, K. (2014). Strong uniform consistency rates of some characteristics of the conditional distribution estimator in the functional single-index model, *Appl. Math (Warsaw)*. **41**, 301-322.

Bouchentouf, A. A., Hamel, E.H., Rabhi, A. and Soltani, S. (2014). Nonparametric estimation of hazard function with functional explicatory variable in single functional index, *Journal of Applied Mathematics and Statistics* . **1**, 20-41.

Bouchentouf, A. A., Hamza, A., 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**(3), 82-101.

Bouchentouf, A.A., Mekkaoui, S. and Rabhi, A. (2015). Strong uniform consistency rates of conditional quantiles for time series data in the single functional index model, *New*

Trends in Mathematical Sciences (NTMSCI). **3**, 181-198.

Burba, F., Ferraty, F., Vieu, P. (2008). Convergence de l'estimateur à noyau des k plus proches voisns en régression fonctionnelle non-paramétrique. *C. R. Acad. Sci . Paris*. **346**, 339-342.

Cardot, H., Crambes, C., Sarda, P. (2004). Spline estimation of conditional quantiles for functional covariates. *C. R. Acad. Sci.Paris*.**339**, 141-144.

Cardot, H., Ferraty, F., Mas, A., Sarda, P. (2003). Testing hypotheses in the functional linear model. *Scand. J. Stat*, **30**, 241-255.

Cardot, H., Ferraty, F., Sarda, P. (1999). Functional linear model. *Stat. Probab. Lett*.**45**, 11-22.

Crambes, C., Delsol, L., Laksaci, A. (2008). Lp errors for robust estimators in functional nonparametric regression. *J. of Nonparametric Statistics*. **20**, 573-598.

Cuevas, A., Febrero, M., Fraiman, R. (2004). An anova test for functional data. *Computational Statistics & Data Analysis*.**47**, 111-122.

Cui, X., Härdle, W.K. and Zhu, L. (2011). The EFM approach for single-index models, *Annals of Statistics*, **39** , 1658-1688.

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

Daoudi, h. and al (2019). asymptotic normality of the nonparameric conditional density fonction estimate with variable quasi-associad data. **3**, 0973-7022.

Daoudi, h. and al(2020) . asymptotic normality of a conditional hazad fonction estimate in the single index for quasi-associad data.*Commun. Statist. Theo. Meth*. **3**,513-530.

Delecroix, M, Härdle, W., Hristache, M. (1999). M-estimateurs semiparamétriques dans les modèles à direction révélatrice unique. *Bull. Belg. Math. Soc*, **6**. 161-185.

Delecroix, M, Härdle, W. Hristache, M. (2003). Efficient estimation in conditional single-index regression. *J. Multivariate Anal* . **86**, 213-226.

Delicado, P. (2007). Functional k-sample problem when data are density functions, Computational Statistics & Data Analysis. **22**, 391-440.

Delsol, L. (2007). Régression non paramétrique fonctionnelle: expression asymptotique des moments, *Ann. I.S.U.P*. **LI**, 43-67.

Delsol, L. (2009). Advances on asymptotic normality in nonparametric functional Time

Series Analysis Statistics. *Statistics*. **43**, 13-33.

Delsol, L. (2011). Nonparametric methods for α-mixing functional random variables.*In the oxford handbook of functional data analysis* .

Demongeot, J., Laksaci, A., Rachdi, M. and Rahmani, S. (2014). On the local linear modelization of the conditional distribution forfunctional data. *Sankhya*. **76**, 328- 355.

Dabo-niang S, Vieu P(2007). On the using of modal curves for radar waveforms classification. *Comput Statist Data Anal.* ,**51**,48-78.

Erbas B, Akram M, Gertig DM, English D, Hopper JL, Kavanagh AM, Hyndman RJ(2010). Using functional data analysis models to estimate future time trends in age-specific breast cancer mortality for the United States and England-Wales. *J Epidemiol*.**20**,159165.

Ferraty, F. (2010). Special issue on statistical methods and problems in infinite dimensional spaces. *J. Multivariate Anal.* **101**, 305-490.

Ferraty, F., Laksaci, A. and Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models, *Stat. Inf. Stochastic Process.* **9**, 47-76.

Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables,*J. Statist. Plann. and Infer*. **140**, 335-352.

Ferraty, F., Mas, A., Vieu, P. (2007). Advances in nonparametric regression for functional variables. A*Aust. and New Zeal. J. of Statist*.**49**, 1-20.

Ferraty, F. Peuch, A. and Vieu, P. (2003). Modèle à indice fonctionnel simple, *C. R. Acad. Sci.Paris*. **336**, 1025-1028.

Ferraty, F., Rabhi, A. and Vieu, P. (2005). Conditional Quantiles for Functionally Dependent Data with Application to the Climatic El Niño Phenomeno, *Sankhyā: The Indian Journal of Statistics, Special Issue on Quantile Regression and Related Methods*.**67** , 399-417.

Ferraty, F., Rabhi, A., Vieu, P. (2008). Estimation non paramétrique de la fonction de hasard avec variable explicative fonctionnelle. *Rom. J. Pure and Applied Math.* **52**, 1-18.

Ferraty, F., Romain, Y. (2011). The Oxford Handbook of Functional Data Analysis. Oxford University Press.

Ferraty, F., Vieu, P. (2000). Dimension fractale et estimation de la régression dans des espaces vectoriels semi-normés. *C. R. Acad. Sci.Paris*. **330**, No.2, 139-142.

Ferraty, F., Vieu, P. (2002). The functional nonparametric model and application to spectrometric data. *Computational Statistics and Data Analysis*. **17**, 545-564.

Ferraty, F. and Vieu, P. (2003). Curves discrimination : a nonparametric functional approach. Special issue in honour of Stan Azen: a birthday celebration. *Computational Statistics and Data Analysis*. **44**, 161-173.

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

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

Ferraty, F., Vieu, P. (2011). Kernel regression estimation for functional data. In the Oxford Handbook of Functional Data Analysis (Ed. F. Ferraty and Y. Romain). *Oxford University Press*.

Ferraty, F., Alejandro Quintela, Vieu, P.(2011). Specification test for conditional distribution with functional data Published online by Cambridge University Press.

Harrison A, Ryan W, Hayes K(2007). Functional data analysis of joint coordination in the development of vertical jump performance. *Sports Biomech* . **6**,199-214.

Hall, P. (1989). On projection pursuit regression. *Ann. Statist*. **17**, 573-588.

Hall, P., Heckman, N. E. (2002). Estimating and depicting the structure of a distribution of random functions. *Biometrika*. **89**, 145-158.

H"ardle, W. (1990). Applied nonparametric regression. Cambridge Univ. Press, UK.

Härdle, W., Hall, P., Ichumira, H., (1993). Optimal smoothing in single-index models, *Ann. Statist*.**21**.

Härdle, W. and Marron, J.S. (1985). Optimal bandwidth selection in nonparametric regression function estimation, *Ann. Statist*. **13**, 1465-1481.

Härdle, W., Marron, J. S. (1990). Semiparametric comparison of regression curves. *The Annals of Statistics*. **18**, 59-63.

Hastie, T., Buja, A., Tibshirani, R. (1995). Penalized discriminant analysis. *Ann. Stat*.**23**, 73-102.

Hristache, M., Juditsky, A., Spokoiny, V. (2001). Direct estimation of the index coefficient in the single-index model. *Ann. Statist*. **29**, 595-623.

Huber, P. J. (1985). Projection pursuit. *Ann. Statist*. **13**, 435-475.

Hyndman RJ, ShangHL(2010). Rainbow plots, bagplots, and boxplots for functional Data. *J Comput Graph Stat*. **19**,29-45.

Hermanussen M. Auxology(2010). An update. Horm Res Paediatr ,**74**,153-164.

Hall, Peter and Wolff, Rodney C. L. and Yao, Qiwei (1999). Methods for estimating a conditional distribution function. *Journal of the American Statistical Association*. **94**, 154-163. **H**utchinson RA, McLellan PJ, Ramsay JO, Sulieman H, Bacon DW (2004). Investigating the impact of operating parameters on molecular weight distributions using functional regression. Macromol Symp , **206**,495-508

Hyndman RJ., Ullah, MS (2007). Robust forecasting of mortality and fertility rates: a functional data approach, *Comput Stat Data Anal*, **51**, 4942–4956.

Hyndman RJ, Shang HL(2010). Rainbow plots, bagplots, and boxplots for functional Data. J Comput Graph Stat , **19**,29-45.

Hermanussen M: Auxology(2010). An update. Horm Res Paediatr , **74**,153-164.

Hyndman R, Booth H(2008). Stochastic population forecasts using functional data models for mortality, fertility and migration. *Int J Forecast* , **24**,323-342.

Illian JB, Prosser JI, Baker KL, Rangel-Castro JL(2009). Functional principal component data analysis: a new method for analysing microbial community fingerprints. *J Microbiol Methods* . **79**,89-95.

Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *Journal of Econometrics*. **58**, 71-120.

Ikeda T, Dowd M, Martin JL(2008). Application of functional data analysis to investigate seasonal progression with interannual variability in plankton abundance in the bay of Fundy,*Canada. Estuar Coast Shelf Sci* . **24**,445-455.

Guo W(2004). Functional data analysis in longitudinal settings using smoothing splines. *Stat Methods Med Res*. **13**,49-62.

Gannoun, A., Saracco, J., and Yu, K. (2003). Nonparametric prediction by conditional median and quantiles. *J. Statist. Plann. Infer*. **117**, 207-223.

Gasser, T., Hall, P., Presnell, B. (1998). Nonparametric estimation of the mode of a distribution of random curves. *J. R. Stat. Soc., Ser. B, Stat. Methodol*. **60**, 681-691.

Geenens, G. (2011). Curse of dimensionality and related issues in nonparametric functional regression, *Statistics Surveys*. **05**, 30-43.
Gouttard S, Prastawa M, Bullitt E, Lin W, Gerig G(2009). Constrained data decomposition and regression for analyzing healthy aging from fiber tract diffusion properties. *Med Image Comput Comput Assist Interv*. **12**,321-328.
Goia, A., and Vieu, P. (2015), A Partitioned Single Functional Index Model, *Computational Statistics*. **30**, 673-692.
Kolmogorov, A. N., Tikhomirov, V .M. (1959). ϵ-entropy and ϵ-capacity. Uspekhi Mat. Nauk **14**, 3-86.
Kuelbs, J., Li, W. (1993). Metric entropy and the small ball problem for Gaussian measures. *J. Funct. Anal.* **116**, 133-157.
Koenig LL, Lucero JC, Perlman E (2008). Speech production variability in fricatives of children and adults. *J Acoust Soc Am* . **124**,3158-3170.
Kim SB, Rattakorn P, Peng YB (2010). An effective clustering procedure of neuronal response profiles in graded thermal stimulation. *Expert Syst Appl*. **37**,5818-5826.
Laksaci, A., Mechab, M. (2010). Estimation non parametrique de la fonction de hasard avec variable explicative fonctionnelle cas des donnees spatiales. *Rev. Roumaine, Math. Pures Appl*. **55**, 35-51.
Laksaci, A., Mechab, M. (2014).condtional hazard estimate for fonctinal random fields.*Journal of Statistical Theory and Practrice*.**8**,192-220.
Laksaci, A., Rachdi, M. and Rahmani, S., (2013). Spatial modelization : local linear estimation of the conditional distribution for functional data. *Spat. Statist*. **6**, 1-23.
Laksaci, A., Madani, F. and Rachdi, M. (2013). Kernel conditional density estimation when the regressor is valued in a semi metric space. *Commun. in Statit.-Theory and Meth*. **42**, 3544-3570.
Leao, D., Fragoso, M., Ruffino, P.(2004). Regular conditional probability, integration of probability and Radon spaces. *Proyecciones*. **23**, 15-29.
Ling, N., and Vieu, P. (2018). Nonparametric modelling for functional data: Selected Survey and Tracks for Future, *Statistics*. **52**, 934-949.
Ling, N.X., and Xu, Q. (2012), Asymptotic Normality of Conditional Density Estimation in the Single Index Model for Functional Time Series Data, *Statistics & Probability Letters*.

82, 2235-2243.

Lee S, Byrd D, KrivokapicÌ J(2006). Functional data analysis of prosodic effects on articulatory timing.*J Acoust Soc Am*. **119**.1666-1671.

Laukaitis A (2005). Functional data analysis for clients segmentation tasks. *Eur J Oper Res*.**163**,210-216.

Laukaitis A (2008). Functional data analysis for cash flow and transactions intensity continuous-time prediction using Hilbert-valued autoregressive processes. *Eur J Oper Res*. **185**,1607-1614.

Lucero JC(2005). Comparison of measures of variability of speech movement trajectories using synthetic records. *J Speech Lang Hear Res*. **48**,336-344.

Mahiddine, A., Bouchentouf, A. and Rabhi, A. (2014). Nonparametric estimation of some characteristics of the conditional distribution in single functional index model. *Malaya J. Math*. **2**, 392-410.

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

Merouan and al(2018).Quadratic error of the conditional hazard fonction in the local linear estimation for fonctional data.*Africa Statistika*.**13**,1759-1777

Massim, I. and Mechab, B. (2016). Local linear estimation of the conditional hazard function, *International Journal of Statistics & Economics*. **17**, 1-11.

Meiring W(2007). Oscillations and time trends in stratospheric ozone levels. *J Am Stat Assoc*. **102**,788-802.

Meyer PM, Zeger SL, Harlow SD, Sowers M, Crawford S, Luborsky JL, Janssen I, McConnell DS, Randolph JF, Weiss G (2007). Characterizing daily urinary hormone profiles for women at midlife using functional data analysis. *Am J Epidemiol*. **165**,936-945.

Muller HG, Wu S, Diamantidis AD, Papadopoulos NT, Carey JR(2009). Reproduction is adapted to survival characteristics across geographically isolated medfly populations. *Proc Biol Sci*. **276**,4409-4416.

Ould-Saïd, E.,Cai, Z. (2005). Strong uniform consistency of nonparametrice stimation of the censored conditional mode function. *J. Nonparam. Statist*. **17**, 797-806.

Ould-Saïd, E. (2006). A strong uniform convergence rate of kernel conditional quantile estimator under random censorship. *Statist. Probab. Lett.* **76**, 579-586.

Parsen ,E.(1962). on estimation of a probability density function mode, *Ann. Math. Statist.* **33**,1065-1076.

Quintela-del-Rio, A. (2006). Nonparametric estimation of the maximum hazard under dependence conditions, *Statist. Probab. Lett.* **76**, 1117-1124.

Quintela-del-Rio, A. (2007). Plug-in bandwidth selection in kernel hazard estimation from dependent data, *Comput. Stat. Data Anal.* **51**, 5800-5812.

Quintela-del-Rio, A. (2008). Hazard function given a functional variable: Non-parametric estimation under strong mixing conditions, *J. Nonparametr. Stat.* **20**, 413-430.

Quintela-del-Rio, A. (2010). On non-parametric techniques for area-characteristic seismic hazard parameters. *Geophys. J. Int.* **180**, 339-346.

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

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

Rachdi, M., Vieu, P. (2007). Nonparametric regression for functional data: automatic smoothing parameter selection, *J. Stat. Plan. Infer.* **137**, 2784-2801.

Rafael Izbichi andAnn B Lee(2016) nonparametric conditional density estimation in a high-dimensional regression setting.*Journal of Computational and Graphical Statistcs*.**25**,1297-1316.

Ramsay, J.O. (2000). Functional components of variation in handwriting. *Journal of the American Statistical Association*. **95**, 9-15.

Ramsay, J. O., Silverman, B. W. (2002). Applied functional data analysis. *Methods and case studies Spinger-Verlag, New York* .

Rosenblatt, M. (1969). Conditional probability density and regression estimators. *In multivariate analysis II, Ed. P.R. Krishnaiah. Academic Press, New York and London* .

Roislien J, Skare O, Gustavsen M, Broch NL, Rennie L, Opheim A(2009). Simultaneous estimation of effects of gender, age and walking speed on kinematic gait data. *Gait*

Posture. **30**,441-445.

Sang, P., and Cao, J. (2020), Functional Single-Index Quantile Regression Models, *Statistics and Computing*. **30**, 771-781.

Sarda, P., Vieu, P. (2000). Kernel Regression. Smoothing and Regression: Approaches, Computation, and application. Ed. M.G. schimek, *Wiley Series in Probability and Statistics*.43-70,

Sarda, P., Vieu, P. (1989). Empirical distribution function for mixing random variables, Application in nonparametric hazard estimation. *Statistics*. **20**, No.4, 559-571.

Sarda, P. and Vieu, Ph. (1991). Smoothing parameter selection in hazard estimation, *Statist. Proba. Let*. **11**, 429-434.

Shang, H.-L. (2018). Estimation of a functional single index model with dependenterrors and unknown error density, *Communications in Statistics-Simulation and Computation*. 1-23.

Singpurwalla, N. D. and Wong, M. Y. (1983). Estimation of the failure rate a survey of non-parametric methods. Non-Bayesian methods, *Commun. Stat. Theory Math*. **12**, 559-588.

Song JJ, Lee H, Morris JS, Kang S(2007). Clustering of time-course gene expression data using functional data analysis. *Comp Bio Chem* .**31**,265-274.

Tanner, M., Wong, W. H. (1983). The estimation of the hazard function from randomly censored data by the kernel methods. *Ann. Statist*.**11**, 989-993.

Theodoros, N., Yannis G. Y. (1997). Rates of convergence of estimate, Kolmogorov entropy and the dimensioanlity reduction principle in regression. *The Annals of Statistics*. **25**, 2493-2511.

Tabti, H., Ait Saidi, A. (2018). Estimation and simulation of conditional hazard function in the quasi-associated framework when the observations are linked via a functional single-index structure, *Journal Communications in Statistics-Theory and Methods*. **47**, 816-838.

Torres JM, Garcia Nieto PJ, Alejano L, Reyes AN(2010). Detection of outliers in gas emissions from urban areas using functional data analysis, *J Hazard Mater***186**,144-149.

Ullah.S and C.F Finch,(2013). Applications of functional data analysis: A systematic review,*BMC Medical Research Methodology* . 13-43

Ullah S., Finch, CF. (2010). Functional data modelling approach for analysing and predicting trends in incidence rates-An application to falls injury, *Osteoporos Int*, **21**. 2125-2134.

Watson, G. S. (1964). Smooth regression analysis. *Sankhyã Ser. A*. **26**, 359-372.

Watson, G. S. and Leadbetter, M. R. (1964). Hazard analysis I, *Biometrika*. **51**, 175-184.

Wu P, Müller (2010) H. Functional embedding for the classification of gene expression profiles.*Bioinformatics*. **26**,509-517.

West RM, Harris K, Gilthorpe MS, Tolman C, Will EJ(2007).
Functional data analysis applied to a randomized controlled clinical trial in hemodialysis patients describes the variability of patient responses in the control of renal anemia. *J Am Soc Nephrol* , **18**,2371-2376.

Xianzhu Xiong and Meijuan,(2020). Non parametric estimation of the conditional density function with right-censored and dependent data.

Youndjé, É.,Sarda,P., Vieu, P. (1996). Optimal smooth hazard estimates, *Test*.**5**, 379-394.

Zhu HT, Styner M, Tang NS, Liu ZX, Lin WL, Gilmore JH(2010). Functional regression analysis of DTI tract statistics. IEEE Trans Med Imaging , **29**,1039-1049.

Zhang Y, Muller HG, Carey JR, Papadopoulos NT(2006). Behavioral trajectories as predictors in event history analysis: male calling behavior forecasts medfly longevity. *Mech Ageing Dev* . **127**,680-686.