Dédicace

Du fond du coeur, à ceux que j'ai le plus chers, à mon père et ma mère.

Aussi éphémère que les rèves et les ombres,
aussi capricieuse que la rosée et l'éclaire,
telle est la vie.
Proverb. La Chine

Remerciements

Tout d'abord je tiens à remercier mon Dieu de m'avoir donné le rêve pour que ce travail existe, le courage pour le commencer, la patience pour le terminer, l'amour par ceux qui m'ont aidé d'une manière ou d'une autre pour mener à bien ce modeste travail, et enfin la mémoire pour le remercier.

Par ce modeste travail, j'adresse un énorme merci à mes parents, leurs existence m'a fournis mon existence, ma continuation, ma convergence vers les belles limites. Les mots me manquaent vraiment pour leurs dire à quel point je leurs suis reconnaissant et combien je tiens à leurs.

Du coeur au coeur je remercie mes frères et soeurs pour leur existence, vraiment leur existence m'a fait: …….. tous.

Je tiens à exprimer mon profond sentiment de respect et de reconnaissance à mon directeur et codirecteur de thèse les Professeurs ***Abbès Rabhi*** *et* ***Abderrahmane Yousfate*** *pour leur encadrement et leurs encouragements durant toute la période de la réalisation de ce travail. Je témoigne ici ma profonde gratitude et remerciements au professeur* ***Abbès Rabhi*** *dont j'ai profité de sa présence et son aide tout au long de préparation de cette thèse. Sa disponibilité et sa présence de chaque instant de recherche m'ont fait preuve de beaucoup compréhension et illustration dans les moments nécessaires, il n'a jamais cessé de me soutenir et de m'encourager.*

Et sans oublier de remercier l'ensemble de mes professeurs qui m'ont accompagnés durant toutes mes études: recherche et enseignemant. En particulier le Professeur ***Abderrahmane Yousfate*** *de son enseignement et sa richesse de ses conseils d'une façon continue, je le remercie infiniment. Je tiens aussi à exprimer mes remerciements particuliers à Mademoiselle* ***Amina Angelika Bouchentouf*** *et* ***Khadidja Nedjadi*** *de m'avoir accorder ses expériences, patience, gentillesses et ses conseils, elles m'ont vraiment encouragé.*

Mes remerciements vont aussi aux membres de jury pour avoir accepté la charge d'examiner et d'évaluer cette thèse.

Je remercie sincèrement le professeur ***Samir Bennaissa*** *pour l'honneur qu'il me fait en présidant ce jury.*
Je remercie vivement le Docteur ***Abdeldjabbar Kandouci*** *pour la confiance dont il me fait preuve en faisant parties de ce jury.*
Je voudrais aussi remercier le Docteur ***Fethi Madani*** *pour l'intérêt qu'il a bien voulu accorder à mes travaux en acceptant de participer au jury.*

Ma gratitude va au doctresse ***Malika DALI-Youcef*** *d'avoir accepté de rapporter sur mes travaux, je suis très reconnaissant pour sa présence à mon jury.*

Et tous ceux qui j'ai l'oublier de remercier.

RÉSUMÈ

Dans cette thèse, nous nous proposons d'étudier quelques paramètres fonctionnels premièrement, nous nous proposons d'étudier le problème de la modélisation non paramétrique lorsque les variables statistiques sont des courbes. Plus précisément, nous nous intéressons à des problèmes de prévisions à partir d'une variable explicative à valeurs dans un espace de dimension infinie (espace fonctionnel) et nous cherchons à développer des alternatives à la méthode de régression. Le maximum ou encore le point à haut risque d'une fonction de risque conditionnel est un paramètre d'un grand intérêt en statistique, notamment dans l'analyse de risque séismique, car il constitue le risque maximal de survenance d'un tremblement de terre dans un intervalle de temps donné. Au moyen d'estimations non paramétriques basées sur les techniques de noyau de convolution de la première dérivée de la fonction de hasard conditionnel, nous établissons le comportement asymptotique d'un taux de hasard d'une variable explicative fonctionnelle ainsi que la normalité asymptotique de la valeur maximale pour un processus indépendant.

Deuxièment lorsque les données sont générées à partir d'un modèle de régression à indice simple. Nous étudions deux paramètres fonctionnels.

Dans un premier temps nous supposons que la variable explicative est à valeurs dans un espace de Hilbert (dimension infinie) et nous considérons l'estimation de la distribution conditionnelle ainsi que les déivées successives de la densité conditionnelle par la méthode de noyau. Nous traitons les propriétés asymptotiques de cet estimateur dans le cas indépendant. Dans notre cas où les observations sont indépendantes identiquement distribuées (i.i.d), nous obtenons la convergence ponctuelle et uniforme presque complète avec vitesse de l'estimateur construit. Comme application nous discutons l'impact de ce résultat en prévision non paramétrique fonctionnelle à partir de l'estimation du mode conditionnelle et le quantile conditionnelle ainsi que le risque maximum. Notons que toutes ces propriétés asymptotiques ont été obtenues sous des conditions standard et elles mettent en évidence le phénomène de concentration de la mesure de probabilité de la variable fonctionnelle sur des petites boules.

Nos résultats asymptotiques exploitent bien la structure topologique de l'espace fonctionnel de nos observations et le caractère fonctionnel de nos modèles. En effet, toutes nos vitesses de convergence sont quantifiées en fonction de la concentration de la mesure de probabilité de la variable fonctionnelle, de l'entropie de Kolmogorov et du degré de régularité des modèles.

ABSTRACT

In this thesis, we study the problem of nonparametric modelization when the data are curves. Indeed, we consider real random variable (named response variable) Y and a functional variable (explanatory variable) X. The nonparametric model used to study the relation between explanatory variable and reponse variable. The maximum of the conditional hazard function is a parameter of great importance in statistics, in particular in seismicity studies, because it constitutes the maximum risk of occurrence of an earthquake in a given interval of time. Using the kernel nonparametric estimates based on convolution kernel techniques of the first derivative of the conditional hazard function, we establish the asymptotic behavior of a hazard rate in the presence of a functional explanatory variable and asymptotic normality of the maximum value in the case of independence data.

We propose to study some functional parameters when the data are generated from a model of regression to a single index. We study two functional parameters.

Firstly, we suppose that the explanatory variable takes its values in Hilbert space (infinite dimensional space) and we consider the estimate of the conditional quantile by the kernel method. We establish some asymptotic properties of this estimator in independent case.

As an application we discuss the impact of this result in functional nonparametric prevision for the estimation of the risk maximum.

Note that all these asymptotic properties are obtained under standard conditions and they highlight the phenomenon of concentration proprieties on small balls probability measure of the functional variable.

Our asymptotic results exploit the topological structure of functional space for the observations. Let us note that all the rates of convergence are based on an hypothesis of concentration of the measure of probability of the functional variable on the small balls and also on the Kolmogorov's entropy which measures the number of the balls necessary to cover some set.

Contents

Chapter 1

Introduction.

This chapter is devoted to the presentation of asymptotic notations and results, then at the end a short description of the thesis will be given.

1.1 French abstract

Dans cette thèse, nous nous proposons d'étudier quelques paramètres fonctionnels premièrement nous, nous proposons d'étudier le problème de la modélisation non paramétrique lorsque les variables statistiques sont des courbes. Plus précisément, nous nous intéressons à des problèmes de prévisions à partir d'une variable explicative à valeurs dans un espace de dimension infinie (espace fonctionnel), et nous cherchons à développer des alternatives à la méthode de régression. En effet, nous supposons qu'on dispose d'une variable aléatoire réelle (réponse), souvent notée Y et d'une variable fonctionnelle (explicative), souvent notée X. Le modèle non paramétrique utilisé pour étudier le lien entre X et Y concerne la distribution conditionnelle dont la fonction de répartition (respectivement la densité), notée F (respectivement f), est supposée appartenir à un espace fonctionnel approprié.

Deuxièment lorsque les données sont générées à partir d'un modèle de régression à indice simple. Nous étudions deux paramètres fonctionnels.

Dans un premier temps nous supposons que la variable explicative est à valeurs dans un espace de Hilbert (dimension infinie) et nous considérons l'estimation de la distribution conditionnelle ainsi que les dérivées successives de la densité conditionnelle par la méthode de noyau. Nous traitons les propriétés asymptotiques de cet estimateur dans les deux cas indépendant et dépendant. Pour le cas où les observations sont indépendantes identiquement distribuées (i.i.d), nous obtenons la convergence ponctuelle et uniforme presque complète avec vitesse de l'estimateur construit. Comme application nous discutons l'impact de ce résultat en prévision non paramétrique fonctionnelle à partir de l'estimation de mode conditionnelle et le quantile conditionnelle.

Dans un second nous considérons que nos données ne sont pas indépendantes.

La dépendance est modélisée via la corrélation des variables. Dans ce contexte nous établissons la convergence ponctuelle et uniforme presque complète avec vitesse de l'estimateur construit ainsi que la normalité asymptotique de l'estimateur à noyau de la distribution conditionnelle convenablement normalisée. Nous donnons de manière explicite la variance asymptotique. Notons que toutes ces propriétés asymptotiques ont été obtenues sous des conditions standard et elles mettent en évidence le phénomène de concentration de la mesure de probabilité de la variable fonctionnelle sur des petites boules. Comme application nous discutons l'impact de ce résultat en prévision non paramétrique fonctionnelle à partir de l'estimation du risque maximum.

Nos résultats asymptotiques exploitent bien la structure topologique de l'espace fonctionnel de nos observations et le caractère fonctionnel de nos modèles. En effet, toutes nos vitesses de convergence sont quantifiées en fonction de la concentration de la mesure de probabilité de la variable fonctionnelle, de l'entropie de Kolmogorov et du degré de régularité des modèles.

1.2 Summary

In this thesis, we study the problem of nonparametric modelization when the data are curves. Indeed, we consider real random variable (named response variable) Y and a functional variable (explanatory variable) X. The nonparametric model used to study the relation between X and Y is the conditional distribution function noted F which has a density f. Both F and f are supposed to belong to some suitable functional spaces.

We propose to study some functional parameters when the data are generated from a model of regression to a single index. We study two functional parameters.

Firstly, we suppose that the explanatory variable takes its values in Hilbert space (infinite dimensional space) and we consider the estimate of the conditional quantile by the kernel method. We establish some asymptotic properties of this estimator in both dependent cases.

Secondly, we consider mixing data. In the dependent case we modelize the later via the correlation of variables. For the case where the observations are strong mixing, we obtain the pointwise and uniform almost complete convergence with rate of the estimator and the asymptotic normality of the kernel estimator of the conditional distribution suitably normalized. We give explicitly the asymptotic variance.

As an application we discuss the impact of this result in functional nonparametric prevision for the estimation of the risk maximum.

Note that all these asymptotic properties are obtained under standard conditions and they highlight the phenomenon of concentration proprieties on small balls probability measure of the functional variable.

Our asymptotic results exploit the topological structure of functional space for the observations. Let us note that all the rates of convergence are based on an hypothesis of concentration of the measure of probability of the functional

variable on the small balls and also on the Kolmogorov's entropy which measures the number of the balls necessary to cover some set.

1.2.1 Nonparametric conditional models and functional variables

The functional statistics is a field of current research where it now occupies an important place in statistical research. It has experienced very important development in recent years in which mingle and complement several statistical approaches to priori remote This branch of statistics aims to study data that, because of their structure and the fact that they are collected on very fine grids,can be equated with curves or surfaces, eg functions of time or space. The need to consider what type of data, now frequently encountered under the name of functional data in the literature, is above all a practical need. This is the statistical modeling of data that are supposed of curves observed on all their trajectories. This is practically possible because of the precision of modern measuring devices and large storage capacity offered by current computer systems. It is easy to obtain a discretization very fine of mathematical objects such as curves, surfaces, temperatures observed by satellite images.... This type of variables can be found in many areas, such as meteorology, quantitative chemistry, biometrics, econometrics or medical imaging. Among the reference books on the subject, there may be mentioned the monographs of Ramsay and Silverman (1997, 2002) for the applied aspects, Bosq (2000) for the theoretical aspects, Ferraty and Vieu (2006) for nonparametric study and Ferraty and Romain (2011) for recent developments. In the same context, we refer to Manteiga and Vieu (2007) well as Ferraty (2010). The objective of this section is to make a bibliographic study on conditional nonparametric models considered in this thesis. The objective of this section is to make a bibliographic study on conditional nonparametric models considered in this thesis, allowing to compare our results with those that already exist. However, given the extent of the available literature in this area, we can not make a exhaustive exposed. Thus, we will restrict our bibliographical study to nonparametric models. we refer to Bosq and Lecoutre (1987), Schimek (2000), Sarda and Vieu (2000) and Ferraty and Vieu (2003, 2006) for a wide range of references.

Give an exhaustive list of situations where of such data are encountered is not envisaged, but specific examples of functional data will be addressed in this thesis. However, beyond this practical aspect, it is necessary to provide a theoretical framework for the study of these data. Although functional statistics have the same objectives as the other branches of statistics (data analysis, inference...), the data have this perticularty to take their values in infinite dimensional spaces, and the usual methods of multivariate statistics are here set default.

The all earliest works in which we find this idea of functional data are finally relatively "ancient" Rao (1958) and Tucker (1958) are considering thus the principal components analysis and factor analysis for functional data and even are considering explicitly the functional data as a particular data type.Thereafter, Ramsay (1982) gives off the concept of functional data and raises the issue of adapting the methods of multivariate statistics in this functional frame.

From there, the work to explore the functional statistics begin to multiply, eventually leading today to works making reference on the subject, such as for example monographs Ramsay and Silverman (2002 et 2005), Ferraty and Vieu (2006)...
The estimated hazard rate, because of the variety of its possible applications, is an important issue in statistics. This subject can (and should) be approached from several angles according to on the complexity of the problem: eventual presence of censorship in the observed sample (common phenomenon in medical applications, for example), or else presence of explanatory variables.
Thus, the estimation of a hazard rate with the presence of an explanatory variable functional to single functional index is a current issue to which this work proposes to provide an answer elements.

1.3 Bibliographical context

The problem of the forecast is a very frequent question in statistics. In nonparametric statistics, the principal tool to answer to this question is the regression model. This tool took a considerable rise from the number of publications which are devoted to him, that the explanatory variables are linked, multi or infinity dimension. However, this tool of forecast is not very adapted for some situation. As example, let us quote the case of conditional density dissymmetrical or the case where it comprises several peaks with one of the peaks strictly more important than the others. In these various cases, one can hope that the conditional mode, median or quantiles envisage better than the regression.

1.3.1 On the regression model

The first results in functional nonparametric statistics were developed by Ferraty and Vieu (2000) and they relate to the estimation of the regression function in an explanatory variable of fractal dimension. They established the almost complete convergence of a kernel estimator of the nonparametric model in the i.i.d case. By building on recent developments in the theory of probabilities of small balls, Ferraty and Vieu (2004) have 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) has proved asymptotic normality of the estimator of Ferraty et 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. Convergence in mean square was investigated by Ferraty et al. (2007). Specifically, they have explained the exact asymptotic term of the quadratic error. This result was used by Rachdi and Vieu (2007) for determine a criterion for automatic to selection of the smoothing parameter based on cross-validation. The local version of this criterion has been studied by Benhenni *e*t al. (2007). We find in this article a comparative study between the local and global approach. As works recents bibliographic in regression, we refer the reader to Ferraty and Vieu (2011) well as Delsol (2011). Results on uniform integrability were established by Delsol (2007,2009) and Delsol *e*t al. (2011). Other works were interested to estimating the regression function using different approaches : the method of k nearest neighbors by Burba *e*t al. (2008), robust technical by Azzidine *e*t al. (2008) and Crambes *e*t al. (2008), the estimate by the simplified method of local polynomial by Barrientos-Marin *e*t al. (2010).

1.3.2 On data and functional variable

The statistical problems involved in the modeling and the study of functional random variables for a long time know large advantage in statistics. The first work is based on the discretization of these functional observations in order to be able to adapt traditional multivariate statistical techniques. But, thanks to the progress of the data-processing tool allowing the recovery of increasingly bulky data, an alternative was recently elaborate consisting in treating this type of data in its own dimension, i.e. by preserving the functional character. Indeed, since the Sixties, the handling of the observations in the form of trajectories was the object of several studies in various scientific disciplines such Obhukov (1960), Holmstrom (1963) in climatic, Deville (1974) in econometric, Molenaar and Boosma (1987) and then Kirkpatrick and Heckman (1989) in genetic.
The functional models of regression (parametric or not parametric) are topics which were privileged these last years. Within the linear framework, the contribution of Ramsay and Silverman (1997, 2002) presents an important collection of statistical methods for the functional variables. In the same way, note that Bosq (2000) 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 *e*t 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.
Recently, Cardot *e*t al. (2004) introduced, by a method of regularization, an estimator for the conditionals quantiles, saw as continues linear forms in Hilbert space. Under conditions on the eigenvalues of the covariance operator of the explanatory variable and on the density of conditional law, they gave the speed of norm convergence in L^2 of the built estimator. We return to Cardot *e*t al. (2003) and to Cuevas *e*t al. (2004) for the problem of the test in the functional linear model. Several authors are interested also the answer variable is qualitative, for example, Hastie *e*t al. (1995), Hall and Heckman (2002),....
The study of the nonparametric models of regression is much more than that of the linear case. The results were provided by Ferraty and Vieu (2000). These result were prolonged by Ferraty and Vieu (2002)..., with the problems of the

regression such forecast in the context of time series. By taking again the estimator of Ferraty and Vieu (2004) and by using the property of concentration of the measurement of probability of the functional explanatory variable, Dabo-Niang and Rhomari (2003) studied norm convergence in L^P of regression estimator. They applied their result to the discrimination and the classification of the curves. Other authors were interested if the answer variable is functional using linear model (Bosq and Delecroix (1985), Besse et al. (2000)). Recently, of the first work relating to model presenting at the same time linear and nonparametric aspects were realized by Ferraty et al. (2003), Aït-Saïd et al. (2005, 2008), Ferré and Villa Ferr(2005)...
The first work on the functional variables of distribution estimate was given by Geffroy (1974), Gasser et al. (1998) then Hall and Heckman (2002) were interested in the nonparametric estimate of the distribution mode a functional variable. The estimate of the median of a random variable distribution which takes its values in a Banach space was studied by Cadre (2001). Dabo-Niang (2002) gives an estimator of the density in a space of infinite dimension and established asymptotic results of this estimator, such convergence on average quadratic, almost sure convergence and the asymptotic normality of an estimator of the histogram type. We will also find in this article an application giving the expression of convergence speed in the case of the estimate of the density of a diffusion process relatively to Wiener measure. Ferraty and Vieu studied the nonparametric estimator of the mode of the density of a random variable with values in a semi-norm vector space of infinite dimension. They establish its almost sure convergence and they also apply this result if the measurement of probability of the variable checks a condition of concentration. Several authors were interested in the application of statistical modeling by functional variables on real data. As example, Ferraty and Vieu (2002, 2003) were interested in spectrometric data and with vocal recordings, Besse et al. (2000) with weather data, Gasser et al. (1998) considered medical data, Ferraty *et al.* (2005) considered environmetric and meteorology data where they have gave an example of application to the prediction via the conditional median, together with the determination of prediction intervals...

1.3.3 Concrete problem in statistics for functional variables

In this part we mention a few areas wherein appear the functional data to give an idea of the type of problems that functional statistics solves.

• In biology, we find the first precursor work of (1958) concerning a study of growth curves. More recently, another example is the study of variations of the angle of the knee during walking (Ramsay and Silverman, 2002) and knee movement during exercise under constraint (Abramovich and Angelini (2006), and Antoniadis and Sapatinas (2003). Concerning animal biology, studies of the oviposition of medfly were made by several authors (Chiou *et al.* (2003a, 2003b), Cardot (2007) and Chiou and Müller (2007)). The data consist of curves giving the spawn for each quantity of eggs over time.

• Chemometrics is part of the fields of study that promote the use of methods for functional statistical. Of many existing work on the subject, include Frank and Friedman (1993) , Hastie and Mallows (1993) who have commented on the article by Frank and Friedman (1993) providing an example of the measuring curves log-intensity of a laser radius refracted depending on the angle of refraction. In 2002, Ferraty and Vieu were interested in the study of the percentage of fat in the piece of meat (reponse variable) given the absorption curves of infrared wavelengths of these pieces of meat (explanatory variable).

• Of environment-related applications have been particularly studied by Aneiros-Perez *et al.* (2004) who have worked on a forecasting problem of pollution. These data consist of measurements of peak ozone pollution every day (variable interest) given curves pollutants and meteorological curves before (explanatory variables).

• Climatology is an area where functional data appear naturally. A study of the phenomenon El Niño (hot current in Pacific Ocean) has been realized by Besse *et al.* (2000); Ramsay and Silverman (2005), Ferraty *e*t al. (2005) and Hall and Vial (2006).

• In linguistics, the works have also been realized, particularly concerning voice recognition. Mention may be made, for example Hastie *et al.* (1995), Berlinet *et al.* (2005) or again Ferraty and Vieu (2003, 2006). This works are strongly related to methods of classification when the explanatory variable is a curve. Briefly, the data curves corresponding to records of phonemes spoken by different individuals.A label is associated with each phoneme (reponse variable) and the goal is to establish a classification of these curves using as explanatory variable the recorded curve.

• In the field of graphology, the contribution of functional statistical techniques has again found application. The works on this problem are for example those of Hastie *et al.* (1995) and Ramsay (2000). The latter for example modelize the pen position (abscissa and ordinate versus time) using differential equations.

• The applications to economics are also relatively many. Works have been realized especially by Kneip and Utikal (2001), and recently by Benko *et al.* (2005), based in particular on an analysis of functional principal components.

There are other areas where the functional statistics was employed such as for example processing of sound signals (Lucero, 1999) or recorded by a radar (Hall et al (2001)), the demographic studies (Hyndman and Ullah (2007)),... and the applications in fields as varied as criminology (how to model and compare the evolution of the crime of an individual during time?) Paleo pathology (can you tell an individual if suffering arthritis from the shape of his femur?) The results study in school tests,...

Finally, one may be led to study the functional random variables even if it has available actual initial data independent or multivariate. This is the case when one wants to compare or study functions that can be estimated from the data. Among Typical examples of this type of situation one can evoke comparison of different density functions (see Kneip and Utikal (2001), Ramsay and Silverman (2002), Delicado (2007) and Nerini Ghattas (2007)), functions

regressions (Härdle and Marron (1990), Heckman and Zamar (2000)), the study of the function representing the probability that an individual has to respond to a test according on its "qualities" correctly Ramsay and Silverman (2002)),...
One can imagine that in the future the use of statistical methods functional will be extended to other areas.

1.3.4 On the problematic of single index models

For several years, a increasing interest is worn to models which incorporating of both the parts parametric and nonparametric. Such models type are called semi-parametric model. This consideration is due primarily to problems due to poor specification of some models. Tackle a problem of mis-specification semiparametric way consists in not specify the functional form of some model components. This approach complete those non-parametric models, which can not be useful in small samples, or with a large number of variables. As example, in the classical regression case, the important parameter whose one assumed existence is the regression function of Y knowing the covariate X, denoted $r(x) = \mathbb{E}(Y|X = x)$, $X, Y \in \mathbb{R}^d \times \mathbb{R}$. 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 $O(n^{-k/2k+d})$ deteriorate. 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 writing than the regression function is of the form

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

The models defined 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 model 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 bring to the covariates a dimension in smaller than dimension of the space variable, thus allowing overcome the problem of curse of dimensionality. For example, 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 chronologies series, several semi-parametric approaches have been

proposed. Without pretend to exhaustively, we quote for example: Xia and An (2002) for the index model. A general presentation of this type of model is given in Ichimura *e*t al. (1993) where the convergence and asymptotic normality are obtained. In the case of M-estimators, Delecroix *e*t al. (1999) proves the consistency and asymptotic normality of the estimate the index and they study it's effectiveness. The statistical literature on these methods is rich, quote Huber (1985) and Hall (1989) present an estimation method which consists projecting the density and the regression function on a space of dimension one, to bring a non-parametric estimation for dimensional covariates. This amounts exactly to estimate these functions in a single index model. Attaoui *e*t al. (2011) have established the pointwise and the uniform almost complete convergence (with the 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 the preliminary estimate of the conditional density. recently Mahiddine *e*t al. (2014) have 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. Without pretend to exhaustivity, we quote for example Härdle *e*t al. (1993), Hristache *e*t al. (2001). Based on the regression function, Delecroix *e*t 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 *e*t 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 *e*t al. (2005). Aït Saidi *e*t 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.

1.3.5 On the conditional model

Nonparametric estimation of the conditional density has been widely studied, when the data is real The First related result in nonparametric functional statistic was obtained by Ferraty *e*t al. (2006). They established the almost complete consistency in the independent and identically distributed (i.i.d.) random variables of the kernel estimator of the conditional distribution and the successive derivatives of conditional probability density.

These results have been extend to dependent data by Ferraty *e*t al. (2005) and Ezzahrioui and Ould Saïd (2010). we send back to Cardot *e*t al. (2004) for one approach for linear the conditional quantile statistical functional. The contribution of the thesis on this model is the study of some asymptotic properties related with the nonparametric estimation of the maximum of the conditional hazard function. The asymptotic results (with rates) are precised. The results obtain The results are detailed in Chapter 3 of this thesis. In a functional data setting, the conditioning variable is allowed to take its values in some abstract semi-metric space. In this case, Ferraty et al. (2005) define non-parametric estimators of the conditional density and the conditional distribution. They give the rates of convergence (in an almost complete sense) to the corresponding functions, in a dependence (α-mixing) context. We extend their results by calculating the maximum of the conditional hazard function of these estimates, and establishing their asymptotic normality, considering a particular type of kernel for the vectorial part of the estimates. Because the hazard function estimator is naturally constructed using these two last estimators, the same type of properties is easily derived for it.

1.3.6 On the 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. We prove consistency properties (with rates) in various situations, including censored and/or dependent variables. 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 author has illustrated these asymptotic results by an application on seismic data. We can also look at the recent work of Laksaci et Mechab (2010) on estimating of conditional hazard function for functional data spatially dependent. In this thesis, we deal the nonparametric estimate of the high risk of the conditional hazard function conditional. We establish the asymptotic behavior of a hazard rate in the presence of a random explanatory variable and asymptotic normality of of independence data.

1.4 Local Weithing of Functional Variables

In the finite dimensional case, the local weighting techniques are very popular in the community of nonparametricians because they are very well adapted to nonparametric models. Clearly, local approaches need to have at hand some topological ways for measuring proximity between functional data.
In the finite dimensional case, one of the most common approaches among these local weighting methods is certainly the kernel one. It is impossible to give an exhaustive bibliography about nonparametric methods for finite dimensional variables, but the state of art in this field is well summarized in Schimek (2000) and Akritas and Politis (2003) while a large number of references can be found in

Sarda and Vieu (2000) concerning the kernel methods especially. We will see in this section how kernel smoothing ideas can be adapted to infinite dimensional variables.

The background presented above is sufficient to introduce the kernel local weighting in the functional case. Let $X_1, X_2, \ldots, X_n$ be n f.r.v. valued in $\mathcal{F}$ and let x be a fixed element of $\mathcal{E}$. A naive functional extension of multivariate kernel local weighting ideas would be to transform the n f.r.v. $x_1, x_2, \ldots, x_n$ into the n quantities

$$\frac{1}{V(h)} K\left(\frac{d(x, X_i)}{h}\right),$$

where d is a semi-metric on $\mathcal{F}$, K is a real (asymmetrical) kernel. In this expression $V(h)$ would be the volume of

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

which is the ball, with respect to the topology induced by d, centered at x and of radius h. However, this naive approach requests to define $V(h)$. In other words, this needs to have at hand a measure on $\mathcal{F}$. This is the main difference with real and multivariate cases for which the Lebesgue measure is implicitly used whereas in the functional space $\mathcal{F}$ we do not have such a universally accepted reference measure (see Dabo-Niang and Rhomari (2003) for deeper discussion).

Therefore, in order to free oneself of a choice of particular measure, we build the normalization by using directly the probability distribution of the f.r.v. The *functional* kernel local weighted variables are defined by:

$$\Delta_i = \frac{K\left(\frac{d(x,X_i}{h}\right)}{E\left(K\left(\frac{d(x,X_i}{h}\right)\right)}. \tag{1.1}$$

If we go back quickly to the multivariate case we have, for some constant C depending on K and on the norm $\|.\|$ used $\mathbb{R}^p$,

$$\mathbb{E}K(\|\mathbf{x} - \mathbf{X}_i\|/h) \sim Cf(\mathbf{x})h^p,$$

as long as $\mathbf{X}_i$ has a density f with respect to Lebesgue measure which is continuous and such that $f(x) > 0$ (this kind results known in the literature as the Bochner's type theorem in Collomb (1976) gives a large scope on such results). So, it is clear now that (1.1) is an extension of the multivariate kernel local weighting in the functional framework.

Note that the kernel functions K to be used here necessarily the asymmetrical ones described in multivariate case above. For the sake of simplicity, in the remainder of this work, we will consider only two kinds of kernel for weighting functional variables.

1.5 Various Approaches to the Prediction Problem

Let us start by recalling some notation. Let $(X_i, Y_i) i = 1, \dots n$ be n independent pairs, identically distributed as (X, Y) and valued in $\mathcal{E} \times \mathbb{R}$, where $(\mathcal{E}, d)$ is a semi-metric space (i.e. X is a f.r.v. and d a semi-metric). Let x (resp. y) be a fixed element of $\mathcal{E}$ (resp. $\mathbb{R}$), let $\mathcal{N}_x \subset \mathcal{E}$ be a neighboorhood of x and S be a fixed compact subset of $\mathbb{R}$. Given x, let us denote by $\widehat{y}$ a predicted value for the scalar response.
We propose to predict the scalar response Y from the functional predictor X by using various methods all based on the conditional distribution of Y given X. This leads naturally to focus on some conditional features such as condition expectation, median, mode and quantiles. The regression (nonlinear) operator r of Y on X is defined by

$$r(x) = \mathbb{E}(Y|X = x),$$

and the condition cumulative distribution function (c.d.f) of Y given X is defined by:

$$\forall y \in \mathbb{R}, \quad F_Y^X(x, y) = \mathbb{P}(Y \leq y|X = x).$$

1.6 Kernel Estimators

Once the nonparametric modeling has been introduced, we have to find ways to estimate the various mathematical objects exhibited in the previous models, namely the (nonlinear) operator r, F_Y^X and f_Y^X.

- **Estimating the regression**. We propose for the nonlinear operator r the following functional kernel regression estimator:

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

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:

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

Thus, it is easy to rewrite estimator $\widehat{r}(x)$ as follows:

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

Which is really a weighted average because:

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

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

1.7 Topological considerations

1.7.1 Kolmogorov's entropy

The purpose of this section is to emphasize the topological components of our study. Indeed,as indicated 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,moreover, we have to take into account the uniformity aspect. To this end,let $\mathcal{S}_{\mathcal{F}}$ be a fixed subset of $\mathcal{H}$ of; we consider the following assumption:

$$\forall x \in \mathcal{S}_{\mathcal{F}}, \quad 0 < C\phi(h) \leq \mathbb{P}\left(X \in B(x,h)\right) \leq C'\phi(h) < \infty.$$

We can say that the first contribution of the topological structure of the functional space is viewed through the function ϕ controlling the concentration of the measure of probability of the functional variable on a small ball. Moreover, for the uniform consistency, where the main tool is to cover a subset $\mathcal{S}_{\mathcal{F}}$ with finite number of balls, one introduces an other topological concept defined as follows:

Definition 1.7.1 *Let $\mathcal{S}_{\mathcal{F}}$ be a subset of a semi-metric space $\mathcal{H}$, and let $\varepsilon > 0$ be given. A finite set of points $x_1, x_2, \ldots, x_N$ in $\mathcal{F}$ is called an ε-net for $\mathcal{S}_{\mathcal{F}}$ if $\mathcal{S}_{\mathcal{F}} \subset \bigcup_{k=1}^{N} B(x_k, \varepsilon)$.*

The quantity $\psi_{\mathcal{S}_{\mathcal{F}}}(\varepsilon) = \log\left(N_\varepsilon(\mathcal{S}_{\mathcal{F}})\right)$, where $N_\varepsilon(\mathcal{S}_{\mathcal{F}})$ 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 introduced 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. Therefore, the choice of the topological structure (with other words, the choice of the semi-metric) will play a crucial role when one is looking at uniform (over some subset $\mathcal{S}_{\mathcal{F}}$) of $\mathcal{F}$) asymptotic results. More precisely, we will see thereafter that a good semi-metric can increase the concentration of the probability measure of the functional variable X as well as minimize the ε-entropy of the subset $\mathcal{S}_{\mathcal{F}}$. In an earlier contribution (see, Ferraty et al., 2006) we highlighted the phenomenon of concentration of the probability measure of the functional variable by computing the small ball probabilities in various standard situations. We will devote Section 1.8.3 to discuss the behaviour of the Kolmogorov's ε-entropy in these standard situations. Finally, we invite the readers interested in these two concepts (entropy and small ball probabilities) or/and the use of the Kolmogorov's ε-entropy in dimensionality reduction problems to refer to respectively, Kuelbs and Li (1993) or/and Theodoros and Yannis (1997).

1.8 Description of the thesis

The first thematic of this thesis focuses on the study of quadratic error in statistical nonparametric functional. Recall that one of the main reasons for the craze of nonparametric functional statistical is the solution it offers to the problem of the curse of dimensionality. This well-known non-parametric statistical phenomenon relates to the significant deterioration of the quality of the estimate when the dimension increase. Our study highlights the phenomenon of concentration properties on small balls of the probability measure of the functional variable.

The second problematic addressed is devoted to the study of some functional parameters in models to revelatory index. We treat the conditional cumulative distribution function and the successive derivatives of the conditional density considering a type of data namely full than the i.i.d case. The explanatory variable for functional parameter which is the conditional distribution function is of infinite dimension.

The uniform convergence in functional nonparametric statistic engenders an another problem of dimensionality. Indeed, in a general way the processing of uniform convergence on a given set is related to the number of balls which cover the whole. In finite dimension for a compact set, this number is of the order of r^d where r is the radius of the balls, d is est the dimension of the space. From probabilistic point of view, this relationship is justified by the fact that the probability of the set is bounded above by the number of balls multiplied by r^d which is the Lebesgue measure of a ball of radius r. So, we can say that there is a relationship between the number of balls, the size of the space and the probability measure used. Thus, it is natural to wonder about the uniform convergence rate of the estimators when the dimension is infinite. Of course, this number depends on the topological structure of the space of functional variable considered but the most important issues are :

1. Can we find a compromise between the radius of the ball and the number of balls to ensure uniform convergence of estimators built?
2. Can we optimize the speed of convergence based on considered the topological structure?

The study conducted in the third part of this thesis is an answer to this question and the concept of entropy plays a key role in our approach.

1.8.1 Plan of the thesis

After devoting the first part of the presentation of the asymptotic notations and results as well as the short description of the thesis. Then, this thesis is divided into two parts. The first part interested only on a real response variable and the case of i.i.d observations. In this context, we propose an estimate of the maximum risk, through the nonparametric estimation of the conditional hazard function. The high risk of the conditional hazard function is a parameter of great importance in seismicity studies, because it constitutes the maximum risk of occurrence of an earthquake in a given interval of time. It is shown that the (empirically determined) the high risk of the kernel estimate we establish the asymptotic behavior of a hazard rate in the presence of a random explanatory variable and asymptotic normality of independence data. We state the almost complete convergence (with rates of convergence) for nonparametric estimates of the derivative of the conditional hazard and the maximum risk and we calculate the variance of the conditional density, distribution and hazard estimates, the asymptotic normality of the three estimators considered is developed.

In the second part, we examine the conditional distibution function and we focus on the case of i.i.d observations. We build in this case a kernel estimator for this functional parameter and we state a nonparametric estimation of some characteristics of the conditional distribution where Kernel type estimators for the conditional cumulative distribution function and the successive derivatives of the conditional density in the single functional index model are introduced. We establish the pointwise almost complete convergence and the uniform almost complete convergence (with the rate) of the kernel estimate of this model. Asymptotic properties are stated for each of these estimates, and they are applied to the estimations of the conditional mode and conditional quantiles.

We will finish this section with some prospects research.

1.8.2 Definitions and outils

All through this party, $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are sequences of real random variables, while $(u_n)_{n\in\mathbb{N}}$ is a deterministic sequence of positive real numbers. We will use the notation $(Z_n)_{n\in\mathbb{N}}$ for a sequence of independent and centered r.r.v.

Definition 1.8.1 *One says that $(X_n)_{n\in\mathbb{N}}$ converges almost completely (a.co.) to some r.r.v. X, if and only if*

$$\forall\varepsilon > 0, \quad \sum_{n\in\mathbb{N}} \mathbb{P}\left(|X_n - X| > \varepsilon\right) < \infty,$$

and the almost complete convergence of $(X_n)_{n\in\mathbb{N}}$ to X is denoted by

$$\lim_{n\to\infty} X_n = X, \ \textit{a.co.}$$

Definition 1.8.2 *One says that the rate of almost complete convergence of* $(X_n)_{n\in\mathbb{N}}$ *to* X *is of order* u_n *if and only if*

$$\exists\varepsilon_0 > 0, \qquad \sum_{n\in\mathbb{N}} \mathbb{P}\left(|X_n - X| > \varepsilon_0 u_n\right) < \infty,$$

and we write

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

Proposition 1.8.1 *Assume that* $\lim_{n\to\infty} u_n = 0,\ X_n = O_{a.co.}(u_n)$ *and* $\lim_{n\to\infty} Y_n = l_0$, *a.co., where* l_0 *is a deterministic real number.*

i) We have $X_nY_n = O_{a.co.}(u_n)$*;*

ii) We have $\dfrac{X_n}{Y_n} = O_{a.co.}(u_n)$ *as long as* $l_0 \neq 0$.

Remark 1.8.1 *The almost convergence of* Y_n *to* l_0 *implies that there exists some* $\delta > 0$ *such that*

$$\sum_{n\in\mathbb{N}} \mathbb{P}\left(|Y_n| > \delta\right) < \infty.$$

Now, one suppose $Z_1, \ldots, Z_n$ will be independent r.r.v. with zero mean. As can be seen throughout this party, the statement of almost complete convergence properties needs to find an upper bound for some probabilities involving sum of r.r.v. such as

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} Z_i\right| > \varepsilon\right),$$

where, eventually, the positive real ε decreases with n. In this context, there exists powerful probabilistic tools, generically called *E*xponential Inequalities. The literature contains various versions of exponential inequalities. These inequalities differ according to the various hypotheses checked by the variables Z_i's. We focus here on the so-called Bernstein's inequality. This choice was made because the from of Bernstein's inequality is the easiest for the theoretical developments on functional statistics that have been stated throughout our thesis. Other forms of such exponential inequality can be found in Fuk-Nagaev (1971) (see also Nagaev (1997) and (1998))

Proposition 1.8.2 *Assume that*

$$\forall m \geq 2,\ |\mathbb{E}Z_i^m| \leq (m!/2)(a_i)^2 b^{m-2},$$

and let $(A_n)^2 = (a_1)^2 + \ldots + (a_n)^2$. *Then, we have:*

$$\forall\varepsilon \geq 0,\ \ \mathbb{P}\left(\left|\sum_{i=1}^{n} Z_i\right| \geq \varepsilon A_n\right) \leq 2\exp\left\{-\frac{\varepsilon^2}{2\left(1+\frac{\varepsilon b}{A_n}\right)}\right\}.$$

Corollary 1.8.1 *i) If* $\forall m \geq 2, \exists C_m > 0,\ \mathbb{E}|Z_1^m| \leq C_m a^{2(m-1)}$*, we have*

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

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

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

Remark 1.8.2 *By applying Proposition 1.8.2 with* $A_n = a\sqrt{u_n}$*,* $b = a^2$ *and taking* $\varepsilon = \varepsilon_0\sqrt{u_n}$*, we obtain for some* $C' > 0$*:*

$$\mathbb{P}\left(\frac{1}{n}\Big| \sum_{i=1}^n Z_i \Big| > \varepsilon_0\sqrt{u_n} \right) \leq 2\exp\left\{ -\frac{\varepsilon_0^2 \log n}{2(1+\varepsilon_0\sqrt{u_n})} \right\} \leq 2n^{-C'\varepsilon_0^2}.$$

Corollary 1.8.2 *i) If* $\exists M < \infty,\ |Z_1| \leq M$*, and denoting* $\sigma^2 = \mathbb{E}Z_1^2$*, we have*

$$\forall \varepsilon \geq 0, \quad \mathbb{P}\left(\Big| \sum_{i=1}^n Z_i \Big| \geq n\varepsilon \right) \leq 2\exp\left\{ -\frac{n\varepsilon^2}{2\sigma^2(1+\varepsilon\frac{M}{\sigma^2})} \right\}.$$

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

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

Remark 1.8.3 *By applying Proposition 1.8.2 with* $a_i^2 = \sigma^2$*,* $A_n = n\sigma^2$*, and by choosing* $\varepsilon = \varepsilon_0\sqrt{u_n}$*, we obtain for some* $C' > 0$*:*

$$\mathbb{P}\left(\frac{1}{n}\Big| \sum_{i=1}^n Z_i \Big| > \varepsilon_0\sqrt{u_n} \right) \leq 2\exp\left\{ -\frac{\varepsilon_0^2 \log n}{2(1+\varepsilon_0\sqrt{v_n})} \right\} \leq 2n^{-C'\varepsilon_0^2}.$$

where $v_n = \frac{M u_n}{\sigma_n^2}$.

Definition 1.8.3 *i) A function* K *from* $\mathbb{R}$ *into* $\mathbb{R}^+$ *such that* $\int K = 1$ *is called a kernel of type I if there exist two real constants* $0 < C_1 < C_2 < \infty$ *such that:*

$$C_1 \mathbf{1}_{[0,1]} \leq K \leq C_2 \mathbf{1}_{[0,1]}.$$

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 $-\infty < C_2 < C_1 < 0$:

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

The first kernel family contains the usual discontinuous kernels such as the asymmetrical box one while the second family contains the standard asymmetrical continuous ones (as the triangle, quadratic, ...). Finally, to be in harmony with this definition and simplify our purpose, for local weighting of real random variables we just consider the following kernel-type.

Definition 1.8.4 *A function K from $\mathbb{R}$ into $\mathbb{R}^+$ such that $\int K = 1$ with compact support $[-1,1]$ and such that $\forall u \in (0,1),\ K(u) > 0$ is called a kernel of type 0.*

We can now build the bridge between local weighting and the notation of small ball probabilities. To fix the ideas, consider the simplest kernel among those of type I namely the asymmetrical box kernel. Let x be f.r.v. valued in $\mathcal{F}$ and x be again a fixed element of $\mathcal{F}$. We can write:

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

Keeping in mind the functional kernel local weighted variables (1.1), the probability of the ball $B(x,h)$ appears clearly in the normalization. At this stage it is worth telling why we are saying *small* ball probabilities. In fact, as we will see later on, the smoothing parameter h (also called the *bandwith*) decreases with the size of the sample of the functional variables (more precisely, h tends to zero when n tends to ∞). Thus, when we take n very large, h is close to zero and then $B(x,h)$ is considered as a small ball and $P(X \in B(x,h))$ as a small ball probability.
From now, for all x in $\mathcal{F}$ and for all positive real h, we will use the notation:

$$\phi_x(h) = \mathbb{P}(X \in B(x,h)).$$

This notion of small ball probabilities will play a major role both from theoretical and piratical points of view. Because the notion of ball is strongly linked with the semi-metric d, the choice of this semi-metric will become an important stage. Now, let X be a f.r.v. taking its values in the semi-metric space $(\mathcal{F}, d)$, let x be a fixed element of $\mathcal{F}$, let h be a real positive number and let K be a kernel function.

Lemma 1.8.1 *If K is a kernel of type I, then there exist nonnegative finite real constant C and C' such that:*

$$C\phi_x(h) \leq \mathbb{E}K\left(\frac{d(x,X}{h}\right) \leq C'\phi_x(h).$$

Lemma 1.8.2 *If K is a kernel of type II and if $\phi_x(.)$ satisfies*

$$\exists C_3 > 0,\ \exists \epsilon_0,\ \forall \epsilon < \epsilon_0,\ \int_0^\epsilon \phi_x(u)du > C_3 \epsilon \phi_x(\epsilon),$$

then there exist nonnegative finite real constant C and C' such that, for h small enough:

$$C\phi_x(h) \leq \mathbb{E}K\left(\frac{d(x,X)}{h}\right) \leq C'\phi_x(h).$$

1.8.3 Some examples

We will start (Example 1) by recalling how this notion behaves in unfunctional case (that is when $\mathcal{F} = \mathbb{R}^p$). Then, Examples 2 and 3 are covering special cases of functional process. More interestingly (from statistical point of view) is Example 4 since it allows to construct, in any case, a semi-metric with reasonably "small" entropy.

Example 1.8.1 *(Compact subset in finite dimensional space) : A standard theorem of topology guaranties that for each compact subset $\mathcal{S}_\mathcal{F}$ of $\mathbb{R}^p$ and for each $\varepsilon > 0$ there is a finite ε-net and we have for any $\varepsilon > 0$,*

$$\psi_{\mathcal{S}_\mathcal{F}}(\varepsilon) \leq Cp \log(1/\varepsilon).$$

More precisely, Chate and Courbage (1997) have shown that, for any $\varepsilon > 0$ the regular polyhedron in $\mathbb{R}^p$ with length r can be covered by $\left([2r\sqrt{p}/\varepsilon]+1\right)^p$ balls, where $[m]$ is the largest integer which is less than or equal to m. Thus, the Kolmogorov's ε-entropy of a polyhedron P_r in $\mathbb{R}^p$ with length r is

$$\forall \varepsilon > 0, \quad \psi_{P_r}(\varepsilon) \sim p \log\left([2r\sqrt{p}/\varepsilon]+1\right).$$

Example 1.8.2 *(Closed ball in a Sobolev space): Kolmogorov and Tikhomirov (1959) obtained many upper and lower bounds for the ε-entropy of several functional subsets. A typical result is given for the class of functions $f(t)$ on $T = [0, 2p)$ with periodic boundary conditions and*

$$\frac{1}{2\pi}\int_0^{2\pi} f^2(t)dt + \frac{1}{2\pi}\int_0^{2\pi} f^{(m)^2}(t)dt \leq r.$$

The ε-entropy of this class, denoted $W_2^m(r)$, is

$$\psi_{W_2^m(r)}(\varepsilon) \leq C\left(\frac{r}{\varepsilon}\right)^{1/m}.$$

Example 1.8.3 *(Unit ball of the Cameron-Martin space) : Recently, Van der Vaart and Van Zanten (2007) characterized the Cameron-Martin space associated to a Gaussian process viewed as map in $\mathcal{C}[0,1]$ with the spectral measure μ satisfying*

$$\int \exp(\delta|\lambda|)\,\mu(d\lambda) < \infty,$$

by

$$H = \left\{ t :\mapsto Re\left(\int e^{-it\lambda} h(\lambda) d\mu(\lambda) \right) : \qquad h \in L_2(\mu) \right\},$$

and they show that Kolmogorov's ε-entropy of the unit ball B^{CMW} of this space with respect to the supremum norm $\|.\|_\infty$ is

$$\psi_{B^{CMW}_{\|\cdot\|_\infty}} \sim \left(\log\left(\frac{1}{\varepsilon} \right) \right)^2, \quad as\ \varepsilon \to 0$$

Example 1.8.4 *(Compact subset in a Hilbert space with a projection semi-metric) : The projection-based semi-metrics are constructed in the following way. Assume that $\mathcal{H}$ is a separable Hilbert space, with inner product $< .,. >$ and with orthonormal basis $\{e_1, \ldots, e_j, \ldots\}$, and let k be a fixed integer, $k > 0$. As shown in Lemma 13.6 of Ferraty and Vieu (2006), a semi-metric d_k on $\mathcal{H}$ can be defined as follows*

$$d_k(x, x') = \sqrt{\sum_{j=1}^{k} < x - x', e_j >^2}. \tag{1.2}$$

Let χ be the operator defined from $\mathcal{H}$ into $\mathbb{R}^k$ by

$$\chi(x) = (< x, e_1 >, \ldots, < x, e_k >),$$

and let d_{eucl} be the euclidian distance on $\mathbb{R}^k$, and let us denote by $B_{eucl}(.,.)$ an open ball of $\mathbb{R}^k$ for the associated topology. Similarly, let us note by $B_k(.,.)$ an open ball of $\mathcal{H}$ for the semi-metric d_k. Because χ is a continuous map from $(\mathcal{H}, d_k)$ into $(\mathbb{R}^k, d_{eucl})$, we have that for any compact subset $\mathcal{S}$ of $(\mathcal{H}, d_k), \chi(\mathcal{S})$ is a compact subset of $\mathbb{R}^k$. Therefore, for each $\varepsilon > 0$ we can cover $\chi(\mathcal{S})$ with balls of centers $z_i \in \mathbb{R}^k$:

$$\chi(\mathcal{S}) \subset \cup_{i=1}^{d} B_{eucl}(z_i, r), \quad with \ \ dr^k = C \ \ for\ some \ \ C > 0. \tag{1.3}$$

For $i = 1, \ldots, d$, let x_i be an element of $\mathcal{H}$ such that $\chi(x_i) = z_i$. The solution of the equation $\chi(x) = z_i$ is not unique in general, but just take x_i to be one of these solutions. Because of (1.2), we have that

$$\chi^{-1}(B_{eucl}(z_i, r)) = B_k(x_i, r). \tag{1.4}$$

Finally, (1.3) and (1.4) are enough to show that the Kolmogorov's ε-entropy of $\mathcal{S}$ is

$$\psi_{\mathcal{S}}(\varepsilon) \approx Ck \log\left(\frac{1}{\varepsilon} \right).$$

1.9 Short presentation of the results

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

1.9.1 Notations

Let X be a functional random variable, *frv* its abbreviation. Let (X_i, Y_i) be a sample of independant pairs, each having the same distribution as (X, Y), our aim is to build nonparametric estimates of several functions related with the conditional probability distribution (*cond-cdf*) of Y given $< X, \theta >=< x, \theta >$.

$$\forall y \in \mathbb{R},\ F(\theta, y, x) = \mathbb{P}(Y \le y| < X, \theta >=< x, \theta >).$$

be the *cond-cdf* of Y given $< X, \theta >=< x, \theta >$, for $x \in \mathcal{H}$, which also shows the relationship between X and Y but is often unknown.

We introduce a kernel type estimators for the conditional cumulative distribution function $\widehat{F}(\theta, \cdot, x)$ of $F(\theta, \cdot, x)$ and the conditional density $\widehat{f}^{(j)}(\theta, \cdot, x)$ of $f^{(j)}(\theta, \cdot, x)$ as follows:

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

where K is a kernel, H is a cumulative distribution function

$$\widehat{f}^{(j)}(\theta, y, x) = \frac{h_H^{-1-j} \sum_{i=1}^{n} K(h_K^{-1}(< x - X_i, \theta >)) H^{(j+1)}(h_H^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(h_K^{-1}(< x - X_i, \theta >))},\ \ y \in \mathbb{R}$$

where K is a kernel, H is a *cdf* and $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers and the j^{th}. $f^{(j)}$ (resp. $\widehat{F}^{(j)}$) is the derivate of f (resp. $\widehat{F}$).

In the following, for any $x \in \mathcal{H}$ and $y \in \mathbb{R}$, let $\mathcal{N}_x$ be a fixed neighborhood of x in $\mathcal{H}$, $\mathcal{S}_{\mathbb{R}}$ will be a fixed compact subset of $\mathbb{R}$, and we will use the notation $B_\theta(x, h) = \{X \in \mathcal{H}/0 < | < x - X, \theta > | < h\}$. Our nonparametric models will be quite general in the sense that we will just need the following simple assumption for the marginal distribution of $< \theta, X >$:

$$\mathbb{P}\left(X \in B_\theta(x, h)\right) = \phi_{\theta,x}(h) > 0.$$

The main objective is to study the pointwise almost complete convergence and the uniform almost complete convergence (with the rate) of the kernel estimate of this model.

The nonparametric estimate $\widehat{h}_Y^X(x, y) = \dfrac{\widehat{f}_Y^X(x, y)}{1 - \widehat{F}_Y^X(x, y)}$ of $h_Y^X(x, y) = \dfrac{f_Y^X(x, y)}{1 - F_Y^X(x, y)}$ when the explanatory variable X is valued in a space of eventually infinite dimension. We establish the asymptotic behavior of a hazard rate in the presence of a random explanatory variable and asymptotic normality of independence data.

1.9.2 Results: single functional index model

Theorem 1 *For any fixed y, we have*

$$|\widehat{F}(\theta,y,x)-F(\theta,y,x)| \;=\; O(h_K^{b_1})+O(h_H^{b_2})+O\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right), \qquad a.co.$$

Theorem 2 *For any fixed y, we have, as n goes to infinity*

$$|\widehat{f}^{(j)}(\theta,y,x)-f^{(j)}(\theta,y,x)| = O\left(h_K^{b_1}\right)+O\left(h_H^{b_2}\right)+O\left(\sqrt{\frac{\log n}{nh_H^{2j+1}\phi_{\theta,x}(h_K)}}\right), \; a.co.$$

In the following result we extended the result of the pointwise convergence in uniform case. The study of the uniform consistency is motivated by the fact that the latter is an indispensable tool for studying the asymptotic properties of all estimates of the functional index θ if is unknown. Thus, by strengthening conditions of preceding result by the following topological terms: Let $\mathcal{S}_{\mathbb{R}}$ is subset compact of $\mathbb{R}$ and $\mathcal{S}_{\mathcal{H}}$ (resp. $\Theta_{\mathcal{H}}$, the space of parameters) such as

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

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, one will have the result.

Theorem 3 *For any compact $\mathcal{S}_{\mathbb{R}}$, $\mathcal{S}_{\mathcal{H}}$ and $\Theta_{\mathcal{H}}$, we have:*

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

Theorem 4 *As n goes to infinity, we have*

$$\sup_{\theta\in\Theta_{\mathcal{F}}}\sup_{x\in\mathcal{S}_{\mathcal{F}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}|\widehat{f}^{(j)}(\theta,y,x)-f^{(j)}(\theta,y,x)| = O(h_K^{b_1})+O(h_H^{b_2})+O_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{nh_H^{2j+1}\phi(h_K)}}\right)$$

The demonstration of these results and the conditions will be given in detail in Chapter 2.

1.9.3 Results: Nonparametric estimation of a high risk

Let $\{(X_i,Y_i),\ i=1,\dots,n\}$ be a sample of n random pairs, each one distributed as (X,Y), where the variable X is of functional nature and X is scalar. Formally, we will consider that X is a random variable valued in some semi-metric functional space $\mathcal{F}$, and we will denote by $d(\cdot,\cdot)$ the associated semi-metric. The conditional cumulative distribution of X given $X=x$ is defined for any $y\in\mathbb{R}$ and any $x\in\mathcal{F}$ by

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

while the conditional density, denoted by $f^x(y)$ is defined as the density of this distribution with respect to the Lebesgue measure on $\mathbb{R}$. The conditional hazard is defined as in the non-infinite case.

The conditional density operator $f^x(\cdot)$ is defined by using kernel smoothing methods

$$\widehat{f}^x(y) = \frac{\sum_{i=1}^{n} h_n^{-1} K\left(h_n^{-1} d(x, X_i)\right) H\left(h_n^{-1}(y - Y_i)\right)}{\sum_{i=1}^{n} K\left(h_n^{-1} d(x, X_i)\right)},$$

where K and H are kernel functions and h_n is sequence of smoothing parameter. The conditional distribution operator $F^x(\cdot)$ can be estimated by

$$\widehat{F}^x(y) = \sum_{i=1}^{n} W_{ni}(x) \mathbf{1}_{\{Y_i \leq y\}}, \qquad \forall y \in \mathbb{R}$$

with $\mathbf{1}_{\{\cdot\}}$ being the indicator function and where $W_{ni}(x) = \frac{h_n^{-1} K(h_n^{-1} d(x, X_i))}{\sum_{j=1}^{n} K(h_n^{-1} d(x, X_j))}$, K is a kernel function and h_n is a sequence of positive real numbers which goes to zero as n goes to infinity.

For $x \in \mathcal{F}$, we denote by $h^x(\cdot)$ the conditional hazard function of Y_1 given $X_1 = x$. We assume that $h^x(\cdot)$ is unique maximum and its high risk point is denoted by $\theta(x) := \theta$, which is defined by

$$h^x(\theta(x)) := h^x(\theta) = \max_{y \in \mathcal{S}} h^x(y). \tag{1.5}$$

A kernel estimator of θ is defined as the random variable $\widehat{\theta}(x) := \widehat{\theta}$ which maximizes a kernel estimator $\widehat{h}^x(\cdot)$, that is,

$$\widehat{h}^x(\widehat{\theta}(x)) := \widehat{h}^x(\widehat{\theta}) = \max_{y \in \mathcal{S}} \widehat{h}^x(y) \tag{1.6}$$

Note that the estimate $\widehat{\theta}$ is note necessarily unique and our results are valid for any choice satisfying (3.3). We point out that we can specify our choice by taking

$$\widehat{\theta}(x) = \inf \left\{ t \in \mathcal{S} \text{ such that } \widehat{h}^x(t) = \max_{y \in \mathcal{S}} \widehat{h}^x(y) \right\}.$$

Theorem 5 *We have*

$$\widehat{\theta} - \theta \to 0 \quad a.co. \tag{1.7}$$

Theorem 6 *Under assumption, we have*

$$\sup_{y \in \mathcal{S}} |\widehat{\theta} - \theta| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n h_n^3 \phi_x(h_n)}}\right) \tag{1.8}$$

Theorem 7 *Under conditions, we have* $(\theta \in \mathcal{S}/f^x(\theta), 1 - F^x(\theta) > 0)$

$$\left(nh_n^3\phi_x(h_n)\right)^{1/2}\left(\widehat{h}'^x(\theta) - h'^x(\theta)\right) \xrightarrow{\mathcal{D}} N\left(0, \sigma^2_{h'}(\theta)\right)$$

where $\rightarrow^{\mathcal{D}}$ *denotes the convergence in distribution,*

$$a_l^y = K^l(1) - \int_0^1 \left(K^l(u)\right)' \zeta_0^y(u)du \quad \textit{for } l = 1, 2$$

and

$$\sigma^2_{h'}(\theta) = \frac{a_2^y h^x(\theta)}{\left(a_1^y\right)^2 (1 - F^x(\theta))} \int (H'(t))^2 dt.$$

Theorem 8 *Under conditions, we have* $(\theta \in \mathcal{S}/f^x(\theta), 1 - F^x(\theta) > 0)$

$$\left(nh_n^3\phi_x(h_n)\right)^{1/2}\left(\widehat{\theta} - \theta\right) \xrightarrow{\mathcal{D}} N\left(0, \frac{\sigma^2_{h'}(\theta)}{(h''^x(\theta))^2}\right)$$

with $\sigma^2_{h'}(\theta) = h^x(\theta)\,(1 - F^x(\theta)) \int (H'(t))^2 dt.$

Bibliography

[1] Abramovich, F., Angelini, C. (2006). Bayesian maximum a posteriori multiple testing procedure. *Sankhyã*, **68**, 436-460.

[2] 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.

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

[4] Akritas, M., Politis, D. (2003). (ed.) Recent advances and trends in nonparametric statistics. Elsevier, Amsterdam.

[5] Aneiros-Pérez, G., Cardot, H., Estévez, G., Vieu, P. (2004). Maximum ozone concentration forecasting by functional non-parametric approaches, *Environmetrics*, **15**, 675-685.

[6] Antoniadis, I., Sapatinas, T. (2003). Wavelet methods for continuoustime prediction using Hilbert-valued autoregressive processes. *J. Multivariate. Anal.* **87**, 133-158.

[7] 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**, No.1, 45-53.

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

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

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

[11] Benko, M., Hardle, W. and Kneip, A. (2005). Common functional principal components. SFB 649 Economic Risk Discussion Paper, 2006-2010.

[12] Berlinet, A., Biau, G.and Rouvière, L. (2005). Parameter selection in modified histogram estimates. *Statistics*, **39**, 91-105.

[13] Besse, P., Cardot, H., Stephenson, D. (2000). Autoregressive forecasting of some functional climatic variations. *Scand. J. Stat.* **27**, No.4, 673-687.

[14] Bosq, D., (1991). Modelization, non-parametric estimation and prediction for continuous time processes. In Nonparametric Functional estimation and Related Topics (Spetses, 1990), 509-529, NATO, Adv. Sci. Inst. Ser. C Math. Phys. Sci. **335**, Kluwer Acad. Publ., Dordrecht.

[15] Bosq, D. (2000). Linear process in function space. Lecture notes in Statistics. **149**, Springer-Verlag.

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

[17] Bosq, D., Lecoutre, J. P.(1987). Théorie de l'estimation fonctionnelle. ECONOMICA, Paris.

[18] 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.

[19] Cadre, B. (2001). Convergent estimators for the L_1-median of a Banach valued random variable. *Statistics*, **35**, No.4, 509-521.

[20] Cardot, H. (2007). Conditional functional principal components analysis. *Scand. J. Stat.* **34**, 317-335.

[21] Cardot, H., Crambes, C., Sarda, P. (2004). Spline estimation of conditional quantiles for functional covariates. *C. R. Acad. Sci.* Paris. **339**, No.2, 141-144.

[22] Cardot, H., Ferraty, F., Mas, A., Sarda, P. (2003). Testing hypotheses in the functional linear model. *Scand. J. Stat.* **30**, No.1, 241-255.

[23] Cardot, H., Ferraty, F., Sarda, P. (1999). Functional linear model. *Stat. Probab. Lett.* **45**, No.1, 11-22.

[24] Chate, H., Courbage, M. (1997). Lattice systems. *Physica.* D, **103**, 1-612.

[25] Chiou, J., Müller, H.G. (2007). Diagnostics for functional regression via residual processes. *Computational Statistics and Data Analysis*, **51**, 4849-4863.

[26] Chiou, J. M., Müller, H. G., Wang, J. L. and Carey, J. R. (2003a). A functional multiplicative effects model for longitudinal data, with application to reproductive histories of female medflies. *Statist. Sinica.* **13**, 1119-1133.

[27] Chiou, J. M., Müller, H. G. and Wang, J. L. (2003b). Functional quasi-likelihood regression models with smooth random effects. *J. Royal Statist. Soc. Ser. B.*, **65**, 405-423.

[28] Collomb, G., (1976). Estimation non paramétrique de la régression. (in french). Ph.D. Université Paul Sabatier, Toulouse.

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

[30] Cuevas, A., Febrero, M., Fraiman, R. (2004). An anova test for functional data. *Computational Statistics & Data Analysis*, **47**, No.1, 111-122.

[31] Dabo-Niang, S. (2002). Sur l'estimation fonctionnelle en dimension infinie : Application aux diffusions. Thèse de Doctorat, Université de Paris 6.

[32] 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**, No.1, 75-80.

[33] Delecroix, M, Härdle, W., Hristache, M. (1999). M-estimateurs semi-paramétriques dans les modèles à direction révélatrice unique. *Bull. Belg. Math. Soc.* Simon Stevin, **6**, No.2, 161-185.

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

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

[36] Delsol, L. (2007). Régression non paramétrique fonctionnelle: expression asymptotique des moments, *Ann. I.S.U.P.* **LI**, No.3, 43-67.

[37] Delsol, L. (2009). Advances on asymptotic normality in nonparametric functional Time Series Analysis Statistics. *Statistics*, **43**, 13-33.

[38] Delsol, L. (2011). Nonparametric methods for α-mixing functional random variables. In The Oxford Handbook of Functional Data Analysis (Ed. F. Ferraty and Y. Romain). Oxford University Press.

[39] Ezzahrioui, M., Ould Saïd, E. (2010). Some asymptotic results of a non-parametric conditional mode estimator for functional time series data. *Statist. Neerlandica*, **64**, No.2, 171-201.

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

[41] Ferraty, F. (2010). Special issue on statistical methods and problems in infinite dimensional spaces. *J. Multivariate Anal.* **101**, No.2, 305-490.

[42] 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.

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

[44] 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** No.2, 399-417.

[45] 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.

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

[47] 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.

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

[49] 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.

[50] 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.

[51] Ferraty, F., Vieu, P. (2006). Nonparametric Functional Data Analysis. Theory and Practice. Theory and Practice. Springer-Verlag.

[52] 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.

[53] Ferré, L., Villa, N. (2005). Discrimination de courbes par régression inverse fonctionnelle. *Revue de Statistique Appliquée*, **LIII**, No.1, 39-57.

[54] Frank, I. E., Friedman, J. H. (1993). A statistical view of some chemometrics regression tools (with discussion). Technometrics, **35**, 109-148.

[55] Fuk, D. Kh., Nagaev, S. V. (1971). Probability inequalities for sums of independent random variables, *Theory Prob. Appl.* **16**, 643-660.

[56] 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**, No.4, 681-691.

[57] Geffroy, J. (1974). Sur l'estimation d'une densité dans dans un espace métrique. *C. R. Acad. Sci.*, Paris, Sér. A, **278**, 1449-1452.

[58] Hall P, Poskitt D, Presnell B. (2001) A functional data-analytic approach to signal discrimination. *Technometrics.*; **43** : 1Ű24.

[59] Hall, P. (1989). On projection pursuit regression. *Ann. Statist.* **17**, No.2, 573-588.

[60] Hall, P., Heckman, N. E. (2002). Estimating and depicting the structure of a distribution of random functions. *Biometrika.* **89**, No.1, 145-158.

[61] Hall, P., Vial, C. (2006). Assessing the finite dimensionality of functional data. *Journal of the Royal Statistical Society. Series: B*, **68**, No.4, 689-705.

[62] Härdle, W. (1990). Applied nonparametric regression. Cambridge Univ. Press, UK.

[63] Härdle, W., Hall, P., Ichumira, H., (1993). Optimal smoothing in single-index models, *Ann. Statist.* **21**, 157-178.

[64] Härdle, W., Marron, J. S. (1990). Semiparametric comparison of regression curves. *The Annals of Statistics.* **18**, No.1, 59-63.

[65] Hastie, T., Buja, A., Tibshirani, R. (1995). Penalized discriminant analysis. *Ann. Stat.***23**, No.1, 73-102.

[66] Hastie, T., Mallows, C. (1993). A discussion of A Statistical View of Some Chemometrics Regression Tools by I.E. Frank and J.H. Friedman. *Techno metrics*, **35**, 140-143.

[67] Heckman, N. E. and R. H. Zamar (2000), Comparing the shapes of regression functions. *Biometrika*, **87**, No.1, 135-144.

[68] Holmstrome, I. (1963). On a method for parametric representation of the state of the atmosphere. *Tellus*, **15**, 127-149.

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

[70] Huber, P. J. (1985). Projection pursuit. *Ann. Statist.* **13**, No.2, 435-475.

[71] Hyndman, R. J., Ullah, Md. S. (2007). Robust forecasting of mortality and fertility rates : a functional data approach. *Computational Statistics and Data Analysis*, **51**, 4942-4956.

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

[73] Kirpatrick, M., Heckman, N. (1989). A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *J. Math. Bio.* **27**, No.4, 429-450.

[74] Kneip, A., Utikal, K. J. (2001). Inference for density families using functional principal component analysis. *Journal of the American Statistical Association*, **96**, 519-542.

[75] Kolmogorov, A. N., Tikhomirov, V .M. (1959). ϵ-entropy and ϵ-capacity. Uspekhi Mat. Nauk **14**, 3-86. (Engl Transl. Amer. Math. Soc. Transl. Ser). **2**, 277-364.

[76] Kuelbs, J., Li, W. (1993). Metric entropy and the small ball problem for Gaussian measures. *J. Funct. Anal.* **116**, 133-157.

[77] 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.

[78] Lucero, J. C. (1999). A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset-offset. *J. Acoust. Soc. Am.* **105**, 423-431.

[79] 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**, No.4, 392-410.

[80] Manteiga, W. G., Vieu, P. (2007). Statistics for Functional Data. *Computational Statistics and Data Analysis*, **51**, 4788-4792.

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

[82] Molenaar, P., Boomsma, D. (1987). The genetic analysis of repeated measures: the karhunen-loeve expansion. *Behavior Genetics*, **17**, 229-242.

[83] Nadaraya, E. (1964). On estimating regression. *Theory Prob. Appl.* **10**, 186-196.

[84] Nagaev, S. V. (1997). Some refinements of probabilistic and moment inequalities. *Teor: Veroyatnost. i Primenen (in russian)*, **42**, No.4, 832-838.

[85] Nagaev, S. V. (1998). Some refinements of probabilistic and moment inequalities. *Theory. Probab. Appl.* **42**, No.4, 707-713.

[86] Nerini, D., Ghattas, B. (2007). Classifying densities using functional regression trees: Applications in oceanology. *Computational Statistics and Data Analysis*, **51**, 4984-4993.

[87] Obhukov, V. (1960). The statistically orthogonal expansion of empirical functions. *American Geophysical Union*, 288-291.

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

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

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

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

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

[93] Rao, C. R. (1958). Some statistical methods for comparing growth curves. *Biometrics*, **14**, 1-17.

[94] Ramsay, J. O. (1982). When the data are functions. *Psychometrika*, **47**, No.4, 379-396.

[95] Ramsay, J., Silverman, B. (1997). Functional Data Analysis. Springer-Verlag.

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

[97] Sarda, P., Vieu, P. (2000). Kernel Regression. Smoothing and Regression: Approaches, Computation, and Application. Ed. M.G. Schimek, 43-70, Wiley Series in Probability and Statistics.

[98] Schimek, M. (2000). Smoothing and Regression : Approaches, computation, and application, Ed. M.G. Schimek, Wiley Series in Probability and Statistics.

[99] 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**, No.6, 2493-2511.

[100] Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika*, **23**, 19-23.

[101] Van Der Vaart, A. W., Van Zanten, J. H. (2007). Bayesian inference with rescaled Gaussian process priors. *Electronic Journal of Statistics*, **1**, 433-448.

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

[103] Xia, X. An H. Z. (2002). An projection pursuit autoregression in time series. J. of Time Series Analysis. **20**, No.6, 693-714.

Chapter 2

Functional variable in single functional index model

This chapter is the object of a work subjected for publication in
International Journal of Mathematical Sciences with Computer Applications (Malaya Journal of Matematik)

2.1 Introduction

The single-index models are becoming increasingly popular because of their importance in several areas of science such as econometrics, biostatistics, medicine, financial econometric and so on. The single-index model, a special case of projection pursuit regression, has proven to be a very efficient way of coping with the high dimensional problem in nonparametric regression. Härdle *et al.* [17], Hristache *et al.* [19]. Delecroix et al. [6] have studied the estimation of the single-index approach of regression function and established some asymptotic properties. The recent literature in this domain shows a great potential of these functional statistical methods. The most popular case of functional random variable corresponds to the situation when we observe random curve on different statistical units. The first work in the fixed functional single-model was given by Ferraty *et al.* [10], where authors have obtained almost complete convergence (with the rate) of the regression function in the i.i.d. case. Their results have been extended to dependent case by Aït Saidi *et al.* [1]. Aït Saidi *et al.* [2] studied the case where the functional single-index is unknown. The authors have proposed for this parameter an estimator, based on the the cross-validation procedure.

In the present work we study a single- index modeling in the case of the functional explanatory variable. More precisely, we consider the problem of estimating some characteristics of the conditional distribution of a real variable Y given a functional variable X when the explanation of Y given X is done through its projection on one functional direction. The conditional distribution plays an important role in prediction problems, such as the conditional mode the conditional median or the conditional quantiles. Nonparametric estimation of the conditional density has been widely studied, when the data are real. The first related result in nonparametric functional statistic was obtained by Ferraty *et al.* [13], the authors have established the almost complete convergence (with rate) in the independent and identically distributed (i.i.d.) random variables. The asymptotic normality of this kernel estimator has been studied in the dependent data by Ezzahrioui and Ould Saïd [9].

The goal of this paper is to establish a nonparametric estimation of some characteristics of the conditional distribution where Kernel type estimators for the conditional cumulative distribution function and the successive derivatives of the conditional density in the single functional index model are introduced. We establish the pointwise almost complete convergence and the uniform almost complete convergence (with the rate) of the kernel estimate of this model. Asymptotic properties are stated for each of these estimates, and they are applied to the estimations of the conditional mode and conditional quantiles.
Now, let us outline the paper. At first, in section 2, we present general notations and some conditions necessary for our study, Then, in sections 3 we propose the estimator of the conditional cumulative distribution function and that of the conditional density derivatives, and we give their pointwise almost complete convergence (with rate). Then, in section 4, we study the uniform almost complete convergence of the conditional cumulative distribution function

(resp. the conditional density derivatives) estimator given in section 3. Section 5 is devoted to some applications, in this part, we first consider the problem of the estimation of the conditional mode in functional single-index model, then we investigate the asymptotic properties of the conditional quantile function of a scalar response and functional covariate when the observations are in single functional index model and data are independent and identically distributed (i.i.d.), after that the cross-validation method is given, which is so important in guarding against testing hypotheses suggested by the data, especially where further samples are hazardous, costly or impossible to collect.
In the end, we finish our paper by giving technical proofs of lemmas and corollary (Appendix).

2.2 General notations and conditions

All along the paper, when no confusion will be possible, we will denote by C, C' or/and $C_{\theta,x}$ some generic constant in $\mathbb{R}_+^*$, and in the following, any real function with an integer in brackets as exponent denotes its derivative with the corresponding order.

Let X be a functional random variable, *frv* its abbreviation. Let (X_i, Y_i) be a sample of independant pairs, each having the same distribution as (X, Y), our aim is to build nonparametric estimates of several functions related with the conditional probability distribution (*cond-cdf*) of Y given $< X, \theta > = < x, \theta >$.
Let

$$\forall y \in \mathbb{R},\ F(\theta, y, x) = \mathbb{P}(Y \le y| < X, \theta > = < x, \theta >).$$

be the *cond-cdf* of Y given $< X, \theta > = < x, \theta >$, for $x \in \mathcal{H}$, which also shows the relationship between X and Y but is often unknown.
If this distribution is absolutely continuous with respect to the Lebesgues measure on $\mathbb{R}$, then we will denote by $f(\theta, \cdot, x)$. (*resp.* $f^{(j)}(\theta, \cdot, x)$) the conditional density (*resp.* its j^{th} order derivative) of Y given $< X, \theta > = < x, \theta >$. In Sections 3 and 4, we will give almost complete convergence[1] results (with rates of convergence[2]) for nonparametric estimates of both functions $F(\theta, \cdot, x)$ and $f^{(j)}(\theta, \cdot, x)$.

In the following, for any $x \in \mathcal{H}$ and $y \in \mathbb{R}$, let $\mathcal{N}_x$ be a fixed neighborhood of x in $\mathcal{H}$, $\mathcal{S}_\mathbb{R}$ will be a fixed compact subset of $\mathbb{R}$, and we will use the notation $B_\theta(x, h) = \{X \in \mathcal{H}/0 < | < x - X, \theta > | < h\}$. Our nonparametric models will be quite general in the sense that we will just need the following simple assumption for the marginal distribution of $< \theta, X >$:

(H1) $\mathbb{P}\left(X \in B_\theta(x, h)\right) = \phi_{\theta,x}(h) > 0$,

[1] Recall that a sequence $(T_n)_{n\in\mathbb{N}}$ of random variables is said to converge almost completely to some variable T, if for any $\epsilon > 0$, we have $\sum_n \mathbb{P}(|T_n - T| > \epsilon) < \infty$. This mode of convergence implies both almost sure and in probability convergence (see for instance Bosq and Lecoutre, 1987).

[2] Recall that a sequence $(T_n)_{n\in\mathbb{N}}$ of random variables is said to be of order of complete convergence u_n, if there exists some $\epsilon > 0$ for which $\sum_n \mathbb{P}(|T_n| > \epsilon u_n) < \infty$. This is denoted by $T_n = O(u_n)$, *a.co.* (or equivalently by $T_n = O_{a.co.}(u_n)$).

together with some usual smoothness conditions on the function to be estimated. According to the type of estimation problem to be considered, we will assume either

(H2) $\forall(y_1, y_2) \in \mathcal{S}_\mathbb{R} \times \mathcal{S}_\mathbb{R}$, $\forall(x_1, x_2) \in \mathcal{N}_x \times \mathcal{N}_x$, $|F(\theta, y_1, x_1) - F(\theta, y_2, x_2)| \leq C_{\theta,x}\left(\|x_1 - x_2\|^{b_1} + |y_1 - y_2|^{b_2}\right)$, $b_1 > 0$, $b_2 > 0$,

(H3) $\forall(y_1, y_2) \in \mathcal{S}_\mathbb{R} \times \mathcal{S}_\mathbb{R}$, $\forall(x_1, x_2) \in \mathcal{N}_x \times \mathcal{N}_x$, $|f^{(j)}(\theta, y_1, x_1) - f^{(j)}(\theta, y_2, x_2)| = C_{\theta,x}\left(\|x_1 - x_2\|^{b_1} + |y_1 - y_2|^{b_2}\right)$, $b_1 > 0$, $b_2 > 0$.

2.3 Pointwise almost complete estimation

In this section we give the pointwise almost complete estimation (with rate) of the conditional cumulative distribution as of the successive derivatives of the conditional density.

2.3.1 Conditional cumulative distribution estimation

The purpose of this section is to estimate the *cond-cdf* $F^x(\theta, \cdot, x)$. We introduce a kernel type estimator $\widehat{F}^x(\theta, \cdot, x)$ of $F^x(\theta, \cdot, x)$ as follows:

$$\widehat{F}(\theta, y, x) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1}(< x - X_i, \theta >)\right) H\left(h_H^{-1}(y - Y_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}(< x - X_i, \theta >)\right)}, \tag{2.1}$$

where K is a kernel, H is a cumulative distribution function (*cdf*) and $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers which goes to zero as n tends to infinity, and with the convention $0/0 = 0$. Note that a similar estimate was already introduced in the case where X is a valued in some semi-metric space which can be of infinite dimension by Ferraty *et al.* [12]. In our single functional index context, we need the following conditions for our estimate:

(H4) H is such that, for all $(y_1, y_2) \in \mathbb{R}^2$, $|H(y_1) - H(y_2)| \leq C|y_1 - y_2|$ $\int |t|^{b_2} H^{(1)}(t)dt < \infty$,

(H5) K is a positive bounded function with support $[-1, 1]$,

(H6) $\lim_{n\to\infty} h_K = 0$ with $\lim_{n\to\infty} \frac{\log n}{n\phi_{\theta,x}(h_K)} = 0$,

(H7) $\lim_{n\to\infty} h_H = 0$ with $\lim_{n\to\infty} n^\alpha h_H = \infty$ for some $\alpha > 0$.

• **Comments on the assumptions**

Our assumptions are very standard for this kind of model. Assumptions (H1) and (H5) are the same as those given in Ferraty et al. [10]. Assumptions (H2) and (H3) is a regularity conditions which characterize the functional space of our model and is needed to evaluate the bias term of our asymptotic results. Assumptions (H4) and (H6)-(H7) are technical conditions and are also similar to those done in Ferraty *et al.* [13].

Theorem 2.3.1 *Under the hypotheses (H1), (H2) and (H4)-(H7), and for any fixed y, we have*

$$|\widehat{F}(\theta,y,x)-F(\theta,y,x)| \;=\; O(h_K^{b_1})+O(h_H^{b_2})+O\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right), \qquad a.co. \tag{2.2}$$

Proof. For $i=1,\dots,n$, we consider the quantities $K_i(\theta,x) := K(h_K^{-1}(< x-X_i,\theta >))$ and, for all $y\in\mathbb{R}$ $H_i(y) = H\left(h_H^{-1}(y-Y_i)\right)$ and let $\widehat{F}_N(\theta,y,x)$ (resp. $\widehat{F}_D(\theta,x)$) be defined as

$$\widehat{F}_N(\theta,y,x) \;=\; \frac{1}{n\,\mathbb{E}(K_1(\theta,x))}\sum_{i=1}^{n}K_i(\theta,x)H_i(y) \qquad (\text{resp. } \widehat{F}_D(\theta,x) \;=\; \frac{1}{n\,\mathbb{E}(K_1(\theta,x))}\sum_{i=1}^{n}K_i(\theta,x)).$$

This proof is based on the following decomposition

$$\begin{aligned}\widehat{F}(\theta,y,x)-F(\theta,y,x) \;&=\; \frac{1}{\widehat{F}_D(\theta,x)}\left\{\left(\widehat{F}_N(\theta,y,x)-\mathbb{E}\widehat{F}_N(\theta,y,x)\right)-\left(F(\theta,y,x)-\mathbb{E}\widehat{F}_N(\theta,y,x)\right)\right\}\\ &\quad+\frac{F(\theta,y,x)}{\widehat{F}_D(\theta,x)}\left\{1-\widehat{F}_D(\theta,x)\right\}\end{aligned} \tag{2.3}$$

and on the following intermediate results.

Lemma 2.3.1 *Under the hypotheses (H1) and (H5)-H6), we have*

$$|\widehat{F}_D(\theta,x)-1| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right), \tag{2.4}$$

Corollary 2.3.1 *Under the hypotheses of Lemma 2.3.1, we have*

$$\sum_{n=1}^{\infty}\mathbb{P}\left(|\widehat{F}_D(\theta,x)|\le 1/2\right) \;<\; \infty. \tag{2.5}$$

Lemma 2.3.2 *Under the hypotheses (H1), (H2) and (H4)-(H.6), we have*

$$|F(\theta,y,x)-\mathbb{E}\widehat{F}_N(\theta,y,x)| = O\left(h_K^{b_1}\right)+O\left(h_H^{b_2}\right), \tag{2.6}$$

Lemma 2.3.3 *Under the hypotheses (H1), (H2) and (H4)-(H7), we have*

$$|\widehat{F}_N(\theta,y,x)-\mathbb{E}\widehat{F}_N(\theta,y,x)| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_{\theta,x}(h_K)}}\right), \tag{2.7}$$

■

2.3.2 Estimating successive derivatives of the conditional density

The main objectif of this part is the estimation of successive derivatives of the conditional density of Y given $< X, \theta >=< x, \theta >$, denoted by $f(\theta, \cdot, x)$. It is well known that, in nonparametric statistics, this latter provides an alternative approach to study the links between Y and X and it can be also used, in single index modelling, to estimate the functional index θ if it is unknown.

So, at first, we propose to define the estimator $\widehat{f}^{(j)}(\theta, y, x)$ of $f^{(j)}(\theta, y, x)$ as follows:

$$\widehat{f}^{(j)}(\theta, y, x) = \frac{h_H^{-1-j} \sum_{i=1}^{n} K(h_K^{-1}(< x - X_i, \theta >)) H^{(j+1)}(h_H^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(h_K^{-1}(< x - X_i, \theta >))}, \quad y \in \mathbb{R} \tag{2.8}$$

Similar estimate was already introduced in the case where X is a valued in some semi-metric space which can be of infinite dimension; Ferraty *et al.* [12], then widely studied (see for instance by Attaoui *et al.* [3], for several asymptotic results and references). In addition to the conditions introduced along the previous section, we need the following ones, which are technical conditions and are also similar to those given in Ferraty *et al.* [13]:

$$\text{(H8)} \quad \begin{cases} \forall (y_1, y_2) \in \mathbb{R}^2, \ |H^{(j+1)}(y_1) - H^{(j+1)}(y_2)| \leq C_{\theta,x}|y_1 - y_2| \\ \exists \nu > 0, \ \forall j' \leq j+1, \ \lim_{y\to\infty} |y|^{1+\nu} \ \left|H^{(j'+1)}(y)\right| = 0. \end{cases}$$

$$\text{(H9)} \quad \lim_{n\to\infty} h_K = 0 \text{ with } \lim_{n\to\infty} \frac{\log n}{n h_H^{2j+1} \phi_{\theta,x}(h_K)} = 0.$$

The next result concerns the asymptotic behaviour of the kernel functional estimator $\widehat{f}^{(j)}(\theta, \cdot, x)$ of the j^{th} order derivative of the conditional density function.

Theorem 2.3.2 *Under Assumptions (H1), (H3)-(H5), and (H7)-(H9), and for any fixed y, we have, as n goes to infinity*

$$|\widehat{f}^{(j)}(\theta, y, x) - f^{(j)}(\theta, y, x)| = O\left(h_K^{b_1}\right) + O\left(h_H^{b_2}\right) + O\left(\sqrt{\frac{\log n}{n h_H^{2j+1} \phi_{\theta,x}(h_K)}}\right) \quad a.c.o \tag{2.9}$$

Proof. This result is based on the same kind of decomposition as (3.25). Indeed, we can write:

$$\begin{aligned}\widehat{f}^{(j)}(\theta,y,x)-f^{(j)}(\theta,y,x) &= \frac{1}{\widehat{F}_D(\theta,x)}\left(\widehat{f}_N^{(j)}(\theta,y,x)-\mathbb{E}(\widehat{f}_N^{(j)}(\theta,y,x)\right)-\frac{1}{\widehat{F}_D(\theta,x)}\\ &\quad \left(f^{(j)}(\theta,y,x)-\mathbb{E}\widehat{f}_N^{(j)}(\theta,y,x)\right) \qquad (2.10)\\ &\quad +\frac{f^{(j)}(\theta,y,x)}{\widehat{F}_D(\theta,x)}\left(1-\widehat{F}_D(\theta,x)\right)\end{aligned}$$

where

$$\widehat{f}_N^{(j)}(\theta,y,x)=\frac{1}{n\,h_H^{j+1}\,\mathbb{E}(K_1(\theta,x))}\sum_{i=1}^{n}K_i(\theta,x)H_i^{(j+1)}(y).$$

Then, Theorem 2.3.2 can be deduced from both following lemmas, together with Lemma 2.3.1 and Corollary 2.3.1.

Lemma 2.3.4 *Under the hypotheses (H1), (H2), (H3), (H5) and (H6) we have*

$$|f^{(j)}(\theta,y,x)-\mathbb{E}(\widehat{f}_N^{(j)}(\theta,y,x)|=O\left(h_K^{b_1}\right)+O\left(h_H^{b_2}\right),$$

Lemma 2.3.5 *Under the hypotheses (H1)-(H7), we have*

$$|\widehat{f}_N^{(j)}(\theta,y,x)-\mathbb{E}(\widehat{f}_N^{(j)}(\theta,y,x)|=O_{a.co.}\left(\sqrt{\frac{\log n}{n\,h_H^{2j+1}\,\phi_{\theta,x}(h_K)}}\right),$$

The proofs of the the above lemmas and corollary are given in the same manner as it was done in [13], since they are a special case of the Lemmas 2.3.2, 2.3.3, 2.3.4 and 2.3.5. It suffices to repalce $\widehat{f}^{(j)}(y,x)$ (resp. $f^{(j)}(y,x)$) by $\widehat{f}^{(j)}(\theta,y,x)$ (resp. $f^{(j)}(\theta,y,x)$), and $\widehat{F}_D(x)$, (resp. $F_D(x)$) by $\widehat{F}_D(\theta,x)$ (resp. $F_D(\theta,x)$) with $d(x_1,x_2)=<x_1-x_2,\theta>$ ■

2.4 Uniform almost complete convergence

In this section we derive the uniform version of Theorem 2.3.1 and Theorem 2.3.2. The study of the uniform consistency is motivated by the fact that the latter is an indispensable tool for studying the asymptotic properties of all estimates of the functional index if is unknown. Noting that, in the multivariate case, the uniform consistency is a standard extension of the pointwise one, however, in our functional case, it requires some additional tools and topological conditions (see Ferraty *et al.*, 2009, 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

$$\mathcal{S}_{\mathcal{H}}\subset\bigcup_{k=1}^{d_n^{\mathcal{S}_{\mathcal{H}}}}B(x_k,r_n)\ \text{and}\ \Theta_{\mathcal{H}}\subset\bigcup_{j=1}^{d_n^{\Theta_{\mathcal{H}}}}B(t_j,r_n) \qquad (2.11)$$

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.

2.4.1 Conditional cumulative distribution estimation

In this section we propose to study the uniform almost complete convergence of our estimator defined above (2.1) for this, we need the following assumptions:

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

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

(A2) $\forall (y_1, y_2) \in \mathcal{S}_{\mathbb{R}} \times \mathcal{S}_{\mathbb{R}}, \forall (x_1, x_2) \in \mathcal{S}_{\mathcal{H}} \times \mathcal{S}_{\mathcal{H}}$ and $\forall \theta \in \Theta_{\mathcal{H}}$,

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

(A3) The kernel K satisfy (H3) and Lipschitz's condition holds

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

(A4) For $r_n = O\left(\frac{\log n}{n}\right)$ 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 $\beta > 1$.

Remark 2.4.1 *Note that Assumptions (A1) and (A2) are, respectively, the uniform version of (H1) and (H2). Assumptions (A1) and (A4) are linked with the the topological structure of the functional variable, see Ferraty* et al. *[14].*

Theorem 2.4.1 *Under Assumptions (A1)-(A4) and (H4), as n goes to infinity, we have*

$$\sup_{\theta \in \Theta_{\mathcal{H}}} \sup_{x \in \mathcal{S}_{\mathcal{H}}} \sup_{y \in \mathcal{S}_{\mathbb{R}}} |\widehat{F}(\theta, y, x) - F(\theta, y, x)| = O(h_K^{b_1}) + O(h_H^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}} \right). \tag{2.12}$$

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

Corollary 2.4.1 *Under Assumptions (A1)-(A4) and (H4), as n goes to infinity, we have*

$$\sup_{x \in \mathcal{S}_{\mathcal{H}}} \sup_{y \in \mathcal{S}_{\mathbb{R}}} |\widehat{F}(\theta, y, x) - F(\theta, y, x)| = O(h_K^{b_1}) + O(h_H^{b_2}) + O_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}}{n\phi(h_K)}} \right). \tag{2.13}$$

Clearly The proofs of these two results namely the Theorem 2.4.1 and Corollary 2.4.1 can be deduced from the following intermediate results which are only uniform version of Lemmas 2.3.1-2.3.3 and Corollary 2.3.1.

Lemma 2.4.1 *Under Assumptions (A1), (A3) and (A4), we have as $n \to \infty$*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}|\widehat{F}_D(\theta,x)-1| = O_{a.co}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}+\log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

Corollary 2.4.2 *Under the assumptions of Lemma 2.4.1, we have,*

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

Lemma 2.4.2 *Under Assumptions (A1), (A2) and (H4), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}|F(\theta,y,x)-\mathbb{E}(\widehat{F}_N(\theta,y,x))| = O(h_K^{b_1})+O(h_H^{b_2}). \tag{2.14}$$

Lemma 2.4.3 *Under the assumptions of Theorem 2.4.1, we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathcal{H}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}|\widehat{F}_N(\theta,y,x)-\mathbb{E}\left[\widehat{F}_N(\theta,y,x)\right]| = O_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}+log\ d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right)$$

2.4.2 Estimating successive derivatives of the conditional density

In this part we focus on the study of uniform almost complete convergence of our estimator defined above (2.8). Thus, in addition to the conditions introduced in the section 4, we need the following ones.

(A5) $\forall(y_1,y_2)\in\mathcal{S}_{\mathbb{R}}\times\mathcal{S}_{\mathbb{R}}, \forall(x_1,x_2)\in\mathcal{S}_{\mathcal{F}}\times\mathcal{S}_{\mathcal{F}}$ and $\forall\theta\in\Theta_{\mathcal{F}}$,

$$|f^{(j)}(\theta,y_1,x_1)-f^{(j)}(\theta,y_2,x_2)|\le C\left(\|x_1-x_2\|^{b_1}+|y_1-y_2|^{b_2}\right),$$

(A6) For some $\gamma\in(0,1)$, $\lim_{n\to\infty} n^{\gamma}h_H=\infty$, and for $r_n = O\left(\frac{\log n}{n}\right)$ the sequences $d_n^{\mathcal{S}_{\mathcal{F}}}$ and $d_n^{\Theta_{\mathcal{F}}}$ satisfy:

$$\frac{(log\ n)^2}{nh_H^{2j+1}\phi(h_K)} < log\ d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}} < \frac{nh_H^{2j+1}\phi(h_K)}{\log n},$$

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 $\beta>1$.

Theorem 2.4.2 *Under Hypotheses (A1), (A3) ,(A5)-(A6) and (H8), as n goes to infinity, we have*

$$\sup_{\theta\in\Theta_{\mathcal{F}}}\sup_{x\in\mathcal{S}_{\mathcal{F}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}|\widehat{f}^{(j)}(\theta,y,x)-f^{(j)}(\theta,y,x)| = O(h_K^{b_1})+O(h_H^{b_2})+O_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{nh_H^{2j+1}\phi(h_K)}}\right). \tag{2.15}$$

Proof. This result is based on the same kind of decomposition (2.10), therefore, Theorem 2.4.2 can be deduced from both following lemmas, together with Lemma 2.4.1 and Corollary 2.4.2.

Lemma 2.4.4 *Under Assumptions (A1), (A5) and (H8), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{F}}}\sup_{x\in\mathcal{S}_{\mathcal{F}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}|f^{(j)}(\theta,y,x)-\mathbb{E}(\widehat{f}_N^{(j)}(\theta,y,x))| = O(h_K^{b_1})+O(h_H^{b_2}).$$

Lemma 2.4.5 *Under the assumptions of Theorem 2.4.2, we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{F}}}\sup_{x\in\mathcal{S}_{\mathcal{F}}}\sup_{y\in\mathcal{S}_{\mathbb{R}}}\left|\widehat{f}_N^{(j)}(\theta,y,x)]-\mathbb{E}\left[\widehat{f}_N^{(j)}(\theta,y,x)\right]\right| = O_{a.co.}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}+\log d_n^{\Theta_{\mathcal{F}}}}{nh_H^{2j+1}\phi_{\theta,x}(h_K)}}\right).$$

■

2.5 Applications

2.5.1 The conditional mode in functional single-index model

In this section we will consider the problem of the estimation of the conditional mode in the functional single-index model. Our main aim, here, is to establish the a.co. convergence of the kernel estimator of the conditional mode of Y given $< X,\theta >=< x,\theta >$ denoted by $M_\theta(x)$, uniformly on fixed subset $\mathcal{S}_{\mathcal{H}}$ of $\mathcal{H}$. For this, we assume that $M_\theta(x)$ satisfies on $\mathcal{S}_{\mathcal{H}}$ the following uniform uniqueness property (see, Ould-saïd and Cai [110], for the multivariate case).

(A6) $\forall\varepsilon_0>0$, $\exists\eta>0$, $\forall\varphi:\mathcal{S}_{\mathcal{H}}\longrightarrow\mathcal{S}_{\mathbb{R}}$,

$$\sup_{x\in\mathcal{S}_{\mathcal{H}}}|M_\theta(x)-\varphi(x)|\geq\varepsilon_0\Longrightarrow\sup_{x\in\mathcal{S}_{\mathcal{H}}}|f(\theta,\varphi(x),x)-f(\theta,M_\theta(x),x)|\geq\eta.$$

We estimate the conditional mode $\widehat{M_\theta}(x)$ with a random variabl M_θ such that

$$\widehat{M_\theta}(x)=\arg\sup_{y\in\mathcal{S}_{\mathbb{R}}}\widehat{f}(\theta,y,x). \tag{2.16}$$

Note that the estimate $\widehat{M_\theta}$ is not necessarily unique, and if this is the case all the remaining of our paper will concern any value $\widehat{M_\theta}$ satisfying (2.16). The difficulty of the problem is naturally linked with the flatness of the function

$f(\theta, y, x)$ around the mode M_θ. This flatness can be controled by the number of vanishing derivatives at point M_θ, and this parameter will also have a great influence on the asymptotic rates of our estimates. More precisely, we introduce the following additional smoothness condition.

$$(A7)\begin{cases} f^{(l)}(\theta, M_\theta(x), x) = 0, & \text{if;} 1 \leq l < j \\ \text{and } f^{(j)}(\theta, \cdot, x), & \text{is uniformly continuous on } \mathcal{S}_{\mathbb{R}} \\ \text{such that,} & |f^{(j)}(\theta, \cdot, x)| > C > 0 \end{cases}$$

Theorem 2.5.1 *Under the assumptions of Theorem 2.4.2 hold together with (A6)-(A7) we have*

$$\sup_{x\in\mathcal{S}_{\mathcal{H}}} |\widehat{M_\theta}(x) - M_\theta(x)| = O(h_K^{\frac{b_1}{j}}) + O(h_H^{\frac{b_2}{j}}) + O_{a.co.}\left(\left(\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}}{n^{1-\gamma}\phi(h_K)}\right)^{\frac{1}{2j}}\right)$$

Let us now define the application framework of our results to prediction problem by applying the result in the above Theorem, we obtain the following result.

Corollary 2.5.1 *Under the assumptions of Theorem 3.15, we have as n goes to infinity*

$$\widehat{M_\theta}(x) - M_\theta(x) \longrightarrow 0 \quad a.co.$$

2.5.2 Conditional quantile in functional single-index model

In this part of paper we investigate the asymptotic properties of the conditional quantile function of a scalar response and functional covariate when the observations are from a single functional index model and data are independent and identically distributed (i.i.d.)

We will consider the problem of the estimation of the conditional quantiles. Saying that, we are implicitely assuming the existence of a regular version for the conditional distribution of Y given $< X, \theta >$. Now, let $t_\theta(\alpha)$ be the α-order quantile of the distribution of Y given $< X, \theta >=< x, \theta >$. From the *cond-cdf* $F(\theta, \cdot, x)$, it is easy to give the general definition of the α-order quantile:

$$t_\theta(\alpha) = \inf\{t \in \mathbb{R} : \ F(\theta, t, x) \geq \alpha\}, \qquad \forall \alpha \in (0, 1).$$

In order to simplify our framework and to focus on the main interest of our part (the functional feature of $< X, \theta >$), we assume that $F(\theta, \cdot, x)$ is strictly increasing and continuous in a neighborhood of $t_\theta(\alpha)$. This is insuring unicity of the conditional quantile $t_\theta(\alpha)$ which is defined by:

$$t_\theta(\alpha) \; = \; F^{-1}(\theta, \alpha, x). \tag{2.17}$$

In what remains, we wish to stay in a free distribution framework. This will lead to assume only smoothness restrictions for the *cond-cdf* $F(\theta, \cdot, x)$ through nonparametric modelling (see Section 2).

As by-product of (2.17) and (2.1), it is easy to derive an estimator $\widehat{t}_\theta(\alpha)$ of $t_\theta(\alpha)$:

$$\widehat{t}_\theta(\alpha) \;=\; \widehat{F}^{-1}(\theta, \alpha, x). \tag{2.18}$$

As we will see later on, such an estimator is unique as soon as H is an increasing continuous function.
Naturally, we will estimate this quantile by mean of the conditional distribution estimator studied in previous sections. Here also, as far as we know, the literature on (conditional and/or unconditional) quantile estimation is quite important when the explanatory variable X is real (see for instance Samanta, 1989, for previous results and Berlinet *et al.*, 2001, for recent advances and references). In the functional case, the conditional quantiles for scalar response and a scalar/multivariate covariate have received considerable interest in the statistical literature. For completely observed data, several nonparametric approaches have been proposed, for instance, Gannoun et al., (2003) introduced a smoothed estimator based on double kernel and local constant kernel methods and Berlinet et al., (2001) established its asymptotic normality. Under random censoring, Gannoun et al., (2005) introduced a local linear (LL) regression (see Koenker and Bassett (1978) for the definition) and El Ghouch and Van Keilegom (2009) studied the same LL estimator. Ould-Saïd (2006) constructed a kernel estimator of the conditional quantile under independent and identically distributed (i.i.d.) censorship model and established its strong uniform convergence rate. Liang and De Uña-Àlvarez (2011) established the strong uniform convergence (with rate) of the conditional quantile function under α-mixing assumption.
Recently, many authors are interested in the estimation of conditional quantiles for a scalar response and functional covariate. Ferraty *et al.*, (2005) introduced a nonparametric estimator of conditional quantile defined as the inverse of the conditional cumulative distribution function when the sample is considered as an α-mixing sequence. They stated its rate of almost complete consistency and used it to forecast the well-known El Niño time series and to build confidence prediction bands. Ezzahrioui and Ould-Saïd (2008) established the asymptotic normality of the kernel conditional quantile estimator under α-mixing assumption. Recently, and within the same framework, Dabo-Niang and Laksaci (2012) provided the consistency in L^p norm of the conditional quantile estimator for functional dependent data.

So, in this work we propose to estimate $t_\theta(\alpha)$ by the estimate $\widehat{t}_\theta(\alpha)$ defined as (2.18) or as

$$\widehat{F}(\theta, \widehat{t}_\theta(\alpha), x) = \alpha. \tag{2.19}$$

To insure existence and unicity of this quantile, we will assume that

(A8) $F(\theta, \cdot, x)$ is strictly increasing,

Note that, because H is a *cdf* satisfying (H4), such a value $\widehat{t}_\theta(\alpha)$ is always existing. It could be the case that it is not unique, but if this happens all the remaining of the paper will concern any among all the values $\widehat{t}_\theta(\alpha)$ satisfying (2.19).

In order to insure unicity of $\widehat{t}_\theta(\alpha)$ we will make the following, quite unrestrictive, assumption:

(A9) H is is strictly increasing,

As for the mode estimation problem discussed before, the difficulty occur in estimating the conditional quantile $t_\theta(\alpha)$ is linked with the flatness of the curve of the conditional distribution $F(\theta,\cdot,x)$ around $t_\theta(\alpha)$. More precisely, we will suppose that there exists some integer $j>0$ such that:

$$(A10)\left\{\begin{array}{ll} F^{(l)}(\theta,t_\theta(\alpha),x)=0, & \text{if; } 1\le l<j\\ \text{and } F^{(j)}(\theta,\cdot,x), & \text{is uniformly continuous on; } \mathcal{S}_\mathbb{R}\\ \text{such that,} & |F^{(j)}(\theta,t_\theta(\alpha),x)|>C>0\end{array}\right.$$

Theorem 2.5.2 *If the conditions of Theorem 2.4.2 hold together with (A8)-(A10), we have*

$$\sup_{x\in\mathcal{S}_\mathcal{H}}|\widehat{t}_\theta(\alpha)-t_\theta(\alpha)|=O\left(h_K^{\frac{b_1}{j}}+h_H^{\frac{b_2}{j}}\right)+O\left(\left(\frac{\log d_n^{\mathcal{S}_\mathcal{H}}}{n\ \phi_x(h_K)}\right)^{\frac{1}{2j}}\right),\qquad a.co. \tag{2.20}$$

Proof. Let us write the following Taylor expansion of the function $\widehat{F}(\theta,\cdot,x)$:

$$\begin{aligned}\widehat{F}(\theta,t_\theta(\alpha),x)-\widehat{F}(\theta,\widehat{t}_\theta(\alpha),x) &= \sum_{l=1}^{j-1}\frac{(t_\theta(\alpha)-\widehat{t}_\theta(\alpha))^l}{l!}\widehat{F}^{(l)}(\theta,t_\theta(\alpha),x)\\ &+ \frac{(t_\theta(\alpha)-\widehat{t}_\theta(\alpha))^j}{j!}\widehat{F}^{(j)}(\theta,t^*,x),\end{aligned}$$

where t^* is some point between $t_\theta(\alpha)$ and $\widehat{t}_\theta(\alpha)$. It suffices now to use the first part of condition (A10) to be able to rewrite this expression as:

$$\begin{aligned}\widehat{F}(\theta,t_\theta(\alpha),x)-\widehat{F}(\theta,\widehat{t}_\theta(\alpha),x) &= \sum_{l=1}^{j-1}\frac{(t_\theta(\alpha)-\widehat{t}_\theta(\alpha))^l}{l!}\left(\widehat{f}^{(l-1)}(\theta,t_\theta(\alpha),x)-f^{(l-1)}(\theta,t_\theta(\alpha),x)\right)\\ &+ \frac{(t_\theta(\alpha)-\widehat{t}_\theta(\alpha))^j}{j!}\widehat{f}^{(j-1)}(\theta,t^*,x),\end{aligned}$$

As long as we could be able to check that

$$\exists\tau>0,\ \sum_{j=1}^{n=\infty}\mathbb{P}\left(f^{(j-1)}(\theta,t^*,x)<\tau\right)\ <\ \infty, \tag{2.21}$$

we would have

$$(t_\theta(\alpha) - \widehat{t}_\theta(\alpha))^j = O\left(\widehat{F}(\theta, t_\theta(\alpha), x) - F(\theta, t_\theta(\alpha), x)\right) + O\left(\sum_{l=1}^{j-1}(t_\theta(\alpha) - \widehat{t}_\theta(\alpha))^l(\widehat{f}^{(l-1)}(\theta, t_\theta(\alpha), x) - f^{(l-1)}(\theta, t_\theta(\alpha), x)\right), \text{ a.co.} \tag{2.22}$$

By comparing the rates of convergence given in Theorems 2.4.1 and 3.14, we see that the leading term in right hand side of equation (2.22) is the first one. So we have

$$(t_\theta(\alpha) - \widehat{t}_\theta(\alpha))^j = O_{a.co.}\left(\widehat{F}(\theta, t_\theta(\alpha), x) - F(\theta, t_\theta(\alpha), x)\right),$$

Because of Theorem 3.14, this is enough to get the claimed result, and so (2.21) is the only result that remains to check. This will be done directly by using the uniform continuity of the function $f^{(j-1)}(\theta, \cdot, x)$ given by second part of (A10) together with the third part of (A7) and with the following lemma.

Lemma 2.5.1 *If the conditions of Theorem 2.4.1 hold together with (A8) and (A9), then we have:*

$$\widehat{t}_\theta(\alpha) - t_\theta(\alpha) \rightarrow 0, \text{ a.co.} \tag{2.23}$$

■

2.5.3 The cross-validation method

This part is devoted to another type of application called the cross-validation method, mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice. This method is widely applied, it can be used to compare the performances of different predictive modeling procedures. For instance, in optical character recognition; a mechanical or electronic conversion of scanned or photographed images of typewritten or printed text into machine-encoded/computer-readable text, this later is widely used as a form of data entry from some sort of original paper data source, whether passport documents, invoices, bank statement, receipts, business card, mail, or any number of printed records. It can also be used in variable selection; the process of selecting a subset of relevant features for use in model construction.

After this short introduction let's give an application of the method:

1. The regression operator $\widehat{r}_\theta(x)$ depends on the functional parameter θ, So, a crucial question arises: how to choose the functional index θ? The answer is nontrivial and a firstway consists in extending the standard cross-validation procedure to our functional context. For this, one considers various quadratic distances, namely the averaged squared error

$$\text{ASE}(\theta) = n^{-1}\sum_{j=1}^{n}\left(r_{\theta_0}(X_j) - \widehat{r}_\theta(X_j)\right)^2, \tag{2.24}$$

the integrated squared error

$$\text{ISE}(\theta) = \mathbb{E}\left[(r\theta_0(X_0) - \widehat{r}_\theta(X_0))^2 \,|Z_1, \ldots, Z_n\right], \tag{2.25}$$

and the mean integrated squared error

$$\text{MISE}(\theta) = \mathbb{E}\left[ISE(\theta)\right]. \tag{2.26}$$

The main goal consists in finding a θ which minimizes (in some sense) over Θ_n these quantities. However, because all these quadratic distances depend on the unknown regression operator r_{θ_0}, the criterion used in practice for choosing θ is

$$CV(\theta) = n^{-1} \sum_{j=1}^{n} \left(Y_j - \widehat{r}_\theta^{-j}(X_j)\right)^2 \tag{2.27}$$

where $\widehat{r}_\theta^{-j}$ is the leave-one-out estimate of $r_\theta(x)$, given by

$$\widehat{r}_\theta^{-j}(x) = \frac{(n-1)^{-1} \displaystyle\sum_{\substack{i=1 \\ i \neq j}}^{n} Y_i K\left(h_K^{-1}(< x - X_i, \theta >)\right)}{(n-1)^{-1} \displaystyle\sum_{\substack{i=1 \\ i \neq j}}^{n} K\left(h_K^{-1}(< x - X_i, \theta >)\right)}. \tag{2.28}$$

So, the selection rule will be to choose θ_{CV} which minimizes the so-called cross-validation criterion $CV(\theta)$. Clearly, for a given θ, $CV(\theta)$ is a computable quantity. It measures a quadratic distance between $(Y_1, \ldots, Y_n)$ and its prediction $\widehat{r}_\theta^{-j}(X_1), \ldots, \widehat{r}_\theta^{-j}(X_n)$ when, for each i, $\widehat{r}_\theta^{-i}(.)$ is built without the ith data (X_i, Y_i). So, the method of cross-validation consists in choosing among several candidates θ, the one who is the most adapted to our data set (X_i, Y_i) in terms of prediction. This method is inspired by the cross-validation ideas which have been proposed in various standard nonparametric estimation problems (see [79] for the regression problem, [102] for the density and [131] for the hazard function).

From a practical point of view, some questions arise in order to implement this single-functional index model. What about the identifiability of the model given a sample of observed curves $(x_1, \ldots, x_n)$? How to build the set of functional indexes $\Theta_{\mathcal{F}}$? What about the choice of the bandwidth h?

Emphasizes the good behaviour of this simple cross-validated procedure, even in pathological situations. To see that, one focuses on a favourable case (i.e. $\theta_0 \in \Theta_{\mathcal{F}}$).

First of all, one builds a sample of n curves curves as follows:

$$x_i(t_j) = a_i cos(2\pi t_j) + b_i sin(4\pi t_j) + 2c_i(t_j - 0.25)(t_j - 0.5),$$

where $0 = t_1 < t_2 < \ldots < t_{n-1} < t_n = 1$ are equispaced points, the a_i's, b_i's and c_i's being independent observations uniformly distributed on $[0, 1]$. Once the curves are defined, one simulates a single-functional index model as follows:

- Choose one $\theta_0(\cdot)$.
- Choose one link function $r(\cdot)$.
- Compute the inner products $< \theta_0, x_1 >, \ldots, < \theta_0, x_n >$.
- Generate independently $\varepsilon_1, \ldots, \varepsilon_n$, from a centred Gaussian of variance equal to 0.05 times the empirical variance of $r(< \theta_0, x_1 >), \ldots, r(< \theta_0, x_n >)$ (i.e. signal-to-noise ratio = 0.05).
- Simulate the corresponding responses: $Y_i = r(< \theta_0, x_i >) + \varepsilon_i$.

Finally, the observations $(x_k, Y_k)_{k=1,\ldots,m}$ are used for the learning step and the others (i.e. $(x_l, Y_l)_{l=m+1,\ldots,n}$ allow the computation of the mean square error of prediction:

$$\text{MSEP} = \frac{1}{n-m} \sum_{j=n-m}^{n} \left(Y_j \widehat{r}\,(< \theta_{CV}, x_j >)\right)^2 .$$

In order to highlight the specificity of such a single-functional index model, the obtained predictions are compared with those coming from a pure nonparametric functional data analysis (NPFDA) method (see [13] for details and references therein). Actually, the NPFDA regression method uses the following kernel estimator:

$$\forall x \in \mathcal{H}, \quad \widehat{r}(x) = \frac{\displaystyle\sum_{i=1}^{n} Y_i K\left(h^{-1}(d(X_i, x)\right)}{\displaystyle\sum_{i=1}^{n} K\left(h^{-1}(d(X_i, x)\right)} \tag{2.29}$$

for estimating the regression operator m in the nonparametric model $Y_i = r(X_i) + \varepsilon_i$, for all $i = 1, \ldots, n$, where $d(.,.)$ is a fixed semi-metric.

If one looks at the NPFDA kernel estimator (2.29), it suffices to replace the fixed semi-metric $d(\cdot,\cdot)$ with $d_{\theta_{CV}}(\cdot,\cdot)$. What does this mean? It means that the functional index model can be seen as one way of building an nonparametric functional data analysis (NPFDA) kernel estimator with a data-driven semi-metric. In particular, in pure nonparametric functional models when one has no idea of the semi-metric, the functional index model appears to be a method for performing an adaptative one. The functional index model makes the NPFDA method more flexible. In this sense, the functional index model is not a competitive statistical technique with respect to the NPFDA method, but rather a complementary one.

2. If we wish to predict a real characteristic denoted Y of X_n knowing the curve X_{n-1}, we have to consider the observations (X_i, y_i) where y_i is the characteristic we want to provide at the instant i. For example:

   - If we want to predict the value of the process at time t_j knowing the curve X_{n-1}, we set $Y_i = X_{i+1}(t_j)$.
   - For the sup, we pose $Y_i = \sup_t X_{i+1}(t)$.
   - If we look for the time where the process reaches maximum, we set $Y_i = \arg\sup_t X_{i+1}(t)$.

   By using the conditional mode as a prediction tool, we can predict Y by $\widehat{M_\theta(X_{n-1})}$.

2.6 Appendix

Proof of Lemma 2.4.1 For all $x \in \mathcal{S}_\mathcal{H}$ and $\theta \in \Theta_\mathcal{H}$, we set

$$k(x) = \arg\min_{k\in\{1\ldots r_n\}} \|x - x_k\| \text{ and } j(\theta) =) = \arg\min_{j\in\{1\ldots l_n\}} \|\theta - t_j\|.$$

Let us consider the following decomposition

$$\begin{aligned}
\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\left|\widehat{F}_D(\theta,x) - \mathbb{E}\left(\widehat{F}_D(\theta,x)\right)\right| \leq\; & \underbrace{\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\left|\widehat{F}_D(\theta,x) - (\widehat{F}_D(\theta,x_{k(x)})\right|}_{\Pi_1} \\
& + \underbrace{\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\left|\widehat{F}_D(\theta,x_{k(x)}) - \widehat{F}_D(t_{j(\theta)},x_{k(x)})\right|}_{\Pi_2} \\
& + \underbrace{\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\left|\widehat{F}_D(t_{j(\theta)},x_{k(x)}) - \mathbb{E}\left(\widehat{F}_D(t_{j(\theta)},x_{k(x)})\right)\right|}_{\Pi_3} \\
& + \underbrace{\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\left|\mathbb{E}\left(\widehat{F}_D(t_{j(\theta)},x_{k(x)})\right) - \mathbb{E}\left(\widehat{F}_D(\theta,x_{k(x)})\right)\right|}_{\Pi_4} \\
& + \underbrace{\sup_{x\in\mathcal{S}_\mathcal{H}}\sup_{\theta\in\Theta_\mathcal{H}}\left|\mathbb{E}\left(\widehat{F}_D(\theta,x_{k(x)})\right) - \mathbb{E}\left(\widehat{F}_D(\theta,x)\right)\right|}_{\Pi_5}
\end{aligned}$$

For Π_1 and Π_2, we employe the Hölder continuity condition on K, Cauchy Schwartz's and the Bernstein's inequalities, we get

$$\Pi_1 = O\left(\sqrt{\frac{\log d_n^{S_\mathcal{H}} + \log d_n^{\Theta_\mathcal{H}}}{n\phi(h_K)}}\right), \;\; \Pi_2 = O\left(\sqrt{\frac{\log d_n^{S_\mathcal{H}} + \log d_n^{\Theta_\mathcal{H}}}{n\phi(h_K)}}\right) \quad (2.30)$$

Then, by using the fact that $\Pi_4 \leq \Pi_1$ and $\Pi_5 \leq \Pi_2$, we get for n tending to infinity

$$\Pi_4 = O\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right), \quad \Pi_5 = O\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right) \tag{2.31}$$

Now, we deal with Π_3, for all $\eta > 0$, we have

$$\mathbb{P}\left(\Pi_3 > \eta\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right)\right) \leq$$

$$d_n^{S_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}} \max_{k\in\{1\ldots d_n^{S_{\mathcal{H}}}\}} \max_{j\in\{1\ldots d_n^{\Theta_{\mathcal{H}}}\}} \mathbb{P}\left(\left|\widehat{F}_D(t_{j(\theta)}, x_{k(x)}) - \mathbb{E}\left(\widehat{F}_D(t_{j(\theta)}, x_{k(x)})\right)\right| > \eta\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right)\right).$$

Applying Bernstein's exponential inequality to

$$\frac{1}{\phi(h_K)}\left(K_i\left(t_{j(\theta)}, x_{k(x)}\right) - \mathbb{E}\left(K_i\left(t_{j(\theta)}, x_{k(x)}\right)\right)\right),$$

then under (A7), we get

$$\Pi_3 = O\left(\sqrt{\frac{\log d_n^{S_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

Lastly the result will be easily deduced from the latter together with (2.30) and (2.31).

Proof Corollary 2.4.2 It is easy to see that,
$\inf_{\theta\in\Theta_{\mathcal{H}}} \inf_{x\in\mathcal{S}_{\mathcal{H}}} |\widehat{F}_D(\theta, x)| \leq 1/2 \implies \exists x \in \mathcal{S}_{\mathcal{H}},\ \exists\theta \in \Theta_{\mathcal{H}}$, such that

$$1 - \widehat{F}_D(\theta, x) \geq 1/2 \implies \sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in\mathcal{S}_{\mathcal{H}}} |1 - \widehat{F}_D(\theta, x)| \geq 1/2.$$

We deduce from Lemma 2.4.1 the following inequality

$$\mathbb{P}\left(\inf_{\theta\in\Theta_{\mathcal{H}}} \inf_{x\in\mathcal{S}_{\mathcal{H}}} |\widehat{F}_D(\theta, x)| \leq 1/2\right) \leq \mathbb{P}\left(\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{x\in\mathcal{S}_{\mathcal{H}}} |1 - \widehat{F}_D(\theta, x)| \leq 1/2\right).$$

Consequently,

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{\theta\in\Theta_{\mathcal{H}}} \inf_{x\in\mathcal{S}_{\mathcal{H}}} \widehat{F}_D(\theta, x) < \frac{1}{2}\right) < \infty$$

□

Proof of Lemma 2.4.2 One has

$$\mathbb{E}\widehat{F}_N(\theta,y,x)-F(\theta,y,x) = \frac{1}{\mathbb{E}K_1(x,\theta)}\mathbb{E}\left[\sum_{i=1}^n K_i(x,\theta)H_i(y)\right]-F(\theta,y,x)$$

$$= \frac{1}{\mathbb{E}K_1(x,\theta)}\mathbb{E}\left(K_1(x,\theta)\left[E\left(H_1(y)| < X_1,\theta >\right)-F(\theta,y,x)\right]\right) \quad (2.32)$$

Moreover, we have

$$\mathbb{E}\left(H_1(y)| < X_1,\theta >\right) = \int_{\mathbb{R}} H\left(h_H^{-1}(y-z)\right) f(\theta,z,X_1)dz,$$

now, integrating by parts and using the fact that H is a *cdf*, we obtain

$$\mathbb{E}\left(H_1(y)| < X_1,\theta >\right) = \int_{\mathbb{R}} H^{(1)}(t)F(\theta,y-h_Ht,X_1)dt.$$

Thus, we have

$$\left|\mathbb{E}\left(H_1(y)| < X_1,\theta >\right)-F(\theta,y,x)\right| \leq \int_{\mathbb{R}} H^{(1)}(t)\left|F(\theta,y-h_Ht,X_1)-F(\theta,y,x)\right|dt.$$

Finally, the use of (A2) implies that

$$\left|\mathbb{E}\left(H_1(y)|X_1\right)-F^x(y)\right| \leq C_{\theta,x}\int_{\mathbb{R}} H^{(1)}(t)\left(h_K^{b_1}+|t|^{b_2}h_H^{b_2}\right)dt. \quad (2.33)$$

Because this inequality is uniform on $(\theta,y,x)\in\Theta_{\mathcal{H}}\times\mathcal{S}_{\mathcal{H}}\times\mathcal{S}_{\mathbb{R}}$ and because of (H4), (2.14) is a direct consequence of (2.32), (2.33) and of Corollary 2.4.2.

□

Proof of Lemma 2.4.3 We keep the notation of the Lemma 2.4.1 and we use the compact of $\mathcal{S}_{\mathbb{R}}$, we can write that, for some, $t_1,\ldots,t_{z_n}\in\mathcal{S}_{\mathbb{R}}$, $\mathcal{S}_{\mathbb{R}}\subset\bigcup_{m=1}^{z_n}(y_m-l_n,y_m+l_n)$ with $l_n=n^{-1/2b_2}$ and $z_n\leq Cn^{-1/2b_2}$. Taking $m(y)=\arg\min_{\{1,2,\ldots,z_n\}}|y-t_m|$.

Thus, we have the following decomposition:

$$\begin{aligned}\left|\widehat{F}_N(\theta,y,x)-\mathbb{E}\left(\widehat{F}_N(\theta,y,x)\right)\right| = & \underbrace{\left|\widehat{F}_N(\theta,y,x)-\widehat{F}_N(\theta,y,x_{k(x)})\right|}_{\Gamma_1}\\ & +\underbrace{\left|\widehat{F}_N(\theta,y,x_{k(x)})-\mathbb{E}\left(\widehat{F}_N(\theta,y,x_{k(x)})\right)\right|}_{\Gamma_2}\\ & +2\underbrace{\left|\widehat{F}_N(t_{j(\theta)},y,x_{k(x)})-\widehat{F}_N(t_{j(\theta)},y_{m(y)},x_{k(x)})\right|}_{\Gamma_3}\\ & +2\underbrace{\left|\mathbb{E}\left(\widehat{F}_N(t_{j(\theta)},y,x_{k(x)})\right)-\mathbb{E}\left(\widehat{F}_N(t_{j(\theta)},y_{m(y)},x_{k(x)})\right)\right|}_{\Gamma_4}\\ & +\underbrace{\left|\mathbb{E}\left(\widehat{F}_N(\theta,y,x_{k(x)})\right)-\mathbb{E}\left(\widehat{F}_N(\theta,y,x)\right)\right|}_{\Gamma_5}\end{aligned}$$

$\hookrightarrow$ Concerning Γ_1 we have

$$\left|\widehat{F}_N(\theta,y,x)-\widehat{F}_N(\theta,y,x_{k(x)})\right| \leq \frac{1}{n}\sum_{i=1}^{n}\left|\frac{1}{\mathbb{E}K_1(\theta,x)}K_i(\theta,x)H_i(y)-\frac{1}{\mathbb{E}K_1(\theta,x_{k(x)})}K_i(\theta,x_{k(x)})H_i(y)\right|.$$

We use the Hölder continuity condition on K, the Cauchy-Schwartz inequality, the Bernstein's inequality and the boundness of H (assumption (H4)). This allows us to get:

$$\begin{aligned}\left|\widehat{F}_N(\theta,y,x)-\widehat{F}_N(\theta,y,x_{k(x)})\right| &\leq \frac{C}{\phi(h_K)}\frac{1}{n}\sum_{i=1}^{n}\left|K_i(\theta,x)H_i(y)-K_i(\theta,x_{k(x)})H_i(y)\right| \\ &\leq \frac{C}{\phi(h_K)}\frac{1}{n}\sum_{i=1}^{n}|H_i(y)|\left|K_i(\theta,x)-K_i(\theta,x_{k(x)})\right| \\ &\leq \frac{C'r_n}{\phi(h_K)}\end{aligned}$$

$\hookrightarrow$ Concerning Γ_2, the monotony of the functions $\mathbb{E}\widehat{F}_N(\theta,\cdot,x)$ and $\widehat{F}_N(\theta,\cdot,x)$ permits to write, $\forall m\leq z_n,\ \forall x\in\mathcal{S}_{\mathcal{H}},\ \forall\theta\in\Theta_{\mathcal{H}}$

$$\begin{aligned}\mathbb{E}\widehat{F}_N(\theta,y_{m(y)}-l_n,x_{k(x)}) &\leq \sup_{y\in(y_{m(y)}-l_n,y_{m(y)}+l_n)}\mathbb{E}\widehat{F}_N(\theta,y,x)\leq\mathbb{E}\widehat{F}_N(\theta,y_{m(y)}+l_n,x_{k(x)}) \\ \widehat{F}_N(\theta,y_{m(y)}-l_n,x_{k(x)}) &\leq \sup_{y\in(y_{m(y)}-l_n,y_{m(y)}+l_n)}\widehat{F}_N(\theta,y,x)\leq\widehat{F}_N(\theta,y_{m(y)}+l_n,x_{k(x)})\end{aligned} \tag{2.34}$$

Next, we use the Hölder's condition on $F(\theta,y,x)$ and we show that, for any $y_1,\ y_2\in\mathcal{S}_{\mathbb{R}}$ and for all $x\in\mathcal{S}_{\mathcal{H}},\ \theta\in\Theta_{\mathcal{H}}$

$$\begin{aligned}\left|\mathbb{E}\widehat{F}_N(\theta,y_1,x)-\mathbb{E}\widehat{F}_N(\theta,y_2,x)\right| &= \frac{1}{\mathbb{E}K_1(x,\theta)}\left|\mathbb{E}\left(K_1(x,\theta)F(\theta,y_1,X_1)\right)-\mathbb{E}\left(K_1(x,\theta)F(\theta,y_2,X_1)\right)\right| \\ &\leq C|y_1-y_2|^{b_2}.\end{aligned} \tag{2.35}$$

Now, we have, for all $\eta>0$

$$\mathbb{P}\left(\left|\widehat{F}_N(\theta,y,x_{k(x)})-\mathbb{E}\widehat{F}_N(\theta,y,x_{k(x)})\right|>\eta\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right)$$

$$=$$

$$\mathbb{P}\left(\max_{j\in\{1\dots d_n^{\Theta_{\mathcal{H}}}\}}\max_{k\in\{1\dots d_n^{\mathcal{S}_{\mathcal{H}}}\}}\max_{1\leq m\leq z_n}\left|\widehat{F}_N(\theta,y,x_{k(x)})-\mathbb{E}\widehat{F}_N(\theta,y,x_{k(x)})\right|>\eta\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right)$$

$$\leq$$

$$z_nd_n^{\mathcal{S}_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}\max_{j\in\{1\dots d_n^{\Theta_{\mathcal{H}}}\}}\max_{k\in\{1\dots d_n^{\mathcal{S}_{\mathcal{H}}}\}}\max_{1\leq m\leq z_n}\mathbb{P}\left(\left|\widehat{F}_N(\theta,y,x_{k(x)})-\mathbb{E}\widehat{F}_N(\theta,y,x_{k(x)})\right|>\eta\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{F}}}d_n^{\Theta_{\mathcal{F}}}}{n\phi(h_K)}}\right)$$

$$\leq$$

$$2z_n d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}} \exp\left(-C\eta^2 \log d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}\right)$$

choising $z_n = O\left(l_n^{-1}\right) = O\left(n^{\frac{1}{2b_2}}\right)$, we get

$$\mathbb{E}\left(\left|\widehat{F}_N(\theta, y, x_{k(x)}) - \mathbb{E}\widehat{F}_N(\theta, y, x_{k(x)})\right| > \eta\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right) \leq C' z_n \left(d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}\right)^{1-C\eta^2}$$

putting $C\eta^2 = \beta$ and using (A4), we get

$$\Gamma_2 = O_{a.co}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

$\hookrightarrow$ Concerning the terms Γ_3 and Γ_4, using Lipschitz's condition on the kernel H, one can write

$$\begin{aligned}\left|\widehat{F}_N(t_{j(\theta)}, y, x_{k(x)}) - \widehat{F}_N(t_{j(\theta)}, y_{m(y)}, x_{k(x)})\right| &\leq C\frac{1}{n\phi(h_K)}\sum_{i=1}^{n} K_i(t_{j(\theta)}, x_{k(x)})\left|H_i(y) - H_i(y_{m(y)}\right| \\ &\leq \frac{Cl_n}{nh_H\phi(h_K)}\sum_{i=1}^{n} K_i(t_{j(\theta)}, x_{k(x)}).\end{aligned}$$

Once again a standard exponential inequality for a sum of bounded variables allows us to write

$$\widehat{F}_N(t_{j(\theta)}, y, x_{k(x)}) - \widehat{F}_N(t_{j(\theta)}, y_{m(y)}, x_{k(x)}) = O\left(\frac{l_n}{h_H}\right) + O_{a.co}\left(\frac{l_n}{h_H}\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right).$$

Now, the fact that $\lim_{n\to\infty} n^{\gamma} h_H = \infty$ and $l_n = n^{-1/2b_2}$ imply that:

$$\frac{l_n}{h_H\phi(h_K)} = o\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right),$$

then

$$\Gamma_3 = O_{a.co}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

Hence, for n large enough, we have

$$\Gamma_3 \leq \Gamma_4 = O_{a.co}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}} d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

↪ Concerning Γ_5, we have

$$\mathbb{E}\left(\widehat{F}_N(\theta,y,x_{k(x)})\right)-\mathbb{E}\left(\widehat{F}_N(\theta,y,x)\right)\leq\sup_{x\in\mathcal{S}_{\mathcal{H}}}\left|\widehat{F}_N(\theta,y,x)-\widehat{F}_N(\theta,y,x_{k(x)})\right|,$$

then following similar proof used in the study of Γ_1 and using the same idea as for $\mathbb{E}\left(\widehat{F}_D(\theta,x_{k(x)})\right)-\mathbb{E}\left(\widehat{F}_D(\theta,x)\right)$ we get, for n tending to infinity,

$$\Gamma_5=O_{a.co}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}}d_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right).$$

□

Proof of Lemma 2.4.4. Let $H_i^{(j+1)}(y)=H^{(j+1)}\left(h_H^{-1}(y-Y_i)\right)$, note that

$$\mathbb{E}\widehat{f}_N^{(j)}(\theta,y,x)-f^{(j)}(\theta,y,x)=\frac{1}{h_H^{j+1}\,\mathbb{E}K_1(x,\theta)}\mathbb{E}\left(K_1(x,\theta)\left[\mathbb{E}\left(H_1^{(j+1)}(y)|<X,\theta>\right)-h_H^{j+1}f^{(j)}(\theta,y,x)\right]\right).\tag{2.36}$$

Moreover,

$$\begin{aligned}\mathbb{E}\left(H_1^{(j+1)}(y)|<X,\theta>\right)&=\int_{\mathbb{R}}H^{(j+1)}\left(h_H^{-1}(y-z)\right)f(\theta,z,X)dz,\\&=-\sum_{l=1}^{j}h_H^{l}\left[H^{(j-l+1)}\left(h_H^{-1}(y-z)\right)f^{(l-1)}(\theta,z,X)\right]_{-\infty}^{+\infty}\\&\quad+h_H^{j}\int_{\mathbb{R}}H^{(1)}\left(h_H^{-1}(y-z)\right)f^{(j)}(\theta,z,X)dz.\end{aligned}\tag{2.37}$$

Condition (H8) allows us to cancel the first term in the right side of (2.37) and we can write:

$$\left|\mathbb{E}\left(H_1^{(j+1)}(y)|<X,\theta>\right)-h_H^{j+1}f^{(j)}(\theta,y,x)\right|\leq h_H^{j+1}\int_{\mathbb{R}}H^{(1)}(t)\left|f^{(j)}(\theta,y-h_Ht,X)-f^{(j)}(\theta,y,x)\right|dt.$$

Finally, (A5) allows to write

$$\left|\mathbb{E}\left(H_1^{(j+1)}(y)|<X,\theta>\right)-h_H^{j+1}f^{(j)}(\theta,y,x)\right|\leq C_{\theta,x}h_H^{j+1}\int_{\mathbb{R}}H^{(1)}(t)\left(h_K^{b_1}+|t|^{b_2}h_H^{b_2}\right)dt.\tag{2.38}$$

This inequality is uniform on $(\theta,y,x)\in\Theta_{\mathcal{F}}\times\mathcal{S}_{\mathcal{F}}\times\mathcal{S}_{\mathbb{R}}$, now to finish the proof it is sufficient to use (H4).

□

Proof of Lemma 2.4.5. Let $l_n=n^{-\frac{3}{2}\gamma-\frac{1}{2}}$ and $z_n\leq Cn^{-\frac{3}{2}\gamma-\frac{1}{2}}$.

Consider the following decomposition

$$
\begin{aligned}
\left|\widehat{f}_N^{(j)}(\theta,y,x)-\mathbb{E}\left(\widehat{f}_N^{(j)}(\theta,y,x)\right)\right| = & \underbrace{\left|\widehat{f}_N^{(j)}(\theta,y,x)-\widehat{f}_N^{(j)}(\theta,y,x_{k(x)})\right|}_{\Delta_1} \\
& +\underbrace{\left|\widehat{f}_N^{(j)}(\theta,y,x_{k(x)})-\mathbb{E}\left(\widehat{f}_N^{(j)}(\theta,y,x_{k(x)})\right)\right|}_{\Delta_2} \\
& +2\underbrace{\left|\widehat{f}_N^{(j)}(t_{j(\theta)},y,x_{k(x)})-\widehat{f}_N^{(j)}(t_{j(\theta)},y_{m(y)},x_{k(x)})\right|}_{\Delta_3} \\
& +2\underbrace{\left|\mathbb{E}\left(\widehat{f}_N^{(j)}(t_{j(\theta)},y,x_{k(x)})\right)-\mathbb{E}\left(\widehat{f}_N^{(j)}(t_{j(\theta)},y_{m(y)},x_{k(x)})\right)\right|}_{\Delta_4} \\
& +\underbrace{\left|\mathbb{E}\left(\widehat{f}_N^{(j)}(\theta,y,x_{k(x)})\right)-\mathbb{E}\left(\widehat{f}_N^{(j)}(\theta,y,x)\right)\right|}_{\Delta_5}
\end{aligned}
$$

⇝ Concerning Δ_1, we use the Hölder continuity condition on K, the Cauchy-Schwartz's inequality and the Bernstein's inequality. With theses arguments we get

$$\Delta_1 = O\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{H}}+\log d_n^{\Theta_\mathcal{H}}}{nh_H^{2j+1}\phi(h_K)}}\right).$$

Then using the fact that $\Delta_5 \le \Delta_1$, we obtain

$$\Delta_5 \le \Delta_1 = O\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{H}}+\log d_n^{\Theta_\mathcal{H}}}{nh_H^{2j+1}\phi(h_K)}}\right). \tag{2.39}$$

⇝ For Δ_2, we follow the same idea given for Γ_2, we get

$$\Delta_2 = O\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{H}}+\log d_n^{\Theta_\mathcal{H}}}{nh_H^{2j+1}\phi(h_K)}}\right)$$

⇝ Concerning Δ_3 and Δ_4, Using Lipschitz's condition on the kernel H,

$$\left|\widehat{f}_N^{(j)}(t_{j(\theta)},y,x_{k(x)})-\widehat{f}_N^{(j)}(t_{j(\theta)},y_{m(y)},x_{k(x)})\right| \le \frac{l_n}{h_H^{j+2}\phi(h_k)},$$

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

$$\frac{l_n}{h_H^{j+2}\phi(h_k)} = o\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{H}}+\log d_n^{\Theta_\mathcal{H}}}{nh_H^{2j+1}\phi(h_K)}}\right)$$

So, for n large enough, we have

$$\Delta_3 = O_{a.co}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_\mathcal{H}}+\log d_n^{\Theta_\mathcal{H}}}{nh_H^{2j+1}\phi(h_K)}}\right).$$

And as $\Delta_4 \leq \Delta_3$, we obtain

$$\Delta_4 \leq \Delta_3 = O_{a.co}\left(\sqrt{\frac{\log d_n^{\mathcal{S}_{\mathcal{H}}} + \log d_n^{\Theta_{\mathcal{H}}}}{n h_H^{2j+1} \phi(h_K)}}\right). \tag{2.40}$$

Finally, the lemma can be easily deduced from (2.39) and (2.40)

□

Proof of Lemma 2.5.1. Because of (H4) and (A9) the function $\widehat{F}(\theta, \cdot, x)$ is uniformly continuous and strictly increasing. So, we have:

$$\forall \epsilon > 0,\ \exists \delta(\epsilon) > 0, \forall y, |\widehat{F}(\theta, y, x) - \widehat{F}(\theta, t_\theta(\alpha), x)| \leq \delta(\epsilon) \Rightarrow |y - t_\theta(\alpha)| \leq \epsilon.$$

This leads directly to

$$\begin{aligned}\forall \epsilon > 0,\ \exists \delta(\epsilon) > 0,\ \mathbb{P}\left(|\widehat{t}_\theta(\alpha) - t_\theta(\alpha)| > \epsilon\right) &\leq \mathbb{P}\left(|\widehat{F}(\theta, \widehat{t}_\theta(\alpha), x) - \widehat{F}(\theta, t_\theta(\alpha), x)| \geq \delta(\epsilon)\right) \\ &= \mathbb{P}\left(|F(\theta, t_\theta(\alpha), x) - \widehat{F}(\theta, t_\theta(\alpha), x)| \geq \delta(\epsilon)\right).\end{aligned}$$

Finally, It suffices to use the result of Theorem 2.4.1 to get the claimed result.

□

Bibliography

[1] 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.

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

[3] 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(1), 45-53.

[4] Berlinet, A., Cadre, B., and Gannoun, A. (2001). On the conditional L1-median and its estimation. J. Nonparametr. Statist., 13(5), 631-645.

[5] Dabo-Niang, S. and Laksaci, A. (2012). Nonparametric quantile regression estimation for functional dependent data. Comm. Statist. Theory Methods, 41(7), 1254-1268.

[6] Delecroix, M, Hördle, W., Hristache, M., (2003). Efficient estimation in conditional single-index regression. J. Multivariate Anal. 86, 213-226.

[7] El Ghouch, A. and Van Keilegom, I. (2009). Local linear quantile regression with dependent censored data. Statist. Sinica, 19(4), 1621-1640.

[8] Ezzahrioui, M. and Ould-Saïd, E. (2008). Asymptotic results of a nonparametric conditional quantile estimator for functional time series. Comm. Statist. Theory Methods, 37(16-17), 2735-2759.

[9] Ezzahrioui, M., Ould Saïd, E., (2010). Some asymptotic results of a nonparametric conditional mode estimator for functional time series data. Statist. Neerlandica 64, 171-201 .

[10] Ferraty, F., Peuch, A., Vieu, P., (2003). Modèle à indice fonctionnel simple. C. R. Mathématiques Paris 336, 1025-1028.

[11] 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** No.2, 399-417.

[12] Ferraty, F., Laksaci, A., Vieu, P., (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. Statist. Inf. for Stoch. Proc. 9, 47-76.

[13] Ferraty, F., Vieu, P., (2006). Nonparametric Functional Data Analysis: Theory and Practice. Springer Series in Statistics, Springer, New York.

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

[15] Gannoun, A., Saracco, J., and Yu, K. (2003). Nonparametric prediction by conditional median and quantiles. J. Statist. Plann. Inference, 117(2), 207-223.

[16] Gannoun, A., Saracco, J., Yuan, A., and Bonney, G. E. (2005). Nonparametric quantile regression with censored data. Scand. J. Statist., 32(4), 527-550.

[17] Härdle, W., Hall, P., Ichumira, H., (1993). Optimal smoothing in single-index models, Ann. Statist. 21, 157-178.

[18] Härdle, W. and Marron, J.S., (1985). Optimal bandwidth selection in nonparametric regression function estimation, Ann. Statist. 13, 1465ï£¡1481.

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

[20] Koenker, R. and Bassett, J., G. (1978). Regression quantiles. Econometrica, 46(1), 33-50.

[21] Laksaci, A., Lemdani, M., Ould Saïd, E., (2009). A generalized L1 - approach for a kernel estimator of conditional quantile with functional regressors: Consistency and asymptotic normality. Statist. & Probab. Lett. 79, 1065-1073.

[22] Liang, H.-Y. and de Uña-Àlvarez, J. (2011). Asymptotic properties of conditional quantile estimator for censored dependent observations. Ann. Inst. Statist. Math., 63(2), 267-289.

[23] Marron, J.S. (1987). A comparison of cross-validation techniques in density estimation, Ann. Statist. 15, 152-162.

[24] Ould-Saïd, E.,Cai, Z. (2005). Strong uniform consistency of nonparametrice stimation of the censored conditional mode function. Nonparametric-Statist.17, 797-806.

[25] Ould-Saïd, E. (2006). A strong uniform convergence rate of kernel conditional quantile estimator under random censorship. Statist. Probab. Lett., 76(6), 579-586.

[26] Rosenblatt, M., (1969). Conditional probability density and regression estimators. In Multivariate Analysis II, Ed. P.R. Krishnaiah. Academic Press, New York and London.

[27] Samanta, M. (1989). Non-parametric estimation of conditional quantiles. *Stat. Probab . Lett.* **7**, No.5, 407-412.

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

Chapter 3

Real response and independent case

This chapter is the object of a work subjected for publication in **ProbStat Forum**

3.1 Introduction

The statistical analysis of functional data studies the experiments whose results are generally the curves. Under this supposition, the statistical analysis focuses on a framework of infinite dimension for the data under study. This field of modern statistics has received much attention in the last 20 years, and it has been popularized in the book of Ramsay and Silverman (2005). This type of data appears in many fields of applied statistics: environmetrics (Damon and Guillas, (2002)), chemometrics (Benhenni *et al.*, (2007)), meteorological sciences (Besse *et al.*, (2000)), etc.

The study of the hazard function has been subject to several investigations and many authors considered this function in their investigations. Among others, we refer to Watson and Leadbetter (1964) who were the first to study the nonparametric estimation of the hazard function and proposed a kernel estimate; in the sequel, many authors have been interested in the study of such a function. For instance, Collomb *et al.* (1987) studied the dependent case, Liu and Van Ryzin (1985) were interested in the histogram estimator of the hazard function for censored data, and Youndje *et al.* (1996) proposed a solution to the bandwith selection problem for the kernel hazard estimate and gave properties of the selected bandwith. In the sequel, Quintela (2007) used the plug-in bandwith selection method in the case of a weak dependence on the sample data and a result of asymptotic optimality for the plug-in bandwith is presented. Simulations are done as well to compare this method to the "leave more than one out" cross-validation criterion and either to show that smaller errors and much less sample variability can be reached. Besides, we notice that most of the precursor literature on nonparametric smoothing of hazard function was based on the assumption of independence on the sample variables, which is far from being realistic. This is for instance the case of micro earthquake studies (Rice and Rosenblatt (1976); Estévez *et al.*, 2002). Alternatively, many authors investigated the case of dependent hazard estimation, for instance Sarda and Vieu (1989), Vieu (1991), and Estévez and Quintela (1999), to name a few.

Recently, Quintela (2010) studied the recursive kernel hazard estimation of strong mixing data, by use of the density and the distribution, and established the strong consistency of the proposed estimator and a rate of convergence identical to the one obtained in the independence case.

3.1.1 Hazard and conditional hazard

The estimation of the hazard function is a problem of considerable interest, especially to inventory theorists, medical researchers, logistics planners, reliability engineers and seismologists. The non-parametric estimation of the hazard function has been extensively discussed in the literature. Beginning with Watson and Leadbetter (1964), there are many papers on these topics: Ahmad (1976), Singpurwalla and Wong (1983), etc. We can cite Quintela (2007) for a survey.

The literature on the estimation of the hazard function is very abundant, when observations are vectorial. Cite, for instance, Watson and Leadbetter

(1964), Roussas (1989), Lecoutre and Ould-Saïd (1993), Estèvez *et al.* (2002) and Quintela-del-Rio (2006) for recent references. In all these works the authors consider independent observations or dependent data from time series. The first results on the nonparametric estimation of this model, in functional statistics were obtained by Ferraty *et al.* (2008). They studied the almost complete convergence of a kernel estimator for hazard function of a real random variable dependent on a functional predictor and Laksaci and Mechab (2010) in the case of spatial variables. Asymptotic normality of the latter estimator was obtained, in the case of α- mixing, by Quintela-del-Rio (2008). We refer to Ferraty *et al.* (2010) and Bouchentouf *et al.* (2014) for uniform almost complete convergence of the functional component of this nonparametric model. When hazard rate estimation is performed with multiple variables, the result is an estimate of the conditional hazard rate for the first variable, given the levels of the remaining variables. Many references, practical examples and simulations in the case of non-parametric estimation using local linear approximations can be found in Spierdijk (2008).

Our paper presents some asymptotic properties related with the nonparametric estimation of the maximum of the conditional hazard function. In a functional data setting, the conditioning variable is allowed to take its values in some abstract semi-metric space. In this case, Ferraty *et al.* (2008) define nonparametric estimators of the conditional hazard function. They give the rates of convergence (in an almost complete sense) to the corresponding functions, in a dependence (α-mixing) context. We extend their results by calculating the maximum of the conditional hazard function of these estimates, and establishing their asymptotic normality, considering a particular type of kernel for the functional part of the estimates. Because the hazard function estimator is naturally constructed using these two last estimators, the same type of properties is easily derived for it. Our results are valid in a real (one- and multi-dimensional) context.

If X is a random variable associated to a lifetime (ie, a random variable with values in $\mathbb{R}^+$, the hazard rate of X (sometimes called hazard function, failure or survival rate) is defined at point x as the instantaneous probability that life ends at time x. Specifically, we have:

$$h(x) = \lim_{dx\to 0} \frac{\mathbb{P}\left(X \leq x + dx | X \geq x\right)}{dx}, \qquad (x > 0).$$

When X has a density f with respect to the measure of Lebesgue, it is easy to see that the hazard rate can be written as follows:

$$h(x) = \frac{f(x)}{S(x)} = \frac{f(x)}{1 - F(x)}, \text{ for all } x \text{ such that } F(x) < 1,$$

where F denotes the distribution function of X and $S = 1 - F$ the survival function of X.

In many practical situations, we may have an explanatory variable $X = x$ and the main issue is to estimate the conditional random rate defined as

$$h^x(y) = \lim_{dy \to 0} \frac{\mathbb{P}\left(X \leq y + dy | Y > y, \quad X = x\right)}{dy}, \text{ for } y > 0,$$

which can be written naturally as follows:

$$h^x(y) = \frac{f^x(y)}{S^x(y)} = \frac{f^x(y)}{1 - F^x(y)}, \text{ once } F^x(y) < 1. \tag{3.1}$$

Study of functions $h(\cdot)$ and $h^z(\cdot)$ is of obvious interest in many fields of science (biology, medicine, reliability , seismology, econometrics, ...) and many authors are interested in construction of nonparametric estimators of h.

In this paper we propose an estimate of the maximum risk, through the nonparametric estimation of the conditional hazard function.

The layout of the paper is as follows. Section 3.2 describes the non-parametric functional setting: the structure of the functional data and the mixing conditions, the conditional density, distribution and hazard operators, and the corresponding non-parametric kernel estimators. Section 3.3 states the almost complete convergence[1] (with rates of convergence[2]) for nonparametric estimates of the derivative of the conditional hazard and the maximum risk. In Section 3.4, we calculate the variance of the conditional density, distribution and hazard estimates, the asymptotic normality of the three estimators considered is developed in this Section. Finally, Section 3.5 includes some proofs of technical Lemmas.

3.2 Nonparametric estimation with functional data

Let $\{(X_i, Y_i),\ i = 1, \ldots, n\}$ be a sample of n random pairs, each one distributed as (X, Y), where the variable X is of functional nature and Y is scalar. Formally, we will consider that X is a random variable valued in some semi-metric functional space $\mathcal{F}$, and we will denote by $d(\cdot, \cdot)$ the associated semi-metric. The conditional cumulative distribution of X given $X = x$ is defined for any $y \in \mathbb{R}$ and any $x \in \mathcal{F}$ by

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

while the conditional density, denoted by $f^x(y)$ is defined as the density of this distribution with respect to the Lebesgue measure on $\mathbb{R}$. The conditional hazard is defined as in the non-infinite case (3.1).

In a general functional setting, f, F and h are not standard mathematical objects. Because they are defined on infinite dimensional spaces, the term operators may be a more adjusted in terminology.

[1]Recall that a sequence $(T_n)_{n \in \mathbb{N}}$ of random variables is said to converge almost completely to some variable T, if for any $\epsilon > 0$, we have $\sum_n \mathbb{P}(|T_n - T| > \epsilon) < \infty$. This mode of convergence implies both almost sure and in probability convergence (see for instance Bosq and Lecoutre, (1987)).

[2]Recall that a sequence $(T_n)_{n \in \mathbb{N}}$ of random variables is said to be of order of complete convergence u_n, if there exists some $\epsilon > 0$ for which $\sum_n \mathbb{P}(|T_n| > \epsilon u_n) < \infty$. This is denoted by $T_n = \mathcal{O}(u_n)$, *a.co.* (or equivalently by $T_n = \mathcal{O}_{a.co.}(u_n)$).

3.2.1 The functional kernel estimates

Following in Ferraty *et al.* (2008), the conditional density operator $f^x(\cdot)$ is defined by using kernel smoothing methods

$$\widehat{f}^x(y) = \frac{\sum_{i=1}^{n} h_n^{-1} K\left(h_n^{-1} d(x, X_i)\right) H\left(h_n^{-1}(y - Y_i)\right)}{\sum_{i=1}^{n} K\left(h_n^{-1} d(x, X_i)\right)},$$

where K and H are kernel functions and h_n is sequence of smoothing parameter. The conditional distribution operator $F^x(\cdot)$ can be estimated by

$$\widehat{F}^x(y) = \sum_{i=1}^{n} W_{ni}(x) \mathbf{1}_{\{Y_i \leq y\}}, \qquad \forall y \in \mathbb{R}$$

with $\mathbf{1}_{\{\cdot\}}$ being the indicator function and where $W_{ni}(x) = \frac{h_n^{-1} K(h_n^{-1} d(x,X_i))}{\sum_{j=1}^{n} K(h_n^{-1} d(x,X_j))}$, K is a kernel function and h_n is a sequence of positive real numbers which goes to zero as n goes to infinity.

Consequently, the conditional hazard operator is defined in a natural way by

$$\widehat{h}^x(y) = \frac{\widehat{f}^x(y)}{1 - \widehat{F}^x(y)}.$$

For $x \in \mathcal{F}$, we denote by $h^x(\cdot)$ the conditional hazard function of Y_1 given $X_1 = x$. We assume that $h^x(\cdot)$ is unique maximum and its high risk point is denoted by $\theta(x) := \theta$, which is defined by

$$h^x(\theta(x)) := h^x(\theta) = \max_{y \in \mathcal{S}} h^x(y). \tag{3.2}$$

A kernel estimator of θ is defined as the random variable $\widehat{\theta}(x) := \widehat{\theta}$ which maximizes a kernel estimator $\widehat{h}^x(\cdot)$, that is,

$$\widehat{h}^x(\widehat{\theta}(x)) := \widehat{h}^x(\widehat{\theta}) = \max_{y \in \mathcal{S}} \widehat{h}^x(y) \tag{3.3}$$

where h^x and $\widehat{h}^x$ are defined above.

Note that the estimate $\widehat{\theta}$ is note necessarily unique and our results are valid for any choice satisfying (3.3). We point out that we can specify our choice by taking

$$\widehat{\theta}(x) = \inf \left\{ t \in \mathcal{S} \text{ such that } \widehat{h}^x(t) = \max_{y \in \mathcal{S}} \widehat{h}^x(y) \right\}.$$

As in any non-parametric functional data problem, the behavior of the estimates is controlled by the concentration properties of the functional variable $X = x$.

$$\phi_x(h_n) = \mathbb{P}(X \in B(x, h_n)),$$

where $B(x,h_n)$ being the ball of center x and radius h_n, namely $B(x,h_n) = \mathbb{P}(f \in \mathcal{F}, d(x,f) < h_n)$ (for more details, see Ferraty and Vieu (2006), Chapter 6).

In the following, x will be a fixed point in $\mathcal{F}$, $\mathcal{N}_x$ will denote a fixed neighborhood of x, $\mathcal{S}$ will be a fixed compact subset of $\mathbb{R}^+$. We will led to the hypothesis below concerning the function of concentration ϕ_x

(H1) $\forall h_n > 0,\ 0 < \mathbb{P}(X \in B(x,h_n)) = \phi_x(h_n)$ and $\lim_{h_n \to 0} \phi_x(h_n) = 0$

Note that (H1) can be interpreted as a concentration hypothesis acting on the distribution of the *f.r.v.* of $X = x$.

Our nonparametric models will be quite general in the sense that we will just need the following simple assumption for the marginal distribution of X, and let us introduce the technical hypothesis necessary for the results to be presented. The non-parametric model is defined by assuming that

$$\text{(H2)} \begin{cases} \forall (y_1,y_2) \in \mathcal{S}^2, \forall (x_1,x_2) \in \mathcal{N}_x^2, \text{ for some } b_1 > 0,\ b_2 > 0 \\ |F^{x_1}(y_1) - F^{x_2}(y_2)| \le C_x(d(x_1,x_2)^{b_1} + |y_1 - y_2|^{b_2}), \end{cases}$$

$$\text{(H3)} \begin{cases} \forall (y_1,y_2) \in \mathcal{S}^2, \forall (x_1,x_2) \in \mathcal{N}_x^2, \text{ for some } j = 0,1,\ \nu > 0,\ \beta > 0 \\ |f^{x_1\,(j)}(y_1) - f^{x_2\,(j)}(y_2)| \le C_x(d(x_1,x_2)^{\nu} + |y_1 - y_2|^{\beta}), \end{cases}$$

(H4) $\exists \gamma < \infty, f'^{x}(y) \le \gamma,\ \forall (x,y) \in \mathcal{F} \times \mathcal{S}$,

(H5) $\exists \tau > 0, F^x(y) \le 1 - \tau,\ \forall (x,y) \in \mathcal{F} \times \mathcal{S}$.

(H6) H is differentiable such that

$$\begin{cases} \text{(H6a) } \forall (t_1,t_2) \in \mathbb{R}^2;\ |H^{(j)}(t_1) - H^{(j)}(t_2)| \le C|t_1 - t_2|, \text{ for } j = 0,1 \\ \text{and } H^{(j)} \text{are bounded for } j = 0,1 \\ \text{(H6b) } \int_{\mathbb{R}} t^2 H^2(t)dt < \infty, \\ \text{(H6c) } \int_{\mathbb{R}} |t|^{\beta}(H'(t))^2 dt < \infty \end{cases}$$

(H7) The kernel K is positive bounded function supported on $[0,1]$ and it is of class $\mathcal{C}^1$ on $(0,1)$ such that $\exists C_1,\ C_2,\ -\infty < C_1 < K'(t) < C_2$ for $0 < t < 1$.

(H8) There exists a function $\zeta_0^x(\cdot)$ such that for all $t \in [0,1]$ $\lim_{h_n \to 0} \frac{\phi_x(th_n)}{\phi_x(h_n)} = \zeta_0^x(t)$.

(H9) The bandwidth h_n, small ball probability $\phi_x(h_n)$ satisfying

$$\begin{cases} \text{(H9a) } \lim_{n\to\infty} h_n = 0 \\ \text{(H9b) } \lim_{n\to\infty} \frac{\log n}{n h_n \phi_x(h_n)} = 0 \\ \text{(H9c) } \lim_{n\to\infty} \frac{\log n}{n h_n^{2j+1} \phi_x(h_n)} = 0,\ j = 0,1; \end{cases}$$

Remark 3.2.1 *Assumption (H1) plays an important role in our methodology. It is known as (for small h_n) the "concentration hypothesis acting on the distribution of X" in infinite-dimensional spaces. This assumption is not at all restrictive and overcomes the problem of the non-existence of the probability density function. In many examples, around zero the small ball probability$\phi_x(h_n)$ can be written approximately as the product of two independent functions $\psi(z)$ and $\varphi(h_n)$ as $\phi_x(h_n) = \psi(x)\varphi(h_n) + o(\varphi(h_n))$. This idea was adopted by Masry (2005) who reformulated the Gasser* et al. *(1998) one. The increasing propriety of $\phi_x(\cdot)$ implies that $\zeta^x_{h_n}(\cdot)$ is bounded and then integrable (all the more so $\zeta^x_0(\cdot)$ is integrable).*

Without the differentiability of $\phi_x(\cdot)$, this assumption has been used by many authors where $\psi(\cdot)$ is interpreted as a probability density, while $\varphi(\cdot)$ may be interpreted as a volume parameter. In the case of finite-dimensional spaces, that is $\mathcal{S} = \mathbb{R}^d$, it can be seen that $\phi_x(h_n) = C(d)h_n^d\psi(x) + o(h_n^d)$, where $C(d)$ is the volume of the unit ball in $\mathbb{R}^d$. Furthermore, in infinite dimensions, there exist many examples fulfilling the decomposition mentioned above. We quote the following (which can be found in Ferraty et al. *(2007)):*

1. $\phi_x(h_n) \approx \psi(h_n)h_n^\gamma$ *for some* $\gamma > 0$.

2. $\phi_x(h_n) \approx \psi(h_n)h_n^\gamma \exp\{C/h_n^p\}$ *for some* $\gamma > 0$ *and* $p > 0$.

3. $\phi_x(h_n) \approx \psi(h_n)/|\ln h_n|$.

The function $\zeta^x_{h_n}(\cdot)$ which intervenes in Assumption (H9) is increasing for all fixed h_n. Its pointwise limit $\zeta^x_0(\cdot)$ also plays a determinant role. It intervenes in all asymptotic properties, in particular in the asymptotic variance term. With simple algebra, it is possible to specify this function (with $\zeta_0(u) := \zeta^x_0(u)$ in the above examples by:

1. $\zeta_0(u) = u^\gamma$,

2. $\zeta_0(u) = \delta_1(u)$ *where* $\delta_1(\cdot)$ *is Dirac function,*

3. $\zeta_0(u) = \mathbf{1}_{]0,1]}(u)$.

Remark 3.2.2 *Assumptions (H2) and (H3) are the only conditions involving the conditional probability and the conditional probability density of Y given $X = x$. It means that $F(\cdot|\cdot)$ and $f(\cdot|\cdot)$ and its derivatives satisfy the Hölder condition with respect to each variable. Therefore, the concentration condition (H1) plays an important role. Here we point out that our assumptions are very usual in the estimation problem for functional regressors (see, e.g., Ferraty* et al. *(2008)).*

Remark 3.2.3 *Assumptions (H6), (H7) and (H9) are classical in functional estimation for finite or infinite dimension spaces.*

3.3 Nonparametric estimate of the maximum of the conditional hazard function

Let us assume that there exists a compact $\mathcal{S}$ with a unique maximum θ of h^x on $\mathcal{S}$. We will suppose that h^x is sufficiently smooth (at least of class $\mathcal{C}^2$) and verifies that $h'^x(\theta) = 0$ and $h''^x(\theta) < 0$.

We can write an estimator of the first derivative of the conditional hazard function through the first derivative of the estimator (3.1). Our maximum estimate is defined by assuming that there is some unique $\widehat{\theta}$ on $\mathcal{S}$ such that $0 = \widehat{h'}(\widehat{\theta}) < |\widehat{h'}^x(y)|$ for all $y \in \mathcal{S}$ and $y \neq \widehat{\theta}$

Furthermore, we assume that $\theta \in \mathcal{S}^\circ$, where $\mathcal{S}^\circ$ denotes the interior of $\mathcal{S}$, and that θ satisfies the uniqueness condition, that is; for any $\varepsilon > 0$ and $\mu(x)$, there exists $\xi > 0$ such that $|\theta(x) - \mu(x)| \geq \varepsilon$ implies that $|h^x(\theta(x)) - h^x(\mu(x))| \geq \xi$.

We can write an estimator of the first derivative of the hazard function through the first derivative of the estimator. Our maximum estimate is defined by assuming that there is some unique $\widehat{\theta}$ on $\mathcal{S}^\circ$.

It is therefore natural to try to construct an estimator of the derivative of the function h^x on the basis of these ideas. To estimate the conditional distribution function and the conditional density function in the presence of functional conditional random variable $X = x$.

The kernel estimator of the derivative of the function conditional random functional h^z can therefore be constructed as follows:

$$\widehat{h'}^x(y) = \frac{\widehat{f'}^x(y)}{1 - \widehat{F}^x(y)} + (\widehat{h}^x(y))^2, \tag{3.4}$$

the estimator of the derivative of the conditional density is given in the following formula:

$$\widehat{f'}^x(y) = \frac{\sum_{i=1}^{n} K(h_n^{-1} d(x, X_i)) H'(h_n^{-1}(y - Y_i))}{h_n^2 \sum_{i=1}^{n} K(h_n^{-1} d(x, X_i))}. \tag{3.5}$$

Later, we need assumptions on the parameters of the estimator, ie on K, H, H' and h_n are little restrictive. Indeed, on one hand, they are not specific to the problem estimate of h^x (but inherent problems of F^x, f^x and f'^x estimation), and secondly they consist with the assumptions usually made under functional variables.

Remark 3.3.1 *Generally, the hazard function has a global maximum in the time intervals with values closest to zero, corresponding to the earthquakes with bigger intensity (Vere-Jones, 1970).*

Also, the hazard function can have several local maxima, indicating the times with the highest risk in a certain period (see the examples in Estévez et al.*, (2002)).*

The hypothesis of uniqueness is only established for the sake of clarity. Following our proofs, if several local estimated maxima exist, the asymptotic results remain valid for each of them.

We state the almost complete convergence (withe rates of convergence) of the maximum estimate by the following results:

Theorem 3.3.1 *Under assumption (H1)-(H7) we have*

$$\widehat{\theta} - \theta \to 0 \quad a.co. \tag{3.6}$$

Remark 3.3.2 *The hypothesis of uniqueness is only established for the sake of clarity. Following our proofs, if several local estimated maxima exist, the asymptotic results remain valid for each of them.*

Proof. Because $h'^x(\cdot)$ is continuous, we have, for all $\epsilon > 0$. $\exists\ \eta(\epsilon) > 0$ such that

$$|t - \theta| > \epsilon \Rightarrow |h'^x(t) - h'^x(\theta)| > \eta(\epsilon).$$

Therefore,

$$\mathbb{P}\{|\widehat{\theta} - \theta| \geq \epsilon\} \leq \mathbb{P}\{|h'^x(\widehat{\theta}) - h'^x(\theta)| \geq \eta(\epsilon)\}.$$

We also have

$$|h'^x(\widehat{\theta}) - h'^x(\theta)| \leq |h'^x(\widehat{\theta}) - \widehat{h}'^x(\widehat{\theta})| + |\widehat{h}'^x(\widehat{\theta}) - h'^x(\theta)| \leq \sup_{y\in\mathcal{S}} |\widehat{h}'^x(y) - h'^x(y)|, \tag{3.7}$$

because $\widehat{h}'^x(\widehat{\theta}) = h'^x(\theta) = 0$.

Then, uniform convergence of h'^x will imply the uniform convergence of $\widehat{\theta}$. That is why, we have the following lemma.

Lemma 3.3.1 *Under assumptions of Theorem 3.3.1, we have*

$$\sup_{y\in\mathcal{S}} |\widehat{h}'^x(y) - h'^x(y)| \to 0 \quad a.co. \tag{3.8}$$

■

The proof of this lemma will be given in Appendix.

Theorem 3.3.2 *Under assumption (H1)-(H7), (H9a) and (H9c) we have*

$$\sup_{y\in\mathcal{S}} |\widehat{\theta} - \theta| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n h_n^3 \phi_x(h_n)}}\right) \tag{3.9}$$

Proof. By using Taylor expansion of the function h'^x at the point $\widehat{\theta}$, we obtain

$$h'^x(\widehat{\theta}) = h'^x(\theta) + (\widehat{\theta} - \theta) h''^x(\theta_n^*), \tag{3.10}$$

with θ_n^* a point between θ and $\widehat{\theta}$.

Now, because $h'^x(\theta) = \widehat{h}'^x(\widehat{\theta})$

$$|\widehat{\theta} - \theta| \leq \frac{1}{h''^x(\theta_n^*)} \sup_{y \in \mathcal{S}} |\widehat{h}'^x(y) - h'^x(y)| \tag{3.11}$$

The proof of Theorem will be completed showing the following lemma.

Lemma 3.3.2 *Under the assumptions of Theorem 3.3.2, we have*

$$\sup_{y \in \mathcal{S}} |\widehat{h}'^x(y) - h'^x(y)| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n h_n^3 \phi_x(h_n)}}\right) \tag{3.12}$$

The proof of lemma will be given in the Appendix. ∎

3.4 Asymptotic normality

To obtain the asymptotic normality of the conditional estimates, we have to add the following assumptions:

(H6d) $\displaystyle\int_{\mathbb{R}} (H'(t))^2 dt < \infty,$

(H10) $0 = \widehat{h'}^x(\widehat{\theta}) < |\widehat{h'}^x(y)|, \forall y \in \mathcal{S},\ y \neq \widehat{\theta}$

The following result gives the asymptotic normality of the maximum of the conditional hazard function. Let

$$\mathcal{A} = \{(x, y) : (x, y) \in \mathcal{F} \times \mathbb{R},\ a_2^y F^x(y)\left(1 - F^x(y)\right) \neq 0\}$$

Theorem 3.4.1 *Under conditions (H1)-(H10) we have* $(\theta \in \mathcal{S}/f^x(\theta), 1 - F^x(\theta) > 0)$

$$\left(n h_n^3 \phi_x(h_n)\right)^{1/2} \left(\widehat{h}'^x(\theta) - h'^x(\theta)\right) \xrightarrow{\mathcal{D}} N\left(0, \sigma_{h'}^2(\theta)\right)$$

where $\rightarrow^{\mathcal{D}}$ *denotes the convergence in distribution,*

$$a_l^y = K^l(1) - \int_0^1 \left(K^l(u)\right)' \zeta_0^y(u) du \quad \textit{for } l = 1, 2$$

and

$$\sigma_{h'}^2(\theta) = \frac{a_2^y h^x(\theta)}{\left(a_1^y\right)^2 \left(1 - F^x(\theta)\right)} \int (H'(t))^2 dt.$$

Proof. Using again (3.17), and the fact that

$$\frac{\left(1 - F^x(y)\right)}{\left(1 - \widehat{F}^x(y)\right)\left(1 - F^x(y)\right)} \longrightarrow \frac{1}{1 - F^x(y)}$$

and

$$\frac{\widehat{f'}^x(y)}{\left(1-\widehat{F}^x(y)\right)(1-F^x(y))} \longrightarrow \frac{f'^x(y)}{(1-F^x(y))^2}.$$

The asymptotic normality of $\left(nh_n^3\phi_z(h_n)\right)^{1/2}\left(\widehat{h'}^x(\theta)-h'^x(\theta)\right)$ can be deduced from both following lemmas,

Lemma 3.4.1 *Under Assumptions (H1)-(H2) and (H6)-(H8), we have*

$$(n\phi_x(h_n))^{1/2}\left(\widehat{F}^x(y)-F^x(y)\right) \xrightarrow{\mathcal{D}} N\left(0,\sigma^2_{F^x}(y)\right) \tag{3.13}$$

where

$$\sigma^2_{F^x}(y)=\frac{a_2^y F^x(y)\left(1-F^x(y)\right)}{(a_1^y)^2}.$$

Lemma 3.4.2 *Under Assumptions (H1)-(H3) and (H5)-(H9), we have*

$$(nh_n\phi_x(h_n))^{1/2}\left(\widehat{h}^x(y)-h^x(y)\right) \xrightarrow{\mathcal{D}} N\left(0,\sigma^2_{h^x}(y)\right) \tag{3.14}$$

where

$$\sigma^2_{h^x}(y)=\frac{a_2^y h^x(y)}{(a_1^y)^2\left(1-F^x(y)\right)}\int_{\mathbb{R}} H^2(t)dt.$$

Lemma 3.4.3 *Under Assumptions of Theorem 3.4.1, we have*

$$\left(nh_n^3\phi_x(h_n)\right)^{1/2}\left(\widehat{f'}^x(y)-f'^x(y)\right) \xrightarrow{\mathcal{D}} N\left(0,\sigma^2_{f'^x}(y)\right) \tag{3.15}$$

where

$$\sigma^2_{f'^x(y)}=\frac{a_2^y f^x(y)}{(a_1^y)^2}\int_{\mathbb{R}}(H'(t))^2dt.$$

Lemma 3.4.4 *Under the hypotheses of Theorem 3.4.1, we have*

$$Var\left[\widehat{f'}^x_N(y)\right]=\frac{\sigma^2_{f'^x(y)}}{nh_n^3\phi_x(h_n)}+o\left(\frac{1}{nh_n^3\phi_x(h_n)}\right),$$

$$Var\left[\widehat{F}^x_N(y)\right]=o\left(\frac{1}{nh_n\phi_x(h_n)}\right);$$

and

$$Var\left[\widehat{F}^x_D\right]=o\left(\frac{1}{nh_n\phi_x(h_n)}\right).$$

Lemma 3.4.5 *Under the hypotheses of Theorem 3.4.1, we have*

$$Cov(\widehat{f'}^x_N(y), \widehat{F}^x_D) = o\left(\frac{1}{nh_n^3\phi_x(h_n)}\right),$$

$$Cov(\widehat{f'}^x_N(y), \widehat{F}^x_N(y)) = o\left(\frac{1}{nh_n^3\phi_x(h_n)}\right)$$

and

$$Cov(\widehat{F}^x_D, \widehat{F}^x_N(y)) = o\left(\frac{1}{nh_n\phi_z(h_n)}\right).$$

Remark 3.4.1

It is clear that, the results of lemmas, Lemma 3.4.4 and Lemma 3.4.5 allows to write

$$Var\left[\widehat{F}^x_D - \widehat{F}^x_N(y)\right] = o\left(\frac{1}{nh_n\phi_x(h_n)}\right).$$

The proofs of lemmas, Lemma 3.4.1 can be seen in Belkhir *et al.* (2015), Lemma 3.4.4 and Lemma 3.4.5 see Rabhi *et al.* (2015).

■

Finally, by this last result and (3.10), the following theorem follows:

Theorem 3.4.2 *Under conditions (H1)-(H11) we have* $(\theta \in \mathcal{S}/f^x(\theta), 1-F^x(\theta) > 0)$

$$\left(nh_n^3\phi_x(h_n)\right)^{1/2}\left(\widehat{\theta}-\theta\right) \xrightarrow{\mathcal{D}} N\left(0, \frac{\sigma^2_{h'}(\theta)}{(h''^x(\theta))^2}\right)$$

with $\sigma^2_{h'}(\theta) = h^x(\theta)\left(1-F^x(\theta)\right)\int (H'(t))^2 dt.$

3.5 Proofs of technical lemmas

Proof of lemma 3.3.1 and lemma 3.3.2. Let

$$\widehat{h'}^x(y) = \frac{\widehat{f'}^x(y)}{1-\widehat{F}^x(y)} + (\widehat{h}^x(y))^2, \tag{3.16}$$

with

$$\widehat{h'}^x(y) - h'^x(y) = \left\{\left(\widehat{h}^x(y)\right)^2 - (h^x(y))^2\right\} + \left\{\frac{\widehat{f'}^x(y)}{1-\widehat{F}^x(y)} - \frac{f'^x(y)}{1-F^x(y)}\right\} \tag{3.17}$$

for the first term of (3.17) we can write

$$\left|\left(\widehat{h}^x(y)\right)^2 - (h^x(y))^2\right| \leq \left|\widehat{h}^x(y) - h^x(y)\right|.\left|\widehat{h}^x(y) + h^x(y)\right| \tag{3.18}$$

because the estimator $\widehat{h}^x(\cdot)$ converge a.co. to $h^x(\cdot)$ we have

$$\sup_{y\in\mathcal{S}}\Big|\left(\widehat{h}^x(y)\right)^2 - (h^x(y))^2\Big| \leq 2\Big|h^x(\theta)\Big|\sup_{y\in\mathcal{S}}\Big|\widehat{h}^x(y) - h^x(y)\Big| \tag{3.19}$$

for the second term of (3.17) we have

$$\begin{aligned}\frac{\widehat{f}'^x(y)}{1-\widehat{F}^x(y)} - \frac{f'^x(y)}{1-F^x(y)} &= \frac{1}{(1-\widehat{F}^x(y))(1-F^x(y))}\left\{\widehat{f}'^x(y) - f'^x(y)\right\}\\ &+\frac{1}{(1-\widehat{F}^x(y))(1-F^x(y))}\left\{f'^x(y)\left(\widehat{F}^x(y)-F^x(y)\right)\right\}\\ &-\frac{1}{(1-\widehat{F}^x(y))(1-F^x(y))}\left\{F^x(y)\left(\widehat{f}'^x(y)-f'^x(y)\right)\right\}.\end{aligned}$$

Valid for all $y\in\mathcal{S}$. Which for a constant $C<\infty$, this leads

$$\sup_{y\in\mathcal{S}}\Big|\frac{\widehat{f}'^x(y)}{1-\widehat{F}^x(y)} - \frac{f'^x(y)}{1-F^x(y)}\Big| \leq$$

$$C\frac{\left\{\sup_{y\in\mathcal{S}}\Big|\widehat{f}'^x(y)-f'^x(y)\Big| + \sup_{y\in\mathcal{S}}\Big|\widehat{F}^x(y)-F^x(y)\Big|\right\}}{\inf_{y\in\mathcal{S}}\Big|1-\widehat{F}^x(y)\Big|}. \tag{3.20}$$

Therefore, the conclusion of the lemma follows from the following results:

$$\sup_{y\in\mathcal{S}}|\widehat{F}^x(y)-F^x(y)| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right). \tag{3.21}$$

$$\sup_{y\in\mathcal{S}}|\widehat{f}'^x(y)-f'^x(y)| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{nh_n^3\phi_x(h_n)}}\right). \tag{3.22}$$

$$\sup_{y\in\mathcal{S}}|\widehat{h}^x(y)-h^x(y)| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{nh_n\phi_x(h_n)}}\right). \tag{3.23}$$

$$\exists\delta>0 \;\text{ such that }\; \sum_1^\infty \mathbb{P}\left\{\inf_{y\in\mathcal{S}}|1-\widehat{F}^x(y)|<\delta\right\}<\infty. \tag{3.24}$$

The proofs of (3.21) and (3.22) appear in Ferraty *et al.* (2006), and (3.23) is proven in Ferraty *et al.* (2008).

- Concerning (3.24) by equation (3.21), we have the almost complete convergence of $\widehat{F}^x(y)$ to $F^x(y)$. Moreover,

$$\forall\varepsilon>0 \qquad \sum_{n=1}^\infty \mathbb{P}\left\{|\widehat{F}^x(y)-F^x(y)|>\varepsilon\right\}<\infty.$$

On the other hand, by hypothesis we have $F^x < 1$, i.e.

$$1 - \widehat{F}^x \geq F^x - \widehat{F}^x,$$

thus,

$$\inf_{y\in\mathcal{S}} |1-\widehat{F}^x(y)| \leq (1-\sup_{y\in\mathcal{S}} F^x(y))/2 \Rightarrow \sup_{y\in\mathcal{S}} |\widehat{F}^x(y)-F^x(y)| \geq (1-\sup_{y\in\mathcal{S}} F^x(y))/2.$$

In terms of probability is obtained

$$\mathbb{P}\left\{\inf_{y\in\mathcal{S}} |1-\widehat{F}^x(y)| < (1-\sup_{y\in\mathcal{S}} F^x(y))/2\right\}$$

$$\leq$$

$$\mathbb{P}\left\{\sup_{y\in\mathcal{S}} |\widehat{F}^x(y)-F^x(y)| \geq (1-\sup_{y\in\mathcal{S}} F^x(y))/2\right\} < \infty.$$

Finally, it suffices to take $\delta = (1 - \sup_{y\in\mathcal{S}} F^x(y))/2$ and apply the results (3.21) to finish the proof of the lemma.

■

Proof of lemma 3.4.2. We can write for all $y \in \mathcal{S}$

$$\begin{aligned}
\widehat{h}^x(y) - h^x(y) &= \frac{\widehat{f}^x(y)}{1-\widehat{F}^x(y)} - \frac{f^x(y)}{1-F^x(y)} \\
&= \frac{1}{\widehat{D}^x(y)}\Big\{ \left(\widehat{f}^x(y) - f^x(y)\right) + f^x(y)\left(\widehat{F}^x(y) - F^x(y)\right) \\
&\qquad - F^x(y)\left(\widehat{f}^x(y) - f^x(y)\right)\Big\}, \\
&= \frac{1}{\widehat{D}^x(y)}\Big\{ (1-F^x(y))\left(\widehat{f}^x(y) - f^x(y)\right) \\
&\qquad - f^x(y)\left(\widehat{F}^x(y) - F^x(y)\right)\Big\} \qquad (3.25)
\end{aligned}$$

with $\widehat{D}^x(y) = (1-F^x(y))\left(1-\widehat{F}^x(y)\right)$.

As a direct consequence of the Lemma 3.4.1, the result (3.26) (see Ezzahrioui and Ould-Saïd (2008)) and the expression (3.25), permit us to obtain the asymptotic normality for the conditional hazard estimator.

$$(nh_n\phi_x(h_n))^{1/2}\left(\widehat{f}^x(y) - f^x(y)\right) \xrightarrow{\mathcal{D}} N\left(0, \sigma^2_{f^x}(y)\right) \qquad (3.26)$$

where

$$\sigma^2_{f^x(y)} = \frac{a_2^y f^x(y)}{(a_1^y)^2}\int_{\mathbb{R}} (H(t))^2 dt$$

■

Proof of lemma 3.4.3. For $i = 1, \dots, n$, we consider the quantities $K_i = K\left(h_n^{-1} d(x, X_i)\right)$, $H'_i(y) = H'\left(h_n^{-1}(y - Y_i)\right)$ and let $\widehat{f'}^x_N(y)$ (resp. $\widehat{F}^x_D$) be defined as

$$\widehat{f'}^x_N(y) = \frac{h_n^{-2}}{n\,\mathbb{E}K_1} \sum_{i=1}^{n} K_i H'_i(y) \qquad \text{(resp. } \widehat{F}^x_D = \frac{1}{n\,\mathbb{E}K_1} \sum_{i=1}^{n} K_i\text{)}.$$

This proof is based on the following decomposition

$$\begin{aligned} \widehat{f'}^x(y) - f'^x(y) &= \frac{1}{\widehat{F}^x_D} \left\{ \left(\widehat{f'}^x_N(y) - \mathbb{E}\widehat{f'}^x_N(y) \right) - \left(f'^x(y) - \mathbb{E}\widehat{f'}^x_N(y) \right) \right\} \\ &\quad + \frac{f'^x(y)}{\widehat{F}^x_D} \left\{ \mathbb{E}\widehat{F}^x_D - \widehat{F}^x_D \right\} \end{aligned} \tag{3.27}$$

and on the following intermediate results.

$$\sqrt{n h_n^3 \phi_x(h_n)} \left(\widehat{f'}^x_N(y) - \mathbb{E}\widehat{f'}^x_N(y) \right) \xrightarrow{\mathcal{D}} N\left(0, \sigma^2_{f'^x}(y)\right) \tag{3.28}$$

where $\sigma^2_{f'^x}(y)$ is defined as in Lemma 3.4.3.

$$\lim_{n \to \infty} \sqrt{n h_n^3 \phi_x(h_n)} \left(\mathbb{E}\widehat{f'}^x_N(y) - f'^x(y) \right) = 0 \tag{3.29}$$

$$\sqrt{n h_n^3 \phi_x(h_n)} \left(\widehat{F}^x_D(y) - 1 \right) \xrightarrow{\mathbb{P}} 0, \quad \text{as} \quad n \to \infty. \tag{3.30}$$

- Concerning (3.28).

  By the definition of $\widehat{f'}^x_N(y)$, it follows that

$$\begin{aligned} \Omega_n &= \sqrt{n h_n^3 \phi_x(h_n)} \left(\widehat{f'}^x_N(y) - \mathbb{E}\widehat{f'}^x_N(y) \right) \\ &= \sum_{i=1}^{n} \frac{\sqrt{\phi_x(h_n)}}{\sqrt{n h_n}\mathbb{E}K_1} \left(K_i H'_i - \mathbb{E}K_i H'_i \right) \\ &= \sum_{i=1}^{n} \Delta_i, \end{aligned}$$

  which leads

$$Var(\Omega_n) = n h_n^3 \phi_x(h_n) Var\left(\widehat{f'}^x_N(y) - \mathbb{E}\left[\widehat{f'}^x_N(y) \right] \right). \tag{3.31}$$

  Now, we need to evaluate the variance of $\widehat{f'}^x_N(y)$. For this we have for all $1 \le i \le n$, $\Delta_i(x, y) = K_i(x) H'_i(y)$, so we have

$$\begin{aligned} Var(\widehat{f'}^x_N(y)) &= \frac{1}{\left(n h_n^2 \mathbb{E}[K_1(x)]\right)^2} \sum_{i=1}^{n} \sum_{j=1}^{n} Cov\left(\Delta_i(x, y), \Delta_j(x, y)\right) \\ &= \frac{1}{n\left(h_n^2 \mathbb{E}[K_1(x)]\right)^2} Var\left(\Delta_1(x, y)\right). \end{aligned}$$

Therefore

$$Var\left(\Delta_1(x,y)\right) \leq \mathbb{E}\left(H'^2_1(y)K_1^2(x)\right) \leq \mathbb{E}\left(K_1^2(x)\mathbb{E}\left[H'^2_1(y)|X_1=x\right]\right).$$

Now, by a change of variable in the following integral and by applying (H4) and (H7), one gets

$$\begin{aligned}\mathbb{E}\left(H'^2_1(y)|X_1=x\right) &= \int_{\mathbb{R}} H'^2\left(\frac{y-u}{h_n}\right) f^x(u)du \\ &\leq h_n \int_{\mathbb{R}} H'^2(t)\left(f^x(y-h_n t,x)-f^x(y)\right)dt \\ &\quad +h_n f^x(y)\int_{\mathbb{R}} H'^2(t)dt \\ &\leq h_n^{1+b_2}\int_{\mathbb{R}} |t|^{b_2}H'^2(t)dt + h_n f^x(y)\int_{\mathbb{R}} H'^2(t)dt \\ &= h_n\left(o(1)+f^x(y)\int_{\mathbb{R}} H'^2(t)dt\right). \qquad (3.32)\end{aligned}$$

By means of (3.32) and the fact that, as $n\to\infty$, $\mathbb{E}\left(K_1^2(x)\right) \longrightarrow a_2^y\phi_x(h_n)$, one gets

$$Var\left(\Delta_1(x,y)\right) = a_2^y\phi_x(h_n)h_n\left(o(1)+f^x(y)\int_{\mathbb{R}} H'^2(t)dt\right).$$

So, using (H8), we get

$$\begin{aligned}\frac{1}{n\left(h_n^2\mathbb{E}[K_1(x)]\right)^2}Var\left(\Delta_1(x,y)\right) &= \frac{a_2^y\phi_x(h_n)}{n\left(a_1^y h_n^2\phi_x(h_n)\right)^2}h_n\left(o(1)+f^x(y)\int_{\mathbb{R}} H'^2(t)dt\right) \\ &= o\left(\frac{1}{nh_n^3\phi_x(h_n)}\right) + \frac{a_2^y f^x(y)}{(a_1^y)^2 nh_n^3\phi_x(h_n)}\int_{\mathbb{R}} H'^2(t)dt.\end{aligned}$$

Thus as $n\to\infty$ we obtain

$$\frac{1}{n\left(h_n^2\mathbb{E}[K_1(x)]\right)^2}Var\left(\Delta_1(x,y)\right) \longrightarrow \frac{a_2^y f^x(y)}{(a_1^y)^2 nh_n^3\phi_x(h_n)}\int_{\mathbb{R}} H'^2(t)dt. \quad (3.33)$$

Indeed

$$\sum_{i=1}^n \mathbb{E}\Delta_i^2 = \frac{\phi_x(h_n)}{h_n\mathbb{E}^2K_1}\mathbb{E}K_1^2(H'_1)^2 - \frac{\phi_x(h_n)}{h_n\mathbb{E}^2K_1}\left(\mathbb{E}K_1H'_1\right)^2 = \Pi_{1n}-\Pi_{2n}. \qquad (3.34)$$

As for Π_{1n}, by the property of conditional expectation, we get

$$\Pi_{1n} = \frac{\phi_x(h_n)}{\mathbb{E}^2K_1}\mathbb{E}\left\{K_1^2\int H'^2(t)\left(f'^x(y-th_n)-f'^x(y)+f'^x(y)\right)dt\right\}.$$

Meanwhile, by (H1), (H3), (H7) and (H8), it follows that:

$$\frac{\phi_x(h_n)\mathbb{E}K_1^2}{\mathbb{E}^2K_1} \underset{n\to\infty}{\longrightarrow} \frac{a_2^y}{(a_1^y)^2},$$

which leads

$$\Pi_{1n} \underset{n\to\infty}{\longrightarrow} \frac{a_2^y f^x(y)}{(a_1^y)^2} \int (H'(t))^2 dt, \tag{3.35}$$

Regarding Π_{2n}, by (H1), (H3) and (H6), we obtain

$$\Pi_{2n} \underset{n\to\infty}{\longrightarrow} 0. \tag{3.36}$$

This result, combined with (3.34) and (3.35), allows us to get

$$\lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{E}\Delta_i^2 = \sigma^2_{f'^x}(y) \tag{3.37}$$

Therefore, combining (3.33) and (3.36)-(3.37), (3.28) is valid.

- Concerning (3.29).

  The proof is completed along the same steps as that of Π_{1n}. We omit it here.

- Concerning (3.30). The idea is similar to that given by Belkhir *et al.* (2015).

  By definition of $\widehat{F}_D^x$, we have

$$\sqrt{nh_n^3\phi_x(h_n)}(\widehat{F}_D^x - 1) = \Omega_n - \mathbb{E}\Omega_n,$$

where $\Omega_n = \frac{\sqrt{nh_n^3\phi_x(h_n)}\sum_{i=1}^n K_i}{n\mathbb{E}K_1}$. In order to prove (3.30), similar to Belkhir *et al.* (2015), we only need to proov $Var\ \Omega_n \to 0$, as $n \to \infty$. In fact, since

$$\begin{aligned} Var\ \Omega_n &= \frac{nh_n^3\phi_x(h_n)}{n\mathbb{E}^2K_1}(nVarK_1) \\ &\leq \frac{nh_n^3\phi_x(h_n)}{\mathbb{E}^2K_1}\mathbb{E}K_1^2 \\ &= \Psi_1, \end{aligned}$$

then, using the boundedness of function K allows us to get that:

$$\Psi_1 \leq Ch_n^3\phi_x(h_n) \to 0, \quad \text{as } n \to \infty.$$

It is clear that, the results of (3.21), (3.22), (3.24) and Lemma 3.4.4 permits us

$$\mathbb{E}\left(\widehat{F}_D^x - \widehat{F}_N^x(y) - 1 + F^x(y)\right) \longrightarrow 0,$$

and

$$Var\left(\widehat{F}_D^x - \widehat{F}_N^x(y) - 1 + F^x(y)\right) \longrightarrow 0;$$

then

$$\widehat{F}_D^x - \widehat{F}_N^x(y) - 1 + F^x(y) \xrightarrow{\mathbb{P}} 0.$$

Moreover, the asymptotic variance of $\widehat{F}_D^x - \widehat{F}_N^x(y)$ given in Remark 3.4.1 allows to obtain

$$\frac{nh_n\phi_x(h_n)}{\sigma_{F^x}^2(y)} Var\left(\widehat{F}_D^x - \widehat{F}_N^x(y) - 1 + \mathbb{E}\left(\widehat{F}_N^x(y)\right)\right) \longrightarrow 0.$$

By combining result with the fact that

$$\mathbb{E}\left(\widehat{F}_D^x - \widehat{F}_N^x(y) - 1 + \mathbb{E}\left(\widehat{F}_N^x(y)\right)\right) = 0,$$

we obtain the claimed result.

Therefore, the proof of this Lemma is completed.
■

Bibliography

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

[2] Belkhir, N., Rabhi, A., and Soltani, S. (2015). Exact asymptotic errors of the hazard conditional rate kernel, *Journal of Statistics Applications & Probability Letters. An International Journal*, **2**(3), 191–204.

[3] Benhenni, K., Ferraty, F., Rachdi, M. and Vieu, P. (2007). Local smoothing regression with functional data, *Comput. Statist.* **22**, 353–369.

[4] Besse, P., Cardot, H. and Stephenson, D. (2000). Autoregressive forecasting of some functional climatic variations, *Scand. J. Statist.* **27**, 673–687.

[5] Bosq, D. and Lecoutre, J. P. (1987). Théorie de l'estimation fonctionnelle, *ECONOMICA* (eds), Paris.

[6] Bouchentouf, A. A., Djebbouri, T., Rabhi, A., 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**(4), 301–322.

[7] Collomb, G., Härdle, W. and Hassani, S. (1987). A note on prediction via conditional mode estimation. *J. Statist. Plann. and Inf.*, **15**, 227-236.

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

[9] Estévez, G., Quintela, A. (1999). Nonparametric estimation of the hazard function under dependence conditions. *Commun. Statist. Theor. Meth.*, **28**(10), 2297–2331.

[10] Estévez-Pérez, G., Quintela-del-Rio, A. and Vieu, P. (2002). Convergence rate for cross-validatory bandwidth in kernel hazard estimation from dependent samples, *J. Statist. Plann. Inference.* **104**, 1–30.

[11] Ezzahrioui, M. and Ould-Saïd, E. (2008). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data, *Journal of Nonparametric Statistics*, **20**(1), 3–18.

[12] 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** No.2, 399-417.

[13] 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.

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

[15] Ferraty, F., Mas, A. and Vieu, P. (2007). Advances in nonparametric regression for functional variables, *Australian and New Zealand Journal of Statistics*, **49**, 1–20.

[16] Ferraty, F., Rabhi, A. and Vieu, P. (2008). Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle, *Rev. Roumaine Math. Pures Appl.* **53**, 1–18.

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

[18] Gasser, T., Hall, P. and Presnell, B. (1998). Nonparametric estimation of the mode of a distribution of random curves, *Journal of the Royal Statistical Society, Ser. B.*, **60**, 681–691.

[19] Laksaci, A. and Mechab, B. (2010). Estimation non paramétrique de la fonction de hasard avec variable explicative fonctionnelle cas des données spatiales, *Rev. Rom. J. Pure & Applied Math.*, **55**, 35–51.

[20] Lecoutre, J. P. and Ould-Saïd, E. (1993). Estimation de la densité et de la fonction de hasard conditionnelle pour un processus fortement mélangeant avec censure, *C. R. Math. Acad. Sci. Paris.* **314**, 295–300.

[21] Liu, R., Van Ryzin, J. (1985). A histogram estimator of the hazard rate with censored data, *Ann. Statist.*, **13**, 592–605.

[22] Masry, E. (2005). Non-parametric regression estimation for dependent functional data: Asymptotic normality, *Stoch. Process. Appl.* **115**, 155–177.

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

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

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

[26] Quintela-del-Rio, A. (2010). Recursive kernel hazard estimation of strong mixing data. *Far East J. Theor. Statist.*, **30**(2), 89–105.

[27] 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**(3), 301–323.

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

[29] Rice, J., Rosenblatt, M. (1976). Estimation of the log survival function and hazard function, *Sankhya A*, **36**, 60–78.

[30] Roussas, G. (1989). Hazard rate estimation under dependence conditions, *J. Statist. Plann. Inference.* **22**, 81–93.

[31] Samanta, M. (1989). Non-parametric estimation of conditional quantiles. *Stat. Probab . Lett.* **7**, No.5, 407-412.

[32] Sarda, P., Vieu, P. (1989). Empirical distribution function for mixing random variables, Application in nonparametric hazard estimation. *Statistics*, **20**(4), 559–571.

[33] Singpurwalla, N. D. and Wong, M. Y. (1983). Estimation of the failure rate - A survey of non-parametric methods. Part I: Non-Bayesian methods, *Commun. Stat. Theory Math.* **12**, 559–588.

[34] Spierdijk, L. (2008). Non-parametric conditional hazard rate estimation: A local linear approach, *Comput. Stat. Data Anal.* **52**, 2419–2434.

[35] Vere-Jones, D. (1970). Stochastic models for earthquake occurrence, *J. Roy. Statist. Soc. Ser. B.*, **32**, 1–62.

[36] Vieu, P. (1991). Quadratic errors for nonparametric estimates under dependence. *J. Multivariate Anal.*, **39**, 324–347.

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

[38] Watson, G. S. and Leadbetter, M. R. (1964). Hazard analysis, *Sankhyã*, **26**, 101–116.

[39] Youndjé, É., Sarda, P., Vieu, P. (1996). Optimal smooth hazard estimates, *Test*, **5**, 379–394.

Chapter 4

General Bibliography

Bibliography

General Bibliography

[1] Abramovich, F., Angelini, C. (2006). Bayesian maximum a posteriori multiple testing procedure. *Sankhyã*, **68**, 436-460.

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

[3] 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.

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

[5] Akritas, M., Politis, D. (2003). (ed.) Recent advances and trends in nonparametric statistics. Elsevier, Amsterdam.

[6] Aneiros-Pérez, G., Cardot, H., Estévez, G., Vieu, P. (2004). Maximum ozone concentration forecasting by functional non-parametric approaches, *Environmetrics*, **15**, 675-685.

[7] Antoniadis, I., Sapatinas, T. (2003). Wavelet methods for continuoustime prediction using Hilbert-valued autoregressive processes. *J. Multivariate. Anal.* **87**, 133-158.

[8] 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**, No.1, 45-53.

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

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

[11] Belkhir, N., Rabhi, A., and Soltani, S. (2015). Exact asymptotic errors of the hazard conditional rate kernel. *Journal of Statistics Applications & Probability Letters. An International Journal*, **2**, No.3, 191-204.

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

[13] Benko, M., Hardle, W. and Kneip, A. (2005). Common functional principal components. SFB 649 Economic Risk Discussion Paper, 2006-2010.

[14] Berlinet, A., Biau, G.and Rouvière, L. (2005). Parameter selection in modified histogram estimates. *Statistics*, **39**, 91-105.

[15] Berlinet, A., Cadre, B., and Gannoun, A. (2001). On the conditional L1-median and its estimation. *J. Nonparametr. Statist.* **13**, No.5, 631-645.

[16] Besse, P., Cardot, H., Stephenson, D. (2000). Autoregressive forecasting of some functional climatic variations. *Scand. J. Stat.* **27**, No.4, 673-687.

[17] Bosq, D., (1991). Modelization, non-parametric estimation and prediction for continuous time processes. In Nonparametric Functional estimation and Related Topics (Spetses, 1990), 509-529, NATO, Adv. Sci. Inst. Ser. C Math. Phys. Sci. **335**, Kluwer Acad. Publ., Dordrecht.

[18] Bosq, D. (2000). Linear process in function space. Lecture notes in Statistics. **149**, Springer-Verlag.

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

[20] Bosq, D., Lecoutre, J. P.(1987). Théorie de l'estimation fonctionnelle. ECONOMICA, Paris.

[21] 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**, No.4, 301-322.

[22] 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.

[23] Cadre, B. (2001). Convergent estimators for the L_1-median of a Banach valued random variable. *Statistics*, **35**, No.4, 509-521.

[24] Cardot, H. (2007). Conditional functional principal components analysis. *Scand. J. Stat.* **34**, 317-335.

[25] Cardot, H., Crambes, C., Sarda, P. (2004). Spline estimation of conditional quantiles for functional covariates. *C. R. Acad. Sci.* Paris. **339**, No.2, 141-144.

[26] Cardot, H., Ferraty, F., Mas, A., Sarda, P. (2003). Testing hypotheses in the functional linear model. *Scand. J. Stat.* **30**, No.1, 241-255.

[27] Cardot, H., Ferraty, F., Sarda, P. (1999). Functional linear model. *Stat. Probab. Lett.* **45**, No.1, 11-22.

[28] Chate, H., Courbage, M. (1997). Lattice systems. *Physica.* D, **103**, 1-612.

[29] Chiou, J., Müller, H.G. (2007). Diagnostics for functional regression via residual processes. *Computational Statistics and Data Analysis*, **51**, 4849-4863.

[30] Chiou, J. M., Müller, H. G., Wang, J. L. and Carey, J. R. (2003a). A functional multiplicative effects model for longitudinal data, with application to reproductive histories of female medflies. *Statist. Sinica.* **13**, 1119-1133.

[31] Chiou, J. M., Müller, H. G. and Wang, J. L. (2003b). Functional quasi-likelihood regression models with smooth random effects. *J. Royal Statist. Soc. Ser. B.*, **65**, 405-423.

[32] Collomb, G., (1976). Estimation non paramétrique de la régression. (in french). Ph.D. Université Paul Sabatier, Toulouse.

[33] Collomb, G., Härdle, W. and Hassani, S. (1987). A note on prediction via conditional mode estimation. *J. Statist. Plann. and Inf.* **15**, 227-236.

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

[35] Cuevas, A., Febrero, M., Fraiman, R. (2004). An anova test for functional data. *Computational Statistics & Data Analysis*, **47**, No.1, 111-122.

[36] Dabo-Niang, S. (2002). Sur l'estimation fonctionnelle en dimension infinie : Application aux diffusions. Thèse de Doctorat, Université de Paris 6.

[37] Dabo-Niang, S. and Laksaci, A. (2012). Nonparametric quantile regression estimation for functional dependent data. *Commun. Statist. Theo. Meth.* **41**, No.7, 1254-1268.

[38] 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**, No.1, 75-80.

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

[40] Delecroix, M, Härdle, W., Hristache, M. (1999). M-estimateurs semi-paramétriques dans les modèles à direction révélatrice unique. *Bull. Belg. Math. Soc.* Simon Stevin, **6**, No.2, 161-185.

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

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

[43] Delsol, L. (2007). Régression non paramétrique fonctionnelle: expression asymptotique des moments, *Ann. I.S.U.P.* **LI**, No.3, 43-67.

[44] Delsol, L. (2009). Advances on asymptotic normality in nonparametric functional Time Series Analysis Statistics. *Statistics*, **43**, 13-33.

[45] Delsol, L. (2011). Nonparametric methods for α-mixing functional random variables. In The Oxford Handbook of Functional Data Analysis (Ed. F. Ferraty and Y. Romain). Oxford University Press.

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

[47] El Ghouch, A. and Van Keilegom, I. (2009). Local linear quantile regression with dependent censored data. *Statist. Sinica*, **19**, No.4, 1621-1640.

[48] Estévez, G., Quintela, A. (1999). Nonparametric estimation of the hazard function under dependence conditions. *Commun. Statist. Theor. Meth.* **28**(10), 2297-2331.

[49] Estévez-Pérez, G., Quintela-del-Rio, A. and Vieu, P. (2002). Convergence rate for cross-validatory bandwidth in kernel hazard estimation from dependent samples, *J. Statist. Plann. Infer.* **104**, 1-30.

[50] Ezzahrioui, M. and Ould-Saïd, E. (2008). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data, *Journal of Nonparametric Statistics*, **20**, No.1, 3-18.

[51] Ezzahrioui, M., Ould Saïd, E. (2010). Some asymptotic results of a nonparametric conditional mode estimator for functional time series data. *Statist. Neerlandica*, **64**, No.2, 171-201.

[52] Ferraty, F. (2010). Special issue on statistical methods and problems in infinite dimensional spaces. *J. Multivariate Anal.* **101**, No.2, 305-490.

[53] 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.

[54] 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.

[55] 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.

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

[57] 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** No.2, 399-417.

[58] 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.

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

[60] 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.

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

[62] 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.

[63] 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.

[64] Ferraty, F., Vieu, P. (2006). Nonparametric Functional Data Analysis. Theory and Practice. Theory and Practice. Springer-Verlag.

[65] 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.

[66] Ferré, L., Villa, N. (2005). Discrimination de courbes par régression inverse fonctionnelle. *Revue de Statistique Appliquée*, **LIII**, No.1, 39-57.

[67] Frank, I. E., Friedman, J. H. (1993). A statistical view of some chemometrics regression tools (with discussion). Technometrics, **35**, 109-148.

[68] Fuk, D. Kh., Nagaev, S. V. (1971). Probability inequalities for sums of independent random variables, *Theory Prob. Appl.* **16**, 643-660.

[69] Gannoun, A., Saracco, J., and Yu, K. (2003). Nonparametric prediction by conditional median and quantiles. *J. Statist. Plann. Infer.* **117**, No.2, 207-223.

[70] Gannoun, A., Saracco, J., Yuan, A., and Bonney, G. E. (2005). Nonparametric quantile regression with censored data. S*cand. J. Statist.* **32**, No.4, 527-550.

[71] 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**, No.4, 681-691.

[72] Geffroy, J. (1974). Sur l'estimation d'une densité dans dans un espace métrique. *C. R. Acad. Sci.*, Paris, Sér. A, **278**, 1449-1452.

[73] Hall, P. (1989). On projection pursuit regression. *Ann. Statist.* **17**, No.2, 573-588.

[74] Hall, P., Heckman, N. E. (2002). Estimating and depicting the structure of a distribution of random functions. *Biometrika.* **89**, No.1, 145-158.

[75] Hall P, Poskitt D, Presnell B. (2001) A functional data-analytic approach to signal discrimination. *Technometrics.*; **43** : 1Ű24.

[76] Hall, P., Vial, C. (2006). Assessing the finite dimensionality of functional data. *Journal of the Royal Statistical Society. Series: B*, **68**, No.4, 689-705.

[77] Härdle, W. (1990). Applied nonparametric regression. Cambridge Univ. Press, UK.

[78] Härdle, W., Hall, P., Ichumira, H., (1993). Optimal smoothing in single-index models, *Ann. Statist.* **21**, 157-178.

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

[80] Härdle, W., Marron, J. S. (1990). Semiparametric comparison of regression curves. *The Annals of Statistics.* **18**, No.1, 59-63.

[81] Hastie, T., Buja, A., Tibshirani, R. (1995). Penalized discriminant analysis. *Ann. Stat.***23**, No.1, 73-102.

[82] Hastie, T., Mallows, C. (1993). A discussion of A Statistical View of Some Chemometrics Regression Tools by I.E. Frank and J.H. Friedman. *Techno metrics*, **35**, 140-143.

[83] Heckman, N. E. and R. H. Zamar (2000), Comparing the shapes of regression functions. *Biometrika*, **87**, No.1, 135-144.

[84] Holmstrome, I. (1963). On a method for parametric representation of the state of the atmosphere. *Tellus*, **15**, 127-149.

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

[86] Huber, P. J. (1985). Projection pursuit. *Ann. Statist.* **13**, No.2, 435-475.

[87] Hyndman, R. J., Ullah, Md. S. (2007). Robust forecasting of mortality and fertility rates : a functional data approach. *Computational Statistics and Data Analysis*, **51**, 4942-4956.

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

[89] Kirpatrick, M., Heckman, N. (1989). A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *J. Math. Bio.* **27**, No.4, 429-450.

[90] Kneip, A., Utikal, K. J. (2001). Inference for density families using functional principal component analysis. *Journal of the American Statistical Association*, **96**, 519-542.

[91] Koenker, R. and Bassett, J., G. (1978). Regression quantiles. *Econometrica*, **46**, No.1, 33-50.

[92] Kolmogorov, A. N., Tikhomirov, V .M. (1959). ϵ-entropy and ϵ-capacity. Uspekhi Mat. Nauk **14**, 3-86. (Engl Transl. Amer. Math. Soc. Transl. Ser). **2**, 277-364.

[93] Kuelbs, J., Li, W. (1993). Metric entropy and the small ball problem for Gaussian measures. *J. Funct. Anal.* **116**, 133-157.

[94] Laksaci, A., Lemdani, M., Ould Saïd, E., (2009). A generalized L1 - approach for a kernel estimator of conditional quantile with functional regressors: Consistency and asymptotic normality. *Statist. and Probab. Lett.* **79**, 1065-1073.

[95] 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.

[96] Lecoutre, J. P. and Ould-Saïd, E. (1993). Estimation de la densité et de la fonction de hasard conditionnelle pour un processus fortement mélangeant avec censure, *C. R. Math. Acad. Sci.* Paris, **314**, 295-300.

[97] Liang, H.-Y. and de Uña-Àlvarez, J. (2011). Asymptotic properties of conditional quantile estimator for censored dependent observations. *Ann. Inst. Statist. Math.* **63**, No.2, 267-289.

[98] Liu, R., Van Ryzin, J. (1985). A histogram estimator of the hazard rate with censored data, *Ann. Statist.* **13**, 592-605.

[99] Lucero, J. C. (1999). A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset-offset. *J. Acoust. Soc. Am.* **105**, 423-431.

[100] 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**, No.4, 392-410.

[101] Manteiga, W. G., Vieu, P. (2007). Statistics for Functional Data. *Computational Statistics and Data Analysis*, **51**, 4788-4792.

[102] Marron, J.S. (1987). A comparison of cross-validation techniques in density estimation, *Ann. Statist.* **15**, 152-162.

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

[104] Molenaar, P., Boomsma, D. (1987). The genetic analysis of repeated measures: the karhunen-loeve expansion. *Behavior Genetics*, **17**, 229-242.

[105] Nadaraya, E. (1964). On estimating regression. *Theory Prob. Appl.* **10**, 186-196.

[106] Nagaev, S. V. (1997). Some refinements of probabilistic and moment inequalities. *Teor: Veroyatnost. i Primenen (in russian)*, **42**, No.4, 832-838.

[107] Nagaev, S. V. (1998). Some refinements of probabilistic and moment inequalities. *Theory. Probab. Appl.* **42**, No.4, 707-713.

[108] Nerini, D., Ghattas, B. (2007). Classifying densities using functional regression trees: Applications in oceanology. *Computational Statistics and Data Analysis*, **51**, 4984-4993.

[109] Obhukov, V. (1960). The statistically orthogonal expansion of empirical functions. *American Geophysical Union*, 288-291.

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

[111] Ould-Saïd, E. (2006). A strong uniform convergence rate of kernel conditional quantile estimator under random censorship. *Statist. Probab. Lett.* **76**, No.6, 579-586.

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

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

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

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

[116] Quintela-del-Rio, A. (2010b). Recursive kernel hazard estimation of strong mixing data. *Far East J. Theor. Statist.* **30**(2), 89-105.

[117] 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**, No.3, 301-323.

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

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

[120] Ramsay, J., Silverman, B. (1997). Functional Data Analysis. Springer-Verlag.

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

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

[123] Rao, C. R. (1958). Some statistical methods for comparing growth curves. *Biometrics*, **14**, 1-17.

[124] Ramsay, J. O. (1982). When the data are functions. *Psychometrika*, **47**, No.4, 379-396.

[125] Rice, J., Rosenblatt, M. (1976). Estimation of the log survival function and hazard function, *Sankhyã: A*, **36**, 60-78.

[126] Rosenblatt, M. (1969). Conditional probability density and regression estimators. In Multivariate Analysis II, Ed. P.R. Krishnaiah. Academic Press, New York and London.

[127] Roussas, G. (1989). Hazard rate estimation under dependence conditions, *J. Statist. Plann. Inference.* **22**, 81-93.

[128] Sarda, P., Vieu, P. (2000). Kernel Regression. Smoothing and Regression: Approaches, Computation, and Application. Ed. M.G. Schimek, 43-70, Wiley Series in Probability and Statistics.

[129] Samanta, M. (1989). Non-parametric estimation of conditional quantiles. *Stat. Probab . Lett.* **7**, No.5, 407-412.

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

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

[132] Schimek, M. (2000). Smoothing and Regression : Approaches, computation, and application, Ed. M.G. Schimek, Wiley Series in Probability and Statistics.

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

[134] Spierdijk, L. (2008). Non-parametric conditional hazard rate estimation: A local linear approach, *Comput. Stat. Data Anal.* **52**, 2419-2434.

[135] 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**, No.6, 2493-2511.

[136] Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika*, **23**, 19-23.

[137] Van Der Vaart, A. W., van Zanten, J. H. (2007). Bayesian inference with rescaled Gaussian process priors. *Electronic Journal of Statistics*, **1**, 433-448.

[138] Vere-Jones, D. (1970). Stochastic models for earthquake occurrence, *J. Roy. Statist. Soc. Ser. B.* **32**, 1-62.

[139] Vieu, P. (1991). Quadratic errors for nonparametric estimates under dependence. *J. Multivariate Anal.* **39**, 324-347.

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

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

[142] Watson, G. S. and Leadbetter, M. R. (1964). Hazard analysis, *Sankhyã*, **26**, 101-116.

[143] Xia, X. An H. Z. (2002). An projection pursuit autoregression in time series. J. of Time Series Analysis. **20**, No.6, 693-714.

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