N° d'ordre :

REPUBLIQUE ALGERIENNE DEMOCRATIQUE & POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR & DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE DJILLALI LIABES
FACULTE DES SCIENCES EXACTES
SIDI BEL ABBES

THÈSE DE DOCTORAT

Présentée par

Mr. Yassine Amine HAMMOU

Spécialité : *Mathématiques*
Option : *Statistiques*

Intitulée

CONTRIBUTION À L'ESTIMATION NON PARAMÉTRIQUE POUR LES MODÈLES DE SURVIE CONDITIONNELLES

Soutenue le 23/05/2017
Devant le jury composé de :

Président :
Mr. Boubaker MECHAB, Maître de Conférences A à l'Université de Sidi Bel Abbès

Examinateurs :
Mr. Abdeldjebbar KANDOUCI, Professeur à l'Université de Saïda
Mr. Fethi MADANI, Maître de Conférences A à l'Université de Saïda

Directeur de thèse :
Mr. Abbès RABHI, Maître de Conférences A à l'Université de Sidi Bel Abbès

Propriétés Asymptotiques pour des Estimateurs Non Paramétriques à Covariable Fonctionnelle

Je dédie ce modeste travail :

A mes très chers parents qui ont tout fait pour que je réussisse dans mes études et ma vie professionnelle et qui voient aujourd'hui leurs efforts, leurs encouragements, leurs conseils et surtout leurs sacrifices couronnés par cette thèse de Doctorat. Ils ont veillé à mon éducation avec infiniment d'amour et d'affectation. Que Dieu me permette de leur rendre au moins une partie aussi infime soit-elle, de tout ce que je leur dois.

A ma très chère femme " Fadia " pour son soutien, sa patience, son dévouement, sa confiance indéfectible en moi et tous les bienfaits qu'elle m'apporte quotidiennement.

A mon très cher fils adoré " Rassim Djawed " qui illumine ma vie chaque jour que Dieu fait et qui me donne du courage pour continuer à avancer.

A mes deux chers frères " Sofiane Lotfi " et " Adil Kamel Nassim " ainsi qu'à leurs épouses et fils pour leur présence à mes cotés, en témoignage des liens fraternels solides qui nous unissent. Aucune dédicace ne saurait leur exprimer mes sentiments les plus chaleureux. Je leurs souhaite beaucoup de bonheur et de réussite dans leur vie.

A mes chers oncles et mes chères tantes d'Oran que je n'oublie pas, pour leurs prières et leur bienveillance, qui furent d'un apport considérable.

A ma tante de SBA et toute ma famille de près ou de loin qui m'ont beaucoup aidé, encouragé et soutenu.

A mes grands-parents maternelle et paternelle qui me manquent terriblement, pour tous les souvenirs merveilleux avec eux, ancrés à jamais dans ma mémoire, que Dieu les accueille dans son vaste paradis.

A tous ceux qui me sont chers et proches et qui m'ont aidé de près ou de loin.

" A tous ceux qui ont sacrifié leur temps pour la science et à tous ceux qui ont utilisé la science pour le bien et la prospérité de l'humanité. "

Remerciements

Avant tout je remercie le bon Dieu tout puissant qui m'a le plus aidé et qui m'a tout donné pour que je puisse entreprendre et mener à bien ces travaux de thèse. Grâce à lui j'ai réussi.

Ensuite, je tiens à exprimer ici ma profonde gratitude ainsi que mes sincères remerciements à mon directeur de thèse Dr. Abbès RABHI qui m'a fait l'honneur de diriger ce travail malgré les occupations et les responsabilités qu'il assumait, pour son soutien constant son encadrement et ses encouragements durant toute la période de la réalisation de cette thèse. Il a réussi à me transmettre le plaisir de la recherche et la rigueur intellectuelle pour mener à bien ce travail. Il m'a initié à la recherche en estimation fonctionnelle. Il m'a guidé et encouragé avec beaucoup de compétences. Que ce travail soit le modeste témoignage de ma très haute considération et ma profonde reconnaissance. Merci infiniment, Docteur, pour votre disponibilité, vos conseils et tout l'intérêt que vous avez accordé à mes travaux. Par la même occasion, qu'il trouve en ces mots l'expression de mon admiration et mon profond respect pour ses grandes qualités scientifiques et surtout humaines.

Je remercie l'ensemble de mes Professeurs, collègues et amis du Laboratoire de Mathématiques de l'Université "Djillali LIABES" de Sidi-Bel-Abbès qui m'ont apporté aide, soutien, sympathie et amitié ainsi que tous ceux pour qui le mot science rime encore avec conscience.

Un grand merci aussi et beaucoup de reconnaissance à mon ancien camarade de classe, mon ami intime et frère Dr. Soufiane MOKEDDEM, Maître de Conférences A, en Mathématiques à l'université de Sidi-Bel-Abbès, pour son soutien constant et chaleureux, pour ces orientations, son aide et pour son amitié et loyauté indéfectible. Je lui souhaite beaucoup de réussite et de bonheur dans sa vie.

Que les membres de jury trouvent ici mes remerciements les plus vifs pour avoir accepté d'honorer par leur jugement ce travail : Dr. Boubaker MECHAB, je suis très sensible à l'honneur que vous me faites en acceptant d'être le président du jury de cette thèse. Dr. Fethi MADANI et Dr. Abdeldjebbar KANDOUCI qui m'ont fait l'honneur de prendre connaissance de ce travail et d'en être examinateurs. Qu'ils trouvent ici l'expression de ma profonde reconnaissance.

Enfin, j'adresse mes remerciements à l'ensemble des personnes présentes à cette soutenance.

" Aller jusqu'au bout, ce n'est pas seulement résister, mais aussi se laisser aller. "

A. CAMUS

Résumé

Dans cette thèse, nous proposons d'étudier quelques paramètres fonctionnels. Premièrement 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 nous sommes intéressés à l'estimation de la fonction du hasard conditionnelle ainsi que le taux du hasard maximal, dont nous donnons nos premiers résultats lorsque l'échantillon considéré est non nécessairement i.i.d.

Dans un second temps nous supposons que la variable explicative est à valeurs dans un espace semi métrique (dimension infinie) et nous considérons l'estimation de la fonction de hasard conditionnelle par la méthode de noyau. Nous traitons les propriétés asymptotiques de cet estimateur dans le cas dépendant. Pour le cas où les observations sont dépendantes, 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 risque maximal.

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 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) X and a functional variable (explanatory variable) Z. The nonparametric model used to study the relation between Z and X 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.
Secondly we propose to study some functional parameters when the data are generated from a model of regression. We study two functional parameters.
Firstly, we are interested in the conditional hazard function estimation as the asymptotic normality, the results are given in the case when the variables are dependent.
Secondly, we suppose that the explanatory variable is valued in metric space (infinite dimension) and we consider the conditional hazard function estimation via kernel approach. We establish its asymptotic properties ; pointwise and almost surely convergence (with rate) in dependent case. As an application we discuss the impact of this result in functional non parametric prediction from the estimation maximum of the conditional hazard function. In the case, we establish the pointwise and almost surely convergence (with rate) of the kernel estimator of the conditional hazard function.
The maximum 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. using the kernel nonparametric estimates of the first derivative of the conditional hazard function, we establish uniform convergence properties and asymptotic normality of an estimate of the maximum in the context of strong mixing dependence. 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.
As far as we know, the problem of estimating the conditional hazard in the functional single index parameter for censored data was not attacked. In general the nonparametric estimation under censored data is new in the statistical literature. What doubtless makes, the originality of this thesis.

Table des matières

Chapitre 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 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 (1997, 2002) for the applied aspects, Bosq (2000) for the theoretical aspects, Ferraty and Vieu (2006) for non-parametric studyet 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) anf 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 peculiarity 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 the 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), présence éventuelle de dépendance entre les variables observées (phénomène commun dans les applications sismiques ou économétriques, par exemple) 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.

1.2 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.

1.2.1 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 et 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) andt Delsol et al. (2011). Other works were interested to estimating the regression function using different approaches : the method of k nearest neighbors by Burba et al. (2008), robust technical by Azzidine et al. (2008) and Crambes et al. (2008), the estimate by the simplified method of local polynomial by Barrientos-Marin *e*t al. (2010).

1.2.2 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. Norm convergence in L^2 for a regularized version (spline) of the preceding estimator was established by the same authors in 2000.

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 *e*t al. (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 et al. (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, 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). More recently, 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). 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 et al. (2004) 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, Rabhi and Vieu (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...

1.2.3 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 (2007). concerning animal biology, studies of the oviposition of medfly were made by several authors (Chiou, Müller, Wang and Carey (2003), Chiou, Müller and Wang (2003), Cardot (2007) and Chiou et 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). D'autres articles parmi lesquels Ferraty and Vieu (2002), Ferré and Yao (2005), Ferraty *e*t al. (2006), Ferraty and Vieu (2006), Aneiros-Pérez and Vieu (2007), Ferraty, Mas and Vieu (2007) and Mas and Pumo (2007) they proposed and applied other methods to meet this problematic.

• Of environment-related applications have been particularly studied by Aneiros-Perez, Cardot, Estevez-Perez and Vieu (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 Cardot and Stephenson (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 Buja and Tibshirani (1995), Berlinet Biau and Rouvière (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 Buja and Tibshirani (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 rerecently by Benko, Härdle and Kneip (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.

1.2.4 On the conditional distribution

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 et al *et* 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 et al. (2005) and Ezzahrioui and Ould Saïd (2010). we send back to Cardot et al. (2004) for one approach for linear the conditional quantile statistical functional. The contribution of the thesis on this model is the study of the squared error and the uniform convergence on arguments to simple functional index of the estimator of the conditional distribution function and the conditional density. The asymptotic results (with rates) are precised. The results obtain The results are detailed in Chapter 2 of this thesis. These are the first consistent results available in the literature of estimating the conditional distribution function and conditional hazard function in the functional single index parameter for complete (uncensored) data and/or censored.

1.2.5 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 conditional hazard function, when the covariate is functional and the observations are linked with a single-index structure. We establish the pointwise almost complete convergence and the uniform almost complete convergence (with the rate) of the kernel estimate of this model in various situations, including censored and non-censored data. These first uniform results are detailed in the chapter 3.

1.2.6 On analysis of survival data

Survival analysis is the name of a collection of statistical techniques that is concerned with the modeling of lifetime data. These methods are used to describe, quantify and understand the stochastic behavior of time-to-events. In survival analysis we use the term "failure" for the occurrence of the event of interest (even though the event may actually be a "success", such as recovery from therapy). On the other hand the term "survival time" specifies the length of time taken for failure to occur, usually denoted T, that is assumed to be a positive random variable. Survival analysis methods have been used in a number of applied fields, such as medicine, public health, biology, epidemiology, engineering, economics, finance, social sciences, psychology and demography. The analysis of failure time data usually means addressing one of three problems : the estimation of survival functions, the comparison of treatments or survival functions, and the assessment of covariate effects or the dependence of failure time on explanatory variables.

The survival function at time t is defined as

$$S(t) = \mathbb{P}(T > t) = \int_t^{\infty} f(u)du = 1 - F(t) \tag{1.1}$$

where f and F are the density and distribution function of T, respectively, and it can be interpreted as the proportion of the population that survives up to time t. The empirical survival function is a non-parametric estimator of the unconditional survival function for complete data and is given by

$$\widehat{S}(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{t_i > t\}} = 1 - \widehat{F}(t)$$

The conditional survival function is the probability that the individual will be alive at time t given a time-fixed covariate, z_0 :

$$S(t|z_0) = \mathbb{P}(T > t|Z = z_0)$$

where Z is the covariate and z_0 is a fixed value. Not only are the lifetime and its covariate random variables unknown, but usually the conditional survival function is also unknown and needs to be estimated. There are many reasons that make it difficult to get complete data in studies involving survival times. A study is often finished before the death of all patients, and we may keep only the information that some patients are still alive at the end of the study, not observing when they really die. In the presence of censored data, the time to event is unknown, and all we know is that the survival time has occurred before, between or after certain time points. This obviates the need for inference methods for censored data.

When the failure time is observed completely, there are numerous methods to make non parametric inference on its conditional distribution. For instance Nadaraya (1964) and Watson (1964) proposed a nonparametric estimator (NW) to estimate the conditional expectation $\mu(z_0) = \mathbb{E}(T|Z = z_0)$

as as a locally weighted average using a kernel function. Beran (1981) extended the Kaplan-Meier estimator and proposed a method for non-parametric estimation (generalized Kaplan-Meier) of the conditional survival function for right-censored data. Turnbull (1976) proposed a nonparametric estimator of the unconditional survival function under interval-censoring.

Our objectives in this thesis are mainly to present simple non-parametric or semiparametric approaches to estimate the conditional hazard function when the data are generated from a model of regression to a single index under complete and/or censored data.

1.2.7 On The Hazard Function

An alternative characterization of the distribution of T is given by the hazard function, or instantaneous rate of occurrence of the event, defined as

$$h(t) = \lim_{\Delta t \to 0} \frac{\mathbb{P}(t < T \leq t + \Delta t\ ,\ T \geq t)}{\Delta t} \quad [t > 0]$$

The numerator of this expression is the conditional probability that the event will occur in the interval $(t, t + \Delta t)$ given that it has not occurred before, and the denominator is the width of the interval. Dividing one by the other we obtain a rate of event occurrence per unit of time. Taking the limit as the width of the interval goes down to zero, we obtain an instantaneous rate of occurrence.

The conditional probability in the numerator may be written as the ratio of the joint probability that T is in the interval $(t, t+\Delta t)$ and $T > t$ (which is, of course, the same as the probability that t is in the interval), to the probability of the condition $T > t$. The former may be written as $f(t)\Delta t$ for small Δt, while the latter is $S(t)$ by definition. Dividing by Δt and passing to the limit gives the useful result

$$h(t) = \frac{f(t)}{S(t)} \tag{1.2}$$

which some authors give as a definition of the hazard function. In words, the rate of occurrence of the event at duration t equals the density of events at t, divided by the probability of surviving to that duration without experiencing the event.

Note from Equation (1.1) that $-f(t)$ is the derivative of $S(t)$. This suggests rewriting Equation (1.2) as

$$h(t) = -\frac{d}{dt} \log S(t). \tag{1.3}$$

If we now integrate from 0 to t and introduce the boundary condition $S(0) = 1$ (since the event is sure not to have occurred by duration 0), we can solve the above expression to obtain a formula for the probability of surviving to duration t as a function of the hazard at all durations up to t :

$$S(t) = \exp\left\{-\int_0^t h(u)du\right\}. \tag{1.4}$$

This expression should be familiar to demographers. The integral in curly brackets in this equation is called the cumulative hazard (or cumulative risk) and is denoted

$$H(t) = \int_0^t h(u)du. \tag{1.5}$$

You may think of $H(t)$ as the sum of the risks you face going from duration 0 to t.

These results show that the survival and hazard functions provide alternative but equivalent characterizations of the distribution of T. Given the survival function, we can always differentiate to to obtain the density and then calculate the hazard using Equation (1.2). Given the hazard, we can always integrate to obtain the cumulative hazard and then exponentiate to obtain the survival function using Equation (1.4). An example will help fix ideas.

Example 1.2.1 *The simplest possible survival distribution is obtained by assuming a constant risk over time, so the hazard is*

$$h(t) = \lambda$$

for all t. The corresponding survival function is

$$S(t) = \exp(\lambda t).$$

This distribution is called the exponential distribution with parameter λ. The density may be obtained multiplying the survivor function by the hazard to obtain

$$f(t) = \lambda \exp(-\lambda t).$$

The mean turns out to be $1/\lambda$. This distribution plays a central role in survival analysis, although it is probably too simple to be useful in applications in its own right.

1.2.8 Convergence notions

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 1.2.1 *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

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

Definition 1.2.2 *One says that the rate of almost omplete 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 1.2.1 *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 1.2.1 *The almost convergence of Y_n to l_0 implies that there existe 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 convegence properties needs to find an upper bound for some probabilities involving sum of r.r.v. such as

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

where, eventually, the positive real ε decreases with n. In this context, there exists powerful probabilistic tools, generically called Exponential 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 1.2.2 *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 :*

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

Corollary 1.2.1 *i) If* $\forall m \geq 2, \exists C_m > 0,\ \mathbb{E}|Z_1^m| \leq C_m a^{2(m-1)}$, *we have*

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

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 1.2.2 *By applying Proposition 1.2.2 with* $A_n = a\sqrt{u_n}$, $b = a^2$ *and taking* $\varepsilon = \varepsilon_0\sqrt{u_n}$, *we obtain for some* $C' > 0$ *:*

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

Corollary 1.2.2 *i) If* $\exists M < \infty,\ |Z_1| \leq M$, *and denoting* $\sigma^2 = \mathbb{E}Z_1^2$, *we have*

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

ii) Assume that the variables depend on n (that is, $Z_i = Z_{i,n}$) and are such that $\exists M = M_n < \infty$, $|Z_1| \le M$ 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 1.2.3 *By applying Proposition 1.2.2 with $a_i^2 = \sigma^2$, $A_n = n\sigma^2$, and by choising $\varepsilon = \varepsilon_0\sqrt{u_n}$, we obtain for some $C' > 0$:*

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

Where $v_n = \frac{Mu_n}{\sigma_n^2}$

1.3 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.

1.3.1 Why Use Kernel Methods for Functional Data ?

Kernel methods are well-known and intensively used by the community of nonparametricians because they are a useful way to do local weithing. We start by recalling shortly what is kernel local weighting in the real and multivariate cases before extending it to the functional context.

1. **Real Case**

As it well known, kernel local weighting is based on a kernel function (classically denoted by K) and on a smoothing parameter, which is called bandwidth and usually denoted by h. If is a fixed real number, the kernel local weithing transforms n r.r.v. $X_1, X_2, \ldots, X_n$ into $\Delta_1, \Delta_2, \ldots, \Delta_n$ such that :

$$\Delta_i = \Delta_i(x, h, K) = \frac{1}{h} K \left(\frac{x - X_i}{h} \right).$$

The main ideas of the local weithing around x is to attribute at each r.r.v. X_i a weight taking into account the distance between x and X_i ; the more X_i is distant from x, the smaller is the weighting.

Before going on, let us recall what is a kernel function exactly in this simplest situation. In fact, there exists a large variety of kernels. Any density function can be considered as a kernel, but even unnecessary positive functions can be used Gasser and Müller (1979). To simplify our purpose, we consider at this stage only positive and symmetrical kernels which are the most classical ones.

To precise the notion of kernel local weighting, let us consider the Box kernel and rewrite the Δ_i's as follows :

$$\Delta_i = \frac{1}{h} \mathbf{1}_{[x-h, x+h]}(X_i).$$

In this situation, the local feature of the weighting appears obvious since the r.r.v. outside the rang $[x - h, x + h]$ are ignored. In addition, the normalization $1/h$ is proportional to the size of the set $[x-h, x+h]$ on which the X_i's are taken into account. These points are not only true for the Box kernel, but are shared by any compactely supported kernels.

2. **Multivariate Case**

In multivariate situations one is observing n random vectors $\mathbf{X}_1, \ldots, \mathbf{X}_n$ valued in $\mathbb{R}^p$. The previous kernel local weighting can be extended easily to this situation. To that end, it suffices to consider a multivariate kernel K^* which will be a functions from $\mathbb{R}^p$ into $\mathbb{R}$. The first (natural) way to do that is to define K^* as a product of p real kernel functions $K_1, K_2, \ldots, K_p$:

$$\forall \mathbf{u} =^t (u_1, \ldots, u_p) \in \mathbb{R}^p, \; K^*(\mathbf{u}) = K_1(u_1) \times K_2(u_2) \times \ldots \times K_p(u_p).$$

As pointed out in Härdle and Müller (2000), a second way consists in combining a real kernel function H with a norm (denoted by $\|.\|$) in $\mathbb{R}^p$ as follows :

$$\forall \mathbf{u} \in \mathbb{R}^p,\ K^*(\mathbf{u}) = K(\|\mathbf{u}\|).$$

Note that if $K_1 = K_2 = \ldots = K_p = \mathbf{1}_{[-1,1]}$ and if $\|\cdot\|$ is the supremum norm, both approaches coincide by taking $K = \mathbf{1}_{[-1,1]}$. Moreover, because $\|\mathbf{u}\|$ is always a positive quantity, the real kernel K should have a positive support (i.e., $\{v \in \mathbb{R}\ \text{ such that } K(v) > 0\} \subset \mathbb{R}^+$). This leads to use asymmetrical functions for the kernel K.

Now, let us discuss how this can be interpreted in terms of local weighting. Indeed, what happens is very similar to the real case. Let $\mathbf{x}$ be a fixed vector of $\mathbb{R}^p$. The multivariate kernel local weighting consists in transforming the n random vectors $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ into the n variables $\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2, \ldots, \boldsymbol{\Delta}_n$:

$$\Delta_i = \frac{1}{h^p} K^* \left(\frac{\mathbf{x} - \mathbf{X}_i}{h} \right).$$

If we consider compactly supported kernels, it appears clearly that the Δ_i are locally weighted transformations of the variables $\mathbf{X}_i$ is out of some neighborhood of $\mathbf{x}$. Moreover, the normalization $1/h^p$ is proportional to the volume of the set on which the $\mathbf{X}_i$'s are taken into account.

3. **Functional Case**

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 (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 :

$$\Delta_i = \frac{K\left(\frac{d(x,X_i}{h}\right)}{E\left(K\left(\frac{d(x,X_i}{h}\right)\right)}. \tag{1.6}$$

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 resultis known in the literature as the Bochner's type theorem and Collomb (1976) gives a large scope on such results). So, it is clear now that (1.6) 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.

Definition 1.3.1

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 1.3.2 *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 (1.6), 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 pratical points of view. Because tne 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 1.3.1 *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 1.3.2 *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 :

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

1.4 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 :

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

In addition, if the probability distribution of Y given X is absolutely continuous with respect to the Lebesgue measure, we note $f_Y^X(x, y)$ the value of the corresponding density function at (x, y). Note that under a differentiabillity assumption on $F_Y^X(x, .)$, this functional conditional density can be written as

$$\forall y \in \mathbb{R},\quad f_Y^X(x, y) = \frac{\partial}{\partial y} F_Y^X(x, y). \tag{1.7}$$

For these two last definitions, we are implicitly assuming that there exists a regular version of this conditional probability. This assumption will be done implicitly as long as we will need to introduce this condional cdf $F_Y^X(x, y)$ or the conditional density $f_Y^X(x, y)$.

It is clear that each of these nonlinear operators gives information about the link between X, Y and thus can be useful for predicting method. The first way to construct such a prediction is obtained directly from the regression operator by putting :

$$\widehat{y} = \widehat{r}(x),$$

$\widehat{r}$ being an estimator of r. The second one consists of considering the median $m(x)$ of the conditional c.f.d. F_Y^X :

$$m(x) = \inf\{y \in \mathbb{R},\ F_Y^X(x,y) \geq 1/2\},$$

and to use as predictor :

$$\widehat{y} = \widehat{m}(x),$$

where $\widehat{m}(x)$ is an estimator of this functional conditional median $m(x)$. Note that such a conditional median estimate will obviously depend on some previous estimation of the nonlinear operator F_Y^X. Finally, the third predictor is based directly on the mode $\theta(x)$ of the conditional density of Y given X :

$$\theta(x) = \arg\sup_{y \in S} f_Y^X(x,y).$$

This definition assumes implicitly that $\theta(x)$ exists on S. The predictor is defined by :

$$\widehat{y} = \widehat{\theta}(x),$$

where $\widehat{\theta}(x)$ is an estimator of this functional conditional mode $\theta(x)$. Once again note that this conditional mode estimate will directly depend on some previous estimation of the nonlinear operator f_Y^X.

1.5 Kernel Estimators

Once the nonparametric modelling 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 :

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

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.

– **Estimating the conditional c.d.f.**. We focus now on the estimator $\widehat{F}_Y^X$ of the conditional c.d.f. F_Y^X, but let us first explain how we can extend the idea previously used for the construction of the kernel regression estimator. Clearly, $F_Y^X = \mathbb{P}\left(Y \leq y | X = x\right)$ can be expressed in terms of conditional expectation :

$$F_Y^X = \mathbb{E}\left(\mathbf{1}_{(-\infty, y]}(Y) | X = x\right)$$

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

$$\widetilde{F}_Y^X(x, y) = \frac{\sum_{i=1}^{n} K\left(h^{-1}d(x, X_i)\right) \mathbf{1}_{(-\infty, y]}(Y_i)}{\sum_{i=1}^{n} K\left(h^{-1}d(x, X_i)\right)}.$$

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

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

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

$$\widehat{F}_Y^X(x,y) = \frac{\sum_{i=1}^n K\left(h^{-1}d(x,X_i)\right) H(g^{-1}(y-Yi)}{\sum_{i=1}^n K\left(h^{-1}d(x,X_i)\right)}, \tag{1.8}$$

where g is a strictly positive real number (depending on n). To fix the ideas, let us consider K_0 as a kernel of type 0 see Definition (1.3.2). In this case, H is a c.f.d. and the quantity $H(g^{-1}(y-Y_i))$ acts as a local weighting : when Y_i is less than y the quantity $H(g^{-1}(y-Y_i))$ is large and the more Y_i is above y, the smaller the quantity $H(g^{-1}(y-Y_i))$. Moreover.
It is clear that the parameter g acts as the bandwidth h. The smoothness of the function $\widehat{F}_Y^X(x,.)$ is controled both by the smoothing parameter g and by the regularity of the c.d.f. H. The idea to build such a smooth c.d.f. estimate was introduced by Azzalini and Reiss (1981). The roles of the other parameters invoved in this functional kernel c.d.f. estimate (i.e. the roles of K and h) are the same as in the regression setting. From this conditional c.d.f. estimate (1.8), one can attack the prediction problem by defining a kernel estimator of the funvtional conditional median $m(x)$ as follows :

$$\widehat{m}(x) = \inf\{y \in \mathbb{R},\ \widehat{F}_Y^X(x,y) \geq 1/2\}.$$

More generally, we can also define from (1.8) a kernel estimator of the functional conditional quantiles $t_\alpha(x)$, for any α in $[1, 1/2]$, as follows :

$$\widehat{t}_\alpha(x) = \inf\{y \in \mathbb{R},\ \widehat{F}_Y^X(x,y) \geq \alpha\}.$$

– **Estimating the conditional density**. It is know that, under some differentiability assumption, the conditional density function can be obtained by derivating the conditional c.d.f. (see (1.7)). Since we have now at hand some estimator $\widehat{F}_Y^X$ of F_Y^X, it is natural to propose the following estimate :

$$\widehat{f}_Y^X(x,y) = \frac{\partial}{\partial y}\widehat{F}_Y^x.$$

Assuming the differentiability of H, we have

$$\frac{\partial}{\partial y}\widehat{F}_Y^x = \frac{\displaystyle\sum_{i=1}^{n} K\left(h^{-1}d(x,X_i)\right)\frac{\partial}{\partial y}H(g^{-1}(y-Y_i))}{\displaystyle\sum_{i=1}^{n} K\left(h^{-1}d(x,X_i)\right)},$$

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

$$\widehat{f}_Y^X(x,y) = \frac{\sum_{i=1}^{n} K\left(h^{-1}d(x,X_i)\right)\frac{1}{g}H'(g^{-1}(y-Y_i))}{\sum_{i=1}^{n} K\left(h^{-1}d(x,X_i)\right)}.$$

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

$$\widehat{f}_Y^X(x,y) = \frac{\displaystyle\sum_{i=1}^{n} K\left(h^{-1}d(x,X_i)\right)\frac{1}{g}HK_0(g^{-1}(y-Y_i))}{\displaystyle\sum_{i=1}^{n} K\left(h^{-1}d(x,X_i)\right)}.$$

This kind of estimate has been widely studied in the un-functional setting, that is, in the setting when X is changed into a finite dimensional variable. Concerning the parameters involved in the functional part of the estimate (namely, the roles of K and h) are the same as in the regression setting discussed just before while those involved in the un-functional part (namely, K_0 and g) are acting exactly as K and h, respectively as a weight function and as a smoothing factor.

To end, note that we can easily get the following kernel functional conditional mode estimator of $\theta(x)$:

$$\widehat{\theta}(x) = \arg\sup_{y\in S}\widehat{f}_Y^X(x,y).$$

1.6 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 hazard function considering two types of data namely full data and censored right into a type of correlation which is none other than the i.i.d case.The explanatory variable for functional parameter which is the conditional hazard 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.

1.6.1 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 reponse variable and the case of i.i.d observations. In this context, we study the mean square convergence of kernel estimators of the conditional distribution function and the conditional density. Then, we derive results on the estimator of the conditional hazard function. In the second part, we examine the conditional hazard function and we focus on the situation where the covariate is uncensored and/or right-censored and always in case of i.i.d observations. We build in this case a kernel estimator for this functional parameter.We establish the pointwise almost complete convergence and the uniform almost complete convergence (with the rate) of this estimator. The interest of our study is to show how the estimation of the conditional density can be used to obtain an estimate of the simple functional index if it is unknown. More specifically, this parameter can be estimated by the method pseudo-maximum the likelihood which is based on the preliminary estimation of the conditional density.

We will finish this section with some prospects research.

1.7 Short presentation of the results

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

1.7.1 Notations

Let (X, Y) a random pair where Y is valued in $\mathbb{R}$ and X is valued in some semi-normed vector space $(\mathcal{F}, d(.;.))$ which can be of infinite dimension. We will say that X is a functional random variable and we will use the abbreviation *frv*.

For $x \in \mathcal{F}$, we will denote the *cond-cdf* of Y given $X = x$ (respect. the conditional survival function) by

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

$$(\text{resp.} \quad S^x(y) = 1 - F^x(y))$$

If this distribution is absolutely continuous with respect to the Lebesgues measure on $\mathbb{R}$, then we will denote by f^x the conditional density of Y given $X = x$.

Let $(X_i, Y_i)_{i=1,\dots,n}$ be the be the statitical sample of pairs which are identically distributed like (X, Y), but not necessarily independant.

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

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

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

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.

In the following (x, y) will be a fixed point in $\mathbb{R} \times \mathcal{F}$ and $N_x \times N_y$ will denote a fixed neighborhood of (x, y), S will be a fixed compact subset of $\mathbb{R}$, and we will use the notation $B(x, h) = \{x' \in \mathcal{F}/d(x', x) < h\}$.

1.8 The model

We consider a random pair (X, Y) where Y is valued in $\mathbb{R}$ and X is valued in some infinite dimensional semi-metric vector space $(\mathcal{F}, d(., .))$. Let (X_i, Y_i), $i = 1, \ldots, n$ be the statistical sample of pairs which are identically distributed like (X, Y), but not necessarily independent. From now on, X is called functional random variable f.r.v. Let x be fixed in $\mathcal{F}$ and let $F_{X|Y}(\cdot, x)$ be the conditional cumulative distribution function (cond-cdf) of Y given $X = x$, namely :

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

Let $Q_{Y|X}(\gamma)$ be the $\gamma-$order quantile of the distribution of Y given $X = x$. From the cond-cdf $F_{Y|X}(., x)$, it is easy to give the general definition of the γ-order quantile :

$$Q(\gamma | X = x) \equiv Q_{Y|X}(\gamma) = \inf\left\{t : F_{Y|X}(t, x) \geq \gamma\right\}, \quad 0 \leq \gamma \leq 1 \tag{1.9}$$

Then, the definition of conditional quantile implies that

$$F_{Y|X}\left(Q_{Y|X}(\gamma)\right) = \gamma.$$

On differentiating partially w.r.t. γ we get

$$f_{Y|X}\left(Q_{Y|X}(\gamma)\right) = \frac{1}{\frac{\partial}{\partial \gamma}\left(Q_{Y|X}(\gamma)\right)}.$$

Thus, the condition quantile density function can be written as follows (see Xiang (1995))

$$\begin{aligned} q(\gamma | X = x) \equiv q_{Y|X}(\gamma) &= \frac{\partial}{\partial \gamma}\left(Q_{Y|X}(\gamma)\right) \\ &= \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)} \end{aligned} \tag{1.10}$$

Let us now, define the kernel estimator $\widehat{F}_{Y|X}(.,x)$ of $F_{Y|X}(.,x)$

$$\widehat{F}_{Y|X}(x,y)=\frac{\sum_{i=1}^{n}K\left(h_K^{-1}d(x,X_i)\right)H\left(h_H^{-1}(y-Y_i)\right)}{\sum_{i=1}^{n}K\left(h_K^{-1}d(x,X_i)\right)} \tag{1.11}$$

where K is a kernel function, H a cumulative distribution function and $h_K=h_{K,n}$ (resp. $h_H=h_{H,n}$) a sequence of positive real numbers. Note that using similar ideas, Roussas (1969) introduced some related estimate but in the special case when X is real, while Samanta (1989) produced previous asymptotic study. It is easy to derive an estimator $\widehat{Q}_{Y|X}$ of $Q_{Y|X}$:

$$\widehat{Q}_{Y|X}(\gamma)=\inf\left\{t:\widehat{F}_{Y|X}(t,x)\geq\gamma\right\}=\widehat{F}_{Y|X}^{-1}\left(Q_{Y|X}(\gamma)\right) \tag{1.12}$$

Let

$$\widehat{F}_{Y|X}^{(j)}(x,y)=\frac{h_H^{-j}\sum_{i=1}^{n}K\left(h_K^{-1}d(x,X_i)\right)H^{(j)}\left(h_H^{-1}(y-Y_i)\right)}{\sum_{i=1}^{n}K\left(h_K^{-1}d(x,X_i)\right)} \tag{1.13}$$

be the jth successive derivative of $\widehat{F}_{Y|X}(x,y)$, $f_{Y|X}(x,y)$ is conditional density function, such that $f_{Y|X}(x,y)=F_{Y|X}^{(1)}(x,y)$.

The kernel type quantile density estimator$K(\cdot)$. The kernel is satisfying the following properties :

The smooth estimator of the conditional quantile density functional defined as follows :

$$\widehat{q}_{Y|X}(\gamma)=\frac{1}{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right)}. \tag{1.14}$$

where $\widehat{f}_{Y|X}(x,y)$ is a conditional kernel density estimator of $f_{Y|X}(x,y)$ and $\widehat{Q}_{Y|X}(\gamma)$ is the conditional empirical estimator of the conditional quantile function $Q_{Y|X}(\gamma)$. Let

$$\widehat{f}_{Y|X}(x,y)=\frac{h_H^{-1}\sum_{i=1}^{n}K\left(h_K^{-1}d(x,X_i)\right)H^{(1)}\left(h_H^{-1}(y-Y_i)\right)}{\sum_{i=1}^{n}K\left(h_K^{-1}d(x,X_i)\right)},$$

and

$$\widehat{f}^{j}_{Y|X}(x,y) = \frac{h_H^{-j-1}\sum_{i=1}^{n} K\left(h_K^{-1}d(x,X_i)\right)H^{(j+1)}\left(h_H^{-1}(y-Y_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}d(x,X_i)\right)}$$

$h_K = h_{K,n}$ (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 $0/0 = 0$.

Theorem 1 *Let $q_{Y|X}(\gamma)$ be the conditional density function corresponding to a density function $f_{Y|X}\left(Q_{Y|X}(\gamma)\right)$ and $\widehat{q}_{Y|X}(\gamma)$ denote the estimator of $q_{Y|X}(\gamma)$. Then as n tends to infinity, we have*

$$\sup_{\gamma}\left|\widehat{q}_{Y|X}(\gamma) - q_{Y|X}(\gamma)\right| \longrightarrow 0 \qquad a.co.$$

Theorem 2 *Let $q_{Y|X}(\gamma)$ be the conditional density function corresponding to a density function $f_{Y|X}\left(Q_{Y|X}(\gamma)\right)$ and $\widehat{q}^{1}_{Y|X}(\gamma)$ the proposed estimator of $q_{Y|X}(\gamma)$, the conditional quantile density function. Then as n tends to infinity, we have*

$$\sup_{\gamma}\left|\widehat{q}^{1}_{Y|X}(\gamma) - q_{Y|X}(\gamma)\right| \longrightarrow 0 \qquad a.co.$$

Theorem 3 *Suppose that F is continuous. For $0 < \gamma < 1$, we have*

$$\sqrt{n\phi_x(h_K)}\left(\widehat{q}^{1}_{Y|X}(\gamma) - q_{Y|X}(\gamma)\right)$$

is asymptotically normal with mean zero and variance $\sigma^2(\gamma)$ where

$$\sigma^2(\gamma) = \frac{n\phi_x(h_K)}{h_H^4}\mathbb{E}\left(\int_0^1 dH^*(\gamma,v)\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)\right)^2.$$

1.8.1 Dependance structure

We assume the sample data $(X_i, Z_i)_{1\leq i\leq n}$ to be dependent and to satisfy the strong mixing condition (α-mixing), introduced by Rosenblatt (1956), defined as : let $\mathbb{N}^*$ denotes the set of positive integers, and for any i and j in $\mathbb{N}^*\cup\infty$, $(i \leq j)$, define $\mathcal{F}_i^j$ to be σ algebra spanned by the variables $(z_i, x_i)\cdots(z_j, x_j)$. The sequence (Z_i, X_i) is said to be α mixing if there exist mixing coefficients $\alpha(k)$ such that $|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq \alpha(k)$, for any sets A and B, that are, respectively, $\mathcal{F}_i^m$ measurable $\mathcal{F}_{m+k}^{\infty}$ measurable (k, m positive integers), and $\alpha(k)\downarrow 0$.

1.8.2 The functional kernel estimates

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

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

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

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

with $\mathbf{1}_{\{\cdot\}}$ being the indicator function and where $W_{ni}(z) = \frac{h_n^{-1} K(h_n^{-1} d(z,Z_i))}{\sum_{j=1}^{n} K(h_n^{-1} d(z,Z_j))}$, 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

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

For $z \in \mathcal{F}$, we denote by $h^z(\cdot)$ the conditional hazard function of X_1 given $Z_1 = z$. We assume that $h^z(\cdot)$ is unique maximum and its high risk point is denoted by $\theta(z) := \theta$, which is defined by

$$h^z(\theta(z)) := h^z(\theta) = \max_{x \in \mathcal{S}} h^z(x). \tag{1.15}$$

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

$$\widehat{h}^z(\widehat{\theta}(z)) := \widehat{h}^z(\widehat{\theta}) = \max_{x \in \mathcal{S}} \widehat{h}^z(x) \tag{1.16}$$

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

Note that the estimate $\widehat{\theta}$ is note necessarily unique. We point out that we can specify our choice by taking

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

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

$$\phi_z(h) = \mathbb{P}(Z \in B(z,h)),$$

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

In the following, z will be a fixed point in $\mathcal{F}$, $\mathcal{N}_z$ will denote a fixed neighborhood of z, $\mathcal{S}$ will be a fixed compact subset of $\mathbb{R}^+$.

1.9 Maximum of the conditional hazard function

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

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

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 $\mu(z)$, there exists $\xi > 0$ such that $|\theta(z) - \mu(z)| \geq \varepsilon$ implies that $|h^z(\theta(z)) - h^z(\mu(z))| \geq \xi$.

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 $\mathcal{S}^\circ$.

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

$$\widehat{h'}^z(x) = \frac{\widehat{f'}^z(x)}{1 - \widehat{F}^z(x)} + (\widehat{h}^z(x))^2, \tag{1.17}$$

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

$$\widehat{f'}^z(x) = \frac{\sum_{i=1}^{n} h_n^{-2} K(h_n^{-1} d(z, Z_i)) H'(h_n^{-1}(x - X_i))}{\sum_{i=1}^{n} K(h_n^{-1} d(z, Z_i))}. \tag{1.18}$$

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^z (but inherent problems of F^z, f^z and f'^z estimation), and secondly they consist with the assumptions usually made under functional variables.

Theorem 4 *We have*

$$\widehat{\theta} - \theta \to 0 \quad a.co. \tag{1.19}$$

Lemma 1 *We have*

$$\sup_{x\in\mathcal{S}} |\widehat{h}'^z(x) - h'^z(x)| \to 0 \quad a.co. \tag{1.20}$$

Theorem 5 *We have*

$$\sup_{x\in\mathcal{S}} |\widehat{\theta} - \theta| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{nh_n^3\phi_z(h_n)}}\right) \tag{1.21}$$

Lemma 2 *We have*

$$\sup_{x\in\mathcal{S}} |\widehat{h}'^z(x) - h'^z(x)| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{nh_n^3\phi_z(h_n)}}\right) \tag{1.22}$$

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

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

Theorem 6 *Under conditions, we have* $(\theta \in \mathcal{S}/f^z(\theta), 1 - F^z(\theta) > 0)$

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

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

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

and

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

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

Lemma 3 *Under Assumptions, we have*

$$(n\phi_z(h_n))^{1/2}\left(\widehat{F}^z(x)-F^z(x)\right)\xrightarrow{\mathcal{D}}N\left(0,\sigma^2_{F^z}(x)\right) \tag{1.23}$$

where

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

Lemma 4 *We have*

$$(nh_n\phi_z(h_n))^{1/2}\left(\widehat{h}^z(x)-h^z(x)\right)\xrightarrow{\mathcal{D}}N\left(0,\sigma^2_{h^z}(x)\right) \tag{1.24}$$

where

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

Lemma 5 *Under Assumptions, we have*

$$\left(nh_n^3\phi_z(h_n)\right)^{1/2}\left(\widehat{f'}^z(x)-f'^z(x)\right)\xrightarrow{\mathcal{D}}N\left(0,\sigma^2_{f'^z}(x)\right) \tag{1.25}$$

where

$$\sigma^2_{f'^z(x)}=\frac{a_2^xf^z(x)}{\left(a_1^x\right)^2}\int_{\mathbb{R}}(H'(t))^2dt$$

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

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

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

References

Abramovich, F., Angelini, C. (2006). Bayesian maximum a posteriori multiple testing procedure. Sankhya, 68 : 436-460.

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

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

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

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

Aneiros-Perez, G., Vieu, P. (2007). Semi-functional partial linear regression. Statist. Proba. Lett. **76**, 1102-1110.

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

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.

Azzalini, A. (1981) A note on the estimation of a distribution function and quantiles by kernel method. Biometrika, **60**, 326-328.

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

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

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

Benko, M., Hardle, W. and Kneip, A. (2005). Common functional principal components. SFB 649 Economic Risk Discussion Paper, 2006-010.

Beran, R. (1981). Nonparametric regression with randomly censored survival data. Technical Report, Dept. Statisticcs, Univ. California, Berkely.

Berlinet, A., Biau, G. et Rouviére, L. (2005).Parameter selection in modified histogram estimates. Statistics, **39**, 91-105.

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

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.

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

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

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

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

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.

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

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

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.

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

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

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

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

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

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.

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

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

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

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

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.

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

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

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

Delsol, L. (2007). Régression non paramétrique fonctionnelle : expression asymptotique des moments, Ann. I.S.U.P. Vol LI, **3**, 43-67.

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

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

Delsol, L., Ferraty, F., Vieu, P. (2011). Structural test in regression on functional variables. J. Multivariate Anal. **102**, No.3, 422-447.

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

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.

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

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.

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

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

Ferraty, F., Rabhi, A., Vieu, P. (2005). Conditional quantiles for functionally dependent data with application to the climatic El Phenomenon. Sankhya, **67**, 378-399.

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

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

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

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

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

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

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

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.

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

Ferré, L., Yao, A. F. (2005). Smoothed Functional Sliced Inverse Regression. Statist. Sinica **15**, 665-683.

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

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

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.

Gasser, T., Müller, H. G. (1979). Kernel estimation of regression function. In : T. Gasser and M. Rosenblatt (eds.) Smoothing techniques for curve estimation. Springer, Heidelberg, 23-68.

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

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

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

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.

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

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

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

HM00 Härdle, W., Müller, M. (2000). Multivariate and semiparametric regression. In : M. Schimek (Ed) Smoothing and regression ; Approaches, Computation, and Application. Wiley Series in Probability and Statistics, Wiley, New York, 357-392.

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

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

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

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

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

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

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

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

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.

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

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 (1961).

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

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.

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

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

Mas, A., Pumo, B. (2007). The ARHD model. J. Statist. Plann. Inference. **137**, 538-553.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

R81 Reiss, R.-D. (1981). Nonparametric estimation of smooth distribution functions. Scand. J. Statist. **8**, 116-119.

Rossi, L., Lubello, C., Poggiali, E., Griffini, O. (2002) . Analysis of a clariflocculation process with a photometric dispersion analyser (PDA2000). Water Supply. **2**, No.5-6, 57-63.

R69 Roussas, G. (1969). Nonparametric estimation of the transition distribution function of a Markov process. Annals of Math. Stat. **40**, 1386-1400.

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

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.

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

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.

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

Turnbull, B. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. J. R. S. Soc. Series : B **38**, 290-295.

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

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

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

Chapitre 2

Nonparametric Estimation of a Conditional Quantile Density Function for Time Series Data

Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.12785/jsap/040102

Nonparametric Estimation of a Conditional Quantile Density Function for Time Series Data

Amina Angelika Bouchentouf[1], *Yassine Hammou*[1], *Khadidja Nedjadi*[2] *and Abbes Rabhi*[1]

[1] Laboratory of Mathematics, University of Sidi Bel Abbes, PO. BOX. 89, Sidi Bel Abbes 22000, Algeria.
[2] Stochastic Models, Statistics and Applications Laboratory, Moulay Tahar University, Saida P.O.Box 138 En-Nasr Saida 20 000 Algeria.

Received: 14 Jul. 2014, Revised: 27 Nov. 2014, Accepted: 5 Dec. 2014
Published online: 1 Mar. 2015

Abstract: The aim of this paper is to estimate nonparametrically the conditional quantile density function. A non-parametric estimator of a conditional quantile function density is presented, its asymptotic properties are derived via the estimation of the conditional distribution, as of the conditional quantile in the case of dependent data. To obtain the asymptotic properties we consider some concentration hypotheses acting on the distribution of the conditional functional variable.

Keywords: Conditional quantile, Conditional quantile density function, Functional variable, Kernel density estimators, α-mixing

The problem of quantile estimation has a very long history, estimating quantiles of any distribution is an important part of Statistics. This allows to derive many applications in various fields as chemistry, geophysics, medicine, meteorology,.... On the other hand, functional random variables are becoming more and more important. The recent literature in this domain shows the great potential of these new functional statistical methods. The most popular case of functional random variable corresponds to the situation when we observe random curve on different statistical units. Such data are called Functional Data. Many multivariate statistical technics, mainly parametric in the functional model terminology, have been extended to functional data and good overviews on this topic can be found in Ramsay [21,22] and or Bosq [3].

More recently, nonparametric methods taking into account functional variables have been developed with very interesting practical motivations dealing with environmetrics (see Damon and Guillas [5], Fernandez et al. [7], Aneiros et al. [1]), chemometrics (see Ferraty and Vieu [9]), meteorological sciences (see Besse et al. [2], Hall and Heckman [15]), speech recognization problem (see Ferraty and Vieu [10]), radar range profile (see Hall et al. [16], Dabo-Niang et al. [4]), medical data (see Gasser et al. [14]), ...

Estimating the conditional quantile constitutes an important statistical topic. It is used to build predictive intervals, as a prediction method by the conditional median and to determine reference curves, predictive intervals etc. It has been widely studied, when the explanatory variable lies within a finite-dimension space (see, e.g., Gannoun et al. [13] and the references therein).

Jones [17] estimated the quantile density function by kernel means, via two alternative approaches. One is the derivative of the kernel quantile estimator, the other is essentially the reciprocal of the kernel density estimator, he gave ways in which the former method has certain advantages over the latter. In his paper, Jones discussed various closely related smoothing issues.

Soni et al. [20] defined a new nonparametric estimator of quantile density function and studied its asymptotic properties are studied. The comparison of the proposed estimator has been made with estimators given in [17].

The goal of this paper is to estimate nonparametrically the conditional quantile density function. A non-parametric estimator of a conditional quantile function density is presented, its asymptotic properties are derived via the estimation of the conditional distribution, as of the conditional quantile in dependent data. In a nonparametric context, it is known that the rate of convergence decreases with the dimension of the space in which the conditional variable is valued. But here, the conditional variable takes its values in an infinite dimensional space. So to override this problem is to consider

* Corresponding author e-mail: rabhi_abbes@yahoo.fr

some concentration hypotheses acting on the distribution of the conditional functional variable which allows to obtain the asymptotic properties.

1 The model

We consider a random pair (X,Y) where Y is valued in $\mathbb{R}$ and X is valued in some infinite dimensional semi-metric vector space $(\mathscr{F},d(.,.))$. Let (X_i,Y_i), $i=1,\ldots,n$ be the statistical sample of pairs which are identically distributed like (X,Y), but not necessarily independent. From now on, X is called functional random variable f.r.v. Let x be fixed in $\mathscr{F}$ and let $F_{Y|X}(\cdot,x)$ be the conditional cumulative distribution function (cond-cdf) of Y given $X=x$, namely:

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

Let $Q_{Y|X}(\gamma)$ be the $\gamma-$order quantile of the distribution of Y given $X=x$. From the cond-cdf $F_{Y|X}(.,x)$, it is easy to give the general definition of the γ-order quantile:

$$Q(\gamma|X=x)\equiv Q_{Y|X}(\gamma)=\inf\left\{t:F_{Y|X}(t,x)\geq\gamma\right\},\ \ 0\leq\gamma\leq 1 \tag{1}$$

Then, the definition of conditional quantile implies that

$$F_{Y|X}\left(Q_{Y|X}(\gamma)\right)=\gamma.$$

On differentiating partially w.r.t. γ we get

$$f_{Y|X}\left(Q_{Y|X}(\gamma)\right)=\frac{1}{\frac{\partial}{\partial\gamma}\left(Q_{Y|X}(\gamma)\right)}.$$

Parzen [19] and Jones [17] defined the quantile density function as the derivative of $Q(\gamma)$, that is, $q(\gamma)=Q'(\gamma)$. Note that the sum of two quantile density functions is again a quantile density function. Thus, the conditional quantile density function can be written as follows (see [26])

$$\begin{aligned} q(\gamma|X=x)\equiv q_{Y|X}(\gamma) &= \frac{\partial}{\partial\gamma}\left(Q_{Y|X}(\gamma)\right)\\ &= \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)} \end{aligned} \tag{2}$$

Let us now, define the kernel estimator $\widehat{F}_{Y|X}(.,x)$ of $F_{Y|X}(.,x)$

$$\widehat{F}_{Y|X}(x,y)=\frac{\sum_{i=1}^{n}K\left(h_K^{-1}d(x,X_i)\right)H\left(h_H^{-1}(y-Y_i)\right)}{\sum_{i=1}^{n}K\left(h_K^{-1}d(x,X_i)\right)} \tag{3}$$

where K is a kernel function, H a cumulative distribution function and $h_K=h_{K,n}$ (resp. $h_H=h_{H,n}$) a sequence of positive real numbers. Note that using similar ideas, Roussas [23] introduced some related estimate but in the special case when X is real, while Samanta [24] produced previous asymptotic study. As a by-product of (1) and (3), it is easy to derive an estimator $\widehat{Q}_{Y|X}$ of $Q_{Y|X}$:

$$\widehat{Q}_{Y|X}(\gamma)=\inf\left\{t:\widehat{F}_{Y|X}(t,x)\geq\gamma\right\}=\widehat{F}^{-1}_{Y|X}\left(Q_{Y|X}(\gamma)\right) \tag{4}$$

Let

$$\widehat{F}^{(j)}_{Y|X}(x,y)=\frac{h_H^{-j}\sum_{i=1}^{n}K\left(h_K^{-1}d(x,X_i)\right)H^{(j)}\left(h_H^{-1}(y-Y_i)\right)}{\sum_{i=1}^{n}K\left(h_K^{-1}d(x,X_i)\right)} \tag{5}$$

be the jth successive derivative of $\widehat{F}_{Y|X}(x,y)$, $f_{Y|X}(x,y)$ is conditional density function, such that $f_{Y|X}(x,y)=F^{(1)}_{Y|X}(x,y)$.

Nair and Sankaran [18] defined the hazard quantile function as follows:

$$H(\gamma) = h(Q(\gamma)) = \frac{f(Q(\gamma))}{S(Q(\gamma))} = ((1-\gamma)q(\gamma))^{-1}.$$

Thus hazard rate of two populations would be equal if and only if their corresponding quantile density functions are equal. This has been used to construct tests for testing equality of failure rates of two independent samples. Now, from this definition, let us introduce the γ-order conditional quantile of the conditional hazard function

$$H_{Y|X}(\gamma) = h_{Y|X}\left(Q_{Y|X}(\gamma)\right) = \frac{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}{S_{Y|X}\left(Q_{Y|X}(\gamma)\right)} = \left((1-\gamma)q_{Y|X}(\gamma)\right)^{-1}.$$

The smooth estimator of the conditional quantile density functional defined as follows:

$$\widehat{q}_{Y|X}(\gamma) = \frac{1}{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right)}. \tag{6}$$

where $\widehat{f}_{Y|X}(x,y)$ is a conditional kernel density estimator of $f_{Y|X}(x,y)$ and $\widehat{Q}_{Y|X}(\gamma)$ is the conditional empirical estimator of the conditional quantile function $Q_{Y|X}(\gamma)$. Let

$$\widehat{f}_{Y|X}(x,y) = \frac{h_H^{-1}\sum_{i=1}^{n} K\left(h_K^{-1}d(x,X_i)\right) H^{(1)}\left(h_H^{-1}(y-Y_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}d(x,X_i)\right)},$$

and

$$\widehat{f}^{(j)}_{Y|X}(x,y) = \frac{h_H^{-j-1}\sum_{i=1}^{n} K\left(h_K^{-1}d(x,X_i)\right) H^{(j+1)}\left(h_H^{-1}(y-Y_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1}d(x,X_i)\right)}$$

$h_K = h_{K,n}$ (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 $0/0 = 0$.

Let's now derive the asymptotic properties of our conditional quantile density function.

(H1) $\forall h > 0$, $\mathbb{P}(d(x,X) \le h) = \mathbb{P}(X \in B(x,h)) = \phi_x(h) > 0$, (with $B(x,h)$ the ball of center x and radius h)

(H2) $\sup_{i\neq j}\mathbb{P}((X_i,X_j) \in B(x,h)\times B(x,h)) = \mathbb{P}(W_i \le h, W_j \le h) \le \psi_x(h)$, where $\psi_x(h) \to 0$ as $h \to 0$. Furthermore, we assume that $\psi_x(h) = O(\phi_x^2(h))$.

(H3) H is such that, for all $(y_1,y_2) \in \mathbb{R}^2$, $|H(y_1) - H(y_2)| \le C|y_1 - y_2|$ and its first derivative $H^{(1)}$ verifies $\int |t|^{b_2} H^{(1)}(t)dt < \infty$,

(H4) K is a function with support $(0,1)$ such that $0 < C_1 < K(t) < C_2 < \infty$,

(H5) $\lim_{n\to\infty} h_K = 0$ with $\lim_{n\to\infty} \frac{\log n}{n\phi_x(h_K)} = 0$

(H6) $\exists j > 0$, $\forall l$, $1 \le l < j$, $f^{(l)}_{Y|X}\left(Q_{Y|X}(\gamma)\right) = 0$ and $\left|f^{(j)}_{Y|X}\left(Q_{Y|X}(\gamma)\right)\right| > 0$

2 Main result

2.1 Estimation of conditional quantile density function

Theorem 1. *Let $q_{Y|X}(\gamma)$ be the conditional density function corresponding to a density function $f_{Y|X}\left(Q_{Y|X}(\gamma)\right)$ and $\widehat{q}_{Y|X}(\gamma)$ denote the estimator of $q_{Y|X}(\gamma)$. Then under the assumptions (H1)-(H6) and as n tends to infinity, we have*

$$\sup_{\gamma}\left|\widehat{q}_{Y|X}(\gamma) - q_{Y|X}(\gamma)\right| \longrightarrow 0 \qquad a.co.$$

Proof. At first let us consider

$$\begin{aligned}
\widehat{q}_{Y|X}(\gamma) &= \frac{1}{\widehat{f}_{Y|X}(\widehat{Q}_{Y|X}(\gamma))} \\
&= \frac{1}{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right) - f_{Y|X}\left(Q_{Y|X}(\gamma)\right) + f_{Y|X}\left(Q_{Y|X}(\gamma)\right)} \\
&= \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}\left(\frac{1}{1+\frac{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right)-f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}}\right)
\end{aligned}$$

Then, we get

$$\begin{aligned}
\widehat{q}_{Y|X}(\gamma) &= \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}\left(1-\frac{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right) - f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}\right) \\
&+ \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}\left(\frac{\left(\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right) - f_{Y|X}\left(Q_{Y|X}(\gamma)\right)\right)^2}{f^2_{Y|X}\left(Q_{Y|X}(\gamma)\right)}\right) \\
&- \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}\left(\frac{\left(\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right) - f_{Y|X}\left(Q_{Y|X}(\gamma)\right)\right)^3}{f^3_{Y|X}\left(Q_{Y|X}(\gamma)\right)} + \dots\right),
\end{aligned}$$

hence

$$\begin{aligned}
\widehat{q}_{Y|X}(\gamma) - q_{Y|X}(\gamma) &= \frac{-\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right) + f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}{f^2_{Y|X}\left(Q_{Y|X}(\gamma)\right)} \\
&+ \frac{\left(\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right) - f_{Y|X}\left(Q_{Y|X}(\gamma)\right)\right)^2}{f^3_{Y|X}\left(Q_{Y|X}(\gamma)\right)} \\
&- \frac{\left(\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right) - f_{Y|X}\left(Q_{Y|X}(\gamma)\right)\right)^3}{f^4_{Y|X}\left(Q_{Y|X}(\gamma)\right)} + \dots
\end{aligned}$$

With

$$\begin{aligned}
\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right) = {} & \widehat{f}_{Y|X}\left(Q_{Y|X}(\gamma)\right) + \left(\widehat{Q}_{Y|X}(\gamma) - Q_{Y|X}(\gamma)\right)\widehat{f'}_{Y|X}\left(Q_{Y|X}(\gamma)\right) + \\
& \frac{\left(\widehat{Q}_{Y|X}(\gamma) - Q_{Y|X}(\gamma)\right)^2 \widehat{f}''_{Y|X}(Q_{Y|X}(\gamma))}{2!} + \\
& \frac{\left(\widehat{Q}_{Y|X}(\gamma) - Q_{Y|X}(\gamma)\right)^3 \widehat{f^{(3)}}_{Y|X}(Q_{Y|X}(\gamma))}{3!} + \dots
\end{aligned}$$

Therefore

$$\begin{aligned}
\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right) - f_{Y|X}\left(Q_{Y|X}(\gamma)\right) = {} & \widehat{f}_{Y|X}\left(Q_{Y|X}(\gamma)\right) - f_{Y|X}\left(Q_{Y|X}(\gamma)\right) + \\
& \left(\widehat{Q}_{Y|X}(\gamma) - Q_{Y|X}(\gamma)\right)\widehat{f}'_{Y|X}\left(Q_{Y|X}(\gamma)\right) + \\
& \frac{\left(\widehat{Q}_{Y|X}(\gamma) - Q_{Y|X}(\gamma)\right)^2 \widehat{f}''_{Y|X}\left(Q_{Y|X}(\gamma)\right)}{2!} + \\
& \frac{\left(\widehat{Q}_{Y|X}(\gamma) - Q_{Y|X}(\gamma)\right)^3 \widehat{f}^{(3)}_{Y|X}\left(Q_{Y|X}(\gamma)\right)}{3!} + \dots
\end{aligned} \tag{7}$$

Now, it rests to show the following convergence

$$\sup_{\gamma}\left|\widehat{Q}_{Y|X}(\gamma)-Q_{Y|X}(\gamma)\right|\underset{n\to\infty}{\longrightarrow}0 \qquad a.co.$$

$$\sup_{\gamma}|\widehat{f}_{Y|X}\left(Q_{Y|X}(\gamma)\right)-f_{Y|X}\left(Q_{Y|X}(\gamma)\right)|\underset{n\to\infty}{\longrightarrow}0 \qquad a.co.$$

It was shown in [8] the following results

$$\sup_{\gamma}|\widehat{Q}_{Y|X}(\gamma)-Q_{Y|X}(\gamma)|\underset{n\to\infty}{\longrightarrow}0 \; a.co.$$

$$\widehat{Q}_{Y|X}(\gamma)-Q_{Y|X}(\gamma)=O\left(h_K^{\frac{b_1}{j}}+h_H^{\frac{b_2}{j}}\right)+O_{a.co.}\left(\frac{\log n}{n\phi_x(h_K)}\right)^{\frac{1}{2j}},$$

and

$$|\widehat{F}^{(j)}_{Y|X}(x,y)-F^{(j)}_{Y|X}(x,y)|=O\left(h_K^{b_1}+h_H^{b_2}\right)+O_{a.co.}\left(\sqrt{\frac{\log n}{nh_H^{2j-1}\phi_x(h_K)}}\right) \tag{8}$$

Note that, $\widehat{f}_{Y|X}=\widehat{F}^{(1)}_{Y|X}$, so applying (8) for $j=1$, we get

$$|\widehat{f}_{Y|X}\left(Q_{Y|X}(\gamma)\right)-f_{Y|X}\left(Q_{Y|X}(\gamma)\right)|=O\left(h_K^{b_1}+h_H^{b_2}\right)+O_{a.co.}\left(\sqrt{\frac{\log n}{nh_H\phi_x(h_K)}}\right), \tag{9}$$

Based on data $X_1,X_2,\dots,X_n$, we propose a smooth estimator of the conditional quantile density function, Ferraty *e*t al. [8] proposed a kernel-type estimator of conditional quantile which is is a conditional version of Parzen's estimator in the univariate case (see Parzen [19]).

For an appropriate kernel function H' and a bandwidth sequence h_H. We suggest an estimator of $q_{Y|X}(\gamma)$;

$$\widehat{q}^{1}_{Y|X}(\gamma)=\frac{1}{h_H}\int_0^1\frac{H'\left(h_H^{-1}(v-\gamma)\right)}{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)}dv \tag{10}$$

The next theorem proves consistency of the proposed estimator of the conditional quantile density function.

Theorem 2.*Let $q_{Y|X}(\gamma)$ be the conditional density function corresponding to a density function $f_{Y|X}\left(Q_{Y|X}(\gamma)\right)$ and $\widehat{q}^{1}_{Y|X}(\gamma)$ given by (10) the proposed estimator of $q_{Y|X}(\gamma)$, the conditional quantile density function. Then under hypotheses (H1)-(H6) as n tends to infinity, we have*

$$\sup_{\gamma}\left|\widehat{q}^{1}_{Y|X}(\gamma)-q_{Y|X}(\gamma)\right|\longrightarrow 0 \qquad a.co.$$

Proof.

(10) gives the estimator of the conditional quantile density function $q(\gamma)$ as

$$\widehat{q}^{1}_{Y|X}(\gamma)=\frac{1}{h_H}\int_0^1\frac{H'\left(h_H^{-1}(v-\gamma)\right)}{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)}dv.$$

Hence

$$\begin{aligned}
\widehat{q}^{1}_{Y|X}(\gamma)-q_{Y|X}(\gamma) &= \frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)}dv - q_{Y|X}(\gamma)\\
&= \frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)}dv - \frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(v)\right)}dv\\
&+ \frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(v)\right)}dv - \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}\\
&= -\frac{1}{h_H}\int_0^1 H'\left(h_H^{-1}(v-\gamma)\right)\left[\frac{1}{f_{Y|X}\left(Q_{Y|X}(v)\right)} - \frac{1}{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)}\right]dv\\
&+ \frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(v)\right)}dv - \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}\\
&= -\frac{1}{h_H}\int_0^1 H'\left(h_H^{-1}(v-\gamma)\right)\left[\frac{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right) - f_{Y|X}\left(Q_{Y|X}(v)\right)}{\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right) f_{Y|X}\left(Q_{Y|X}(v)\right)}\right]dv\\
&+ \frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(v)\right)}dv - \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}.
\end{aligned}$$

Using Theorem 1, $\sup_{\gamma}\left|\widehat{q}_{Y|X}(\gamma)-q_{Y|X}(\gamma)\right|\underset{n\to\infty}{\longrightarrow}0$ a.co. Hence the above expression asymptotically reduces to

$$\begin{aligned}
&-\frac{1}{h_H}\int_0^1 H'\left(h_H^{-1}(v-\gamma)\right)\left(q_{Y|X}(v)\right)^2\left[\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right) - f_{Y|X}\left(Q_{Y|X}(v)\right)\right]dv +\\
&\qquad \frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(v)\right)}dv - \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}\\
&= -\frac{1}{h_H}\int_0^1 H'\left(h_H^{-1}(v-\gamma)\right)q_{Y|X}(v)\left[\widehat{f}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)\widehat{q}_{Y|X}(v) - f_{Y|X}\left(Q_{Y|X}(v)\right)q_{Y|X}(v)\right]dv +\\
&\qquad \frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(v)\right)}dv - \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}.
\end{aligned}$$

Since $dF_{Y|X}(\left(Q_{Y|X}(v)\right) = f(\left(Q_{Y|X}(v)\right)q_{Y|X}(v)dv$, hence

$$\begin{aligned}
\widehat{q}^{1}_{Y|X}(\gamma)-q_{Y|X}(\gamma) &= -\frac{1}{h_H}\int_0^1 H'\left(h_H^{-1}(v-\gamma)\right)q_{Y|X}(v)\left[d\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right) - dF_{Y|X}\left(Q_{Y|X}(v)\right)\right]dv\\
&+\frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(v)\right)}dv - \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}.
\end{aligned}$$

Writing $H^*(\gamma,v) = H'\left(h_H^{-1}(v,\gamma)\right)q_{Y|X}(v)$ and integrating by parts in the first integral, we get

$$\begin{aligned}
\widehat{q}^{1}_{Y|X}(\gamma)-q_{Y|X}(\gamma) &= \left[-\frac{1}{h_H}\left(H^*(v,\gamma)\right)\left(\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right) - F_{Y|X}\left(Q_{Y|X}(v)\right)\right)\right]_0^1\\
&+\frac{1}{h_H}\int_0^1 dH^*(\gamma,v)\left[\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right) - F_{Y|X}\left(Q_{Y|X}(v)\right)\right]dt\\
&+\frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(v)\right)}dv - \frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}.
\end{aligned}$$

Since $F_{Y|X}\left(Q_{Y|X}(0)\right)=\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(0)\right)$ and $F_{Y|X}\left(Q_{Y|X}(1)\right)=\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(1)\right)$, the above expression transforms to

$$\begin{aligned}\widehat{q}^{1}_{Y|X}(\gamma)-q_{Y|X}(\gamma) = &-\frac{1}{h_H}\int_0^1 dH^*(\gamma,v)\left[F_{Y|X}\left(Q_{Y|X}(v)\right)-\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)\right]dt\\ &+\frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(v)\right)}dv-\frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}.\end{aligned}$$

Putting $h_H^{-1}(v-\gamma)=z$ and using (2),

$$\frac{1}{h_H}\int_0^1 \frac{H'\left(h_H^{-1}(v-\gamma)\right)}{f_{Y|X}\left(Q_{Y|X}(v)\right)}dv-\frac{1}{f_{Y|X}\left(Q_{Y|X}(\gamma)\right)}=\frac{1}{h_H}\int_{-\gamma/h_H}^{1-\gamma/h_H} H'(z)q_{Y|X}(\gamma+zh_H)\,dz-q_{Y|X}(\gamma). \tag{11}$$

Using Taylor series expansion, we can write

$$q_{Y|X}(\gamma+zh_H)-q_{Y|X}(\gamma)=\sum_{l=1}^{j-1}\frac{(zh_H)^l}{l!}q^{(l)}_{Y|X}(\gamma)+\frac{(zh_H)^j}{j!}q^{(j)}_{Y|X}(\gamma^*)$$

where $\gamma<\gamma^*<\gamma+zh_H$, assuming higher derivatives of $q_{Y|X}(\gamma)$ exist and are bounded.

Hence (11) can be written as

$$\int_{-\gamma/h_H}^{1-\gamma/h_H} H'(z)\left(q_{Y|X}(\gamma)+\sum_{l=1}^{j-1}\frac{(zh_H)^l}{l!}q^{(l)}_{Y|X}(\gamma)+\frac{(zh_H)^j}{j!}q^{(j)}_{Y|X}(\gamma^*)\right)dv-q_{Y|X}(\gamma). \tag{12}$$

For $n\longrightarrow\infty$, $h_K\longrightarrow 0$, (12) converges to $\int_{\mathbb{R}} H'(z)q_{Y|X}(\gamma)dz-q_{Y|X}(\gamma)$ which equals zero as $\int_{\mathbb{R}} H'(z)=1$.

This gives

$$\widehat{q}^{1}_{Y|X}(\gamma)-q_{Y|X}(\gamma)=\frac{1}{h_H^2}\int_0^1 dH^*(\gamma,v)\left[\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)-F_{Y|X}\left(Q_{Y|X}(v)\right)\right]dv. \tag{13}$$

Since $\sup_v\left|\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)-F_{Y|X}\left(Q_{Y|X}(v)\right)\right|\underset{n\to\infty}{\longrightarrow}0$, hence $\sup_v\left|\widehat{q}^{1}_{Y|X}(v)-q_{Y|X}(v)\right|\underset{n\to\infty}{\longrightarrow}0$.

The following theorem proves asymptotic normality of the proposed estimator.

2.2 Asymptotic normality

In this section we give the asymptotic normality of $\widehat{q}^{1}_{Y|X}(\gamma)$.

Theorem 3. *Suppose that F is continuous. Assume that $K(\cdot)$ satisfies the conditions $(H1)-(H6)$ given in section 2. For $0<\gamma<1$, we have*

$$\sqrt{n\phi_x(h_K)}\left(\widehat{q}^{1}_{Y|X}(\gamma)-q_{Y|X}(\gamma)\right)$$

is asymptotically normal with mean zero and variance $\sigma^2(\gamma)$ where

$$\sigma^2(\gamma)=\frac{n\phi_x(h_K)}{h_H^2}\mathbb{E}\left(\int_0^1 dH^*(\gamma,v)\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)\right)^2.$$

Proof.

Using (13), we have

$$\sqrt{n\phi_x(h_K)}\left(\widehat{q}^{1}_{Y|X}(\gamma)-q_{Y|X}(\gamma)\right)=\frac{\sqrt{n\phi_x(h_K)}}{h_H^2}\int_0^1 dH^*(\gamma,v)\left[\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)-F_{Y|X}\left(Q_{Y|X}(v)\right)\right]dv.$$

Using the results of Ezzahrioui and Ould Saïd [6], for $0<\gamma<1$,

$(n\phi_x(h_K))^{1/2}\left(\widehat{Q}_{Y|X}(\gamma)-Q_{Y|X}(\gamma)\right)$ is asymptotically normal with mean zero and variance $\xi^2\left(x,Q_{Y|X}(\gamma)\right)=\frac{a_2^x}{(a_1^x)^2}\frac{F_{Y|X}\left(Q_{Y|X}(\gamma)\right)\left(1-F_{Y|X}\left(Q_{Y|X}(\gamma)\right)\right)}{f_{Y|X}^2\left(Q_{Y|X}(\gamma)\right)}$.

$(n\phi_x(h_K))^{1/2}\left(\widehat{F}_{Y|X}(x,y)-F_{Y|X}(x,y)\right)$ is asymptotically normal with mean zero and variance $\sigma_F^2=\frac{a_1^x}{(a_1^x)^2}F_{Y|X}(x,y)\left(1-F_{Y|X}(x,y)\right)$.

We have also, $(n\phi_x(h_K)))^{1/2}\left(\widehat{F}_{Y|X}(\widehat{Q}_{Y/X}(t))-F_{Y|X}(Q_{Y/X}(t))\right)$ is asymptotically normal with mean zero and variance $\sigma^2(x)$.

With

$$\sigma^2(x)=\frac{\gamma(1-\gamma)a_2^x(x)}{(f_{y/x}(Q_{Y|X}(\gamma)))^2a_1^x(x)}$$

$$a_j^x(x)=K^j(1)-\int_0^1(K^j)'(s)\beta_x(s)ds \quad j=1,2.$$

and

$$\forall s\in[0,1],\ \lim_{h\to 0}\phi_x(sh)/\phi_x(h)=\beta_x(s).$$

Since $\frac{d}{d\gamma}F_{Y|X}\left(Q_{Y|X}(\gamma)\right)=1$, $(n\phi_x(h_K))^{1/2}\left[\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(\gamma)\right)-F_{Y|X}\left(Q_{Y|X}(\gamma)\right)\right]$ is asymptotically normal with mean zero and variance $\xi^2\left(x,Q_{Y|X}(\gamma)\right)$.

Using Delta method and Slutsky's theorem (Serfling [25]), we get that $\sqrt{n\phi_x(h_K)}\left(\widehat{q}^1_{Y|X}(\gamma)-q_{Y|X}(\gamma)\right)$ is asymptotically normal with mean zero and variance $\sigma^2(\gamma)=\frac{n\phi_x(h_K)}{h_H^2}\mathbb{E}\left(\int_0^1 dH^*(\gamma,v)\widehat{F}_{Y|X}\left(\widehat{Q}_{Y|X}(v)\right)\right)^2$.

The expression of $\sigma^2(\gamma)$ in the above theorem cannot be simplified analytically and one can estimate it using bootstrapping.

Acknowledgement

The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

[1] G. Aneiros Pérez, H. Cardot, G. Estévez Pérez and P. Vieu, Maximum ozone forecasting by functional nonparametric approach. Environmetrics **15**, 675-685 (2004).
[2] P. Besse, H. Cardot, and D. Stephenson, Autoregressive forecasting of some functional climatic variations. Scand. J. Statist, **27**, 673-687 (2000).
[3] D. Bosq, Linear Processes in Function Spaces. Lecture Notes in Statistics **149** , Springer-Verlag, Berlin, (2000).
[4] S. Dabo-Niang, F. Ferraty and P. Vieu, Nonparametric unsupervised classification of satellite wave altimeter forms. In Computational Statistics, J. Antoch, ed., Physica Verlag, Berlin, 879-887 (2004).
[5] J. Damon and S. Guillas, The inclusion of exogenous variables in functional autoregressive ozone forecasting. Environmetrics **13**, 759-774 (2002).
[6] M. Ezzahrioui and E. Ould-Saïd, Asymptotic results of a nonparametric conditional quantile estimator for functional time series. Comm. Statist. Theory Methods. **37**(16-17), 2735-2759 (2008).
[7] B. Fernández de Castro, S. Guillas and W. González Manteiga, Functional samples and bootstrap for predicting sulfur dioxide levels, Technometrics, **47**(2), 212-222 (2005).
[8] F. Ferraty, A. Rabhi and P. Vieu, 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**(2), 378-398 (2005).
[9] F. Ferraty and P. Vieu, The functional nonparametric model and application to spectrometric data. Computat. Statist. **17**, 545-564 (2002).
[10] F. Ferraty and P. Vieu, Curves discrimination: a nonparametric functional approach. Comput. Statist. Data Anal. **44**, 161-173 (2003a).

[11] F. Ferraty and P. Vieu, Functional nonparametric statistics: a double infinite dimensional framework. In Recent Advances and Trends in Nonparametric Statistics; M. Akritas and D. Politis, eds. Elsevier, Amsterdam (2003b).
[12] F. Ferraty and P. Vieu, Nonparametric methods for functional data, with applications in regression, time-series prediction and curves discrimination. J. Non- parametr. Statist. **16**, 11-126 (2004).
[13] A. Gannoun, J. Saracco and K. Yu, Nonparametric prediction by condi- tional median and quantiles. J. Statist. Plann. Inference. **117**, 207-223 (2003).
[14] T. Gasser, P. Hall and B. Presnell, Nonparametric estimation of the mode of a distribution of random curves. J. Roy. Statist. Soc. Ser. B. **60**, 681-691 (1998).
[15] P. Hall and N. Heckman, Estimating and depicting the structure of the distribution of random functions, Biometrika, **89**, 145-158 (2002).
[16] P. Hall, P. Poskitt and D. Presnell, A functional data-analytic approach to signal discrimination. Technometrics **35**, 140-143 (2001).
[17] M. C. Jones, Estimating densities, quantiles, quantile densities and density quantiles. Ann. Inst. Statist. Math. **44**(4), 721-727 (1992).
[18] N. U. Nair and P. G. Sankaran, Quantile based reliability analysis. Communications in Statistics Theory and Methods **38**, 222-232 (2009).
[19] E. Parzen, Nonparametric statistical data modelling. Journal of American Statistical Association **74**, 105-122 (1979).
[20] S. Pooja, D. Isha and J. Kanchan, Nonparametric estimation of quantile density function. Computational Statistics and Data analysis **56**, 3876-3886 (2012).
[21] J. O. Ramsay and B. W. Silverman, Functional Data Analysis. Conditional quantiles for functionally dependent data **21**; Springer, Berlin, 1997.
[22] J. O. Ramsay and B. W. Silverman, Applied Functional Data Analysis. Springer-Verlag, New York, (2002).
[23] G. Roussas, Nonparametric estimation of the transition distribution function of a Markov process. Ann. Math. Statist. **40**, 1386-1400 (1969).
[24] M. Samanta, Nonparametric estimation of conditional quantiles. Statist. Prob and Lett. **7**, 407-412 (1989).
[25] R. J. Serfling, Approximation Theorem of Mathematical Statistics. John Wiley, New York, (1980).
[26] X. Xiang, Estimation of conditional quantile density function. Journal of Nonparametric Statistics **4**, 309-316 (1995).

Chapitre 3

Nonparametric estimation of the maximum of conditional hazard function under dependence conditions for functional data

Afrika Statistika

ISSN 2316-090X

Afrika Statistika
Vol. 10(1), 2015, pages 777–794.
DOI: http://dx.doi.org/1016929.as.2015.777.69

Nonparametric estimation of the maximum of conditional hazard function under dependence conditions for functional data

Abbes Rabhi[‡,1], **Yassine Hammou**[‡] **and Tayeb Djebbouri**[†]

[‡]Mathematics Laboratory, University of Sidi Bel Abbes P.O.Box. 89, Sidi Be1l Abbes 22000, Algeria

[†]Stochastic Models, Statistics and Applications Laboratory, Moulay Tahar University, Saida P.O.Box 138 En-Nasr Saida 20 000 Algeria

Received January 31, 2015; Accepted October 25, 2015

Abstract. 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 a strong mixing process.

Résumé. 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 basés 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 mélangeant.

Key words: Almost complete convergence; Asymptotic normality; Conditional hazard function; Functional data; Nonparametric estimation; Small ball probability; Strong mixing processes.
AMS 2010 Mathematics Subject Classification : Primary: 62F12; Secondary: 62G20; 62M09.

[1]Corresponding author Abbes Rabhi: rabhi_abbes@yahoo.fr
Yassine Hammou: hammouy@yahoo.fr; Tayeb Djebbouri: tdjebbouri20@gmail.com

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 popularised 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.

From a theoretical point of view, a sample of functional data can be involved in many different statistical problems, such as: classification and principal components analysis (PCA) (1986,1991) or longitudinal studies, regression and prediction (Benhenni *et al.*, 2007; Cardot *et al.*, 1999). The recent monograph by Ferraty *et al.* (2007) summarizes many of their contributions to the non-parametric estimation with functional data; among other properties, consistency of the conditional density, conditional distribution and regression estimates are established in the i.i.d. case under dependence conditions (strong mixing). Almost complete rates of convergence are also obtained, and different techniques are applied to several examples of functional data samples. Related work can be seen in the paper of Masry (2005), where the asymptotic normality of the functional nonparametric regression estimate is proven, considering strong mixing dependence conditions for the sample data. For automatic smoothing parameter selection in the regression setting, see Rachdi and Vieu (2007).

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 (1964a), there are many papers on these topics: Ahmad (1976), Singpurwalla and Wong (1983), etc.We can cite Quintela-del-Rio, A. (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 (1964b), Roussas (1989), Lecoutre and Ould-Saïd (1992), Estévez-Pérez *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, A. (2008). We refer to Ferraty *et al.* (2010) and Mahiddine *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 non-parametric 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.* (2007) 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 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 $Z = z$ and the main issue is to estimate the conditional random rate defined as

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

which can be written naturally as follows:

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

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 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 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 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 5 includes some proofs of technical Lemmas.

2. Nonparametric estimation with dependent functional data

Let $\{(Z_i, X_i),\ i = 1, \ldots, n\}$ be a sample of n random pairs, each one distributed as (Z, X), where the variable Z is of functional nature and X is scalar. Formally, we will consider that Z 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 $Z = z$ is defined for any $x \in \mathbb{R}$ and any $z \in \mathcal{F}$ by

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

while the conditional density, denoted by $f^z(x)$ 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 (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.

2.1. Dependance structure

We assume the sample data $(X_i, Z_i)_{1\leq i\leq n}$ to be dependent and to satisfy the strong mixing condition (α-mixing), introduced by Rosenblatt (1956), defined as:

let $\mathbb{N}^*$ denotes the set of positive integers, and for any i and j in $\mathbb{N}^* \cup \infty$, $(i \leq j)$, define $\mathcal{F}_i^j$ to be σ algebra spanned by the variables $(z_i, x_i) \cdots (z_j, x_j)$. The sequence (Z_i, X_i) is said to be α mixing if there exist mixing coefficients $\alpha(k)$ such that $|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq \alpha(k)$, for any sets A and B, that are, respectively, $\mathcal{F}_i^m$ measurable $\mathcal{F}_{m+k}^\infty$ measurable (k, m positive integers), and $\alpha(k) \downarrow 0$.

This is the weakest condition used in studies of dependent samples (for example, the ARMA process, generated by a continuous white noise verifies it). The reader can see Doukhan (1994) for a more complete discussion of the strong mixing condition.

[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 $T_n = \mathcal{O}_{a.co.}(u_n)$).

2.2. The functional kernel estimates

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

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

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

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

with $\mathbf{1}_{\{\cdot\}}$ being the indicator function and where $W_{ni}(z) = \frac{h_n^{-1} K(h_n^{-1} d(z, Z_i))}{\sum_{j=1}^n K(h_n^{-1} d(z, Z_j))}$, 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

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

For $z \in \mathcal{F}$, we denote by $h^z(\cdot)$ the conditional hazard function of X_1 given $Z_1 = z$. We assume that $h^z(\cdot)$ is unique maximum and its high risk point is denoted by $\theta(z) := \theta$, which is defined by

$$h^z(\theta(z)) := h^z(\theta) = \max_{x \in \mathcal{S}} h^z(x). \tag{2}$$

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

$$\widehat{h}^z(\widehat{\theta}(z)) := \widehat{h}^z(\widehat{\theta}) = \max_{x \in \mathcal{S}} \widehat{h}^z(x) \tag{3}$$

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

Note that the estimate $\widehat{\theta}$ is note necessarily unique and our results are valid for any choice satisfying (3). We point out that we can specify our choice by taking

$$\widehat{\theta}(z) = \inf \left\{ t \in \mathcal{S} \; ; \; \widehat{h}^z(t) = \max_{x \in \mathcal{S}} \widehat{h}^z(x) \right\}.$$

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

$$\phi_z(h) = \mathbb{P}(Z \in B(z, h)),$$

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

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

(H0) $\forall h > 0,\ 0 < \mathbb{P}(Z \in B(z,h)) = \phi_z(h)$ and $\lim_{h\to 0}\phi_z(h) = 0$

(H1) $(Z_i, X_i)_{i\in\mathbb{N}}$ is an α-mixing sequence whose the coefficients of mixture verify:

$$\exists a > 0,\ \exists c > 0 :\ \forall n \in \mathbb{N},\ \alpha(n) \le cn^{-a}.$$

(H2) $0 < \max_{i\neq j}\psi_{i,j}(h) = \sup_{i\neq j}\mathbb{P}((Z_i, Z_j) \in B(z,h)\times B(z,h)) = \mathcal{O}\left(\frac{(\phi_x(h))^{(a+1)/a}}{n^{1/a}}\right)$.

Note that (H0) can be interpreted as a concentration hypothesis acting on the distribution of the *f.r.v.* of Z, whereas (H2) concerns the behavior of the joint distribution of the pairs (Z_i, Z_j). In fact, this hypothesis is equivalent to assume that, for n large enough

$$\sup_{i\neq j}\frac{\mathbb{P}((Z_i, Z_j) \in B(z,h)\times B(z,h))}{\mathbb{P}(Z \in B(z,h))} \le C\left(\frac{\phi_x(h)}{n}\right)^{1/a}.$$

This is one way to control the local asymptotic ratio between the joint distribution and its margin. Remark that the upper bound increases with a. In other words, more the dependence is strong, more restrictive is (H2). The hypothesis (H1) specifies the asymptotic behavior of the α-mixing coefficients.

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

$$\text{(H3)}\begin{cases}\forall (x_1,x_2)\in\mathcal{S}^2, \forall(z_1,z_2)\in\mathcal{N}_z^2,\ \text{for some } b_1>0,\ b_2>0\\ |F^{z_1}(x_1)-F^{z_2}(x_2)| \le C_z(d(z_1,z_2)^{b_1}+|x_1-x_2|^{b_2}),\end{cases}$$

$$\text{(H4)}\begin{cases}\forall (x_1,x_2)\in\mathcal{S}^2, \forall(z_1,z_2)\in\mathcal{N}_z^2,\ \text{for some } j=0,1,\ \nu>0,\ \beta>0\\ |f^{z_1\,(j)}(x_1)-f^{z_2\,(j)}(x_2)| \le C_z(d(z_1,z_2)^{\nu}+|x_1-x_2|^{\beta}),\end{cases}$$

(H5) $\exists\gamma < \infty, f'^z(x) \le \gamma,\ \forall(z,x)\in\mathcal{F}\times\mathcal{S}$,

(H6) $\exists\tau > 0, F^z(x) \le 1-\tau,\ \forall(z,x)\in\mathcal{F}\times\mathcal{S}$.

(H7) H is differentiable such that

$$\begin{cases}\text{(H7a) } \forall(t_1,t_2)\in\mathbb{R}^2;\ |H^{(j)}(t_1)-H^{(j)}(t_2)| \le C|t_1-t_2|,\ \text{ for } j=0,1\\ \text{and } H^{(j)}\text{are bounded for } j=0,1\\ \text{(H7 b) } \int_{\mathbb{R}} t^2H^2(t)dt<\infty,\\ \text{(H7c) } \int_{\mathbb{R}}|t|^{\beta}(H'^2dt<\infty\end{cases}$$

(H8) The kernel K is positive bounded function supported on $[0,1]$ and it is of class $\mathcal{C}^1$ on $(0,1)$ such that $\exists C_1,\ C_2,\ -\infty < C_1 < K'(t) < C_2 < 0$ for $0 < t < 1$

(H9) There exists a function $\zeta_0^z(\cdot)$ such that for all $t\in[0,1]$ $\lim_{h\to 0}\frac{\phi_z(th)}{\phi_z(h)} = \zeta_0^z(t)$.

(H10) The bandwidth h_n, small ball probability $\phi_z(h_n)$ and arithmetical α mixing coefficient with order $a > 3$ satisfying

$$\begin{cases}\text{(H10a)}\exists C>0,\ h_n^{2j+1}\phi_z(h_n) \ge \frac{C}{n^{2/(a+1)}},\ \text{ for } j=0,1\\ \text{(H1 0b) }\left(\frac{\phi_z(h_n)}{n}\right)^{1/a}+\phi_z(h_n) = o\left(\frac{1}{n^{2/(a+1)}}\right),\ \text{ for } j=0,1\\ \text{(H10c) }\lim_{n\to\infty}h_n=0,\quad\text{and}\quad \lim_{n\to\infty}\frac{\log n}{nh_n^{2j+1}\phi_x(h_n)}=0,\ j=0,1;\end{cases}$$

Remark 1. Assumption (H0) plays an important role in our methodology. It is known as (for small h) the "concentration hypothesis acting on the distribution of X" in infi- nite-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$\phi_z(h)$ can be written approximately as the product of two independent functions $\psi(z)$ and $\varphi(h)$ as $\phi_z(h) = \psi(z)\varphi(h) + o(\varphi(h))$. This idea was adopted by Masry (2005) who reformulated the Gasser *et al.* (1998) one. The increasing proprety of $\phi_z(\cdot)$ implies that $\zeta_h^z(\cdot)$ is bounded and then integrable (all the more so $\zeta_0^z(\cdot)$ is integrable).

Without the differentiability of $\phi_z(\cdot)$, 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 $\phi_z(h) = C(d)h^d\psi(z) + oh^d)$, 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. $\phi_z(h) \approx \psi(h)h^\gamma$ for som $\gamma > 0$.
2. $\phi_z(h) \approx \psi(h)h^\gamma \exp\{C/h^p\}$ for som $\gamma > 0$ and $p > 0$.
3. $\phi_z(h) \approx \psi(h)/|\ln h|$.

The function $\zeta_h^z(\cdot)$ which intervenes in Assumption (H9) is increasing for all fixed h. Its pointwise limit $\zeta_0^z(\cdot)$ 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 $\zeta_0(u) := \zeta_0^z(u)$ 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 2. Assumptions (H3) and (H4) are the only conditions involving the conditional probability and the conditional probability density of Z given 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 (H0) 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.* (2007)).

Remark 3. Assumptions (H7), (H8) and (H10) are classical in functional estimation for finite or infinite dimension spaces.

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^z on $\mathcal{S}$. We will suppose that h^z is sufficiently smooth (at least of class $\mathcal{C}^2$) and verifies that $h'^z(\theta) = 0$ and $h''^z(\theta) < 0$.

We can write an estimator of the first derivative of the conditional hazard function through the first derivative of the estimator (1). Our maximum estimate is defined by assuming that there is some unique $\widehat{\theta}$ on $\mathcal{S}$ such that $0 = \widehat{h'}(\widehat{\theta}) < |\widehat{h'}^z(x)|$ for all $x \in \mathcal{S}$ and $x \neq \widehat{\theta}$.

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 $\mu(z)$, there exists $\xi > 0$ such that $|\theta(z) - \mu(z)| \geq \varepsilon$ implies that $|h^z(\theta(z)) - h^z(\mu(z))| \geq \xi$.

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^z 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 $Z = z$.

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

$$\widehat{h'}^z(x) = \frac{\widehat{f'}^z(x)}{1 - \widehat{F}^z(x)} + (\widehat{h}^z(x))^2, \tag{4}$$

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

$$\widehat{f'}^z(x) = \frac{\sum_{i=1}^n h_n^{-2} K(h_n^{-1} d(z, Z_i)) H'(h_n^{-1}(x - X_i))}{\sum_{i=1}^n K(h_n^{-1} d(z, Z_i))}. \tag{5}$$

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^z (but inherent problems of F^z, f^z and f'^z estimation), and secondly they consist with the assumptions usually made under functional variables.

Remark 4. 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-Pérez *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 1. *Under assumption (H0)-(H8) we have*

$$\widehat{\theta} - \theta \to 0 \quad a.co. \tag{6}$$

Remark 5. 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 $h'^z(\cdot)$ is continuous, we have, for all $\epsilon > 0$. $\exists\ \eta(\epsilon) > 0$ such that

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

Therefore,

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

We also have

$$|h'^z(\widehat{\theta}) - h'^z(\theta)| \leq |h'^z(\widehat{\theta}) - \widehat{h}'^z(\widehat{\theta})| + |\widehat{h}'^z(\widehat{\theta}) - h'^z(\theta)| \leq \sup_{x\in\mathcal{S}} |\widehat{h}'^z(x) - h'^z(x)|, \tag{7}$$

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

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

Lemma 1. *Under assumptions of Theorem 1, we have*

$$\sup_{x\in\mathcal{S}} |\widehat{h}'^z(x) - h'^z(x)| \to 0 \quad a.co. \tag{8}$$

The proof of this lemma will be given in Appendix.

Theorem 2. *Under assumption (H1)-(H8) and (H10c) we have*

$$\sup_{x\in\mathcal{S}} |\widehat{\theta}-\theta| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n h_n^3 \phi_z(h_n)}}\right) \tag{9}$$

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

$$h'^z(\widehat{\theta}) = h'^z(\theta) + (\widehat{\theta}-\theta) h''^z(\theta_n^*), \tag{10}$$

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

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

$$|\widehat{\theta}-\theta| \leq \frac{1}{h''^z(\theta_n^*)} \sup_{x\in\mathcal{S}} |\widehat{h}'^z(x) - h'^z(x)| \tag{11}$$

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

Lemma 2. *Under the assumptions of Theorem 2, we have*

$$\sup_{x\in\mathcal{S}} |\widehat{h}'^z(x) - h'^z(x)| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n h_n^3 \phi_z(h_n)}}\right) \tag{12}$$

The proof of lemma will be given in the Appendix.

4. Asymptotic normality

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

(H7d) $\int_{\mathbb{R}}(H'^2 dt < \infty$.
(H11) $0 = \widehat{h'}^z(\widehat{\theta}) < |\widehat{h'}^z(x)|), \forall x \in \mathcal{S},\ x \neq \widehat{\theta}$.

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

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

Theorem 3. *Under conditions (H0)-(H11) we have* $(\theta \in \mathcal{S}/f^z(\theta), 1 - F^z(\theta) > 0)$

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

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

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

and

$$\sigma^2_{h'}(\theta) = \frac{a_2^x h^z(\theta)}{\left(a_1^x\right)^2 \left(1 - F^z(\theta)\right)} \int (H'^2 dt.$$

Proof. Using again (17), and the fact that

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

and

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

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

Lemma 3. *Under Assumptions (H0)-(H3) and (H7)-(H9), we have*

$$\left(n\phi_z(h_n)\right)^{1/2}\left(\widehat{F}^z(x) - F^z(x)\right) \xrightarrow{\mathcal{D}} N\left(0, \sigma^2_{F^z}(x)\right) \tag{13}$$

where

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

Lemma 4. *Under Assumptions (H0)-(H4) and (H6)-(H10), we have*

$$(nh_n\phi_z(h_n))^{1/2}\left(\widehat{h}^z(x)-h^z(x)\right)\xrightarrow{\mathcal{D}}N\left(0,\sigma^2_{h^z}(x)\right) \tag{14}$$

where

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

Lemma 5. *Under Assumptions of Theorem 3, we have*

$$\left(nh_n^3\phi_z(h_n)\right)^{1/2}\left(\widehat{f'}^z(x)-f'^z(x)\right)\xrightarrow{\mathcal{D}}N\left(0,\sigma^2_{f'^z}(x)\right) \tag{15}$$

where

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

The proofs of Lemma (3) can be seen in Laksaci *et al.* (2011).

Finally, by this last result and (10), the following theorem follows:

Theorem 4. *Under conditions (H1)-(H11) we have* $(\theta\in\mathcal{S}/f^z(\theta),1-F^z(\theta)>0)$

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

with $\sigma^2_{h'}(\theta)=h^z(\theta)\left(1-F^z(\theta)\right)\int H'^2dt.$

5. Proofs of technical lemmas

*Proof (Lemma **1** and Lemma **2**).* Let

$$\widehat{h}'^z(x)=\frac{\widehat{f}'^z(x)}{1-\widehat{F}^z(x)}+(\widehat{h}^z(x))^2, \tag{16}$$

with

$$\widehat{h}'^z(x)-h'^z(x)=\underbrace{\left\{\left(\widehat{h}^z(x)\right)^2-(h^z(x))^2\right\}}_{\Gamma_1}+\underbrace{\left\{\frac{\widehat{f}'^z(x)}{1-\widehat{F}^z(x)}-\frac{f'^z(x)}{1-F^z(x)}\right\}}_{\Gamma_2} \tag{17}$$

for the first term of (17) we can write

$$\left|\left(\widehat{h}^z(x)\right)^2-(h^z(x))^2\right|\le\left|\widehat{h}^z(x)-h^z(x)\right|.\left|\widehat{h}^z(x)+h^z(x)\right| \tag{18}$$

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

$$\sup_{x\in\mathcal{S}}\left|\left(\widehat{h}^z(x)\right)^2-(h^z(x))^2\right|\le 2\left|h^z(\theta)\right|\sup_{x\in\mathcal{S}}\left|\widehat{h}^z(x)-h^z(x)\right| \tag{19}$$

for the second term of (17) we have

$$\frac{\widehat{f}'^z(x)}{1-\widehat{F}^z(x)} - \frac{f'^z(x)}{1-F^z(x)} = \frac{1}{(1-\widehat{F}^z(x))(1-F^z(x))}\left\{\widehat{f}'^z(x) - f'^z(x)\right\}$$
$$+ \frac{1}{(1-\widehat{F}^z(x))(1-F^z(x))}\left\{f'^z(x)\left(\widehat{F}^z(x) - F^z(x)\right)\right\}$$
$$- \frac{1}{(1-\widehat{F}^z(x))(1-F^z(x))}\left\{F^z(x)\left(\widehat{f}'^z(x) - f'^z(x)\right)\right\}.$$

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

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

$$C\frac{\left\{\sup_{x\in\mathcal{S}}\left|\widehat{f}'^z(x) - f'^z(x)\right| + \sup_{x\in\mathcal{S}}\left|\widehat{F}^z(x) - F^z(x)\right|\right\}}{\inf_{x\in\mathcal{S}}\left|1-\widehat{F}^z(x)\right|}. \tag{20}$$

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

$$\sup_{x\in\mathcal{S}}|\widehat{F}^z(x) - F^z(x)| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_z(h_n)}}\right) \tag{21}$$

$$\sup_{x\in\mathcal{S}}|\widehat{f}'^z(x) - f'^z(x)| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{nh_n^3\phi_z(h_n)}}\right) \tag{22}$$

$$\sup_{x\in\mathcal{S}}|\widehat{h}^z(x) - h^z(x)| = \mathcal{O}\left(h_n^{b_1}\right) + \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{nh_n\phi_z(h_n)}}\right) \tag{23}$$

$$\exists\delta > 0 \;\text{ such that }\; \sum_1^\infty \mathbb{P}\left\{\inf_{y\in\mathcal{S}}|1-\widehat{F}^z(x)| < \delta\right\} < \infty. \tag{24}$$

The proofs of (21) and (22) appear in Ferraty *et al.* (2007) and (23) is proven in Ferraty *et al.* (2008).

- Concerning (24) by equation (21), we have the almost complete convergence of $\widehat{F}^z(x)$ to $F^z(x)$. Moreover,

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

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

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

thus,

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

In terms of probability is obtained

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

Finally, it suffices to take $\delta = (1 - \sup_{x\in\mathcal{S}} F^z(x))/2$ and apply the results (21) to finish the proof of the lemma.

Proof (Lemma 4). We can write for all $x \in \mathcal{S}$

$$\begin{aligned}
\widehat{h}^z(x) - h^z(x) &= \frac{\widehat{f}^z(x)}{1-\widehat{F}^z(x)} - \frac{f^z(x)}{1-F^z(x)} \\
&= \frac{1}{\widehat{D}^z(x)}\Big\{ \left(\widehat{f}^z(x) - f^z(x)\right) + f^z(x)\left(\widehat{F}^z(x) - F^z(x)\right) \\
&\qquad - F^z(x)\left(\widehat{f}^z(x) - f^z(x)\right) \Big\}, \\
&= \frac{1}{\widehat{D}^z(x)}\Big\{ (1-F^z(x))\left(\widehat{f}^z(x) - f^z(x)\right) \\
&\qquad - f^z(x)\left(\widehat{F}^z(x) - F^z(x)\right) \Big\} \qquad (25)
\end{aligned}$$

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

As a direct consequence of the Lemma 3, the result (26) (see Ezzahrioui and Ould-Saïd, 2010) and the expression (25), permit us to obtain the asymptotic normality for the conditional hazard estimator.

$$(nh_n\phi_z(h_n))^{1/2}\left(\widehat{f}^z(x) - f^z(x)\right) \xrightarrow{\mathcal{D}} N\left(0, \sigma^2_{f^z}(x)\right) \qquad (26)$$

where

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

Proof (Lemma 5). For $i = 1, \ldots, n$, we consider the quantities $K_i = K\left(h_n^{-1}d(z, Z_i)\right)$, $H'_i(x) = H'\left(h_n^{-1}(x - X_i)\right)$ and let $\widehat{f'}^Z_N(x)$ (resp. $\widehat{F}^Z_D$) be defined as

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

This proof is based on the following decomposition

$$\widehat{f'}^z(x) - f'^z(x) = \frac{1}{\widehat{F}_D^z}\left\{\left(\widehat{f'}_N^z(x) - \mathbb{E}\widehat{f'}_N^z(x)\right) - \left(f'^z(x) - \mathbb{E}\widehat{f'}_N^z(x)\right)\right\} + \frac{f'^z(x)}{\widehat{F}_D^z}\left\{\mathbb{E}\widehat{F}_D^z - \widehat{F}_D^z\right\} \tag{27}$$

and on the following intermediate results.

$$\sqrt{nh_n^3\phi_z(h_n)}\left(\widehat{f'}_N^z(x) - \mathbb{E}\widehat{f'}_N^z(x)\right) \xrightarrow{\mathcal{D}} N\left(0, \sigma^2_{f'^z}(x)\right) \tag{28}$$

where $\sigma^2_{f'^z}(x)$ is defined as in Lemma 5.

$$\lim_{n\to\infty}\sqrt{nh_n^3\phi_z(h_n)}\left(\mathbb{E}\widehat{f'}_N^z(x) - f'^z(x)\right) = 0 \tag{29}$$

$$\sqrt{nh_n^3\phi_z(h_n)}\left(\widehat{F}_D^z(x) - 1\right) \xrightarrow{\mathbb{P}} 0, \quad \text{as} \quad n\to\infty. \tag{30}$$

- Concerning (28). By the definition of $\widehat{f'}_N^z(x)$, it follows that

$$\sqrt{nh_n^3\phi_z(h_n)}\left(\widehat{f'}_N^z(x) - \mathbb{E}\widehat{f'}_N^z(x)\right) = \sum_{i=1}^n \frac{\sqrt{\phi_z(h_n)}}{\sqrt{nh_n}\mathbb{E}K_1}\left(K_iH'_i - \mathbb{E}K_iH'_i\right) = \sum_{i=1}^n \Delta_i,$$

which leads

$$\sum_{i=1}^n \mathbb{E}\Delta_i^2 = \frac{\phi_z(h_n)}{h_n\mathbb{E}^2K_1}\mathbb{E}K_1^2(H'_1)^2 - \frac{\phi_z(h_n)}{h_H\mathbb{E}^2K_1}\left(\mathbb{E}K_1H'_1\right)^2 = \Pi_{1n} - \Pi_{2n}. \tag{31}$$

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

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

Meanwhile, by (H0), (H4), (H8) and (H9), it follows that:

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

which leads

$$\Pi_{1n} \underset{n\to\infty}{\longrightarrow} \frac{a_2^x f^z(x)}{(a_1^x)^2}\int (H'^2dt, \tag{32}$$

Regarding Π_{2n}, by (H0), (H4) and (H7), we obtain

$$\Pi_{2n} \underset{n\to\infty}{\longrightarrow} 0. \tag{33}$$

This result, combined with (31) and (32), allows us to get

$$\lim_{n\to\infty}\sum_{i=1}^n \mathbb{E}\Delta_i^2 = \sigma^2_{f'^z}(x). \tag{34}$$

Secondly, by the boundedness of H', we have

$$\begin{aligned}\mathbb{E}\left(|\Delta_i\Delta_j|\right) &\leq \frac{C\phi_z(h_n)}{n\mathbb{E}^2K_1}\left(K_iK_j + \mathbb{E}K_iK_j\right)\\ &\leq \frac{C}{nh_n}\left\{\left(\frac{\phi_z(h_n)}{n}\right)^{1/a} + \phi_z(x)(h_n)\right\}, \quad \forall i\neq j.\end{aligned}$$

Then, taking

$$\delta_n = \max_{1\leq i\neq j\leq n}\left\{\mathbb{E}\left(|\ \Delta_i\Delta_j|\right)\right\} = \frac{C}{nh_n}\left(\left(\frac{\phi_z(h_n)}{n}\right)^{1/a} + \phi_z(x)(h_n)\right).$$

Leads

$$nm_n\delta_n = \frac{Cm_n}{h_n}\left(\left(\frac{\phi_z(h_n)}{n}\right)^{1/a} + \phi_z(x)(h_n)\right). \tag{35}$$

Similarly, the boundedness of H' and K allows us to take $C_i = \mathcal{O}\left(\frac{1}{\sqrt{nh_n^3\phi_z(h_n)}}\right)$, which implies that

$$\left(\sum_{j=m_n+1}^{\infty}\alpha(j)\right)\sum_{i=1}^{n}C_i^2 \leq \frac{C}{h_n\phi_z(h_n)}\int_{t\geq m_n}t^{-a}dt = \frac{C}{h_n\phi_z(h_n)}\frac{m_n^{-a+1}}{a-1}. \tag{36}$$

Then, the sum of the right side of (35) and (36) is of type $Am_n + Bm_n^{-a+1}$, by talking $m_n = (A/B)^{-1/a} = \{(a-1)\phi_z(h_n)((\frac{\phi_z(h_n)}{n})^{1/a} + \phi_z(h_n))\}^{-1/a} \to \infty$, it is clear that, under conditions (H10a) and (H10b), combining (35) and (36) allows us to get

$$nm_n\delta_n = o(1), \tag{37}$$

and

$$\left(\sum_{j=m_n+1}^{\infty}\alpha(j)\right)\sum_{i=1}^{n}C_i^2 = o(1), \tag{38}$$

respectively. Finally, by choosing $\varrho_n = \sqrt{\frac{nh_n^3\phi_z(h_n)}{\log n}}$, under (H10a) again and $a>3$, we have

$$\frac{\varrho_n}{\sqrt{n}} = o(1) \tag{39}$$

and

$$\begin{aligned}\frac{n}{\varrho_n}\alpha(\varepsilon\varrho_n) &\leq C\frac{(\log n)^{(a+1)/2}}{n^{(a-1)/2}(h_n^3\phi_z(h_n))^{(a+1)/2}}\\ &\leq C\frac{(\log n)^{(a+1)/2}}{n^{(a-3)/2}} \to 0 \quad \text{as } n\to\infty.\end{aligned}$$

Therefore, combining (33)-(39) with Corollary 2.2 in Liebscher (2001), (28) is valid.

- Concerning (29). The proof is completed along the same steps as that of Π_{1n}. We omit it here.

- Concerning (30). The idea is similar to that given by Ferraty *et al.* (2007).
  By definition of $\widehat{F}_D^z(x)$, we have

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

where $\Omega_n=\frac{\sqrt{nh_n^3\phi_z(h_n)}\sum_{i=1}^n K_i}{n\mathbb{E}K_1}$. In order to prove (30), similar to Ferraty *et al.* (2007), we only need to proov $Var\ \Omega_n\to 0$, as $n\to\infty$. In fact, since

$$\begin{aligned}
Var\ \Omega_n &= \frac{nh_n^3\phi_z(h_n)}{n\mathbb{E}^2K_1}\left(nVarK_1+\sum_{1\le i}\sum_{j\le n}cov(K_i,K_j)\right)\\
&\le \frac{nh_n^3\phi_z(h_n)}{\mathbb{E}^2K_1}\mathbb{E}K_1^2+\frac{nh_n^3\phi_z(h_n)}{n\mathbb{E}^2K_1}\sum\sum_{0\le|i-j|\le v_n}cov(K_i,K_j)\\
&+\frac{nh_n^3\phi_z(h_n)}{n\mathbb{E}^2K_1}\sum\sum_{0\le|i-j|\ge v_n}cov(K_i,K_j)\\
&=\Psi_1+\Psi_2+\Psi_3,
\end{aligned}$$

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

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

Meanwhile, by (H0) and (H1), it follows that

$$\Psi_2\le v_nh_n^3\left\{\left(\frac{\phi_z(h_n)}{n}\right)^{1/a}+\phi_z(h_n)\right\}. \tag{40}$$

Finally, using the Davydov-Rio's inequality in Rio (2000) for mixing processes leads to

$$|cov(K_i,K_j|\le C\alpha(|i-j|),$$

for all $i\ne j$. Then, we have

$$\begin{aligned}
\Psi_3 &\le \frac{h_n^3\phi_z(h_n)}{n\mathbb{E}^2K_1}n^2C\alpha(|i-j|)\le C\frac{h_n^3\phi_z(h_n)}{n\mathbb{E}^2K_1}n^2v_n^{-a+1}\\
&\le Ch_n^3nv_n^{-a+1}.
\end{aligned} \tag{41}$$

Since the right side of (40) and (41) is also of type $Av_n+Bv_n^{-a+1}$, by choosing $v_n=[n^{-1}((\frac{\phi_z(h_n)}{n})^{1/a}+\phi_z(h_n))]^{-1/a}\to\infty$ and simple calculations, we get that $\Psi_2\to 0$ and $\Psi_3\to 0$ as $n\to\infty$, respectively. Therefore, the proof of this result is completed.

Therefore, the proof of this lemma is completed.

References

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

Benhenni, K., Ferraty, F., Rachdi, M. and Vieu, P., 2007. Local smoothing regression with functional data. *Comput. Statist.* **22**, 353-369.

Benhenni, K., Hedli-Griche, S., Rachdi M. and Vieu, P., 2008. Consistency of the regression estimator with functional data under long memory conditions. *J. of Statist. Probab. Lett.* **78**(8) , 1043-1049.

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

Besse, P. and Ramsay, J.O., 1986. Principal component analysis of sampled curves, *Psychometrika.* **51**, 285-311.

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

Cardot, H., Ferraty, F. and Sarda, P., 1999. Functional linear model. *Statist. Probab. Lett.* **45**, 11-22.

Damon, J. and Guillas, S., 2002. The inclusion of exogenous variables in functional autoregressive ozone forecasting. *Environmetrics.* **13**, 759-774.

Doukhan, P., 1994. *Mixing: Properties and Examples.* Lecture Notes in Statist. 85, Springer-Verlag, NewYork.

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.

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.

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

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.

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.

Ferraty, F., Rabhi, A. and Vieu, P., 2005. Conditional quantiles for functional dependent data with application to the climatic El Ninõ phenomenon. *Sankhyã.* **67**, 378-399.

Ferraty, F., Rabhi, A. and Vieu, P., 2008. t Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle. *Rev. Roumaine Math. Pures Appl.* **53**, 1-18.

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

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.

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.

Laksaci, A., Lemdani, M. and Ould-Saïd, E., 2011. Asymptotic results for an L^1 -norm kernel estimator of the conditional quantile for functional dependent data with application to climatology, *Sankhyã: The Indian Journal of Statistics* **73**(1)-A, 125-141.

Lecoutre, J.P. and Ould-Saïd, E., 1992. 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.

Liebscher, E. 2001. Central limit theorem for α-mixing triangular arrays with applications to nonparemetric statistics, *Mathematical Methods of Statistics*, **10**(2), 194-214.

Mahiddine, A., Bouchentouf, A.A. and Rabhi, A., 2014. Nonparametric estimation of some characteristics of the conditional distribution in single functional index model, *Malaya Journal of Matematik*, , to appear .

Masry, E. 2005. Non-parametric regression estimation for dependent functional data: Asymptotic normality. *Stoch. Process. Appl.* **115**, 155-177.

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

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

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

Rachdi, M. and Vieu, P. 2007. Non-parametric regression for functional data: Automatic smoothing parameter selection. *J. Stat. Plan. Inference.* **137**, 2784-2801.

Ramsay, J.O. and Silverman, B.W., 2005. *Functional Data Analysis, 2nd ed.* Springer-Verlag, NewYork.

Rice, J. and Silverman, B.W., 1991. Estimating the mean and covariance structure nonparametrically when the data are curves. *J.R. Stat. Soc.* Ser. B. 53 , 233-243.

Rio, E., 2000. Théorie asymptotique des processus aléatoires faiblements dépendant. *Mathématiques Application*, **22**(3), 331-334.

Rosenblatt, M. 1956. A central limit theorem and a strong mixing condition. *Proc. Nat.Acad. Sci.* **42**, 43-47.

Roussas, G. 1989. Hazard rate estimation under dependence conditions. *J. Statist. Plann. Inference.* **22**, 81-93.

Singpurwalla, N.D. and Wong, M.Y., 1983. Estimation of the failure rate - A survey of nonparametric methods. Part I: Non-Bayesian methods. *Commun. Stat. Theory Math.* **12**, 559-588.

Spierdijk, L., 2008. Non-parametric conditional hazard rate estimation: A local linear approach. *Comput. Stat. Data Anal.* **52**, 2419-2434.

Vere-Jones, D., 1970. Stochastic models for earthquake occurrence. *J. Roy. Statist. Soc.* Ser. B. **32**, 1–62

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

Watson, G. S. and Leadbetter, M. R., (1964). Hazard analysis II. *Sankhyã.* Ser. A 26 1964 101-116.

General Bibliography

Abramovich, F., Angelini, C. (2006). Bayesian maximum a posteriori multiple testing procedure. Sankhya, 68 : 436-460.

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

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

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

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

Aneiros-Perez, G., Vieu, P. (2007). Semi-functional partial linear regression. Statist. Proba. Lett. **76**, 1102-1110.

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

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.

Azzalini, A. (1981) A note on the estimation of a distribution function and quantiles by kernel method. Biometrika, **60**, 326-328.

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

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

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

Benko, M., Hardle, W. and Kneip, A. (2005). Common functional principal components. SFB 649 Economic Risk Discussion Paper, 2006-010.

Beran, R. (1981). Nonparametric regression with randomly censored survival data. Technical Report, Dept. Statisticcs, Univ. California, Berkely.

BERLINET A., BIAU G.et ROUVI'ERE L.(2005).Parameter selection in modified histogram estimates. Statistics, **39**, 91-105.

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

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.

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

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

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

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

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.

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

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

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.

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

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

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

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

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

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.

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

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

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

Dabo-Niang, S. (2002). Estimation de la densité dans un espace de dimension infinie : application aux diffusions. C. R. Acad. Sci., Paris. **334**, 213-216.

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

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.

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

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

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

Delsol, L. (2007). Régression non paramétrique fonctionnelle : expression asymptotique des moments, Ann. I.S.U.P. Vol LI, **3**, 43-67.

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

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

Delsol, L., Ferraty, F., Vieu, P. (2011). Structural test in regression on functional variables. J. Multivariate Anal. **102**, No.3, 422-447.

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

Engle, R. F., Russell, J. (1988). Autoregressive conditional duration : A new model for irregulary spaced transaction data. Econometrica. **66**, 1127-1162.

Estévez-Pérez, G. (2002). On convergence rates for quadratic errors in kernel hazard estimation. Statist. Probab. Lett. **57**, 231-241.

Estévez-Pérez, G., Quintela-del-Rio, A., Vieu, P. (2002). Convergence rate for cross-validatory bandwidth in kernel hazard estimation from dependent samples. J. Statist. Plann. Inference **104**, 1-30.

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.

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

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.

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.

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

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

Ferraty, F., Rabhi, A., Vieu, P. (2005). Conditional quantiles for functionally dependent data with application to the climatic El Phenomenon. Sankhya, **67**, 378-399.

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

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

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

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

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

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

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

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.

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

Ferré, L., Yao, A. F. (2005). Smoothed Functional Sliced Inverse Regression. Statist. Sinica **15**, 665-683.

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

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

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.

Gasser, T., Müller, H. G. (1979). Kernel estimation of regression function. In : T. Gasser and M. Rosenblatt (eds.) Smoothing techniques for curve estimation. Springer, Heidelberg, 23-68.

Gefeller, O., Michels, P. (1992). A review on smoothing methods for the estimation of the hazard rate based on kernel functions. In Y. Dodge and J. Whittaker (Eds). Computational Statistics. Physica-Verlag, 459-464.

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

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

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

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.

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

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

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

HM00 Härdle, W., Müller, M. (2000). Multivariate and semiparametric regression. In : M. Schimek (Ed) Smoothing and regression ; Approaches, Computation, and Application. Wiley Series in Probability and Statistics, Wiley, New York, 357-392.

Hassani, S., Sarda, P., Vieu, P. (1986). Approche non-paramétrique en théorie de la fiabilité. Rev. Statist. Appl. **35**, 27-41.

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

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

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

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

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

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

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

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

Izenman, A. (1991). Developments in nonparametric density estimation. J. Amer. Statist. Assoc. **86**, 205-224.

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.

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

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 (1961).

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

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.

Lecoutre, J. P., Ould-Saïd, E. (1995). Hazard rate estimation for strong mixing and censored processes. J. Nonparametric Statist. **5**, 83-89.

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

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

Mahiddine, A., Bouchentouf, A. A., Rabhi, A. Nonparametric estimation of some characteristics of the conditional distribution in single functional index model. Accepted in Malaya Journal of Matematik (MJM).

Mas, A., Pumo, B. (2007). The ARHD model. J. Statist. Plann. Inference. **137**, 538-553.

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

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

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

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

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

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

Nielson, J. P., Linton, O. (1995). Kernel estimation in a nonparametric marker dependent hazard model. Ann. Statist. **23**, 1735-1748.

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

Padgett, W. J. (1988). Nonparametric estimation of density and hazard rate functions when samples are censored. In P.R. Krishnaiah and C.R. Rao (Eds), *Handbook of Statistics*, **7**, 313-331. Elsevier Science Publishers.

Pascu, M., Vaduva, I. (2003). Nonparametric estimation of the hazard rate : a survey. Rev. Roumaine Math. Pures &t Appl. **48**, 173-191.

Patil, P. N., Wells, M. T., Maron, J. S. (1994). Some heuristics of kernel based estimators of ratio functions. J. Nonparametr. Statist. **4**, 203-209.

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

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

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

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

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

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

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

Reiss, R.-D. (1981). Nonparametric estimation of smooth distribution functions. Scand. J. Statist. **8**, 116-119.

Rossi, L., Lubello, C., Poggiali, E., Griffini, O. (2002) . Analysis of a clariflocculation process with a photometric dispersion analyser (PDA2000). Water Supply. **2**, No.5-6, 57-63.

Roussas, G. (1969). Nonparametric estimation of the transition distribution function of a Markov process. Annals of Math. Stat. **40**, 1386-1400.

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

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.

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

Singpurwalla, N. Wong, M. Y. (1983). Estimation of the failure rate - a survey of nonparametric models. Part I : Non-Bayesian methods. Comm. Statist. Theory Methods **12**, 559-588.

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

Theodoros, N., Yannis G. Y. (1997) Rates of convergence of estimate, Kolmogorov entropy and the dimensioanlity reduction principle in regression. The Annals of Statistics, **25**, No.6, 2493-2511.

Troutman, J. L. (1996). Variatonal Calculus and Optimal Control. Springer-Verlag.

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

Turnbull, B. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. J. R. S. Soc. Series : B **38**, 290-295.

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

Van Keilegom, I. Veraverbeke, N. (2001). Hazard rate estimation in nonparametric regression with censored data. Ann. Inst. Statist. Math., **53**, 730-745.

Vieu, P. (1991). Quadratic errors for nonparametric estimates under dependence. J.Multivariate. Anal. **39**, 324-347.

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

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