Dedication

This thesis is dedicated to

My mother, who sincerely raised me with here caring and offered me unconditional love, a very special thank for the myriad of ways in which, throughout my life, you have actively supported me in my determination to find and realize my potential.

My sisters specially Karima , brothers and the husbands of my sisters who have supported me all the way since the beginning of my study.

To the best children of family and spatially my daughter ***Amira Kamar.***

To those who have been deprived from their right to study and to all those who believe in the richness of learning.

Acknowledgments

First of all, I thank **ALLAH** for His help and blessing.
This thesis would not have been possible without the invaluable support and guidance of so many people. To whom I am greatly indebted:

I want to express my great appreciation and deepest gratitude to my supervisor, **Mr.ABBES** ***Rabhi*** without whom this thesis would not have been emerged. It was a great honor and an immense pleasure to be able to accomplish this work under his direction. I could benefit from his sound advice and careful guidance, his availability, his kindness and his constant support to carry out this work. Again thank you !

I want to thank very warmly **Mr.ABDELDJEBBAR** ***Kandouci*** for chairing my committee members and for his help,. I heartily thank **Mr.FETHI** ***Madani*** for his help advice and support throughout the course of this thesis.
I am very honored that **Mr.TOUFIK** ***Guendouzi*** , MISS.**AMINA** ***Angelika Bouchentouf*** agreed to participate in my jury, thank you for your time and effort in reviewing this work.

Warm thanks to all the members of laboratory of stochastic models, statistic and applications, for their kindness and expertise.

Many thanks to all my friends and colleague: ***S. Idrisi, Z. Goumidi, M. Kadrouci, K. mehdi, D. Bennafla, K. Rouane*** , for their encouragement, backing, and understanding, in particular during the time of preparation of this work. I hope to get the chance to return the favor.

The most important persons are saved for last: My mother and my sisters ***Karima, Houaria, Rekia,***,special thank to my brother ***bendjaloule*** who supported me all the period of my study. deserve special mention for their inseparable support and prayers. I am deeply and forever indebted to them for their love and encouragement throughout my entire life.

I would like to thank everybody who was important to the successful realization of thesis, as well as expressing my apology that I could not mention personally one by one.

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 l'erreur quadratique, dont nous donnons nos premiers résultats lorsque l'échantillon considéré est i.i.d. en premier lieu et fortement mélangeant en deuxième lieu.
Dans un second temps nous supposons que la variable explicative est à valeurs dans un espace de Hilbert (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 indépendant. Pour le cas où les observations sont indépendantes identiquement distribuées (i.i.d.), nous obtenons la convergence ponctuelle et uniforme presque complète avec vitesse de l'estimateur construit. Comme application nous discutons l'impact de ce résultat en prévision non paramétrique fonctionnelle à partir de l'estimation de mode conditionnelle. Les données incomplètes sont modélisées via la présence de la censure à droite des variables. Dans ce contexte nous établissons la convergence ponctuelle et uniforme presque complète avec vitesse de l'estimateur construit de l'estimateur à noyau de la fonction de hasard conditionnelle. Nos résultats asymptotiques exploitent bien la structure topologique de l'espace fonctionnel de nos observations et le caractère fonctionnel de nos modèles. En effet, toutes nos vitesses de convergence sont quantifiées en fonction de la concentration de la mesure de probabilité de la variable fonctionnelle, de l'entropie de Kolmogorov et du degré de régularité des modèles.
Autant que l'on sache le problème de l'estimation non paramétrique de la fonction de hasard conditionnelle de modèle d'indice fonctionnel sur des données complètes et/ou censurées n'a pas été abordé. En général, l'estimation sur des données censurées à indice fonctionnel est récente dans la littérature statistique.

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

Contents

Chapter 1

Introduction.

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

1.0.1 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 [19] for the theoretical aspects, Ferraty and Vieu [69] for non-parametric studyet Ferraty and Romain [64] for recent developments. In the same context, we refer to Manteiga and Vieu [112] well as Ferraty [57]. 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 [21], Schimek [147], Sarda and Vieu [146] anf Ferraty and Vieu ([67], [69]) 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 functional data are finally relatively "ancient" Rao [137] and Tucker [155] 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 [132] 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 ([135] and [136]), Ferraty and Vieu [69]...

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 (for instance common phenomenon in medical applications), possible presence of dependency between the observed variables (for instance a common phenomenon in applications seismic or econometric) 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.

1.1 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.

1.1.1 On the regression model

The first results in functional nonparametric statistics were developed by Ferraty and Vieu [65] 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 [68] 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 [114] has proved asymptotic normality of the estimator of Ferraty et Vieu [68] for the regression function. The reader can find in the book of Ferraty and Vieu [69], a wide range of applications of the regression function in functional statistics. Convergence in mean squared was investigated by Ferraty *a*nd al. [60]. Specifically, they have explained the exact asymptotic term of the quadratic error. This result was used by Rachdi and Vieu [130] 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 *a*nd al. [13]. 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 [64] well as Delsol [49]. Results on uniform integrability were established by Delsol ([47],[48]) andt Delsol *a*nd al. [50]. Other works were interested to estimating the regression function using different approaches : the method of k nearest neighbors by Burba *a*nd al. [25], robust technical by Azzidine *a*nd al. [11] and Crambes *a*nd al. [37], the estimate by the simplified method of local polynomial by Barrientos-Marin *a*nd al. [12].

1.1.2 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 [122], Holmstrom [95] in climatic, Deville [51] in econometric, Molenaar and Boosma [115] and then Kirkpatrick and Heckman [103] 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, ([134], [135]) presents an important collection of statistical methods for the functional variables. In the same way, note that Bosq [19] 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 *a*nd al. [27] built an estimator for the model of the Hilbertien linear regression similar to Bosq estimator [23] 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 [19].

Recently, Cardot *a*nd al. [29] 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 *a*nd al. [30] and to Cuevas *a*nd al. [38] for the problem of the test in the functional linear model. Several authors are interested also the answer variable is qualitative, for example, Hastie *a*nd al. [92], Hall *a*nd al. [84],....
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 [65]. These result were prolonged by Ferraty *a*nd al. [66]..., with the problems of the regression such forecast in the context of time series. By taking again the estimator of Ferraty and Vieu [68] and by using the property of concentration of the measurement of probability of the functional explanatory variable, Niang and Rhomari [41] 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 [22], Besse *a*nd al. [18]). Recently, of the first work relating to model presenting at the same time linear and nonparametric aspects were realized by Ferraty *a*nd al. [67], Aït-Saïd *a*nd al. ([3], [4]), Ferré and Villa Ferr[74]...
The first work on the functional variables of distribution estimate was given by Geffroy [81]. More recently, Gasser *e*t al. [78] then Hall and Heckman [84] 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 [26]. Niang [42] 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 *a*nd al. [68] 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 ([66], [67]) were interested in spectrometric data and with vocal recordings, Besse *a*nd al. [18] with weather data, Gasser *a*nd al. [78] considered medical data, Ferraty, Rabhi and Vieu [71] considered environmetrics 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...

1.1.3 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, [135]) and knee movement during exercise under constraint (Abramovich and Angelini [1], and Antoniadis and Sapatinas [8]. concerning animal biology, studies of the oviposition of medley were made by several authors (Chiou, Müller, Wang and Carey [34], Chiou, Müller and Wang [35], Cardot [28] and Chiou et Müller [33]). 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 [76] , Hastie and Mallows [93] who have commented on the article by Frank and Friedman [76] providing an example of the measuring curves log-intensity of a laser radius refracted depending on the angle of refraction. In [66], 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). Other articles like Ferraty and Vieu [66], Ferré and Yao [75], Ferraty *a*nd al. [58], Ferraty and Vieu [69], Aneiros-Pérez and Vieu [7], Ferraty, Mas and Vieu [60] and Mas and Pumo [113] 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 [6] 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 Ni*ñ*o (hot current in Pacific Ocean) has been realized by Besse Cardot and Stephenson [18]; Ramsay and Silverman [131], Ferraty *a*nd al. [62] and Hall and Vial [86].

• In linguistics, the works have also been realized, particularly concerning voice recognition. Mention may be made, for example Hastie Buja and Tibshirani [92], Berlinet Biau and Rouvière [16] or again Ferraty and Vieu ([67], [69]). 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 [92] and Ramsay [133]. The latter for example Modeling 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 [104], and rerecently by Benko, Härdle and Kneip [14], 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, [110]) or recorded by a radar (Hall and al [83]), the demographic studies (Hyndman and Ullah [98]),... 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 [104], Ramsay and Silverman [135], Delicado [46] and Nerini Ghattas [117]), functions regressions (Härdle and

Marron [89], Heckman and Zamar [94]), 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 [135]),...
One can imagine that in the future the use of statistical methods functional will be extended to other areas.

1.1.4 On the problematic of single index models

For several years, a increasing interest is worn to models which incorporating of both the parts parametric and nonparametric. Such models type are called semi-parametric model. This consideration is due primarily to problems due to poor specification of some models. Tackle a problem of miss-specification semiparametric way consists in not specify the functional form of some model components. This approach complete those non-parametric models, which can not be useful in small samples, or with a large number of variables. As example, in the classical regression case, the important parameter whose one assumed existence is the regression function of Y knowing the covariate X, denoted $r(x) = \mathbb{E}(Y|X = x)$, $X, Y \in \mathbb{R}^d \times \mathbb{R}$. For this model, the non-parametric method considers only regularity assumptions on the function r. Obviously, this method has some drawbacks. One can cite the problem of curse of dimensionality. This problem appears when the number of regressors d increases, the rate of convergence of the nonparametric estimator r which is supposed k times differentiable is $O(n^{-k/2k+d})$ deteriorate. The second drawback is the lack of means to quantify the effect of each explanatory variable. To alleviate in these drawbacks, an alternative approach is naturally provided by the semi-parametric model which supposes the introduction of a parameter on the regressors, by writing than the regression function is of the form

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

The models defined are known in the literature as the single-index models.
These models allow to obtain a compromise between parametric models, generally too restrictive and nonparametric model where the rate of convergence of the estimators deteriorate quickly in the presence of a large number of explanatory variables. In this area, types different of models have been studied in the literature : amongst the most famous, there may be mentioned additive models, partially linear models or single index models. The idea of these models, in the case of estimating the conditional density or regression consists in bring to the covariates a dimension in smaller than dimension of the space variable, thus allowing overcome the problem of curse of dimensionality. For example, for example, in the partially linear model, we decompose the quantity to be estimated, into a linear part and a functional part. This latter quantity does not pose estimation problem since it's expressed as a function of explanatory variables of finite dimension, thus avoiding the problems associated with curse of dimensionality. in order to treat the problem of curse of dimensionality in the case chronologies series, several semi-parametric approaches have been proposed. Without pretend to exhaustivity, we quote for example: Xia and An [161] for the index model. A general presentation of this type of model is given in Ichimura and al. [100] where the convergence and asymptotic normality are obtained. In the case of M-estimators, Delecroix and al. [44] proves the consistency and asymptotic normality

of the estimate the index and they study it's effectiveness. The statistical literature on these methods is rich, quote Huber [97] and Hall [85] present an estimation method which consists projecting the density and the regression function on a space of dimension one, to bring a non-parametric estimation for dimensional covariates. This amounts exactly to estimate these functions in a single index model. Attaoui and al. [9] have established the pointwise and the uniform almost complete convergence (with the rate) of the kernel estimate of this model. The interest of their study is to show how the estimate of the conditional density can be used to obtain an estimate of the simple functional index if the latter is unknown. More precisely, this parameter can be estimated by pseudo-maximum likelihood method which is based the preliminary estimate of the conditional density. recently Mahiddine and al. [111] have established the pointwise almost complete convergence and the uniform almost complete convergence (with the rate) of some characteristics of the conditional distribution and the successive derivatives of the conditional density when the observations are linked with a single-index structure and they are applied to the estimations of the conditional mode and conditional quantiles.
The single-index approach is widely applied in econometrics as a reasonable compromise between nonparametric and parametric models. Such kind of modelization is intensively studied in the multivariate case. Without pretend to exhaustivity, we quote for example Härdle et al. [88], Hristache and al. [96]. Based on the regression function, Delecroix and al. [45] studied the estimation of the single-index and established some asymptotic properties. The literature is strictly limited in the case where the explanatory variable is functional (that is a curve). The first asymptotic properties in the fixed functional single-model were obtained by Ferraty and al. [67]. They established the almost complete convergence, in the i.i.d. case, of the link regression function of this model. Their results were extended to dependent case by Aït Saidi and al. [3]. Aït Saidi and al. [4] studied the case where the functional single-index is unknown. They proposed an estimator of this parameter, based on the cross-validation procedure.

1.1.5 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 and al and al. [69]. 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 and al. [71] and Ezzahrioui and Ould Saïd [56]. we send back to Cardot et al. [29] 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 precise. The results obtain The results are detailed in Chapter 3 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.

1.1.6 On the conditional hazard function

The literature on estimating the conditional hazard function is relatively restricted into functional statistics. The article by Ferraty *e*t al. [63] 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 [128]. The latter established the asymptotic normality of the estimator proposed by Ferraty *a*nd al. [72].

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 [107] 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.

1.1.7 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

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

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 [116] and Watson [160] 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 [156] 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.

1.1.8 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

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

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

$$h(t) = \frac{f(t)}{S(t)} \tag{2}$$

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) that $-f(t)$ is the derivative of $S(t)$. This suggests rewriting Equation (2) as

$$h(t) = -\frac{d}{dt}\log S(t). \tag{3}$$

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:

$$S(t) = \exp\left\{-\int_0^t h(u)du\right\}. \tag{4}$$

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

$$H(t) = \int_0^t h(u)du. \tag{5}$$

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 (2). Given the hazard, we can always integrate to obtain the cumulative hazard and then exponentiate to obtain the survival function using Equation (4). An example will help fix ideas.

Example 1.1.1 *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.*

1.1.9 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 1.1.2 *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 1.1.3 *One says that the rate of almost complete convergence of $(X_n)_{n\in\mathbb{N}}$ to X is of order u_n if and only if*

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

and we write

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

Proposition 1.1.4 *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 1.1.5 *The almost convergence of Y_n to l_0 implies that there exists some $\delta > 0$ such that*

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

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

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

where, eventually, the positive real ε decreases with n. In this context, there exists powerful probabilistic tools, generically called 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 [77] (see also Nagaev [117] and [118])

Proposition 1.1.6 *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 1.1.7 *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 1.1.8 *By applying Proposition 1.1.6 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 1.1.9 *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* $\exists M = M_n < \infty$*,* $|Z_1| \leq M$ *and define* $\sigma_n^2 = \mathbb{E}Z_1^2$*. If* $u_n = n^{-1}\sigma_n^2 \log n$ *verifies* $\lim_{n\to\infty} u_n = 0$*, and if* $M/\sigma_n^2 < C < \infty$*, then we have:*

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

Remark 1.1.10 *By applying Proposition 1.1.6 with* $a_i^2 = \sigma^2$*,* $A_n = n\sigma^2$*, and by choosing* $\varepsilon = \varepsilon_0 \sqrt{u_n}$*, we obtain for some* $C' > 0$*:*

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

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

In the remainder of this work, we will consider only two kinds of kernel for weighting functional variables.

Definition 1.1.11 *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 1.1.12 *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 (**??**), 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 practical points of view. Because the notion of ball is strongly linked with the semi-metric d, the choice of this semi-metric will become an important stage.
Now, let X be a f.r.v. taking its values in the semi-metric space $(\mathcal{F}, d)$, let x be a fixed element of $\mathcal{F}$, let h be a real positive number and let K be a kernel function.

Lemma 1.1.13 *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 1.1.14 *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:

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

1.2 Various Approaches to the Prediction Problem

Let us start by recalling some notation. Let $(X_i, Y_i)i = 1, \ldots n$ 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 differentiability assumption on $F_Y^X(x, .)$, this functional conditional density can be written as

$$\forall y \in \mathbb{R},\quad f_Y^X(x, y) = \frac{\partial}{\partial y}F_Y^X(x, y). \tag{6}$$

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 conditional 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.

1.3 Kernel Estimators

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

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

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

where K is an asymmetrical kernel and h (depending on n) is a strictly positive real. It is a functional extension of the familiar Nadaraya-Watson estimate (see Nadaraya [116] and Watson [159] which was previously introduced for finite dimensional nonparametric regression (see Härdle [87] 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 [142] and Samanta [145] 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:

$$\widehat{F}_Y^X(x,y) = \frac{\sum_{i=1}^{n} K\left(h^{-1}d(x,X_i)\right) H(g^{-1}(y-Y_i))}{\sum_{i=1}^{n} K\left(h^{-1}d(x,X_i)\right)}, \tag{7}$$

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 (1.1.12). 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.

It is clear that the parameter g acts as the bandwidth h. The smoothness of the function $\widehat{F}_Y^X(x,.)$ is controlled 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 [10] and Reiss [138]. The roles of the other parameters involved in this functional kernel c.d.f. estimate (i.e. the roles of K and h) are the same as in the regression setting. From this conditional c.d.f. estimate (**??**).

- **Estimating the conditional density**. It is know that, under some differentiability assumption, the conditional density function can be obtained by derivation the conditional c.d.f. (see (6)). 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

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

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:

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

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 $\theta(x)$:

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

1.4 Topological considerations

1.4.1 Kolmogorov's entropy

The purpose of this section is to emphasize the topological components of our study. Indeed,as indicated in Ferraty and Vieu [69], all the asymptotic results in nonparametric statistics for functional variables are closely related to the concentration properties of the probability measure of the functional variable X. Here,we have more over to take into account the uniformity aspect. To this end,let $\mathcal{S}_\mathcal{F}$ be a fixed subset of $\mathcal{H}$ of; we consider the following assumption:

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

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

Definition 1.4.1 *Let $\mathcal{S}_\mathcal{F}$ be a subset of a semi-metric space $\mathcal{H}$, and let $\varepsilon > 0$ be given. A finite set of points $x_1, x_2, \ldots, x_N$ in $\mathcal{F}$ is called an ε-net for $\mathcal{S}_\mathcal{F}$ if $\mathcal{S}_\mathcal{F} \subset \bigcup_{k=1}^N B(x_k,\varepsilon)$.*

The quantity $\psi_{\mathcal{S}_\mathcal{F}}(\varepsilon) = \log\left(N_\varepsilon(\mathcal{S}_\mathcal{F})\right)$, where $N_\varepsilon(\mathcal{S}_\mathcal{F})$ is the minimal number of open balls in $\mathcal{F}$ of radius ε which is necessary to cover $\mathcal{S}_\mathcal{F}$, is called the Kolmogorov's ε-entropy of the set $\mathcal{S}_\mathcal{F}$.

This concept was introduced by Kolmogorov in the mid-1950's (see, Kolmogorov and Tikhomirov, [105]) and it represents a measure of the complexity of a set, in sense that, high entropy means that much information is needed to describe an element with an accuracy ε. Therefore, the choice of the topological structure (with other words, the choice of the semi-metric) will play a crucial role when one is looking at uniform (over some subset $\mathcal{S}_{\mathcal{F}}$) of $\mathcal{F}$) asymptotic results. More precisely, we will see thereafter that a good semi-metric can increase the concentration of the probability measure of the functional variable X as well as minimize the ε-entropy of the subset $\mathcal{S}_{\mathcal{F}}$. In an earlier contribution (see, Ferraty *a*nd al., [69]) we highlighted the phenomenon of concentration of the probability measure of the functional variable by computing the small ball probabilities in various standard situations. We will devote Section 1.4.2 to discuss the behavior of the Kolmogorov's ε-entropy in these standard situations. Finally, we invite the readers interested in these two concepts (entropy and small ball probabilities) or/and the use of the Kolmogorov's ε-entropy in dimensionality reduction problems to refer to respectively, Kuelbs and Li [106] or/and Theodoros and Yannis [151].

1.4.2 Some examples

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

Example 1.4.2 *(Compact subset in finite dimensional space) : A standard theorem of topology guaranties that for each compact subset $\mathcal{S}_{\mathcal{F}}$ of $\mathbb{R}^p$ and for each $\varepsilon > 0$ there is a finite ε-net and we have for any $\varepsilon > 0$,*

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

More precisely, Chate and Courbage [32] have shown that, for any $\varepsilon > 0$ the regular polyhedron in $\mathbb{R}^p$ with length r can be covered by $\left([2r\sqrt{p}/\varepsilon] + 1\right)^p$ balls, where $[m]$ is the largest integer which is less than or equal to m. Thus, the Kolmogorov's ε-entropy of a polyhedron P_r in $\mathbb{R}^p$ with length r is

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

Example 1.4.3 *(Closed ball in a Sobolev space) : Kolmogorov and Tikhomirov [105] obtained many upper and lower bounds for the ε-entropy of several functional subsets. A typical result is given for the class of functions $f(t)$ on $T = [0, 2p)$ with periodic boundary conditions and*

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

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

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

Example 1.4.4 *(Unit ball of the Cameron-Martin space) : Recently, Van der Vaart and Van Zanten [153] characterized the Cameron-Martin space associated to a Gaussian process viewed as map in* $\mathcal{C}[0,1]$ *with the spectral measure* μ *satisfying*

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

by

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

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

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

Example 1.4.5 *(Compact subset in a Hilbert space with a projection semi-metric) : The projection-based semi-metrics are constructed in the following way. Assume that* $\mathcal{H}$ *is a separable Hilbert space, with inner product* $< .,. >$ *and with orthonormal basis* $\{e_1, \ldots, e_j, \ldots\}$*, and let* k *be a fixed integer,* $k > 0$*. As shown in Lemma 13.6 of Ferraty and Vieu [69], a semi-metric* d_k *on* $\mathcal{H}$ *can be defined as follows*

$$d_k(x,x') = \sqrt{\sum_{i=1}^{k} < x - x', e_j >^2}. \tag{8}$$

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

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

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

$$\chi(\mathcal{S}) \subset \cup_{i=1}^{d} B_{eucl}(z_i, r), \quad \text{with } dr^k = C \quad \text{for some } C > 0. \tag{9}$$

For $i = 1, \ldots, d$*, let* x_i *be an element of* $\mathcal{H}$ *such that* $\chi(x_i) = z_i$*. The solution of the equation* $\chi(x) = z_i$ *is not unique in general, but just take* x_i *to be one of these solutions. Because of (8), we have that*

$$\chi^{-1}(B_{eucl}(z_i, r)) = B_k(x_i, r). \tag{10}$$

Finally, (9) and (10) are enough to show that the Kolmogorov's ε*-entropy of* $\mathcal{S}$ *is*

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

1.5 Description of the thesis

The first thematic of this thesis focuses on the study of quadratic error in statistical non-parametric 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 and mixing 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.

1.5.1 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 repones 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.

1.6 Short presentation of the results

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

1.6.1 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 statistical sample of pairs which are identically distributed like (X, Y), but not necessarily independent.

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

$B(x,h) = \{x' \in \mathcal{F}/d(x',x) < h\}$. Our nonparametric models will be quite general in the sense that we will just need the following simple assumption for the marginal distribution of X:

$$C_B^2(\mathcal{F}\times\mathbb{R}) = \left\{ \begin{array}{l} \varphi : \mathcal{F}\times\mathbb{R} \longrightarrow \mathbb{R} \\ (x,y) \longrightarrow \varphi(x,y) \qquad such \;\; as: \\ \forall z \in N_x, \varphi(z,.) \in C^2(N_y) \;\; and \; \left(\varphi(.,y), \dfrac{\partial^2\varphi(.,y)}{\partial y^2}\right) \in C_B^1(x)\times C_B^1(x), \end{array} \right\}$$

where $C_B^1(x)$ is the set of continuously differentiable functions in the meaning of Gateaux on N_x (see Troutman [154] for this type of differentiability), which the derivative operator of order 1 at point x is bounded on the unit ball $B(0,1)$ the functional space $\mathcal{F}$.

We then construct the conditional hazard function of Y knowing $X=x$ as follows:

$$\forall x \in \mathcal{F}, \quad \forall y \in \mathbb{R} \qquad h_Y^X(x,y) = \frac{f_Y^X(x,y)}{1-F_Y^X(x,y)} = \frac{f_Y^X(x,y)}{S_Y^X(x,y)}$$

The main objective is to study the the nonparametric estimate $\widehat{h}_Y^X(x,y) = \dfrac{\widehat{f}_Y^X(x,y)}{1-\widehat{F}_Y^X(x,y)}$ of $h_Y^X(x,y) = \dfrac{f_Y^X(x,y)}{1-F_Y^X(x,y)}$ when the explanatory variable X is valued in a space of eventually infinite dimension. We give precise asymptotic evaluations of the quadratic error of this estimator.

1.6.2 Nonparametric models

In the following x will be a fixed point in $\mathcal{F}$, N_x will denote a fixed neighborhood of x, S will be a fixed compact subset of $\mathbb{R}$.
Our nonparametric models will be quite general in the sense that we will just need.
together with some usual smoothness condition on the function to be estimated. According to the type of estimation problem to be considered, we will assume either
$\exists \tau < \infty, f_Y^X(x,y) \leq \tau, \forall (x,y) \in \mathcal{F}\times\mathcal{S}$,
$\exists \beta > 0, F_Y^X(x,y) \leq 1-\beta, \forall (x,y) \in \mathcal{F}\times\mathcal{S}$.

1.6.3 Results: i.i.d. Case

Theorem 1.6.1

$$\begin{aligned} \text{MSE}\, \widehat{h}_Y^X(x,y) &\equiv \mathbb{E}\left[(\widehat{h}_Y^X(x,y) - h_Y^X(x,y))\right]^2 \\ &\equiv B_n(x,y) + \frac{\sigma_h^2(x,y)}{nh_H\phi_x(h_K)} + o(h_H^2) + o(h_K) + o\left(\frac{1}{n\,h_H\phi_x(h_K)}\right) \end{aligned}$$

where

$$B_n(x,y) = \frac{(B_H^f(x,y) - h_Y^X(x,y)B_H^F(x,y))h_H^2 + (B_K^f(x,y) - h_Y^X(x,y)B_K^F(x,y))h_K}{1-F_Y^X(x,y)}$$

with

$$\begin{aligned}
B_H^f(x,y) &= \frac{1}{2}\frac{\partial^2 f^x(y)}{\partial y^2}\int t^2 H'(t)dt,\\
B_K^f(x,y) &= \frac{\int_{B(0,1)} K(\|v\|)D_x f_Y^X(x,y)[v]g(x,v)d\mu(v)}{\int_{B(0,1)} K(\|v\|)g(x,v)d\mu(v)}\\
B_H^F(x,y) &= \frac{1}{2}\frac{\partial^2 F_Y^X(x,y)}{\partial y^2}\int t^2 H(t)dt\\
B_K^F(x,y) &= \frac{\int_{B(0,1)} K(\|v\|)D_x F_Y^X(x,y)[v]g(x,v)d\mu(v)}{\int_{B(0,1)} K(\|v\|)g(x,v)d\mu(v)}.
\end{aligned}$$

and

$$\sigma_h^2(x,y) = \frac{\beta_2 h_Y^X(x,y)}{(\beta_1^2(1-F_Y^X(x,y))} \quad (\textit{with } \beta_j = \int_{B(0,1)} K^j(\|v\|)g(x,v)d\mu(v),\ \textit{for},\ j=1,\,2).$$

Theorem 1.6.2 *Assume some hypotheses, then we have for any $x \in \mathcal{A}$,*

$$\left(\frac{nh_H\phi_x(h_K)}{\sigma_h^2(x,y)}\right)^{1/2}\left(\widehat{h}_Y^X(x,y) - h_Y^X(x,y) - B_n(x,y)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad \textit{as} \quad n\to\infty.$$

where

$$\mathcal{A} = \{x\in\mathcal{F},\ f_Y^X(x,y)(1-F_Y^X(x,y)) \neq 0\}$$

and $\xrightarrow{\mathcal{D}}$ means the convergence in distribution.

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

1.6.4 Results: dependent Case

Theorem 1.6.3 ** Under assumptions, we have*

$$\mathbb{E}\left[\widehat{h'}^x(y) - h'^x(y)\right]^2 = B_n^2(x,y) + \frac{\sigma_{h'}^2(x,y)}{nh_H^3\phi_x(h_K)} + o(h_H^4 + h_K) + o\left(\frac{1}{n\,h_H^3\phi_x(h_K)}\right),$$

where

$$B_n(x,y) = \frac{(B_H^{f'} - h'^x(y)B_H^F)h_H^2 + (B_K^{f'} - h'^x(y)B_K^F)h_K}{1-F^x(y)}$$

with

$$\begin{aligned}
B_H^{f'}(x,y) &= \frac{1}{2}\frac{\partial^2 f^x(y)}{\partial y^2}\int t^2 H"(t)dt\\
B_K^{f'}(x,y) &= h_K\Phi_0'(0)\frac{\left(K(1) - \int_0^1 (sK'(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K"(s)\beta_x(s)ds\right)}\\
B_H^F(x,y) &= \frac{1}{2}\frac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H'(t)dt\\
B_K^F(x,y) &= h_K\Psi_0'(0)\frac{\left(K(1) - \int_0^1 (sK(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K'(s)\beta_x(s)ds\right)}.
\end{aligned}$$

and

$$\sigma^2_{h'}(x,y) = \frac{\beta_2 h^x(y)}{(\beta_1^2(1-F^x(y)))}\int (H''(t))^2 dt \;\; (\beta_j = K^j(1) - \int_0^1 (K^j)"(s)\beta_x(s)ds, \; for, \; j=1,\,2),$$

This part contains results on the asymptotic normality of $\widehat{h}^X(y)$ and $\widehat{h'}^X(y)$. Let us assume that h^X is sufficiently smooth (at least of class $\mathcal{C}^2$).

Theorem 1.6.4 *Under some hypotheses, then we have for any $x \in \mathcal{A}$,*

$$\left(\frac{nh_H^3\phi_x(h_K)}{\sigma^2_{h'}(x,y)}\right)^{1/2}\left(\widehat{h'}^x(y) - h'^x(y) - B_n(x,y)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad n\to\infty.$$

where

$$\mathcal{A} = \{x\in\mathcal{F},\; f^x(y)(1-F^x(y)) \neq 0\}$$

and $\xrightarrow{\mathcal{D}}$ means the convergence in distribution.

1.6.5 Results: functional single index

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

Let

$$\forall x\in\mathbb{R},\; F(\theta,x,Z) = \mathbb{P}(X\leq x| < Z,\theta >=< z,\theta >),$$

be the *cond-cdf* of X given $< Z,\theta >=< z,\theta >$, for $z\in\mathcal{H}$, which also shows the relationship between Z and X but is often unknown.
If this distribution is absolutely continuous with respect to the Lebesgues measure on $\mathbb{R}$, then we will denote by $f(\theta,\cdot,z)$ the conditional density of X given $< Z,\theta >=< z,\theta >$.

In the following, for any $z\in\mathcal{H}$ and $y\in\mathbb{R}$, let $\mathcal{N}_z$ be a fixed neighborhood of z in $\mathcal{H}$, $\mathcal{S}_\mathbb{R}$ will be a fixed compact subset of $\mathbb{R}$, and we will use the notation

$$\phi_{\theta,z}(h) = \mathbb{P}\left(Z\in B_\theta(z,h)\right) = \mathbb{P}\left(Z\in\{z'\in\mathcal{H}, 0 < | < z-z',\theta > |) < h\}\right).$$

In order to ensure the identifiability of model, We suppose that ;F is twice differentiable (w.r.t.) x and θ, such as $< \theta, e_1 >= 1$ e_1 Is the first vector Of the orthonormal basis of $\mathcal{H}$. Clearly, under this Condition, we have, for all $x\in\mathcal{H}$,

$$F_1\left(.| < \theta_1, x >\right) = F_2\left(.| < \theta_2, x >\right) \quad\Longrightarrow\quad F_1\equiv F_2 \text{ and } \theta_1=\theta_2.$$

Let $(Z_i, X_i)_{i=1,\dots,n}$ be the be the statistical sample of pairs which are identically distributed like (Z, X), but not necessarily independent.

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

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

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

with the convention $0/0=0$. Note that a similar estimate was already introduced in the case where X is a valued in some semi-metric space which can be of infinite dimension by Ferraty *et al.* [69].
We then construct the conditional hazard function of X knowing $<\theta,Z>=<\theta,z>$ as follows:

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

1.6.6 Case of non censored data

Let $(X_i,Z_i)_{1\leq i\leq n}$ be random variables, each of them follows the same law of a couple (X,Z) where X is valued in $\mathbb{R}$ and Z has values in the Hilbert space $(\mathcal{H},<\cdot;\cdot>)$. In this section we will suppose that X_i and Z_i are observed. From now, z denotes a fixed element of the functional space $\mathcal{H}$, $\mathcal{N}_z$ denotes a fixed neighborhood of z and $\mathcal{S}_{\mathbb{R}}$ is a fixed compact of $\mathbb{R}^+$.
The non-parametric model on the estimated function h^Z will be determined by the regularity conditions on the conditional distribution of X knowing Z, we have: $\forall(x_1,x_2)\in\mathcal{S}_{\mathbb{R}}^2$, $\forall(z_1,z_2)\in\mathcal{N}_z^2$

$$|F(\theta,x_1,z_1)-F(\theta,x_2,z_2)|\leq A_{\theta,z}\left(\|z_1,z_2\|^{b_1}+|x_1-x_2|^{b_2}\right),$$
$$|f(\theta,x_1,z_1)-f(\theta,x_2,z_2)|\leq A_{\theta,z}\left(\|z_1,z_2\|^{b_1}+|x_1-x_2|^{b_2}\right);\quad b_1>0,\ b_2>0.$$

Theorem 1.6.5 *we have:*

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

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

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

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

Theorem 1.6.6 *For any compact $\mathcal{S}_{\mathbb{R}}$, $\mathcal{S}_{\mathcal{H}}$ and $\Theta_{\mathcal{H}}$, we have:*

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

The demonstration of these results and the conditions will be given in detail in Chapter 3.

1.6.7 Results for censored data

Estimation of the hazard function when the data are censored is an important problem in medical research. So, in practice, in medical applications, it can be in the presence of variables censored. This problem is usually modeled by considering a positive variable called C, and the observed random variables are not the couples (X_i, Z_i) but only the (T_i, Δ_i, Z_i) where $T_i = \min(X_i,\ C_i)$ and $\Delta_i = I_{X_i \le C_i}$. In the following we use the notations $F_1(\theta,\cdot,Z)$ and $f_1(\theta,\cdot,Z)$ to describe the distribution function and conditional density of C knowing Z and we use the notation $S_1(\theta,\cdot,Z) = 1 - F_1(\theta,\cdot,Z)$.

The aim of this section, is to adapt these ideas as part of an explanatory variable Z functional, and build a kernel estimator function type of conditional random $h(\theta,\cdot,Z)$ adapted to the censored data. If we introduce the notation $L(\theta,\cdot,Z) = 1 - S_1(\theta,\cdot,Z)S(\theta,\cdot,Z)$ and $\varphi(\theta,\cdot,Z) = f(\theta,\cdot,Z)S_1(\theta,\cdot,Z)$, we can reformulate the expression for the complete data as follows:

$$h(\theta,t,Z) = \frac{\varphi(\theta,t,Z)}{1-L(\theta,t,Z)},\ \forall t,\ L(\theta,t,Z) < 1. \tag{11}$$

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

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

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

and

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

Finally the hazard function estimator is given as:

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

Theorem 1.6.7 *We have:*

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

Thereafter we propose to study the uniform almost complete convergence of our estimator defined above

Theorem 1.6.8 *We have*

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

Chapter 2

Real response and dependent condition

This chapter is the object of two works subjected for publication in:
Journal of Statistical Theory and Applications
And Journal of Acta Universitatis Apulensis.

Conditional risk estimate for functional data under strong mixing conditions

2.1 Introduction

We consider the problem of nonparametric estimation of the conditional hazard function for functional mixing data. More precisely, given a strictly stationary random variables $Z_i = (X_i, Y_i)_{i\in\mathbb{N}}$, we investigate a kernel estimate of the conditional hazard function of univariate response variable Y_i given the functional variable X_i . The principal aim of this chapter is to give the mean squared convergence rate and to prove the asymptotic normality of the proposed estimator.

2.2 The model

Consider $Z_i = (X_i, Y_i), i \in \mathbb{N}$ be a $\mathcal{F} \times \mathbb{R}$-valued measurable strictly stationary process, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, where $(\mathcal{F}, d)$ is a semi-metric space.
In the following x will be a fixed point in $\mathcal{F}$ and $\mathbb{N}_x$ will denote a fixed neighborhood of x. We assume that the regular version of the conditional probability of Y given X exists. Moreover, we suppose that, for all $z \in \mathbb{N}_x$ the conditional distribution function of Y given $X = z, F^z(.)$, is 3-times continuously differentiable and we denote by f^z its conditional density with respect to Lebesgue's measure over $\mathbb{R}$. In this chapter, we consider the problem of the nonparametric estimation of the conditional hazard function defined, for all $y \in \mathbb{R}$ such that $F^x(y) < 1$, by

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

In our spatial context, we estimate this function by

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

where

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

and

$$\widehat{f}^x(y) = \frac{h_H^{-1}\sum_{i=1}^n K(h_K^{-1}d(x,X_i))H'(h_H^{-1}(y-Y_i))}{\sum_{i=1}^n K(h_K^{-1}d(x,X_i))}, \forall y \in \mathbb{R}$$

We can write an estimator of the first derivative of the hazard function through the first derivative of the estimator.
It is therefore natural to try to construct an estimator of the derivative of the function h^X on the basis of these ideas. To estimate the conditional distribution function and the conditional density function in the presence of functional conditional random variable X.
The kernel estimator of the derivative of the function conditional random functional h'^X can therefore be constructed as follows:

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

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

$$\widehat{f}'^X(y) = \frac{\sum_{i=1}^n h_H^{-2}K(h_K^{-1}d(X,X_i))H''(h_H^{-1}(y-Y_i))}{\sum_{i=1}^n K(h_K^{-1}d(X,X_i))} \tag{2}$$

Later, we need assumptions on the parameters of the estimator, ie on K, H, H', h_H and h_K are little restrictive. Indeed, on one hand, they are not specific to the problem estimate of h^X (but inherent problems of F^X, f^X and f'^X estimation), and secondly they consist with the assumptions usually made under functional variables, with $K$is the kernel, H is a given continuously differentiable distribution function,$h_K = h_{K,n}$(resp. $h_H = h_{H,n}$) is a sequence of positive real numbers and H' is the derivative of H. Furthermore, the estimator $\hat{h'}^x(y)$ can we written as

$$\widehat{h'}^x(y) = \frac{\widehat{f'}^x_N(y)}{\widehat{F}^x_D - \widehat{F}^x_N(y)} + \left(\frac{\widehat{f}^x_N(y)}{\widehat{F}^x_D - \widehat{F}^x_N(y)}\right)^2 \tag{3}$$

where

$$\widehat{F}^x_D = \frac{1}{n\mathbb{E}[K_1]}\sum_{i=1}^n K(h_K^{-1}d(x,X_i)), K_1 = K(h_K^{-1}d(x,X_1))$$

$$\widehat{F}^x_N(y) = \frac{1}{n\mathbb{E}[K_1]}\sum_{i=1}^n K(h_K^{-1}d(x,X_i))H(h_H^{-1}(y-Y_i))$$

$$\widehat{f}^x_N(y) = \frac{1}{nh_H\mathbb{E}[K_1]}\sum_{i=1}^n K(h_K^{-1}d(x,X_i))H'(h_H^{-1}(y-Y_i))$$

$$\widehat{f'}^x_N(y) = \frac{1}{nh_H^2\mathbb{E}[K_1]}\sum_{i=1}^n K(h_K^{-1}d(x,X_i))H''(h_H^{-1}(y-Y_i)).$$

2.3 Notations and hypotheses

All along the chapter, when no confusion is possible, we will denote by C and C' some strictly positive generic constants. In order to establish our asymptotic results we need the following hypotheses, for all $r > 0$ and $i \in \mathbb{N}$:

$(H_0)\mathbb{P}(X \in B(x,r)) =: \phi_x(r) > 0, where B(x,r) = \{x' \in \mathcal{F}/d(x,x') < r\}.$

$(H_1)(X_i, Y_i)_i$ is α-mixing sequence whose the coefficients of mixture verify:

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

$(H_2) 0 < \sup_{i \neq j} \mathbb{P}((X_i, X_j) \in B(x,h) \times B(x,h)) = \mathcal{O}\left(\frac{(\phi_x(h))^{(a+1)/a}}{n^{1/a}}\right).$

Note that (H_0) can be interpreted as a concentration hypothesis acting on the distribution of the f.r.v. X, whereas (H_2) concerns the behavior of the joint distribution of the pairs (X_i, X_j). In fact, this hypothesis is equivalent to assume that, for n large enough

$$\sup_{i \neq j} \frac{\mathbb{P}((X_i, X_j) \in B(x,h) \times B(x,h))}{\mathbb{P}(X \in B(x,h))} \leq C \left(\frac{\phi_x(h)}{n}\right)^{1/a}.$$

(H_3) for $l \in \{0, 2\}$, the functions $\Psi_l(s) = \mathbb{E}[\frac{\partial^l F^X(y)}{\partial y^l} - \frac{\partial^l F^x(y)}{\partial y^l} | d(x,X) = s]$ and $\Phi_l(s) = \mathbb{E}[\frac{\partial^l f^X(y)}{\partial y^l} - \frac{\partial^l f^x(y)}{\partial y^l} | d(x,X) = s]$ are derivable at $s = 0$.

(H_4) The bandwidth h_K as $n \longrightarrow \infty$ satisfies:

$$h_K \downarrow 0, \forall t \in [0,1] \lim_{h_K \to 0} \frac{\phi_x(th_K)}{\phi_x(h_K)} = \beta_x(t) \text{ and } nh_H^3 \phi_x(h_K) \to \infty \ .$$

$(H5)$ The kernel K from $\mathbb{R}$ into $\mathbb{R}^+$ is a differentiable function supported on $[0, 1]$. Its derivative K' exists and is such that there exist two constants C and C' with $-\infty < C < K'(t) < C' < 0$ for $0 \leq t \leq 1$.

$(H6)$ H has even bounded derivative function supported on $[0, 1]$ that verifies

$$\int_{\mathbb{R}} t^2 H'(t) dt < \infty \quad and \quad \int_{\mathbb{R}} | t |^{b_2} (H^{(2)})^2(t) dt < \infty.$$

$(H7)$ There exist sequences of integers (u_n) and (v_n) increasing to infinity such that $(u_n + v_n) \leq n$, satisfying

$(i) v_n = o((nh_H^3\phi_x(h_K))^{1/2})$ and $(\frac{n}{h_H^3\phi_x(h_K)})^{1/2}\alpha(v_n) \longrightarrow 0 \quad as \quad n \longrightarrow 0,$
$(ii) q_n v_n = o((nh_H^3\phi_x(h_K))^{1/2})$ and $q_n(\frac{n}{h_H^3\phi_x(h_K)})^{1/2}\alpha(v_n) \longrightarrow 0 \quad as \quad n \longrightarrow \infty,$
where q_n is the largest integer such that $q_n(u_n + v_n) \leq n$.

2.3.1 Remarks on the assumptions

Remark 3.1. Assumption (H_0) plays an important role in our methodology. It is known as (for small h) the "concentration hypothesis acting on the distribution of X " in infinite-dimensional spaces. This assumption is not at all restrictive and overcomes the problem of the non-existence of the probability density function. In many examples, around zero the small ball probability $\phi_x(h)$ can be written approximately as the product of two independent functions $\Psi(x)$ and $\varphi(h)$ as $\phi_x(h) = \Psi(x)\varphi(h) + o(\varphi(h))$. This idea was adopted by Masry [114] who reformulated the Gasser et al. [78] one. The increasing propriety of $\phi_x(.)$ implies that $\xi_h^x(.)$ is bounded and then integrable (all the more so $\xi_0^x(.)$ is integrable).
Without the differentiability of $\phi_x(.)$, this assumption has been used by many authors where $\Psi(.)$ is interpreted as a probability density, while $\varphi(.)$ may be interpreted as a volume parameter. In the case of finite-dimensional spaces, that is $\mathcal{L} = \mathbb{R}^d$, it can be seen that $\phi_x(h) = C(d)h^d\Psi(x) + oh^d)$, 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. [60]):

$$(1)\phi_x(h) \approx \Psi(h)h^\gamma \quad \text{for some} \quad \gamma > 0.$$

$$(2)\phi_x(h) \approx \Psi(h)h^\gamma \exp\{C/h^p\} \quad \text{for some} \quad \gamma > 0 \quad and \quad p > 0.$$

$$(3)\phi_x(h) \approx \Psi(h)/ \mid \ln h \mid .$$

The function $\zeta_h^x(.)$ which intervenes in Assumption (H4) is increasing for all fixed h. Its pointwise limit $\zeta_0^x(.)$ 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 $\zeta_0(u) := \zeta_0^x(u)$ in the above examples by:

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

$$(2)\zeta_0(u)\delta_1(u) \quad where \quad \delta_1(.)\text{is Dirac function},$$

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

Assumption (H2) is classical and permits to make the variance term negligible.
Remark 3.2. Assumptions (H3) is a regularity condition which characterize the functional space of our model and is needed to evaluate the bias.

Remark 3.3. Assumptions (H5) and (H6) are classical in functional estimation for finite or infinite dimension spaces.

■ **First**

We purpose to study the L^2-consistency and the asymptotic normality of the nonparametric estimate $\hat{h'}^x$ of h'^x when the random filed $(Z_i, i \in \mathbb{N})$ satisfies the Previous mixing condition mentioned.

2.3.2 Mean squared convergence

The first result concerns the L^2-consistency of $\hat{h}'^x(y)$.

Theorem 3.1. Under assumptions (H0)-(H6), we have

$$\mathbb{E}\left[\hat{h}'^x(y) - h'^x(y)\right]^2 = B_n^2(x,y) + \frac{\sigma_{h'}^2(x,y)}{nh_H^3\phi_x(h_K)} + o\Big(h_H^4 + h_K\Big) + o\Big(\frac{1}{nh_H^3\phi_x(h_K)}\Big),$$

where

$$B_n(x,y) = \frac{(B_H^{f'} - h'^x(y)B_H^F)h_H^2 + (B_H^{f'} - h'^x(y)B_K^F)h_K}{1 - F^x(y)}$$

with

$$\begin{aligned}
B_H^{f'}(x,y) &= \tfrac{1}{2}\tfrac{\partial^2 f^x(y)}{\partial y^2}\int t^2 H''(t)dt \\
B_K^{f'}(x,y) &= h_K\Phi_0'(0)\tfrac{(K(1)-\int_0^1 (sK'(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 sK''(s)\beta_x(s)ds)} \\
B_H^F(x,y) &= \tfrac{1}{2}\tfrac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H'(t)dt \\
B_K^F(x,y) &= h_K\Psi_0'(0)\tfrac{(K(1)-\int_0^1 (sK(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 sK'(s)\beta_x(s)ds)}.
\end{aligned}$$

and

$$\sigma_{h'}^2(x,y) = \frac{\beta_2 h^x(y)}{(\beta_1^2(1 - F^x(y)))}\int (H''(t))^2 dt(\beta_j = K^j(1) - \int_0^1 (K^j)''(s)\beta_x(s)ds, for, j = 1,2),$$

Proof. By using the same decomposition used in (Theorem 3.1 Rabhi et al. [129], P.408), we show that the proof of Theorem 3.1 can be deduced from the following intermediates results:
Lemma 3.1. Under the hypotheses of Theorem 3.1, we have

$$\mathbb{E}[\hat{f}_N'^x(y)] - f'^x(y) = B_H^{f'}(x,y)h_H^2 + B_K^{f'}(x,y)h_K + o(h_H^4) + o(h_K)$$

and

$$\mathbb{E}[\hat{F}_N^x(y)] - F^x(y) = B_H^F(x,y)h_H^2 + B_K^F(x,y)h_K + o(h_H^2) + o(h_K).$$

Remark 3.4. Observe that, the result of this lemma permits to write

$$[\mathbb{E}\hat{F}_N^x(y) - F^x(y)] = \mathbb{O}(h_H^2 + h_K)$$

and

$$[\mathbb{E}\hat{f}'^{x}_{N}(y) - f'^{x}(y)] = \mathbb{O}(h_H^4 + h_K).$$

Lemma 3.2. Under the hypotheses of Theorem (3.1), we have

$$Var[\hat{f}'^{x}_{N}(y)] = \frac{\sigma^2_{f'}(x,y)}{nh_H^3\phi_x(h_K)} + o\left(\frac{1}{nh_H^3\phi_x(h_K)}\right),$$

$$Var[\hat{F}^{x}_{N}(y)] = o\left(\frac{1}{nh_H\phi_x(h_K)}\right)$$

and

$$Var[\hat{F}^{x}_{D}] = o\left(\frac{1}{nh_H\phi_x(h_K)}\right).$$

where $\sigma^2_{f'}(x,y) := f^x(y)\int(H''(t))^2dt$.

Lemma 3.3. Under the hypotheses of Theorem (3.1), we have

$$Cov(\hat{f}'^{x}_{N}(y), \hat{F}^{x}_{D}) = o\left(\frac{1}{nh_H^3\phi_x(h_K)}\right),$$

$$Cov(\hat{f}'^{x}_{N}(y), \hat{F}^{x}_{N}(y)) = o\left(\frac{1}{nh_H^3\phi_x(h_K)}\right)$$

and

$$Cov(\hat{F}^{x}_{D}, \hat{F}^{x}_{N}(y)) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right).$$

Remark 3.5. It is clear that, the results of Lemmas (3.2 and 3.3) allows to write

$$Var(\hat{F}^{x}_{D} - \hat{F}^{x}_{N}) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right)$$

2.3.3 Asymptotic normality

This section contains results on the asymptotic normality of $\hat{h}^X(y)$ and $\hat{h}'^X(y)$. Let us assume that h^X is sufficiently smooth (at least of class $\mathcal{C}^2$). We can write an estimator of the first derivative of the hazard function through the first derivative of the estimator. Later, we need assumptions on the parameters of the estimator, ie on K,H,H' h_H and h_K are little restrictive. Indeed, on one hand, they are not specific to the problem estimate of h^X (but inherent problems of F^X, f^X and f'^X estimation), and secondly they consist with the assumptions usually made under functional variables.

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

(H8) H' is twice differentiable.

(H9) The bandwidth h_H and h_K , small ball probability $\phi_z(h)$ and arithmetical α mixing coefficient with order $a > 3$ satisfying

$$\left\{\begin{array}{l}(H9a)\exists C>0, h_H^{2j+1}\phi_z(h_k)\geq \frac{c}{n^{2/(a+1)}}, for \quad j=0,1\\(H9b)\left(\frac{\phi_z(h_k)}{n}\right)^{1/a}+\phi_z(h_k)=0\left(\frac{1}{n^{2/(a+1)}}\right), for \quad j=0,1\\(H9c)\lim_{n\longrightarrow\infty}h_K=0, \lim_{n\longrightarrow\infty}h_H=0, and \lim_{n\longrightarrow\infty}\frac{\log n}{nh_H^{2j+1}\phi_x(h_K)}=0, j=0,1;\end{array}\right\}$$

Theorem 3.2. Assume that (H0)-(H9) hold, then we have for any $x \in \mathcal{A}$,

$$\left(\frac{nh_H^3\phi_x(h_K)}{\sigma_{h'}^2(x,y)}\right)^{1/2}\left(\hat{h}'^x(y)-h'^x(y)-B_n(x,y)\right)\xrightarrow{\mathbb{D}}\mathcal{N}(0,1) \quad as \quad n\longrightarrow\infty.$$

where

$$\mathcal{A}=\{x\in\mathcal{F}, f^x(y)(1-F^x(y))\neq 0\}$$

and $\xrightarrow{\mathbb{D}}$ means the convergence in distribution.
Obviously, if one imposes some additional assumptions on the function $\phi_x(.)$ and the bandwidth parameters (h_K and h_H) we can improved our asymptotic normality by removing the bias term $B_n(x,y)$.

Corollary 3.1.

Under the hypotheses of Theorem 3.2 and if the bandwidth parameters (h_K and h_H) and if the function $\phi_x(h_K)$ satisfies:

$$\lim_{n\to\infty}(h_H^4+h_K)\sqrt{n\phi_x(h_K)}=0$$

we have

$$\left(\frac{nh_H^3\phi_x(h_K)}{\sigma_{h'}^2(x,y)}\right)^{1/2}\left(\hat{h}'^x(y)-h'^x(y)\right)\xrightarrow{\mathbb{D}}\mathcal{N}(0,1) \quad as \quad n\longrightarrow\infty.$$

Proof of Theorem and Corollary. We consider the decomposition

$$\begin{aligned}\hat{h}^x(y)-h^x(y) &= \tfrac{1}{\hat{F}_D^x-\hat{F}_N^x(y)}[\hat{f}_N^x(y)-\mathbb{E}\hat{f}_N^x(y)]\\&+ \tfrac{h^x(y)}{\hat{F}_D^x-\hat{F}_N^x(y)}(\mathbb{E}\hat{F}_N^x(y)-F^x(y))\\&+ \tfrac{1}{\hat{F}_D^x-\hat{F}_N^x(y)}(\mathbb{E}\hat{f}_N^x(y)-f^x(y))\\&+ \tfrac{h^x(y)}{\hat{F}_D^x-\hat{F}_N^x(y)}(1-\mathbb{E}\hat{F}_N^x(y)-(\hat{F}_D^x-\hat{F}_N^x(y)))\end{aligned} \tag{4}$$

and let

$$\hat{h}'^X(y)=\frac{\hat{f}'^X(y)}{1-\hat{F}^Y(y)}+(\hat{h}^X(y))^2, \tag{5}$$

with

$$\hat{h}'^X(y) - h'^X(y) = \{(\hat{h}^X(y))^2 - (h^X(y))^2\} + \left\{\frac{\hat{f}'^X(y)}{1-\hat{F}^X(y)} - \frac{f'^X(y)}{1-F^X(y)}\right\} \quad (6)$$

for the first term of (6) we can write

$$\mid (\hat{h}^X(y))^2 - (h^X(y))^2 \mid \leq \mid \hat{h}^X(y) - h^X(y) \mid . \mid \hat{h}^X(y) + h^X(y) \mid \quad (7)$$

because the estimator $\hat{h}^X(.)$ converge a.co. to $h^X(.)$ we have

$$\sup_{y\in\mathcal{L}} \mid (\hat{h}^X(y))^2 - (h^X(y))^2 \mid \leq 2 \mid h^X(y) \mid \sup_{y\in\mathcal{L}} \mid \hat{h}^X(y) - h^X(y) \mid$$

$\mathcal{L}$ will be a fixed compact subset of $\mathbb{R}^+$, for the second term of (6) we have

$$\begin{aligned}\frac{\hat{f}'^X(y)}{1-\hat{F}^X(y)} - \frac{f'^X(y)}{1-F^X(y)} &= \frac{1}{(1-\hat{F}^X(y))(1-F^X(y))}\{\hat{f}'^X(y) - f'^X(y)\} \\ &+ \frac{1}{(1-\hat{F}^X(y))(1-F^X(y))}\{f'^X(y)(\hat{F}^X(y) - F^X(y))\} \\ &+ \frac{1}{(1-\hat{F}^X(y))(1-F^X(y))}\{F^X(y)(\hat{f}'^X(y) - f'^X(y))\}\end{aligned}$$

Therefore, Theorem 3.2 and corollary 3.1 are consequence of Lemma3.1, remark (3.4) and the following results.

Lemma3.4. Under the hypotheses of theorem 3.2

$$\left(\frac{nh_H^3\phi_x(h_K)}{\sigma_{f'}^2(x,y)}\right)^{1/2}\left(\hat{f}_N'^x(y) - \mathbb{E}[\hat{f}_N'^x(y)]\right) \longrightarrow \mathbb{N}(0,1).$$

Lemma3.5. Under Assumptions (H0)-(H6) and (H8), we have

$$(n\phi_x(h_K))^{1/2}(\hat{F}^X(y) - F^X(y)) \xrightarrow{\mathbb{D}} \mathcal{N}(0, \sigma_{F^X}^2(y)) \quad (8)$$

where

$$\sigma_{F^X}^2(y) = \frac{\beta_2 F^X(y)(1-F^X(y))}{\beta_1^2}$$

Lemma3.6. Under the hypotheses of Theorem 3.2

$$\hat{F}_D^x - \hat{F}_N^x(y) \to 1 - F^x(y) \quad in \quad probability$$

and

$$\left(\frac{nh_H\phi_x(h_K)}{\sigma_h^2(x,y)}\right)^{1/2}\left(\hat{F}_D^x - \hat{F}_N^x(y) - 1 + \mathbb{E}[\hat{F}_N^x(y)]\right) = Op(1).$$

Lemma3.7. Under Assumptions (H0)-(H7), we have

$$(nh_H\phi_x(h_K))^{1/2}(\hat{h}^X(y) - h^X(y)) \xrightarrow{\mathbb{D}} \mathcal{N}(0, \sigma_{h^X}^2(y)) \quad (9)$$

where

$$\sigma_{h^X}^2(y) = \frac{\beta_2 h^X(y)}{\beta_1^2(1-F^X(y))}\int_{\mathbb{R}}(H'(t))^2 dt$$

The proofs of Lemma 3.5 can be seen in Ezzahrioui and Ould-Saïd [56].

2.4 Applications

In this section we emphasize the potential impact of our work by studying its practical interest in some important statistical problems. Moreover, in order to show the easily implementation of our approach on a concrete cases, we discuss in the second part of this section the practical utilization of our model in risk analysis.

• On the choices of the bandwidths parameters: As all smoothing by a kernel method, the choice of bandwidths parameters has crucial role in determining the performance of the estimators. The mean quadratic error given in Theorem (3.1) is a basic ingredient to solve this problem. Usually, the ideal theoretical choices are obtained by minimizing this error. Here, we have explicated its leading term which is

$$B_n^2(x,y) + \frac{\sigma_{h'}^2(x,y)}{nh_H^3\phi_x(h_K)}.$$

Then, the smoothing parameters minimizing this leading term is asymptotically optimal with respect the L^2-error. However, the practical utilization of this criterium requires some additional computational efforts. More precisely, it requires the estimation of the unknown quantities $\Psi_0^{'}, \Phi_0^{'}, f^{'x}(y)$ and $F^x(y)$.Clearly, all these estimations can be obtained by using a pilots estimators of the conditional distribution function $F^x(y)$ and of the conditional density $f^{'x}(y)$. Such estimations are possible by using the kernel methods, with a separate choice of the bandwidths parameters between both models. More preciously, for the conditional density, we propose to adopt, to the functional case, the bandwidths selectors studied by Bouraine et al. [24] by considering the following criterion

$$CVPDF = \frac{1}{n}\sum_{i=1} W_1(X_i)\int \hat{f'}^{X_i^{-i^2}}(y)W_2(y)dy - \frac{2}{n}\sum_{i=1} \hat{f'}^{X_i^{-i}}(Y_i)W_1(X_i)W_2(Y_i) \qquad (10)$$

while, for the the conditional distribution function we can use the cross-validation rule proposed by De Gooijer and Gannoun [82] (in vectorial case)

$$CVCDF = \frac{1}{n}\sum_{k,l\in l_n} [I_{Y_k\le Y_l} - \hat{F}^{X_k^{-k}}(Y_l)]^2 W(X_k)$$

where W_1, W_2 and W are some suitable trimming functions and

$$\hat{F}^{X_k^{-k}}(Y_l) = \frac{\sum_{i\in I_{n,\varsigma n}^{k,l}} K(h_K^{-1}d(X_k,X_i))H(h_H^{-1}(Y_l-Y_i))}{\sum_{i\in I_{n,\varsigma n}^{k,l}} K(h_K^{-1}d(X_k,X_i))}$$

and

$$\widehat{f'}^{X_i^{-i}}(y) = \frac{h_H^{-2}\sum_{j\in I_{n,\varsigma n}^i} K(h_K^{-1}d(X_i,X_j))H''(h_H^{-1}(y-Y_j))}{\sum_{j\in I_{n,\varsigma n}^i}^n K(h_K^{-1}d(X_i,X_j))}$$

with

$$\begin{cases} I^{k,l}_{n,\varsigma_n} = \{\text{i such that} \quad |i-k| \geq \varsigma_n \quad \text{and} \quad |i-l| \geq \varsigma_n \\ \text{and} I^{i}_{n,\varsigma_n} = \{j \quad \text{such that} \quad |j-i| \geq \varsigma_n\}. \end{cases}$$

Of course, we can also adopt another selection methods, such that, the parametric bootstrap method, proposed by Hall et al. [83] and Hyndman et al. [99] for, respectively, the conditional cumulative distribution function and the conditional density in the finite dimensional case. Nevertheless, a data-driven method allows to overcome this additional computation is very important in practice and is one of the natural prospects of the present work.

• Confidence intervals: The main application of Theorem 3.2 is to build confidence band for the true value of $h'^{x}(y)$. Similarly to the previous application, the practical utilization of our result in this topic requires the estimation of the quantity $\sigma^2_{h'}(x,y)$. A plug-in estimate for the asymptotic standard deviation $\sigma^2_{h'}(x,y)$ can be obtained by using the estimators $\widehat{f'}^{x}(y)$ and $\widehat{F}^{x}(y)$ of $f'^{x}(y)$ and$F'^{x}(y)$. Then we get

$$\widehat{\sigma}^2_{h'}(x,y) := \frac{\widehat{\beta}_2 \widehat{f}'^{x}(y)}{\left(\widehat{\beta}^2_1(1-\widehat{F}^{x}(y))^2\right)}$$

where

$$\begin{cases} \widehat{\beta}_1 = \frac{1}{n\phi_x(h_k)} \sum_{i=1}^{n} K(h_k^{-1} d(x,X_i)) \\ \text{and} \quad \widehat{\beta}_2 = \frac{1}{n\phi_x(h_k)} \sum_{i=1}^{n} K^2(h_k^{-1} d(x,X_i)) \end{cases}$$

Clearly, the function $\phi_x(.)$ does not appear in the calculation of the confidence interval by simplification. More precisely, we obtain the following approximate $(1-\zeta)$ confidence band for $h'^{x}(y)$

$$\widehat{h}'^{x}(y) \pm t_{1-\zeta/2} \times \left(\frac{\widehat{\sigma}^2_{h'}(x,y)}{nh_h^3 \phi_x(h_k)}\right)^{1/2}$$

where $t_{1-\zeta/2}$ denotes the $1-\zeta/2$ quantile of the standard normal distribution.

2.5 Appendix

In the following, we will denote $\forall i$

$$K_i = K(h_H^{-1} d(x,X_i)), \quad H_i = H(h_H^{-1}(y-Y_i)) \quad \textit{and} \quad H''_i = H''(h_H^{-1}(y-Y_i)).$$

Proof of Lemma 3.1. Firstly, for $\mathbb{E}[\widehat{f}'^{x}_N(y)]$, we start by writing

$$\mathbb{E}[\widehat{f}'^{x}_N(y)] = \frac{\mathbb{E}\left[K_1 \mathbb{E}[h_H^{-2} H''_1|X]\right]}{\mathbb{E}[K_1]} \quad \textit{with} \quad h_H^{-2}\mathbb{E}[H''_1|X] = \int_{\mathbb{R}} H''(t) f^X(y - h_H t) dt.$$

The latter can be re-written, by using a Taylor expansion under (H3), as follows

$$h_H^{-2}\mathbb{E}[H_1''|X] = f^X(y) + \frac{h_H^2}{2}\bigg(\int t^2 H''(t)dt\bigg)\frac{\partial^2 f^X(y)}{\partial^2 y} + o(h_H^2).$$

Thus, we get

$$\mathbb{E}[\widehat{f}_N'^{x}(y)] = \frac{1}{\mathbb{E}[K_1]}\bigg(\mathbb{E}\left[\frac{h_H^2}{2}K_1\frac{\partial^2 f^X(y)}{\partial^2 y}\right]\int t^2 H''(t)dt\bigg) + \frac{1}{\mathbb{E}[K_1]}(\mathbb{E}[K_1 f^X(y)] + o(h_H^2)).$$

Let $\Psi_l(.,y) := \frac{\partial^l f(y)}{\partial^l y}$: for $l \in \{0.2\}$, since $\Phi_l(0) = 0$, we have

$$\begin{aligned}
\mathbb{E}[K_1\Psi_l(X,y)] &= \Psi_l(x,y)\mathbb{E}[K_1] + \mathbb{E}[K_1(\Psi_l(X,y) - \Psi_l(x,y))] \\
&= \Psi_l(x,y)\mathbb{E}[K_1] + \mathbb{E}[K_1(\Phi_l(d(x,X))] \\
&= \Psi_l(x,y)\mathbb{E}[K_1] + \Phi_l'(0)\mathbb{E}[d(x,X)K_1] + o(\mathbb{E}[d(x,X)K_1]).
\end{aligned}$$

So,

$$\begin{aligned}
\mathbb{E}[\widehat{f'}_N^x(y)] &= f^x(y) + \tfrac{h_H^2}{2}\tfrac{\partial^2 f^x(y)}{\partial y^2}\textstyle\int t^2 H''(t)dt + o\bigg(h_H^2\tfrac{\mathbb{E}[d(x,X)K_1]}{\mathbb{E}[K_1]}\bigg) \\
&+ \Phi_0'(0)\tfrac{\mathbb{E}[d(x,X)K_1]}{\mathbb{E}[K_1]} + o\bigg(\tfrac{\mathbb{E}[d(x,X)K_1]}{\mathbb{E}[K_1]}\bigg).
\end{aligned}$$

Similarly to Ferraty et al. [60] we show that

$$\frac{1}{\phi_x(h_K)}\mathbb{E}[d(x,X)K_1] = h_K\bigg(K(1) - \int_0^1 (sK(s))'\beta_x(s)ds + o(1)\bigg)$$

and

$$\frac{1}{\phi_x(h_K)}\mathbb{E}[K_1] = K(1) - \int_0^1 K'(s)\beta_x(s)ds + o(1).$$

Hence,

$$\begin{aligned}
\mathbb{E}[\widehat{f'}_N^x(y)] &= f^x(y) + \tfrac{h_H^2}{2}\tfrac{\partial^2 f^x(y)}{\partial y^2}\textstyle\int t^2 H''(t)dt \\
&+ h_K\Phi_0'(0)\tfrac{(K(1)-\int_0^1(sK(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)} + o(h_H^2) + o(h_K).
\end{aligned}$$

Secondly, concerning $\mathbb{E}[\widehat{F}_N^x(y)]$, we write by an integration by part

$$\mathbb{E}[\widehat{F}_N^x(y)] = \frac{1}{\mathbb{E}[K_1]}\mathbb{E}[K_1\mathbb{E}[H_1|X]] \quad with \quad \mathbb{E}[H_1|X] = \int_{\mathbb{R}} H'(t)F^X(y - h_H t)dt.$$

The same steps used to studying $\mathbb{E}[\widehat{f}_N^x(y)]$ can be followed to prove that

$$\begin{aligned}\mathbb{E}[\widehat{F}_N^x(y)] &= F^x(y) + \tfrac{h_H^2}{2}\tfrac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H'(t)dt \\ &+ h_K \Psi_0'(0)\tfrac{(K(1)-\int_0^1 (sK(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)} + o(h_H^2) + o(h_K).\end{aligned}$$

Proof of Lemma 3.2. For the first quantity $Var[\widehat{f}_N^x(y)]$, we have

$$s_n^2 = Var[\widehat{f'}_N^x(y)] = \frac{1}{(nh_H^2\mathbb{E}[K_1(x)])^2} Var\left[\sum_{i=1}\Gamma_i(x)\right]$$

where

$$\Gamma_i(x) = K_i(x)H_i''(y) - \mathbb{E}[K_i(x)H_i''(y)].$$

Thus

$$\begin{aligned}Var[\widehat{f}_N^x(y)] &= \tfrac{1}{(nh_H^2\mathbb{E}[K_1])^2}\underbrace{\sum_{i\neq j} cov(\Gamma_i(x),\Gamma_j(x))}_{s_n^{cov}} + \underbrace{\sum_{i=1}^n Var(\Gamma_i(x))}_{s_n^{var}} \\ &= \tfrac{Var[\Gamma_1]}{n(h_H^2\mathbb{E}[K_1])^2} + \tfrac{1}{(nh_H^2\mathbb{E}[K_1])^2}\sum_{i\neq j} Cov(\Gamma_i,\Gamma_j).\end{aligned}$$

Let us calculate the quantity $Var[\Gamma_1(x)]$. We have:

$$\begin{aligned}Var[\Gamma_1(x)] &= \mathbb{E}[K_1^2(x)H_1''^2(y)] - \Big(\mathbb{E}[K_1(x)H_1''(y)]\Big)^2 \\ &= \mathbb{E}[K_1^2(x)]\tfrac{\mathbb{E}[K_1^2(x)H_1''^2(y)]}{\mathbb{E}[K_1^2(x)]} \\ &- (\mathbb{E}[K_1(x)])^2\left(\tfrac{\mathbb{E}[K_1(x)H_1''(y)]}{\mathbb{E}[K_1(x)]}\right)^2.\end{aligned}$$

So, by using the same arguments as those used in pervious lemma we get

$$\begin{aligned}\tfrac{1}{\phi_x(h_K)}\mathbb{E}[K_1^2(x)] &= K^2(1) - \int_0^1 (K^2(s))'\beta_x(s)ds + o(1) \\ \tfrac{\mathbb{E}[K_1^2(x)H_1''^2(y)]}{\mathbb{E}[K_1^2(x)]} &= h_H^2 f^x(y)\int H''^2(t)dt + o(h_H^2) \\ \tfrac{\mathbb{E}[K_1(x)H_1''(y)]}{\mathbb{E}[K_1(x)]} &= h_H^2 f^x(y) + o(h_H^2)\end{aligned}$$

which implies that

$$Var[\Gamma_i(x)] = h_H^2\phi_x(h_K)f^x(y)\int H''^2(t)dt\left(K^2(1) - \int_0^1 (K^2(s))'\beta_x(s)ds\right) + o(h_H^2\phi_x(h_K)). \quad (11)$$

Now, let us focus on the covariance term. To do that, we need to calculate the asymptotic behavior of quantity defined as

$$\sum_{i\neq j} |cov(\Gamma_i(x), \Gamma_j(x))| = \sum_{1\leq |i-j|\leq c_n} |cov(\Gamma_i(x), \Gamma_j(x))| = J_{1,n} + J_{2,n}.$$

with $c_n \to \infty$, as $n \to \infty$.
for all (i, j) we write

$$cov(\Gamma_i(x), \Gamma_j(x)) = \mathbb{E}[K_i(x)K_j(x)H_i''(y)H_j''(y)] - (\mathbb{E}[K_i(x)H_i''(y)])^2$$

and we use the fact that

$$\mathbb{E}[H_i''(y)H_j''(y)|(X_i, X_j)] = \mathcal{O}(h_H^4); \forall i \neq j, \mathbb{E}[H_i''(y)|X_i] = \mathcal{O}(h_H^2); \forall i.$$

For $J_{1,n}$: by means of the integral realized above and under (H2) and (H5), we get

$$\mathbb{E}[K_iK_jH_i''H_j''] \leq Ch_H^4\mathbb{P}[(X_i, X_j) \in B(x, h_K) \times B(x, h_K)]$$

and

$$\mathbb{E}[K_i(x)H_i''(y)] \leq Ch_H^2\mathbb{P}(X_i \in B(x, h_K)).$$

It follows that, the hypothesis (H0), (H2) and (H5), imply that

$$cov(\Gamma_i(x), \Gamma_j(x)) \leq Ch_H^2\phi_x(h_K)\left(\phi_x(h_K) + \left(\frac{\phi_x(h_K)}{n}\right)^{1/a}\right)$$

So

$$J_{1,n} \leq C\left(nc_nh_H^4\left(\frac{\phi_x(h_K)}{n}\right)^{1/a}\phi_x(h_K)\right).$$

Hence

$$J_{1,n} = \mathcal{O}\left(nc_nh_H^4\left(\frac{\phi_x(h_K)}{n}\right)^{1/a}\phi_x(h_K)\right).$$

On the other hand, these covariances can be controlled by mean of the usual Davydov-Rios's covariance inequality for mixing processes (see Rio [140], formula 1.12a). Together with (H1), this inequality leads to:

$$\forall i \neq j, \quad |Cov(D_i(x), D_j(x))| \leq C|i-j|^{-a}.$$

By the fact, $\sum_{k\geq c_n+1} k^{-a} \leq \int_{C_n}^{\infty} t^{-a}dt = \frac{c_n^{-a+1}}{a-1}$, we get by applying (H1),

$$J_{2,n} \leq \sum_{|i-j|\geq c_n+1} |i-j|^{-a} \leq \frac{nc_n^{-a+1}}{a-1}$$

Thus, by using the following classical technique (see Bosq [20]), we can write

$$s_n^{cov} = \sum_{0<|i-j|\leq u_n} |Cov(\Gamma_i(x), \Gamma_j(x))| + \sum_{|i-j|>u_n} |Cov(\Gamma_i(x), \Gamma_j(x))|.$$

Thus

$$s_n^{cov} \leq C_n\left(c_n h_H^4 \left(\frac{\phi_x(h_K)}{n}\right)^{1/a} \phi_x(h_K) + \frac{c_n^{-a+1}}{a-1}\right)$$

Choosing $c_n = h_H^{-4}\left(\frac{\phi_x(h_K)}{n}\right)^{-1/a}$, and owing to the right inequality in (H7(ii)), we can deduce

$$s_n^{cov} = o(nh_H^2\phi_x(h_K)). \tag{12}$$

Finally,

$$\begin{aligned} s_n^2 &= o(nh_H^2\phi_x(h_K)) + \mathcal{O}(nh_H^2\phi_x(h_K)) \\ &= \mathcal{O}(nh_H^2\phi_x(h_K)) \end{aligned}$$

In conclusion, we have

$$\begin{aligned} Var[\widehat{f'}^x_N(y)] &= \frac{f^x(y)\int H''^2(t)dt}{nh_H^4\phi_x(h_K)}\left(\frac{(K^2(1)-\int_0^1(K^2(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)^2}\right) \\ &+ o\left(\frac{1}{nh_H^2\phi_x(h_K)}\right) \end{aligned} \tag{13}$$

Now, for $\widehat{F}_N^x(y)$, (resp. $\widehat{F}_D^x$) we replace $H_i''(y)$ by $H_i(y)$ (resp. by 1) and we follow the same ideas, under the fact that $H \leq 1$

$$\begin{aligned} Var[\widehat{F}_N^x(y)] &= \frac{F^x(y)}{n\phi_x(h_K)}\left(\int H'^2(t)dt\right)\left(\frac{(K^2(1)-\int_0^1(K^2(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)^2}\right) \\ &+ o\left(\frac{1}{n\phi_x(h_K)}\right). \end{aligned}$$

and

$$Var[\widehat{F}_D^x] = \frac{1}{n\phi_x(h_K)}\left(\frac{(K^2(1)-\int_0^1(K^2(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)^2}\right) + o\left(\frac{1}{n\phi_x(h_K)}\right).$$

This yields the proof.

Proof of Lemma 3.3. The proof of this lemma follows the same steps as the previous Lemma. For this, we keep the same notation and we write

$$
\begin{aligned}
Cov(\widehat{f'}^x_N(y), \widehat{F}^x_N(y)) &= \tfrac{1}{nh_H^2(\mathbb{E}[K_1(x)])^2} Cov(\Gamma_1(x), \Delta_1(x)) \\
&+ \tfrac{1}{n^2h_H^2(\mathbb{E}[K_1(x)])^2} \sum_{i \neq j} Cov(\Gamma_i(x), \Delta_j(x))
\end{aligned}
$$

Where

$$\Delta_i(x) = K_i(x)H_i(y) - \mathbb{E}[K_i(x)H_i(y)].$$

For the first term, we have under (H4)

$$
\begin{aligned}
Cov(\Gamma_1(x), \Delta_1(x)) &= \mathbb{E}[K_1^2(x)H_1(y)H_1''(y)] - \mathbb{E}[K_1(x)H_1(y)]\mathbb{E}[K_1(x)H_1''(y)] \\
&= \mathcal{O}(h_H^2\phi_x((h_k)) + \mathcal{O}(h_H^2\phi_x^2((h_k)) \\
&= \mathcal{O}(h_H^2\phi_x((h_k))
\end{aligned}
$$

Therefore,

$$\frac{1}{nh_H^2(\mathbb{E}[K_1(x)])^2} Cov(\Gamma_1(x), \Delta_1(x)) = \mathcal{O}\left(\frac{1}{n\phi_x(h_K)}\right) = o\left(\frac{1}{nh_H^2\phi_x(h_K)}\right) \tag{14}$$

So, by using similar arguments as those invoked in the proof of Lemma 3.2, and we use once again the boundedness of K and H, and the fact that (H1) and (H6) imply that

$$\mathbb{E}(H_i''(y)|X_i) = \mathcal{O}(h_H^2).$$

Moreover, the right part of (H7(ii)) implies that

$$Cov(\Gamma_i(x), \Delta_j(x)) = \mathcal{O}\left(h_H^2\phi_x(h_K)\left(\frac{\phi_x(h_K)}{n}\right)^{1/a} + \phi_x(h_K)\right),$$

Meanwhile, using the Davydov-Rio's inequality in Rio [140] for mixing processes leads to

$$|Cov(\Gamma_i(x), \Delta_j(x))| \leq C\alpha(|i-j|) \leq C|i-j|^{-a},$$

we deduce easily that for any $c_n > 0$:

$$\sum_{i \neq j} Cov(\Gamma_i(x), \Delta_j(x)) = \mathcal{O}\left(nc_nh_H^2\phi_x(h_K)\left(\frac{\phi_x(h_K)}{n}\right)^{1/a} + \phi_x(h_K)\right) + \mathcal{O}(nh_H^2c_n^{-a}).$$

It suffices now to take $c_n = h_H^{-2}\left(\frac{\phi_x(h_K)}{n}\right)^{1/a}$ to get the following expression for the sum of the covariances:

$$\sum_{i \neq j} Cov(\Gamma_i(x), \Delta_j(x)) = o(n\phi_x(h_K)). \tag{15}$$

From (14) and (15) we deduce that

$$Cov(\widehat{f'}_N^x(y), \widehat{F}_N^x(y)) = o\left(\frac{1}{nh_H^2\phi_x(h_K)}\right)$$

The same arguments can be used to shows that

$$Cov(\widehat{f'}_N^x(y), \widehat{F}_D^x) = o\left(\frac{1}{nh_H^2\phi_x(h_K)}\right)$$

and

$$Cov(\widehat{F}_N^x(y), \widehat{F}_D^x) = o\left(\frac{1}{n\phi_x(h_K)}\right).$$

Proof of Lemma 3.4. Let

$$S_n = \sum_{i=1}^{n} \Lambda_i(x)$$

Where

$$\Lambda_i(x) := \frac{\sqrt{h_H\phi_x(h_K)}}{h_H\mathbb{E}[K_1(x)]}\Gamma_i(x). \tag{16}$$

Obviously, we have

$$\sqrt{nh_H^3\phi_x(h_K)}[\sigma_{f'}(x,y)]^{-1}(\widehat{f'}_N^x(y) - \mathbb{E}\widehat{f'}_N^x(y)) = (n(\sigma_{f'}(x,y))^2)^{-1/2}S_n.$$

Thus, the asymptotic normality of $(n(\sigma_{f'}(x,y))^2)^{-1/2}S_n$, is sufficient to show the proof of this Lemma. This last is shown by the blocking method, where the random variables Λ_i are grouped into blocks of different sizes defined.
We consider the classical big- and small-block decomposition. We split the set {1,2,...,n} into $2k_n + 1$ subsets with large blocks of size u_n and small blocks of size v_n and put

$$k_n := \left[\frac{n}{u_n + v_n}\right].$$

Assumption (H7)(ii) allows us to define the large block size by

$$u_n := \left[\left(\frac{nh_H^3\phi_x(h_K)}{q_n}\right)^{1/2}\right].$$

Using Assumption (H7) and simple algebra allows us to prove that

$$\frac{v_n}{u_n} \to 0, \quad \frac{u_n}{n} \to 0, \quad \frac{u_n}{\sqrt{nh_H^3\phi_x(h_K)}} \to 0, \quad and \quad \frac{n}{u_n}\alpha(v_n) \to 0 \tag{17}$$

Now, let Υ_j,Υ'_j and Υ''_j be defined as follows:

$$\Upsilon_j = \sum_{i=j(u+v)+1}^{j(u+v)+u} \Lambda_i(x), 0 \leq j \leq k+1$$

$$\Upsilon'_j = \sum_{i=j(u+v)+u+1}^{(j+1)(u+v)+u} \Lambda_i(x), 0 \leq j \leq k+1$$

$$\Upsilon''_j = \sum_{i=k(u+v)+1}^{n} \Lambda_i(x), 0 \leq j \leq k+1$$

Clearly, we can write

$$S_n := \sum_{j=0}^{k-1} \Upsilon_j + \sum_{j=0}^{k-1} \Upsilon'_j + \Upsilon''_k r =: S'_n + S''_n + S'''_n.$$

We prove that

$$(i)\frac{1}{n}\mathbb{E}(S''_n)^2 \to 0, \quad (ii)\frac{1}{n}\mathbb{E}(S'''_n)^2 \to 0, \tag{18}$$

$$|\mathbb{E}\{exp(itn^{-1/2}S'_n)\} - \prod_{j=0}^{k-1}\mathbb{E}\{exp(itn^{-1/2}\Upsilon_j)\}| \to 0, \tag{19}$$

$$\frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}(\Upsilon_j^2) \longrightarrow \sigma^2_{f'(x,y)}, \tag{20}$$

$$\frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}\left(\Upsilon_j^2 \mathbb{I}_{\{|\Upsilon_j|>\varepsilon\sqrt{n\sigma^2_{f'}(x,y)}\}}\right) \longrightarrow 0, \tag{21}$$

for every $\varepsilon > 0$.
Expression (18) show that the terms S''_n and S'''_n are negligible, while Equations (19) and (20) show that the Υ_j are asymptotically independent, verifying that the sum of their variances tends to $\sigma^2_{f'}(x,y)$. Expression (21) is the Lindeberg-Feller's condition for a sum of independent terms. Asymptotic normality of S_n is a consequence of Equations (18)-(21).

- **Proof of (18)** Because $\mathbb{E}(\Lambda_j) = 0, \forall j$, we have that

$$\mathbb{E}(S''_n)^2 = Var\left(\sum_{j=0}^{k-1}\Upsilon'_j\right) = \sum_{j=0}^{k-1} Var(\Upsilon'_j) + \sum_{0\leq i<j\leq k-1} Cov(\Upsilon'_i, \Upsilon'_j) := \Pi_1 + \Pi_2.$$

By the second-order stationarity we get

$$\begin{aligned} Var(\Upsilon'_j) &= Var\Bigg(\sum_{i=j(u_n+v_n)+u_n+1}^{(j+1)(u_n+v_n)} \Lambda_i(x)\Bigg) \\ &= v_n Var(\Lambda_1(x)) + \sum_{i\neq j}^{v_n} Cov(\Lambda_i(x), \Lambda_j(x)). \end{aligned}$$

Then

$$\begin{aligned} \tfrac{\Pi_1}{n} &= \tfrac{kv_n}{n} Var(\Lambda_1(x)) + \tfrac{1}{n}\sum_{j=0}^{k-1}\sum_{i\neq j}^{v_n} Cov(\Lambda_i(x), \Lambda_j(x)) \\ &\leq \tfrac{kv_n}{n}\left\{\tfrac{\phi_x(h_k)}{h_H\mathbb{E}^2K_1(x)} Var(\Gamma_1(x))\right\} + \tfrac{1}{n}\sum_{i\neq j}^{n} |Cov(\Lambda_i(x), \Lambda_j(x))| \\ &\leq \tfrac{kv_n}{n}\left\{\tfrac{1}{h_H\phi_x(h_k)} Var(\Lambda_1(x))\right\} + \tfrac{1}{n}\sum_{i\neq j}^{n} |Cov(\Lambda_i(x), \Lambda_j(x))| \end{aligned}$$

Simple algebra gives us

$$\frac{kv_n}{n} \cong \left(\frac{n}{u_n+v_n}\right)\frac{v_n}{n} \cong \frac{v_n}{u_n+v_n} \cong \frac{v_n}{u_n} \longrightarrow 0 \quad as \quad n \longrightarrow \infty.$$

Using Equation (12) we have

$$\lim_{n\longrightarrow\infty} \frac{\Pi_1}{n} = 0 \tag{22}$$

Now, let us turn to Π_2/n. We have

$$\begin{aligned} \tfrac{\Pi_2}{n} &= \tfrac{1}{n}\sum_{i=0_{i\neq j}}^{k-1}\sum_{j=0}^{k-1} Cov(\Upsilon_i(x), \Upsilon_j(x)) \\ &= \tfrac{1}{n}\sum_{i=0_{i\neq j}}^{k-1}\sum_{j=0}^{k-1}\sum_{l_1=1}^{v_n}\sum_{l_2}^{v_n} Cov(\Lambda_{m_j+l_1}, \Lambda_{m_j+l_2}) \end{aligned}$$

with $m_i = i(u_n+v_n)+v_n$. As $i \neq j$, we have $|m_i - m_j + l_1 - l_2| \geq u_n$. It follows that

$$\frac{\Pi_2}{n} \leq \frac{1}{n}\sum_{i=1_{|i-j|\geq u_n}}^{n}\sum_{j=1}^{n} Cov(\Lambda_i(x), \Lambda_j(x)),$$

Then

$$\lim_{n\longrightarrow\infty} \frac{\Pi_2}{n} = 0. \tag{23}$$

By Equations (22) and (23) we get Part(i) of the Equation(18).
We turn to (ii), we have

$$\begin{aligned} \tfrac{1}{n}\mathbb{E}(S''_n)^2 &= \tfrac{1}{n}Var(\Upsilon''_K) \\ &= \tfrac{\mathcal{V}_n}{n}Var(\Lambda_1(x)) + \tfrac{1}{n}\sum_{i=1_{i\neq j}}^{\mathcal{V}_n}\sum_{j=1}^{\mathcal{V}_n} Cov(\Lambda_i(x),\Lambda_j(x)) \end{aligned}$$

where $\mathcal{V}_n = n - k_n(u_n + v_n)$; by the definition of k_n, we have $\mathcal{V}_n \leq u_n + v_n$.
Then

$$\frac{1}{n}\mathbb{E}(S''_n)^2 \leq \frac{u_n + v_n}{n}Var(\Lambda_1(x)) + \frac{1}{n}\sum_{i=1_{i\neq j}}^{\mathcal{V}_n}\sum_{j=1}^{\mathcal{V}_n} Cov(\Lambda_i(x),\Lambda_j(x))$$

and by the definition of u_n and v_n we achieve the proof of (ii) of Equation (18).

• **Proof of (19)** We make use of Volkonskii and Rozanov's lemma (see the appendix in Masry, [114]) and the fact that the process (X_i, X_j) is strong mixing.
Note that $\Upsilon_a is \mathcal{F}_{i_a}^{j_a}$-mesurable with $i_a = a(u_n + v_n) + 1$ and $j_a = a(u_n + v_n) + u_n$; hence, with $V_j = \exp(itn^{-1/2}\Upsilon_j)$ we have

$$\begin{aligned} |\mathbb{E}\left\{\exp(itn^{-1/2}S'_n)\right\} - \prod_{j=0}^{k-1}\mathbb{E}\left\{\exp(itn^{-1/2}\Upsilon_j)\right\}| &\leq 16k_n\alpha(v_n+1) \\ &\cong \tfrac{n}{v_n}\alpha(v_n+1) \end{aligned}$$

which goes to zero by the last part of Equation (17). Now we establish Equation (20).

• **Proof of (20)** Note that $Var(S'_n) \to \sigma^2_{f'}(x,y)$ by equation (18) and since $Var(S'_n) \to \sigma^2_{f'}(x,y)$ (by the definition of the Λ_i and Equation (13)). Then because

$$\mathbb{E}(S'_n)^2 = Var(S'_n) = \sum_{j=0}^{k-1} Var(\Upsilon_j) + \sum_{i=0_{i\neq j}}^{k-1}\sum_{j=0}^{k-1} Cov(\Upsilon_i,\Upsilon_j),$$

all we have to prove is that the double sum of covariances in the last equation tends to zero. Using the same arguments as those previously used for Π_2 in the proof of first term of Equation (18)we obtain by replacing v_n by u_n we get

$$\frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}(\Upsilon_j^2 5) = \frac{ku_n}{n}Var(\Lambda_1) + o(1).$$

As $Var(\Lambda_1) \longrightarrow \sigma^2_{f'}(x,y)$ and $k_{un}/n \longrightarrow 1$, we get the result.
Finally, we prove Equation (21).

• **Proof of (21)** Recall that
$$\Upsilon_j = \sum_{i=j(u_n+v_n)+1}^{j(u_n+v_n)+u_n} \Lambda_i.$$
Making use Assumptions (H5) and (H6), we have
$$|\Lambda_i| \leq C(h_H^2\phi_x(h_k))^{-1/2}$$
thus
$$|\Upsilon_j| \leq Cu_n(h_H^2\phi_x(h_k))^{-1/2},$$
which goes to zero as n goes to infinity by Equation (17). Then for n large enough, the set $\{|\Upsilon_j| > \varepsilon(n\sigma_{f'}^2(x,y))^{-1/2}\}$ becomes empty, this completes the proof and therefore that of the asymptotic normality of $(n(\sigma_{f'}(x,y))^2)^{-1/2}S_n$,

• **Proof of Lemmas 3.6.** It is clear that, the result of Lemma 3.1 and Lemma 3.2 permits us
$$\mathbb{E}(\widehat{F}_D^x - \widehat{F}_N^x - 1 + F^x(y)) \longrightarrow 0$$
and
$$Var(\widehat{F}_D^x - \widehat{F}_N^x - 1 + F^x(y)) \longrightarrow 0$$
then
$$\widehat{F}_D^x - \widehat{F}_N^x - 1 + F^x(y) \stackrel{\mathbb{P}}{\longrightarrow} 0$$
Moreover, the asymptotic variance of $\widehat{F}_D^x - \widehat{F}_N^x$ given in remark (3.5) allows to obtain
$$\frac{nh_H\phi_x(h_K)}{\sigma_h(x,y)^2} Var(\widehat{F}_D^x - \widehat{F}_N^x - 1 + \mathbb{E}(\widehat{F}_N^x(y))) \longrightarrow 0.$$
By combining result with the fact that
$$\mathbb{E}(\widehat{F}_D^x - \widehat{F}_N^x - 1 + \mathbb{E}(\widehat{F}_N^x(y))) = 0$$
we obtain the claimed result.

• **Proof of Lemmas 3.7.** The proof is based on decomposition (4). Therefore, Lemma 3.7 is consequence of a special case of the lemmas Lemma 3.1 with Lemma 3.4 (it suffices to replace $\widehat{f'}_N^x(y)$ and $\widehat{f}_N^x(y)$ and $f^x(y)$) Remark 3.4 and Lemma 3.6.

■ **Second**

Our main purpose is to study the L^2-consistency and the asymptotic normality of the non-parametric estimate $\hat{h}^x$ of h^x when the random filed $(Z_i, i \in \mathbb{N})$ satisfies the Previous mixing condition mentioned with change:
We add at the (H3) $\Phi_l'(s) = \mathbb{E}[\frac{\partial^l f'^X(y)}{\partial y^l} - \frac{\partial^l f'^x(y)}{\partial y^l} | d(x, X) = s]$ are derivable at $s = 0$.
and the fourth and seventh conditions are becoming as follows:
(H_4) The bandwidth h_K satisfies for $j = 0, 1$:

$$h_K \downarrow 0, \forall t \in [0,1] \lim_{h_K \to 0} \frac{\phi_x(th_K)}{\phi_x(h_K)} = \beta_x(t) \text{ and } nh_H^{2j+1}\phi_x(h_K) \to \infty \text{ as } n \longrightarrow \infty.$$

$(H7)$ There exist sequences of integers (u_n) and (v_n) increasing to infinity such that $(u_n + v_n) \leq n$, satisfying for j=0,1

$$(i) v_n = o((nh_H^{2j+1}\phi_x(h_K))^{1/2}) \text{ and } (\tfrac{n}{h_H^{2j+1}\phi_x(h_K)})^{1/2}\alpha(v_n) \longrightarrow 0 \quad as \quad n \longrightarrow 0,$$
$$(ii) q_n v_n = o((nh_H^{2j+1}\phi_x(h_K))^{1/2}) \text{ and } q_n(\tfrac{n}{h_H^{2j+1}\phi_x(h_K)})^{1/2}\alpha(v_n) \longrightarrow 0 \quad as \quad n \longrightarrow 0,$$

where q_n is the largest integer such that $q_n(u_n + v_n) \leq n$.

Furthermore, the estimator $\hat{h}^x(y)$ can written as:

$$\widehat{h}^x(y) = \frac{\hat{f}_N^x(y)}{\hat{F}_D^x - \hat{F}_N^x(y)} \tag{24}$$

where

$$\hat{F}_D^x = \tfrac{1}{n\mathbb{E}[K_1]} \sum_{i=1}^n K(h_K^{-1} d(x, X_i)), K_1 = K(h_K^{-1} d(x, X_1))$$

$$\hat{F}_N^x(y) = \tfrac{1}{n\mathbb{E}[K_1]} \sum_{i=1}^n K(h_K^{-1} d(x, X_i)) H(h_H^{-1}(y - Y_i))$$

$$\hat{f}_N^x(y) = \tfrac{1}{nh_H\mathbb{E}[K_1]} \sum_{i=1}^n K(h_K^{-1} d(x, X_i)) H'(h_H^{-1}(y - Y_i))$$

2.5.1 Mean squared convergence

The first result concerns the L^2-consistency of $\hat{h}^x(y)$.

Theorem 1. Under assumptions (H0)-(H6), we have

$$\mathbb{E}\left[\hat{h}^x(y) - h^x(y)\right]^2 = B_n^2(x, y) + \frac{\sigma_h^2(x, y)}{nh_H\phi_x(h_K)} + o(h_H^4) + o(h_K) + o\left(\frac{1}{nh_H\phi_x(h_K)}\right),$$

and

$$\int_{\mathbb{R}} \mathbb{E}\left[\hat{h}^x(y) - h^x(y)\right]^2 dx = \int_{\mathbb{R}} B_n^2(x, y) dx + \int_{\mathbb{R}} \frac{\sigma_h^2(x, y)}{nh_H\phi_x(h_K)} dx + o\left(\frac{1}{nh_H\phi_x(h_K)}\right),$$

where

$$B_n(x,y) = \frac{(B_H^f - h^x(y)B_H^F)h_H^2 + (B_K^f - h^x(y)B_K^F)h_K}{1 - F^x(y)}$$

with

$$\begin{aligned}
B_H^f(x,y) &= \tfrac{1}{2}\tfrac{\partial^2 f^x(y)}{\partial y^2}\int t^2 H'(t)dt \\
B_K^f(x,y) &= h_K\Phi_0'(0)\tfrac{(K(1)-\int_0^1 (sK(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)} \\
B_H^F(x,y) &= \tfrac{1}{2}\tfrac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H^{'}(t)dt \\
B_K^F(x,y) &= h_K\Psi_0'(0)\tfrac{(K(1)-\int_0^1 (sK(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)}.
\end{aligned}$$

and

$$\sigma_h^2(x,y) = \frac{\beta_2 h^x(y)}{(\beta_1^2(1 - F^x(y)))} \quad (with \quad \beta_j = K^j(1) - \int_0^1 (K^j)'(s)\beta_x(s)ds, for, j = 1,2),$$

Proof.

By using the same decomposition used in (Theorem 3.1 Rabhi et al. [129], P.408), we show that the proof of Theorem 1 can be deduced from the following intermediates results:

Lemma 2. Under the hypotheses of Theorem (1), we have

$$\mathbb{E}[\hat{f}_N^x(y)] - f^x(y) = B_H^f(x,y)h_H^2 + B_K^f(x,y)h_K + o(h_H^2) + o(h_K)$$

and

$$\mathbb{E}[\hat{F}_N^x(y)] - F^x(y) = B_H^F(x,y)h_H^2 + B_K^F(x,y)h_K + o(h_H^2) + o(h_K).$$

Remark 1. Observe that, the result of this lemma permits to write

$$[\mathbb{E}\hat{F}_N^x(y) - F^x(y)] = o(h_H^2) + \mathbb{O}(h_K)$$

and

$$[\mathbb{E}\hat{f}_N^x(y) - f^x(y)] = o(h_H^2 + h_K).$$

Lemma 3. Under the hypotheses of Theorem (1), we have

$$Var[\hat{f}_N^x(y)] = \frac{\sigma_f^2(x,y)}{nh_H\phi_x(h_K)} + o\left(\frac{1}{nh_H\phi_x(h_K)}\right),$$

$$Var[\hat{F}_N^x(y)] = o\left(\frac{1}{nh_H\phi_x(h_K)}\right)$$

and

$$Var[\hat{F}_D^x] = o\left(\frac{1}{nh_H\phi_x(h_K)}\right).$$

where $\sigma_f^2(x,y) := f^x(y)\int H'^2(t)dt$.

Lemma 4. Under the hypotheses of Theorem (1), we have

$$Cov(\hat{f}_N^x(y), \hat{F}_D^x) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right),$$

$$Cov(\hat{f}_N^x(y), \hat{F}_N^x(y)) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right)$$

and

$$Cov(\hat{f}_D^x, \hat{F}_N^x(y)) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right).$$

Remark 2. It is clear that, the results of Lemmas (3 and 4) allows to write

$$Var(\hat{F}_D^x - \hat{F}_N^x) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right)$$

2.5.2 Asymptotic normality

This section contains results on the asymptotic normality of $\hat{h}^x(y)$ and $\hat{h}'^x(y)$. Let us assume that h^Z is sufficiently smooth (at least of class $\mathcal{C}^2$).
We can write an estimator of the first derivative of the hazard function through the first derivative of the estimator.
It is therefore natural to try to construct an estimator of the derivative of the function h^X on the basis of these ideas. To estimate the conditional distribution function and the conditional density function in the presence of functional conditional random variable X.
The kernel estimator of the derivative of the function conditional random functional h^Z can therefore be constructed as follows:

$$\widehat{h}'^X(y) = \frac{\widehat{f}'^X(y)}{1-\widehat{F}^Y(y)} + (\widehat{h}^X(y))^2, \tag{25}$$

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

$$\hat{f}'^X(y) = \frac{\sum_{i=1}^n h_H^{-2} K(h_K^{-1} d(X,X_i)) H''(h_H^{-1}(y-Y_i))}{\sum_{i=1}^n K(h_K^{-1} d(X,X_i))} \tag{26}$$

Later, we need assumptions on the parameters of the estimator, ie on K, H, H', h_H and h_K are little restrictive. Indeed, on one hand, they are not specific to the problem estimate of h^X (but inherent problems of F^X, f^X and f'^X estimation), and secondly they consist with the assumptions usually made under functional variables.
To obtain the asymptotic normality of the conditional estimates, we have to add the following

assumptions:
(H8) H' is twice differentiable.
(H9) The bandwidth h_H and h_K , small ball probability $\phi_z(h)$ and arithmetical α mixing coefficient with order $a > 3$ satisfying

$$\left\{\begin{array}{l} (H9a) \exists C>0, h_H^{2j+1}\phi_z(h_k) \geq \frac{c}{n^{2/(a+1)}}, for \quad j=0,1 \\ (H9b)\left(\frac{\phi_z(h_k)}{n}\right)^{1/a} + \phi_z(h_k) = 0\left(\frac{1}{n^{2/(a+1)}}\right), for \quad j=0,1 \\ (H9c) \lim_{n\longrightarrow\infty} h_K = 0, \lim_{n\longrightarrow\infty} h_H = 0, and \lim_{n\longrightarrow\infty} \dfrac{\log n}{n h_H^{2j+1}\phi_x(h_K)} = 0, j=0,1; \end{array}\right\}$$

Theorem 5. Assume that (H0)-(H7) hold,and if the following inequalities

$$\exists \eta > 0, \quad C, C' > 0 \quad such \quad that \quad C n^{\frac{3-a}{a+1}+\eta} \leq h_H \phi_x(h_K) \quad and \quad \phi_x(h_K) \leq C' n^{\frac{1}{1-a}} \tag{27}$$

are verified with $a > (5+\sqrt{17})/2$, then we have for any $x \in \mathcal{A}$

$$\left(\frac{n h_H \phi_x(h_K)}{\sigma_h^2(x,y)}\right)^{1/2} \left(\hat{h}^x(y) - h^x(y) - B_n(x,y)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad n \longrightarrow \infty.$$

where

$$\mathcal{A} = \{x \in \mathcal{F}, f^x(y)(1-F^x(y)) \neq 0\}$$

and $\xrightarrow{\mathcal{D}}$ means the convergence in distribution.
Obviously, if one imposes some additional assumptions on the function $\phi_x(.)$ and the bandwidth parameters (h_K and h_H) we can improved our asymptotic normality by removing the bias term $B_n(x,y)$.

Corollary 6.
Under the hypotheses of Theorem 5 and if the bandwidth parameters (h_K and h_H) and if the function $\phi_x(h_K)$ satisfies:

$$\lim_{n\to\infty} (h_H^2 + h_K)\sqrt{n\phi_x(h_K)} = 0$$

we have

$$\left(\frac{n h_H \phi_x(h_K)}{\sigma_h^2(x,y)}\right)^{1/2} \left(\hat{h}^x(y) - h^x(y)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad n \quad \longrightarrow \infty.$$

Proof of Theorem 5 and Corollary 6. We consider the decomposition

$$\begin{aligned} \hat{h}^x(y) - h^x(y) &= \tfrac{1}{\hat{F}_D^x - \hat{F}_N^x(y)}[\hat{f}_N^x(y) - \mathbb{E}\hat{f}_N^x(y)] \\ &+ \tfrac{1}{\hat{F}_D^x - \hat{F}_N^x(y)}\{h^x(y)(\mathbb{E}\hat{F}_N^x(y) - F^x(y)) + (\mathbb{E}\hat{f}_N^x(y) - f^x(y)) \\ &+ \tfrac{h^x(y)}{\hat{F}_D^x - \hat{F}_N^x(y)}(1 - \mathbb{E}\hat{F}_N^x(y) - (\hat{F}_D^x - \hat{F}_N^x(y))) \end{aligned} \tag{28}$$

Therefore, theorem (5) and corollary (6) are consequence of lemma (2), remark (1) and the following results.

Lemma 7. Under the hypotheses of theorem 5

$$\left(\frac{nh_H\phi_x(h_K)}{\sigma_f^2(x,y)}\right)^{1/2}\left(\hat{f}_N^x(y)-\mathbb{E}[\hat{f}_N^x(y)]\right)\longrightarrow \mathbb{N}(0,1).$$

Lemma 8. Under the hypotheses of Theorem 5

$$\hat{F}_D^x-\hat{F}_N^x(y)\to 1-F^x(y)\quad in\quad probability$$

and

$$\left(\frac{nh_H\phi_x(h_K)}{\sigma_h^2(x,y)}\right)^{1/2}\left(\hat{F}_D^x-\hat{F}_N^x(y)-1+\mathbb{E}[\hat{F}_N^x(y)]\right)=Op(1).$$

Theorem 9 Under Assumptions (H0)-(H9), then we have for any $x\in\mathcal{A}$,

$$(nh_H^3\phi_x(h_K))^{1/2}(\hat{h}'^X(y)-h'^X(y))\xrightarrow{\mathcal{D}}\mathcal{N}(0,\sigma_{h'}^2(y))$$

$$a_l^z=K^l(1)-\int_0^1(K^l(u))'\beta_0^z(u)du\quad for\quad l=1,2$$

and

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

Proof. Let

$$\hat{h}'^X(y)=\frac{\hat{f}'^X(y)}{1-\hat{F}^X(y)}+(\hat{h}^X(y))^2, \tag{29}$$

with

$$\hat{h}'^X(y)-\hat{h}^X(y)=\{(\hat{h}^X(y))^2-(h^X(y))^2\}+\left\{\frac{\hat{f}'^X(y)}{1-\hat{F}^X(y)}-\frac{f'^X(y)}{1-F^X(y)}\right\} \tag{30}$$

Using again (30), and fact that

$$\frac{(1-F^X(y))}{(1-\widehat{F}^X(y))(1-F^X(y))}\longrightarrow\frac{1}{1-F^X(y)}$$

and

$$\frac{\widehat{f}'^X(y)}{(1-\widehat{F}^X(y))(1-F^X(y))}\longrightarrow\frac{f'^X(y)}{(1-F^X(y))^2}$$

The asymptotic normality of $(nh_H^3\phi_x(h_K))^{1/2}(\hat{h}'^X(y) - h'^X(y))$ can be deduced from both following lemmas and corollary 6.

Lemma 10. Under assumptions (H0)-(H6)and (H8), we have

$$(n\phi_x(h_K))^{1/2}(\widehat{F}^X(y) - F^X(y)) \xrightarrow{\mathcal{D}} \mathbb{N}(0, \sigma^2_{F^X}(y)) \tag{31}$$

where

$$\sigma^2_{F^X}(y) = \frac{a_2^z F^X(y)(1 - F^X(y))}{(a_1^z)^2}$$

Lemma 11. Under assumptions of theorem 3.3, we have

$$(nh_H^3\phi_x(h_K))^{1/2}(\widehat{f}'^X(y) - f'^X(y)) \xrightarrow{\mathcal{D}} \mathbb{N}(0, \sigma^2_{f'^X}(y)) \tag{32}$$

where

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

The proofs of lemma 10 can be seen in ezzahrioui and ould- said [56].

2.6 Discussions and Applications

In this section we emphasize the potential impact of our work by studying its practical interest in some important statistical problems. Moreover, in order to show the easily implementation of our approach on a concrete cases, we discuss in the second part of this section the practical utilization of our model in risk analysis.

2.6.1 Some derivatives

• On the choices of the bandwidths parameters: As all smoothing by a kernel method, the choice of bandwidths parameters has crucial role in determining the performance of the estimators. The mean quadratic error given in Theorem (1) is a basic ingredient to solve this problem. Usually, the ideal theoretical choices are obtained by minimizing this error. Here, we have explicated its leading term which is

$$B_n^2(x,y) + \frac{\sigma_h^2(x,y)}{nh_H\phi_x(h_K)}.$$

Then, the smoothing parameters minimizing this leading term is asymptotically optimal with respect the L^2-error. However, the practical utilization of this criterium requires some additional computational efforts. More precisely, it requires the estimation of the unknown quantities $\Psi_0', \Phi_0', f^x(y)$ and $F^x(y)$.Clearly, all these estimations can be obtained by using a pilots estimators of the conditional distribution function $F^x(y)$ and of the conditional density $f^x(y)$. Such estimations are possible by using the kernel methods, with a separate choice of the bandwidths parameters between both models. More preciously, for the conditional density, we propose to adopt, to the functional case, the bandwidths selectors studied by Bouraine et al. [24] by considering the following criterion

$$CVPDF = \frac{1}{n}\sum_{i=1} W_1(X_i)\int \hat{f}^{X_i^{-i}2}(y)W_2(y)dy - \frac{2}{n}\sum_{i=1} \hat{f}^{X_i^{-i}}(Y_i)W_1(X_i)W_2(Y_i) \quad (33)$$

while, for the the conditional distribution function we can use the cross-validation rule proposed by De Gooijer and Gannoun [82] (in vectorial case)

$$CVCDF = \frac{1}{n}\sum_{k,l\in I_n} [I_{Y_k\le Y_l} - \hat{F}^{X_k^{-k}}(Y_l)]^2 W(X_k)$$

where W_1, W_2 and W are some suitable trimming functions and

$$\hat{F}^{X_k^{-k}}(Y_l) = \frac{\sum_{i\in I^{k,l}_{n,\varsigma_n}} K(h_K^{-1}d(X_k,X_i))H(h_H^{-1}(Y_l - Y_i))}{\sum^n_{i\in I^{k,l}_{n,\varsigma_n}} K(h_K^{-1}d(X_k,X_i))}$$

and

$$\widehat{f}^{X_i^{-i}}(y) = \frac{h_H^{-1}\sum_{j\in I^i_{n,\varsigma_n}} K(h_K^{-1}d(X_i,X_j))H'(h_H^{-1}(y - Y_j))}{\sum_{j\in I^i_{n,\varsigma_n}} K(h_K^{-1}d(X_i,X_j))}$$

with

$$\begin{cases} I^{k,l}_{n,\varsigma_n} = \{\text{i such that} \quad |i-k|\ge \varsigma_n \quad \text{and} \quad |i-l|\ge\varsigma_n\} \\ I^i_{n,\varsigma_n} = \{j \quad \text{such that} \quad |j-i|\ge\varsigma_n\}. \end{cases}$$

Of course, we can also adopt another selection methods, such that, the parametric bootstrap method, proposed by Hall et al. [83] and Hyndman et al. [99] for, respectively, the conditional cumulative distribution function and the conditional density in the finite dimensional case. Nevertheless, a data-driven method allows to overcome this additional computation is very important in practice and is one of the natural prospects of the present work.

•**Confidence intervals**: The main application of Theorem 5 is to build confidence band for the true value of $h^x(y)$. Similarly to the previous application, the practical utilization of our result in this topic requires the estimation of the quantity $\sigma_h^2(x,y)$. A plug-in estimate for the asymptotic standard deviation $\sigma_h^2(x,y)$ can be obtained by using the estimators $\widehat{f}^x(y)$ and $\widehat{F}^x(y)$ of $f^x(y)$ and $F^x(y)$. Then we get

$$\widehat{\sigma}_h^2(x,y) := \frac{\widehat{\zeta}_2\widehat{f}^x(y)}{\left(\widehat{\zeta}_1^2(1-\widehat{F}^x(y))^2\right)}$$

where

$$\begin{cases} \widehat{\zeta}_1 = \frac{1}{n\phi_x(h_k)}\sum_{i=1}^n K(h_k^{-1}d(x,X_i)) \\ \text{and} \quad \widehat{\zeta}_2 = \frac{1}{n\phi_x(h_k)}\sum_{i=1}^n K^2(h_k^{-1}d(x,X_i)) \end{cases}$$

Clearly, the function $\phi_x(.)$ does not appear in the calculation of the confidence interval by simplification. More precisely, we obtain the following approximate $(1-\zeta)$ confidence band for $h^x(y)$

$$\widehat{h}^x(y) \pm t_{1-\zeta/2} \times \left(\frac{\widehat{\sigma}_h^2(x,y)}{nh_h\phi_x(h_k)}\right)^{1/2}$$

where $t_{1-\zeta/2}$ denotes the $1-\zeta/2$ quantile of the standard normal distribution.

2.7 Appendix

In the following, we will denote $\forall i$

$$K_i = K(h_H^{-1}d(x,X_i)), \quad H_i = H(h_H^{-1}(y-Y_i)) \quad and \quad H_i' = H'(h_H^{-1}(y-Y_i)).$$

Proof of Lemma 2. Firstly, for $\mathbb{E}[\widehat{f}_N^x(y)]$, we start by writing

$$\mathbb{E}[\widehat{f}_N^x(y)] = \frac{\mathbb{E}\left[K_1\mathbb{E}[h_H^{-1}H_1'|X]\right]}{\mathbb{E}[K_1]} \quad with \quad h_H^{-1}\mathbb{E}[H_1'|X] = \int_{\mathbb{R}} H'(t)f^X(y-h_Ht)dt.$$

The latter can be re-written, by using a Taylor expansion under (H3), as follows

$$h_H^{-1}\mathbb{E}[H_1'|X] = f^X(y) + \frac{h_H^2}{2}\left(\int t^2H'(t)dt\right)\frac{\partial^2 f^X(y)}{\partial^2 y} + o(h_H^2).$$

Thus, we get

$$\mathbb{E}[\widehat{f}_N^x(y)] = \frac{1}{\mathbb{E}[K_1]}\left(\mathbb{E}[K_1f^X(y)] + \left(\int t^2H'(t)dt\right)\mathbb{E}\left[K_1\frac{\partial^2 f^X(y)}{\partial^2 y}\right] + o(h_H^2)\right).$$

Let $\Psi_l(.,y) := \frac{\partial^l f(y)}{\partial^l y}$: for $l \in \{0.2\}$, since $\Phi_l(0) = 0$, we have

$$\begin{aligned}
\mathbb{E}[K_1\Psi_l(X,y)] &= \Psi_l(x,y)\mathbb{E}[K_1] + \mathbb{E}[K_1(\Psi_l(X,y) - \Psi_l(x,y))] \\
&= \Psi_l(x,y)\mathbb{E}[K_1] + \mathbb{E}[K_1(\Phi_l(d(x,X))] \\
&= \Psi_l(x,y)\mathbb{E}[K_1] + \Phi_l'(0)\mathbb{E}[d(x,X)K_1] + o(\mathbb{E}[d(x,X)K_1]).
\end{aligned}$$

So,

$$\begin{aligned}
\mathbb{E}[\widehat{f}_N^x(y)] &= f^x(y) + \tfrac{h_H^2}{2}\tfrac{\partial^2 f^x(y)}{\partial y^2}\textstyle\int t^2H'(t)dt + o\left(h_H^2\tfrac{\mathbb{E}[d(x,X)K_1]}{\mathbb{E}[K_1]}\right) \\
&+ \Phi_0'(0)\tfrac{\mathbb{E}[d(x,X)K_1]}{\mathbb{E}[K_1]} + o\left(\tfrac{\mathbb{E}[d(x,X)K_1]}{\mathbb{E}[K_1]}\right).
\end{aligned}$$

Similarly to Ferraty et al. [60] we show that

$$\frac{1}{\phi_x(h_K)}\mathbb{E}[d(x,X)K_1] = h_K\Big(K(1) - \int_0^1 (sK(s))'\beta_x(s)ds + o(1)\Big)$$

and

$$\frac{1}{\phi_x(h_K)}\mathbb{E}[K_1] = K(1) - \int_0^1 K'(s)\beta_x(s)ds + o(1).$$

Hence,

$$\begin{aligned}\mathbb{E}[\widehat{f}_N^x(y)] &= f^x(y) + \tfrac{h_H^2}{2}\tfrac{\partial^2 f^x(y)}{\partial y^2}\int t^2 H'(t)dt \\ &+ h_K\Phi_0'(0)\tfrac{(K(1)-\int_0^1(sK(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)} + o(h_H^2) + o(h_K).\end{aligned}$$

Secondly, concerning $\mathbb{E}[\widehat{F}_N^x(y)]$, we write by an integration by part

$$\mathbb{E}[\widehat{F}_N^x(y)] = \frac{1}{\mathbb{E}[K_1]}\mathbb{E}[K_1\mathbb{E}[H_1|X]] \quad with \quad \mathbb{E}[H_1|X] = \int_{\mathbb{R}} H'(t)F^X(y-h_Ht)dt.$$

The same steps used to studying $\mathbb{E}[\widehat{f}_N^x(y)]$ can be followed to prove that

$$\begin{aligned}\mathbb{E}[\widehat{F}_N^x(y)] &= F^x(y) + \tfrac{h_H^2}{2}\tfrac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H'(t)dt \\ &+ h_K\Psi_0'(0)\tfrac{(K(1)-\int_0^1(sK(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)} + o(h_H^2) + o(h_K).\end{aligned}$$

Proof of Lemma 3. For the first quantity $Var[\widehat{f}_N^x(y)]$, we have

$$s_n^2 = Var[\widehat{f}_N^x(y)] = \frac{1}{(nh_H\mathbb{E}[K_1(x)])^2}Var\left[\sum_{i=1}\Gamma_i(x)\right]$$

where

$$\Gamma_i(x) = K_i(x)H_i'(y) - \mathbb{E}[K_i(x)H_i'(y)].$$

Thus

$$\begin{aligned}Var[\widehat{f}_N^x(y)] &= \tfrac{1}{(nh_H\mathbb{E}[K_1])^2}\underbrace{\sum_{i\neq j} cov(\Gamma_i(x),\Gamma_j(x))}_{s_n^{cov}} + \underbrace{\sum_{i=1}^n Var(\Gamma_i(x))}_{s_n^{var}} \\ &= \tfrac{Var[\Gamma_1]}{n(h_H\mathbb{E}[K_1])^2} + \tfrac{1}{(nh_H\mathbb{E}[K_1])^2}\sum_{i\neq j} Cov(\Gamma_i,\Gamma_j).\end{aligned}$$

Let us calculate the quantity $Var[\Gamma_1(x)]$. We have:

$$\begin{aligned} Var[\Gamma_1(x)] &= \mathbb{E}[K_1^2(x)H_1'^2(y)] - \Big(\mathbb{E}[K_1(x)H_1'(y)]\Big)^2 \\ &= \mathbb{E}[K_1^2(x)]\tfrac{\mathbb{E}[K_1^2(x)H_1'^2(y)]}{\mathbb{E}[K_1^2(x)]} \\ &- (\mathbb{E}[K_1(x)])^2\Big(\tfrac{\mathbb{E}[K_1(x)H_1'(y)]}{\mathbb{E}[K_1(x)]}\Big)^2. \end{aligned}$$

So, by using the same arguments as those used in pervious lemma we get

$$\begin{aligned} \tfrac{1}{\phi_x(h_K)}\mathbb{E}[K_1^2(x)] &= K^2(1) - \int_0^1 (K^2(s))'\beta_x(s)ds + o(1) \\ \tfrac{\mathbb{E}[K_1^2(x)H_1'^2(y)]}{\mathbb{E}[K_1^2(x)]} &= h_H f^x(y)\int H'^2(t)dt + o(h_H) \\ \tfrac{\mathbb{E}[K_1(x)H_1'(y)]}{\mathbb{E}[K_1(x)]} &= h_H f^x(y) + o(h_H) \end{aligned}$$

which implies that

$$Var[\Gamma_i(x)] = h_H\phi_x(h_K)f^x(y)\int H'^2(t)dt\Big(K^2(1) - \int_0^1 (K^2(s))'\beta_x(s)ds\Big) + o(h_H\phi_x(h_K)). \quad (34)$$

Now, let us focus on the covariance term. To do that, we need to calculate the asymptotic behavior of quantity defined as

$$\sum_{i\neq j}|cov(\Gamma_i(x),\Gamma_j(x))| = \sum_{1\leq|i-j|\leq c_n}|cov(\Gamma_i(x),\Gamma_j(x))| = J_{1,n} + J_{2,n}.$$

with $c_n \to \infty$, as $n \to \infty$.
for all (i,j) we write

$$cov(\Gamma_i(x),\Gamma_j(x)) = \mathbb{E}[K_i(x)K_j(x)H_i'(y)H_j'(y)] - (\mathbb{E}[K_i(x)H_i'(y)])^2$$

and we use the fact that

$$\mathbb{E}[H_i'(y)H_j'(y)|(X_i,X_j)] = \mathcal{O}(h_H^2); \forall i\neq j, \mathbb{E}[H_i'(y)|X_i] = \mathcal{O}(h_H); \forall i.$$

For $J_{1,n}$: by means of the integral realized above and under (H2) and (H5), we get

$$\mathbb{E}[K_iK_jH_i'H_j'] \leq Ch_H^2\mathbb{P}[(X_i,X_j)\in B(x,h_K)\times B(x,h_K)]$$

and

$$\mathbb{E}[K_i(x)H_i'(y)] \leq Ch_H\mathbb{P}(X_i\in B(x,h_K)).$$

It follows that, the hypothesis (H0), (H2) and (H5), imply that

$$cov(\Gamma_i(x),\Gamma_j(x)) \leq Ch_H^2\phi_x(h_K)\left(\phi_x(h_K) + \left(\frac{\phi_x(h_K)}{n}\right)^{1/a}\right)$$

So

$$J_{1,n} \leq C\left(nc_n h_H^2 \left(\frac{\phi_x(h_K)}{n} \right)^{1/a} \phi_x(h_K) \right).$$

Hence

$$J_{1,n} = \mathcal{O}\left(nc_n h_H^2 \left(\frac{\phi_x(h_K)}{n} \right)^{1/a} \phi_x(h_K) \right).$$

On the other hand, these covariances can be controled by mean of the usual Davydov-Rios's covariance inequality for mixing processes (see Rio [140], formula 1.12a). Together with (H1), this inequality leads to:

$$\forall i \neq j, \quad |Cov(D_i(x), D_j(x))| \leq C|i-j|^{-a}.$$

By the fact, $\sum_{k \geq c_n+1} k^{-a} \leq \int_{C_n}^{\infty} t^{-a} dt = \frac{c_n^{-a+1}}{a-1}$, we get by applying (H1),

$$J_{2,n} \leq \sum_{|i-j| \geq c_n+1} |i-j|^{-a} \leq \frac{nc_n^{-a+1}}{a-1}$$

Thus, by using the following classical technique (see Bosq [20]), we can write

$$s_n^{cov} = \sum_{0<|i-j|\leq u_n} |Cov(\Gamma_i(x), \Gamma_j(x))| + \sum_{|i-j|>u_n} |Cov(\Gamma_i(x), \Gamma_j(x))|.$$

Thus

$$s_n^{cov} \leq C_n \left(c_n h_H^2 \left(\frac{\phi_x(h_K)}{n} \right)^{1/a} \phi_x(h_K) + \frac{c_n^{-a+1}}{a-1} \right)$$

Choosing $c_n = h_H^{-2} \left(\frac{\phi_x(h_K)}{n} \right)^{-1/a}$, and owing to the right inequality in (H7(ii)), we can deduce

$$s_n^{cov} = o(nh_H \phi_x(h_K)). \tag{35}$$

Finally,

$$\begin{aligned} s_n^2 &= o(nh_H\phi_x(h_K)) + \mathcal{O}(nh_H\phi_x(h_K)) \\ &= \mathcal{O}(nh_H\phi_x(h_K)) \end{aligned}$$

In conclusion, we have

$$\begin{aligned} Var[\widehat{f}_N^x(y)] &= \frac{f^x(y)\int H'^2(t)dt}{nh_H\phi_x(h_K)} \left(\frac{(K^2(1)-\int_0^1 (K^2(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)^2} \right) \\ &+ o\left(\frac{1}{nh_H\phi_x(h_K)} \right) \end{aligned} \tag{36}$$

Now, for $\widehat{F}_N^x(y)$, (resp. $\widehat{F}_D^x$) we replace $H_i'(y)$ by $H_i(y)$ (resp. by 1) and we follow the same ideas, under the fact that $H \leq 1$

$$\begin{aligned} Var[\widehat{F}_N^x(y)] &= \tfrac{F^x(y)}{n\phi_x(h_K)}\left(\int H'^2(t)dt\right)\left(\tfrac{(K^2(1)-\int_0^1 (K^2(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)^2}\right) \\ &+ o\left(\tfrac{1}{n\phi_x(h_K)}\right). \end{aligned}$$

and

$$Var[\widehat{F}_D^x] = \frac{1}{n\phi_x(h_K)}\left(\frac{(K^2(1)-\int_0^1 (K^2(s))'\beta_x(s)ds)}{(K(1)-\int_0^1 K'(s)\beta_x(s)ds)^2}\right) + o\left(\frac{1}{n\phi_x(h_K)}\right).$$

This yields the proof.

Proof of Lemma 4. The proof of this lemma follows the same steps as the previous Lemma. For this, we keep the same notation and we write

$$\begin{aligned} Cov(\widehat{f}_N^x(y), \widehat{F}_N^x(y)) &= \tfrac{1}{nh_H(\mathbb{E}[K_1(x)])^2} Cov(\Gamma_1(x), \Delta_1(x)) \\ &+ \tfrac{1}{n^2h_H(\mathbb{E}[K_1(x)])^2}\sum_{i\neq j} Cov(\Gamma_i(x), \Delta_j(x)) \end{aligned}$$

Where

$$\Delta_i(x) = K_i(x)H_i(y) - \mathbb{E}[K_i(x)H_i(y)].$$

For the first term, we have under (H4)

$$\begin{aligned} Cov(\Gamma_1(x), \Delta_1(x)) &= \mathbb{E}[K_1^2(x)H_1(y)H_1'(y)] - \mathbb{E}[K_1(x)H_1(y)]\mathbb{E}[K_1(x)H_1'(y)] \\ &= \mathcal{O}(h_H\phi_x((h_k)) + \mathcal{O}(h_H\phi_x^2((h_k)) \\ &= \mathcal{O}(h_H\phi_x((h_k)) \end{aligned}$$

Therefore,

$$\frac{1}{nh_H(\mathbb{E}[K_1(x)])^2} Cov(\Gamma_1(x), \Delta_1(x)) = \mathcal{O}\left(\frac{1}{n\phi_x(h_K)}\right) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right) \tag{37}$$

So, by using similar arguments as those invoked in the proof of Lemma 3, and we use once again the boundedness of K and H, and the fact that (H1) and (H6) imply that

$$\mathbb{E}(H_i'(y)|X_i) = \mathcal{O}(h_H).$$

Moreover, the right part of (H7(ii)) implies that

$$Cov(\Gamma_i(x), \Delta_j(x)) = \mathcal{O}\left(h_H\phi_x(h_K)\left(\frac{\phi_x(h_K)}{n}\right)^{1/a} + \phi_x(h_K)\right),$$

Meanwhile, using the Davydov-Rio's inequality in Rio [140] for mixing processes leads to

$$|Cov(\Gamma_i(x), \Delta_j(x))| \leq C\alpha(|i-j|) \leq C|i-j|^{-a},$$

we deduce easily that for any $c_n > 0$:

$$\sum_{i\neq j} Cov(\Gamma_i(x), \Delta_j(x)) = \mathcal{O}\left(nc_n h_H \phi_x(h_K)\left(\frac{\phi_x(h_K)}{n}\right)^{1/a} + \phi_x(h_K)\right) + \mathcal{O}(nh_H c_n^{-a}).$$

It suffices now to take $c_n = h_H^{-1}\left(\frac{\phi_x(h_K)}{n}\right)^{1/a}$ to get the following expression for the sum of the covariances:

$$\sum_{i\neq j} Cov(\Gamma_i(x), \Delta_j(x)) = o(n\phi_x(h_K)). \tag{38}$$

From (37) and (38) we deduce that

$$Cov(\widehat{f}_N^x(y), \widehat{F}_N^x(y)) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right)$$

The same arguments can be used to shows that

$$Cov(\widehat{f}_N^x(y), \widehat{F}_D^x) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right)$$

and

$$Cov(\widehat{F}_N^x(y), \widehat{F}_D^x) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right).$$

Proof of Lemma 7. Let

$$S_n = \sum_{i=1}^n \Lambda_i(x)$$

Where

$$\Lambda_i(x) := \frac{\sqrt{h_H\phi_x(h_K)}}{h_H\mathbb{E}[K_1(x)]}\Gamma_i(x). \tag{39}$$

Obviously, we have

$$\sqrt{nh_H\phi_x(h_K)}[\sigma_f(x,y)]^{-1}(\widehat{f}_N^x(y) - \mathbb{E}\widehat{f}_N^x(y)) = (n(\sigma_f(x,y))^2)^{-1/2}S_n.$$

Thus, the asymptotic normality of $(n(\sigma_f(x,y))^2)^{-1/2}S_n$, is sufficient to show the proof of this Lemma. This last is shown by the blocking method, where the random variables Λ_i are grouped into blocks of different sizes defined. We consider the classical big- and small-block decomposition. We split the set {1,2,...,n} into $2k_n + 1$ subsets with large blocks of size u_n and small blocks of size v_n and put

$$k_n := \left[\frac{n}{u_n + v_n}\right].$$

Assumption (H7(ii)) allows us to define the large block size by

$$u_n := \left[\left(\frac{nh_H\phi_x(h_K)}{q_n}\right)^{1/2}\right].$$

Using Assumption (H7) and simple algebra allows us to prove that

$$\frac{v_n}{u_n} \to 0, \quad \frac{u_n}{n} \to 0, \quad \frac{u_n}{\sqrt{nh_H\phi_x(h_K)}} \to 0, \quad and \quad \frac{n}{u_n}\alpha(v_n) \to 0 \tag{40}$$

Now, let Υ_j,$\Upsilon_j^{'}$ and $\Upsilon_j^{''}$ be defined as follows:

$$\Upsilon_j = \sum_{i=j(u+v)+1}^{j(u+v)+u} \Lambda_i(x), \quad 0 \leq j \leq k+1$$

$$\Upsilon_j^{'} = \sum_{i=j(u+v)+u+1}^{(j+1)(u+v)+u} \Lambda_i(x), \quad 0 \leq j \leq k+1$$

$$\Upsilon_j^{''} = \sum_{i=k(u+v)+1}^{n} \Lambda_i(x), \quad 0 \leq j \leq k+1$$

Cleary, we can write

$$S_n := \sum_{j=0}^{k-1} \Upsilon_j + \sum_{j=0}^{k-1} \Upsilon_j^{'} + \Upsilon_k^{''}r =: S_n^{'} + S_n^{''} + S_n^{'''}.$$

We prove that

$$(i)\frac{1}{n}\mathbb{E}(S_n^{''})^2 \to 0, \quad (ii)\frac{1}{n}\mathbb{E}(S_n^{'''})^2 \to 0, \tag{41}$$

$$|\mathbb{E}\{exp(itn^{-1/2}S_n^{'})\} - \prod_{j=0}^{k-1}\mathbb{E}\{exp(itn^{-1/2}\Upsilon_j)\}| \to 0, \tag{42}$$

$$\frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}(\Upsilon_j^2) \longrightarrow \sigma^2_{f(x,y)}, \tag{43}$$

$$\frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}\left(\Upsilon_j^2\mathbb{I}_{\{|\Upsilon_j|>\varepsilon\sqrt{n\sigma_f^2(x,y)}\}}\right)\longrightarrow 0, \tag{44}$$

for every $\varepsilon>0$.
Expression (41) show that the terms S_n'' and S_n''' are negligible, while Equations (42) and (43) show that the Υ_j are asymptotically independent, verifying that the sum of their variances tends to $\sigma_f^2(x,y)$. Expression (44) is the Lindeberg-Feller's condition for a sum of independent terms. Asymptotic normality of S_n is a consequence of Equations (41)-(44).

- **Proof of (41)** Because $\mathbb{E}(\Lambda_j)=0,\forall j$, we have that

$$\mathbb{E}(S_n'')^2 = Var\left(\sum_{j=0}^{k-1}\Upsilon_j'\right)=\sum_{j=0}^{k-1}Var(\Upsilon_j')+\sum_{0\le i<j\le k-1}Cov(\Upsilon_i',\Upsilon_j'):=\Pi_1+\Pi_2.$$

By the second-order stationarity we get

$$\begin{aligned}Var(\Upsilon_j') &= Var\left(\sum_{i=j(u_n+v_n)+u_n+1}^{(j+1)(u_n+v_n)}\Lambda_i(x)\right)\\ &= v_nVar(\Lambda_1(x))+\sum_{i\neq j}^{v_n}Cov(\Lambda_i(x),\Lambda_j(x)).\end{aligned}$$

Then

$$\begin{aligned}\tfrac{\Pi_1}{n} &= \tfrac{kv_n}{n}Var(\Lambda_1(x))+\tfrac{1}{n}\sum_{j=0}^{k-1}\sum_{i\neq j}^{v_n}Cov(\Lambda_i(x),\Lambda_j(x))\\ &\le \tfrac{kv_n}{n}\left\{\tfrac{\phi_x(h_k)}{h_H\mathbb{E}^2K_1(x)}Var(\Gamma_1(x))\right\}+\tfrac{1}{n}\sum_{i\neq j}^{n}|Cov(\Lambda_i(x),\Lambda_j(x))|\\ &\le \tfrac{kv_n}{n}\left\{\tfrac{1}{h_H\phi_x(h_k)}Var(\Lambda_1(x))\right\}+\tfrac{1}{n}\sum_{i\neq j}^{n}|Cov(\Lambda_i(x),\Lambda_j(x))|\end{aligned}$$

Simple algebra gives us

$$\frac{kv_n}{n}\cong\left(\frac{n}{u_n+v_n}\right)\frac{v_n}{n}\cong\frac{v_n}{u_n+v_n}\cong\frac{v_n}{u_n}\longrightarrow 0\quad as\quad n\longrightarrow\infty.$$

Using Equation (35) we have

$$\lim_{n\longrightarrow\infty}\frac{\Pi_1}{n}=0 \tag{45}$$

Now, let us turn to Π_2/n. We have

$$\begin{aligned}\frac{\Pi_2}{n} &= \frac{1}{n}\sum_{i=0_{i\neq j}}^{k-1}\sum_{j=0}^{k-1} Cov(\Upsilon_i(x),\Upsilon_j(x))\\ &= \frac{1}{n}\sum_{i=0_{i\neq j}}^{k-1}\sum_{j=0}^{k-1}\sum_{l_1=1}^{v_n}\sum_{l_2}^{v_n} Cov(\Lambda_{m_j+l_1},\Lambda_{m_j+l_2})\end{aligned}$$

with $m_i = i(u_n+v_n)+v_n$. As $i\neq j$, we have $|m_i - m_j + l_1 - l_2| \geq u_n$. It follows that

$$\frac{\Pi_2}{n} \leq \frac{1}{n}\sum_{i=l_{|i-j|\geq u_n}}^{n}\sum_{j=1}^{n} Cov(\Lambda_i(x),\Lambda_j(x)),$$

Then

$$\lim_{n\longrightarrow\infty}\frac{\Pi_2}{n} = 0. \tag{46}$$

By Equations (45) and (46) we get Part(i) of the Equation(41).
We turn to (ii), we have

$$\begin{aligned}\frac{1}{n}\mathbb{E}(S_n''')^2 &= \frac{1}{n}Var(\Upsilon_K'')\\ &= \frac{\mathcal{V}_n}{n}Var(\Lambda_1(x)) + \frac{1}{n}\sum_{i=1_{i\neq j}}^{\mathcal{V}_n}\sum_{j=1}^{\mathcal{V}_n} Cov(\Lambda_i(x),\Lambda_j(x))\end{aligned}$$

where $\mathcal{V}_n = n - k_n(u_n+v_n)$; by the definition of k_n, we have $\mathcal{V}_n \leq u_n+v_n$.
Then

$$\frac{1}{n}\mathbb{E}(S_n''')^2 \leq \frac{u_n+v_n}{n}Var(\Lambda_1(x)) + \frac{1}{n}\sum_{i=1_{i\neq j}}^{\mathcal{V}_n}\sum_{j=1}^{\mathcal{V}_n} Cov(\Lambda_i(x),\Lambda_j(x))$$

and by the definition of u_n and v_n we achieve the proof of (ii) of Equation (41).

- **Proof of (42)** We make use of Volkonskii and Rozanov's lemma (see the appendix in Masry, [114]) and the fact that the process (X_i, X_j) is strong mixing.

Note that $\Upsilon_a is \mathcal{F}_{i_a}^{j_a}$-mesurable with $i_a = a(u_n+v_n)+1$ and $j_a = a(u_n+v_n)+u_n$; hence, with $V_j = \exp(itn^{-1/2}\Upsilon_j)$ we have

$$\begin{aligned}|\mathbb{E}\left\{\exp(itn^{-1/2}S_n')\right\} - \prod_{j=0}^{k-1}\mathbb{E}\left\{\exp(itn^{-1/2}\Upsilon_j)\right\}| &\leq 16k_n\alpha(v_n+1)\\ &\cong \frac{n}{u_n}\alpha(v_n+1)\end{aligned}$$

which goes to zero by the last part of Equation (40). Now we establish Equation (43).

• **Proof of (43)**Note that $Var(S'_n) \to \sigma_f^2(x,y)$ by equation (41) and since $Var(S_n) \to \sigma_f^2(x,y)$ (by the definition of the Λ_i and Equation (36)). Then because

$$\mathbb{E}(S'_n)^2 = Var(S'_n) = \sum_{j=0}^{k-1} Var(\Upsilon_j) + \sum_{i=0_{i\neq j}}^{k-1} \sum_{j=0}^{k-1} Cov(\Upsilon_i, \Upsilon_j),$$

all we have to prove is that the double sum of covariances in the last equation tends to zero. Using the same arguments as those previously used for Π_2 in the proof of first term of Equation (41)we obtain by replacing v_n by u_n we get

$$\frac{1}{n}\sum_{j=0}^{k-1} \mathbb{E}(\Upsilon_j^2) = \frac{ku_n}{n} Var(\Lambda_1) + o(1).$$

As $Var(\Lambda_1) \longrightarrow \sigma_f^2(x,y)$ and $ku_n/n \longrightarrow 1$, we get the result.
Finally, we prove Equation (43).

• **Proof of (44)** Recall that

$$\Upsilon_j = \sum_{i=j(u_n+v_n)+1}^{j(u_n+v_n)+u_n} \Lambda_i.$$

Making use Assumptions (H5) and (H6), we have

$$|\Lambda_i| \leq C(h_H\phi_x(h_k))^{-1/2}$$

thus

$$|\Upsilon_j| \leq Cu_n(h_H\phi_x(h_k))^{-1/2},$$

which goes to zero as n goes to infinity by Equation (40). Then for n large enough, the set $\{|\Upsilon_j| > \varepsilon(n\sigma_f^2(x,y))^{-1/2}\}$ becomes empty, this completes the proof and therefore that of the asymptotic normality of $(n(\sigma_f(x,y))^2)^{-1/2}S_n$,

• **Proof of Lemmas 8.** It is clear that, the result of Lemma (3.1) and Lemma (3.2) permits us

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

and

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

then

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

Moreover, the asymptotic variance of $\widehat{F}_D^x - \widehat{F}_N^x$ given in remark (2) allows to obtain

$$\frac{nh_H\phi_x(h_K)}{\sigma_h(x,y)^2} Var(\widehat{F}_D^x - \widehat{F}_N^x - 1 + \mathbb{E}(\widehat{F}_N^x(y))) \longrightarrow 0.$$

By combining result with the fact that

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

we obtain the claimed result.

• **Proof of Lemmas 11.** For $i = 1, ..., n$, we consider the quantities $K_i = K(h_K^{-1} d(z, Z_i))$, $H_i'' = H''(h_H^{-1}(x - X_i))$, and let $\widehat{f'}_N^Z(x)$ (resp. $\widehat{F}_D^Z$) be defined as:

$$\widehat{f'}_N^Z(x) = \frac{h_H^{-2}}{n\mathbb{E}K_1} \sum_{i=1}^{n} K_i H_i''(x) \quad (resp \quad \widehat{F}_D^Z = \frac{1}{n\mathbb{E}K_1} \sum_{i=1}^{n} K_i).$$

This proof is based on the following decomposition

$$\begin{aligned} \widehat{f'}^X(y) - f'^X(y) &= \tfrac{1}{\widehat{F}_D^X} \{(\widehat{f'}_N^X(y) - \mathbb{E}\widehat{f'}_N^X(y)) - (f'^X(y) - \widehat{f'}_N^X(y))\} \\ &+ \tfrac{f'^X(y)}{\widehat{F}_D^X} \{\mathbb{E}\widehat{F}_D^X - \widehat{F}_D^X\} \end{aligned} \tag{47}$$

and the following intermediate results.

$$\sqrt{nh_H^3 \phi_x(h_K)} (\widehat{f'}_N^X(y) - \mathbb{E}\widehat{f'}_N^X(y)) \xrightarrow{\mathbb{D}} \mathcal{N}(0, \sigma_{f'^x}^2(y)) \tag{48}$$

where $\sigma_{f'^x}^2(y)$ is define as lemma 11.

$$\lim_{n \longrightarrow \infty} \sqrt{nh_H^3 \phi_x(h_K)} (\mathbb{E}\widehat{f'}_N^X(y) - \widehat{f'}^X(y)) = 0 \tag{49}$$

$$\sqrt{nh_H^3 \phi_x(h_K)} (\widehat{F}_D^X(y) - 1) \xrightarrow{\mathbb{P}} 0, \quad as \quad n \longrightarrow \infty. \tag{50}$$

• Concerning (48). by definition of $\widehat{f'}_N^X(y)$, it follows that

$$\sqrt{nh_H^3 \phi_x(h_K)} (\widehat{f'}_N^X(y) - \mathbb{E}\widehat{f'}_N^X(y)) = \sum_{i=1}^{n} \frac{\sqrt{\phi_x(h_K)}}{\sqrt{nh_H}\mathbb{E}K_1} (K_i H_i'' - \mathbb{E}K_i H_i'') = \sum_{i=1}^{n} \Delta_i,$$

which leads

$$\sum_{i=1}^{n} \mathbb{E}\Delta_i^2 = \frac{\phi_x(h_K)}{h_H \mathbb{E}^2 K_1} \mathbb{E}K_1^2 (H_1'')^2 - \frac{\phi_x(h_K)}{h_H \mathbb{E}^2 K_1} (\mathbb{E}K_1 H_1'')^2 = \Pi_{1n} - \Pi_{2n}. \tag{51}$$

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

$$\Pi_{1n} = \frac{\phi_x(h_K)}{\mathbb{E}^2 K_1} \mathbb{E}\{K_1^2 \int H''^2(t)(f'^X(y - th_H) - f'^X(y) + f'^X(y))dt\}.$$

Meanwhile, by (H0), (H3), (H4) and (H5), it follows that:

$$\frac{\phi_x(h_K)\mathbb{E}K_1^2}{\mathbb{E}^2 K_1} \longrightarrow_{n\longrightarrow\infty} \frac{a_2^z}{(a_1^z)^2},$$

which leads

$$\Pi_{1n} \longrightarrow_{n\longrightarrow\infty} \frac{a_2^z f^X(y)}{(a_1^z)^2}\int (H''(t))^2 dt, \tag{52}$$

Regarding Π_{2n}, by (H0), (H3), and (H6), we obtain

$$\Pi_{2n} \longrightarrow_{n\longrightarrow\infty} 0. \tag{53}$$

This result, combined with (51) and (52), allows us to get

$$\lim_{n\longrightarrow\infty}\sum_{i=1}^{n}\mathbb{E}\Delta_i^2 = \sigma^2_{f'^X}(y) \tag{54}$$

Secondly, by the boundedness of H", we have

$$\begin{aligned}\mathbb{E}(\mid \Delta_i\Delta_j \mid) &\leq \tfrac{C\phi_x(h_K)}{n\mathbb{E}^2K_1}(K_iK_j + \mathbb{E}K_iK_j)\\ &\leq \tfrac{C}{nh_H}\left\{\left(\tfrac{\phi_x(h_K)}{n}\right)^{1/a} + \phi_x(h_K)\right\}, \forall i\neq j.\end{aligned}$$

Then taking

$$\delta_n = \max_{1\leq i\neq j\leq n}\{\mathbb{E}(\mid \Delta_i\Delta_j \mid)\} = \frac{C}{nh_H}\left(\left(\frac{\phi_x(h_K)}{n}\right)^{1/a} + \phi_x(h_K)\right).$$

leads

$$nm_n\delta_n = \frac{Cm_n}{h_H}\left(\left(\frac{\phi_x(h_K)}{n}\right)^{1/a} + \phi_x(h_K)\right). \tag{55}$$

Similarly, the boundedness of H" and K allows us to take $C_i = O\left(\frac{1}{\sqrt{nh_H^3\phi_x(h_K)}}\right)$, which implies that

$$\left(\sum_{j=m_n+1}^{\infty}\alpha(j)\right)\sum_{i=1}^{n}C_i^2 \leq \frac{C}{h_H\phi_x(h_K)}\int_{t\geq m_n} t^{-a}dt = \frac{C}{h_H\phi_x(h_K)}\frac{m_n^{-a+1}}{a-1}. \tag{56}$$

Then, the sum of right side of of (55) and (56) is of type $Am_n + Bm_n^{-a+1}$, by talking

$$m_n = (A/B)^{-1/a} = \{(a-1)\phi_x(h_K)((\frac{\phi_x(h_K)}{n})^{1/a} + \phi_x(h_K))\}^{-1/a} \longrightarrow \infty,$$

it is clear that, under condition (H9a) and (H9b), combining (55) and (56) allows us to get

$$nm_n\delta_n = o(1), \tag{57}$$

and

$$\left(\sum_{j=m_n+1}^{\infty}\alpha(j)\right)\sum_{i=1}^{n}C_i^2 = o(1), \tag{58}$$

respectively. Finally, by choosing $\varrho_n = \sqrt{\frac{nh_H^3\phi_x(h_K)}{\log n}}$, under (H9a) again and $a > 3$, we have

$$\frac{\varrho_n}{\sqrt{n}} = o(1), \tag{59}$$

and

$$\begin{aligned}\frac{n}{\varrho_n}\alpha(\varepsilon\varrho_n) &\leq C\frac{(\log n)^{(a+1)/2}}{n^{(a-1)/2}(h_H^3\phi_x(h_K))^{(a+1)/2}}\\ &\leq C\frac{(\log n)^{(a+1)/2}}{n^{(a-3)/2}} \longrightarrow 0 \quad as \quad n\longrightarrow\infty\end{aligned}$$

.

Therefore, combing (53)-(59) with corollary 2.2 in liebscher [109] (48) is valid.

• Concerning (49). The proof is completed along the same stapes as that of Π_{1n}. We omit it here.

• Concerning (50). The idea is similar to that given by Ferraty et al. [63]
by definition of $\widehat{F}_D^X(y)$, we have

$$\sqrt{nh_H^3\phi_x(h_K)}(\widehat{F}_D^X(y) - 1) = \Omega_n - \mathbb{E}\Omega_n,$$

where $\Omega_n = \frac{\sqrt{nh_H^3\phi_x(h_K)}\sum_{i=1}^{n}K_i}{n\mathbb{E}K_1}$. In order to prove (50), similar to Ferraty et al.[63], we only need to proov $Var\Omega_n \longrightarrow 0, \quad as \quad n\longrightarrow\infty$. In fact, since

$$\begin{aligned}Var\Omega_n &= \frac{nh_H^3\phi_x(h_K)}{n\mathbb{E}^2K_1}\left(nVarK_1 + \sum_{1\leq i}\sum_{j\leq n}Cov(K_i,K_j)\right)\\ &\leq \frac{nh_H^3\phi_x(h_K)}{\mathbb{E}^2K_1}\mathbb{E}K_1^2 + \frac{nh_H^3\phi_x(h_K)}{n\mathbb{E}^2K_1}\sum\sum_{0\leq|i-j|\leq v_n}Cov(K_i,K_j)\\ &+ \frac{nh_H^3\phi_x(h_K)}{n\mathbb{E}^2K_1}\sum\sum_{0\leq|i-j|\geq v_n}Cov(K_i,K_j)\\ &= \Psi_1+\Psi_2+\Psi_3,\end{aligned}$$

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

$$\Psi_1 \leq Ch_H^3 \phi_x(h_K) \longrightarrow 0 \quad as \quad n \longrightarrow \infty.$$

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

$$\Psi_2 \leq v_n h_H^3 \left\{ (\frac{\phi_x(h_K)}{n})^{1/a} + \phi_x(h_K) \right\}. \tag{60}$$

Finally, using the Davydov-Rio's inequality in Rio [140] for mixing processes leads to

$$\mid Cov(K_i, K_j) \mid \leq C\alpha(\mid i - j \mid),$$

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

$$\begin{aligned} \Psi_3 &\leq \tfrac{h_H^3 \phi_x(h_K)}{n\mathbb{E}^2 K_1} n^2 C\alpha(\mid i - j \mid) \leq C \tfrac{h_H^3 \phi_x(h_K)}{n\mathbb{E}^2 K_1} n^2 v_n^{-a+1} \\ &\leq Ch_H^3 n v_n^{-a+1}. \end{aligned} \tag{61}$$

Since the right side (60) and (61) is also of type $Av_n + Bv_n^{-a+1}$, by choosing $v_n = [n^{-1}(\left(\frac{\phi_x(h_K)}{n}\right)^{1/a} + \phi_x(h_K))]^{1/a} \longrightarrow \infty$ and simple calculations, we get that $\Psi_2 \longrightarrow 0$ and $\Psi_3 \longrightarrow 0$ as $n \longrightarrow \infty$, respectively.

Therefore, the proof of this result is completed.

Chapter 3

Real response and independent case

This chapter is the object of a work subjected for publication in Journal of Statistics Applications & Probability Letters.

Exact Asymptotic Errors of the Hazard Conditional Rate Kernel

3.1 Introduction

This chapter deals with a scalar response conditioned by a functional random variable. The main goal is to estimate nonparametrically Kernel type estimator for the conditional hazard function. Finally, asymptotic properties of this estimator are stated bias the exact expression involved in the leading terms of the quadratic error and we investigate the asymptotic normality of the kernel conditional hazard function estimator.

3.2 General Notations and Conditions

We consider a random pair (X, Y) where Y is valued in $\mathbb{R}$ and X is valued in some semi normed vector space $(F, \| . \|)$ which can be of infinite dimension . We will say that X is a functional random variable and we will use the abbreviation frv. From a sample of independent pairs (X_i, Y_i), each having the same distribution as (X, Y), our aim is to study convergence mean square of the estimator of the conditional hazard function of a real random variable conditional on one variable functional. The nonparametric estimate of function related with the conditional probability distribution (cond-cdf) of Y given $X = x$. For $x \in \mathbb{F}$, we assume that the regular version of the conditional probability of Y given $X = x$ exists denoted by F_Y^X and has a bounded density with respect to Lebesgue measure over R, denoted by f_Y^X. In the following (x, y) will be a fixed point in $\mathbb{R} \times \mathbb{F}$ and $N_x \times S_{\mathbb{R}}$ will denote a fixed neighborhood of (x, y), S_R will be a fixed compact subset of $\mathbb{R}$, and we will use the notation $B(x, h) = \{x' \in \mathbb{F} / \parallel x' - x \parallel < h\}$. Our nonparametric models will be quite general in the sense that we will just need the following simple assumption for the marginal distribution of X:

$$C_B^2(\mathcal{F} \times \mathbb{R}) = \left\{ \begin{array}{l} \varphi : \mathcal{F} \times \mathbb{R} \to \mathbb{R} \\ (x, y) \to \varphi(x, y) such \quad as : \\ \forall z \in N_x, \varphi(z, .) \in C^2(N_y) and \left(\varphi(., y), \frac{\partial^2 \varphi(., y)}{\partial y^2}\right) \in C_B^1(x) \times C_B^1(x), \end{array} \right\} \quad (1)$$

where $C_B^1(x)$ is the set of continuously differentiable functions to sens of Gteaux on N_x (see Troutman [154] for this type of differentiability), which the derivative operator of order 1 at point x is bounded on the unit ball $B(0,1)$ the functional space $\mathbb{F}$. Given i.i.d. observations $(X_1, Y_1), ..., (X_n, Y_n)$ of (X, Y), the kernel estimate of the conditional distribution $F_Y^X(x, y)$ denoted $\hat{F}_Y^X(x, y)$,is defined by:

$$\hat{F}_Y^X(x,y) = \frac{\sum_{i=1}^n K(h_K^{-1} \parallel x - X_i \parallel) H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^n K(h_K^{-1} \parallel x - X_i \parallel)},$$

withe the convention $\frac{0}{0} = 0$, The function K is kernel, H is *cdf* and $h_K = h_{K,n}$ (resp $h_H = h_{H,n}$)is sequence of positive real number. Note that from this estimator, we derive an estimator for the density conditional, denoted $\hat{f}_Y^X(x, y)$ defined by:

$$\hat{f}_Y^X(x,y) = \frac{h_H^{-1} \sum_{i=1}^n K(h_K^{-1} \parallel x - X_i \parallel) H'(h_H^{-1}(y - Y_i))}{\sum_{i=1}^n K(h_K^{-1} \parallel x - X_i \parallel)},$$

where H' is kernel (is derivative of H). We then construct the conditional hazard function of y knowing $X = x$ as follows:

$$\forall x \in \mathbb{F}, \quad \forall y \in \mathbb{R} \quad h_Y^X(x,y) = \frac{f_Y^X(x,y)}{1 - F_Y^X(x,y)} = \frac{f_Y^X(x,y)}{S_Y^X(x,y)} \tag{2}$$

The main objective is to study the nonparametric estimate $\hat{h}_Y^X(x, y)$ of $h_Y^X(x, y)$. Furthermore,$h_Y^X(x, y)$ the estimator can we written as

$$\hat{h}_Y^X(x,y) = \frac{\hat{f}_Y^X(x,y)}{1 - \hat{F}_Y^X(x,y)} = \frac{\hat{f}_N(x,y)}{\hat{f}_D(x) - \hat{g}_N(x,y)}, \tag{3}$$

Where

$$\hat{f}_D(x) = \frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K(h_K^{-1} \parallel x - X_i \parallel), K_1(x) = K(h_K^{-1} \parallel x - X_i \parallel),$$

$$\hat{g}_N(x,y) = \frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K(h_K^{-1} \parallel x - X_i \parallel) H(h_H^{-1}(y - Y_i)),$$

$$\hat{f}_N(x,y) = \hat{g}_N^{(1)}(x,y) = \frac{1}{nh_H\mathbb{E}[K_1(x)]} \sum_{i=1}^n K(h_K^{-1} \parallel x - X_i \parallel) H'(h_H^{-1}(y - Y_i)),$$

where H' is the derivative of H, when the explanatory variable$X$ is valued in a space of eventually infinite dimension.We give precise asymptotic evaluations of the quadratic error of this estimator.

3.3 Asymptotic Properties

As with any problem of nonparametric estimation, the dimension of the space $\mathbb{F}$ plays an important role in the properties of concentration of the variable X. Thus, when this dimension is not necessarily finite, the probability functions defined by small balls

$$\phi_x(h) = \mathbb{P}(X \in B(x,h)) = \mathbb{P}(X \in \{x' \in \mathbb{F}, \parallel x - x' \parallel < h\}), \tag{$\star$}$$

Then to establish the convergence in mean square of the estimator $\hat{h}_Y^X(x,y)$ to $h_Y^X(x,y)$ and the asymptotic normality of the kernel conditional hazard function estimator, we introduce the following assumptions, let b_1 and b_2 be two positive numbers; such that:

$(H1)$ for all $r > 0$, the random variable $Z = r^{-1}(x - X)$ is absolutely continuous relative in the measure μ. His density $w(r, x, v)$ is strictly positive on $B(0,1)$ and can be written as:

$$w(r,x,v) = \phi(r)g(x,v) + o(\phi(r)) \quad for \quad all \quad v \in B(0,1), \tag{4}$$

where

$-\phi$ is an increasing function with values in $\mathbb{R}^+$.

$-g$ is defined on $\mathbb{F} \times \mathbb{F}$, with values in $\mathbb{R}^+$ where $0 < \int_{B(0,1)} g(x,v)d\mu(v) < \infty$.

(H2) The kernel K from $\mathbb{R}$ into $\mathbb{R}^+$ is a differentiable function supported on $[0,1]$ its derivative K' exist and such that there exist tow constants C and C' with $-\infty < C < K'(t) < C' < 0 \quad for \quad 0 \leq t \leq 1$.

(H3) H' is a kernel bounded, integrable, positive, symmetric such that:

$$\int H'(t)dt = 1, \int t^2 H'(t)dt < \infty, \quad \int_{\mathbb{R}} \mid t \mid^{b_2} H'(t)dt < \infty,$$

where

$$H(x) = \int_{-\infty}^{x} H'(t)dt \quad (see \quad Feraty \quad and \quad Vieu \quad [8])$$

(H4)The bandwidth h_K satisfies:

$$h_K \downarrow 0, \forall t \in [0,1] \qquad \lim_{h_K \longrightarrow 0} \frac{\phi_x(th_K)}{\phi_x(h_K)} = \beta^x_{h_K}(t) \quad and \quad nh_H\phi_x(h_K) \longrightarrow \infty \quad n \longrightarrow \infty.$$

$$(H5)\begin{cases} \nexists\tau < \infty, f_Y^X(x,y) \leq \tau, \forall (x,y) \in \mathcal{F} \times \mathcal{L}_{\mathbb{R}}, and; \\ \forall (x_1,x_2) \in N_x^2, \forall (y_1,y_2) \in \mathcal{L}_{\mathbb{R}}^2, \mid f_Y^X(x_1,y_1) - f_Y^X(x_2,y_2) \mid \leq C_x(\parallel x_1 - x_2 \parallel^{b_1} + \mid y_1 - y_2 \mid^{b_2}). \end{cases}$$

$$(H6)\begin{cases} \exists\beta > 0, F_Y^X(x,y) \leq 1 - \beta, \forall (x,y) \in \mathcal{F} \times \mathcal{L}_{\mathbb{R}}, and; \\ \forall (x_1,x_2) \in N_x^2, \forall (y_1,y_2) \in \mathcal{L}_{\mathbb{R}}^2, \mid F_Y^X(x_1,Y_1) - F_Y^X(x_2,y_2) \mid \leq C_x(\parallel x_1 - x_2 \parallel^{b_1} + \mid y_1 - y_2 \mid^{b_2}). \end{cases}$$

3.3.1 Mean Squared Convergence

The result concerns the L^2-consistency $\hat{h}_Y^X(x,y)$.

Theorem 3.3.1 *Under hypotheses (H1)-(H6) and if* $F_Y^X(x,y)$ *(resp.* $f_Y^X(x,y) \in C_B^2(\mathbb{F} \times \mathbb{R})$

$$MSE\hat{h}_Y^X(x,y) \equiv \mathbb{E}\left[(\hat{h}_Y^X(x,y) - h_Y^x(X,y)\right]^2$$

then

$$\equiv B_n(x,y) + \frac{\sigma_h^2(x,y)}{nh_n\phi_x(h_n)} + o(h_h^2) + o(h_K) + o\left(\frac{1}{nh_n\phi_x(h_k)}\right)$$

where

$$B_n(x,y) = \frac{(B_H^f(x,y) - h_Y^X(x,y)B_H^F(x,y))h_H^2 + (B_K^f(x,y) - h_Y^X(x,y)B_K^F(x,y))h_K}{1 - F_Y^X(x,y)},$$

with

$$\begin{aligned}
B_H^f(x,y) &= \tfrac{1}{2}\tfrac{\partial^2 f^X(y)}{\partial y^2}\int t^2 H'(t)dt,\\
B_K^f(x,y) &= \tfrac{\int_{B(0,1)} K(\|v\|)D_x f_Y^X(x,y)[v]g(x,v)d\mu(v)}{\int_{B(0;1)} K(\|v\|)g(x,v)d\mu(v)}\\
B_H^F(x,y) &= \tfrac{1}{2}\tfrac{\partial^2 F_Y^X(x,y)}{\partial y^2}\int t^2 H(t)dt,\\
B_K^F(x,y) &= \tfrac{\int_{B(0,1)} K(\|v\|)D_x F_Y^x(x,y)[v]g(x,v)d\mu(v)}{\int_{B(0,1)} K(\|v\|)g(x,v)d\mu(v)}.
\end{aligned}$$

and $\sigma_h^2(x,y) = \frac{\beta_2 h_Y^X(x,y)}{(\beta_1^2(1-F_Y^X(x,y))}$ *(with* $\beta_j = \int_{B(0,1)} K^j(\| v \|)g(x,v)d\mu(v),\ for, j = 1,2$*).*

Proof. This proof is based on the decomposition

$$\begin{aligned}
\hat{h}_Y^X(x,y) - h_Y^X(x,y) &= \tfrac{\hat{f}_Y^X(x,y)}{1-\hat{F}_Y^X(x,y)} - \tfrac{f_Y^X(x,y)}{1-F_Y^X(X,y)}\\
&= \tfrac{1}{1-\hat{F}_Y^X(x,y)}\left[(\hat{f}_Y^X(x,y) - f_Y^X(x,y)) + \tfrac{f_Y^X(x,y)}{1-F_Y^X(x,y)}(\hat{F}_Y^X(x,y) - F_Y^X(x,y))\right]\\
&= \tfrac{1}{\hat{f}_D^X(x)-\hat{g}_N(x,y)}\left(\hat{f}_N(x,y) - \mathbb{E}\hat{f}_N(x,y)\right)\\
&+ \tfrac{h_Y^X(x,y)}{\hat{f}_D^X(x)-\hat{g}_N(x,y)}\left(\mathbb{E}\hat{g}_N(x,y) - F_Y^X(x,y)\right)\\
&+ \tfrac{1}{\hat{f}_D^X(x)-\hat{g}_N(x,y)}\left(\mathbb{E}\hat{f}_N(x,y) - f_Y^X(x,y)\right)\\
&+ \tfrac{h_Y^X(x,y)}{\hat{f}_D^X(x)-\hat{g}_N(x,y)}\left(1 - \mathbb{E}\hat{g}_N(x,y) - (\hat{f}_D(x) - \hat{g}_N(x,y))\right)
\end{aligned} \tag{5}$$

where D_x means the derivative with respect to x. Hence:

$$\mid \hat{h}_Y^X(x,y) - h_Y^X(x,y) \mid \leq \frac{1}{\mid 1 - \hat{F}_Y^X(x,y) \mid}\left\{\mid \hat{f}_Y^X(x,y) - f_Y^X(X,y) \mid + \mid h_Y^X(x,y)\left(\hat{F}_Y^X(x,y) - F_Y^X(X,y) \mid\right)\right.$$

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

$$\mid \hat{h}_Y^X(x,y) - h_Y^X(x,y) \mid \leq C\frac{\mid \hat{f}_Y^X(x,y) - f_Y^X(X,y) \mid + \mid \hat{F}_Y^X(x,y) - F_Y^X(X,y) \mid}{\mid 1 - \hat{F}_Y^X(x,y) \mid},$$

Then, Theorem (3.3.1) can be deduced from both lemmas above Lemma (3.3.2) and Lemma (3.3.1).

Lemma 3.3.2 *Under hypotheses* $(H1)-(H6)$ *and if* $f_Y^X(X,y) \in C_B^2(\mathbb{F}\times\mathbb{R})$ *then:*

$$\mathbb{E}[\hat{f}_Y^X(x,y)-f_Y^X(X,y)]^2 = B_H^f(x,y)h_H^2+B_K^f(x,y)h_K+\frac{\sigma_f^2(x,y)}{nh_H\phi(h_K)}+o(h_H^2)+\left(\frac{1}{nh_H\phi(h_K)}\right), \quad (6)$$

where

$$\sigma_f^2(x,y) = \frac{(f_Y^X(x,y))\left(\int_{B(0,1)} K^2(\| v \|)g(x,v)d\mu(v)\right)\int H'^2(t)dt}{\left(\int_{B(0,1)} K(\| v \|)g(x,v)d\mu(v)\right)^2},$$

Lemma 3.3.3 *Under hypotheses* $(H1)-(H6)$ *and if* $F_Y^X(x,y) \in C_B^2(\mathbb{F}\times\mathbb{R})$ *then:*

$$\mathbb{E}[(\hat{F}_Y^X(x,y)-F_Y^X(x,y))]^2 = B_H^F(x,y)h_H^2+B_K^F(x,y)h_K+\frac{\sigma_F^2(x,y)}{n\phi(h_K)}+o(h_H^2)+o(h_K)+o(\frac{1}{n\phi(h_K)}), \quad (7)$$

with

$$\sigma_F^2(x,y) = \frac{F_Y^X(x,y)(1-(F_Y^X(x,y))\left(\int_{B(0,1)} K^2(\| v \|)g(x,v)d\mu(v)\right)}{\left(\int_{B(0,1)} K(\| v \|)g(x,v)d\mu(v)\right)^2},$$

Remark 3.3.4 *Observe that, the result of this lemmas Lemma 3.3.2 and Lemma 3.3.3 permits to write*

$$\left[\mathbb{E}\hat{g}_N(x,y) - F_Y^X(x,y)\right] = O(h_H^2) + O(h_K)$$

and

$$\left[\mathbb{E}\hat{f}_N(x,y) - f_Y^X(x,y)\right] = O(h_H^2) + O(h_K).$$

Proof of Lemma (3.3.2) According to the previous decomposition is demonstrated by a separate calculation of both parties, party bias and variance for part two quantities, as the squared error can be expressed as

$$\mathbb{E}[(\hat{f}_Y^X(x,y) - f_Y^X(x,y))^2] = [\mathbb{E}(\hat{f}_Y^X(x,y)) - f_Y^X(X,y)]^2 + Var[\hat{f}_Y^X(x,y)].$$

We define the quantities $K_i(x) = K(h_K^{-1} \parallel x - X_i \parallel), H_i'(y) = H'(h_H^{-1}(y - Y_i))$ for all $i = 1, ..., n$.
We will calculate both sides of this equation (party bias and variance part) to arrive at the calculation of $\mathbb{E}[\hat{f}_Y^X(x,y) - f_Y^X(x,y)]^2$.

We come at the following to writing:

$$\hat{f}_Y^X(x,y) = \frac{\hat{f}_N(x,y)}{\mathbb{E}\hat{f}_D(x)}\left[1 - \frac{\hat{f}_D(x) - \mathbb{E}\hat{f}_D(x)}{\mathbb{E}\hat{f}_D(x)}\right] + \frac{\left(\hat{f}_D(x) - \mathbb{E}\hat{f}_D(x)\right)^2}{(\mathbb{E}\hat{f}_D(x))^2}\hat{f}_Y^X(x,y),$$

from which we draw:

$$\mathbb{E}\hat{f}_Y^X(x,y) = \frac{\mathbb{E}\hat{f}_N(x,y)}{\mathbb{E}\hat{f}_D(x)} - \frac{A_1}{(\mathbb{E}\hat{f}_D(x))^2} + \frac{A_2}{(\mathbb{E}\hat{f}_D(x))^2},$$

as

$$A_1 = \mathbb{E}\hat{f}_N(x,y)(\hat{f}_D(x) - \mathbb{E}\hat{f}_D(x)) = cov(\hat{f}_N(x,y), \hat{f}_D(x))$$

and

$$A_2 = \mathbb{E}(\hat{f}_D(x) - \mathbb{E}\hat{f}_D(x))^2\hat{f}_Y^X(x,y).$$

Can be written as

$$\begin{aligned}
\hat{f}_Y^X(x,y) - f_Y^X(x,y) &= \left(\frac{\hat{f}_N(x,y)}{\mathbb{E}\hat{f}_D(x)} - f_Y^X(x,y)\right) \\
&- \frac{(\hat{f}_N(x,y) - \mathbb{E}\hat{f}_N(x,y))(\hat{f}_D(x) - \mathbb{E}\hat{f}_D(x))}{(\mathbb{E}\hat{f}_D(x))^2} \\
&- \frac{(\mathbb{E}\hat{f}_N(x,y))(\hat{f}_D(x) - \mathbb{E}\hat{f}_D(x))}{(\mathbb{E}\hat{f}_D(x))^2} \\
&+ \frac{(\hat{f}_D(x) - \mathbb{E}\hat{f}_D(x))^2}{(\mathbb{E}\hat{f}_D(x))^2}\hat{f}_Y^X(x,y),
\end{aligned} \tag{8}$$

which implies

$$\begin{aligned}\mathbb{E}[\hat{f}_Y^X(x,y)] - f_Y^X(x,y) &= \Big((\mathbb{E}\hat{f}_D(x))^{-1}\mathbb{E}(\hat{f}_N(x,y)) - f_Y^X(x,y)\Big) - \Big((\mathbb{E}\hat{f}_D(x))^{-2}cov(\hat{f}_N(x,y),\hat{f}_D(y))\Big) \\ &+ (\mathbb{E}\hat{f}_D(x))^{-2}\mathbb{E}\Big(\hat{f}_D(x) - \mathbb{E}\hat{f}_D(x)\Big)^2 \widehat{f_Y^X}(x,y) \\ &= \Big((\mathbb{E}\hat{f}_D(x))^{-1}\mathbb{E}(\hat{f}_N(x,y)) - f_Y^X(x,y)\Big) - \Big(\mathbb{E}\hat{f}_D(x)\Big)^{-2}A_1 + \Big(\mathbb{E}\hat{f}_D(x)\Big)^{-2}A_2\end{aligned}$$

Now you need to write each of these terms and calculate three integrals corresponding to them by a change of variable of type $z = (x-u)/h$. Regarding the term A_2 as the kernel H' is bounded and since K is positive, we can bounded $\hat{f}_Y^X(x,y)$ by a constant $C > 0$, as $\hat{f}_Y^X(x,y) \le C/h_n$, hence

$$\begin{aligned}\mathbb{E}[\hat{f}_Y^X(x,y)] - f_Y^X(x,y) &= \Big((\mathbb{E}\hat{f}_D(x))^{-1}\mathbb{E}(\hat{f}_N(x,y)) - f_Y^X(x,y)\Big) - \Big((\mathbb{E}\hat{f}_D(x))^{-2}cov(\hat{f}_N(x,y),\hat{f}_D(x))\Big) \\ &+ (\mathbb{E}\hat{f}_D(x))^{-2}Var(\hat{f}_D(x))O(h_H^{-1}).\end{aligned}$$

For the par dispersion we inspire techniques Sarda and Vieu [146]and Bosq Lecoutre [21] and by under expression 8, we find that

$$\begin{aligned}Var[\hat{f}_Y^X(x,y)] &= \tfrac{Var[\hat{f}_N(x,y)]}{(\mathbb{E}\hat{f}_D(x))^2} - 2\tfrac{[\mathbb{E}\hat{f}_N(x,y))]cov[\hat{f}_N(x,y),\hat{f}_D(y))]}{(\mathbb{E}\hat{f}_D(x))^3} \\ &+ Var(\hat{f}_D(x))\tfrac{(\mathbb{E}\hat{f}_N(x,y))^2}{(\mathbb{E}\hat{f}_D(x))^4} + O\Big(\tfrac{1}{nh_n\phi(h_n)}\Big).\end{aligned} \tag{9}$$

Finally, Lemma (3.3.2) is a consequence of Corollaries below

Corollary 3.3.5 *Under conditions of Lemma 3.3.2 we have*

$$\frac{\mathbb{E}\hat{f}_N(x,y)}{\mathbb{E}\hat{f}_D(x)} - f_Y^X(x,y) = B_H^f(x,y)h_H^2 + B_K^f(x,y)h_K + O(h_H^2) + O(h_K).$$

Corollary 3.3.6 *Under conditions of Lemma 3.3.2 we have*

$$var[\hat{f}_N(x,y)] = \frac{1}{nh_H\phi(h_K)}\frac{\int_{B(0,1)}K^2(\| v \|)g(x,v)d\mu(v)}{(\int_{B(0,1)}K(\| v \|)g(x,v)d\mu(v))^2}\Big(f_Y^X(x,y)\int H'^2(t)dt\Big) + O\Big(\frac{1}{nh_H\phi(h_K)}\Big)$$

Corollary 3.3.7 *Under conditions of Lemma 3.3.2 we have*

$$cov[\hat{f}_N(x,y),\hat{f}_D(x)] = \frac{1}{n\phi(h_K)}(f_Y^X(x,y))\int_{B(0,1)} K^2(\parallel v \parallel)g(x,v)d\mu(v) + O\left(\frac{1}{n\phi(h_K)}\right)$$

Corollary 3.3.8 *Under conditions of Lemma 3.3.2 we have*

$$Var[\hat{f}_D(x)] = \frac{\int_{B(0,1)} K^2(\parallel v \parallel)g(x,v)d\mu(v)}{n\phi(h_K)} + O\left(\frac{1}{n\phi(h_K)}\right)$$

Proof of Corollary 3.3.5 By definition of $\hat{f}_N(x,y)$ we have

$$\begin{aligned}
\mathbb{E}\hat{f}_N(x,y) &= \tfrac{1}{nh_H\phi(h_K)}\textstyle\sum_{i=1}^n \mathbb{E}(K_i(x)H_i'(y)) \\
&= \tfrac{1}{h_n\phi(h_n)}\mathbb{E}\left[K_1(x)H_1(\tfrac{y-Y_i}{h_H})\right] \qquad (10) \\
&= \tfrac{1}{h_H\phi(h_K)}\mathbb{E}(K_1(x)[\mathbb{E}(H_1'(h_n^{-1}(y-Y_i)|X))])
\end{aligned}$$

for the calculation of $E(H_1'(h_H^{-1}(y-Y_i)|X))$ considering the change of variable $t = h_H^{-1}(y-z)$, we have

$$\begin{aligned}
E(H_1'(h_H^{-1}(y-Y_i)|X)) &= \tfrac{1}{h_H}\textstyle\int H'\left(\tfrac{y-z}{h_H}\right)f^x(z)dz \\
&= \textstyle\int H'(t)f^x(y-h_Ht)dt
\end{aligned}$$

Just develop the function $f_Y^X(y-h_Ht)$ in the neighborhood of y, which is possible since $f_Y^X(x,.)$ being a function of class C^2 in y, then, we can use the Taylor expansion of the function f_Y^X:

$$f_Y^X(y-h_Ht) = f_Y^X(x,y) - h_Ht\frac{\partial f_Y^X(x,y)}{\partial y} + \frac{h_H^2t^2}{2}\frac{\partial^2 f_Y^X(x,y)}{\partial y^2} + o(h_H^2)$$

which gives, under the assumption (H3)

$$\mathbb{E}(H_1'\backslash X) = f_Y^X(x,y) + \frac{h_H^2t^2}{2}\frac{\partial^2 f_Y^X(x,y)}{\partial y^2}\int t^2H'(t)dt + o(h_H^2).$$

We replace in equation (10) found

$$\mathbb{E}\hat{f}_N(x,y) = \frac{1}{h_H\phi(h_K)}\left[\mathbb{E}(K_1(x))f_Y^X(x,y) + \frac{h_H^2t^2}{2}\int t^2H'(t)dt\mathbb{E}\left(K_1(x)\frac{\partial^2 f_Y^X(x,y)}{\partial y^2}\right)\right] + o(h_H^2) \qquad (11)$$

To simplify the writing of this equation we set $\psi_l(.,y) = \frac{\partial^l f_Y^X(x,y)}{\partial y^l}, l \in \{0,2\}$.
The function $\psi_l(.,y)$ defined on the functional space $\mathbb{F}$ denotes the one or other of the two

functions $\psi_0(.,y) = f_Y^X(x,y)$ et $\psi_2(.,y) = \frac{\partial^2 f_Y^X(x,y)}{\partial y^2}$.
The kernel K is assumed compact support, then, for all $l \in \{0,2\}$ we have

$$\begin{aligned} E(K_1\psi_l(X,y)) &= \mathbb{E}K(h_K^{-1} \parallel x-X \parallel)\psi_l(x-h_K(h_K^{-1}(x-X)),y) \\ &= \int_{B(0,1)} K(\parallel v \parallel)\psi_l(x-h_Kv,y)w(h_K,x,v)d\mu(v). \end{aligned}$$

The function $\psi_l(.,y)$ is of class C^1 in the neighborhood of x, then

$$\psi_l(x-h_Kv,y) = \psi_l(x,y) - h_K\frac{\partial\psi_l(x,y)[v]}{\partial x} + o(h_K)$$

and we find that

$$\begin{aligned} E(K_1\psi_l(X,y)) &= \psi_l(x,y)\int_{B(0,1)} K(\parallel v \parallel)w(h_K,x,v)d\mu(v) \\ &- h_K\int_{B(0,1)} K(\parallel v \parallel)\tfrac{\partial\psi_l(x,y)[v]}{\partial x}w(h_K,x,v)d\mu(v) \\ &+ o(h_K)\int_{B(0,1)} K(\parallel v \parallel)w(h_n,x,v)d\mu(v) \end{aligned}$$

Therefore we have

$$\begin{aligned} \mathbb{E}\hat{f}_N(x,y) &= \tfrac{1}{h_H\phi(h_K)}[\psi_0(x,y)\int_{B(0,1)} K(\parallel v \parallel)w(h_K,x,v)d\mu(v) \\ &- h_K\int_{B(0,1)} K(\parallel v \parallel)\tfrac{\partial\psi_0(x,y)[v]}{\partial x}w(h_K,x,v)d\mu(v) \\ &+ \tfrac{h_H^2}{2}\int t^2H'(t)dt(\psi_2(x,y)\int_{B(0,1)} K(\parallel v \parallel)w(h_K,x,v)d\mu(v) \\ &- h_K\int_{B(0,1)} K(\parallel v \parallel)\tfrac{\partial\psi_2(x,y)[v]}{\partial x}w(h_K,x,v)d\mu(v))] + o(h_H^2) + o(h_K). \end{aligned}$$

multiplying by $g(x,v)$, adding and subtracting the two terms

$$\begin{aligned} \mathbb{E}\hat{f}_N(x,y) &= \tfrac{1}{h_n\phi(h_K)}\psi_0(x,y)\int_{B(0,1)} K(\parallel v \parallel)w(h_K,x,v)d\mu(v) \\ &- h_K\int_{B(0,1)} K(\parallel v \parallel)\tfrac{\partial\psi_0(x,y)[v]}{\partial x}g(x,v)d\mu(v) \\ &- h_K\int_{B(0,1)} K(\parallel v \parallel)\tfrac{\partial\psi_0(x,y)[v]}{\partial x}\left(\tfrac{w(h_K,x,v)}{h_H\phi(h_K)} - g(x,v)\right)d\mu(v) \\ &+ \tfrac{h_H^2}{2}\int t^2H'(t)dt[\tfrac{1}{\phi(h_K)}\psi_2(x,y)\int_{B(0,1)} K(\parallel v \parallel)w(h_K,x,v)d\mu(v) \\ &- h_K\int_{B(0,1)} K(\parallel v \parallel)\tfrac{\partial\psi_2(x,y)[v]}{\partial x}g(x,v)d\mu(v) \\ &- h_K\int_{B(0,1)} K(\parallel v \parallel)\tfrac{\partial\psi_2(x,y)[v]}{\partial x}\left(\tfrac{w(h_K,x,v)}{h_H\phi(h_K)} - g(x,v)\right)d\mu(v)] + o(h_H^2+h_k). \end{aligned}$$

Thus

$$\begin{aligned}
\mathbb{E}\hat{f}_N(x,y) &= \tfrac{1}{h_H\phi(h_K)}\psi_0(x,y)\int_{B(0,1)} K(\parallel v \parallel)w(h_K,x,v)d\mu(v)\\
&- h_k\int_{B(0,1)} K(\parallel v \parallel)\tfrac{\partial\psi_0(x,y)[v]}{\partial x}g(x,v)d\mu(v)\\
&+ \tfrac{h_H^2}{2}\int t^2H'(t)dt[\tfrac{1}{h_H\phi(h_K)}\psi_2(x,y)\int_{B(0,1)} K(\parallel v \parallel)w(h_K,x,v)d\mu(v)] + o(h_H^2+h_k).
\end{aligned}$$

On the other hand we have

$$\mathbb{E}\hat{f}_D(x) = \frac{\mathbb{E}K_1}{\phi(h_K)} = \frac{1}{\phi(h_K)}\int_{B(0,1)} K(\parallel v \parallel)w(h_K,x,v)d\mu(v). \tag{12}$$

by substituting in the formula for $\mathbb{E}f_N(x,y)$ it follows that

$$\begin{aligned}
\mathbb{E}f_N(x,y) &= \psi_0(x,y)(\mathbb{E}\hat{f}_D(x)) - h_K\int_{B(0,1)} K(\parallel v \parallel)\tfrac{\partial\psi_0(x,y)[v]}{\partial x}g(x,v)d\mu(v)\\
&+ \tfrac{h_H^2}{2}\int t^2H'(t)dt[(\mathbb{E}\hat{f}_D(x))\psi_2(x,y)] + o(h_H^2) + o(h_K).
\end{aligned}$$

Using the hypothesis (H1), equation (12) can be expressed as

$$\mathbb{E}\hat{f}_D(x) = \int_{B(0,1)} K(\parallel v \parallel)g(x,v)d\mu(v) + o(1) \tag{13}$$

Finally we arrive at

$$\begin{aligned}
(\mathbb{E}\hat{f}_D(x))^{-1}\mathbb{E}[\hat{f}_N(x,y)] - f_Y^X(x,y) &= -h_K\frac{\int_{B(0,1)} K(\|v\|)\frac{\partial f^x(y)[v]}{\partial x}g(x,v)d\mu(v)}{\int_{B(0,1)} K(\|v\|)h(x,v)d\mu(v)}\\
&+ \tfrac{h_H}{2}\tfrac{\partial^2 f^x(y)[v]}{\partial y^2}\int t^2H'(t)dt + o(h_H^2) + o(h_K^2).
\end{aligned} \tag{14}$$

Proof of Corollary 3.3.6 By definition of $\hat{f}_N(x,y)$ we have

$$\begin{aligned}
Var\bigg(\hat{f}_N(x,y)\bigg) &= \tfrac{1}{(n(h_H\phi(h_K))^2}\sum_{i=1}^n Var(K_i(x)H_i'(y))\\
&= \tfrac{1}{n(h_H\phi(h_K))^2}Var(K_1(x)H_1'(x))\\
&= \tfrac{1}{n(h_H\phi(h_K))^2}(\mathbb{E}(K_1(x)H_1'(y))^2 - (\mathbb{E}(K_1(x)H_1'(y))^2)\\
&= \tfrac{1}{n(h_H\phi(h_K))^2}\mathbb{E}(K_1(x)H_1'(y))^2 - n^{-1}\bigg(\tfrac{(\mathbb{E}(K_1(x)H_1'(y)))}{h_H\phi(h_K)}\bigg)^2.
\end{aligned}$$

By Corollary 3.3.5 and equation (13) we have $\frac{(\mathbb{E}K_1(X)H_1'(y))}{h_H\phi(h_K)} = \mathbb{E}\hat{f}_N(x,y) = O(1)$, and the fact that

$$Var\bigg(\hat{f}_N(x,y)\bigg) = \frac{1}{n(h_H\phi(h_K)^2}\mathbb{E}(K_1(x)H_1'(y))^2 + o\bigg(\frac{1}{nh_H\phi(h_K)}\bigg).$$

Just now evaluate the quantity $\mathbb{E}(K_1(x)H_1'(y))^2$. Indeed, the proof is similar to the one used for previous lemma, by conditioning x and considering the usual change of variables $(y-z)/h_H = t$ we obtain

$$\begin{aligned}\mathbb{E}(K_1(x)H_1'(y))^2 &= \mathbb{E}(K_1(x)^2\mathbb{E}(H_1'^2(y)|X)) \\ &= \tfrac{1}{h_H^2}\mathbb{E}\bigg(K_1(x)^2\int H'^2(\tfrac{y-z}{h_H})f^x(z)dz\bigg) \\ &= \tfrac{1}{h_H}\mathbb{E}\bigg(K_1^2(x)\int H'^2(t)f^x(y-h_Ht)dt\bigg),\end{aligned}$$

by a Taylor expansion of the order 1 from y we show that for n large enough

$$f_Y^X(x, y-h_Ht) = f_Y^X(x,y) + o(h_H) = f_Y^X(x,y) + O(1)$$

Hence

$$\mathbb{E}(K_1(x)H_1'(y))^2 = \frac{1}{h_H}\int H'^2(t)dt\mathbb{E}(K_1^2(x)f_Y^X(x,y)) + o\bigg(\frac{1}{h_H}\bigg)$$

The same way and with the same techniques used in the above proof of Corollary 3.3.5, we show that it suffices now to estimate the amount $\mathbb{E}(K_1(x)H_1'(y))^2$. Indeed, for a demonstration similar to the proof lemma, in conditioning by X and considering the usual change of variable $(y-z)/h_H = t$ we find that:

$$\begin{aligned}\mathbb{E}(K_1^2(x)f_Y^X(x,y)) &= \mathbb{E}K^2(h_K^{-1}\parallel x-X\parallel)f(x-h_n(h_K^{-1}(x-X)),y) \\ &= \textstyle\int_{B(0,1)}K^2(\parallel v\parallel)f_Y^X(x-h_Kv,y)w(h_K,x,v)d\mu(v) \\ &= \phi(h_n)f_Y^X(x,y)\textstyle\int_{B(0,1)}K^2(\parallel v\parallel)g(x,y)d\mu(v) + O(\phi(h_K)).\end{aligned}$$

such that $\parallel v\parallel = h_K^{-1}\parallel x-X\parallel$, this allows us to conclude

$$\mathbb{E}(K_1(x)H_1'(y))^2 = \frac{1}{h_H}\int H'^2(t)dt\bigg(\phi(h_K)f_Y^X(x,y)\int_{B(0,1)}K^2(\parallel v\parallel)g(x,y)d\mu(v)\bigg) + O\bigg(\frac{\phi(h_K)}{h_H}\bigg).$$

The hypothesis (H3) implies that the kernel H is square summable, therefore

$$Var\bigg(\hat{f}_N(x,y)\bigg) = \frac{1}{(n(h_H\phi(h_K))}\bigg[f_Y^X(x,y)\int H'^2(t)dt\int_{B(0,1)}K^2(\parallel v\parallel)g(x,y)d\mu(v)\bigg] + O\bigg(\frac{1}{n(h_H\phi(h_K)}\bigg)$$

Proof of Corollary 3.3.7: By definition of $\hat{f}_N(x,y)$ and $\hat{f}_D(x)$ we obtain

$$\begin{aligned} Cov(\hat{f}_N(x,y),\hat{f}_D(x)) &= \tfrac{1}{n(h_H\phi(h_K))^2}Cov(K_1(x)H_1'(y),K_1(x)) \\ &= \tfrac{1}{n(h_H\phi(h_K))^2}(\mathbb{E}K_1^2(x)H_1'(y)-\mathbb{E}K_1(x)H_1'(y)\mathbb{E}K_1(x)) \\ &= \tfrac{\mathbb{E}K_1^2(x)H_1'(y)}{n(h_H\phi(h_K))^2}-\left(\tfrac{\mathbb{E}K_1(x)H_1'(y)}{n(h_H\phi(h_K))^2}\right)\left(\tfrac{\mathbb{E}K_1(x))}{n(h_H\phi(h_K))^2}\right) \end{aligned}$$

The proof of this Corollary is very similar to the one used for Corollary 3.3.5. To do this, replace K_1^2 with K_1 then using the fact that $\frac{(\mathbb{E}K_1(X)H_1(y))}{\phi(h_K)}=O(1)$ and $\frac{(\mathbb{E}K_1(X))}{\phi(h_K)}=O(1)$ we deduce that

$$Cov(\hat{f}_N(x,y),\hat{f}_D(x))=\frac{1}{n\phi(h_K)}(f_Y^X(x,y))\int_{B(0,1)}K^2(\parallel v\parallel)g(x,y)d\mu(v)+O\left(\frac{1}{n\phi(h_K)}\right). \quad (15)$$

Proof of Corollary 3.3.8: By definition of $\hat{f}_D(x)$ we have

$$\begin{aligned} Var(\hat{f}_D(x)) &= \tfrac{1}{n(\phi(h_K))^2}(Var(K_1)) \\ &= \tfrac{\mathbb{E}K_1^2(x)}{n(\phi(h_K))^2}-n^{-1}\left(\tfrac{\mathbb{E}K_1(x)}{\phi(h_K)}\right) \\ &= \tfrac{\int_{B(0,1)}K^2(\|v\|)g(x,v)d\mu(v)}{n(\phi(h_K))}+O\left(\tfrac{1}{n\phi(h_K)}\right) \end{aligned} \quad (16)$$

.

This allows us to complete the proof of Lemma 3.3.2.

Proof of Lemma 3.3.3: The calculation of the squared error of the conditional distribution is with the same techniques used in the previous lemma 3.3.2 by a separate calculation of two parts: part bias and some variance for the two quantities, as the squared error the conditional distribution can be expressed as

$$\mathbb{E}[(\widehat{F}_Y^X(x,y)-F_Y^X(x,y))^2]=[\mathbb{E}(\widehat{F}_Y^X(x,y))-F_Y^X(x,y)]^2+Var[\hat{F}_Y^X(x,y)].$$

Finally, Lemma 3.3.3 can be deduced from following corollaries

Corollary 3.3.9 *Under hypotheses (H1)-(H6), we have*

$$\frac{\mathbb{E}\hat{g}_N(x,y)}{\mathbb{E}\hat{f}_D(x)}-F_Y^X(x,y)=B_H^F(x,y)h_H^2+B_K^F(x,y)h_K+o(h_H^2)+o(h_K),$$

Corollary 3.3.10 *Under hypotheses (H1)-(H6), we have*

$$Var[\hat{g}_N(x,y)]=\frac{\int_{B(0,1)}K^2(\parallel v\parallel)g(x,v)d\mu(v)}{n\phi(h_K)}\left(F_Y^X(x,y)\int H^2(t)dt\right)+O\left(\frac{1}{n\phi(h_K)}\right),$$

Corollary 3.3.11 *Under hypotheses (H1)-(H6), we have*

$$Cov[\hat{g}_N(x,y), \hat{f}_D(x)] = \frac{1}{n\phi(h_K)}(F_Y^X(x,y)) \int_{B(0,1)} K^2(\| v \|)g(x,v)d\mu(v) + O\left(\frac{1}{n\phi(h_K)}\right)$$

Remark 3.3.12 *It is clear that, the results of Corollaries Corollary 3.3.6-3.3.8 and Corollary 3.3.10-3.3.11 allows to write*

$$Var[\hat{f}_D(x) - \hat{g}_N(x,y)] = o\left(\frac{1}{nh_H\phi_x(h_K)}\right).$$

3.3.2 Asymptotic Normality

This section contains results on the asymptotic normality of $\widehat{h}_Y^X(x,y)$.

Theorem 3.3.13 *Assume that (H1)-(H6) hold, and if the following equation* $(\star)$ *is verified, then we have for any* $x \in \mathcal{A}$,

$$(\frac{nh_H\phi_x(h_K)}{\sigma_h^2(x,y)})^{1/2}(\widehat{h}_Y^X(x,y) - h_Y^X(x,y) - B_n(x,y)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad n \to \infty,$$

where

$$A = \{x \in \mathcal{F}, f_Y^X(x,y)(1 - F_Y^X(x,y)) \neq 0\},$$

and $\xrightarrow{\mathcal{D}}$ *means the convergence in distribution.*

Evidently, if one imposes some additional assumptions on the function $\phi_x(\cdot)$ and the bandwidth parameters (h_K and h_H) our asymptotic normality can be improved by removing the bias term $B_n(x,y)$.

Corollary 3.3.14 *Under the hypotheses of Theorem 3.3.13 and if the bandwidth parameters* (h_K *and* h_H) *and if the function* $\phi_x(h_K)$ *satisfies:*

$$lim_{n\to\infty}(h_H^2 + h_K)\sqrt{n\phi_x(h_K)} = 0,$$

we have

$$(\frac{nh_H\phi_x(h_K)}{\sigma_h^2(x,y)})^{1/2}(\widehat{h}_Y^X(x,y) - h_Y^X(x,y)) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad as \quad n \to \infty,$$

Proof Consider the decomposition

$$\widehat{h}_Y^X(x,y) - h_Y^X(x,y) = \frac{1}{\widehat{f}_D(x) - \widehat{g}_N(x,y)}(\widehat{f}_N(x,y) - \mathbb{E}\widehat{f}_N(x,y))$$

$$+\frac{1}{\widehat{f}_D(x) - \widehat{g}_N(x,y)}\{h_Y^X(x,y)(\mathbb{E}\widehat{g}_N(x,y) - F_Y^X(x,y)) + (\mathbb{E}\widehat{f}_N(x,y) - f_Y^X(x,y))\}$$

$$+\frac{h_Y^X(x,y)}{\widehat{f}_D(x) - \widehat{g}_N(x,y)}\{1 - \mathbb{E}\widehat{g}_N(x,y)(\widehat{f}_D(x) - \widehat{g}_N(x,y))\}$$

Therefore, Theorem 3.3.13 and Corollary 3.3.14 are a consequence of Lemma 3.3.15, Lemma 3.3.16 and the following results.

Lemma 3.3.15 *Under the hypotheses of Theorem 3.3.13*

$$(\frac{nh_H\phi_x(h_K)}{\sigma_f^2(x,y)})^{1/2}(\widehat{f}^N(x,y)-\mathbb{E}\widehat{f}^N(x,y))\to\mathcal{N}(0,1)$$

Lemma 3.3.16 *Under the hypotheses of Theorem 3.3.13*

$$\widehat{F}_D(x)-\widehat{g}_N(x,y)\longrightarrow 1-F_Y^X(x,y)\quad in\quad probability,$$

and

$$(\frac{nh_H\phi_x(h_K)}{\sigma_h^2(x,y)})^{1/2}(\widehat{f}_D(x)-\widehat{g}_N(x,y)-1+\widehat{g}_N(x,y))=OP(1).$$

Proof of Lemma 3.3.15 Define

$$\Gamma_i(x,y)=\frac{\sqrt{\phi_x(h_K)}}{\sqrt{nh_H}\mathbb{E}[K_1(x)]}(\Delta_i(x,y)-\mathbb{E}[\Delta_i(x,y)]),$$

and

$$\Omega_n=\sum_{i=1}^n\Gamma_i(x,y).$$

Thus

$$\Omega_n=\sqrt{nh_H\phi_x(h_K)}(\widehat{f}_N(x,y)-\mathbb{E}\widehat{f}_N(x,y))$$

So, our claimed result is now

$$\Omega_n\to\mathcal{N}(0,\sigma_f^2(x,y)).$$

Therefore, we have

$$Var(\Omega_n)=nh_H\phi_x(h_K)Var(\widehat{f}^N(x,y)-\mathbb{E}[\widehat{f}_N(x,y)])$$

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

$$Var(\widehat{f}_N(x,y))=\frac{1}{(nh_H\mathbb{E}[K_1(x)])^2}\sum_{i=1}^n\sum_{j=1}^n Cov(\Delta_i(x,y),\Delta_j(x,y))=\frac{1}{n(h_H\mathbb{E}[K_1(x)])^2}Var(\Delta_1(x,y)).$$

Therefore

$$Var(\Delta_1(x,y))\leq\mathbb{E}(H_1^{'2}(y)K_1^2(x))\leq\mathbb{E}(K_1^2(x)\mathbb{E}[H_1^{'2}|X_1])$$

Now, by a change of variable in the following integral and by applying (H3) and (H5), one gets

$$\begin{aligned}\mathbb{E}[H_1^{'2}|X_1] &= \int_{\mathbb{R}}H^{'2}\left(\frac{\|y-u\|}{h_H}\right)f_Y^X(x,u)du\\ &\leq h_H\int_{\mathbb{R}}H^{'2}(t)(f_Y^X(y-h_Ht,x)-f_Y^X(x,y))dt+h_Hf_Y^X(x,y)\int_{\mathbb{R}}H^{'2}(t)dt\\ &\leq h_H^{1+b_2}\int_{\mathbb{R}}|t|^{b_2}H^{'2}(t)dt+h_Hf_Y^X(x,y)\int_{\mathbb{R}}H^{'2}(t)dt\\ &= h_H\left(o(1)+f_Y^X(x,y)\int_{\mathbb{R}}H^{'2}(t)dt\right)\end{aligned}\qquad(17)$$

By means of (17) and the fact that, as $n \to \infty, \mathbb{E}(K_1^2(x)) \to \beta_2 \phi_x(h_K)$ one gets

$$Var(\Delta_1(x,y)) = \beta_2 \phi_x(h_H) h_H \left(o(1) + f_X^Y(x,y) \int_{\mathbb{R}} H'^2(t)dt \right)$$

So, using (H4), we get

$$\frac{1}{n(h_H \mathbb{E}[K_1(x)])^2} Var(\Delta_1(x,y)) = \frac{\beta_2 \phi_x(h_K)}{n(\beta_1 h_H \phi_x(h_K))^2} h_H \left(o(1) + f_Y^X(x,y) \int_{\mathbb{R}} H'^2(t)dt \right)$$

$$= O\left(\frac{1}{nh_H \phi_x(h_K)} \right) + \frac{\beta_2 f_Y^X(x,y)}{\beta_1^2 nh_H \phi_x(h_K)} \int_{\mathbb{R}} H'^2(t)dt$$

Thus as $n \to \infty$ we obtain

$$\frac{1}{(nh_H \mathbb{E}[K_1(x)])^2} Var(\Delta_1(x,y)) \to \frac{\beta_2 f_Y^X(x,y)}{\beta_1^2 nh_H \phi_x(h_K)} \int_{\mathbb{R}} H'^2(t)dt. \tag{18}$$

Finally, the proof of Lemma is completed by using result (18), to get

$$Var(\Omega_n) \to \frac{\beta_2}{\beta_1^2} f_Y^X(x,y) \int_{\mathbb{R}} H'^2(t)dt = \sigma_f^2(x,y)$$

Proof of Lemma 3.3.16

It is clear that, the result of Corollary 3.3.6, Corollary 3.3.8 and Corollary 3.3.10 permits us

$$\mathbb{E}(\widehat{F}_D(x) - \widehat{g}_N(x,y) - 1 + F_Y^X(x,y)) \to 0,$$

and

$$Var(\widehat{F}_D(x) - \widehat{g}_N(x,y) - 1 + F_Y^X(x,y)) \to 0,$$

then

$$\widehat{F}_D(x) - \widehat{g}_N(x,y) - 1 + F_Y^X(x,y) \xrightarrow{\mathcal{P}} 0.$$

Moreover, the asymptotic variance of $\widehat{f}_D(x) - \widehat{g}_N(x,y)$ given in Remark 3.3.12 allows to obtain

$$\frac{nh_H \phi_x(h_K)}{\sigma_h(x,y)^2} Var(\widehat{f}_D(x,y) - \widehat{g}_N(x,y) - 1 + \mathbb{E}(\widehat{g}_N(x,y))) \to 0.$$

By combining result with the fact that

$$\mathbb{E}(\widehat{f}_D(x) - \widehat{g}_N(x,y) - 1 + \mathbb{E}(\widehat{g}_N(x,y))) = 0.$$

Finally, wee obtain the claimed result.

3.4 Remarques and Commentary

1. The hypothesis (H1) on the functional variable X can be divided into two parts:

   - (i)The first part is rarely used in non-parametric statistical functional, because it requires the introduction of a reference measurement of the functional space. However, in this chapter the objective that we impose this condition. In other words, it allows us to achieve a natural generalization of the squared error obtained by Vieu [158] in the vector case. The hypothesis (H1) is not very restrictive. Indeed, the first part of this hypothesis is verified, when, for example X is a diffusion process satisfying standard conditions (see Niang [40]).
   - (ii) The second part (5) is less restrictive than the following condition, given for all$(r, v) \in \mathbb{R}_*^+ \times B(0, 1)(x fixed)$:

     $$\exists C_1, C_2 > 0, 0 < C_1\phi(r)g(x, v) \leq w(r, x, v) \leq C_2\phi(r)g(x, v),$$

     which is a classic property in functional analysis. Note that, this assumption is used to describe the phenomenon of concentration of the probability measure of the explanatory variable X, since we have:

     $$\mathbb{P}(X \in B(x, r)) = \int_{B_{(0;1)}} w(r, x, v)d\mu(v) = \phi(r) \int_{B(0,1)} g(x, v)d\mu(v) + o(\phi(r)) > 0.$$

     This is a simple asymptotic separation of variables. This condition is designed to be able to adapt traditional techniques of the case if different multi functional, even if the reference measure μ does not have the same properties of the Lebesgue measure, such as translation invariance and homogeneity. In the case of finite dimension, the hypothesis (H1) is satisfied when the density of the explanatory variable X is of class C^1 and strictly positive. Indeed, the density of $Z = r^{-1}(x - X)$ and $w(r, x, v) = r^p f(x - rv)$, where f is the density of X and p dimension, therefore $w(r, x, v) = r^p f(x) + o(r^p)$.

2. In this chapter, we chose a condition of derivability as our goal is to find an expression for the rate of convergence explicitly, asymptotically exact and keeps the usual form of the squared error (see Vieu [158]). However, if one proceeds by a Lipschitz condition for example the conditional density of type:

   $$\forall(y_1, y_2) \in \mathcal{S}_\mathbb{R} \times \mathcal{S}_\mathbb{R}, \forall(x_1, x_2) \in N_x \times N_x, |f^{x_1}(y_1) - f^{x_2}(y_2)| \leq A_x((d(x_1, x_2)^2) + |y_1 - y_2|^2),$$

   which is less restrictive than the condition (2), we obtain a result for the conditional distribution and conditional density respectively for example of type:

$$\mathbb{E}[(\widehat{F}_Y^X(x,y) - F_Y^X(x,y))^2] = \mathcal{O}(h_H^4 + h_K^4) + \mathcal{O}\left(\frac{1}{n\phi(h_k)}\right),$$

$$\mathbb{E}[(\widehat{f}_Y^X(x,y) - f_Y^X(x,y))^2] = \mathcal{O}(h_H^4 + h_K^4) + \mathcal{O}\left(\frac{1}{nh_H\phi(h_k)}\right),$$

But such an expression (implicitly) the rate of convergence will not allow us to properly determine the smoothing parameter. In other words, this condition of differentiability is a good compromise to obtain an explicit expression for the rate of convergence. Note that this condition is often taken in the case of finite dimension.

3. The dimensionality of the observations (resp. model) is used in the expression of the rate of convergence of the two lemmas Lemma 3.3.2 and Lemma 3.3.3. We find the "dimensionality" of the model in the way, while the "dimensionality" of the variable in the functional dispersion bias the property of concentration of the probability measure of the functional variable which is closely related to the topological structure of the functional space of the explanatory variable. Ours asymptotique results highlights the importance of the concentration properties on small balls of the probability measure of the underlying functional variable. This highlights the role of semi-metric the quality of our estimate. A suitable choice of this parameter allows us to an interesting solution to the problem of curse of dimensionality. (see [63]). Another argument has a dramatic effect in our estimation. This is the smoothing parameter h_K (resp. h_H). The term of our rate of convergence, decomposed into two main parts: part bias proportional to h_K (resp. h_H), and part dispersion inversely proportional to h_K (resp. h_H)(f is an increasing function depending on the h_K), makes this relatively easy choice minimizing the main part of this expression to determine this parameter.

Chapter 4

Functional explanatory variable in single functional index

The work given in this chapter has been published in Journal of Applied Mathematics and Statistics

In this chapter we deal with nonparametric estimate of the conditional hazard function, when the covariate is functional. Kernel type estimators for the conditional hazard function of a scalar response variable Y given a Hilbertian random variable X are introduced, where 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. The rates of convergence emphasize the crucial role played by the small ball probabilities with respect to the distribution of the explanatory functional variable.

4.1 Setting the Problem

4.1.1 Bibliographic 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:

$$h(x) = \lim_{\triangle x \to 0} \frac{P(X \leq x + \triangle x | X \geq x)}{\triangle x} \qquad (x > 0) \tag{1}$$

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:

$$h(x) = \frac{f(x)}{S(x)} \quad for \quad all \quad x \quad such \quad that \quad F(x) < 1, \tag{2}$$

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 and the main issue is to estimate the conditional random rate defined as

$$h^Z(x) = \lim_{\triangle x \to 0} \frac{P(X \leq x + \triangle x | X \geq x, Z)}{\triangle x} \qquad for (x > 0)$$

which can be written naturally as follows:

$$h^Z(x) = \frac{f^Z(x)}{S^Z(x)}, once F^Z(x) < 1. \tag{3}$$

Study of functions h and h^Z 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 . One of the most common techniques for building estimators of h (respectively h^Z) is based on (2) (resp. (3)) and consist in studying a quotient between the estimator of f (respectively f^Z) and that of S (respectively, S^Z). Patil et al. [125] presented an overview of these estimation techniques. Nonparametric methods based on the ideas of the convolution kernel, which are known for their good behavior in density estimation (conditional or not) problems are widely used in nonparametric estimation of hazard function. A wide range of literature in this area is provided by bibliographic reviews Singpurwalla and Wong [148] Hassani et al. [91], Izenman [101], Gefeller and Michels [79], Pascu and Vaduva [124], and Ferraty et al.

4.1.2 Conditional hazard in the case of explanatory functional

The progress of data collection methods offers opportunities for statisticians to provide increasingly observations of functional variables. Works of Ramsay and Silverman [131] and Ferraty and Vieu [69] offer a wide range of statistics methods, parametric or nonparametric, recently developed to treat various estimation problems which occur in functional random variables (ie with values in a space of in..inite dimension). Until now such statistical developments for functional variables in single functional index does not exist in the context of estimating a hazard function.

Let $(X_i, Z_i)_{1\leq i\leq n}$ be n random variables, identically distributed as the random pair (X, Z) with values in $\mathbb{R}\times\mathcal{H}$, where $\mathcal{H}$ is a separable real Hilbert space with the norm $\|.\|$ generated by an inner product $<.,.>$. We consider the semi metric d_θ, associated to the single index $\theta\in\mathcal{H}$,defined by $\forall z_1, z_2\in\mathcal{H} : d_\theta(z_1, z_2) := |<z_1 - z_2, \theta>|$. Under such topological structure and for a fixed functional θ, we suppose that the conditional hazard function of X given $Z = z$ denoted by $h^z(.)$ exists and is given by

$$\forall x\in\mathbb{R}, h^z_\theta(x) =: h(x| <z, \theta>).$$

Clearly, the identifiability of the model is assured, and we have for all $z\in\mathcal{H}$

$$h_1(.| <z, \theta_1>) = h_2(.| <z, \theta_2>) \Longrightarrow h_1 \equiv h_2 \quad and \quad \theta_1 = \theta_2.$$

For more details see Ait Saidi et al. [3]. In the following, we denote by $h(\theta, ., Z)$, the conditional hazard function of X given $<z, \theta>$.

The objective of this chapter is to study a model in which the conditional random explanatory variable Z is not necessarily real or multi-dimensional but only assumed to be values in an abstract space $\mathcal{H}$ provided a scalar product $<.,.>$. As with any problem of non-parametric estimation, the dimension of the space $\mathcal{H}$ plays an important role in the properties of concentration of the variable X. Thus, when the dimension is not necessarily finite, probability functions defined by small balls of:

$$\phi_{\theta,z}(h) = \mathbb{P}(Z\in B_\theta(z, h)) = \mathbb{P}(Z\in\{z'\in\mathcal{H}, 0 < |<z - z', \theta>| < h\}),$$

intervene directly in the asymptotic behavior of any functional non-parametric estimator (see Ferraty and Vieu [69]). The asymptotic results presented later in this chapter on the estimation of the function $h(\theta, x, Z)$ does not escape this rule.

From now, z denotes a fixed element of the functional space $\mathcal{H}, \mathcal{N}_z$ denotes a fixed neighborhood of z and $\mathcal{S}_{\mathbb{R}}$ is a fixed compact of $\mathbb{R}_+$. Now, we should make some assumptions on the concentration function $\phi_{\theta,z}(h)$

- (H1) $\forall h > 0, \phi_{\theta,z}(h) > 0$

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

- (H2) $\exists A_{\theta,z} < \infty, \exists b_1, b_2 > 0, \forall (x_1, x_2) \in S_{\mathbb{R}}^2, \forall (z_1, z_2) \in N_z^2$:

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

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

- (H3) $\exists \nu < \infty, \forall (x, z') \in S_{\mathbb{R}} \times N_z, f(\theta, x, z') \leq \nu$
- (H4) $\exists \beta > 0, \forall (x, z') \in S_{\mathbb{R}} \times N_z, F(\theta, x, z') \leq 1 - \beta$

4.1.3 Construction of the estimator in the case of non-censured data

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

Recent advances in non-parametric statistics for functional variables, as presented in Ferraty and Vieu [69] show that the techniques based on convolution kernels are easily transposed to the context of functional variables. Moreover, these kernel's techniques have good properties in the problems of estimation of hazard function when the variables are of finite-dimensional. The reader may consult the work Ferraty et al. [69] which is a pioneering paper on the subject and that of Quintela-del-Rio [127] for the most recent results in this area.

Therefore, drawing on these ideas, it is natural to try to construct an estimator of the function $h(\theta, ., Z)$. To estimate the conditional distribution function and the conditional density in the presence of functional the variable Z, Mahiddine et al. [111] proposed the following functional kernel estimators:

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

and

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

where K is a kernel,H is a distribution function and $h_K = h_{K,n}(resp.h_H = h_{H,n})$ is a sequence of positive real numbers a kernel estimator of the functional conditional hazard function $h(\theta, ., Z)$ may therefore be constructed in the following way:

$$\hat{h}(\theta, x, Z) = \frac{\hat{f}(\theta, x, Z)}{1 - \hat{F}(\theta, x, Z)} \quad (4)$$

The assumptions we need later for the parameters of the estimator, ie on K, H, h_H and h_K are not restrictive.

Indeed, on one hand, they are not specific to the problem of estimating $h(\theta, x, Z)$ (but rather inherent to the estimation problems of $F(\theta, x, Z)$ and $f(\theta, x, Z)$, and in other hand they correspond to the assumptions usually made in the context of non-functional variables. More precisely, we introduce the following conditions which guarantee the good behavior of the estimators $\hat{F}(\theta, x, Z)$ and $\hat{f}(\theta, x, Z)$ (see Ferraty and Vieu [69]):

- $(H5)$ H is a bounded Lipschitz continuous function, such that

$$\int H'(t)dt = 1, \int |t|^{b_2} H(t)dt < \infty \quad \text{and} \quad \int H^2(t)dt < \infty$$

- $(H6)$ K is positive bounded function with support $[-1, 1]$.
- $(H7)$ The bandwidth h_K has to satisfy

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

- $(H8)$ The bandwidth h_H has to satisfy

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

Under these general conditions, we will establish in 4.3.1 the pointwise convergence of the kernel estimator $\widehat{h}(\theta, x, z)$ of the functional conditional hazard function $h(\theta, x, z)$ when the observed sample is not censored. In section 4.3.2, these results will be generalized to censored variables.

4.1.4 Estimation in censored case

Estimation of the hazard function when the data are censored is an important problem in medical research. So, in practice, in medical applications, it can be in the presence of variables censored. This problem is usually modeled by considering a positive variable called C, and the observed random variables are not the couples (X_i, Z_i) but only the (T_i, Δ_i, Z_i) where $T_i = \text{minimize}(X_i, C_i)$ and $\Delta_i = I_{X_i \le C_i}$. In the following we use the notations $F_1(\theta, ., Z)$ and $f_1(\theta, ., Z)$ to describe the distribution function and conditional density of C knowing Z and we use the notation $S_1(\theta, ., Z) = 1 - F_1(\theta, ., Z)$. Models such censorship where abundantly studied in the literature for real or multi-dimensional random variables, and in the nonparametric case kernel's techniques are particularly used (see Tanner and Wong [150] Padgett [123] Lecoutre and Ould-Said [108] and van Keilegom Veraverbeke [102]), for functional variables see Ferraty et al., and Laksaci and Mechab [107] in the case of spatial variables.

The aim of this section, is to adapt these ideas as part of an explanatory variable Z functional, and build a kernel estimator function type of conditional random $h(\theta, ., Z)$ adapted to the censored data. If we introduce the notation $L(\theta, ., Z) = 1 - S_1(\theta, ., Z)S(\theta, ., Z)$ and $\varphi(\theta, ., Z) = f(\theta, ., Z)S_1(\theta, ., Z)$, we can reformulate the expression (3) as follows:

$$h(\theta, t, Z) = \frac{\varphi(\theta, t, Z)}{1 - L(\theta, t, Z)}, \forall t, L(\theta, t, Z) < 1. \tag{5}$$

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

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

and

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

Finally the hazard function estimator is given as:

$$\widetilde{h}(\theta, t, Z) = \frac{\widehat{\varphi}(\theta, t, Z)}{1 - \widehat{L}(\theta, t, Z)}. \tag{6}$$

In addition to the assumptions introduced in section 4.2.3, we need additional conditions. These assumptions are identical to those found in the classical literature for non-functional variables (see previous references), these additional hypotheses are as follows:

- $(H9)$ Conditionally to Z, the variables X and C are independent;
- $(H10)$ $\exists A_{\theta,z} < \infty, \exists b_1, b_2 > 0, \forall (t_1, t_2) \in \mathcal{S}_{\mathbb{R}}^2, \forall (z_1, z_2) \in \mathcal{N}_z^2$:

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

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

- $(H11)$ $\exists \mu < \infty, \varphi(\theta, t, z') < \mu, \forall (t, z') \in \mathbb{R}_+ \times \mathcal{N}_z$,
- $(H12)$ $\exists \eta > 0, L(\theta, t, z') \le 1 - \eta, \forall (t, z') \in \mathbb{R}_+ \times \mathcal{N}_z$.

Under these very general conditions, we establish in Section 4.3.1 the rates of convergence of the kernel estimator $\widetilde{h}(\theta, ., z)$ of the functional conditional Hazard function $h(\theta, ., z)$ when couples of variables $(X_i, Z_i)_{i=1,...n}$ are independents. In section 4.3.2 these results will be generalized by dispensing with the condition of censored data.

4.2 Pointwise Almost Complete Convergence

4.2.1 Case of non censored data

We begin by studying statistical samples satisfying a classical assumption of independence, couples (X_i, Z_i) are iid

Theorem 4.2.1 *Under hypotheses (H1)-(H8), we have:*

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

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

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

hence

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

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

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

And conventionally (see for instance the Proposition A6ii of Ferraty and Vieu [69]) the announced result follows directly from the following properties:

$$\sup_{x\in\mathcal{S}_\mathbb{R}} |F(\theta,x,z) - \hat{F}(\theta,x,z)| = \mathcal{O}(h_K^{b_1} + h_H^{b_2}) + \mathcal{O}_{a.c.o}\left(\sqrt{\frac{logn}{n\phi_{\theta,z}(h_K)}}\right) \qquad (7)$$

and

$$\sup_{x\in\mathcal{S}_\mathbb{R}} |f(\theta,x,z)-\hat{f}(\theta,x,z)| = \mathcal{O}(h_K^{b_1}+h_H^{b_2}) + \mathcal{O}_{a.c.o}\left(\sqrt{\frac{logn}{nh_H\phi_{\theta,z}(h_K)}}\right) \tag{8}$$

and from the next result which is a consequence of property (7).

Corollary 4.2.2 *Under the conditions of Theorem 4.3.1, we have*

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

The results (7) and (8) are known results (see for instance Ferraty and Vieu [69], Propositions 6.19 and 6.20).

4.2.2 Estimation with censored data

The goal now is to take these asymptotic properties in the broader context of a censored sample as described in Section 4.2.4. We will begin in this section by discussing the case censored. Obviously, obtaining these results require more sophisticated than those presented under uncensored technical developments. To ensure a good readability in this Section 4.3.2, the presentation of these technical details will later in Paragraph 5. We begin by studying statistical samples satisfying a standard assumption of independence, ie. triples (X_i, C_i, Z_i) are i.i.d.

Theorem 4.2.3 *Under assumptions* $(H1)-(H2), and (H5)-(H12)$, *we have:*

$$\sup_{t\in\mathcal{S}_\mathbb{R}} |\tilde{h}(\theta,t,z)-h(\theta,t,z)| = \mathcal{O}(h_K^{b_1}+h_H^{b_2}) + \mathcal{O}_{a.c.o}\left(\sqrt{\frac{logn}{nh_H\phi_{\theta,z}(h_K)}}\right)$$

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

$$\begin{aligned}
\sup_{t\in\mathcal{S}_\mathbb{R}} |\tilde{h}(\theta,t,z)-h(\theta,t,z)| &\leq \frac{1}{\inf_{t\in\mathcal{S}_\mathbb{R}}|1-\hat{L}(\theta,t,z)|}\left\{\sup_{t\in\mathcal{S}_\mathbb{R}}|\tilde{\varphi}(\theta,t,z)-\varphi(\theta,t,z)|\right\} \\
&\quad + \frac{\sup_{t\in\mathcal{S}_\mathbb{R}}|h(\theta,t,z)|}{\inf_{t\in\mathcal{S}_\mathbb{R}}|1-\hat{L}(\theta,t,z)|}\left\{\sup_{t\in\mathcal{S}_\mathbb{R}}|\tilde{L}(\theta,t,z)-L(\theta,t,z)|\right\} \qquad (9) \\
&\leq C\sup_{t\in\mathcal{S}_\mathbb{R}}\frac{\left\{|\tilde{\varphi}(\theta,t,z)-\varphi(\theta,t,z)|+|L(\theta,t,z)-\tilde{L}(\theta,t,z)|\right\}}{\inf_{t\in\mathcal{S}_\mathbb{R}}|1-\hat{L}(\theta,t,z)|}
\end{aligned}$$

which is obtained from (3) and (5) proceeding as to establish (17). Since $\hat{L}(\theta,t,Z)$ is none other than the kernel estimator of the conditional distribution function of T knowing Z is obtained directly (see Ferraty and Vieu [69], Proposition 6.19) that:

$$\sup_{t\in\mathcal{S}_{\mathbb{R}}} |\tilde{L}(\theta,t,z) - L(\theta,t,z)| = \mathcal{O}(h_K^{b_1} + h_H^{b_2}) + \mathcal{O}_{a.c.o}\left(\sqrt{\frac{logn}{n\phi_{\theta,z}(h_K)}}\right) \tag{10}$$

The proprieties of the estimator $\hat{\varphi}(\theta,.,Z)$ are given in Lemma 4.3.4, the desired result is obtained directly from (9)-(12).

Lemma 4.2.4 *Under hypotheses of theorem 4.3.3, we have:*

$$\sup_{t\in\mathcal{S}_{\mathbb{R}}} |\tilde{\varphi}(\theta,t,z) - \varphi(\theta,t,z)| = \mathcal{O}(h_K^{b_1} + h_H^{b_2}) + \mathcal{O}_{a.c.o}\left(\sqrt{\frac{logn}{nh_H\phi_{\theta,z}(h_K)}}\right) \tag{11}$$

The next result which is a consequence of property (10).

Corollary 4.2.5 *Under the conditions of Theorem 4.3.3, we have*

$$\exists\delta > 0 \quad \textit{such that} \quad \sum_{n=1}^{\infty} \mathbb{P}\{\inf_{x\in\mathcal{S}_{\mathbb{R}}} |1 - \hat{L}(\theta,x,z)| < \delta\} < \infty. \tag{12}$$

4.3 Uniform Almost Complete Convergence

In this party we derive the uniform version of Theorem 4.3.1. The study of the uniform consistency is motivated by the fact that the latter is an indispensable tool for studying the asymptotic properties of all estimates of the functional index if is unknown. Noting that, in the multivariate case, the uniform consistency is a standard extension of the pointwise one, however, in our functional case, it requires some additional tools and topological conditions (see Ferraty et al. [63], for more discussion on the uniform convergence in nonparametric functional statistics). Thus, in addition to the conditions introduced previously, we need the following ones. Firstly, consider

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

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

4.3.1 Non censored data

Thereafter we propose to study the uniform almost complete convergence of our estimator defined above (4) for this, we need the following assumptions:

- $(A1)$ There exists a differentiable function $\phi(.)$ such that $\forall z \in \mathcal{S}_{\mathcal{H}}$ and for all $\theta \in \Theta_{\mathcal{H}}$,

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

- $(A2)$ $\forall (x_1, x_2) \in \mathcal{S}_{\mathbb{R}} \times \mathcal{S}_{\mathbb{R}}, \qquad \forall (z_1, z_2) \in \mathcal{S}_{\mathbb{H}} \times \mathcal{S}_{\mathbb{H}} \quad and \quad \forall \theta \in \Theta_{\mathcal{H}}$

$$|F(\theta, x_1, z_1) - F(\theta, x_2, z_2)| \leq A(\|z_1, z_2\|^{b_1} + |x_1 - x_2|^{b_2})$$

$$|f(\theta, x_1, z_1) - f(\theta, x_2, z_2)| \leq A(\|z_1, z_2\|^{b_1} + |x_1 - x_2|^{b_2})$$

- $(A3)$ $\exists \nu < \infty, \forall (x, z') \in \mathcal{S}_{\mathbb{R}} \times \mathcal{N}_z, \forall \theta \in \Theta_{\mathcal{H}}, f(\theta, x, z') \leq \nu$
- $(A4)$ $\exists \beta > 0, \forall (x, z') \in \mathcal{S}_{\mathbb{R}} \times \mathcal{N}_z, \forall \theta \in \Theta_{\mathcal{H}}, F(\theta, x, z') \leq 1 - \beta$
- $(A5)$ The kernel K satisfy (H3) and Lipschitz's condition holds

$$|K(u) - K(v)| \leq C\|u - v\|,$$

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

$$\frac{(logn)^2}{n\phi(h_K)} < logd_n^{\mathcal{S}_{\mathcal{H}}} + logd_n^{\Theta_{\mathcal{H}}} < \frac{n\phi(h_K)}{logn}$$

and

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

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

$$\frac{(logn)^2}{nh_H\phi(h_K)} < logd_n^{\mathcal{S}_{\mathcal{F}}} + logd_n^{\Theta_{\mathcal{F}}} < \frac{nh_H\phi(h_K)}{logn}$$

and

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

Remark 4.3.1 *Note that Assumptions (A1)-(A4) are, respectively, the uniform version of (H1)-(H4). Assumptions (A1) and (A6) are linked with the the topological structure of the functional variable, see Ferraty et al. [58].*

Theorem 4.3.2 *Under hypotheses (A1)-(A7) and (H5), we have:*

$$\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{z\in\mathcal{S}_{\mathcal{H}}} \sup_{x\in\mathcal{S}_{\mathbb{R}}} |\hat{h}(\theta, x, z) - h(\theta, x, z)| = \mathcal{O}(h_K^{b_1} + h_H^{b_2}) + \mathcal{O}_{a.c.o}\left(\sqrt{\frac{logd_n^{\mathcal{S}_{\mathcal{H}}} + logd_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right)$$

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

Corollary 4.3.3 *Under Assumptions (A1)-(A7) and (H4), as n goes to infinity, we have*

$$\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathbb{R}}}|\hat{h}(\theta,x,z)-h(\theta,x,z)| = \mathcal{O}(h_K^{b_1}+h_H^{b_2})+\mathcal{O}_{a.c.o}\left(\sqrt{\frac{logd_n^{\mathcal{S}_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right)$$

Proof of theorem 4.4.2. Clearly The proofs of these two results namely the Theorem 4.4.2 and Corollary 4.4.3 can be deduced from the following intermediate results which are only uniform version of properties (7) and (8).

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathbb{R}}}|\hat{F}(\theta,x,z)-F(\theta,x,z)| = \mathcal{O}(h_K^{b_1}+h_H^{b_2})+\mathcal{O}_{a.c.o}\left(\sqrt{\frac{logd_n^{\mathcal{S}_{\mathcal{H}}}+logd_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right) \quad (13)$$

and

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{x\in\mathcal{S}_{\mathbb{R}}}|\hat{f}(\theta,x,z)-f(\theta,x,z)| = \mathcal{O}(h_K^{b_1}+h_H^{b_2})+\mathcal{O}_{a.c.o}\left(\sqrt{\frac{logd_n^{\mathcal{S}_{\mathcal{H}}}+logd_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right) \quad (14)$$

from the next result which is a consequence of property (13).

Corollary 4.3.4 *Under the conditions of Theorem 4.4.2, we have*

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

The results (13) and (14) are known results (see for example Mahiddine et al. [111]).

4.3.2 Censored data

Thereafter we propose to study the uniform almost complete convergence of our estimator defined above (6) for this, we need the following assumptions:

- *(A2a)* $\forall(t_1,t_2)\in\mathcal{S}_{\mathbb{R}}\times\mathcal{S}_{\mathbb{R}}$ and $\forall(z_1,z_2)\in\mathcal{S}_{\mathcal{H}}\times\mathcal{S}_{\mathcal{H}}$ and $\forall\theta\in\Theta_{\mathcal{H}}$,
$$|L(\theta,t_1,z_1)-L(\theta,t_2,z_2)|\le A(\|z_1,z_2\|^{b_1}+|t_1-t_2|^{b_2}),$$
$$|\varphi(\theta,t_1,z_1)-\varphi(\theta,t_2,z_2)|\le A(\|z_1,z_2\|^{b_1}+|t_1-t_2|^{b_2});$$
- *(A3a)* $\exists\nu<\infty,\forall(t,z')\in\mathcal{S}_{\mathbb{R}}\times\mathcal{N}_z,\forall\theta\in\Theta_{\mathcal{H}},\varphi(\theta,t,z')\le\nu$,
- *(A4a)* $\exists\beta>0,\forall(t,z')\in\mathcal{S}_{\mathbb{R}}\times\mathcal{N}_z,\forall\theta\in\Theta_{\mathcal{H}},L(\theta,t,z')\le1-\beta$.

Theorem 4.3.5 *Under hypotheses (A1), (A5)-(A7) and (A2a)-(A4a), we get:*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\tilde{h}(\theta,t,z)-h(\theta,t,z)| = \mathcal{O}(h_K^{b_1}+h_H^{b_2})+\mathcal{O}_{a.c.o}\left(\sqrt{\frac{logd_n^{\mathcal{S}_{\mathcal{H}}}+logd_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right)$$

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

Corollary 4.3.6 *Under Assumptions (A1), (A5)-(A7), (A2a)-(A4a) and (H4), as n goes to infinity, we have*

$$\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\tilde{h}(\theta,t,z)-h(\theta,t,z)| = \mathcal{O}(h_K^{b_1}+h_H^{b_2})+\mathcal{O}_{a.c.o}\left(\sqrt{\frac{logd_n^{\mathcal{S}_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right)$$

Proof of theorem 4.4.5. The result is based on the decomposition (9). Clearly The proofs of these two results namely the Theorem 4.4.5 and Corollary 4.4.6 can be deduced from the following intermediate results which are only uniform version of properties (10) and (11). The properties of the estimators $\hat{L}(\theta,.,z)$ and $\hat{\varphi}(\theta,.,z)$ are given in the following Lemma 4.7. Finally, the desired result is obtained directly from (9), (15), (16).

Lemma 4.3.7 *Under hypotheses of Theorem 4.4.5, we have:*

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\hat{L}(\theta,t,z)-L(\theta,t,z)| = \mathcal{O}(h_K^{b_1}+h_H^{b_2})+\mathcal{O}_{a.c.o}\left(\sqrt{\frac{logd_n^{\mathcal{S}_{\mathcal{H}}}+logd_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right) \quad (15)$$

and

$$\sup_{\theta\in\Theta_{\mathcal{H}}}\sup_{z\in\mathcal{S}_{\mathcal{H}}}\sup_{t\in\mathcal{S}_{\mathbb{R}}}|\hat{\varphi}(\theta,t,z)-\varphi(\theta,t,z)| = \mathcal{O}(h_K^{b_1}+h_H^{b_2})+\mathcal{O}_{a.c.o}\left(\sqrt{\frac{logd_n^{\mathcal{S}_{\mathcal{H}}}+logd_n^{\Theta_{\mathcal{H}}}}{nh_H\phi(h_K)}}\right) \quad (16)$$

The next result which is a consequence of property (15).

Corollary 4.3.8 *Under the conditions of Theorem 4.4.5, we have*

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

Sketch of Proof of Lemma 4.4.7

The proof of (15) is based on some results depending on the following decomposition;

$$\begin{aligned}\hat{L}(\theta,t,z)-L(\theta,t,z) &= \frac{1}{\hat{\varphi}_D(\theta,z)}\left\{\left(\hat{L}_N(\theta,t,z)-\mathbb{E}\hat{L}_N(\theta,t,z)\right)-\left(L(\theta,t,z)-\mathbb{E}\hat{L}_N(\theta,t,z)\right)\right\}\\ &+ \frac{L(\theta,t,z)}{\hat{\varphi}_D(\theta,z)}\{1-\hat{\varphi}_D(\theta,z)\}\end{aligned} \quad (17)$$

Then the rest of the proof is similar the one given in Mahiddine and al.[111], where, it is sufficient to replace $\hat{F}_D(\theta,z), F(\theta,t,z)$ and $\mathbb{E}(\hat{F}_N(\theta,t,z))$ (Lemma 6, corollary 3 and Lemma 7) by $\hat{\varphi}_D(\theta,z), L(\theta,t,z)$ and $\mathbb{E}(\hat{L}_N(\theta,t,z))$ respectively.

Then the rest is deduced directly from Lemma 4.4.10, Lemma 4.4.11 and Corollary 4.4.9.

Corollary 4.3.9 *Under Assumptions (A1), (A5) and (A6), we have as $n \to \infty$*

$$\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{z\in\mathcal{S}_{\mathcal{H}}} |\hat{\varphi}_D(\theta,z) - 1| = \mathcal{O}_{a.c.o}\left(\sqrt{\frac{logd_n^{\mathcal{S}_{\mathcal{H}}} + logd_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right) \tag{18}$$

and

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{\theta\in\Theta_{\mathcal{H}}} \inf_{z\in\mathcal{S}_{\mathcal{H}}} \hat{\varphi}_D(\theta,z) < \frac{1}{2}\right) < \infty. \tag{19}$$

Lemma 4.3.10 *Under Assumptions (A1), (A2) and (H5), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{z\in\mathcal{S}_{\mathcal{H}}} \sup_{t\in\mathcal{S}_{\mathbb{R}}} |L(\theta,t,z) - \mathbb{E}(\hat{L}_N(\theta,t,z))| = \mathcal{O}(h_K^{b_1} + h_H^{b_2}) \tag{20}$$

Lemma 4.3.11 *Under assumptions (A1), (A5)-(A7) and (A2a)-(A4a) we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{H}}} \sup_{z\in\mathcal{S}_{\mathcal{H}}} \sup_{t\in\mathcal{S}_{\mathbb{R}}} |\hat{L}_N(\theta,t,z) - \mathbb{E}[\hat{L}_N(\theta,t,z)]| = \mathcal{O}_{a.c.o}\left(\sqrt{\frac{logd_n^{\mathcal{S}_{\mathcal{H}}} + logd_n^{\Theta_{\mathcal{H}}}}{n\phi(h_K)}}\right) \tag{21}$$

• Concerning (16) the proof is based at first on the following decomposition;

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

The rest is deduced directly from Lemma 4.4.12, Lemma 4.4.13 and Corollary 4.4.9.

Lemma 4.3.12 *Under Assumptions (A1), (A2a) and (H5), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{F}}} \sup_{z\in\mathcal{S}_{\mathcal{F}}} \sup_{t\in\mathcal{S}_{\mathbb{R}}} |\varphi(\theta,t,z) - \mathbb{E}(\hat{\varphi}_N(\theta,t,z))| = \mathcal{O}(h_K^{b_1} + h_H^{b_2}) \tag{22}$$

Lemma 4.3.13 *Under the assumptions (A1), (A5) ,(A2a), (A7) and (H5), we have, as n goes to infinity*

$$\sup_{\theta\in\Theta_{\mathcal{F}}} \sup_{z\in\mathcal{S}_{\mathcal{F}}} \sup_{t\in\mathcal{S}_{\mathbb{R}}} |\hat{\varphi}_N(\theta,t,z) - \mathbb{E}[\hat{\varphi}_N(\theta,t,z)]| = \mathcal{O}_{a.c.o}\left(\sqrt{\frac{logd_n^{\mathcal{S}_{\mathcal{F}}} + logd_n^{\Theta_{\mathcal{F}}}}{nh_H\phi_{\theta,z}(h_K)}}\right) \tag{23}$$

4.4 Proof of technical lemmas

In what follows C and c denote generic strictly positive real constants. Furthermore, the following notation are introduced:

$$K_i(\theta,z) = K(h_K^{-1}(< z - Z_i, \theta >)), H_i^{'}(t) = H^{'}(h_H^{-1}(t - T_i)),$$

$$\hat{\varphi}_N(\theta,t,z) = \frac{1}{nh_H \mathbb{E}K_1(\theta,z)} \sum_{i=1}^{n} K_i(\theta,z) H_i^{'}(t)\Delta_i,$$

$$\hat{\varphi}_D(\theta,z) = \frac{1}{n\mathbb{E}K_1(z)} \sum_{i=1}^{n} K_i(\theta,z),$$

$$V_i = \frac{1}{\mathbb{E}K_1(\theta,z)} K_i(\theta,z),$$

$$W_i = \frac{1}{h_H \mathbb{E}K_1(\theta,z)} K_i(\theta,z) H_i^{'}(t)\Delta_i,$$

Proof of corollary 4.3.2. It is clear that

$$\begin{aligned} \inf_{x\in S_{\mathbb{R}}} |1 - \hat{F}(\theta,x,z)| &\leq \Big(1 - \sup_{x\in S_{\mathbb{R}}} F(\theta,x,z)\Big)/2 \\ &\Rightarrow \sup_{x\in S_{\mathbb{R}}} |\hat{F}(\theta,x,z) - F(\theta,x,z)| \geq \Big(1 - \sup_{x\in S_{\mathbb{R}}} F(\theta,x,z)\Big)/2. \end{aligned}$$

Which implies that

$$\begin{aligned} &\sum_{n=1} \mathbb{P}\Big\{ \inf_{x\in S_{\mathbb{R}}} |1 - \hat{F}(\theta,x,z)| \leq \Big(1 - \sup_{x\in S_{\mathbb{R}}} F(\theta,x,z)\Big)/2 \Big\} \\ &\leq \sum_{n=1} \mathbb{P}\Big\{ \sup_{x\in S_{\mathbb{R}}} |\hat{F}(\theta,x,z) - F(\theta,x,z)| \geq \Big(1 - \sup_{x\in S_{\mathbb{R}}} F(\theta,x,z)\Big)/2 \Big\} < \infty. \end{aligned}$$

We deduce from property (7) that

$$\sum_{n=1} \mathbb{P}\Big\{ \inf_{x\in S_{\mathbb{R}}} |1 - \hat{F}(\theta,x,z)| \leq \Big(1 - \sup_{x\in S_{\mathbb{R}}} F(\theta,x,z)\Big)/2 \Big\} < \infty.$$

This proof is achieved by taking $\delta = (1 - \sup_{x\in S_{\mathbb{R}}} F(\theta,x,z))/2$ which is strictly positive.

Proof of lemma 4.3.4. By using the following decomposition:

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

and under the proposition A6ii de Ferraty and Vieu [69], the result of lemma 4.3.4 will result directly following three properties:

$$| \hat{\varphi}_D(\theta, z) - 1 |= \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{nh_H\phi_{\theta,z}(h_K)}}\right), \tag{24}$$

$$\sup_{t\in\mathcal{S}_\mathbb{R}} | \mathbb{E}\hat{\varphi}_N(\theta, t, z) - \varphi(\theta, t, z) |= \mathcal{O}(h_K^{b_1} + h_H^{b_2}), \tag{25}$$

and

$$\frac{1}{\hat{\varphi}_D(z)} \sup_{t\in S_\mathbb{R}} | \hat{\varphi}_N(\theta, t, z) - \mathbb{E}\hat{\varphi}_N(\theta, t, z) |= \mathcal{O}_{a.co.}\left(\sqrt{\frac{\log n}{nh_H\phi_{\theta,z}(h_K)}}\right). \tag{26}$$

- ***Proof of (24).*** It suffices to note that we can write

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

with

$$| V_i |= \mathcal{O}\left(\frac{1}{\phi_{\theta,z}(h)}\right), \tag{27}$$

and

$$\mathbb{E}V_i^2 = \mathcal{O}\left(\frac{1}{\phi_{\theta,z}(h)}\right). \tag{28}$$

By applying an exponential inequality bonded variables (for example corollary A9i of Ferraty and Vieu [69])and taking into account the results (27) and (28), we arrive at

$$\mathbb{P}\left[| \hat{\varphi}_D(\theta, z) - \mathbb{E}\hat{\varphi}_D(\theta, z) |> \varepsilon\sqrt{\frac{\log n}{n\phi_{\theta,z}(h_K)}}\right] = \mathcal{O}\left(n^{-C\varepsilon^2}\right).$$

Now simply choose ε large enough to get the results (24).

- ***Proof of (25)***. We have, for any $t\epsilon S_\mathbb{R}$:

$$\begin{aligned}\mathbb{E}\hat{\varphi}_N(\theta, t, z) &= \frac{1}{h_H\mathbb{E}K_1(\theta, z)}\mathbb{E}(K_1(\theta, z)H_1^{'}(t)\Delta_1) \\ &= \frac{1}{h_H\mathbb{E}K_1(\theta, z)}\mathbb{E}(K_1(\theta, z)\mathbb{E}(H_1^{'}(t)I_{X_1\leq C_1} |< Z_1, \theta >)) \\ &= \frac{1}{h_H\mathbb{E}K_1(\theta, z)}\mathbb{E}(K_1(z)\mathbb{E}(H_1(t)S_1(\theta, X_1, Z_1) |< Z_1, \theta >)),\end{aligned} \tag{29}$$

the last equality arising of conditional independence between $C_1 and X_1$ introduced into [H9], furthermore we have

$$\begin{aligned}
\mathbb{E}(H_1(t)S_1(\theta,X_1,z) \mid < Z_1,\theta >) &= \int H^{'}\big(\frac{t-u}{h_H}\big)S_1(\theta,u,Z_1)f(\theta,u,Z_1)du \\
&= h_H \int H^{'}(v)\varphi(\theta,t-vh_H,Z_1)dv \\
&= h_H(\varphi(\theta,t,z)+o(h_H^{b_2}+h_K^{b_1})),
\end{aligned} \tag{30}$$

the last equality resulting from the property of Lipschitz function φ^z introduced in [H10] and the fact that $H^{'}$ is probability density. It should be noted, again because of the condition [H10], that them o() involved in result (30) are uniform for $t\epsilon S_{\mathbb{R}}$. Thus, the result (25) is an immediate consequence of (29) and (30).

• ***Proof of [26].*** The compactness of the set $S_{\mathbb{R}}$ can be covered the u_n disjoint intervals as follows:

$$S_{\mathbb{R}} \subset \bigcup_{m=1}^{u_n} [\tau_m - l_n, \tau_m + l_n[,$$

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

$$\exists C>0, \exists \alpha>0, l_n = Cu_n^{-1} = n^{-\alpha}. \tag{31}$$

For each $t \in S_{\mathbb{R}}$ noting τ_t the single τ_m such as $t \in [\tau_m - l_n, \tau_m + l_n[$. Finally, (26) can be easily deduced from the following results:

$$\frac{1}{\hat{\varphi}_D(\theta,z)} \sup_{t\in S_{\mathbb{R}}} \mid \hat{\varphi}_N(\theta,t,z) - \hat{\varphi}_N(\theta,\tau_t,z) \mid = \mathbb{O}_{a.co.}\Bigg(\sqrt{\frac{\log n}{nh_H\phi_{\theta,z}(h_K)}}\Bigg), \tag{32}$$

$$\frac{1}{\hat{\varphi}_D(\theta,z)} \sup_{t\in S_{\mathbb{R}}} \mid \mathbb{E}\hat{\varphi}_N(\theta,t,z) - \mathbb{E}\hat{\varphi}_N(\theta,\tau_t,z) \mid = \mathbb{O}_{a.co.}\Bigg(\sqrt{\frac{\log n}{nh_H\phi_{\theta,z}(h_K)}}\Bigg), \tag{33}$$

and

$$\frac{1}{\hat{\varphi}_D(\theta,z)} \sup_{t\in S_{\mathbb{R}}} \mid \hat{\varphi}_N(\theta,\tau_t,z) - \mathbb{E}\hat{\varphi}_N(\theta,\tau_t,z) \mid = \mathbb{O}_{a.co.}\Bigg(\sqrt{\frac{\log n}{nh_H\phi_{\theta,z}(h_K)}}\Bigg), \tag{34}$$

• ***Proof of (32).*** Because of the condition [H5], there is exist a finite constant C such that for all $t \in S_{\mathbb{R}}$

$$\begin{aligned}
\mid \hat{\varphi}_N(\theta,t,z) - \hat{\varphi}_N(\theta,\tau_t,z) \mid &= \tfrac{1}{nh_H\mathbb{E}K_1(\theta,z)} \textstyle\sum_{i=1}^n \Delta_i K_i(z)(H_i^{'}(t) - H_i^{'}(\tau_t)) \\
&\leq \tfrac{C}{nh_H\mathbb{E}K_1(\theta,z)} \textstyle\sum_{i=1}^n K_i(\theta,z)\frac{|t-\tau_t|}{h_H} \\
&\leq C\hat{\varphi}_D(\theta,z)l_nh_H^{-2}.
\end{aligned} \tag{35}$$

By using (31) and choosing α large enough, we obtain directly (32).

Proof of (33). This result is obtained directly from (24) and (35) using proposition A6ii of Ferraty and Vieu [69].

• ***Proof of (34).*** Obtaining (34) is based on the use of an exponential inequality. Specifically, it suffices to note that we can write

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

with

$$| W_i |= \mathbb{O}\left(\frac{1}{h_H \phi_{\theta,z}(h)}\right), \tag{36}$$

and

$$\begin{aligned}
\mathbb{E}W_i^2 &= \frac{1}{h_H^2(\mathbb{E}K_1(\theta,z))^2}\mathbb{E}K_i^2(\theta,z)H_i^{'2}(t)\Delta_i^2 \\
&\leq C\frac{1}{h_H^2(\mathbb{E}K_1(\theta,z))^2}\mathbb{E}(K_i^2(\theta,z)\mathbb{E}(H_i^{'2}(t))| < Z_i, \theta >) \\
&\leq C\frac{1}{h_H \phi_{\theta,z}(h)^2}\mathbb{E}(K_i^2(\theta,z)\int \frac{1}{h_H}H^{'}\left(\frac{t-u}{h_H}\right)^2 f(\theta,u,Z_i)du) \\
&= \mathcal{O}\left(\frac{1}{h_H\phi_{\theta,z}(h)}\right).
\end{aligned} \tag{37}$$

By using the condition (31) we arrive at

$$\begin{aligned}
&\mathbb{P}\left[\sup_{x\in S_{\mathbb{R}}} | \quad \hat{\varphi}_N(\theta,\tau_t,z) - \mathbb{E}\hat{\varphi}_N(\theta,\tau_t,z) |> \varepsilon\sqrt{\frac{\log n}{nh_H\phi_{\theta,z}(h_K)}}\right] \\
&\leq n^{\alpha}\max_{m=1,\ldots u_n}\mathbb{P}\left[| \hat{\varphi}_N(\theta,\tau_m,z) - \mathbb{E}\hat{\varphi}_N(\theta,\tau_m,z) |> \varepsilon\sqrt{\frac{\log n}{nh_H\phi_{\theta,z}(h_K)}}\right].
\end{aligned} \tag{38}$$

Moreover, by applying an exponential inequality to bounded variables (for example the corollary A9i by Ferraty and Vieu [69]) and taking into account the result (36)and (37), we arrive at

$$\mathbb{P}\left[|\hat{\varphi}_N(\theta,\tau_m,z) - \mathbb{E}\hat{\varphi}_N(\theta,\tau_m,z)| > \varepsilon\sqrt{\frac{\log n}{nh_H\phi_{\theta,z}(h_K)}}\right] = \mathcal{O}(n^{-C\varepsilon^2}). \tag{39}$$

It suffices now to choose ϵ large enough to directly obtain the desired result from (38) and from (39).
The results (32), (33) and (34) are sufficient to conclude the proof of the result (26).

Finally, lemma 4.3.4 is a consequence of (24), (25) and (26) and decomposition (5).

Proof of corollary 4.4.9.

• Concerning (18) for all $z \in S_H$ and $\theta \in \Theta_H$, we set

$K(z) = \arg\text{minimize}_{k\in\{1\dots r_n\}}\|z - z_k\|$ and $j(\theta) = \arg\text{minimize}_{j\in\{1\dots l_n\}}\|\theta - t_j\|$.

Let us consider the following decomposition

$$\begin{aligned}
\sup_{\theta\in S_H}\sup_{\Theta\in S_H} \mid \hat{\varphi}_D(\theta,z) - \mathbb{E}(\hat{\varphi}_D(\theta,z)) \mid \leq & \underbrace{\sup_{\theta\in S_H}\sup_{\Theta\in S_H} \mid \hat{\varphi}_D(\theta,z) - (\hat{\varphi}_D(\theta,z_{k(z)})) \mid}_{\Pi_1} \\
& + \underbrace{\sup_{\theta\in S_H}\sup_{\Theta\in S_H} \mid \hat{\varphi}_D(\theta,z_{k(z)}) - \hat{\varphi}_D(t_{j(\theta)},z_{k(z)}) \mid}_{\Pi_2} \\
& + \underbrace{\sup_{\theta\in S_H}\sup_{\Theta\in S_H} \mid \hat{\varphi}_D(t_{j(\theta)},z_{k(z)}) - \mathbb{E}(\hat{\varphi}_D(t_{j(\theta)},z_{k(z)})) \mid}_{\Pi_3} \\
& + \underbrace{\sup_{\theta\in S_H}\sup_{\Theta\in S_H} \mid \mathbb{E}(\hat{\varphi}_D(t_{j(\theta)},z_{k(z)})) - \mathbb{E}(\hat{\varphi}_D(\theta,z_{k(z)})) \mid}_{\Pi_4} \\
& + \underbrace{\sup_{\theta\in S_H}\sup_{\Theta\in S_H} \mid \mathbb{E}(\hat{\varphi}_D(\theta,z_{k(z)})) - \mathbb{E}(\hat{\varphi}_D(\theta,z)) \mid}_{\Pi_5}
\end{aligned}$$

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

$$\Pi_1 = \mathcal{O}\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{n\phi(h_k)}}\right), \quad \Pi_2 = \mathcal{O}\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{n\phi(h_k)}}\right) \tag{40}$$

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

$$\Pi_4 = \mathcal{O}\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{n\phi(h_k)}}\right), \quad \Pi_5 = \mathcal{O}\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{n\phi(h_k)}}\right) \tag{41}$$

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

$$\mathbb{P}\left(\Pi_3 > n\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{n\phi(h_k)}}\right)\right)$$

$$\leq d_n^{S_H} d_n^{\Theta_H} \max_{k\in\{1\ldots d_n^{S_H}\}} \max_{j\in\{1\ldots d_n^{\Theta_H}\}} \mathbb{P}\left(\mid \hat{\varphi}_D(t_{j(\theta)}, z_{k(z)}) - \mathbb{E}(\hat{\varphi}_D(t_{j(\theta)}, z_{k(z)})) \mid > n \left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{n\phi(h_k)}} \right) \right).$$

Applying Bernstein's exponential inequality to

$$\frac{1}{\phi(h_K)}(K_i(t_{j(\theta)}, z_{k(z)}) - \mathbb{E}(K_i(t_{j(\theta)}, z_{k(z)}))),$$

Then under [A7], we get

$$\Pi_3 = \mathcal{O}\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{n\phi(h_k)}} \right).$$

Lastly the result will be easily deduced from the later together with (40) and (41).
• Concerning (19) it easy to see that,

$$\begin{aligned} \inf_{\theta\in\Theta_{\mathbb{H}}} \inf_{z\in S_{\mathbb{H}}} \mid \hat{\varphi}_D(\theta, z) \mid \quad &\leq \quad 1/2 \Longrightarrow \exists z \in S_H, \exists\theta \in \Theta_H, \text{such that} \\ & 1 - \hat{\varphi}_D(\theta, z) \geq 1/2 \Longrightarrow \sup_{\theta\in\Theta_{\mathbb{H}}} \sup_{z\in S_{\mathbb{H}}} \mid 1 - \hat{\varphi}_D(\theta, z) \mid \leq 1/2. \end{aligned}$$

We deduce from (18) the following inequality

$$\mathbb{P}\left(\inf_{\theta\in\Theta_{\mathbb{H}}} \inf_{z\in S_{\mathbb{H}}} \mid \hat{\varphi}_D(\theta, z) \mid \leq 1/2 \right) \leq \mathbb{P}\left(\sup_{\theta\in\Theta_{\mathbb{H}}} \sup_{z\in S_{\mathbb{H}}} \mid 1 - \hat{\varphi}_D(\theta, z) \mid \geq 1/2. \right)$$

Consequently,

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\inf_{\theta\in\Theta_{\mathbb{H}}} \inf_{z\in S_{\mathbb{H}}} \hat{\varphi}_D(\theta, z) < 1/2 \right) < \infty$$

Proof of lemma 4.4.7.
• Concerning [20], one has

$$\begin{aligned} \mathbb{E}\hat{L}_N(\theta, t, z) - L(\theta, t, z) &= \tfrac{1}{\mathbb{E}K_1(z,\theta)} \mathbb{E}\left[\sum_{i=1}^n K_i(z, \theta) H_i(t)\right] - L(\theta, t, z) \\ &= \tfrac{1}{\mathbb{E}K_1(z,\theta)} \mathbb{E}(K_1(z, \theta)[\mathbb{E}(H_1(t) | < Z_1, \theta >) - L(\theta, t, z)]). \end{aligned} \quad (42)$$

Moreover, we have

$$\mathbb{E}(H_1(t) | < Z_1, \theta >) = \int_{\mathbb{R}} H(h_H^{-1}(t - z)) f(\theta, z, Z_1) dz,$$

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

$$\mathbb{E}(H_1(t)| < Z_1, \theta >) = \int_{\mathbb{R}} H^{'}(t) L(\theta, t - h_H t, Z_1) dt.$$

Thus, we have

$$| \mathbb{E}(H_1(t)| < Z_1, \theta >) - L(\theta, t, z) | \leq \int_{\mathbb{R}} H^{(1)}(t) \mid L(\theta, t - h_H t, Z_1) - L(\theta, t, z) \mid dt.$$

Finally, the use of [A2] implies that

$$| \mathbb{E}(H_1(t)| < Z_1, \theta >) - L(\theta, t, z) | \leq C \int_{\mathbb{R}} H^{'}(t)(h_K^{b_1} + \mid t \mid^{b_2} h_H^{b_2}) dt. \tag{43}$$

Because this inequality is uniform on $(\theta, t, z) \in \Theta_H \times S_H \times S_{\mathbb{R}}$ and because of [H5], [20] is a direct consequence of [42], [43] and [19].

Concerning [21], we keep the notation of the corollary 4.4.9 and we use the compact of $S_{\mathbb{R}}$, we can write that, for some, $t_1, \ldots.., t_{u_n} \in S_{\mathbb{R}}, S_{\mathbb{R}} \subset \bigcup_{m=1}^{u_n} (t_m - l_n, t_m + l_n)$ with $L_n = n^{-\frac{1}{2b_2}}$ and $u_n \leq C n^{\frac{1}{2b_2}}$. Taking $m(t) = \arg minimize_{\{1,2,\ldots,u_n\}} \mid t - t_m \mid$. Thus, we have the following decomposition:

$$\begin{aligned}
| \hat{L}_N(\theta, t, z) - \mathbb{E}(\hat{L}_N(\theta, t, z)) | \quad \leq \quad & \underbrace{| \hat{L}_N(\theta, t, z) - \hat{L}_N(\theta, t, z_{k(z)}) |}_{\Gamma_1} \\
+ \quad & \underbrace{| \hat{L}_N(\theta, t, z_{k(z)}) - \mathbb{E}(\hat{L}_N(\theta, t, z_{k(z)})) |}_{\Gamma_2} \\
+ \quad & 2\underbrace{| \hat{L}_N(t_{j(\theta)}, t, z_{k(z)}) - \hat{L}_N(t_{j(\theta)}, t_{m(t)}, z_{k(z)}) |}_{\Gamma_3} \\
+ \quad & 2\underbrace{| \mathbb{E}(\hat{L}_N(t_{j(\theta)}, t, z_{k(z)}) - \mathbb{E}(\hat{L}_N(t_{j(\theta)}, t_{m(t)}, z_{k(z)})) |}_{\Gamma_4} \\
+ \quad & \underbrace{| \mathbb{E}(\hat{L}_N(\theta, t, z_{k(z)}) - \mathbb{E}(\hat{L}_N(\theta, t, z) |)}_{\Gamma_5}
\end{aligned}$$

$\hookrightarrow$ Concerning Γ_1 we have

$$| \hat{L}_N(\theta, t, z) - \hat{L}_N(\theta, t, z_{k(z)}) | \leq \frac{1}{n} \sum_{i=1}^{n} | \frac{1}{\mathbb{E}K_1(\theta, z)} K_i(\theta, z) H_i(t) - \frac{1}{\mathbb{E}K_1(\theta, z_{K(z)})} K_i(\theta, z_{K(z)}) H_i(t) | .$$

We use the Hölder continuity condition on K, the Cauchy-Schwartz inequality, the Bernsttein's inequality and the boundedness of H [assumption[H5]]. This allows us to get:

$$\begin{aligned} \mid \hat{L}_N(\theta,t,z)-\hat{L}_N(\theta,t,z_{k(z)})\mid &\leq \frac{C}{\phi(h_k)}\frac{1}{n}\sum_{i=1}^n \mid K_i(\theta,z)H_i(t)-K_i(\theta,z_{K(z)})H_i(t)\mid \\ &\leq \frac{C}{\phi(h_k)}\frac{1}{n}\sum_{i=1}^n \mid H_i(t)\mid\mid K_i(\theta,z)-K_i(\theta,z_{K(z)})\mid \\ &\leq \frac{C' r_n}{\phi(h_k)} \end{aligned}$$

$\hookrightarrow$ Concerning Γ_2, the monotony of the functions $\mathbb{E}\hat{L}_N(\theta,.,z)$ and $\hat{L}_N(\theta,.,z)$ permits to write $\forall m\leq u_n, \forall z\in S_{\mathbb{H}}, \forall\theta\in\Theta_H$

$$\mathbb{E}\hat{L}_N(\theta,t_{m(t)}-l_n,z_{k(z)})\leq \sup_{t\in(t_{m(t)}-l_n,t_{m(t)}+l_n)}\mathbb{E}\hat{L}_N(\theta,t,z)\leq \mathbb{E}\hat{L}_N(\theta,t_{m(t)}+l_n,z_{k(z)})$$

$$\hat{L}_N(\theta,t_{m(t)}-l_n,z_{k(z)})\leq \sup_{t\in(t_{m(t)}-l_n,t_{m(t)}+l_n)}\hat{L}_N(\theta,t,z)\leq \hat{L}_N(\theta,t_{m(t)}+l_n,z_{k(z)}).$$

Next we use the Hölder's condition on $L(\theta,t,z)$ and we show that, for any $t_1,t_2\in S_{\mathbb{R}}$ and for all $z\in S_{\mathbb{H}},\theta\in\Theta_H$

$$\begin{aligned} \mid \mathbb{E}\hat{L}_N(\theta,t_1,z)-\mathbb{E}\hat{L}_N(\theta,t_2,z)\mid &= \frac{1}{\mathbb{E}K_1(z,\theta)}\mid \mathbb{E}(K_1(z,\theta)L(\theta,t_1,Z_1))-\mathbb{E}(K_1(z,\theta)L(\theta,t_2,Z_1))\mid \\ &\leq C\mid t_1-t_2\mid^{b_2}. \end{aligned}$$

Now, we have, for all $n>0$

$$\mathbb{P}\left(\sup_{j\in\{1\ldots d_n^{\Theta_H}\}}\sup_{k\in\{1\ldots d_n^{S_H}\}}\sup_{1\leq m\leq u_n}\mid \hat{L}_N(\theta,t,z_{k(z)})-\mathbb{E}\hat{L}_N(\theta,t,z_{k(z)}))\mid > n\sqrt{\frac{\log d_n^{S_H}d_n^{\Theta_H}}{n\phi(h_K)}}\right)$$

$$=$$

$$\mathbb{P}\left(\max_{j\in\{1\ldots d_n^{\Theta_H}\}}\max_{k\in\{1\ldots d_n^{S_H}\}}\max_{1\leq m\leq u_n}\mid \hat{L}_N(\theta,t,z_{k(z)})-\mathbb{E}\hat{L}_N(\theta,t,z_{k(z)}))\mid > n\sqrt{\frac{\log d_n^{S_H}d_n^{\Theta_H}}{n\phi(h_K)}}\right)$$

$$\leq$$

$$u_n d_n^{S_H}d_n^{\Theta_H}\max_{j\in\{1\ldots d_n^{\Theta_H}\}}\max_{k\in\{1\ldots d_n^{S_H}\}}\max_{1\leq m\leq u_n}\mathbb{P}\left(\mid \hat{L}_N(\theta,t,z_{k(z)})-\mathbb{E}\hat{L}_N(\theta,t,z_{k(z)}))\mid > n\sqrt{\frac{\log d_n^{S_H}d_n^{\Theta_H}}{n\phi(h_K)}}\right)$$

$$\leq$$

$$2u_n d_n^{S_H}d_n^{\Theta_H}\exp(-c\eta^2\log d_n^{S_H}d_n^{\Theta_H})$$

choosing $u_n=O(l_n^{-1})=O(n^{\frac{1}{2b_2}})$, we get

$$\mathbb{E}\left(\mid \hat{L}_N(\theta, t, z_{k(z)}) - \mathbb{E}\hat{L}_N(\theta, t, z_{k(z)}) \mid > n\sqrt{\frac{\log d_n^{S_H} d_n^{\Theta_H}}{n\phi(h_K)}}\right) \leq C' u_n (d_n^{S_H} d_n^{\Theta_H})^{1-Cn^2}$$

putting $C\eta^2 = \beta$ and using [A4], we get

$$\Gamma_2 = \mathbb{O}\left(\sqrt{\frac{\log d_n^{S_\varsigma} d_n^{\Theta_\varsigma}}{n\phi(h_K)}}\right).$$

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

$$\begin{aligned} \mid \hat{L}_N(t_{j(\theta)}, t, z_{k(z)}) - \hat{L}_N(t_{j(\theta)}, t_{m(t)}, z_{k(z)}) \mid &\leq C\tfrac{1}{n\phi(h_K)} \textstyle\sum_{i=1}^n K_i(t_{j(\theta)}, z_{k(z)}) \mid H_i(t) - H_i(t_{m(t)}) \mid \\ &\leq \tfrac{Cl_n}{nh_H\phi(h_K)} \textstyle\sum_{i=1}^n K_i(t_{j(\theta)}, z_{k(z)}). \end{aligned}$$

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

$$\hat{L}_N(t_{j(\theta)}, t, z_{k(z)}) - \hat{L}_N(t_{j(\theta)}, t_{m(t)}, z_{k(z)}) = \mathbb{O}\left(\frac{l_n}{h_H}\right) + \mathbb{O}_{a.co}\left(\frac{l_n}{h_H}\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right).$$

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

$$\frac{l_n}{h_H\phi(h_K)} = o\left(\sqrt{\frac{\log d_n^{S_H} d_n^{\Theta_H}}{n\phi(h_K)}}\right),$$

then

$$\Gamma_3 = \mathbb{O}_{a.co}\left(\sqrt{\frac{\log d_n^{S_H} d_n^{\Theta_H}}{n\phi(h_K)}}\right).$$

Hence, for n large enough, we have

$$\Gamma_3 \leq \Gamma_4 = \mathbb{O}_{a.co}\left(\sqrt{\frac{\log d_n^{S_H} d_n^{\Theta_H}}{n\phi(h_K)}}\right).$$

↪ Concerning Γ_5, we have

$$\mathbb{E}(\hat{L}_N(\theta, t, z_{k(z)})) - \mathbb{E}(\hat{L}_N(\theta, t, z)) \leq \sup_{z\in S_H} \mid \hat{L}_N(\theta, t, z) - \hat{L}_N(\theta, t, z_{k(z)}) \mid,$$

then following similar proof used in the study of τ_1 and using the same idea as for $\mathbb{E}(\hat{\varphi}(\theta, z_{k(z)})) - \mathbb{E}(\hat{\varphi}(\theta, z))$ we get, for n tending to infinity,

$$\Gamma_5 = \mathbb{O}_{a.co}\left(\sqrt{\frac{\log d_n^{S_H} d_n^{\Theta_H}}{n\phi(h_K)}}\right).$$

The proof of these for points to ones given in Mahiddine et al [111], so it is sufficient to replace $\hat{F}_D(\theta, z), F(\theta, t, z)$ and $\mathbb{E}(\hat{F}_N(\theta, t, z))$ [lemma 6,corollary 3 and lemma 7] by $\hat{\varphi}_D(\theta, z), L(\theta, t, z)$ and $\mathbb{E}(\hat{L}_N(\theta, t, z))$ respectively.

• Concerning [22], let $H_i^{'}(t) = H^{'}(h_H^{-1}(t - T_i))$, note that

$$\mathbb{E}\hat{\varphi}_N(\theta, t, z) - \varphi(\theta, t, z) = \frac{1}{h_H \mathbb{E}K_1(z, \theta)} \mathbb{E}(K_1(z, \theta)[\mathbb{E}(H_1^{'}(t) I_{X_1 \leq C_1} | < Z_1, \theta >) - h_H \varphi(\theta, t, z)]).$$

Moreover,

$$\begin{aligned}
\mathbb{E}(H_1^{'}(t) S_1(\theta, X_1, z) | < Z_1, \theta >) &= \textstyle\int_{\mathbb{R}} H^{'}(h_H^{-1}(t - w)) S_1(\theta, w, Z_1) f(\theta, w, Z_1) dw, \\
&= h_H \textstyle\int_{\mathbb{R}} H^{'}(h_H^{-1}(t - w)) \varphi(\theta, w, Z_1) dw, \\
&= h_H \textstyle\int_{\mathbb{R}} H^{'}(v) \varphi(\theta, t - v h_H, Z_1) dv.
\end{aligned}$$

Under condition [H10] we can write:

$$| \mathbb{E}(H_1^{'}(t) S_1(\theta, X_1, z) | < Z_1, \theta > - h_H \varphi(\theta, t, Z)) | \leq h_H \int_{\mathbb{R}} H^{'}(t) | \varphi(\theta, t - h_H t, Z_1) - \varphi(\theta, t, Z) | dt.$$

Finally, [A2a] allows to write

$$| \mathbb{E}(H_1^{'}(t) S_1(\theta, X_1, z) | < Z_1, \theta > - h_H \varphi(\theta, t, Z)) | \leq C h_H \int_{\mathbb{R}} H^{'}(t) (h_K^{b_1} + | t |^{b_2} h_H^{b_2}) dt.$$

This inequality is uniform on $(\theta, t, Z) \in \Theta_H \times S_H \times S_{\mathbb{R}}$, now to finish the proof it is sufficient to use [H5].

• Concerning [23], let us keep the definition of $K(z)$ [resp. $j(\theta)$] as in corollary 4.4.9. The compactness of $S_{\mathbb{R}}$ permits to write that $S_{\mathbb{R}} \subset \bigcup_{m=1}^{u_n} (t_m - l_n, t_m + l_n)$ with $L_n = n^{-\frac{3}{2}\gamma - \frac{1}{2}}$ and $u_n \leq C n^{\frac{3}{2}\gamma + \frac{1}{2}}$. Taking $m(t) = \arg minimize_{\{1, \dots, u_n\}} | t - t_m |$. consider the following decomposition:

$$
\begin{aligned}
\mid \hat{\varphi}_N(\theta,t,z) - \mathbb{E}(\hat{\varphi}_N(\theta,t,z)) \mid \;\leq\;& \underbrace{\mid \hat{\varphi}_N(\theta,t,z) - \hat{\varphi}_N(\theta,t,z_{k(z)}) \mid}_{\Delta_1} \\
&+ \underbrace{\mid \hat{\varphi}_N(\theta,t,z_{k(z)}) - \mathbb{E}(\hat{\varphi}_N(\theta,t,z_{k(z)})) \mid}_{\Delta_2} \\
&+ 2\underbrace{\mid \hat{\varphi}_N(t_{j(\theta)},t,z_{k(z)}) - \hat{\varphi}_N(t_{j(\theta)},t_{j(t)},z_{k(z)}) \mid}_{\Delta_3} \\
&+ 2\underbrace{\mid \mathbb{E}(\hat{\varphi}_N(t_{j(\theta)},t,z_{k(z)}) - \mathbb{E}(\hat{\varphi}_N(t_{j(\theta)},t_{j(t)},z_{k(z)})) \mid}_{\Delta_4} \\
&+ \underbrace{\mid \mathbb{E}(\hat{\varphi}_N(\theta,t,z_{k(z)}) - \mathbb{E}(\hat{\varphi}_N(\theta,t,z) \mid}_{\Delta_5}
\end{aligned}
$$

$\rightsquigarrow$ Concerning Δ_1, we use the Hölder continuity condition on K, the Cauchy- Shchwartz's inequality and the Bernstein inequality. With theses arguments we get

$$\Delta_1 = \mathbb{O}\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{nh_H\phi(h_K)}}\right).$$

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

$$\Delta_5 \leq \Delta_1 = \mathbb{O}\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{nh_H\phi(h_K)}}\right). \tag{44}$$

$\rightsquigarrow$ For Δ_2, we follow the same idea given for Γ_2, we get

$$\Delta_2 = \mathbb{O}\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{nh_H\phi(h_K)}}\right).$$

$\rightsquigarrow$ Concerning Δ_3 and Δ_4, using Lipschitz's condition on the kernel H,

$$\mid \hat{\varphi}_N(t_{j(\theta)},t,z_{k(z)}) - \hat{\varphi}_N(t_{j(\theta)},t_{m(t)},z_{k(z)}) \mid \leq \frac{l_n}{h_H^2\phi(h_K)},$$

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

$$\frac{l_n}{h_H^2\phi(h_K)} = o\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{nh_H\phi(h_K)}}\right),$$

So, for n large enough, we have

$$\Delta 3 = \mathbb{O}_{a.co}\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{nh_H\phi(h_K)}}\right).$$

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

$$\Delta_4 \leq \Delta_3 = \mathbb{O}_{a.co}\left(\sqrt{\frac{\log d_n^{S_H} + \log d_n^{\Theta_H}}{nh_H\phi(h_K)}}\right). \tag{45}$$

Finally, the lemma can be easily deduced from (44) and (45).

Chapter 5

Comments and prospect

In this thesis we are interested in the estimation of a functional parameter in the conditional models. We treat the estimation of the conditional hazard function and we give the explicit expression of the terms asymptotically dominant of biais and variance with the rate of convergence, asymptotically exact in the two types of correlations namely the i.i.d case and the case of the dependent variables with asymptotic normality one keeping the form usual of the quadratic error.

The work developed in this thesis offers many prospects in short and long terms. Concerning the short-term prospects:

1. The asymptotic normality of our estimators can enable us to make tests and to build confidence intervals. For example we can plan to use the same ideas of E. Masry [114] concerning the conditional quantile as well as the conditional hazard function, he got results on this problem in regression and should be possible.

2. We can also consider while making an adaptation of the tools developed by Niang and Rhomari [41] to study a norm convergence L^p of our estimators in depending case and the ergodic case , another possible prospect is to obtain the rates of convergence concerning the quadratic error and the quadratic error integrated.

3. The choice of the smoothing parameter: We can generalize the result of E. Youndjé (1993) on the choice of the smoothing parameter for the estimation of the Conditional hazard function (rachdi and Vieu [130] have already addressed this problem in the case of regression functional). Obtaining the results of mean square convergence would have very useful in that sense.

Other topics can be approached in the long term as conditioning by p functional variables or combination linear of these p functional variables, among other work the estimation of conditional quantile and the conditional hazard function for functional explanatory variable open several prospects. For example, we can consider another estimator by using another method like the kernel estimation (Fourier, ondelettes...), or the research of optimal rates, On another plan, concerning the hypothesis of mixture, we can also plan to obtain speeds of convergence for ergodic data by making assumption of ergodicity, or still for truncated data and censured at the same time. Another possible prospect is to suppose not only the explanatory functional variable but also the variable of interest.

Bibliography

[1] F. Abramovich, C. Angelini, Bayesian maximum a posteriori multiple testing procedure.*Sankhya,* 68 (2006)436-460. Uniform strong convergence of the generalized failure rate estimate, *Bull. Math. Statist,* 17 (1976) 77-84.

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

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

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

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

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

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

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

[9] S.Attaoui, A.Laksaci, E.Ould-Saïd, a note on the conditional density estimate in the single functional index model. *Statist. Probab. Lett,* 81 (2011) No.1, 45-53.

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

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

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

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

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

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

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

[17] P. Besse, J.O. Ramsay, Principal component analysis of sampled curves, *Psychometrika,* 51 (1986) 285-311.

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

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

[20] D. Bosq, Nonparametric Statistics for Stochastic Processes. Estimation and Prediction, *Lecture Notes in Statistics. Springer-Verlag, Berlin, New York, Heidelberg,* 2(1998).

[21] D. Bosq, J. P. Lecoutre, Théorie de léstimation fonctionnelle *Ed. Economica.*1987.

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

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

[24] M. Bouraine, A. *At* Saidi, F. Ferraty and P. Vieu, Choix optimal de l'indice multi-fonctionnel: methode de validation croisée, *R. Roumaine Math. Pures et Appl.* 50(5) (2010) 355-367.

[25] F.Burba, F.Ferraty, P.Vieu, Convergence de l'estimateur à noyau des k plus proches voisns en régression fonctionnelle non-paramétrique. *C. R. Acad. Sci., Paris,* 246 (2008) 339-342.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[41] S.Dabo-Niang, N.Rhomari. Estimation non paramétrique de la régression avec variable explicative dans un espace métrique, *C. R. Acad. Sci., Paris,* 336 (2003) No.1, 75-80.

[42] S. Dabo-Niang, Sur léstimation fonctionnelle en dimension infinie : Application aux diffusions, *Thése de Doctorat, Universit de Paris 6* (2002).

[43] J. Damon, S. Guillas, The inclusion of exogenous variables in functional autoregressive ozone forecasting, *Environmetrics,* 13 (2002) 759-774.

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

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

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

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

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

[49] L.Delsol . 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,*(2011).

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

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

[52] T. Elisa, L. John Wenyu Wang, Statistical Methods for Survival Data Analysis Third Edition, John Wiley, 2003

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

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

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

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

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

[58] F. Ferraty, A. Laksaci, A. Vieu, Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statist. Inference Stochastic Process,* 9 (2006), 47-76.

[59] F. Ferraty, A. Laksaci, A. Tadj, P. Vieu, Rate of uniform consistency for nonparametric estimates with functional variables. *Journal ofStatisticalPlanning and Inference,* 140 (2010), 335-352.

[60] F. Ferraty, A. Mas and P. Vieu, Advances in nonparametric regression for functional variables, *Australian and New Zealand Journal of Statistics,* 49 (2007) 1-20.

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

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

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

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

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

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

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

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

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

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

[71] F. Ferraty. A. Rabhi, and P. Vieu, Conditional Quantiles for Functionally Dependent Data with Application to the Climatic El Nino Phenomeno, Sankhyā: *The Indian Journal of Statistics,* 67(2) (2005) 378-398.

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

[73] F. Ferraty and P. Vieu, Non-parametric Functional Data Analysis, *Springer-Verlag, NewYork,* 2006.

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

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

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

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

[78] T. Gasser, P. Hall and B. Presnell, Nonparametric estimation of the mode of a distribution of random curves, *Journal of the Royal Statistical Society, Ser. B.,* 60 (1998) 681-691.

[79] O. Gefeller, P. Michels, A review on smoothing methods for the estimation of the hazard rate based on kernel functions. *Computational Statistics,* (1992),459-464.

[80] O. Gefeller, P. Michels, 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,* (1992)459-464.

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

[82] J. De Gooijer and A. Gannoun, Nonparametric conditional predictive regions for time series, *Comput.Statist. Data Anal.* 33 (2000) 259-257.

[83] P. Hall, R. C. L. Wolef and Q. Yao, Methods for estimating a conditional distribution function, *An.Statist. Assoc.,* 94 (1999) 154-163.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[99] R. J. Hyndman, D. M. Bashtannyk and G. K. Grunwald, Estimating and visualizing conditional densities, *J. Comput. Graph. Statist.* 5 (1996) 315-336.

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

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

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

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

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

[105] A. N. Kolmogorov, V .M. Tikhomirov, (1959). ϵ-entropy and ϵ-capacity. Uspekhi Mat. Nauk **14**, 3-86.*Engl Transl. Amer. Math. Soc. Transl. Ser,* 2, 277-364 (1961).

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

[107] A. Laksaci, M.Mechab, Estimation non parametrique de la fonction de hasard avec variable explicative fonctionnelle cas des donnees spatiales. Rev: *Roumaine, Math Pures Appl,* 55(2010), 35-51.

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

[109] E. Liebscher, Central limit theorem for a-mixing triangular arrays with applications to nonparemetric statistics, *Mathematical Methods of Statistics.* 102 (2001), 194-214.

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

[111] A. Mahiddine, A. A. Bouchentouf, A. Rabhi, Nonparametric estimation of some characteristics of the conditional distribution in single functional index model. *Malaya Journal of Matematik* ,2(4) (2014) 392-410.

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

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

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

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

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

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

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

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

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

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

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

[123] W. J. Padgett, Nonparametric estimation of density and hazard rate functions when samples are censored. *Elsevier Science Publishers* .7(1988), 313-331,

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

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

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

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

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

[129] A. Rabhi, S. Benaissa, E. H. Hamel and B. Mechab, Mean square error of the estimator of the conditional hazard function, *Appl. Math. (Warsaw),* 40(4) (2013) 405-420.

[130] M. Rachdi and P. Vieu, Non-parametric regression for functional data: Automatic smoothing parameter selection, *J. Stat. Plan. Inf.* 137 (2007) 2784-2801.

[131] J.O. Ramsay, B.W. Silverman, Functional Data Analysis, *Springer-Verlag, Berlin, New York, Heidelberg,* 2(2005).

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

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

[134] J. O. Ramsay, B. W. Silverman, Functional Data Analysis.*Springer-Verlag,* (1997)

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

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

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

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

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

[140] E. Rio, Théorie asymptotique des processus aléatoires dépendants (in french). Mathematiques et Applications. 31, *Springer-Verlag, Berlin, New York, Heidelberg,* 2000.

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

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

[143] G. G.Roussas, Hazard rate estimation under dependence conditions, *J. Statist. Plann. Inf.* 22 (1989) 81-93.

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

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

[146] P. Sarda, P.Vieu, Kernel regression. In: M. Schimek (ed.), Smoothing and regression. Approaches, computation and application, Wiley Series in Probability and Statistics, *Wiley, New York,*(2000), 43-70.

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

[148] N. D. Singpurwalla, M. Y. Wong, Estimation of the failure rate-Asurvey of non-parametric methods. Part I: Non-Bayesian methods, *Commun. Stat. Theory Meth.* 12 (1983) 559-588.

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

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

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

[152] L. T. Tran, S. Yakowitz, Nearest neighbor estimators for rom fields, *J. Multivariate. Anal.* 44 (1993) 23-46.

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

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

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

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

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

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

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

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

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