

QUADRATIC ERROR OF THE ESTIMATION OF THE HAZARD FUNCTION CONDITIONAL IN NONPARAMETRIC FUNCTIONAL MODEL

Naouel Belkhir*

Abu Bakr Belkaid University of Tlemcen
Laboratory of Mathematics
ALGERIA

Abbes Rabhi

University of Sidi Bel Abbes
Laboratory of Mathematics
ALGERIA

Sara Soltani

Moulay Tahar University of Saida
Stochastic Models, Statistics and Applications
Saida/ALGERIA

Abstract: This paper 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.

Key words: Functional data, Kernel conditional hazard function, Kernel estimation, Nonparametric Estimation, Probabilities of small balls.

History: Submitted: 18 August 2014; Revised: 1 November 2015; Accepted: 4 November 2015

1. Introduction

The estimated hazard rate, because of the variety of its possible applications, is an important issue in statistics.

This topic can (and should) be approached from several angles depending on the complexity of the problem: presence of censoring in the observed sample (for example, common phenomenon in medical applications), presence of dependence between the observed variables (for example, common phenomenon in applications such as seismic or econometric) or presence of explanatory variables. Many techniques have been studied in the literature to deal with these situations but all deal only with random explanatory variables real and multidimensional.

Technical advances in collection and data storage can have more often statistical functional: curves, images, tables, ... The data are modeled as realizations of a random variable taking values in an abstract space of infinite dimension, and the scientific community was naturally interested in recent years the development of statistical tools capable of handling this type of sample.

Thus, estimating a hazard rate in the presence of functional explanatory variable is a topical issue. In this context, the first results were obtained by Ferraty *et al.* [6]. They studied the almost complete convergence of a kernel estimator of the conditional hazard function assuming i.i.d observations and the case of observations mixing for complete data and censored. The estimators that we define are based on the techniques of convolution kernel.

* Corresponding author. E-mail address: n.belkhir@yahoo.com (N. Belkhir)

The study of functions (the hazard function and the conditional hazard function) is of obvious interest in many scientific fields (biology, medicine, reliability, seismology, econometrics, ...), and many authors have studied the construction of nonparametric estimators of hazard function. One of the most common techniques for constructing estimators of the hazard function (respectively the hazard function conditional) is to study a quotient of the density estimator (respectively the conditional density) and an estimator of S (respectively the conditional survival function). The article by Patil *et al.* [12] presented an overview of estimation techniques. The non-parametric methods based on the ideas of convolution kernel, which are known for their good behavior problems in density estimation (conditional or not), and are widely used in nonparametric estimation of hazard function. A wide range of literature in this area is provided by the literature reviews of Singpurwalla and Wong [15], Hassani *et al.* [9], Izenman [10], Gefeller and Michels [8] and Pascu and Vaduva [11].

Advances in data collection processes have the immediate consequence of the opportunity for statisticians to have more and more observations of functional variables. The works of Ramsay and Silverman [13] and Ferraty and Vieu [5] offer a wide range of statistical methodologies, parametric or not, recently developed to treat various problems of estimation are carried out in functional random variables (ie with values in an infinite dimensional space).

The objective of this paper is to study a model in which the conditional random explanatory variable X is not necessarily real or multi-dimensional but only supposed to be with values in an abstract space \mathcal{F} semi-normed.

As with any problem of nonparametric estimation, the dimension of the space \mathcal{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

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

intervene directly in the asymptotic behavior of any estimator nonparametric functional (see Ferraty *et al.* [4]). The asymptotic results that we present later in this article on convergence in mean square of the conditional hazard function will not escape this rule.

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 $(\mathcal{F}, \|\cdot\|)$ which can be of infinite dimension with semi-norm $\|\cdot\|$ so that $d(x, x') = \|x - x'\|$. 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 $x \in \mathcal{F}$, we assume that the regular version of the conditional probability of Y given $X = x$ exists denoted by $F^x(y)$ and has a bounded density with respect to Lebesgue measure over \mathbb{R} , denoted by $f^x(y)$. In the following (x, y) will be a fixed point in $\mathcal{F} \times \mathbb{R}$ and $N_x \times \mathcal{S}_{\mathbb{R}}$ will denote a fixed neighborhood of (x, y) , $\mathcal{S}_{\mathbb{R}}$ will be a fixed compact subset of \mathbb{R} , and we will use the notation $B(t, h) = \{t' \in \mathcal{F} / \|t' - t\| < h\}$. 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 :

$$C_B^2(\mathcal{F} \times \mathbb{R}) = \left\{ \begin{array}{l} \varphi : \mathcal{F} \times \mathbb{R} \longrightarrow \mathbb{R} \\ (x, y) \longrightarrow \varphi(x, y) \quad \text{such as:} \\ \forall z \in N_x, \varphi(z, \cdot) \in C^2(\mathcal{S}_{\mathbb{R}}) \text{ and } \left(\varphi(\cdot, y), \frac{\partial^2 \varphi(\cdot, y)}{\partial y^2} \right) \in C_B^1(x) \times C_B^1(x), \end{array} \right\} \quad (2.1)$$

where $C_B^1(x)$ is the set of continuously differentiable functions to sens of Gteaux on N_x (see Troutman [16] 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} . Given i.i.d. observations $(X_1, Y_1), \dots, (X_n, Y_n)$ of (X, Y) , where the variable X is of functional nature and Y is scalar. Formally, we will consider that X is a random variable valued in some semi normed functional space \mathcal{F} , and we will denote by $\|\cdot\|$ the associated semi norm defined above. The conditional cumulative distribution of Y given $X = x$ is defined for any $y \in \mathbb{R}$ and any $x \in \mathcal{F}$ by

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

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

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

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

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

with the convention $\frac{0}{0} = 0$, where K and H are kernel functions and h_n is sequence of smoothing parameter. The conditional distribution operator $F^x(\cdot)$ can be estimated by

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

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

$$\forall x \in \mathcal{F}, \quad \forall y \in \mathbb{R} \quad h^x(y) = \frac{f^x(y)}{1 - F^x(y)} = \frac{f^x(y)}{S^x(y)} \tag{2.2}$$

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

3. Asymptotic Properties

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

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

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

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

To establish the convergence in mean square of the estimator $\widehat{h}^x(y)$ to $h^x(y)$, we introduce the following assumptions:

$$(H0) \quad \forall h > 0, \quad 0 < \mathbb{P}(Z \in B(x, h)) = \phi_x(h) \text{ and } \lim_{h \rightarrow 0} \phi_x(h) = 0$$

(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:

$$w(r, x, v) = \phi(r)g(x, v) + o(\phi(r)) \text{ for all } v \in B(0, 1), \quad (3.1)$$

where

— ϕ is an increasing function with values in \mathbb{R}^+ .

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

(H2) The kernel K with compact support (0.1) satisfying $0 < A_3 < K(t) < A_4 < \infty$,

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

$$\int H(t) dt = 1, \quad \int t^2 H(t) dt < \infty,$$

(H4) $\lim_{n \rightarrow \infty} h_n = 0$ and $\lim_{n \rightarrow \infty} nh_n \phi(h_n) = \infty$,

(H5) $\exists \tau < \infty, f^x(y) \leq \tau, \forall (x, y) \in \mathcal{F} \times \mathcal{S}_{\mathbb{R}}$

(H6) $\exists \beta > 0, F^x(y) \leq 1 - \beta, \forall (x, y) \in \mathcal{F} \times \mathcal{S}_{\mathbb{R}}$.

4. Remarks

REMARK 1 (NOTES ON VARIABLE FUNCTIONAL). : The hypothesis (H1) on the functional variable X can be divided into two parts:

1. 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 paper 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 [17] 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 [3]).

2. The second part (3.1) is less restrictive than the following condition, given for all $(r, v) \in \mathbb{R}_*^+ \times B(0, 1)$ (x fixed):

$$\exists C_1, C_2 > 0, \quad 0 < C_1 \phi(r)g(x, v) \leq w(r, x, v) \leq C_2 \phi(r)g(x, v).$$

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:

$$\begin{aligned} \mathbb{P}(X \in B(x, r)) &= \int_{B(0,1)} w(r, x, v) d\mu(v) \\ &= \phi(r) \int_{B(0,1)} g(x, v) d\mu(v) + o(\phi(r)) > 0. \end{aligned}$$

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 $w(r, x, v) = r^p f(x) + o(r^p)$.

REMARK 2 (NOTES ON NON-PARAMETRIC MODEL). In this paper, 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 [17]). 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 (2.1), we obtain a result for the conditional distribution and conditional density respectively for example of type:

$$\mathbb{E} \left[(\widehat{F}^x(y) - F^x(y))^2 \right] = \mathcal{O}(h_n^4) + \mathcal{O} \left(\frac{1}{n\phi(h_n)} \right),$$

$$\mathbb{E} \left[(\widehat{f}^x(y) - f^x(y))^2 \right] = \mathcal{O}(h_n^4) + \mathcal{O} \left(\frac{1}{nh_n\phi(h_n)} \right).$$

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.

5. Main Results

5.1. Mean Squared Convergence

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

THEOREM 1. Under hypotheses (H0)-(H6) and if $F^x(y)$ (resp. $f^x(y)$) $\in C_B^2(\mathcal{F} \times \mathbb{R})$ then

$$\begin{aligned} \text{MSE} \widehat{h}^x(y) &\equiv \mathbb{E} \left[(\widehat{h}^x(y) - h^x(y))^2 \right] \\ &\equiv B_n(x, y) + \frac{\sigma_h^2(x, y)}{nh_n\phi_x(h_n)} + o(h_n^4) + o \left(\frac{1}{nh_n\phi_x(h_n)} \right) \end{aligned}$$

where

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

with

$$\begin{aligned} B_H^f(x, y) &= \frac{1}{2} \frac{\partial^2 f^x(y)}{\partial y^2} \int t^2 H(t) dt, \\ B_K^f(x, y) &= \frac{\int_{B(0,1)} K(\|v\|) D_x f^x(y)[v] g(x, v) d\mu(v)}{\int_{B(0,1)} K(\|v\|) g(x, v) d\mu(v)} \\ B_F(x, y) &= \frac{1}{2} \frac{\partial^2 F^x(y)}{\partial y^2} \int t^2 H(t) dt \\ B_K^F(x, y) &= \frac{\int_{B(0,1)} K(\|v\|) D_x F^x(y)[v] g(x, v) d\mu(v)}{\int_{B(0,1)} K(\|v\|) g(x, v) d\mu(v)}. \end{aligned}$$

and

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

where D_x means the derivative with respect to x .

T his proof is based on the decomposition

$$\begin{aligned}
 \widehat{h}^x(y) - h^x(y) &= \frac{\widehat{f}^x(y)}{1 - \widehat{F}^x(y)} - \frac{f^x(y)}{1 - F^x(y)} \\
 &= \frac{1}{1 - \widehat{F}^x(y)} \left[(\widehat{f}^x(y) - f^x(y)) + \frac{f^x(y)}{1 - F^x(y)} (\widehat{F}^x(y) - F^x(y)) \right] \\
 &= \frac{1}{\widehat{f}_D(x) - \widehat{g}_N(x, y)} \left(\widehat{f}_N(x, y) - \mathbb{E}\widehat{f}_N(x, y) \right) \\
 &\quad + \frac{1}{\widehat{f}_D(x) - \widehat{g}_N(x, y)} (\mathbb{E}\widehat{g}_N(x, y) - F^x(y)) \\
 &\quad + \frac{1}{\widehat{f}_D(x) - \widehat{g}_N(x, y)} \left[(\mathbb{E}\widehat{f}_N(x, y) - f^x(y)) \right. \\
 &\quad \left. + \frac{1}{\widehat{f}_D(x) - \widehat{g}_N(x, y)} \left(1 - \mathbb{E}\widehat{g}_N(x, y) - (\widehat{f}_D(x) - \widehat{g}_N(x, y)) \right) \right] \quad (5.1)
 \end{aligned}$$

where

$$\begin{aligned}
 \widehat{f}_D(x) &:= \frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K(h_K^{-1}\|x - X_i\|), \quad K_1(x) = K(h_K^{-1}\|x - X_i\|) \\
 \widehat{g}_N(x, y) &:= \frac{1}{n\mathbb{E}[K_1(x)]} \sum_{i=1}^n K(h_K^{-1}\|x - X_i\|) \mathbf{1}_{\{Y_i \leq y\}} \\
 \widehat{f}_N(x, y) &:= \widehat{g}_N^{(1)}(x, y) := \frac{1}{nh_H\mathbb{E}[K_1(x)]} \sum_{i=1}^n K(h_K^{-1}\|x - X_i\|) H(h_H^{-1}(y - Y_i)).
 \end{aligned}$$

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

$$\left| \widehat{h}^x(y) - h^x(y) \right| \leq C \frac{\left| \widehat{f}^x(y) - f^x(y) \right| + \left| \widehat{F}^x(y) - F^x(y) \right|}{\left| 1 - \widehat{F}^x(y) \right|},$$

Then, Theorem (1) can be deduced from both lemmas above Lemma (1) and Lemma (2).

LEMMA 1. Under hypotheses (H0)-(H5) and if $f^x(y) \in C_B^2(\mathcal{F} \times \mathbb{R})$ then:

$$\begin{aligned}
 \mathbb{E} \left[(\widehat{f}^x(y) - f^x(y)) \right]^2 &= B_H^2(x, y) h_n^4 + B_K^2(x, y) h_n^2 + \frac{\sigma_f^2(x, y)}{nh_n \phi(h_n)} \\
 &\quad + o(h_n^4) + o\left(\frac{1}{nh_n \phi(h_n)}\right) \quad (5.2)
 \end{aligned}$$

where

$$\sigma_f^2(x, y) = \frac{(f^x(y)) \left(\int_{B(0,1)} K^2(\|v\|) g(x, v) d\mu(v) \right) \int H^2(t) dt}{\left(\int_{B(0,1)} K(\|v\|) g(x, v) d\mu(v) \right)^2},$$

LEMMA 2. Under hypotheses (H0)-(H4), (H6) and if $F^x(y) \in C_B^2(\mathcal{F} \times \mathbb{R})$ then:

$$\begin{aligned}
 \mathbb{E} \left[(\widehat{F}^x(y) - F^x(y)) \right]^2 &= B_F^2(x, y) h_n^4 + B_K^2(x, y) h_n^2 + \frac{\sigma_F^2(x, y)}{n\phi(h_n)} \\
 &\quad + o(h_n^4) + o(h_n^2) + o\left(\frac{1}{n\phi(h_n)}\right) \quad (5.3)
 \end{aligned}$$

with

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

REMARK 3. Observe that, the result of this lemmas Lemma 1 and Lemma 2 permits to write

$$[\mathbb{E}\widehat{g}_N(x, y) - F^x(y)] = \mathcal{O}(h_n)$$

and

$$[\mathbb{E}\widehat{f}_N(x, y) - f^x(y)] = \mathcal{O}(h_n^2).$$

[Proof of Lemma (1)] 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

$$\mathbb{E} \left[(\widehat{f}^x(y) - f^x(y))^2 \right] = \left[\mathbb{E} \left(\widehat{f}^x(y) \right) - f^x(y) \right]^2 + Var \left[\widehat{f}^x(y) \right].$$

We define the quantities $K_i(x) = K(h_K^{-1}\|x - X_i\|)$, $H_i(y) = H(h_H^{-1}(y - Y_i))$ for all $i = 1, \dots, n$ and we set

$$\widehat{g}_N(x, y) = \frac{1}{n\phi(h_n)} \sum_{i=1}^n K_i(x) \mathbf{1}_{\{Y_i \leq y\}}, \quad \widehat{f}_D(x) = \frac{1}{n\phi(h_K)} \sum_{i=1}^n K_i(x)$$

and

$$\widehat{f}_N(x, y) = \widehat{g}_N^{(1)}(x, y) = \frac{1}{nh_n\phi(h_n)} \sum_{i=1}^n K_i(x)H_i(y)$$

We will calculate both sides of this equation (party bias and variance part) to arrive at the calculation of $\mathbb{E} \left[\widehat{f}^x(y) - f^x(y) \right]^2$.

We come at the following to writing:

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

from which we draw:

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

as

$$A_1 = \mathbb{E}\widehat{f}_N(x, y) \left(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x) \right) = Cov(\widehat{f}_N(x, y), \widehat{f}_D(x))$$

and

$$A_2 = \mathbb{E} \left(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x) \right)^2 \widehat{f}^x(y)$$

Can be written as

$$\begin{aligned} \widehat{f}^x(y) - f^x(y) &= \left(\frac{\widehat{f}_N(x, y)}{\mathbb{E}\widehat{f}_D(x)} - f^x(y) \right) \\ &\quad - \frac{\left(\widehat{f}_N(x, y) - \mathbb{E}\widehat{f}_N(x, y) \right) \left(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x) \right)}{\left(\mathbb{E}\widehat{f}_D(x) \right)^2} \end{aligned}$$

$$\begin{aligned}
& - \frac{\left(\mathbb{E}\widehat{f}_N(x, y)\right) \left(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x)\right)}{\left(\mathbb{E}\widehat{f}_D(x)\right)^2} \\
& + \frac{\left(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x)\right)^2}{\left(\mathbb{E}\widehat{f}_D(x)\right)^2} \widehat{f}^x(y)
\end{aligned} \tag{5.4}$$

which implies

$$\begin{aligned}
\mathbb{E}\left[\widehat{f}^x(y)\right] - f^x(y) &= \left(\left(\mathbb{E}\widehat{f}_D(x)\right)^{-1}\mathbb{E}\left(\widehat{f}_N(x, y)\right) - f^x(y)\right) - \left(\left(\mathbb{E}\widehat{f}_D(x)\right)^{-2}\text{Cov}\left(\widehat{f}_N(x, y), \widehat{f}_D(x)\right)\right) \\
&+ \left(\mathbb{E}\widehat{f}_D(x)\right)^{-2} \mathbb{E}\left(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x)\right)^2 \widehat{f}^x(y) \\
&= \left(\left(\mathbb{E}\widehat{f}_D(x)\right)^{-1}\mathbb{E}\left(\widehat{f}_N(x, y)\right) - f^x(y)\right) - \left(\mathbb{E}\widehat{f}_D(x)\right)^{-2} A_1 + \left(\mathbb{E}\widehat{f}_D(x)\right)^{-2} A_2.
\end{aligned}$$

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 $\widehat{f}^x(y)$ by a constant $C > 0$, as $\widehat{f}^x(y) \leq C/h_n$, hence

$$\begin{aligned}
\mathbb{E}\left[\widehat{f}^x(y)\right] - f^x(y) &= \left(\left(\mathbb{E}\widehat{f}_D(x)\right)^{-1}\mathbb{E}\left(\widehat{f}_N(x, y)\right) - f^x(y)\right) - \left(\mathbb{E}\widehat{f}_D(x)\right)^{-2}\text{Cov}\left(\widehat{f}_N(x, y), \widehat{f}_D(x)\right) \\
&+ \left(\mathbb{E}\widehat{f}_D(x)\right)^{-2} \text{Var}\left(\widehat{f}_D(x)\right) O(h_n^{-1}).
\end{aligned}$$

For the par dispersion we inspire techniques Sarda and Vieu [14] and Bosq Lecoutre [1] and by under expression (5.4), we find that

$$\begin{aligned}
\text{Var}\left[\widehat{f}^x(y)\right] &= \frac{\text{Var}\left[\widehat{f}_N(x, y)\right]}{\left(\mathbb{E}\widehat{f}_D(x)\right)^2} - 2 \frac{\left[\mathbb{E}\widehat{f}_N(x, y)\right] \text{Cov}\left[\widehat{f}_N(x, y), \widehat{f}_D(x)\right]}{\left(\mathbb{E}\widehat{f}_D(x)\right)^3} \\
&+ \text{Var}\left(\widehat{f}_D(x)\right) \frac{\left(\mathbb{E}\widehat{f}_N(x, y)\right)^2}{\left(\mathbb{E}\widehat{f}_D(x)\right)^4} + o\left(\frac{1}{nh_n\phi(h_n)}\right).
\end{aligned} \tag{5.5}$$

Finally, Lemma (1) is a consequence of Corollaries below

COROLLARY 1. *Under conditions of Lemma 1 we have*

$$\frac{\widehat{f}_N(x, y)}{\mathbb{E}\widehat{f}_D(x)} - f^x(y) = B_H^f(x, y)h_n^2 + B_K^f(x, y)h_n + o(h_n^2).$$

COROLLARY 2. *Under conditions of Lemma 1 we have*

$$\text{Var}\left[\widehat{f}_N(x, y)\right] = \frac{\int_{B(0,1)} K^2(\|v\|)g(x, v)d\mu(v)}{nh_n\phi(h_n)} \left(f^x(y) \int H^2(t)dt\right) + o\left(\frac{1}{nh_n\phi(h_n)}\right)$$

COROLLARY 3. *Under conditions of Lemma 1 we have*

$$\text{Cov}\left[\widehat{f}_N(x, y), \widehat{f}_D(x)\right] = \frac{1}{n\phi(h_n)} (f^x(y)) \int_{B(0,1)} K^2(\|v\|)g(x, v)d\mu(v) + o\left(\frac{1}{n\phi(h_n)}\right)$$

COROLLARY 4. Under conditions of Lemma 1 we have

$$\text{Var} \left[\widehat{f}_D(x) \right] = \frac{\int_{B(0,1)} K^2(\|v\|)g(x, v)d\mu(v)}{n\phi(h_n)} + o\left(\frac{1}{n\phi(h_n)}\right)$$

[Proof of Corollary 1] By definition of $\widehat{f}_N(x, y)$ we have

$$\begin{aligned} \mathbb{E}\widehat{f}_N(x, y) &= \frac{1}{nh_n\phi(h_n)} \sum_{i=1}^n \mathbb{E}(K_i(x)H_i(y)) \\ &= \frac{1}{h_n\phi(h_n)} \mathbb{E} \left[K_1(x)H_1\left(\frac{y - Y_i}{h_n}\right) \right] \\ &= \frac{1}{h_n\phi(h_n)} \mathbb{E} \left(K_1(x) \left[\mathbb{E}(H_1(h_n^{-1}(y - Y_i)|X)) \right] \right) \end{aligned} \quad (5.6)$$

for the calculation of $\mathbb{E}(H_1(h_n^{-1}(y - Y_i)|X))$ considering the change of variable $t = h_n^{-1}((y - z)$, we have

$$\begin{aligned} \mathbb{E}(H_1(h_n^{-1}(y - Y_i)|X)) &= \frac{1}{h_n} \int H\left(\frac{y - z}{h_n}\right) f^x(z) dz \\ &= \int H(t) f^x(y - h_n t) dt \end{aligned}$$

Just develop the function $f^x(y - h_n t)$ in the neighborhood of y , which is possible since $f^x(\cdot)$ being a function of class C^2 in y , then, we can use the Taylor expansion of the function $f^x(\cdot)$:

$$f^x(y - h_n t) = f^x(y) - h_n t \frac{\partial f^x(y)}{\partial y} + \frac{h_n^2 t^2}{2} \frac{\partial^2 f^x(y)}{\partial y^2} + o(h_n^2)$$

which gives, under the assumption (H3)

$$\mathbb{E}(H_1|X) = f^x(y) + \frac{h_n^2 t^2}{2} \frac{\partial^2 f^x(y)}{\partial y^2} \int t^2 H(t) dt + o(h_n^2).$$

We replace in equation (5.6) found

$$\mathbb{E}\widehat{f}_N(x, y) = \frac{1}{h_n\phi(h_n)} \left[\mathbb{E}(K_1(x)f^x(y)) + \frac{h_n^2 t^2}{2} \int t^2 H(t) dt \mathbb{E}\left(K_1(x) \frac{\partial^2 f^x(y)}{\partial y^2}\right) \right] + o(h_n^2) \quad (5.7)$$

To simplify the writing of this equation we set $\psi_l(\cdot, y) = \frac{\partial^l f^x(y)}{\partial y^l}$, $l \in \{0, 2\}$.

The function $\psi_l(\cdot, y)$ defined on the functional space \mathcal{F} denotes the one or other of the two functions $\psi_0(\cdot, y) = f^x(y)$ et $\psi_2(\cdot, y) = \frac{\partial^2 f^x(y)}{\partial y^2}$.

The kernel K is assumed compact support, then, for all $l \in \{0, 2\}$ we have

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

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

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

and we find that

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

Therefore we have

$$\begin{aligned}\mathbb{E}\widehat{f}_N(x, y) &= \frac{1}{h_n\phi(h_n)} \left[\psi_0(x, y) \int_{B(0,1)} K(\|v\|)w(h_n, x, v)d\mu(v) \right. \\ &\quad - h_n \int_{B(0,1)} K(\|v\|) \frac{\partial\psi_0(x, y)[v]}{\partial x} w(h_n, x, v)d\mu(v) \\ &\quad + \frac{h_n^2}{2} \int t^2 H(t) dt \left(\psi_2(x, y) \int_{B(0,1)} K(\|v\|)w(h_n, x, v)d\mu(v) \right. \\ &\quad \left. \left. - h_n \int_{B(0,1)} K(\|v\|) \frac{\partial\psi_2(x, y)[v]}{\partial x} w(h_n, x, v)d\mu(v) \right) \right] + o(h_n^2).\end{aligned}$$

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

$$\begin{aligned}\mathbb{E}\widehat{f}_N(x, y) &= \frac{1}{h_n\phi(h_n)} \psi_0(x, y) \int_{B(0,1)} K(\|v\|)w(h_n, x, v)d\mu(v) \\ &\quad - h_n \int_{B(0,1)} K(\|v\|) \frac{\partial\psi_0(x, y)[v]}{\partial x} g(x, v)d\mu(v) \\ &\quad - h_n \int_{B(0,1)} K(\|v\|) \frac{\partial\psi_0(x, y)[v]}{\partial x} \left(\frac{w(h_n, x, v)}{h_n\phi(h_n)} - g(x, v) \right) d\mu(v) \\ &\quad + \frac{h_n^2}{2} \int t^2 H(t) dt \left[\frac{1}{\phi(h_n)} \psi_2(x, y) \int_{B(0,1)} K(\|v\|)w(h_n, x, v)d\mu(v) \right. \\ &\quad - h_n \int_{B(0,1)} K(\|v\|) \frac{\partial\psi_2(x, y)[v]}{\partial x} g(x, v)d\mu(v) \\ &\quad \left. - h_n \int_{B(0,1)} K(\|v\|) \frac{\partial\psi_2(x, y)[v]}{\partial x} \left(\frac{w(h_n, x, v)}{h_n\phi(h_n)} - g(x, v) \right) d\mu(v) \right] + o(h_n^2).\end{aligned}$$

$$\begin{aligned}\mathbb{E}\widehat{f}_N(x, y) &= \frac{1}{h_n\phi(h_n)} \psi_0(x, y) \int_{B(0,1)} K(\|v\|)w(h_n, x, v)d\mu(v) \\ &\quad - h_n \int_{B(0,1)} K(\|v\|) \frac{\partial\psi_0(x, y)[v]}{\partial x} g(x, v)d\mu(v) \\ &\quad + \frac{h_n^2}{2} \int t^2 H(t) dt \left[\frac{1}{h_n\phi(h_n)} \psi_2(x, y) \int_{B(0,1)} K(\|v\|)w(h_n, x, v)d\mu(v) \right] + o(h_n^2).\end{aligned}$$

On the other hand we have

$$\mathbb{E}\widehat{f}_D(x) = \frac{\mathbb{E}K_1}{\phi(h_n)} = \frac{1}{\phi(h_n)} \int_{B(0,1)} K(\|v\|)w(h_n, x, v)d\mu(v). \quad (5.8)$$

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

$$\begin{aligned}\mathbb{E}f_N(x, y) &= \psi_0(x, y)(\mathbb{E}\widehat{f}_D(x)) - h_n \int_{B(0,1)} K(\|v\|) \frac{\partial\psi_0(x, y)[v]}{\partial x} g(x, v)d\mu(v) \\ &\quad + \frac{h_n^2}{2} \int t^2 H(t) dt \left[(\mathbb{E}\widehat{f}_D(x))\psi_2(x, y) \right] + o(h_n^2).\end{aligned}$$

Using the hypothesis (H1), equation (5.8) can be expressed as

$$\mathbb{E}\widehat{f}_D(x) = \int_{B(0,1)} K(\|v\|)g(x, v)d\mu(v) + o(1). \quad (5.9)$$

Finally we arrive at

$$\begin{aligned} (\mathbb{E}\widehat{f}_D(x))^{-1}\mathbb{E}[\widehat{f}_N(x, y)] - f^x(y) &= -h_n \frac{\int_{B(0,1)} K(\|v\|) \frac{\partial f^x(y)[v]}{\partial x} g(x, v)d\mu(v)}{\int_{B(0,1)} K(\|v\|)h(x, v)d\mu(v)} \\ &\quad + \frac{h_n}{2} \frac{\partial^2 f^x(y)[v]}{\partial y^2} \int t^2 H(t)dt + o(h_n^2). \end{aligned} \quad (5.10)$$

[Proof of Corollary 2] By definition of $\widehat{f}_N(x, y)$ we have

$$\begin{aligned} Var(\widehat{f}_N(x, y)) &= \frac{1}{(n(h_n\phi(h_n)))^2} \sum_{i=1}^n Var(K_i(x)H_i(y)) \\ &= \frac{1}{n(h_n\phi(h_n))^2} Var(K_1(x)H_1(x)) \\ &= \frac{1}{n(h_n\phi(h_n))^2} (\mathbb{E}(K_1(x)H_1(y))^2 - (\mathbb{E}(K_1(x)H_1(y)))^2) \\ &= \frac{1}{n(h_n\phi(h_n))^2} \mathbb{E}(K_1(x)H_1(y))^2 - n^{-1} \left(\frac{\mathbb{E}(K_1(x)H_1(y))}{h_n\phi(h_n)} \right)^2. \end{aligned}$$

By Corollary 1 and equation (5.9) we have $\frac{\mathbb{E}K_1(x)H_1(y)}{h_n\phi(h_n)} = \mathbb{E}\widehat{f}_N(x, y) = \mathcal{O}(1)$, and the fact that

$$Var(\widehat{f}_N(x, y)) = \frac{1}{n(h_n\phi(h_n))^2} \mathbb{E}(K_1(x)H_1(y))^2 + o\left(\frac{1}{nh_n\phi(h_n)}\right).$$

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 $(y - z)/h_n^{-1} = t$ we obtain

$$\begin{aligned} \mathbb{E}(K_1(x)H_1(y))^2 &= \mathbb{E}(K_1(x)^2 E(H_1^2(y)|X = x)) \\ &= \frac{1}{h_n^2} \mathbb{E}\left(K_1(x)^2 \int H^2\left(\frac{y-z}{h_n}\right) f^x(z) dz\right) \\ &= \frac{1}{h_n} \mathbb{E}\left(K_1^2(x) \int H^2(t) f^x(y - h_n t) dt\right), \end{aligned}$$

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

$$f^x(y - h_n t) = f^x(y) + \mathcal{O}(h_n) = f^x(y) + o(1).$$

Hence

$$\mathbb{E}(K_1(x)H_1(y))^2 = \frac{1}{h_n} \int H^2(t) dt \mathbb{E}(K_1^2(x) f^x(y)) + o\left(\frac{1}{h_n}\right).$$

The same way and with the same techniques used in the above proof of Corollary 1, 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 $(y - z)/h_n^{-1} = t$ we find that:

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

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

$$\mathbb{E}(K_1(x)H_1(y))^2 = \frac{1}{h_n} \int H^2(t) dt \left(\phi(h_n) f^x(y) \int_{B(0,1)} K^2(\|v\|) g(x, v) d\mu(v) \right) + o\left(\frac{\phi(h_n)}{h_n}\right).$$

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

$$Var\left(\widehat{f}_N(x, y)\right) = \frac{1}{(nh_n\phi(h_n))} \left[f^x(y) \int H^2(t) dt \int_{B(0,1)} K^2(\|v\|) g(x, v) d\mu(v) \right] + o\left(\frac{1}{nh_n\phi(h_n)}\right)$$

[Proof of Corollary 3] By definition of $\widehat{f}_N(x, y)$ and $\widehat{f}_D(x)$ we obtain

$$\begin{aligned}Cov\left(\widehat{f}_N(x, y), \widehat{f}_D(x)\right) &= \frac{1}{n(h_n\phi(h_n))^2} Cov(K_1(x)H_1(y), K_1(x)) \\ &= \frac{1}{n(h_n\phi(h_n))^2} (\mathbb{E}K_1^2(x)H_1(y) - \mathbb{E}K_1(x)H_1(y)\mathbb{E}K_1(x)) \\ &= \frac{\mathbb{E}K_1^2(x)H_1(y)}{n(h_n\phi(h_n))^2} - \left(\frac{\mathbb{E}K_1(x)H_1(y)}{n(h_n\phi(h_n))^2}\right) \left(\frac{\mathbb{E}K_1(x)}{n(h_n\phi(h_n))^2}\right).\end{aligned}$$

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

$$Cov\left(\widehat{f}_N(x, y), \widehat{f}_D(x)\right) = \frac{1}{n\phi(h_n)} (f^x(y)) \int_{B(0,1)} K^2(\|v\|) g(x, v) d\mu(v) + o\left(\frac{1}{n\phi(h_n)}\right). \quad (5.11)$$

[Proof of Corollary 4] By definition of $\widehat{f}_D(x)$ we have

$$\begin{aligned}Var\left(\widehat{f}_D(x)\right) &= \frac{1}{n(\phi(h_n))^2} (Var(K_1)) \\ &= \frac{\mathbb{E}K_1^2(x)}{n(\phi(h_n))^2} - n^{-1} \left(\frac{\mathbb{E}K_1(x)}{\phi(h_n)}\right)^2 \\ &= \frac{\int_{B(0,1)} K^2(\|v\|) g(x, v) d\mu(v)}{n(\phi(h_n))} + o\left(\frac{1}{n\phi(h_n)}\right).\end{aligned} \quad (5.12)$$

This allows us to complete the proof of Lemma 1.

[Proof of Lemma 2] The calculation of the squared error of the conditional distribution is with the same techniques used in the previous theorem (1) 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

$$\mathbb{E}\left[(\widehat{F}^x(y) - F^x(y))^2\right] = \left[\mathbb{E}\left(\widehat{F}^x(y) - F^x(y)\right)\right]^2 + Var\left[\widehat{F}^x(y)\right].$$

For $i = 1, \dots, n$, we consider the quantities $K_i(x) = K(h_n^{-1}\|x - X_i\|)$ be defined as

$$\widehat{g}_N(x, y) = \frac{1}{n\phi(h_n)} \sum_{i=1}^n K_i(x) \mathbf{1}_{\{Y_i \leq y\}}, \quad \widehat{f}_D(x) = \frac{1}{n\phi(h_n)} \sum_{i=1}^n K_i(x).$$

Finally, Lemma 2 can be deduced from following corollary

COROLLARY 5. Under hypotheses (H0)-(H4) and (H6), we have

$$\begin{aligned} \frac{\widehat{g}_N(x, y)}{\mathbb{E}\widehat{f}_D(x)} - F^x(y) &= B_F(x, y)h_n^2 + B_K^F(x, y)h_n + o(h_n^2), \\ \text{Var} [\widehat{g}_N(x, y)] &= \frac{\int_{B(0,1)} K^2(\|v\|)g(x, v)d\mu(v)}{n\phi(h_n)} \left(F^x(y) \int H^2(t)dt \right) + o\left(\frac{1}{n\phi(h_n)}\right), \\ \text{Cov} [\widehat{g}_N(x, y), \widehat{f}_D(x)] &= \frac{1}{n\phi(h_n)} (F^x(y)) \int_{B(0,1)} K^2(\|v\|)g(x, v)d\mu(v) + o\left(\frac{1}{n\phi(h_n)}\right). \end{aligned}$$

REMARK 4 (NOTES ON THE SQUARED ERROR). The "dimensionality" of the observations (resp. model) is used in the expression of the rate of convergence of the two theorems (1) and (2). 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 [4]). 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) (ϕ 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.

References

- [1] Bosq, D. and Lecoutre, J. P.,(1987). *Thorie de l'estimation fonctionnelle Ed. Economica*.
- [2] Dabo-Niang, S., (2002). Estimation de la densit dans un espace de dimension infinie : application aux diffusions, C. R., Math., Acad. Sci. Paris, **334**, 213-216.
- [3] Dabo-Niang, S., (2002). Sur l'estimation fonctionnelle en dimension infinie : Application aux diffusions, These de Doctorat, Universit de Paris 6.
- [4] Ferraty, F., Rabhi, A., Vieu P., (2005). Conditional Quantiles for Dependent Functional Data with Application to the Climatic El Niño Phenomenon. *Sankhyā*, **67**, 378-398.
- [5] Ferraty, F. and Vieu, P., (2006). Nonparametric Functional Data Analysis. *Theory and Practice, Springer Series in Statistics*. Springer, New York.
- [6] Ferraty, F., A., Rabhi, Vieu,P., (2008). Estimation nonparamtrique de la fonction de hasard avec variables explicatives fonctionnelles, *Revue Roumaine de mathematique*.
- [7] Gasser, T. and Hall, P.,Presnell, B., (1998). Nonparametric estimation of the mode of a distribution of random curves, *Journal of the Royal Statistical Society, Ser. B.*, **60**, 681-691.
- [8] Gefeller, O. and Michels, P., (1992). A review on smoothing methods for the estimation of the hazard rate based on kernel functions, In Y. Dodge and J. Whittaker (Eds), *Computational Statistics*, Physica-Verlag, 459-464.
- [9] Hassani, S., Sarda, P., Vieu, P., (1986). Approche non-paramtrique en thorie de la fiabilit: revue bibliographique, *Rev. Statist. Appl.*, **35**, 27-41.

- [10] Izenman, A. , (1991). Developments in nonparametric density estimation, *J. Amer. Statist. Assoc.*, **86**, 205-224.
- [11] Pascu, M. , Vaduva, M., Vaduva I., (2003). Nonparametric estimation of the hazard rate: a survey, *Rev. Roumaine Math. Pures Appl.*, **48**, 173-191.
- [12] Patil, P.N., Wells, M.T., Maron, J.S., (1994). Some heuristics of kernel based estimators of ratio functions, *J. Nonparametr. Statist.*, **4**, 203-209.
- [13] Ramsay, J. and Silverman, B., (2005). *Functional Data Analysis, 2nd Ed.*, Springer Series in Statistics. Springer, New York.
- [14] Sarda, P. and Vieu, P., (2000). Kernel regression. In: M. Schimek (ed.), *Smoothing and regression. Approaches, computation and application*, Wiley Series in Probability and Statistics, Wiley, New York, 43-70.
- [15] Singpurwalla, N. and Wong, M.Y., (1983). Estimation of the failure rate - a survey of nonparametric models. Part I: Non-Bayesian methods, *Comm. Statist. Theory Methods*, **12**, 559-588.
- [16] Troutman, J. L., (1996). *Variational Calculus and Optimal Control*, Springer-Verlag.
- [17] Vieu, P., (1991). Quadratic errors for nonparametric estimates under dependence, *J. Multivariate Anal.*, **39**, 324-347.