APPLICATIONES MATHEMATICAE
40,4 (2013), pp. 405–420

Abbes Rabhi, Samir Benaissa, El Hadj Hamel,
Boubaker Mechab (Sidi Bel Abbes)

MEAN SQUARE ERROR OF THE ESTIMATOR OF THE CONDITIONAL HAZARD FUNCTION

Abstract. This paper deals with a scalar response conditioned by a functional random variable. The main goal is to estimate the conditional hazard function. An asymptotic formula for the mean square error of this estimator is calculated considering as usual the bias and variance.

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 (common for example in medical applications), presence of dependence between the observed variables (common in seismic or econometric applications) or presence of explanatory variables. Many techniques have been studied in the literature to deal with these situations but all deal only with random real and multidimensional explanatory variables.

Now, we encounter more frequently functional data such as curves or images. The data are modeled as realizations of a random variable taking values in an abstract space of infinite dimension. 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. [4]. 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.

2010 *Mathematics Subject Classification*: 62G05, 62G20, 62N02.
Key words and phrases: nonparametric estimation, conditional hazard function, probabilities of small balls, functional variable.

DOI: 10.4064/am40-4-2

The estimators that we define are based on the techniques of convolution kernel. The study of 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 the hazard function and conditional hazard function. One of the most common techniques for constructing such estimators is to study the 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. [10] presented an overview of estimation techniques.

Nonparametric methods are based on the idea of convolution kernel. Such kernels are used with success in estimating both density functions and hazard functions. A wide range of literature in this area is provided by the overviews of Singpurwalla and Wong [13], Hassani et al. [7], Izenman [8], Gefeller and Michels [6] and Pascu and Vaduva [9].

Ramsay and Silverman [11] and Ferraty et al. [4] recently developed methods of nonparametric estimation for functional random variables (i.e. with values in an infinite-dimensional space).

The objective of this paper is to study a model in which a conditional random explanatory variable X is not necessarily real or multidimensional but only supposed to have values in an abstract seminormed space $\mathcal{F}$.

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

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

play a crucial role in asymptotic formulas. The role of small ball probabilities is widely presented in Chapter 13.2 of Ferraty and Vieu [5].

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 seminormed vector space $(\mathcal{F}, d(\cdot\,;\cdot))$, which may be of infinite dimension. We will say that X is a *functional random variable* and we will use the abbreviation *frv*. For a sample of independent pairs (X_i, Y_i), each having the same distribution as (X,Y), our aim is to study mean square convergence of the estimator of the conditional hazard function of a real random variable conditional on one functional variable. Nonparametric estimation of the function is related to the conditional probability distribution (*cond-cdf*) of Y given X. For $x \in \mathcal{F}$, we assume that the regular version of the conditional probability of Y given X exists, denoted by F_Y^X, and has a bounded density with respect to Lebesgue measure over $\mathbb{R}$, denoted by f_Y^X.

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

Our nonparametric models will be quite general in the sense that we will just need a simple assumption on the marginal distribution of X. Set

$$C_B^2(\mathcal{F} \times \mathbb{R}) = \Big\{\varphi : \mathcal{F} \times \mathbb{R} \to \mathbb{R} \mid \forall z \in N_x,\ \varphi(z, \cdot) \in C^2(N_y) \text{ and } \Big(\varphi(\cdot, y), \frac{\partial^2 \varphi(\cdot, y)}{\partial y^2}\Big) \in C_B^1(x) \times C_B^1(x)\Big\}, \tag{1}$$

where $C_B^1(x)$ is the set of those continuously Gateaux differentiable functions on N_x (see Troutman [14] for this type of differentiability) for which the derivative operator of order 1 at a point x is bounded on the unit ball $B(0, 1)$ of the functional space $\mathcal{F}$. Given i.i.d. observations $(X_1, Y_1), \dots, (X_n, Y_n)$ of (X, Y), the kernel estimate of the conditional distribution $F_Y^X(x, y)$, denoted $\widehat{F}_Y^X(x, y)$, is defined by

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

with the convention $\frac{0}{0} = 0$. The function K is a kernel, H is a *cfd* and $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers. As an estimator of $\widehat{f}_Y^X(x, y)$ we take a derivative of $\widehat{F}_Y^X(x, y)$, namely

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

where H' is a kernel (it is the derivative of H). According to our notation the conditional hazard function is given by

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

The main objective is to study the nonparametric estimator

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

of $h_Y^X(x, y)$ when the explanatory variable X is valued in a space of possibly infinite dimension. We give precise asymptotic evaluations of the quadratic error of this estimator.

3. Asymptotic properties. To establish the mean square convergence of the estimator $\widehat{h}_Y^X(x, y)$ to $h_Y^X(x, y)$, we introduce the following assumptions:

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

$$w(r, x, v) = \phi(r)g(x, v) + o(\phi(r)) \quad \text{for all } v \in B(0,1), \tag{3}$$

where
- ϕ is an increasing function with values in $\mathbb{R}^+$,
- g is defined on $\mathcal{F} \times \mathcal{F}$ with values in $\mathbb{R}^+$ and

$$0 < \int\limits_{B(0,1)} g(x, v)\, d\mu(v) < \infty.$$

(H2) The kernel K has compact support and

$$0 < A_3 < K(t) < A_4 < \infty,$$

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

$$H(x) = \int\limits_{-\infty}^{x} H'(t)\, dt, \qquad \int H'(t)\, dt = 1, \qquad \int t^2 H'(t)\, dt < \infty,$$

(H4) $\lim\limits_{n\to\infty} h_K = 0$, $\lim\limits_{n\to\infty} h_H = 0$ and $\lim\limits_{n\to\infty} nh_H\phi(h_K) = \infty$,

(H5) $\exists \tau < \infty$, $f_Y^X(x, y) \le \tau$, $\forall (x, y) \in \mathcal{F} \times \mathcal{S}$,

(H6) $\exists \beta > 0$, $F_Y^X(x, y) \le 1 - \beta$, $\forall (x, y) \in \mathcal{F} \times \mathcal{S}$.

We establish the following results.

THEOREM 3.1. *Under the hypotheses* (H1)–(H6) *and if* $f_Y^X(x, y) \in C_B^2(\mathcal{F} \times \mathbb{R})$ *then*

$$\mathbb{E}[(\widehat{f}_Y^X(x, y) - f_Y^X(x, y))^2] = B_H^2(x, y)h_H^4 + B_K^2(x, y)h_K^2 + \frac{V_{HK}(x, y)}{nh_H\phi(h_K)} + o(h_H^4) + o(h_K^2) + o\left(\frac{1}{nh_H\phi(h_K)}\right) \tag{4}$$

with

$$B_H(x, y) = \frac{1}{2h_H}\,\frac{\partial^2 f^x(y)}{\partial y^2}\int t^2 H'(t)\, dt,$$

$$B_K(x, y) = \frac{\int_{B(0,1)} K(\|v\|) D_x f_Y^X(x, y)[v] g(x, v)\, d\mu(v)}{\int_{B(0,1)} K(\|v\|) g(x, v)\, d\mu(v)},$$

$$V_{HK}(x, y) = \frac{f_Y^X(x, y)\big(\int_{B(0,1)} K^2(\|v\|) g(x, v)\, d\mu(v)\big)\int H^2(t)\, dt}{\big(\int_{B(0,1)} K(\|v\|) g(x, v)\, d\mu(v)\big)^2}.$$

THEOREM 3.2. *Under the hypotheses* (H1)–(H6) *and if* $F_Y^X(x,y) \in C_B^2(\mathcal{F}\times\mathbb{R})$ *then*

$$(5)\quad \mathbb{E}[(\widehat{F}_Y^X(x,y)-F_Y^X(x,y))^2] = B_H'^2(x,y)h_H^4 + B_K'^2(x,y)h_K^2 + \frac{V_{HK}'(x,y)}{n\phi(h_K)} + o(h_H^4)+o(h_K^2)+o\biggl(\frac{1}{n\phi(h_K)}\biggr)$$

with

$$B_H'(x,y) = \frac{1}{2}\frac{\partial^2 F_Y^X(x,y)}{\partial y^2}\int t^2 H(t)\,dt,$$

$$B_K'(x,y) = \frac{\int_{B(0,1)} K(\|v\|)D_x F_Y^X(x,y)[v]g(x,v)\,d\mu(v)}{\int_{B(0,1)} K(\|v\|)g(x,v)\,d\mu(v)},$$

$$V_{HK}'(x,y) = \frac{F_Y^X(x,y)\bigl(\int_{B(0,1)} K^2(\|v\|)g(x,v)\,d\mu(v)\bigr)\int H^2(t)\,dt}{\bigl(\int_{B(0,1)} K(\|v\|)g(x,v)\,d\mu(v)\bigr)^2},$$

where D_x means the derivative with respect to x.

Proof of Theorem 3.1. We have to calculate separately the bias and the variance since

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

We define the quantities

$$K_i(x) = K(h_K^{-1}\|x-X_i\|),\qquad H_i'(y) = H'(h_H^{-1}(y-Y_i))\qquad \text{for all } i=1,\dots,n$$

and we set

$$\widehat{g}_N(x,y) = \frac{1}{n\phi(h_K)}\sum_{i=1}^n K_i(x)H_i'(y),\qquad \widehat{f}_D(x) = \frac{1}{n\phi(h_K)}\sum_{i=1}^n K_i(x)$$

and

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

where H' is the derivative of H.

By a straightforward calculation we obtain

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

from which we deduce

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

with

$$A_1 = \mathbb{E}\widehat{f}_N(x,y)(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x)) = \mathrm{Cov}(\widehat{f}_N(x,y), \widehat{f}_D(x)),$$
$$A_2 = \mathbb{E}(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x))^2 \widehat{f}_Y^X(x,y).$$

We can write

$$\begin{aligned}(6)\qquad \widehat{f}_Y^X(x,y) - f_Y^X(x,y) = {} & \left(\frac{\widehat{f}_N(x,y)}{\mathbb{E}\widehat{f}_D(x)} - f_Y^X(x,y)\right) \\ & - \frac{(\widehat{f}_N(x,y) - \mathbb{E}\widehat{f}_N(x,y))(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x))}{(\mathbb{E}\widehat{f}_D(x))^2} \\ & - \frac{(\mathbb{E}\widehat{f}_N(x,y))(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x))}{(\mathbb{E}\widehat{f}_D(x))^2} \\ & + \frac{(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x))^2}{(\mathbb{E}\widehat{f}_D(x))^2} \widehat{f}_Y^X(x,y),\end{aligned}$$

which implies

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

Now we 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 bound $\widehat{f}_Y^X(x,y)$ by a constant $C > 0$, as $\widehat{f}_Y^X(x,y) \le C/h_H$; hence

$$\begin{aligned}(7)\qquad &\mathbb{E}[\widehat{f}_Y^X(x,y)] - f_Y^X(x,y) \\ &= \big((E\widehat{f}_D(x))^{-1}\mathbb{E}(\widehat{f}_N(x,y)) - f_Y^X(x,y)\big) - \big((\mathbb{E}\widehat{f}_D(x))^{-2}\mathrm{Cov}(\widehat{f}_N(x,y)), \widehat{f}_D(x)\big) \\ &\qquad + (\mathbb{E}\widehat{f}_D(x))^{-2}\,\mathrm{Var}(\widehat{f}_D(x))O(h_H^{-1}).\end{aligned}$$

By results of Sarda and Vieu [12] and Bosq and Lecoutre [1] we obtain

$$\begin{aligned}(8)\qquad \mathrm{Var}[\widehat{f}_Y^X(x,y)] = {} & \frac{\mathrm{Var}[\widehat{f}_N(x,y)]}{(\mathbb{E}\widehat{f}_D(x))^2} - 2\frac{[\mathbb{E}\widehat{f}_N(x,y)]\mathrm{Cov}[\widehat{f}_N(x,y), \widehat{f}_D(x)]}{(\mathbb{E}\widehat{f}_D(x))^3} \\ & + \mathrm{Var}(\widehat{f}_D(x))\frac{(\mathbb{E}\widehat{f}_N(x,y))^2}{(\mathbb{E}\widehat{f}_D(x))^4} + o\left(\frac{1}{nh_H\phi(h_K)}\right).\end{aligned}$$

Finally, Theorem 3.1 is a consequence of the lemmas below. ∎

LEMMA 3.1. *Under the conditions of Theorem* 3.1 *we have*

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

LEMMA 3.2. *Under the conditions of Theorem* 3.1 *we have*

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

LEMMA 3.3. *Under the conditions of Theorem* 3.1 *we have*

$$\mathrm{Cov}[\widehat{f}_N(x,y),\widehat{f}_D(x)] = \frac{1}{n\phi(h_K)}f_Y^X(x,y)\int\limits_{B(0,1)} K^2(\|v\|)g(x,v)\,d\mu(v) + o\bigg(\frac{1}{n\phi(h_K)}\bigg).$$

LEMMA 3.4. *Under the conditions of Theorem* 3.1 *we have*

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

Proof of Lemma 3.1. By definition of $\widehat{f}_N(x,y)$ we have

$$\begin{aligned}\mathbb{E}\widehat{f}_N(x,y) &= \frac{1}{nh_H\phi(h_K)}\sum_{n=1}^{\infty}\mathbb{E}(K_i(x))H_i'(y) \\ &= \frac{1}{h_H\phi(h_K)}\bigg[\mathbb{E}K_1(x)H_1'\bigg(\frac{y-Y_i}{h_H}\bigg)\bigg] \\ &= \frac{1}{h_H\phi(h_K)}\mathbb{E}(K_1(x)[\mathbb{E}(H_1'(h_H^{-1}(y-Y_i)/X))]).\end{aligned} \tag{9}$$

To calculate $\mathbb{E}(H_1'(h_H^{-1}(y-Y_i))/X)$, considering the change of variable $t = h_H^{-1}(y-z)$, we have

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

We can use the Taylor expansion of the function f_Y^X:

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

which gives, under the assumption (H3),

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

We insert this in (9) to obtain

$$\mathbb{E}\widehat{f}_N(x,y) = \frac{1}{\phi(h_K)}\bigg[\mathbb{E}(K_1(x) f_Y^X(X,y)) + \frac{h_H^2 t^2}{2}\int t^2 H'(t)\,dt \\ \mathbb{E}\bigg(K_1(x)\frac{\partial^2 f_Y^X(X,y)}{\partial y^2}\bigg)\bigg] + o(h_H^2). \tag{10}$$

To simplify the writing, set

$$\psi_l(\cdot,y) = \frac{\partial^l f_Y^X(x,y)}{\partial y^l}, \qquad l \in \{0,2\}.$$

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

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

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

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

and we find that

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

Therefore

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

$$- h_K \int_{B(0,1)} K(\|v\|) \frac{\partial \psi_2(x,y)[v]}{\partial x} w(h_K,x,v)\, d\mu(v) \Big) \Big]$$
$$+ o(h_H^2) + o(h_K).$$

By (H1), we set $w(h_K,x,v) = \phi(h_K)g(x,v) + o(\phi(h_K))$. Then

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

Hence

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

On the other hand,

$$\mathbb{E}\widehat{f}_D(x) = \frac{\mathbb{E}K_1}{\phi(h_K)} = \frac{1}{\phi(h_K)} \int_{B(0,1)} K(\|v\|)w(h_K,x,v)\, d\mu(v). \tag{11}$$

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

$$\mathbb{E} f_N(x,y) = \psi_0(x,y)(\mathbb{E}\widehat{f}_D(x)) - h_K \int\limits_{B(0,1)} K(\|v\|)\frac{\partial\psi_0(x,y)[v]}{\partial x} g(x,v)\,d\mu(v) + \frac{h_H^2}{2}\int t^2 H'(t)\,dt\,[(\mathbb{E}\widehat{f}_D(x))\psi_2(x,y)] + o(h_H^2) + o(h_K).$$

Using the hypothesis (H1), equation (11) can be expressed as

$$\mathbb{E}\widehat{f}_D(x) = \int\limits_{B(0,1)} K(\|v\|) g(x,v)\,d\mu(v) + o(1). \tag{12}$$

Finally we arrive at

$$\begin{aligned}(\mathbb{E}\widehat{f}_D(x))^{-1}\mathbb{E}[\widehat{f}_N(x,y)] - f_Y^X(x,y) &= -h_K\frac{\int_{B(0,1)} K(\|v\|)\frac{\partial f^x(y)[v]}{\partial x} g(x,v)\,d\mu(v)}{\int_{B(0,1)} K(\|v\|)h(x,v)\,d\mu(v)} \\ &\quad + \frac{h_H}{2}\,\frac{\partial^2 f^x(y)[v]}{\partial y^2}\int t^2 H'(t)\,dt + o(h_H^2) + o(h_K^2). \quad \blacksquare\end{aligned} \tag{13}$$

Proof of Lemma 3.2. By definition of $\widehat{f}_N(x,y)$ we have

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

By Lemma 3.1 and (12) we have

$$\frac{\mathbb{E}(K_1(x)H_1'(y))}{h_H\phi(h_K)} = \mathbb{E}\widehat{f}_N(x,y) = O(1),$$

and

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

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

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

By Taylor expansion of order 1 at y we show that for n large enough

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

Hence

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

In the same way and with the same techniques used in the above proof of Lemma 3.1, we show that it suffices now to estimate the amount $\mathbb{E}(K_1(x)H_1'(y))^2$. Indeed, similar to the previous proof, conditioning on X and considering the usual change of variable $(y-z)/h_H = t$ we find that

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

with $\|v\| = h_K^{-1}\|x-X\|$; this allows us to conclude that

$$\begin{aligned}&\mathbb{E}(K_1(x)H_1'(y))^2 \\ &\quad= \frac{\int H'^2(t)\,dt}{h_H}\Big(\phi(h_K)f_Y^X(x,y)\int_{B(0,1)} K^2(\|v\|)g(x,v)\,d\mu(v)\Big) + o\Big(\frac{\phi(h_K)}{h_H}\Big).\end{aligned}$$

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

$$\begin{aligned}&\mathrm{Var}(\widehat{f}_N(x,y)) \\ &\quad= \frac{1}{nh_H\phi(h_K)}\Big[f_Y^X(x,y)\int H'^2(t)\,dt\int_{B(0,1)} K^2(\|v\|)g(x,v)\,d\mu(v)\Big] \\ &\qquad + o\Big(\frac{1}{nh_H\phi(h_K)}\Big). \ \blacksquare\end{aligned}$$

Proof of Lemma 3.3. By definition of $\widehat{f}_N(x,y)$ and $\widehat{f}_D(x)$ we obtain

$$\begin{aligned}\mathrm{Cov}(\widehat{f}_N(x,y),\widehat{f}_D(x)) &= \frac{1}{n(h_H\phi(h_K))^2}\mathrm{Cov}(K_1(x)H_1'(y),K_1(x))\\ &= \frac{1}{n(h_H\phi(h_K))^2}\big(\mathbb{E}K_1^2(x)H_1'(y)-\mathbb{E}K_1(x)H_1'(y)\mathbb{E}K_1(x)\big)\\ &= \frac{\mathbb{E}K_1^2(x)H_1'(y)}{n(h_H\phi(h_k))^2}-\frac{\mathbb{E}K_1(x)H_1'(y)}{n(h_H\phi(h_K))^2}\,\frac{\mathbb{E}K_1(x)}{n(h_H\phi(h_K))^2}.\end{aligned}$$

The proof of this lemma is very similar to the one of Lemma 3.1. We just replace H_1^2 with H_1, then using the fact that

$$\frac{\mathbb{E}(K_1(x)H_1(y))}{\phi(h_K)}=O(1) \quad \text{and} \quad \frac{\mathbb{E}(K_1(x))}{\phi(h_K)}=O(1)$$

we deduce that

$$\begin{aligned}(14)\qquad &\mathrm{Cov}(\widehat{f}_N(x,y),\widehat{f}_D(x))\\ &= \frac{1}{n\phi(h_K)}(f_Y^X(x,y))\int\limits_{B(0,1)}K^2(\|v\|)g(x,v)\,d\mu(v)+o\bigg(\frac{1}{n\phi(h_K)}\bigg). \qquad \blacksquare\end{aligned}$$

Proof of Lemma 3.4. By definition of $\widehat{f}_D(x)$ we have

$$\begin{aligned}(15)\qquad \mathrm{Var}(\widehat{f}_D(x)) &= \frac{1}{n\phi(h_K)^2}\mathrm{Var}(K_1)\\ &= \frac{\mathbb{E}K_1^2(x)}{n\phi(h_K)^2}-n^{-1}\bigg(\frac{\mathbb{E}K_1(x)}{\phi(h_K)}\bigg)\\ &= \frac{\int_{B(0,1)}K^2(\|v\|)g(x,v)\,d\mu(v)}{n\phi(h_K)}+o\bigg(\frac{1}{n\phi(h_K)}\bigg).\end{aligned}$$

This allows us to complete the proof of Theorem 3.1. ■

Proof of Theorem 3.2. The calculation of the squared error of the conditional distribution is with the same techniques used in the proof of Theorem 3.1.

We calculate two parts: the bias and the variance. The squared error of the conditional distribution can be expressed as

$$\begin{aligned}&\mathbb{E}[(\widehat{F}_Y^X(x,y)-F_Y^X(x,y))^2]\\ &\qquad = [\mathbb{E}(\widehat{F}_Y^X(x,y))-F_Y^X(x,y)]^2+\mathrm{Var}[\widehat{F}_Y^X(x,y)].\end{aligned}$$

For $i=1,\dots,n$, we consider the quantities

$$K_i(x)=K(h_K^{-1}\|x-X_i\|),$$

$$H_i(y)=H(h_H^{-1}(y-Y_i))$$

and define

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

Finally, Theorem 3.2 can be deduced from the following lemmas.

Lemma 3.5. *Under the hypotheses* (H1)–(H6) *we have*

$$\frac{\widehat{g}_N(x,y)}{\mathbb{E}\widehat{f}_D(x)} - F_Y^X(x,y) = B'_H(x,y)h_H^2 + B'_K(x,y)h_K + o(h_H^2) + o(h_K).$$

Lemma 3.6. *Under the hypotheses* (H1)–(H6) *we have*

$$\mathrm{Var}[\widehat{g}_N(x,y)] = \frac{\int_{B(0,1)} K^2(\|v\|) g(x,v)\, d\mu(v)}{n\phi(h_K)} \Big(F_Y^X(x,y) \int H^2(t)\, dt \Big) + o\bigg(\frac{1}{n\phi(h_K)}\bigg).$$

Lemma 3.7. *Under the hypotheses* (H1)–(H6) *we have*

$$\mathrm{Cov}[\widehat{g}_N(x,y), \widehat{f}_D(x)] = \frac{1}{n\phi(h_K)} F_Y^X(x,y) \int\limits_{B(0,1)} K^2(\|v\|) g(x,v)\, d\mu(v) + o\bigg(\frac{1}{n\phi(h_K)}\bigg). \ \blacksquare$$

Theorem 3.3. *Under the hypotheses* (H1)–(H6) *and if* $F_Y^X(x,y)$, $f_Y^X(x,y) \in C_B^2(\mathcal{F} \times \mathbb{R})$ *then*

$$\begin{aligned}\mathrm{MSE}\, \widehat{h}_Y^X(x,y) &\equiv \mathbb{E}[(\widehat{h}_Y^X(x,y) - h_Y^X(x,y))^2] \\ &\le \mathbb{E}[(\widehat{f}_Y^X(x,y) - f_Y^X(x,y))^2] + \mathbb{E}[(\widehat{F}_Y^X(x,y) - F_Y^X(x,y))^2].\end{aligned}$$

Proof of Theorem 3.3. This proof is based on the decomposition

$$(16)\qquad |\widehat{h}_Y^X(x,y) - h_Y^X(x,y)|$$
$$\begin{aligned}&= \frac{1}{1-\widehat{F}_Y^X(x,y)} \bigg[(\widehat{f}_Y^X(x,y) - f_Y^X(x,y)) + \frac{f_Y^X(x,y)}{1-F_Y^X(x,y)} (\widehat{F}_Y^X(x,y) - F_Y^X(x,y)) \bigg] \\ &\le \frac{1}{1-\widehat{F}_Y^X(x,y)} \bigg[(\widehat{f}_Y^X(x,y) - f_Y^X(x,y)) + \frac{\tau}{\beta} (\widehat{F}_Y^X(x,y) - F_Y^X(x,y)) \bigg] \\ &\le (\widehat{f}_Y^X(x,y) - f_Y^X(x,y)) + \frac{\tau}{\beta} (\widehat{F}_Y^X(x,y) - F_Y^X(x,y)).\end{aligned}$$

Therefore

$$\mathbb{E}|\widehat{h}_Y^X(x,y) - h_Y^X(x,y)|^2$$
$$\leq \mathbb{E}\left[(\widehat{f}_Y^X(x,y) - f_Y^X(x,y)) + \frac{\tau}{\beta}(\widehat{F}_Y^X(x,y) - F_Y^X(x,y))\right]^2.$$

Now Theorem 3.3 can be deduced from Theorems 3.1 and 3.2. ∎

4. Remarks and comments

Notes on the functional variable. The hypothesis (H1) on the functional variable X can be divided into two parts:

(i) The first part is rarely used in nonparametric statistical functionals, because it requires the introduction of a reference measurement of the functional space. However, in this paper we impose this condition. It allows us to achieve a natural generalization of the squared error obtained by Vieu [15] in the vector case.

(ii) The second part (3) is a classic property in functional analysis. This is a simple asymptotic separation of variables. This condition is designed to be able to adapt traditional techniques of the case of different functionals, even if the reference measure μ does not have the properties of the Lebesgue measure, such as translation invariance and homogeneity.

Note that the hypothesis (H1) is used to describe the phenomenon of concentration of the probability measure of the explanatory variable X, because

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

Note also that the first part of the hypothesis (H1) is satisfied when, for example, X is a diffusion process satisfying the standard conditions (see Dabo-Niang [2]). 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 $Z = r^{-1}(x - X)$ is $w(r,x,v) = r^p f(x - rv)$, where f is the density of X and p the dimension, therefore $w(r,x,v) = r^p f(x) + o(r^p)$.

Notes on the nonparametric model. In this paper, we choose a differentiability condition, as our goal is to find an expression for the rate of convergence that is explicit, asymptotically exact and keeps the usual form of the squared error (see Vieu [15]). However, if one assumes a Lipschitz condition, for example the conditional density of the type

$$\forall(y_1, y_2) \in N_y \times N_y, \forall(x_1, x_2) \in N_x \times N_x,$$
$$|f^{x_1}(y_1) - f^{x_2}(y_2)| \leq A_x((d(x_1, x_2)^2) + |y_1 - y_2|^2)$$

which is less restrictive than the condition (1), we obtain a result for the conditional distribution and conditional density respectively for example of the type

$$\mathbb{E}[(\widehat{F}_Y^X(x,y) - F_Y^X(x,y))^2] = O(h_H^4 + h_K^4) + O\left(\frac{1}{n\phi(h_k)}\right),$$
$$\mathbb{E}[(\widehat{f}_Y^X(x,y) - f_Y^X(x,y))^2] = O(h_H^4 + h_K^4) + O\left(\frac{1}{nh_H\phi(h_k)}\right).$$

But such an expression (implicit) for the rate of convergence will not allow us to properly determine the smoothing parameter. In other words, our differentiability condition 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.

Notes on the squared error. The "dimensionality" of the observations (resp. model) is used in the expression of the rate of convergence of Theorems 3.1 and 3.2. We find the dimensionality of the model in such a way that the dimensionality of the variable in the functional dispersion biases 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. Our asymptotical results highlight the importance of the concentration properties on small balls of the probability measure of the underlying functional variable. This stresses the role of the semimetric. A suitable choice of this parameter allows us to obtain an interesting solution to the problem of curse of dimensionality (see [3]). 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 decomposes into two main parts: the bias part proportional to h_K (resp. h_H), and the dispersion part inversely proportional to h_K (resp. h_H) (ϕ is an increasing function depending on h_K). This makes it relatively easy to choose the parameter so as to minimize the main part of this expression.

Acknowledgments. The authors are grateful to the anonymous referee for his/her comments which have contributed to a nicer presentation of this work.

References

[1] D. Bosq et J.-P. Lecoutre, *Théorie de l'estimation fonctionnelle*, Economica, 1987.

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

[3] F. Ferraty, A. Rabhi and P. Vieu, *Conditional quantiles for dependent functional data with application to the climatic El Niño phenomenon*, Sankhyā 67 (2005), 378–398.

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

[5] F. Ferraty and P. Vieu, *Nonparametric Functional Data Analysis. Theory and Practice*, Springer, New York, 2006.

[6] O. Gefeller and 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.

[7] S. Hassani, P. Sarda et P. Vieu, *Approche non-paramétrique en théorie de la fiabilité: revue bibliographique*, Rev. Statist. Appl. 35 (1986), 27–41.

[8] A. Izenman, *Recent developments in nonparametric density estimation*, J. Amer. Statist. Assoc. 86 (1991), 205–224.

[9] M. Pascu and I. Văduva, *Nonparametric estimation of the hazard rate: a survey*, Rev. Roumaine Math. Pures Appl. 48 (2003), 173–191.

[10] P. N. Patil, M. T. Wells and J. S. Marron, *Some heuristics of kernel based estimators of ratio functions*, J. Nonparametric Statist. 4 (1994), 203–209.

[11] J. Ramsay and B. Silverman, *Functional Data Analysis*, 2nd ed., Springer, New York, 2005.

[12] P. Sarda and P. Vieu, *Kernel regression*, in: M. Schimek (ed.), Smoothing and Regression. Approaches, Computation, and Application, Wiley, New York, 2000, 43–70.

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

[14] J. L. Troutman, *Variatonal Calculus and Optimal Control*, Springer, 1996.

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

Abbes Rabhi, Samir Benaissa, El Hadj Hamel, Boubaker Mechab
Laboratory of Mathematics
Faculty of Science
University of Sidi Bel Abbes P.B. 89
Sidi Bel Abbes 22000, Algeria
E-mail: rabhi_ abbes@yahoo.fr
benaissamir@yahoo.fr
rbe0223@yahoo.fr
mechaboub@yahoo.fr

Received on 14.11.2011;
revised version on 8.11.2013 (2110)