Journal of Statistics Applications & Probability Letters
An International Journal

http://dx.doi.org/10.12785/jsapl/020303

Exact Asymptotic Errors of the Hazard Conditional Rate Kernel

Naouel Belkhir[1,*], *Abbes Rabhi*[2] *and Sara Soltani*[3]

[1] Laboratory of Mathematics Abu Bakr Belkaid University of Tlemcen, Algeria
[2] Laboratory of Mathematics, University of Sidi Bel Abbes, Algeria
[3] Stochastic Models, Statistics and Applications Moulay Tahar University of Saida, Saida, Algeria

Received: 16 Feb. 2015, Revised: 2 Aug. 2015, Accepted: 3 Aug. 2015
Published online: 1 Sep. 2015

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 and we investigate the asymptotic normality of the kernel conditional hazard function estimator.

Keywords: Asymptotic normality, functional data, Kernel conditional hazard function, kernel estimation, nonparametric estimation, probabilities of small balls.

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.* [9]. 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.

In the case where the data is incomplete, Lemdani and Ould-Saïd [14], they give the asymptotic mean integrated squared error and the mean squared error for the kernel estimator of the hazard rate from truncated and censored data.

Recently, Djebbouri *et al.* [5] studied the mean squared convergence rate and are proved the asymptotic normality for functional mixing data case of this estimate.

The estimators that we define are based on the techniques of convolution kernel. 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.* [17] presented an overview of estimation techniques. The non-parametric methods based on the ideas of convolution kernel, which are known for their

* Corresponding author e-mail: n.belkhir@yahoo.com

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 [20], Hassani *et al.* [12], Izenman [13], Gefeller and Michels [11] and Pascu and Vaduva [16].

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 [18] and Ferraty and Vieu [8] 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) and/or random variables real. In this context Arfi [1] is study the almost sure convergence of the kernel type estimator of the hazard function is shown under $\widetilde{\rho}$-mixing condition with censored data over a sequence of compact sets which increases to $\mathbb{R}^d$.

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 $\mathscr{F}$ semi-normed.

As with any problem of nonparametric estimation, the dimension of the space $\mathscr{F}$ 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

$$\phi_x(h) = \mathbb{P}(X \in B(x,h)) = \mathbb{P}\left(X \in \{x' \in \mathscr{F}, \|x - x'\| < h\}\right), \tag{1}$$

intervene directly in the asymptotic behavior of any estimator nonparametric functional (see Ferraty *et al.* [7]). 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 $(\mathscr{F}, \|\cdot\|)$ 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 independant pairs (X_i, Y_i), each having the same distribution as (X,Y), our aim is to study the mean square convergence 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. For $x \in \mathscr{F}$, 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 $\mathbb{R} \times \mathscr{F}$ and $N_x \times \mathscr{S}_\mathbb{R}$ will denote a fixed neighborhood of (x,y), $\mathscr{S}_\mathbb{R}$ will be a fixed compact subset of $\mathbb{R}$, and we will use the notation $B(x,h) = \{x' \in \mathscr{F} / \|x' - x\| < h\}$. 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:

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

where $C_B^1(x)$ is the set of continuously differentiable functions to sens of Gâteaux on N_x (see Troutman [21] 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 $\mathscr{F}$. 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:

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

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

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

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

$$\forall x \in \mathscr{F}, \ \forall y \in \mathbb{R} \qquad h_Y^X(x,y) = \frac{f_Y^X(x,y)}{1 - F_Y^X(x,y)} = \frac{f_Y^X(x,y)}{S_Y^X(x,y)}. \tag{3}$$

The main objective is to study the the nonparametric estimate $\widehat{h}_Y^X(x,y)$ of $h_Y^X(x,y)$.

Furthermore, the estimator $\widehat{h}_Y^X(x,y)$ can we written as

$$\widehat{h}_Y^X(x,y) = \frac{\widehat{f}_Y^X(x,y)}{1 - \widehat{F}_Y^X(x,y)} = \frac{\widehat{f}_N(x,y)}{\widehat{f}_D(x) - \widehat{g}_N(x,y)}, \tag{4}$$

where

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

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 asymptotic properties

To establish the convergence in mean square of the estimator $\widehat{h}_Y^X(x,y)$ 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:

$$w(r,x,v) = \phi(r)g(x,v) + o(\phi(r)) \text{ for all } v \in B(0,1), \tag{5}$$

where
(i) ϕ is an increasing function with values in $\mathbb{R}^+$.
(ii) g is defined on $\mathscr{F} \times \mathscr{F}$, 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' exists and is such that there exist two constants C and C' with $-\infty < C < K'(t) < C' < 0$ for $0 \le t \le 1$.

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

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

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

(H4) The bandwidth h_K satisfies:

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

(H5) $\begin{cases} \exists \tau < \infty, f_Y^X(x,y) \le \tau, \forall (x,y) \in \mathscr{F} \times \mathscr{S}_{\mathbb{R}}, \text{ and;} \\ \forall (x_1,x_2) \in N_x^2, \ \forall (y_1,y_2) \in \mathscr{S}_{\mathbb{R}}^2, \ |f_Y^X(x_1,y_1) - f_Y^X(x_2,y_2)| \le C_x\left(\|x_1 - x_2\|^{b_1} + |y_1 - y_2|^{b_2}\right). \end{cases}$

(H6) $\begin{cases} \exists \beta > 0, F_Y^X(x,y) \le 1 - \beta, \forall (x,y) \in \mathscr{F} \times \mathscr{S}_{\mathbb{R}}, \text{ and;} \\ \forall (x_1,x_2) \in N_x^2, \ \forall (y_1,y_2) \in \mathscr{S}_{\mathbb{R}}^2, \ |F_Y^X(x_1,y_1) - F_Y^X(x_2,y_2)| \le C_x\left(\|x_1 - x_2\|^{b_1} + |y_1 - y_2|^{b_2}\right). \end{cases}$

3.1 Remarks on the assumptions

Remark 3.1. Assumption (1) 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 $\psi(z)$ and $\varphi(h)$ as $\phi_z(h) = \psi(z)\varphi(h) + o(\varphi(h))$.

This idea was adopted by Masry [15] who reformulated the Gasser *et al.* [10] one. The increasing property of $\phi_x(.)$ implies that $\zeta_h^x(.)$ is bounded and then integrable (all the more so $\zeta_0^x(.)$ 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 $\mathscr{S} = \mathbb{R}^d$, 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.* [6]):

1. $\phi_x(h) \approx \psi(h)h^\gamma$ for som $\gamma > 0$.
2. $\phi_x(h) \approx \psi(h)h^\gamma \exp\{C/h^p\}$ for som $\gamma > 0$ and $p > 0$.
3. $\phi_x(h) \approx \psi(h)/|\ln h|$.

The function $\beta_h^x(.)$ which intervenes in Assumption (H4) is increasing for all fixed h. Its pointwise limit $\beta_0^x(.)$ 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 $\beta_0(u) := \beta_0^x(u)$ in the above examples by:

1. $\beta_0(u) = u^\gamma$,
2. $\beta_0(u) = \delta_1(u)$ where $\delta_1(.)$ is Dirac function,
3. $\beta_0(u) = \mathbf{1}_{]0,1]}(u)$.

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

Theorem 3.1. Under the hypothesis (H1)-(H6) and if $F_Y^X(x,y)$ (resp. $f_Y^X(x,y)) \in C_B^2(\mathscr{F}\times\mathbb{R})$ then

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

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

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

and

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

Proof. This proof is based on the decomposition

$$\begin{aligned}
\widehat{h}_Y^X(x,y) - h_Y^X(x,y) &= \frac{\widehat{f}_Y^X(x,y)}{1-\widehat{F}_Y^X(x,y)} - \frac{f_Y^X(x,y)}{1-F_Y^X(x,y)} \\
&= \frac{1}{1-\widehat{F}_Y^X(x,y)}\left[(\widehat{f}_Y^X(x,y) - f_Y^X(x,y)) + \frac{f_Y^X(x,y)}{1-F_Y^X(x,y)}(\widehat{F}_Y^X(x,y) - F_Y^X(x,y))\right] \\
&= \frac{1}{\widehat{f}_D^X(x) - \widehat{g}_N(x,y)}\left(\widehat{f}_N(x,y) - \mathbb{E}\widehat{f}_N(x,y)\right) + h_Y^X(x,y)\left(\mathbb{E}\widehat{g}_N(x,y) - F_Y^X(x,y)\right) \\
&\quad + \frac{1}{\widehat{f}_D^X(x) - \widehat{g}_N(x,y)}\left[\left(\mathbb{E}\widehat{f}_N(x,y) - f_Y^X(x,y)\right) + h_Y^X(x,y)\left(1 - \mathbb{E}\widehat{g}_N(x,y) - \left(\widehat{f}_D(x) - \widehat{g}_N(x,y)\right)\right)\right], \quad (6)
\end{aligned}$$

where D_x means the derivative with respect to x.

Hence

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

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

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

Then, Theorem 3.1 can be deduced from both lemmas above Lemma 3.1 and Lemma 3.2. □.

Lemma 3.1. Under the hypothesis (H1)-(H6) and if $f_Y^X(x,y) \in C_B^2(\mathscr{F}\times\mathbb{R})$ then:

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

where

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

Lemma 3.2 Under the hypothesis (H1)-(H6) and if $F_Y^X(x,y) \in C_B^2(\mathscr{F}\times\mathbb{R})$ then:

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

with

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

Remark 3.2. Observe that, the result of this lemmas Lemma 3.1 and Lemma 3.2 permits to write

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

and

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

Proof of Lemma 3.1. 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

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

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 $i = 1,\ldots,n$.

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

We come at the following to writing:

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

from which we draw:

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

as

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

Can be written as

$$\begin{aligned}\widehat{f}_Y^X(x,y) - f_Y^X(x,y) &= \left(\frac{\widehat{f}_N(x,y)}{\mathbb{E}\widehat{f}_D(x)} - f_Y^X(x,y)\right) - \frac{\left(\widehat{f}_N(x,y) - \mathbb{E}\widehat{f}_N(x,y)\right)\left(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x)\right)}{\left(\mathbb{E}\widehat{f}_D(x)\right)^2} \\ &- \frac{\left(\mathbb{E}\widehat{f}_N(x,y)\right)\left(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x)\right)}{\left(\mathbb{E}\widehat{f}_D(x)\right)^2} + \frac{\left(\widehat{f}_D(x) - \mathbb{E}\widehat{f}_D(x)\right)^2}{\left(\mathbb{E}\widehat{f}_D(x)\right)^2}\widehat{f}_Y^X(x,y), \end{aligned} \tag{7}$$

which implies

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

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

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

For the par dispersion we inspire techniques Sarda and Vieu [19]and Bosq Lecoutre [2] and by under expression (7), we find that

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

Finally, Lemma 3.1 is a consequence of Corollaries below. □.

Corollary 3.1. Under the conditions of Lemma 3.1 we have

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

Corollary 3.2. Under the conditions of Lemma 3.1 we have

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

Corollary 3.3. Under the conditions of Lemma 3.1 we have

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

Corollary 3.4. Under the conditions of Lemma 3.1 we have

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

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

$$\begin{aligned}\mathbb{E}\widehat{f}_N(x,y) &= \frac{1}{nh_H\phi(h_K)}\sum_{n=1}^{\infty}\mathbb{E}(K_i(x))H_i'(y)\\ &= \frac{1}{h_H\phi(h_K)}\left[\mathbb{E}K_1(x)H_1'\left(\frac{y-Y_i}{h_H}\right)\right]\\ &= \frac{1}{h_H\phi(h_K)}\mathbb{E}\left(K_1(x)\left[\mathbb{E}\left(H_1'(h_H^{-1}(y-Y_i)/X)\right)\right]\right),\end{aligned} \tag{8}$$

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

$$\mathbb{E}\left(H_1'(h_H^{-1}(y-Y_i)/X)\right)=\frac{1}{h_H}\int H'\left(\frac{y-z}{h_H}\right)f^x(z)dz=\int H'(t)f^x(y-h_Ht)dt.$$

Just develop the function $f_Y^X(x,y-h_Ht)$ 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:

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

which gives, under the assumption (H3)

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

We replace in equation (8) found

$$\mathbb{E}\widehat{f}_N(x,y)=\frac{1}{h_H\phi(h_K)}\left[\mathbb{E}\left(K_1(x)f_Y^X(x,y)\right)+\frac{h_H^2t^2}{2}\int t^2H'(t)dt\ \mathbb{E}\left(K_1(x)\frac{\partial^2 f_Y^X(x,y)}{\partial y^2}\right)\right]+o(h_H^2). \tag{9}$$

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

The function $\psi_l(\cdot,y)$ defined on the functional space $\mathscr{F}$ denotes the one or other of the two functions $\psi_0(\cdot,y)=f_Y^X(x,y)$ and $\psi_2(\cdot,y)=\dfrac{\partial^2 f_Y^X(x,y)}{\partial y^2}$.

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

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

The function $\psi_l(\cdot,y)$ 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

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

Therefore we have

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

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

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

Thus

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

On the other hand we have

$$\mathbb{E}\widehat{f}_D(x) = \frac{\mathbb{E}K_1}{\phi(h_K)} = \frac{1}{\phi(h_K)}\int_{B(0,1)}K(\|v\|)w(h_K,x,v)d\mu(v), \tag{10}$$

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

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

Using the hypothesis (H1), equation (10) can be expressed as

$$\mathbb{E}\widehat{f}_D(x) = \int_{B(0,1)}K(\|v\|)g(x,v)d\mu(v) + o(1). \tag{11}$$

Finally we arrive at

$$\begin{aligned}(\mathbb{E}\widehat{f}_D(x))^{-1}\mathbb{E}\left[\widehat{f}_N(x,y)\right] - f_Y^X(x,y) = & -h_K\frac{\int_{B(0,1)}K(\|v\|)\ \frac{\partial f^x(y)[v]}{\partial x}g(x,v)d\mu(v)}{\int_{B(0,1)}K(\|v\|)h(x,v)d\mu(v)}\\ & +\frac{h_H}{2}\frac{\partial^2 f^x(y)[v]}{\partial y^2}\int t^2H'(t)dt + o(h_H^2) + o(h_K^2).\end{aligned} \tag{12}$$

□.

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

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

By Corollary 3.1 and equation (11) we have $\frac{(\mathbb{E}K_1(x)H_1'(y))}{h_H\phi(h_K)} = \mathbb{E}\widehat{f}_N(x,y) = \mathscr{O}(1)$, and the fact that

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

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

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

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

$$f_Y^X(x,y-h_Ht) = f_Y^X(x,y) + \mathscr{O}(h_H) = f_Y^X(x,y) + o(1).$$

Hence

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

The same way and with the same techniques used in the above proof of Corollary 3.1, 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 $(y-z)/h_H^{-1} = t$ we find that:

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

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

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

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

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

$\square$.

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

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

The proof of this Corollary is very similar to the one used for Corollary 3.1. To do this, replace H_1^2 with H_1 then using the fact that $\frac{(\mathbb{E}K_1(x)H_1(y))}{\phi(h_K)}=\mathscr{O}(1)$ and $\frac{(\mathbb{E}K_1(x))}{\phi(h_K)}=\mathscr{O}(1)$ we deduce that

$$Cov\left(\widehat{f}_N(x,y),\widehat{f}_D(x)\right)=\frac{1}{n(\phi(h_k)}(f_Y^X(x,y))\int_{B(0,1)}K^2(\|v\|)g(x,v)d\mu(v)+o\left(\frac{1}{n\phi(h_K)}\right). \tag{13}$$

□.

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

$$\begin{aligned}Var\left(\widehat{f}_D(x)\right) &= \frac{1}{n(\phi(h_K)^2}(Var(K_1))\\ &= \frac{\mathbb{E}K_1^2(x)}{n(\phi(h_K)^2}-n^{-1}\left(\frac{\mathbb{E}K_1(x)}{\phi(h_K)}\right)\\ &= \frac{\int_{B(0,1)}K^2(\|v\|)g(x,v)d\mu(v)}{n(\phi(h_K)}+o\left(\frac{1}{n\phi(h_K)}\right).\end{aligned} \tag{14}$$

This allows us to complete the proof of Lemma 3.1. □.

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

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

Finally, Lemma 3.2 can be deduced from following corollaries.

Corollary 3.5. Under the hypotheses (H1)-(H6) we have

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

Corollary 3.6. Under the hypotheses (H1)-(H6) we have

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

Corollary 3.7. Under the hypotheses (H1)-(H6) we have

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

□.

Remark 3.3. It is clear that, the results of Corollaries Corollary 3.2-3.4 and Corollary 3.6-3.7 allows to write

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

3.2 Asymptotic normality

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

Theorem 3.2. Assume that (H1)-(H6) hold, and if the following equation (1) is verified, then we have for any $x \in \mathscr{A}$,

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

where

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

and $\xrightarrow{\mathscr{D}}$ 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 3.8. Under the hypotheses of Theorem 4.1 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

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

Proof. Consider the decomposition

$$\begin{aligned}\widehat{h}_Y^X(x,y) - h_Y^X(x,y) = & \frac{1}{\widehat{f}_D(x) - \widehat{g}_N(x,y)}\left(\widehat{f}_N(x,y) - E\,\widehat{f}_N(x,y)\right)\\ & + \frac{1}{\widehat{f}_D(x) - \widehat{g}_N(x,y)}\left\{h_Y^X(x,y)\left(\mathbb{E}\widehat{g}_N(x,y) - F_Y^X(x,y)\right) + \left(\mathbb{E}\,\widehat{f}_N(x,y) - f_Y^X(x,y)\right)\right\}\\ & + \frac{h_Y^X(x,y)}{\widehat{f}_D(x) - \widehat{g}_N(x,y)}\left\{1 - \mathbb{E}\widehat{g}_N(x,y) - \left(\widehat{f}_D(x) - \widehat{g}_N(x,y)\right)\right\}. \end{aligned} \quad (15)$$

Therefore, Theorem 3.2 and Corollary 3.8 are a consequence of Lemma 3.1, Lemma 3.2 and the following results.

Lemma 3.3. Under the hypotheses of Theorem 3.2

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

Lemma 3.4. Under the hypotheses of Theorem 3.2

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

and

$$\left(\frac{nh_H\phi_x(h_K)}{\sigma_h^2(x,y)}\right)^{1/2}\left(\widehat{f}_D(x) - \widehat{g}_N(x,y) - 1 + \mathbb{E}[\widehat{g}_N(x,y)]\right) = o_{\mathbb{P}}(1).$$

□.

Proof of Lemma 3.3. Define

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

and

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

Thus

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

So, our claimed result is now

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

Therefore, we have

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

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

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

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

$$\begin{aligned}
\mathbb{E}\left(H_1'^2(y)|X_1\right) &= \int_{\mathbb{R}} H'^2\left(\frac{\|y-u\|}{h_H}\right) f_Y^X(x,u)du \\
&\leq h_H\int_{\mathbb{R}} H'^2(t)\left(f_Y^X(y-h_Ht,x) - f_Y^X(x,y)\right)dt + h_H f_Y^X(x,y)\int_{\mathbb{R}} H'^2(t)dt \\
&\leq h_H^{1+b_2}\int_{\mathbb{R}} |t|^{b_2}H'^2(t)dt + h_H f_Y^X(x,y)\int_{\mathbb{R}} H'^2(t)dt \\
&= h_H\left(o(1) + f_Y^X(x,y)\int_{\mathbb{R}} H'^2(t)dt\right). \qquad (16)
\end{aligned}$$

By means of (16) and the fact that, as $n \to \infty$, $\mathbb{E}\left(K_1^2(x)\right) \longrightarrow \beta_2\phi_x(h_K)$, one gets

$$Var(\Delta_1(x,y)) = \beta_2\phi_x(h_K)h_H\left(o(1) + f(x,y)\int_{\mathbb{R}} H'^2(t)dt\right).$$

So, using (H4), we get

$$\begin{aligned}
\frac{1}{n(h_H\mathbb{E}[K_1(x)])^2}Var(\Delta_1(x,y)) &= \frac{\beta_2\phi_x(h_K)}{n(\beta_1h_H\phi_x(h_K))^2}h_H\left(o(1) + f_Y^X(x,y)\int_{\mathbb{R}} H'^2(t)dt\right) \\
&= o\left(\frac{1}{nh_H\phi_x(h_K)}\right) + \frac{\beta_2 f_Y^X(x,y)}{\beta_1^2nh_H\phi_x(h_K)}\int_{\mathbb{R}} H'^2(t)dt.
\end{aligned}$$

Thus as $n \to \infty$ we obtain

$$\frac{1}{n(h_H\mathbb{E}[K_1(x)])^2}Var(\Delta_1(x,y)) \longrightarrow \frac{\beta_2 f_Y^X(x,y)}{\beta_1^2nh_H\phi_x(h_K)}\int_{\mathbb{R}} H'^2(t)dt. \qquad (17)$$

Finally, the proof of Lemma is completed by using result (17), to get

$$Var(\Omega_n) \longrightarrow \frac{\beta_2}{\beta_1^2}f_Y^X(x,y)\int_{\mathbb{R}} H'^2(t)dt = \sigma_f^2(x,y).$$

□.

Proof of Lemma 3.4.

It is clear that, the result of Corollary 3.2, Corollary 3.4 and Corollary 3.6 permits us

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

and

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

then

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

Moreover, the asymptotic variance of $\widehat{f}_D(x) - \widehat{g}_N(x,y)$ given in Remark 3.3 allows to obtain

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

By combining result with the fact that

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

Finally, wee obtain the claimed result. $\square$.

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 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 [22] 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 [4]).
   (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):

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

   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:
   $\mathbb{P}(X \in B(x,r)) = \int_{B(0;1)} w(r,x,v)d\mu(v) = \phi(r)\int_{B(0,1)} g(x,v)d\mu(v) + o(\phi(r)) > 0.$
   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 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 [22]). However, if one proceeds by a Lipschitz condition for example the conditional density of type:

$$\forall (y_1,y_2) \in \mathscr{S}_{\mathbb{R}} \times \mathscr{S}_{\mathbb{R}}, \forall (x_1,x_2) \in N_x \times N_x,$$

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

   which is less restrictive than the condition (2), we obtain a result for the conditional distribution and conditional density respectively for example of type:

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

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

   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 3.1 and Lemma 3.2. 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 [7]). 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.

Acknowledgement

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

References

[1] M. Arfi, Nonparametric Estimation for the Hazard Function. Communications in Statistics-Theory and Methods **42**(14), 2543-2550 (2013).
[2] D. Bosq, J.P. Lecoutre, Thorie de l'estimation fonctionnelle. Ed. Economica, 1987.
[3] S. Dabo-Niang, Estimation de la densit dans un espace de dimension infinie : application aux diffusions. C. R., Math., Acad. Sci. Paris **334**, 213-216 (2002).
[4] S. Dabo-Niang, Sur l'estimation fonctionnelle en dimension infinie : Application aux diffusions. Thse de Doctorat, Universit de Paris 6, 2002.
[5] T. Djebbouri, E.H. Hamel, A. Rabhi, On conditional hazard function estimate for functional mixing data. New Trends in Mathematical Sciences (NTMSCI) **3**(2), 79-95 (2015).
[6] F. Ferraty, A. Mas, P. Vieu, Advances in nonparametric regression for functional variables. Australian and New Zealand Journal of Statistics **49**, 1-20 (2007).
[7] F. Ferraty, A. Rabhi and P. Vieu, Conditional Quantiles for Dependent Functional Data with Application to the Climatic El Nïno Phenomenon. Sankhyã: The Indian Journal of Statistics, Special Issue on Quantile Regression and Related Methods **67**, 378-399 (2005).
[8] F. Ferraty and P. Vieu, Nonparametric Functional Data Analysis. Theory and Practice. Springer Series in Statistics. Springer, New York, 2006.
[9] F. Ferraty, A. Rabhi and P. Vieu, Estimation nonparamétrique de la fonction de hasard avec variables explicatives fonctionnelles. Revue Roumaine de mathématique **52**(2), 1-18 (2008).
[10] T. Gasser, P. Hall, B. Presnell, Nonparametric estimation of the mode of a distribution of random curves, Series B. Journal of the Royal Statistical Society, **60**, 681-691 (1998).
[11] 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, pp. 459-464, Physica-Verlag, 1992.
[12] S. Hassani, P. Sarda et P. Vieu, Approche non-paramtrique en thorie de la fiabilit: revue bibliographique. Rev. Statist. Appl. **35**, 27-41 (1986).
[13] A. Izenman, Developments in nonparametric density estimation. J. Amer. Statist. Assoc. **86**, 205-224 (1991).
[14] M. Lemdani, E. Ould-Saïd, Asymptotic behavior of the hazard rate kernel estimator under truncated and censored data. Communications in StatisticsTheory and Methods **36**(1), 155-173 (2007).
[15] E. Masry, Non-parametric regression estimation for dependent functional data: Asymptotic normality. Stoch. Process. Appl. **115**, 155-177 (2005).
[16] M. Pascu, M. and I. Vaduva, Nonparametric estimation of the hazard rate: a survey. Rev. Roumaine Math. Pures Appl. **48**, 173-191 (2003).
[17] P.N. Patil, M.T. Wells and J.S. Maron, Some heuristics of kernel based estimators of ratio functions. J. Nonparametr. Statist. **4**, 203-209 (1994).
[18] J. Ramsay and B. Silverman, Functional Data Analysis, 2nd Ed. Springer Series in Statistics. Springer, New York, 2005.
[19] 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, pp. 43-70, 2000.
[20] 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**, 559-588 (1983).
[21] J.L. Troutman. Variatonal Calculus and Optimal Control. Springer-Verlag, 1996.
[22] P. Vieu, Quadratic errors for nonparametric estimates under dependence. J.Multivariate Anal. **39**, 324-347 (1991).