REVSTAT – Statistical Journal
Volume 12, Number 3, November 2014, 273–297

THE k NEAREST NEIGHBORS ESTIMATION OF THE CONDITIONAL HAZARD FUNCTION FOR FUNCTIONAL DATA

Authors: Mohammed Kadi Attouch
– Lab. de Statistique Processus Stochastiques, Univ. Djillali Liabès, Sidi Bel Abbès, BP 89, Sidi Bel Abbès 22000, Algeria
attou_kadi@yahoo.fr

Fatima Zohra Belabed
– Lab. de Statistique Processus Stochastiques, Univ. Djillali Liabès, Sidi Bel Abbès, BP 89, Sidi Bel Abbès 22000, Algeria
zahira_bell@yahoo.fr

Received: July 2012 Revised: June 2013 Accepted: October 2013

Abstract:

• In this paper, we study the nonparametric estimator of the conditional hazard function using the k nearest neighbors (k-NN) estimation method for a scalar response variable given a random variable taking values in a semi-metric space. We give the almost complete convergence (its corresponding rate) of this estimator and we establish the asymptotic normality. Then the effectiveness of this method is exhibited by a comparison with the kernel method estimation given in Ferraty *et al.* ([12]) and Laksaci and Mechab ([15]) in both cases simulated data and real data.

Key-Words:

• *functional data; nonparametric regression; k-NN estimator; the conditional hazard function; rate of convergence; random bandwidth; asymptotic normality.*

AMS Subject Classification:

• 62G05, 62G08, 62G20, 62G35.

1. INTRODUCTION

The conditional hazard function remains an indispensable tool in survival analysis and many other fields (medicine, reliability or seismology).

The nonparametric estimation of this function in the case of multivariate data is abundant. The first works date back to Waston and Leadbetter ([31]), they introduce the hazard estimate method, since, several results have been developed, see for example, Roussas ([26]) (for previous works), Li and Tran ([18]) (for recent references). The literature has paid quite some attention to nonparametric hazard rate estimation when the data are functional. The first work which deals with this question is Ferraty *et al.* ([12]). They established the almost complete convergence of the kernel estimate of the conditional hazard function in the independent case. This result was extended to the dependent case by Quintela-del-Río ([23]), he treats the almost complete convergence, the mean quadratic convergence and the asymptotic normality of this estimate. The uniform version of the almost complete convergence (with rate) in the i.i.d. case was obtained by Ferraty *et al.* ([10]). Recently, Laksaci and Mechab ([16]) consider the spatial case. The almost complete convergence rate of an adapted estimate of this model are given.

Estimating the conditional hazard function is closely related to the conditional density, and for the last one, the bandwidth selection is very important for the performance of an estimate. The bandwidth must not be too large, so as to prevent over-smoothing, i.e. substantial bias, and must not be too small either, so as prevent detecting the underlying structure. Particularly, in nonparametric curve estimation, the smoothing parameter is critical for the performance.

Starting from this point of view, this work deals with the nonparametric estimation with k nearest neighbors method k-NN, more precisely we consider a kernel estimator of the hazard function constructed from a local window to take into account the exact k nearest neighbors with real response variable Y and functional curves X.

The k nearest neighbor or k-NN estimator is a weighted average of response variables in the neighborhood of x. The existent bibliography of the k-NN method estimation dates back to Royall ([27]) and Stone ([30]) and has received, since, continuous developments (Mack ([20]) derived the rates of convergence for the bias and variance as well as asymptotic normality in the multivariate case, Collomb ([4]) studied different types of convergence (probability, a.s., a.co.) of the estimator of the regression function. Devroye ([6]) obtained the strong consistency and the uniform convergence. For the functional data studies, the k-NN kernel estimate was first introduced in the monograph of Ferraty and Vieu ([13]), Burba *et al.* ([2]) obtained the rate of almost complete convergence of the regression function using the k-NN method for independent data and the asymptotic normality of robust nonparametric regression function was established in Attouch and Benchikh ([1]).

This paper is organized as follows. In Section 2 we present the model and the k-NN estimator. Section 3, is dedicated to fix notations, hypotheses and the presentation of the main results, the almost complete convergence and the asymptotic normality. Section 4 is devoted to some applications in several problems of nonparametric statistics. Some technical auxiliary results are deployed in Section 5, subsequently, in Section 6, we show the proofs of our main result.

2. MODELS AND ESTIMATORS

Let $(X_i, Y_i)_{i=\overline{1,n}}$ be an independent sequence identically distributed (i.i.d.) as (X, Y) which is a random pair valued in $\mathcal{E}\times\mathbb{R}$. Here $(\mathcal{E}, d)$ is a semi-metric space. $\mathcal{E}$ is not necessarily of a finite dimension, and we do not suppose the existence of a density for the functional random variable X.

Our goal, in this article, is to estimate the conditional hazard function defined by:

$$h^X(Y) \;=\; \frac{f^X(Y)}{1 - F^X(Y)}\;, \tag{2.1}$$

where

$f^X(Y)$ is the conditional density function of Y given X,
$F^X(Y)$ is the conditional distribution function of Y given X.

For a fixed $x \in \mathcal{E}$, the k-NN kernel estimator of $h^x(Y{=}y)$ is given by:

$$\widehat{h}^x_{k-NN}(Y{=}y) \;=\; \widehat{h}^x(y) \;=\; \frac{\widehat{f}^x(y)}{1 - \widehat{F}^x(y)}\;, \tag{2.2}$$

with

$$\begin{aligned} F^x(y) \;&=\; \mathbb{P}\big[Y{\le}\, y \,/\, X{=}\,x\big] \\ &=\; \mathbb{E}\big[1\!\!1_{]-\infty,y]} \,/\, X{=}\,x\big] \\ &=\; r\big(1\!\!1_{]-\infty,y]}\big)\;, \end{aligned}$$

where $r(\cdot)$ is the regression function defined in Ferraty and Vieu ([13]). Therefore:

$$\widetilde{F}^x(y) \;=\; \widehat{r}\big(1\!\!1_{]-\infty,y]}\big) \;=\; \frac{\displaystyle\sum_{i=1}^n 1\!\!1_{]-\infty,y]}\, K\big(H_n^{-1} d(x, X_i)\big)}{\displaystyle\sum_{i=1}^n K\big(H_n^{-1} d(x, X_i)\big)}\;.$$

Finally, by Roussas ([25]), Samanta ([28]) and Ferraty and Vieu ([13]), the estimator of the conditional distribution function is given by

$$\text{(2.3)} \qquad \widehat{F}^x(y) \;=\; \frac{\displaystyle\sum_{i=1}^{n} K\big(H_n^{-1} d(x,X_i)\big)\, R\big(g_n^{-1}(y-Y_i)\big)}{\displaystyle\sum_{i=1}^{n} K\big(H_n^{-1} d(x,X_i)\big)}\,, \qquad \forall\, y \in \mathbb{R}\,,$$

where K is an asymmetrical kernel, H_n is a positive random variable, defined as follows:

$$\text{(2.4)} \qquad H_n(x) \;=\; \min\left\{ h \in \mathbb{R}^+ / \sum_{i=1}^{n} 1\!\!1_{B(x,h)}(X_i) = k \right\},$$

with

$$B(x,h) \;=\; \Big\{ x' \in \mathcal{E};\; d(x,x') < h \Big\}\,.$$

R is a distribution function and $(g_n)_{n\in\mathbb{N}}$ is a sequence of strictly positive real numbers (depending on n). Under a differentiability assumption of $\widehat{F}^x(y)$, we can obtain the conditional density function by differentiating the conditional distribution function, then we have

$$\widehat{f}^x(y) \;=\; \frac{\partial}{\partial y}\widehat{F}^x(y)$$

and then

$$\text{(2.5)} \qquad \widehat{f}^x(y) \;=\; \frac{\displaystyle\sum_{i=1}^{n} K\big(H_n^{-1} d(x,X_i)\big)\, g_n^{-1} R'\big(g_n^{-1}(y-Y_i)\big)}{\displaystyle\sum_{i=1}^{n} K\big(H_n^{-1} d(x,X_i)\big)}\,.$$

In parallel, in order to emphasize differences between the k-NN method and the traditional kernel approach, we define the estimator of the conditional hazard function Ferraty *et al.* ([12]) by:

$$\text{(2.6)} \qquad \widehat{h}^x_{\text{kernel}}(y) \;=\; \frac{\widehat{f}^x_{\text{kernel}}(y)}{1-\widehat{F}^x_{\text{kernel}}(y)}\,,$$

with

$$\text{(2.7)} \qquad \widehat{f}^x_{\text{kernel}}(y) \;=\; \frac{\displaystyle\sum_{i=1}^{n} K\big(h_n^{-1} d(x,X_i)\big)\, g_n^{-1} R'\big(g_n^{-1}(y-Y_i)\big)}{\displaystyle\sum_{i=1}^{n} K\big(h_n^{-1} d(x,X_i)\big)}$$

and

$$\widehat{F}^x_{\text{kernel}}(y) \;=\; \frac{\displaystyle\sum_{i=1}^{n} K\big(h_n^{-1} d(x, X_i)\big)\, R\big(g_n^{-1}(y - Y_i)\big)}{\displaystyle\sum_{i=1}^{n} K\big(h_n^{-1} d(x, X_i)\big)}\;, \tag{2.8}$$

where K is a kernel, R is a distribution function and $(h_n)_{n\in\mathbb{N}}$, $(g_n)_{n\in\mathbb{N}}$ are sequences of strictly positive numbers.

3. ASYMPTOTIC PROPERTIES OF THE k-NN METHOD

3.1. The almost complete convergence (a.co.)

We focus in the pointwise the almost complete convergence[1] and rate of convergence[2] of the k-NN estimator of the conditional hazard function $\widehat{h}^x(y)$ defined on (2.2).

Before giving the main asymptotic result, we need some assumptions. The first one is about the concentration function $\varphi_x(h)$ and can be interpreted as a small ball probability of the functional variable x given by:

(H1)
$$\begin{aligned}\varphi_x(h) \;&=\; \mathbb{P}\big(X \in B(x,h)\big)\\ &=\; \mathbb{P}\big[X \in \big\{x' \in \mathcal{E};\; d(x,x') < h\big\}\big]\;,\end{aligned}$$

with $\varphi_x(h)$ continuous and strictly increasing in a neighborhood of 0 and $\varphi_x(0) = 0$.

(H2) We also need a kernel K:
The kernel K is a function from $\mathbb{R}$ into $\mathbb{R}^+$, we say that K is a kernel of type I, so that: there exist two real constants C_1, C_2, $0 < C_1 < C_2 < \infty$, such that

$$C_1 \mathbb{1}_{[0,1]} < K < C_2 \mathbb{1}_{[0,1]}\;.$$

[1]Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real random variables. We say that $(X_n)_{n\in\mathbb{N}}$ converges almost completely (a.co.) to some r.r.v. X if and only if:

$$\forall\, \epsilon > 0\;, \qquad \sum_{n=1}^{\infty} \mathbb{P}\big[|X_n - X| > \epsilon\big] < \infty\;.$$

[2]Let $(u_n)_{n\in\mathbb{N}}$ be a sequence of positive real number. We say that $X_n = O_{\text{a.co.}}(u_n)$ if and only if: $\exists\, \epsilon > 0$, so that, $\displaystyle\sum_{n=1}^{\infty} \mathbb{P}\big[|X_n| > \epsilon\, u_n\big] < \infty$. This kind of convergence implies both almost sure convergence and convergence in probability.

K is a kernel of type II, so that: the support of K is $[0,1]$ and if its derivative K' exists on $[0,1]$ and satisfies, for two real constants $-\infty < C_1 < C_2 < 0$,

$$C_1 < K' < C_2 \ .$$

In this case, we also suppose that: $\exists\, C_3 > 0$, $\exists\, \epsilon_0$:

$$\forall\, \epsilon < \epsilon_0 \ , \qquad \int_0^{\epsilon} \varphi_x(u)\, du \ > \ C_3\, \epsilon\, \varphi_x(\epsilon) \ .$$

(H3) R is a differentiable function such that:

$$\exists\, C < \infty \, , \quad \forall\, (x_1, x_2) \in \mathbb{R}^2, \qquad |R'(x_1) - R'(x_2)| \le C|x_1 - x_2| \ .$$
$$R' \quad \text{is of the support compact } \ [-1,1] \ .$$

(H4) $\exists\, \zeta > 0$:

$$\begin{cases} \forall\, (x_1, x_2) \in \mathbb{R}^2 \, , \quad |R(x_1) - R(x_2)| \ \le \ C|x_1 - x_2| \ , \\ \displaystyle\int |t|^{\zeta} R'(t)\, dt \ < \ \infty \ . \end{cases}$$

(H5) $(g_n)_{n\in\mathbb{N}}$ is a strictly positive sequence such that:

$$\begin{cases} \displaystyle\lim_{n\to\infty} g_n = 0 \ , \qquad \exists\, a > 0, \ \ \lim_{n\to\infty} n^a g_n = \infty \ , \\ \displaystyle\lim_{n\to\infty} \frac{\log n}{n\, g_n\, \varphi_x(h)} \ = \ 0 \ . \end{cases}$$

The nonparametric model of the function h^x will be determined by regularity conditions of the conditional distribution of Y given X. These conditions are:

(H6) N_x will denote a fixed neighborhood of x, S will be a fixed compact subset of $\mathbb{R}$:
We will consider two kinds of nonparametric models. The first one is called the "Lipschitz-type" model that is defined:

$$Lip_{\mathcal{E}\times\mathbb{R}} : \begin{cases} f : \mathcal{E}\times\mathbb{R} \to \mathbb{R} \, , \quad \forall\, (x_1, x_2) \in N_x^2 \, , \quad \forall\, (y_1, y_2) \in S^2, \\ \qquad\qquad \exists\, C < \infty \, , \quad \exists\, \alpha, \beta > 0 \, , \\ \big| f(x_1, y_1) - f(x_2, y_2) \big| \ \le \ C\big(d(x_1, x_2)^{\alpha} + |y_1 - y_2|^{\beta} \big) \, . \end{cases}$$

(H7) The second one, called the "Continuity type" model, is defined as:

$$C^0_{\mathcal{E}\times\mathbb{R}} = \Big\{ f : \mathcal{E}\times\mathbb{R} \to \mathbb{R} \, , \ \forall\, x' \in N_x, \ \lim_{d(x,x')\to 0} f(x', y) = f(x, y) \Big\} \, .$$

(H8) Finally, we will consider the conditional moments of the response random variable Y:

$$\forall\, m \ge 2, \quad \mathbb{E}\big[|Y|^m \,/\, X = x \big] \ = \ \sigma_m(x) < \infty \ ,$$

with $\sigma_m(\cdot)$ continuous on x.

Before studying the k-NN estimator, we remind asymptotic properties of $\widehat{h}^x_{\text{kernel}}$ defined by equation (2.6). Ferraty *et al.* ([12]), showed the almost complete convergence of this estimator.

Theorem 3.1.

- *In the "continuity type" model and under the assumptions (H1), (H2), (H6) and (H8) we have:*

$$\widehat{h}^x_{\text{kernel}}(y) \longrightarrow h^x(y) \qquad \text{a.co.}$$

- *Under the "Lipschitz type" model and the hypotheses (H1), (H2), (H3), (H5), (H8), we have:*

$$\widehat{h}^x_{\text{kernel}}(y) - h^x(y) \;=\; O(h_n^\alpha) + O(g_n^\beta) + O\left(\sqrt{\frac{\log n}{n\,\varphi_x(h)}}\right).$$

Now we state the almost complete convergence for the nonparametric k-NN method estimate, defined in (2.2).

Theorem 3.2. *In the "continuity type" model and under the hypotheses (H1), (H2), (H4), (H5) and (H6), suppose that $k = k_n$ is a sequence of positive real numbers such that $\dfrac{k_n}{n} \to 0$ and $\dfrac{\log n}{k_n} \to 0$, then we have:*

$$\lim_{n\to\infty} \widehat{h}^x(y) \;=\; h^x(y) \qquad \text{a.co.}$$

Proof: We consider the following decomposition:

$$\widehat{h}^x(y) - h^x(y) \;=\; \frac{1}{1-\widehat{F}^x(y)}\left[\widehat{f}^x(y) - f^x(y)\right] + h^x(y)\,\frac{1}{1-\widehat{F}^x(y)}\left[\widehat{F}^x(y) - F^x(y)\right]. \tag{3.1}$$

Then the proof of Theorem 3.2 can be deduced from the following intermediate results. □

Lemma 3.1. *Under the hypotheses of Theorem 3.2, we have:*

$$\lim_{n\to\infty} \widehat{f}^x(y) = f^x(y) \qquad \text{a.co.} \tag{3.2}$$

and

$$\lim_{n\to\infty} \widehat{F}^x(y) = F^x(y) \qquad \text{a.co.} \tag{3.3}$$

Lemma 3.2. *Under the hypotheses of Theorem 3.2, we have:*

$$\exists\,\delta > 0\,, \qquad \sum_{n\in\mathbb{N}} \mathbb{P}\left[\left(1 - \widehat{F}^x(y)\right) < \delta\right] < \infty\,. \tag{3.4}$$

Theorem 3.3. *The hypotheses (H1)–(H8) imply*

$$\widehat{h}^x(y) - h^x(y) \;=\; O\left(\varphi_x^{-1}\left(\frac{k}{n}\right)^{\alpha}\right) + O\big(g_n^{\beta}\big) + O\left(\sqrt{\frac{\log n}{k_n\, g_n}}\right) \qquad \text{a.co.}$$

Proof: We consider the decomposition (3.1), and the proof of this Theorem is a consequence of these results. □

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

$$\widehat{f}^x(y) - f^x(y) \;=\; O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O\big(g_n^{\beta}\big) + O\left(\sqrt{\frac{\log n}{k_n\, g_n}}\right) \qquad \text{a.co.} \tag{3.5}$$

Lemma 3.4. *Under the hypotheses of Theorem (3.3), we have:*

$$\widehat{F}^x(y) - F^x(y) \;=\; O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O\big(g_n^{\beta}\big) + O\left(\sqrt{\frac{\log n}{k_n}}\right) \qquad \text{a.co.} \tag{3.6}$$

3.2. Asymptotic normality

This section contains results on the asymptotic normality of $\widehat{h}^x(y)$. For this, we have to add the followings assumptions:

(**H9**) For each sequence $U_n \downarrow 0$ as $n \to \infty$ of positive real numbers, there exists a function $\lambda(\cdot)$ such that:

$$\forall\, t \in [0,1]\,, \quad \lim_{U_n \to \infty} \frac{\varphi_x(t U_n)}{\varphi_x(U_n)} = \lambda(t)\ .$$

(**H10**) $\displaystyle \lim_{n\to\infty} \left(g_n^2 - \varphi_x^{-1}\left(\frac{k}{n}\right)\right)\sqrt{k_n} = 0$ and $\displaystyle \frac{1}{k_n g_n} = o\big(g_n^{\beta}\big)$.

Theorem 3.4. *Assume that (H1), (H9), (H10) hold, then for any $x \in \mathcal{A}$, we have:*

$$\left(\frac{k_n\, g_n}{\sigma_h^2(x,y)}\right)^{1/2} \Big[\widehat{h}^x(y) - h^x(y)\Big] \xrightarrow{\ \mathcal{D}\ } \mathcal{N}(0,1) \qquad \textit{as} \quad n \to \infty\ , \tag{3.7}$$

where

$$\sigma_h^2(x,y) \;=\; \frac{\alpha_2\, h^x(y)}{\alpha_1^2\big(1 - F^x(y)\big)} \tag{3.8}$$

$$\Big(\textit{with:}\ \ \alpha_j = K^j(1) - \int_0^1 (K^j)'(s)\,\lambda(s)\,ds \quad \textit{for}\ \ j = 1,2\Big)\,,$$

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

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

Proof: We consider the decomposition (3.1) and we show that the proof of Theorem (3.4) is a consequence of the following results. □

Lemma 3.5. *Under the hypotheses of Theorem (3.4), we have:*

$$\left(\frac{k_n\, g_n}{\sigma_f^2(x,y)}\right)^{1/2} \left[\widehat{f}^x(y) - f^x(y)\right] \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \qquad \text{as} \quad n \to \infty\ ,$$

where

$$\sigma_f^2(x,y) \;=\; f^x(y) \int R'^2(t)\, dt\ . \tag{3.9}$$

Lemma 3.6. *Under the hypotheses of Theorem 3.4, we have:*

$$\left(\frac{k_n\, g_n}{\sigma_F^2(x,y)}\right)^{1/2} \left[\widehat{F}^x(y) - F^x(y)\right] \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \qquad \text{as} \quad n \to \infty\ ,$$

where

$$\sigma_F^2(x,y) \;=\; F^x(y)\left[1 - F^x(y)\right]. \tag{3.10}$$

Lemma 3.7. *Under the hypotheses of Theorem 3.4, we have:*

$$\left(1 - \widehat{F}^x(y)\right) \to \left(1 - F^x(y)\right) \qquad \textit{in probability}\,.$$

4. APPLICATIONS

4.1. Conditional Confidence Interval

The main application of the Theorem (3.4) is the to build confidence interval for the true value of $h^x(y)$ for a given curve $X = x$. A plug-in estimate for the asymptotic standard deviation $\sigma(x,\theta_x)$ can be obtained using the estimators $\widehat{h}^x(y)$ and $\widehat{F}^x(y)$ of $h^x(y)$, $F^x(y)$ respectively. We get $\widehat{\sigma}(x,y) := \left(\dfrac{\widehat{\alpha_2}\widehat{h}^x(y)}{(\widehat{\alpha_1})^2\,(1-\widehat{F}^x(y))}\right)^{1/2}$.
Then $\widehat{h}^x(y)$ can be used to get the following approximate $(1-\zeta)$ confidence interval for $h^x(y)$

$$\widehat{h}^x(y) \pm t_{1-\zeta/2} \times \left(\frac{\widehat{\sigma}_n^2(x,y)}{g_n k}\right)^{1/2}$$

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

We estimate empirically α_1 and α_2 by

$$\widehat{\alpha_1} = \frac{1}{kg(x)} \sum_{i=1}^{n} K_i \qquad \text{and} \qquad \widehat{\alpha_2} = \frac{1}{kg(x)} \sum_{i=1}^{n} K_i^2 ,$$

where $K_i = K\left(\dfrac{d(x, X_i)}{\phi^{-1}(k/n)}\right)$.

This last estimation is justified by the fact that, under (H1), (H5) and (H6), we have, (see Ferraty and Vieu ([13]) p. 44)

$$\frac{1}{kg(x)} \mathbb{E}[K_1^j] \to \alpha_j , \qquad j = 1, 2 .$$

4.2. A Simulation study

In this section we will show the effectiveness of k-NN method compared to the kernel estimation using simulated data. For this we considered a sample of a diffusion process on interval $[0,1]$, $Z_1(t) = 2 - \cos(\pi t W)$ and $Z_2(t) = \cos(\pi t W)$, where W is the standard normal distribution and take $X(t) = AZ_1(t) + (1-A)Z_2(t)$, where A is random variable Bernoulli distributed. We carried out the simulation with a 200-sample of the curve X which is represented by the following graph:

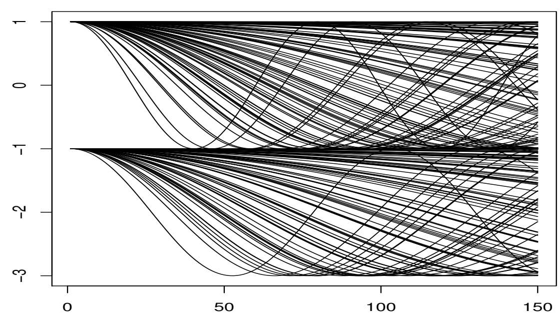

Figure 1: The 200 curves X.

For the scalar response variable, we took $Y = Ar_1(X) + (1-A)r_2(X)$ where r_1 (resp. r_2) is the nonlinear regression model $r_1(X) = 0.25 \times \left(\int_0^1 X'(t)\, dt\right)^2 + \epsilon$, with ϵ is $U([0, 0.5])$ (resp. $r_2(X)$ is the null function). We choose a quadratic kernel K defined by:

$$K(x) = \frac{3}{2}(1 - x^2)\, \mathbb{1}_{[0,1]} .$$

In practice, the semi-metric choice is based on the regularity of the curves X. For this we use the semi-metric defined by the L_2-distance between the q^{th} derivatives of the curves. In order to evaluate the MSE (Mean Square Error) we proceed by the following algorithm:

Step 1. We split our data into two subsets; the first sample, of size $n = 120$ corresponds to the learning sample which will be used, as a sample, to compute our conditional hazard function estimators for the 80 remaining curves (considered as the test sample).

- $(X_j, Y_j)_{j \in J}$ learning sample,
- $(X_i, Y_i)_{i \in I}$ test sample.

Step 2.
- We use the learning sample for computing the hazard function estimator $\widehat{h}_j$, for all $j \in J$.
- We set: $i^* = \arg\min_{j \in J} d(X_i, X_j)$.
- We put: $\forall\, i \in I$,

$$\widehat{T}_i = \widehat{h}^{X_{i^*}}(Y_i) \quad \text{for kernel method}\,,$$
$$\widehat{T}_i = \widehat{h}^{X_{k_{opt}}}(Y_i) \quad \text{for } k\text{-NN method}\,,$$

where

$$X_{i^*}: \text{ is the nearest curve to } X_j\,,$$
$$k_{opt}: \ \arg\min_a (CV(a))\,,$$

with

$$CV(a) \;=\; \frac{1}{n}\left[\sum_{i \in J} \int \left(\widehat{f}^{-i}_{(a,b)}(X_i, y)\right)^2 dy \;-\; 2\sum_{i \in J} \widehat{f}^{-i}_{(a,b)}(X_i, Y_i)\right]$$

and

$$\widehat{f}^{-k}_{(a,b)}(x, y) \;=\; \frac{b^{-1}\sum_{i \in J, i \neq k} K\left(\dfrac{d(x, X_i)}{a}\right) R\left(\dfrac{y - Y_i}{b}\right)}{\sum_{i \in J} K\left(\dfrac{d(x, X_i)}{a}\right)}\,.$$

Step 3. The error used to evaluate this comparison is the mean of square error (MSE) expressed by

$$\frac{1}{\mathrm{card}(I)} \sum_{i \in I} \left| h(Y_i) - \widehat{T}(X_i, Y_i)\right|^2\,,$$

where $\widehat{T}$ designate the estimator used: kernel or k-NN method estimation and h is the true hazard function.

Consequently, the k-NN method gives slightly better results than the kernel method. This is confirmed by the MSE-k-NN$\,=0.8227394$ and MSE-Kernel$\,=$ 1.347982.

4.3. Real data application

To highlight the efficiency and robustness of the method of k nearest neighbors with respect to the kernel method in estimating the conditional hazard function, we will test these two methods in the presence or not of heterogeneous data.

To do this, based on the study of Burba *et al.* (2009) which emphasizes the effect of the nature of the data (homogeneous or heterogeneous) on the quality of the estimate, especially the superiority of the k-nearest neighbors in the presence of very heterogeneous data.

For this purpose, we apply the described algorithm used in the simulation study to some chemiometrical real data available on the site[3], the original of these data (215 selected pieces of meat) comes from a quality control problem in the food industry that controls grease on a sample of finely chopped meat by chemical processes.

The sample of size 215 was split into learning sample of size 205 (with all data), 178 (without the heterogeneous data, 27 values) and testing sample of size 10. Figure 2 displays the curves of learning sample for all data and the curves of learning sample without the heterogeneous one.

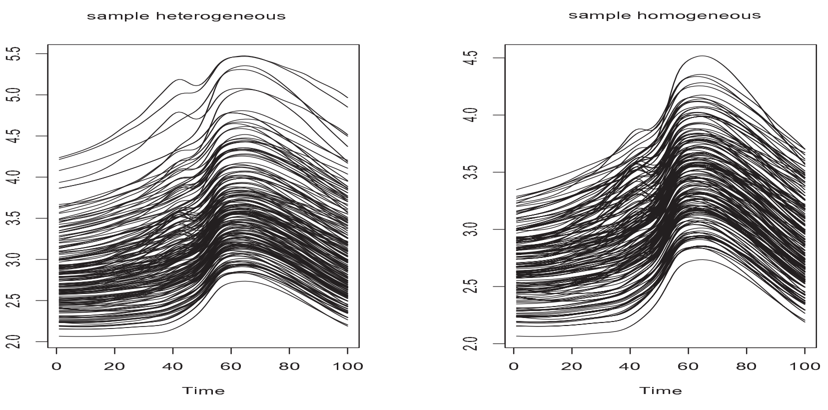

Figure 2: The learning curves.

For our study, we use the standard L^2 semi-metric and a quadratic kernel function K.

We plot the conditional hazard function estimated for the first 3 values of the testing sample, Figure 3 depicts that the k-NN method in presence of hetero-

[3] `http://lib.stat.cmu.edu/datasets/tecator`.

geneous data give a better estimation of the conditional hazard function prediction (regular function) than the kernel method estimation (non-regular function) and when the data are homogeneous the two method give the same result which can be easily seen in Figure 4.

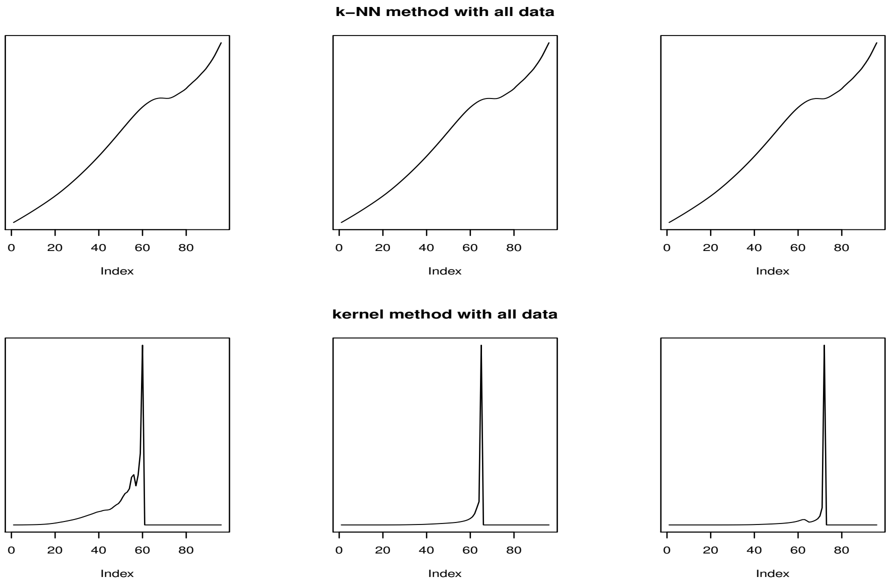

Figure 3: k-NN method (upper panels) *vs* kernel method (lower panels) of conditional hazard function far all data.

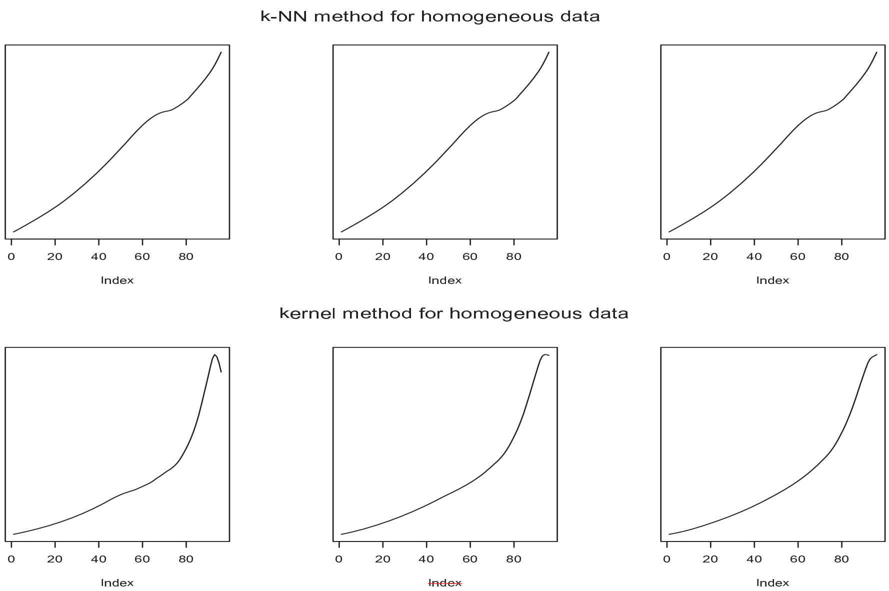

Figure 4: k-NN method (upper panels) *vs* kernel method (lower panels) of conditional hazard function for homogeneous data.

5. GENERAL TECHNICAL TOOLS

Let $(A_i, B_i)_{i\in\mathbb{N}}$ be a sequence of random variables with values in $(\Omega\times\mathbb{R}, \mathcal{A}\otimes\mathcal{B})$, independent but not necessarily identically distributed,where $(\Omega,\mathcal{A})$ is a general measurable space, let $G\colon \mathbb{R}\times\Omega \to \mathbb{R}^+$ a measurable function such that: $\forall\, w, w' \in \mathbb{R}$,

$$w \le w' \implies G(w,z) \le G(w',z)\ , \qquad \forall\, z\in\Omega\ .$$

Let c be a not random positive real number and T a real random variable: we define, $\forall\, n\in\mathbb{N}^*$,

$$C_n(T) \;=\; \frac{\displaystyle\sum_{i=1}^{n} B_i\, G(T, A_i)}{\displaystyle\sum_{i=1}^{n} G(T, A_i)}\ .$$

Lemma 5.1 (Burba *et al.* ([3]))**.** *Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of real random variables and $(u_n)_{n\in\mathbb{N}}$ be a decreasing sequence of positive numbers. If $l = \lim u_n \neq 0$, and if, for all increasing sequence $\beta_n \in\,]0,1[$, there exist two sequences of real random variables $(D_n^+(\beta_n))_{n\in\mathbb{N}}$ and $(D_n^-(\beta_n))_{n\in\mathbb{N}}$:*

(L1) $\forall\, n\in\mathbb{N},\quad D_n^- \le D_n^+$ *and* $\mathbb{1}_{D_n^-\le D_n\le D_n^+} \to 1$ *a.co.*

(L2) $$\frac{\displaystyle\sum_{i=1}^{n} G(D_n^-, A_i)}{\displaystyle\sum_{i=1}^{n} G(D_n^+, A_i)} - \beta_n = O(u_n) \qquad \textit{a.co.}$$

(L3) $C_n(D_n^-) - c = O(u_n)$ *a.co.*
$C_n(D_n^+) - c = O(u_n)$ *a.co.*

Then:

$$C_n(D_n) - c \;=\; O(u_n) \qquad \textit{a.co.}$$

If $l=0$ and if (L1), (L2), (L3) hold for any increasing sequence $\beta_n \in\,]0,1[$ with limit 1, the same result holds.

Lemma 5.2 (Burba *et al.* ([3]))**.** *Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of real random variables and $(v_n)_{n\in\mathbb{N}}$ be a decreasing positive sequence. If $l' = \lim v_n \neq 0$, and if, for all increasing sequence $\beta_n \in\,]0,1[$, there exist two sequences of real random variables $(D_n^+(\beta_n))_{n\in\mathbb{N}}$ and $(D_n^-(\beta_n))_{n\in\mathbb{N}}$:*

(L′1) $\forall\, n\in\mathbb{N},\quad D_n^- \le D_n^+$ *and* $\mathbb{1}_{D_n^-\le D_n\le D_n^+} \to 1$ *a.co.*

$$(\mathbf{L'2}) \quad \frac{\sum_{i=1}^{n} G(D_n^-, A_i)}{\sum_{i=1}^{n} G(D_n^+, A_i)} - \beta_n = o(v_n) \qquad a.co.$$

$$(\mathbf{L'3}) \quad C_n(D_n^-) - c = o(v_n) \qquad a.co.$$
$$C_n(D_n^+) - c = o(v_n) \qquad a.co.$$

Then:

$$C_n(D_n) - c = o(v_n) \qquad a.co.$$

If $l' = 0$ and if (L′1), (L′2), (L′3) hold for any increasing sequence $(\beta_n) \in]0,1[$ with limit 1, the same result holds.

Burba *et al.* ([3]) use in their consistency proof of the k-NN kernel estimate for independent data a Chernoff-type exponential inequality to check conditions (L1) or (L′1).

Lemma 5.3 (Burba *et al.* ([3])). *Let $(X_1, X_2, ..., X_n)$ be independent random variable in $\{0,1\}$. We note $X = \sum_{i=1}^{n} X_i$ and $\mu = \mathbb{E}(X)$: then, $\forall\, \delta > 0$,*

$$\mathbb{P}\big[X > (1+\delta)\,\mu\big] \;<\; \big[e^{\delta}/(1+\delta)^{1+\delta}\big]^{\mu},$$
$$\mathbb{P}\big[X < (1-\delta)\,\mu\big] \;<\; \big[e^{-\delta^2/2\mu}\big].$$

APPENDIX

Proof of Section 3.1

Proof of Lemma 3.1: On one hand, to prove the first result, we apply Lemma 5.2 with:

$$
\text{(A.1)} \qquad \begin{cases} v_n = 1 \ , \\ H_n = D_n \ , \\ \widehat{f}^x(y) = C_n(D_n) \ , \\ f^x(y) = c \ . \end{cases}
$$

Choose $\beta_n \in]0,1[$, (D_n^-) and (D_n^+) such that:

$$
\text{(A.2)} \qquad \begin{cases} \varphi_x(D_n^-) = \sqrt{\beta_n}\,\varphi_x(h) = \sqrt{\beta_n}\,\dfrac{k_n}{n} \ , \\ \varphi_x(D_n^+) = \dfrac{1}{\sqrt{\beta_n}}\,\varphi_x(h) = \dfrac{1}{\sqrt{\beta_n}}\,\dfrac{k_n}{n} \ . \end{cases}
$$

Define

$$
\text{(A.3)} \qquad \begin{cases} h^- = D_n^- = \varphi_x^{-1}\left(\sqrt{\beta_n}\,\dfrac{k_n}{n}\right) , \\ h^+ = D_n^+ = \varphi_x^{-1}\left(\dfrac{1}{\sqrt{\beta_n}}\,\dfrac{k_n}{n}\right) . \end{cases}
$$

Ferraty and Vieu ([13]) proved under the conditions of Theorem 3.1 that:

$$
\frac{1}{n\,\varphi_x(h)} \sum_{i=1}^n K\big(h^{-1} d(x,X_i)\big) \;\to\; 1 \qquad \text{a.co.}
$$

Under the conditions (A.2) and (A.3), we have:

$$
\begin{cases} \dfrac{1}{n\,\varphi_x(D_n^-)} \displaystyle\sum_{i=1}^n K\big((D_n^-)^{-1} d(x,X_i)\big) \to 1 \qquad \text{a.co.} \\ \dfrac{1}{n\,\varphi_x(D_n^+)} \displaystyle\sum_{i=1}^n K\big((D_n^+)^{-1} d(x,X_i)\big) \to 1 \qquad \text{a.co.} \end{cases}
$$

Then:

$$
\frac{\displaystyle\sum_{i=1}^n K\big((D_n^-)^{-1} d(x,X_i)\big)}{\displaystyle\sum_{i=1}^n K\big((D_n^+)^{-1} d(x,X_i)\big)} \;\to\; \beta_n \qquad \text{a.co.}\,,
$$

so that (L′2) is checked. Now by using Lemma (6.15) in Ferraty and Vieu ([13]) under the conditions of Theorem 3.1 and

$$D_n \longrightarrow 0\,, \quad \frac{\log n}{n\varphi_x(D_n)} \longrightarrow 0 \qquad (n\to\infty)\,, \tag{A.4}$$

we have:

$$C_n(D_n^-) \to c \qquad \text{a.co.}$$
$$C_n(D_n^+) \to c \qquad \text{a.co.}$$

so (L′3) is verified. Finally, we check (L′1). The first part is obvious, and the second one that: $\forall\,\epsilon>0$,

$$\sum_{n\ge 0} \mathbb{P}\left[\left|\mathbb{1}_{D_n^- < H_n < D_n^+} - 1\right| > \epsilon\right] < \infty\,.$$

We know that:

$$\mathbb{P}\Big[\left|\mathbb{1}_{D_n^- < H_n < D_n^+} - 1\right| > \epsilon\Big] \,\le\, \underbrace{\mathbb{P}\big[H_n < D_n^-\big]}_{A_1} + \underbrace{\mathbb{P}\big[H_n > D_n^+\big]}_{A_2}\,,$$

$$A_1 \,\le\, \mathbb{P}\left[\sum_{i=1}^n \mathbb{1}_{B(x,D_n^-)} > k_n\right].$$

And by using Lemma 5.3 with

$$\begin{cases} X_i = \mathbb{1}_{B(x,D_n^-)}\,, \\ X = \displaystyle\sum_{i=1}^n \mathbb{1}_{B(x,D_n^-)}\,, \\ \mathbb{P}\big(X_i{=}1\big) \,=\, \varphi_x(D_n^-)\,, \\ \mu \,=\, \mathbb{E}(X) = \displaystyle\sum_{i=1}^n \mathbb{E}\big[\mathbb{1}_{B(x,D_n^-)}\big] \,=\, n\,\varphi_x(D_n^-)\,, \end{cases} \tag{A.5}$$

we get:

$$\begin{aligned} \mathbb{P}\big[H_n < D_n^-\big] \,&<\, \left[e^{\left(\frac{1}{\sqrt{\beta}}-1\right)} \Big/ \left(\frac{1}{\sqrt{\beta}}\right)^{-\frac{1}{\sqrt{\beta}}}\right]^{n\varphi_x(D_n^-)} \\ &<\, n^{\left(-\log\sqrt{\beta_n}\,e^{(1-\sqrt{\beta_n})}\right)\frac{-k_n}{\log n}}\,. \end{aligned}$$

Under the hypotheses (A.4) and as $\sqrt{\beta}(e^{(1-\sqrt{\beta})}) < 1$ then:

$$A_1 \,=\, \mathbb{P}\big[H_n < D_n^-\big] \,<\, \infty\,.$$

Turning now to the study of A_2, we obtain

$$\mathbb{P}\big[H_n > D_n^+\big] \,=\, \mathbb{P}\left[\sum_{i=1}^n \mathbb{1}_{B(x,D_n^+)} < n\sqrt{\beta}\varphi_x(D_n^+)\right]$$

under the modification (A.5), and by applying the Lemma 5.3 we obtain:

$$\begin{aligned}\mathbb{P}\big[H_n > D_n^+\big] &< e^{\frac{-k_n(1-\sqrt{\beta})^2}{2\sqrt{\beta}}}\\ &< \left(n^{\frac{(1-\sqrt{\beta})^2}{2\sqrt{\beta}}}\right)^{\frac{-k_n}{\log n}} .\end{aligned}$$

Since $\dfrac{(1-\sqrt{\beta})^2}{2\sqrt{\beta}} > 0$ and $\dfrac{n\varphi_x(h)}{\log n} \to \infty$ then:

$$\mathbb{P}\big[H_n > D_n^+\big] < \infty .$$

Finally:

$$\mathbb{P}\left[\left|\mathbb{1}_{D_n^- < H_n < D_n^+} - 1\right| > \epsilon\right] < \infty .$$

On the other hand, we prove the second result. For this, we use the preceding steps with:

$$\text{(A.6)}\qquad \begin{cases} v_n = 1 \ , \\ H_n = D_n \ , \\ \widehat{F}^x(y) = C_n(D_n) \ , \\ F^x(y) = c \ . \end{cases}$$ □

Proof of Lemma 3.2: It is clear that:

$$\left|1 - \widehat{F}^x(y)\right| < \frac{1 - F^x(y)}{2} \implies \left|\widehat{F}^x(y) - F^x(y)\right| > \frac{1 - F^x(y)}{2} .$$

Turning now, to the term of probability, we obtain:

$$\mathbb{P}\left[\left|1 - \widehat{F}^x(y)\right| < \frac{1 - F^x(y)}{2}\right] \leq \mathbb{P}\left[\left|\widehat{F}^x(y) - F^x(y)\right| > \frac{1 - F^x(y)}{2}\right],$$

$$\sum_{n\in\mathbb{N}} \mathbb{P}\left[\left|1 - \widehat{F}^x(y)\right| < \frac{1 - F^x(y)}{2}\right] \leq \sum_{n\in\mathbb{N}} \mathbb{P}\left[\left|\widehat{F}^x(y) - F^x(y)\right| > \frac{1 - F^x(y)}{2}\right].$$

For the second term, by result 3.3, we have:

$$\sum_{n\in\mathbb{N}} \mathbb{P}\left[\left|\widehat{F}^x(y) - F^x(y)\right| > \frac{1 - F^x(y)}{2}\right] < \infty .$$

Then, for $\delta = \dfrac{1 - F^x(y)}{2}$, we obtain:

$$\sum_{n\in\mathbb{N}} \mathbb{P}\left[\left|\widehat{F}^x(y) - F^x(y)\right| > \frac{1 - F^x(y)}{2}\right] < \infty .$$ □

Proof of Lemma 3.3: To prove this lemma, we use Lemma 5.1. Choose β_n as an increasing sequence in $]0,1[$ with limit 1. Furthermore, we choose D_n^- and D_n^+ under (A.2), Ferraty and Vieu ([13]) proved under the conditions of Theorem 3.1 that:

$$\widehat{r}_3(x) - \mathbb{E}\left[\widehat{r}_3(x)\right] = O\left(\sqrt{\frac{\log n}{n h_n \varphi_x(h)}}\right),$$

with

$$\widehat{r}_3(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{K\left(h_n^{-1} d(x,X_i)\right) R\left(g_n^{-1}(y-Y_i)\right)}{\mathbb{E}\left[K\left(h_n^{-1} d(x,X_1)\right)\right]},$$

$$\widehat{r}_3(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{K\left(h_n^{-1} d(x,X_i)\right) \Gamma_i(y)}{\mathbb{E}\left[K\left(h_n^{-1} d(x,X_1)\right)\right]},$$

$$\Gamma_i(y) = R\left(g_n^{-1}(y-Y_i)\right).$$

Then,

$$\begin{aligned}
\widehat{r}_3(x) - \mathbb{E}\left[\widehat{r}_3(x)\right] &= \frac{1}{n}\sum_{i=1}^{n} \frac{K\left(h_n^{-1} d(x,X_i)\right)}{\mathbb{E}K\left(h_n^{-1} d(x,X_1)\right)}\,\Gamma_i(y) \\
&\quad - \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\frac{K\left(h_n^{-1} d(x,X_i)\right)}{\mathbb{E}K\left(h_n^{-1} d(x,X_1)\right)}\,\Gamma_i(y)\right] \\
&= \frac{1}{n\mathbb{E}K\left(h_n^{-1} d(x,X_1)\right)} \sum_{i=1}^{n} K\left(h_n^{-1} d(x,X_i)\right)\Gamma_i(y) \\
&\quad - \frac{1}{\mathbb{E}K\left(h_n^{-1} d(x,X_1)\right)}\,\mathbb{E}\left[K\left(h_n^{-1} d(x,X_1)\right)\mathbb{E}(\Gamma_1(y)/X_1)\right] \\
&= \frac{1}{n\mathbb{E}K\left(h_n^{-1} d(x,X_1)\right)} \sum_{i=1}^{n} K\left(h_n^{-1} d(x,X_i)\right)\Gamma_i(y) - \mathbb{E}(\Gamma_1(y)/X_1).
\end{aligned}$$

Using the fact that $\mathbb{E}\left[K\left(h_n^{-1} d(x,X_i)\right)\right] = O(\varphi_x(h)$ (see Ferraty and Vieu ([13]) and under the notations (A.2) and (A.3), we have:

$$\begin{cases}
\dfrac{1}{n\varphi_x(D_n^-)} \displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^-}\right)\Gamma_i(y) = \mathbb{E}(\Gamma_1(y)/X_1) + O\left(\sqrt{\dfrac{\log n}{g_n k_n}}\right), \\[2ex]
\dfrac{1}{n\varphi_x(D_n^+)} \displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^+}\right)\Gamma_i(y) = \mathbb{E}(\Gamma_1(y)/X_1) + O\left(\sqrt{\dfrac{\log n}{g_n k_n}}\right).
\end{cases}$$

By this, we obtain:

$$\frac{\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^-}\right)}{\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^+}\right)} - \beta_n = O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) \qquad \text{a.co.}$$

that (L2) is verified. Now, we apply Lemma (6.15) for Ferraty and Vieu ([13]) under (A.2) and (A.1), we get:

$$C_n(D_n^-) - c = O\left(\varphi_x^{-1}\left(\frac{k}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) \qquad \text{a.co.}$$

$$C_n(D_n^+) - c = O\left(\varphi_x^{-1}\left(\frac{k}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) \qquad \text{a.co.}$$

that verifies condition (L3). □

Proof of Lemma 3.4: To verify this Lemma, we pass by the same steps as before, such that: Ferraty and Vieu ([13]) showed that

$$\widehat{r}_1(x) - 1 = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right),$$

with

$$\widehat{r}_1(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{K\left(h_n^{-1} d(x, X_i)\right)}{\mathbb{E}K\left(h_n^{-1} d(x, X_1)\right)}.$$

Then

$$\frac{1}{n}\sum_{i=1}^{n} K\left(h_n^{-1} d(x, X_i)\right) - \varphi_x(h) = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right)$$

and under the same choice of $h^- = D_n^-$ and $h^+ = D_n^+$ as above, we have:

$$\begin{cases} \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} K\left(\dfrac{d(x, X_i)}{D_n^-}\right) = \sqrt{\beta_n}\dfrac{k_n}{n} + O\left(\sqrt{\dfrac{\log n}{k_n}}\right), \\ \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} K\left(\dfrac{d(x, X_i)}{D_n^+}\right) = \dfrac{1}{\sqrt{\beta_n}}\dfrac{k_n}{n} + O\left(\sqrt{\dfrac{\log n}{k_n}}\right). \end{cases}$$

We get

$$\frac{\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x, X_i)}{D_n^-}\right)}{\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x, X_i)}{D_n^+}\right)} - \beta_n = O\left(\sqrt{\frac{\log n}{k_n}}\right)$$

so that, (L2) is checked. Now we are able to apply Lemma (6.14) in Ferraty and Vieu ([13]) under (A.6), we obtain

$$C_n(D_n^-) - c = O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{k_n}}\right),$$

$$C_n(D_n^+) - c = O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{k_n}}\right),$$

and (L3) is verified. □

Proof of Section 3.2

Proof of Lemma 3.5: We denote:

$$\begin{cases} C_n(H_n) = \widehat{f^x}(y) \;, \\ c = f^x(y) \;. \end{cases} \tag{A.7}$$

Under (A.2) and (A.3), we have:

$$\begin{aligned} &\left(\frac{k_n g_n}{\sigma_f^2(x,y)}\right)^{1/2} \left[C_n(H_n) - c\right] = \\ &= \left(\frac{k_n g_n}{\sigma_f^2(x,y)}\right)^{1/2} \left[C_n(D_n^+) - c\right] + \left(\frac{k_n g_n}{\sigma_f^2(x,y)}\right)^{1/2} \left[C_n(H_n) - C_n(D_n^+)\right] . \end{aligned} \tag{A.8}$$

Then, to establish the asymptotic normality of the conditional density function, we need to show the asymptotic normality of the first term in equation (A.8) and the second term converges a.co. to 0.

For this, we remind that, under the same assumptions as Lemma 3.5, Quintela-del-Río ([23]) in Theorem 5 proved that

$$\left(\frac{k_n g_n}{\sigma_f^2(x,y)}\right)^{1/2} \left[C_n(D_n^+) - c\right] \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \qquad \text{as} \quad n \to \infty \;.$$

On the other hand, by hypothesis (H2) and the fact that $\mathbb{1}_{\{D_n^- \le H_n \le D_n^+\}} \to 1$ where $\dfrac{k_n}{n} \to 0$ (see Burba *et al.* ([3])), we have:

$$C_n(D_n^+) \le C_n(H_n) \le C_n(D_n^-) \;.$$

Using the fact that:

$$\begin{aligned} |C_n(H_n) - C_n(D_n^+)| &\le |C_n(D_n^-) - C_n(D_n^+)| \\ &\le |C_n(D_n^-) - \mathbb{E}\left[C_n(D_n^-)\right]| + |C_n(D_n^+) - \mathbb{E}[C_n(D_n^+)]| \\ &\quad + |\mathbb{E}[C_n(D_n^-)] - \mathbb{E}[C_n(D_n^+)]| \;. \end{aligned} \tag{A.9}$$

For the first term, we can write:

$$|C_n(D_n^-) - \mathbb{E}\left[C_n(D_n^-)\right]| \le |C_n(D_n^-) - c| + |\mathbb{E}\left[C_n(D_n^-)\right] - c|$$

by Lemma (3.3), we have:

$$|C_n(D_n^-) - c| = O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{k_n g_n}}\right)$$

and Quintela-del-Río ([23]) proved that:

$$|\mathbb{E}\left[C_n(D_n^-)\right] - c| \;=\; o(g_n^\beta) + O\left(\frac{1}{k_n}\right). \tag{A.10}$$

Finally, under hypothesis (H10), we obtain the almost complete convergence of the first term of (A.9). And to establish the almost complete convergence of the second term we apply the same steps as before.

Finally for the third term, we have:

$$|\mathbb{E}[C_n(D_n^-)] - \mathbb{E}[C_n(D_n^+)]| \;\le\; |\mathbb{E}[C_n(D_n^-)] - c| + |\mathbb{E}[C_n(D_n^+)] - c|$$

the almost complete convergence to 0 of these two terms is verified in (A.10). □

Proof of Lemma 3.6: To prove this Lemma, we apply the same steps as preceding with:

$$\begin{cases} C_n(H_n) = \widehat{F}^x(y)\ , \\ c = F^x(y)\ . \end{cases} \tag{A.11}$$

□

Proof of Lemma 3.7: It is clear that, the result (3.3) of Lemma (3.1) permits to conclude that:

$$\widehat{F}^x(y) \to F^x(y) \quad \text{in probability}\,.$$

□

ACKNOWLEDGMENTS

The authors are grateful to the Editor and Referee for helpful comments and suggestions that improved the paper.

The first author was supported by “L’Agence Thématique de Recherche en Sciences et Technologie ATRST (Ex ANDRU)” in P.N.R., No.46/15.

REFERENCES

[1] Attouch, M. and Benchikh, T. (2012). Asymptotic distribution of robust k-nearest neighbour estimator for functional nonparametric models, *Mathematic Vesnic*, **64**(4), 275–285.

[2] Burba, F.; Ferraty, F. and Vieu, P. (2009). k-Nearest Neighbour method in functional nonparametric regression, *Journal of Nonparametric Statistics* **21**(4), 453–469.

[3] Burba, F.; Ferraty, F. and Vieu, P. (2008). Convergence de l'estimateur à noyau des k plus proches voisins en régression fonctionnelle non-paramétrique, *C. R. Acad. Sci. Paris*, **346**(5–6), 339–342.

[4] Collomb, G. (1980). Estimation de la régression par la méthode des k plus proches voisins avec noyau: quelques propriétés de convergence ponctuelle, *Statistique Non Parametrique Asymptotique*, J.-P. Raoult, Ed., vol. 821 of *Lecture Notes in Mathematics*, Springer, Berlin/Heidelberg, pp. 159–175.

[5] Collomb, G. (1997). Quelques propriétés de la méthode du noyau pour l'estimation nonparamétrique de la régression en un point fixe, *C. R. Acad. Sci. Paris*, Série 1, **285**, 289–293.

[6] Devroye, L.P. (1978). The uniform convergence of nearest neighbour regression function estimators and their application in optimization, *IEEE Trans. Inform. Theory*, **24**, 142–151.

[7] Demongeot, J.; Laksaci, A.; Madani, F. and Rachdi, M. (2010). Local linear estimation of the conditional density for functional data, *C. R. Math. Acad. Sci. Paris*, **348**(15–16), 931–934.

[8] Demongeot, J.; Laksaci, A.; Madani, F. and Rachdi, M. (2011). A fast functional locally modeled conditional density and mode for functional time-series, *Recent Advances in Functional Data Analysis and Related Topics Contributions to Statistics*, Springer, pp. 85–90.

[9] Demongeot, J.; Laksaci, A.; Madani, F. and Rachdi, M. (in press). Functional data: local linear estimation of the density and its application, *Statistics*, DOI: 10.1080/02331888.2011.568117.

[10] Ferraty, F.; Laksaci A.; Tadj, A. and Vieu, P. (2011). Kernel regression with functional response, *Electronic Journal of Statistics*, **5**, 159–171.

[11] Ferraty, F.; Laksaci A. and Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models, *Statistical Inference for Stochastic Processes*, **9**, 47–76.

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

[13] Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis*, Springer, New York.

[14] Laksaci, A.; Madani, F. and Rachdi, M. (2012). Kernel conditional density estimation when the regressor is valued in a semi-metric space, *Communications Statistics – Theory and Methods*, accepted paper.

[15] Laksaci, A. and Mechab, B. (2012). Conditional hazard estimate for functional random fields, *Communications Statistics – Theory and Methods*, accepted paper.

[16] Laksaci, A. and Mechab, B. (2010). Estimation nonparamétrique de la fonction de hasard avec variable explicative fonctionnelle: cas des données spatiales, *Rev. Roumaine Math. Pures Appl.*, **55**, 35–51.

[17] Lian, H. (2011). Convergence on functional k-nearest neighbor regression estimate with functional responses, *Electronic Journal of Statistics*, **5**, 31–40.

[18] Li, J. and Tran, L.T. (2007). Hazard rate estimation on random fields, *J. Multivariate Anal.*, **98**, 1337–1355.

[19] Maillot, B. and Louani, D. (2008). *Propriétés asymptotiques de quelques estimateurs non-paramétriques pour des variables vectorielles et fonctionnelles*, Thèse de doctorat, Université Paris 6.

[20] Mack, Y.P. (1981). Local properties of k-NN regression estimates, *SIAM J. Algebraic Discrete Methods*, **2**, 311–323.

[21] Muller, S. and Dippon, J. (2011) *k-NN kernel estimate for nonparametric functional regression in time series analysis*, Fachbereich Mathematik, Fakultat Mathematik und Physik (Pfaffenwaldring 57), 014/2011, preprint, University of Stuttgart.

[22] Oliveira, P.E. (2005). *Nonparametric density and regression estimate functional data*, Tech. rep., Departamento de Matemática, Universidade de Coimbra.

[23] Quintela-del-Río, A. (2008). Hazard function given a functional variable: non-parametric estimation under strong mixing conditions, *J. Nonparametric Stat.*, **20**, 413–430.

[24] Ramsay, J. and Silverman, B. (2005). *Functional Data Analysis*, 2nd Ed., Springer Series in Statistics, Springer, New York.

[25] Roussas, G. (1969). Nonparametric estimation of the transition distribution function of a Markov process, *Annals of Mathematical Statistics*, **40**, 1386–1400.

[26] Roussas, G. (1989). Nonparametric estimation in mixing sequences of random variables, *J. Statist. Plann.*, **18**, 135–149.

[27] Royall, R.M. (1966). *A class of nonparametric estimates of a smooth regression function*, Ph.D. Diss., Stanford University.

[28] Samanta, M. (1989). Non-parametric estimation of conditional quantiles, *Statist. Proba. Letters*, **7**, 407–412.

[29] Samanta, M. and Thavaneswaran, A. (1990). Non-parametric estimation of conditional model, *Comm. Statist. Theory and Meth.*, **16**, 4515–4524.

[30] Stone, C.J. (1977). Consistent nonparametric regression, *Ann. Statist.*, **5**, 595–645.

[31] Waston, G.S. and Leadbetter, M.R. (1964). Hazard analysis, *I. Biometrika*, **51**, 175–184.