

# Nonparametric estimation of high risk of the conditional hazard function under independent conditions

Abderrahim Mahiddine<sup>a</sup>, Khadidja Nedjadi<sup>b</sup>, Abbes Rabhi<sup>c</sup>

<sup>a</sup>Mathematics Laboratory, University of Sidi Bel Abbes, P.O.Box. 89, Sidi Bel Abbes 22000, Algeria.

<sup>b</sup>Stochastic Models, Statistics and Applications Laboratory, Moulay Tahar University of Saida, P.O.Box 138 En-Nasr Saida 20 000 Algeria.

<sup>c</sup>Mathematics Laboratory, University of Sidi Bel Abbes, P.O.Box. 89, Sidi Bel Abbes 22000, Algeria.

**Abstract.** The high risk of the conditional hazard function is a parameter of great importance in seismicity studies, because it constitutes the maximum risk of occurrence of an earthquake in a given interval of time. It is shown that the (empirically determined) the high risk of the kernel estimate we establish the asymptotic behavior of a hazard rate in the presence of a random explanatory variable and asymptotic normality of independence data.

## 1. Introduction

The statistical analysis of functional data studies the experiments whose results are generally the curves. Under this supposition, the statistical analysis focuses on a framework of infinite dimension for the data under study. This field of modern statistics has received much attention in the last 20 years, and it has been popularized in the book of Ramsay and Silverman (2005). This type of data appears in many fields of applied statistics: environmetrics (Damon and Guillas, (2002)), chemometrics (Benhenni *et al.*, (2007)), meteorological sciences (Besse *et al.*, (2000)), etc.

The study of the hazard function has been subject to several investigations and many authors considered this function in their investigations. Among others, we refer to Watson and Leadbetter (1964) who were the first to study the nonparametric estimation of the hazard function and proposed a kernel estimate; in the sequel, many authors have been interested in the study of such a function. For instance, Collomb *et al.* (1987) studied the dependent case, Liu and Van Ryzin (1985) were interested in the histogram estimator of the hazard function for censored data, and Youndje *et al.* (1996) proposed a solution to the bandwidth selection problem for the kernel hazard estimate and gave properties of the selected bandwidth. In the sequel, Quintela (2007) used the plug-in bandwidth selection method in the case of a weak dependence on the sample data and a result of asymptotic optimality for the plug-in bandwidth is presented. Simulations are done as well to compare this method to the "leave more than one out" cross-validation criterion and either to show that smaller errors and much less sample variability can be reached. Besides, we notice that most of the precursor literature on nonparametric smoothing of hazard function was based on the assumption

---

2010 *Mathematics Subject Classification.* 62F12, 62G20, 62M09

*Keywords.* Almost complete convergence, Asymptotic normality, Conditional hazard function, Functional data, Small ball probability, Sequence of compact sets.

Received: 08 November 2014; Accepted: 24 December 2015

*Email addresses:* [ma2006ne\\_86@yahoo.fr](mailto:ma2006ne_86@yahoo.fr) (Abderrahim Mahiddine), [khadouj\\_2012@hotmail.com](mailto:khadouj_2012@hotmail.com) (Khadidja Nedjadi), [rabhi\\_abbes@yahoo.fr](mailto:rabhi_abbes@yahoo.fr) (Abbes Rabhi)

of independence on the sample variables, which is far from being realistic. This is for instance the case of micro earthquake studies (Rice and Rosenblatt (1976); Estévez *et al.*, 2002). Alternatively, many authors investigated the case of dependent hazard estimation, for instance Sarda and Vieu (1989), Vieu (1991), and Estvez and Quintela (1999), to name a few.

Recently, Quintela (2010) studied the recursive kernel hazard estimation of strong mixing data, by use of the density and the distribution, and established the strong consistency of the proposed estimator and a rate of convergence identical to the one obtained in the independence case.

### 1.1. Hazard and conditional hazard

The estimation of the hazard function is a problem of considerable interest, especially to inventory theorists, medical researchers, logistics planners, reliability engineers and seismologists. The non-parametric estimation of the hazard function has been extensively discussed in the literature. Beginning with Watson and Leadbetter (1964), there are many papers on these topics: Ahmad (1976), Singpurwalla and Wong (1983), etc. We can cite Quintela (2007) for a survey.

The literature on the estimation of the hazard function is very abundant, when observations are vectorial. Cite, for instance, Watson and Leadbetter (1964), Roussas (1989), Lecoutre and Ould-Saïd (1993), Estévez *et al.* (2002) and Quintela-del-Rio (2006) for recent references. In all these works the authors consider independent observations or dependent data from time series. The first results on the nonparametric estimation of this model, in functional statistics were obtained by Ferraty *et al.* (2008). They studied the almost complete convergence of a kernel estimator for hazard function of a real random variable dependent on a functional predictor and Laksaci and Mechab (2010) in the case of spatial variables. Asymptotic normality of the latter estimator was obtained, in the case of  $\alpha$ -mixing, by Quintela-del-Rio (2008). We refer to Ferraty *et al.* (2010) and Bouchentouf *et al.* (2014) for uniform almost complete convergence of the functional component of this nonparametric model. When hazard rate estimation is performed with multiple variables, the result is an estimate of the conditional hazard rate for the first variable, given the levels of the remaining variables. Many references, practical examples and simulations in the case of non-parametric estimation using local linear approximations can be found in Spierdijk (2008).

Our paper presents some asymptotic properties related with the nonparametric estimation of the maximum of the conditional hazard function. In a functional data setting, the conditioning variable is allowed to take its values in some abstract semi-metric space. In this case, Ferraty *et al.* (2008) define non-parametric estimators of the conditional hazard function. They give the rates of convergence (in an almost complete sense) to the corresponding functions, in a dependence ( $\alpha$ -mixing) context. We extend their results by calculating the maximum of the conditional hazard function of these estimates, and establishing their asymptotic normality, considering a particular type of kernel for the functional part of the estimates. Because the hazard function estimator is naturally constructed using these two last estimators, the same type of properties is easily derived for it. Our results are valid in a real (one- and multi-dimensional) context.

If  $X$  is a random variable associated to a lifetime (ie, a random variable with values in  $\mathbb{R}^+$ , the hazard rate of  $X$  (sometimes called hazard function, failure or survival rate) is defined at point  $x$  as the instantaneous probability that life ends at time  $x$ . Specifically, we have:

$$h(x) = \lim_{dx \rightarrow 0} \frac{\mathbb{P}(X \leq x + dx | X \geq x)}{dx}, \quad (x > 0).$$

When  $X$  has a density  $f$  with respect to the measure of Lebesgue, it is easy to see that the hazard rate can be written as follows:

$$h(x) = \frac{f(x)}{S(x)} = \frac{f(x)}{1 - F(x)}, \quad \text{for all } x \text{ such that } F(x) < 1,$$

where  $F$  denotes the distribution function of  $X$  and  $S = 1 - F$  the survival function of  $X$ .

In many practical situations, we may have an explanatory variable  $X = x$  and the main issue is to estimate the conditional random rate defined as

$$h^x(y) = \lim_{dy \rightarrow 0} \frac{\mathbb{P}(X \leq y + dy | Y > y, X = x)}{dy}, \text{ for } y > 0,$$

which can be written naturally as follows:

$$h^z x(y) = \frac{f^x(y)}{S^x(y)} = \frac{f^x(y)}{1 - F^x(y)}, \text{ once } F^x(y) < 1. \tag{1}$$

Study of functions  $h(\cdot)$  and  $h^z(\cdot)$  is of obvious interest in many fields of science ( biology, medicine, reliability , seismology, econometrics, ... ) and many authors are interested in construction of nonparametric estimators of  $h$ .

In this paper we propose an estimate of the maximum risk, through the nonparametric estimation of the conditional hazard function.

The layout of the paper is as follows. Section 2 describes the non-parametric functional setting: the structure of the functional data and the mixing conditions, the conditional density, distribution and hazard operators, and the corresponding non-parametric kernel estimators. Section 3 states the almost complete convergence<sup>1)</sup> (with rates of convergence<sup>2)</sup>) for nonparametric estimates of the derivative of the conditional hazard and the maximum risk. In Section 4, we calculate the variance of the conditional density, distribution and hazard estimates, the asymptotic normality of the three estimators considered is developed in this Section. Finally, Section 5 includes some proofs of technical Lemmas.

**2. Nonparametric estimation with functional data**

Let  $\{(X_i, Y_i), i = 1, \dots, n\}$  be a sample of  $n$  random pairs, each one distributed as  $(X, Y)$ , where the variable  $X$  is of functional nature and  $Y$  is scalar. Formally, we will consider that  $X$  is a random variable valued in some semi-metric functional space  $\mathcal{F}$ , and we will denote by  $d(\cdot, \cdot)$  the associated semi-metric. The conditional cumulative distribution of  $Y$  given  $X = x$  is defined for any  $y \in \mathbb{R}$  and any  $x \in \mathcal{F}$  by

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

while the conditional density, denoted by  $f^x(y)$  is defined as the density of this distribution with respect to the Lebesgue measure on  $\mathbb{R}$ . The conditional hazard is defined as in the non-infinite case (1).

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

*2.1. The functional kernel estimates*

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

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

where  $K$  and  $H$  are kernel functions and  $h_n$  is sequence of smoothing parameter. The conditional distribution operator  $F^x(\cdot)$  can be estimated by

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

<sup>1)</sup>Recall that a sequence  $(T_n)_{n \in \mathbb{N}}$  of random variables is said to converge almost completely to some variable  $T$ , if for any  $\epsilon > 0$ , we have  $\sum_n \mathbb{P}(|T_n - T| > \epsilon) < \infty$ . This mode of convergence implies both almost sure and in probability convergence (see for instance Bosq and Lecoutre, (1987)).

<sup>2)</sup>Recall that a sequence  $(T_n)_{n \in \mathbb{N}}$  of random variables is said to be of order of complete convergence  $u_n$ , if there exists some  $\epsilon > 0$  for which  $\sum_n \mathbb{P}(|T_n| > \epsilon u_n) < \infty$ . This is denoted by  $T_n = \mathcal{O}(u_n)$ , *a.co.* (or equivalently by  $T_n = \mathcal{O}_{a.co.}(u_n)$ ).

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

Consequently, the conditional hazard operator is defined in a natural way by

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

For  $z \in \mathcal{F}$ , we denote by  $h^x(\cdot)$  the conditional hazard function of  $Y_1$  given  $X_1 = x$ . We assume that  $h^x(\cdot)$  is unique maximum and its high risk point is denoted by  $\theta(x) := \theta$ , which is defined by

$$h^x(\theta(x)) := h^x(\theta) = \max_{y \in \mathcal{S}} h^x(y). \tag{2}$$

A kernel estimator of  $\theta$  is defined as the random variable  $\widehat{\theta}(x) := \widehat{\theta}$  which maximizes a kernel estimator  $\widehat{h}^x(\cdot)$ , that is,

$$\widehat{h}^x(\widehat{\theta}(x)) := \widehat{h}^x(\widehat{\theta}) = \max_{y \in \mathcal{S}} \widehat{h}^x(y) \tag{3}$$

where  $h^z$  and  $\widehat{h}^z$  are defined above.

Note that the estimate  $\widehat{\theta}$  is not necessarily unique and our results are valid for any choice satisfying (3). We point out that we can specify our choice by taking

$$\widehat{\theta}(x) = \inf \left\{ t \in \mathcal{S} \text{ such that } \widehat{h}^x(t) = \max_{y \in \mathcal{S}} \widehat{h}^x(y) \right\}.$$

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

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

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

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

$$(H1) \quad \forall h_n > 0, 0 < \mathbb{P}(X \in B(x, h_n)) = \phi_x(h_n) \text{ and } \lim_{h_n \rightarrow 0} \phi_x(h_n) = 0$$

Note that (H1) can be interpreted as a concentration hypothesis acting on the distribution of the *f.r.v.* of  $X = x$ .

Our nonparametric models will be quite general in the sense that we will just need the following simple assumption for the marginal distribution of  $X$ , and let us introduce the technical hypothesis necessary for the results to be presented. The non-parametric model is defined by assuming that

$$(H2) \quad \left\{ \begin{array}{l} \forall (y_1, y_2) \in \mathcal{S}^2, \forall (x_1, x_2) \in \mathcal{N}_x^2, \text{ for some } b_1 > 0, b_2 > 0 \\ |F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C_x(d(x_1, x_2)^{b_1} + |y_1 - y_2|^{b_2}), \end{array} \right.$$

$$(H3) \quad \left\{ \begin{array}{l} \forall (y_1, y_2) \in \mathcal{S}^2, \forall (x_1, x_2) \in \mathcal{N}_x^2, \text{ for some } j = 0, 1, \nu > 0, \beta > 0 \\ |f^{x_1(j)}(y_1) - f^{x_2(j)}(y_2)| \leq C_x(d(x_1, x_2)^\nu + |y_1 - y_2|^\beta), \end{array} \right.$$

$$(H4) \quad \exists \gamma < \infty, f^{f^z}(x) \leq \gamma, \quad \forall (x, y) \in \mathcal{F} \times \mathcal{S},$$

$$(H5) \quad \exists \tau > 0, F^x(y) \leq 1 - \tau, \quad \forall (x, y) \in \mathcal{F} \times \mathcal{S}.$$

(H6)  $H$  is differentiable such that

$$\left\{ \begin{array}{l} \text{(H6a)} \quad \forall (t_1, t_2) \in \mathbb{R}^2; |H^{(j)}(t_1) - H^{(j)}(t_2)| \leq C|t_1 - t_2|, \text{ for } j = 0, 1 \\ \text{and } H^{(j)} \text{ are bounded for } j = 0, 1 \\ \text{(H6b)} \quad \int_{\mathbb{R}} t^2 H^2(t) dt < \infty, \\ \text{(H6c)} \quad \int_{\mathbb{R}} |t|^\beta (H'(t))^2 dt < \infty \end{array} \right.$$

(H7) The kernel  $K$  is positive bounded function supported on  $[0, 1]$  and it is of class  $C^1$  on  $(0, 1)$  such that  $\exists C_1, C_2, -\infty < C_1 < K'(t) < C_2 < +\infty$  for  $0 < t < 1$

(H8) There exists a function  $\zeta_0^x(\cdot)$  such that for all  $t \in [0, 1]$   $\lim_{h_n \rightarrow 0} \frac{\phi_x(th_n)}{\phi_x(h_n)} = \zeta_0^x(t)$ .

(H9) The bandwidth  $h_n$ , small ball probability  $\phi_x(h_n)$  satisfying

$$\left\{ \begin{array}{l} \text{(H9a)} \quad \lim_{n \rightarrow \infty} h_n = 0 \\ \text{(H9b)} \quad \lim_{n \rightarrow \infty} \frac{\log n}{nh_n \phi_x(h_n)} = 0 \\ \text{(H9c)} \quad \lim_{n \rightarrow \infty} \frac{\log n}{nh_n^{2j+1} \phi_x(h_n)} = 0, \quad j = 0, 1; \end{array} \right.$$

**Remark 2.1.** Assumption (H1) plays an important role in our methodology. It is known as (for small  $h_n$ ) 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  $\phi_x(h_n)$  can be written approximately as the product of two independent functions  $\psi(z)$  and  $\varphi(h_n)$  as  $\phi_x(h_n) = \psi(x)\varphi(h_n) + o(\varphi(h_n))$ . This idea was adopted by Masry (2005) who reformulated the Gasser et al. (1998) one. The increasing propriety of  $\phi_x(\cdot)$  implies that  $\zeta_{h_n}^x(\cdot)$  is bounded and then integrable (all the more so  $\zeta_0^x(\cdot)$  is integrable).

Without the differentiability of  $\phi_x(\cdot)$ , this assumption has been used by many authors where  $\psi(\cdot)$  is interpreted as a probability density, while  $\varphi(\cdot)$  may be interpreted as a volume parameter. In the case of finite-dimensional spaces, that is  $\mathcal{S} = \mathbb{R}^d$ , it can be seen that  $\phi_x(h_n) = C(d)h_n^d \psi(x) + o(h_n^d)$ , 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. (2007)):

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

The function  $\zeta_{h_n}^x(\cdot)$  which intervenes in Assumption (H9) is increasing for all fixed  $h_n$ . Its pointwise limit  $\zeta_0^x(\cdot)$  also plays a determinant role. It intervenes in all asymptotic properties, in particular in the asymptotic variance term. With simple algebra, it is possible to specify this function (with  $\zeta_0(u) := \zeta_0^x(u)$  in the above examples by:

1.  $\zeta_0(u) = u^\gamma$ ,
2.  $\zeta_0(u) = \delta_1(u)$  where  $\delta_1(\cdot)$  is Dirac function,
3.  $\zeta_0(u) = \mathbf{1}_{]0,1]}(u)$ .

**Remark 2.2.** Assumptions (H2) and (H3) are the only conditions involving the conditional probability and the conditional probability density of  $Y$  given  $X = x$ . It means that  $F(\cdot|\cdot)$  and  $f(\cdot|\cdot)$  and its derivatives satisfy the Hölder condition with respect to each variable. Therefore, the concentration condition (H1) plays an important role. Here we point out that our assumptions are very usual in the estimation problem for functional regressors (see, e.g., Ferraty et al. (2008)).

**Remark 2.3.** Assumptions (H6), (H7) and (H9) are classical in functional estimation for finite or infinite dimension spaces.

### 3. Nonparametric estimate of the maximum of the conditional hazard function

Let us assume that there exists a compact  $\mathcal{S}$  with a unique maximum  $\theta$  of  $h^x$  on  $\mathcal{S}$ . We will suppose that  $h^x$  is sufficiently smooth (at least of class  $\mathcal{C}^2$ ) and verifies that  $h'^x(\theta) = 0$  and  $h''^x(\theta) < 0$ .

We can write an estimator of the first derivative of the conditional hazard function through the first derivative of the estimator (1). Our maximum estimate is defined by assuming that there is some unique  $\hat{\theta}$  on  $\mathcal{S}$  such that  $0 = \hat{h}'(\hat{\theta}) < |\hat{h}'^x(y)|$  for all  $x \in \mathcal{S}$  and  $x \neq \hat{\theta}$

Furthermore, we assume that  $\theta \in \mathcal{S}^\circ$ , where  $\mathcal{S}^\circ$  denotes the interior of  $\mathcal{S}$ , and that  $\theta$  satisfies the uniqueness condition, that is; for any  $\varepsilon > 0$  and  $\mu(z)$ , there exists  $\xi > 0$  such that  $|\theta(x) - \mu(x)| \geq \varepsilon$  implies that  $|h^x(\theta(x)) - h^x(\mu(x))| \geq \xi$ .

We can write an estimator of the first derivative of the hazard function through the first derivative of the estimator. Our maximum estimate is defined by assuming that there is some unique  $\hat{\theta}$  on  $\mathcal{S}^\circ$ .

It is therefore natural to try to construct an estimator of the derivative of the function  $h^x$  on the basis of these ideas. To estimate the conditional distribution function and the conditional density function in the presence of functional conditional random variable  $X = x$ .

The kernel estimator of the derivative of the function conditional random functional  $h^z$  can therefore be constructed as follows:

$$\hat{h}'^x(y) = \frac{\hat{f}'^x(y)}{1 - \hat{F}^x(y)} + (\hat{h}^x(y))^2, \tag{4}$$

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

$$\hat{f}'^x(y) = \frac{\sum_{i=1}^n K(h_n^{-1}d(x, X_i))H'(h_n^{-1}(y - Y_i))}{h_n^2 \sum_{i=1}^n K(h_n^{-1}d(x, X_i))}. \tag{5}$$

Later, we need assumptions on the parameters of the estimator, ie on  $K, H, H'$  and  $h_n$  are little restrictive. Indeed, on one hand, they are not specific to the problem estimate of  $h^x$  (but inherent problems of  $F^x, f^x$  and  $f'^x$  estimation), and secondly they consist with the assumptions usually made under functional variables.

**Remark 3.1.** *Generally, the hazard function has a global maximum in the time intervals with values closest to zero, corresponding to the earthquakes with bigger intensity (Vere-Jones, 1970).*

*Also, the hazard function can have several local maxima, indicating the times with the highest risk in a certain period (see the examples in Estévez et al., (2002)).*

*The hypothesis of uniqueness is only established for the sake of clarity. Following our proofs, if several local estimated maxima exist, the asymptotic results remain valid for each of them.*

We state the almost complete convergence (with the rates of convergence) of the maximum estimate by the following results:

**Theorem 3.2.** *Under assumption (H1)-(H7) we have*

$$\hat{\theta} - \theta \rightarrow 0 \quad a.co. \tag{6}$$

**Remark 3.3.** *The hypothesis of uniqueness is only established for the sake of clarity. Following our proofs, if several local estimated maxima exist, the asymptotic results remain valid for each of them.*

*Proof.* Because  $h'^x(\cdot)$  is continuous, we have, for all  $\epsilon > 0$ .  $\exists \eta(\epsilon) > 0$  such that

$$|t - \theta| > \epsilon \Rightarrow |h'^x(t) - h'^x(\theta)| > \eta(\epsilon).$$

Therefore,

$$\mathbb{P}\{|\hat{\theta} - \theta| \geq \epsilon\} \leq \mathbb{P}\{|h'^x(\hat{\theta}) - h'^x(\theta)| \geq \eta(\epsilon)\}.$$

We also have

$$|h^{'x}(\hat{\theta}) - h^{'x}(\theta)| \leq |h^{'x}(\hat{\theta}) - \hat{h}^{'x}(\hat{\theta})| + |\hat{h}^{'x}(\hat{\theta}) - h^{'x}(\theta)| \leq \sup_{y \in \mathcal{S}} |\hat{h}^{'x}(y) - h^{'x}(y)|, \tag{7}$$

because  $\hat{h}^{'x}(\hat{\theta}) = h^{'x}(\theta) = 0$ .

Then, uniform convergence of  $h^{'x}$  will imply the uniform convergence of  $\hat{\theta}$ . That is why, we have the following lemma.

**Lemma 3.4.** *Under assumptions of Theorem 3.2, we have*

$$\sup_{y \in \mathcal{S}} |\hat{h}^{'x}(y) - h^{'x}(y)| \rightarrow 0 \quad a.co. \tag{8}$$

□

The proof of this lemma will be given in Appendix.

**Theorem 3.5.** *Under assumption (H1)-(H7), (H9a) and (H9c) we have*

$$\sup_{y \in \mathcal{S}} |\hat{\theta} - \theta| = \mathcal{O}(h_n^{b_1}) + \mathcal{O}_{a.co.} \left( \sqrt{\frac{\log n}{nh_n^3 \phi_x(h_n)}} \right) \tag{9}$$

*Proof.* By using Taylor expansion of the function  $h^{'x}$  at the point  $\hat{\theta}$ , we obtain

$$h^{'x}(\hat{\theta}) = h^{'x}(\theta) + (\hat{\theta} - \theta)h^{''x}(\theta_n^*), \tag{10}$$

with  $\theta^*$  a point between  $\theta$  and  $\hat{\theta}$ .

Now, because  $h^{'x}(\theta) = \hat{h}^{'x}(\hat{\theta})$

$$|\hat{\theta} - \theta| \leq \frac{1}{h^{''x}(\theta_n^*)} \sup_{y \in \mathcal{S}} |\hat{h}^{'x}(y) - h^{'x}(y)| \tag{11}$$

The proof of Theorem will be completed showing the following lemma.

**Lemma 3.6.** *Under the assumptions of Theorem 3.5, we have*

$$\sup_{y \in \mathcal{S}} |\hat{h}^{'x}(y) - h^{'x}(y)| = \mathcal{O}(h_n^{b_1}) + \mathcal{O}_{a.co.} \left( \sqrt{\frac{\log n}{nh_n^3 \phi_x(h_n)}} \right) \tag{12}$$

The proof of lemma will be given in the Appendix.

□

#### 4. Asymptotic normality

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

$$(H6d) \int_{\mathbb{R}} (H'(t))^2 dt < \infty,$$

$$(H10) 0 = \hat{h}^{'x}(\hat{\theta}) < |\hat{h}^{'x}(y)|, \forall y \in \mathcal{S}, y \neq \hat{\theta}$$

The following result gives the asymptotic normality of the maximum of the conditional hazard function. Let

$$\mathcal{A} = \{(x, y) : (x, y) \in \mathcal{F} \times \mathbb{R}, a_2^y F^x(y) (1 - F^x(y)) \neq 0\}$$

**Theorem 4.1.** Under conditions (H1)-(H10) we have  $(\theta \in \mathcal{S}/f^x(\theta), 1 - F^x(\theta) > 0)$

$$(nh_n^3 \phi_x(h_n))^{1/2} (\widehat{h}'^x(\theta) - h'^x(\theta)) \xrightarrow{\mathcal{D}} N(0, \sigma_{h'}^2(\theta))$$

where  $\rightarrow^{\mathcal{D}}$  denotes the convergence in distribution,

$$a_l^y = K^l(1) - \int_0^1 (K^l(u))' \zeta_0^y(u) du \quad \text{for } l = 1, 2$$

and

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

*Proof.* Using again (17), and the fact that

$$\frac{(1 - F^x(y))}{(1 - \widehat{F}^x(y))(1 - F^x(y))} \rightarrow \frac{1}{1 - F^x(y)}$$

and

$$\frac{\widehat{f}'^x(y)}{(1 - \widehat{F}^x(y))(1 - F^x(y))} \rightarrow \frac{f'^x(y)}{(1 - F^x(y))^2}.$$

The asymptotic normality of  $(nh_n^3 \phi_x(h_n))^{1/2} (\widehat{h}'^x(\theta) - h'^x(\theta))$  can be deduced from both following lemmas,

**Lemma 4.2.** Under Assumptions (H1)-(H2) and (H6)-(H8), we have

$$(n\phi_x(h_n))^{1/2} (\widehat{F}^x(y) - F^x(y)) \xrightarrow{\mathcal{D}} N(0, \sigma_{F^x}^2(y)) \tag{13}$$

where

$$\sigma_{F^x}^2(y) = \frac{a_2^y F^x(y)(1 - F^x(y))}{(a_1^y)^2}$$

**Lemma 4.3.** Under Assumptions (H1)-(H3) and (H5)-(H9), we have

$$(nh_n \phi_x(h_n))^{1/2} (\widehat{h}^x(y) - h^x(y)) \xrightarrow{\mathcal{D}} N(0, \sigma_{h^x}^2(y)) \tag{14}$$

where

$$\sigma_{h^x}^2(y) = \frac{a_2^y h^x(y)}{(a_1^y)^2 (1 - F^x(y))} \int_{\mathbb{R}} H^2(t) dt$$

**Lemma 4.4.** Under Assumptions of Theorem 4.1, we have

$$(nh_n^3 \phi_x(h_n))^{1/2} (\widehat{f}'^x(y) - f'^x(y)) \xrightarrow{\mathcal{D}} N(0, \sigma_{f'^x}^2(y)) \tag{15}$$

where

$$\sigma_{f'^x}^2(y) = \frac{a_2^y f^x(y)}{(a_1^y)^2} \int_{\mathbb{R}} (H'(t))^2 dt$$



**Lemma 4.5.** Under the hypotheses of Theorem 4.1, we have

$$\begin{aligned} \text{Var} \left[ \widehat{f}'_N(y) \right] &= \frac{\sigma_{f^{t_x}(y)}^2}{nh_n^3 \phi_x(h_n)} + o \left( \frac{1}{nh_n^3 \phi_x(h_n)} \right), \\ \text{Var} \left[ \widehat{F}_N^x(y) \right] &= o \left( \frac{1}{nh_n \phi_x(h_n)} \right); \end{aligned}$$

and

$$\text{Var} \left[ \widehat{F}_D^x \right] = o \left( \frac{1}{nh_n \phi_x(h_n)} \right).$$

**Lemma 4.6.** Under the hypotheses of Theorem 4.1, we have

$$\begin{aligned} \text{Cov}(\widehat{f}'_N(y), \widehat{F}_D^x) &= o \left( \frac{1}{nh_n^3 \phi_x(h_n)} \right), \\ \text{Cov}(\widehat{f}'_N(y), \widehat{F}_N^x(y)) &= o \left( \frac{1}{nh_n^3 \phi_x(h_n)} \right) \end{aligned}$$

and

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

**Remark 4.7.**

It is clear that, the results of lemmas, Lemma 4.5 and Lemma 4.6 allows to write

$$\text{Var} \left[ \widehat{F}_D^x - \widehat{F}_N^x(y) \right] = o \left( \frac{1}{nh_n \phi_x(h_n)} \right)$$

The proofs of lemmas, Lemma 4.2 can be seen in Belkhir *et al.* (2015), Lemma 4.5 and Lemma 4.6 see Rabhi *et al.* (2015). □

Finally, by this last result and (10), the following theorem follows:

**Theorem 4.8.** Under conditions (H1)-(H11) we have  $(\theta \in \mathcal{S}/f^x(\theta), 1 - F^x(\theta) > 0)$

$$(nh_n^3 \phi_x(h_n))^{1/2} (\widehat{\theta} - \theta) \xrightarrow{\mathcal{D}} N \left( 0, \frac{\sigma_{h'}^2(\theta)}{(h''^x(\theta))^2} \right)$$

with  $\sigma_{h'}^2(\theta) = h^x(\theta) (1 - F^x(\theta)) \int (H'(t))^2 dt$ .

### 5. Proofs of technical lemmas

*Proof.* [Proof of lemma 3.4 and lemma 3.6] Let

$$\widehat{h}^{t_x}(y) = \frac{\widehat{f}^{t_x}(y)}{1 - \widehat{F}^x(y)} + (\widehat{h}^x(y))^2, \tag{16}$$

with

$$\widehat{h}^{t_x}(y) - h^{t_x}(y) = \left\{ (\widehat{h}^x(y))^2 - (h^x(y))^2 \right\} + \left\{ \frac{\widehat{f}^{t_x}(y)}{1 - \widehat{F}^x(y)} - \frac{f^{t_x}(y)}{1 - F^x(y)} \right\} \tag{17}$$

for the first term of (17) we can write

$$\left| \left( \widehat{h}^x(y) \right)^2 - \left( h^x(y) \right)^2 \right| \leq \left| \widehat{h}^x(y) - h^x(y) \right| \cdot \left| \widehat{h}^x(y) + h^x(y) \right| \tag{18}$$

because the estimator  $\widehat{h}^x(\cdot)$  converge a.co. to  $h^z(\cdot)$  we have

$$\sup_{y \in \mathcal{S}} \left| \left( \widehat{h}^x(y) \right)^2 - \left( h^x(y) \right)^2 \right| \leq 2 \left| h^x(\theta) \right| \sup_{y \in \mathcal{S}} \left| \widehat{h}^x(y) - h^x(y) \right| \tag{19}$$

for the second term of (17) we have

$$\begin{aligned} \frac{\widehat{f}'^x(y)}{1 - \widehat{F}^x(y)} - \frac{f'^x(y)}{1 - F^x(y)} &= \frac{1}{(1 - \widehat{F}^x(y))(1 - F^x(y))} \left\{ \widehat{f}'^x(y) - f'^x(y) \right\} \\ &+ \frac{1}{(1 - \widehat{F}^x(y))(1 - F^x(y))} \left\{ f'^x(y) \left( \widehat{F}^x(y) - F^x(y) \right) \right\} \\ &- \frac{1}{(1 - \widehat{F}^x(y))(1 - F^x(y))} \left\{ F^x(y) \left( \widehat{f}'^x(y) - f'^x(y) \right) \right\}. \end{aligned}$$

Valid for all  $y \in \mathcal{S}$ . Which for a constant  $C < \infty$ , this leads

$$\begin{aligned} \sup_{y \in \mathcal{S}} \left| \frac{\widehat{f}'^x(y)}{1 - \widehat{F}^x(y)} - \frac{f'^x(y)}{1 - F^x(y)} \right| &\leq \\ C \frac{\left\{ \sup_{y \in \mathcal{S}} \left| \widehat{f}'^x(y) - f'^x(y) \right| + \sup_{y \in \mathcal{S}} \left| \widehat{F}^x(y) - F^x(y) \right| \right\}}{\inf_{y \in \mathcal{S}} \left| 1 - \widehat{F}^x(y) \right|}. \end{aligned} \tag{20}$$

Therefore, the conclusion of the lemma follows from the following results:

$$\sup_{y \in \mathcal{S}} \left| \widehat{F}^x(y) - F^x(y) \right| = \mathcal{O} \left( h_n^{b_1} \right) + \mathcal{O}_{a.co.} \left( \sqrt{\frac{\log n}{n \phi_x(h_n)}} \right) \tag{21}$$

$$\sup_{y \in \mathcal{S}} \left| \widehat{f}'^x(y) - f'^x(y) \right| = \mathcal{O} \left( h_n^{b_1} \right) + \mathcal{O}_{a.co.} \left( \sqrt{\frac{\log n}{n h_n^3 \phi_x(h_n)}} \right) \tag{22}$$

$$\sup_{y \in \mathcal{S}} \left| \widehat{h}^x(y) - h^x(y) \right| = \mathcal{O} \left( h_n^{b_1} \right) + \mathcal{O}_{a.co.} \left( \sqrt{\frac{\log n}{n h_n \phi_x(h_n)}} \right) \tag{23}$$

$$\exists \delta > 0 \text{ such that } \sum_1^\infty \mathbb{P} \left\{ \inf_{y \in \mathcal{S}} \left| 1 - \widehat{F}^x(y) \right| < \delta \right\} < \infty. \tag{24}$$

The proofs of (21) and (22) appear in Ferraty *et al.* (2006), and (23) is proven in Ferraty *et al.* (2008).

- Concerning (24) by equation (21), we have the almost complete convergence of  $\widehat{F}^x(y)$  to  $F^x(y)$ . Moreover,

$$\forall \varepsilon > 0 \quad \sum_{n=1}^\infty \mathbb{P} \left\{ \left| \widehat{F}^x(y) - F^x(y) \right| > \varepsilon \right\} < \infty.$$

On the other hand, by hypothesis we have  $F^x < 1$ , i.e.

$$1 - \widehat{F}^x \geq F^x - \widehat{F}^x,$$

thus,

$$\inf_{y \in \mathcal{S}} |1 - \widehat{F}^x(y)| \leq (1 - \sup_{y \in \mathcal{S}} F^x(y))/2 \Rightarrow \sup_{y \in \mathcal{S}} |\widehat{F}^x(y) - F^x(y)| \geq (1 - \sup_{y \in \mathcal{S}} F^x(y))/2.$$

In terms of probability is obtained

$$\begin{aligned} & \mathbb{P} \left\{ \inf_{y \in \mathcal{S}} |1 - \widehat{F}^x(y)| < (1 - \sup_{y \in \mathcal{S}} F^x(y))/2 \right\} \\ & \leq \\ & \mathbb{P} \left\{ \sup_{y \in \mathcal{S}} |\widehat{F}^x(y) - F^x(y)| \geq (1 - \sup_{y \in \mathcal{S}} F^x(y))/2 \right\} < \infty. \end{aligned}$$

Finally, it suffices to take  $\delta = (1 - \sup_{y \in \mathcal{S}} F^x(y))/2$  and apply the results (21) to finish the proof of the lemma. □

*Proof.* [Proof of lemma 4.3] We can write for all  $x \in \mathcal{S}$

$$\begin{aligned} \widehat{h}^x(y) - h^x(y) &= \frac{\widehat{f}^x(y)}{1 - \widehat{F}^x(y)} - \frac{f^x(y)}{1 - F^x(y)} \\ &= \frac{1}{\widehat{D}^x(y)} \left\{ \left( \widehat{f}^x(y) - f^x(y) \right) + f^x(y) \left( \widehat{F}^x(y) - F^x(y) \right) \right. \\ & \quad \left. - F^x(y) \left( \widehat{f}^x(y) - f^x(y) \right) \right\}, \\ &= \frac{1}{\widehat{D}^x(y)} \left\{ (1 - F^x(y)) \left( \widehat{f}^x(y) - f^x(y) \right) \right. \\ & \quad \left. - f^x(y) \left( \widehat{F}^x(y) - F^x(y) \right) \right\} \end{aligned} \tag{25}$$

with  $\widehat{D}^x(y) = (1 - F^x(y)) (1 - \widehat{F}^x(y))$ .

As a direct consequence of the Lemma 4.2, the result (26) (see Ezzahrioui and Ould-Saïd (2008)) and the expression (25), permit us to obtain the asymptotic normality for the conditional hazard estimator.

$$(nh_n \phi_x(h_n))^{1/2} \left( \widehat{f}^x(y) - f^x(y) \right) \xrightarrow{\mathcal{D}} N(0, \sigma_{f^x}^2(y)) \tag{26}$$

where

$$\sigma_{f^x}^2(y) = \frac{a_2^y f^x(y)}{(a_1^y)^2} \int_{\mathbb{R}} (H(t))^2 dt$$

□

*Proof.* [Proof of lemma 4.4] For  $i = 1, \dots, n$ , we consider the quantities  $K_i = K(h_n^{-1}d(x, X_i))$ ,  $H'_i(y) = H'(h_n^{-1}(y - Y_i))$  and let  $\widehat{f}'^x_N(y)$  (resp.  $\widehat{F}'^x_D$ ) be defined as

$$\widehat{f}'^x_N(y) = \frac{h_n^{-2}}{n \mathbb{E}K_1} \sum_{i=1}^n K_i H'_i(y) \quad (\text{resp. } \widehat{F}'^x_D = \frac{1}{n \mathbb{E}K_1} \sum_{i=1}^n K_i).$$

This proof is based on the following decomposition

$$\begin{aligned} \widehat{f}'^x(y) - f'^x(y) &= \frac{1}{\widehat{F}_D^x} \left\{ \left( \widehat{f}'_N^x(y) - \mathbb{E}\widehat{f}'_N^x(y) \right) - \left( f'^x(y) - \mathbb{E}\widehat{f}'_N^x(y) \right) \right\} \\ &\quad + \frac{f'^x(y)}{\widehat{F}_D^x} \left\{ \mathbb{E}\widehat{F}_D^x - \widehat{F}_D^x \right\} \end{aligned} \tag{27}$$

and on the following intermediate results.

$$\sqrt{nh_n^3\phi_x(h_n)} \left( \widehat{f}'_N^x(y) - \mathbb{E}\widehat{f}'_N^x(y) \right) \xrightarrow{\mathcal{D}} N(0, \sigma_{f'^x}^2(y)) \tag{28}$$

where  $\sigma_{f'^x}^2(x)$  is defined as in Lemma 4.4.

$$\lim_{n \rightarrow \infty} \sqrt{nh_n^3\phi_x(h_n)} \left( \mathbb{E}\widehat{f}'_N^x(y) - f'^x(y) \right) = 0 \tag{29}$$

$$\sqrt{nh_n^3\phi_x(h_n)} \left( \widehat{F}_D^x(y) - 1 \right) \xrightarrow{\mathbb{P}} 0, \text{ as } n \rightarrow \infty. \tag{30}$$

- Concerning (28).

By the definition of  $\widehat{f}'_N^x(y)$ , it follows that

$$\begin{aligned} \Omega_n &= \sqrt{nh_n^3\phi_x(h_n)} \left( \widehat{f}'_N^x(y) - \mathbb{E}\widehat{f}'_N^x(y) \right) \\ &= \sum_{i=1}^n \frac{\sqrt{\phi_x(h_n)}}{\sqrt{nh_n}\mathbb{E}K_1} (K_i H'_i - \mathbb{E}K_i H'_i) \\ &= \sum_{i=1}^n \Delta_i, \end{aligned}$$

which leads

$$Var(\Omega_n) = nh_n^3\phi_x(h_n)Var \left( \widehat{f}'_N^x(y) - \mathbb{E} \left[ \widehat{f}'_N^x(y) \right] \right). \tag{31}$$

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

$$\begin{aligned} Var(\widehat{f}'_N^x(y)) &= \frac{1}{(nh_n^2\mathbb{E}[K_1(x)])^2} \sum_{i=1}^n \sum_{j=1}^n Cov(\Delta_i(x, y), \Delta_j(x, y)) \\ &= \frac{1}{n(h_n^2\mathbb{E}[K_1(x)])^2} Var(\Delta_1(x, y)). \end{aligned}$$

Therefore

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

Now, by a change of variable in the following integral and by applying (H4) and (H7), one gets

$$\begin{aligned}
 \mathbb{E} \left( H'^2_1(y) | X_1 = x \right) &= \int_{\mathbb{R}} H'^2 \left( \frac{y-u}{h_n} \right) f^x(u) du \\
 &\leq h_n \int_{\mathbb{R}} H'^2(t) (f^x(y - h_n t, x) - f^x(y)) dt \\
 &\quad + h_n f^x(y) \int_{\mathbb{R}} H'^2(t) dt \\
 &\leq h_n^{1+b_2} \int_{\mathbb{R}} |t|^{b_2} H'^2(t) dt + h_n f^x(y) \int_{\mathbb{R}} H'^2(t) dt \\
 &= h_n \left( o(1) + f^x(y) \int_{\mathbb{R}} H'^2(t) dt \right). \tag{32}
 \end{aligned}$$

By means of (32) and the fact that, as  $n \rightarrow \infty$ ,  $\mathbb{E} (K^2_1(x)) \rightarrow a^y_2 \phi_x(h_n)$ , one gets

$$\text{Var} (\Delta_1(x, y)) = a^y_2 \phi_x(h_n) h_n \left( o(1) + f^x(y) \int_{\mathbb{R}} H'^2(t) dt \right).$$

So, using (H8), we get

$$\begin{aligned}
 \frac{1}{n (h^2_n \mathbb{E}[K_1(x)])^2} \text{Var} (\Delta_1(x, y)) &= \frac{a^y_2 \phi_x(h_n)}{n (a^y_1 h^2_n \phi_x(h_n))^2} h_n \left( o(1) + f^x(y) \int_{\mathbb{R}} H'^2(t) dt \right) \\
 &= o \left( \frac{1}{nh^3_n \phi_x(h_n)} \right) + \frac{a^y_2 f^x(y)}{(a^y_1)^2 nh^3_n \phi_x(h_n)} \int_{\mathbb{R}} H'^2(t) dt.
 \end{aligned}$$

Thus as  $n \rightarrow \infty$  we obtain

$$\frac{1}{n (h^2_n \mathbb{E}[K_1(x)])^2} \text{Var} (\Delta_1(x, y)) \rightarrow \frac{a^y_2 f^x(y)}{(a^y_1)^2 nh^3_n \phi_x(h_n)} \int_{\mathbb{R}} H'^2(t) dt. \tag{33}$$

Indeed

$$\sum_{i=1}^n \mathbb{E} \Delta_i^2 = \frac{\phi_x(h_n)}{h_n \mathbb{E}^2 K_1} \mathbb{E} K_1^2 (H'_1)^2 - \frac{\phi_x(h_n)}{h_n \mathbb{E}^2 K_1} (\mathbb{E} K_1 H'_1)^2 = \Pi_{1n} - \Pi_{2n}. \tag{34}$$

As for  $\Pi_{1n}$ , by the property of conditional expectation, we get

$$\Pi_{1n} = \frac{\phi_x(h_n)}{\mathbb{E}^2 K_1} \mathbb{E} \left\{ K_1^2 \int H'^2(t) (f'^x(y - th_n) - f'^x(y) + f'^x(y)) dt \right\}.$$

Meanwhile, by (H1), (H3), (H7) and (H8), it follows that:

$$\frac{\phi_x(h_n) \mathbb{E} K_1^2}{\mathbb{E}^2 K_1} \xrightarrow{n \rightarrow \infty} \frac{a^y_2}{(a^y_1)^2},$$

which leads

$$\Pi_{1n} \xrightarrow{n \rightarrow \infty} \frac{a^y_2 f^x(y)}{(a^y_1)^2} \int (H'(t))^2 dt, \tag{35}$$

Regarding  $\Pi_{2n}$ , by (H1), (H3) and (H6), we obtain

$$\Pi_{2n} \xrightarrow[n \rightarrow \infty]{} 0. \tag{36}$$

This result, combined with (34) and (35), allows us to get

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \mathbb{E} \Delta_i^2 = \sigma_{f^x}^2(y) \tag{37}$$

Therefore, combining (33) and (36)-(37), (28) is valid.

- Concerning (29).

The proof is completed along the same steps as that of  $\Pi_{1n}$ . We omit it here.

- Concerning (30). The idea is similar to that given by Belkhir *et al.* (2015).

By definition of  $\widehat{F}_D^x$ , we have

$$\sqrt{nh_n^3 \phi_x(h_n)} (\widehat{F}_D^x - 1) = \Omega_n - \mathbb{E} \Omega_n,$$

where  $\Omega_n = \frac{\sqrt{nh_n^3 \phi_x(h_n)} \sum_{i=1}^n K_i}{n \mathbb{E} K_1}$ . In order to prove (30), similar to Belkhir *et al.* (2015), we only need to prove  $\text{Var } \Omega_n \rightarrow 0$ , as  $n \rightarrow \infty$ . In fact, since

$$\begin{aligned} \text{Var } \Omega_n &= \frac{nh_n^3 \phi_x(h_n)}{n \mathbb{E}^2 K_1} (n \text{Var} K_1) \\ &\leq \frac{nh_n^3 \phi_x(h_n)}{\mathbb{E}^2 K_1} \mathbb{E} K_1^2 \\ &= \Psi_1, \end{aligned}$$

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

$$\Psi_1 \leq Ch_n^3 \phi_x(h_n) \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

It is clear that, the results of (21), (22), (24) and Lemma 4.5 permits us

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

and

$$\text{Var} \left( \widehat{F}_D^x - \widehat{F}_N^x(y) - 1 + F^x(y) \right) \rightarrow 0;$$

then

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

Moreover, the asymptotic variance of  $\widehat{F}_D^x - \widehat{F}_N^x(y)$  given in Remark 4.7 allows to obtain

$$\frac{nh_n \phi_x(h_n)}{\sigma_{F^x}^2(y)} \text{Var} \left( \widehat{F}_D^x - \widehat{F}_N^x(y) - 1 + \mathbb{E} \left( \widehat{F}_N^x(y) \right) \right) \rightarrow 0.$$

By combining result with the fact that

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

we obtain the claimed result.

Therefore, the proof of this Lemma is completed. □

## References

- [1] Ahmad, I. A. (1976). Uniform strong convergence of the generalized failure rate estimate, *Bull. Math. Statist.* **17**, 77–84.
- [2] Belkhir, N., Rabhi, A., and Soltani, S. (2015). Exact asymptotic errors of the hazard conditional rate kernel, *Journal of Statistics Applications & Probability Letters. An International Journal*, **2**(3), 191–204. <http://dx.doi.org/10.12785/jsapl/020303>.
- [3] Benhenni, K., Ferraty, F., Rachdi, M. and Vieu, P. (2007). Local smoothing regression with functional data, *Comput. Statist.* **22**, 353–369.
- [4] Besse, P., Cardot, H. and Stephenson, D. (2000). Autoregressive forecasting of some functional climatic variations, *Scand. J. Statist.* **27**, 673–687.
- [5] Bosq, D. and Lecoutre, J. P. (1987). Théorie de l'estimation fonctionnelle, *ECONOMICA* (eds), Paris.
- [6] Bouchentouf, A. A., Djebbouri, T., Rabhi, A., Sabri, K. (2014). Strong uniform consistency rates of some characteristics of the conditional distribution estimator in the functional single-index model, *Appl. Math. (Warsaw)*, **41**(4), 301–322.
- [7] Collomb, G., Härdle, W. and Hassani, S. (1987). A note on prediction via conditional mode estimation. *J. Statist. Plann. and Inf.*, **15**, 227–236.
- [8] Damon, J. and Guillas, S. (2002). The inclusion of exogenous variables in functional autoregressive ozone forecasting, *Environmetrics*, **13**, 759–774.
- [9] Estévez, G., Quintela, A. (1999). Nonparametric estimation of the hazard function under dependence conditions. *Commun. Statist. Theor. Meth.*, **28**(10), 22972331.
- [10] Estévez-Pérez, G., Quintela-del-Rio, A. and Vieu, P. (2002). Convergence rate for cross-validated bandwidth in kernel hazard estimation from dependent samples, *J. Statist. Plann. Inference.* **104**, 1–30.
- [11] Ezzahrioui, M. and Ould-Saïd, E. (2008). Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data, *Journal of Nonparametric Statistics*, **20**(1), 3–18.
- [12] Ferraty, F., Laksaci, A. and Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models, *Stat. Inf. Stochastic Process.*, **9**, 47–76.
- [13] Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables, *J. Statist. Plann. and Inf.* **140**, 335–352.
- [14] Ferraty, F., Mas, A. and Vieu, P. (2007). Advances in nonparametric regression for functional variables, *Australian and New Zealand Journal of Statistics*, **49**, 1–20.
- [15] Ferraty, F., Rabhi, A. and Vieu, P. (2008). Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle, *Rev. Roumaine Math. Pures Appl.* **53**, 1–18.
- [16] Ferraty, F. and Vieu, P. (2006). Nonparametric functional data analysis, *Springer Series in Statistics, Theory and practice*, Springer, New York.
- [17] Gasser, T., Hall, P. and Presnell, B. (1998). Nonparametric estimation of the mode of a distribution of random curves, *Journal of the Royal Statistical Society, Ser. B.*, **60**, 681–691.
- [18] Laksaci, A. and Mechab, B. (2010). Estimation non paramétrique de la fonction de hasard avec variable explicative fonctionnelle cas des données spatiales, *Rev. Rom. J. Pure & Applied Math.*, **55**, 35–51.
- [19] Lecoutre, J. P. and Ould-Saïd, E. (1993). Estimation de la densité et de la fonction de hasard conditionnelle pour un processus fortement mélangé avec censure, *C. R. Math. Acad. Sci. Paris.* **314**, 295–300.
- [20] Liu, R., Van Ryzin, J. (1985). A histogram estimator of the hazard rate with censored data, *Ann. Statist.*, **13**, 592605.
- [21] Masry, E. (2005). Non-parametric regression estimation for dependent functional data: Asymptotic normality, *Stoch. Process. Appl.* **115**, 155–177.
- [22] Quintela-del-Rio, A. (2006). Nonparametric estimation of the maximum hazard under dependence conditions, *Statist. Probab. Lett.* **76**, 1117–1124.
- [23] Quintela-del-Rio, A. (2007). Plug-in bandwidth selection in kernel hazard estimation from dependent data, *Comput. Stat. Data Anal.* **51**, 5800–5812.
- [24] Quintela-del-Rio, A. (2008). Hazard function given a functional variable: Non-parametric estimation under strong mixing conditions, *J. Nonparametr. Stat.* **20**, 413–430.
- [25] Quintela-del-Rio, A. (2010). Recursive kernel hazard estimation of strong mixing data. *Far East J. Theor. Statist.*, **30**(2), 89105.
- [26] Rabhi, A., Soltani, S. and Traore, A. (2015). Conditional risk estimate for functional data under strong mixing conditions, *Journal of Statistical Theory and Applications (JSTA)*, **14**(3), 301–323.
- [27] Ramsay, J. O. and Silverman, B. W. (2005). Functional Data Analysis, 2nd ed., *Springer-Verlag*, New-York.
- [28] Rice, J., Rosenblatt, M. (1976). Estimation of the log survival function and hazard function, *Sankhya A*, **36**, 6078.
- [29] Roussas, G. (1989). Hazard rate estimation under dependence conditions, *J. Statist. Plann. Inference.* **22**, 81–93.
- [30] Sarda, P., Vieu, P. (1989). Empirical distribution function for mixing random variables, Application in nonparametric hazard estimation. *Statistics*, **20**(4), 559571.
- [31] Singpurwalla, N. D. and Wong, M. Y. (1983). Estimation of the failure rate - A survey of non-parametric methods. Part I: Non-Bayesian methods, *Commun. Stat. Theory Math.* **12**, 559–588.
- [32] Spierdijk, L. (2008). Non-parametric conditional hazard rate estimation: A local linear approach, *Comput. Stat. Data Anal.* **52**, 2419–2434.
- [33] Vere Vere-Jones, D. (1970). Stochastic models for earthquake occurrence, *J. Roy. Statist. Soc. Ser. B.*, **32**, 1–62.
- [34] Vieu, P. (1991). Quadratic errors for nonparametric estimates under dependence. *J. Multivariate Anal.*, **39**, 324347.
- [35] Watson, G. S. and Leadbetter, M. R. (1964). Hazard analysis I, *Biometrika*, **51**, 175–184.
- [36] Watson, G. S. and Leadbetter, M. R. (1964). Hazard analysis, *Sankhyā*, **26**, 101–116.

- [37] Youndjé, É., Sarda, P., Vieu, P. (1996). Optimal smooth hazard estimates, *Test*, **5**, 379394.