



On the strong convergence of the hazard rate and its maximum risk point estimators in presence of censorship and functional explanatory covariate

Komi Agbokou^{1,2,*} and Kossi Essona Gneyou²

^{1,2}Department of Mathematics, Faculty of sciences, University of Lomé, BP 1515, Togo

Received June 18, 2017; Accepted : August 1, 2017

Copyright © 2017, Afrika Statistika and Statistics and The Probability African Society (SPAS). All rights reserved

Abstract. . In the literature much work has been devoted to the non-parametric estimation of survival analysis functions. In this work, we focus on the nonparametric estimation of the conditional hazard rate and the point of its maximum, in the model of right censored data with presence of functional covariate. We establish the almost uniform complete convergence of these estimators at appropriate rates. This generalizes the almost sure convergence obtained in the literature.

Résumé. Dans la littérature beaucoup de travaux ont été consacrés à l'estimation non-paramétrique des fonctions d'analyse de survie. Dans ce travail, nous nous focalisons sur l'estimation non-paramétrique du taux de hasard conditionnel et du point de son maximum, dans le modèle des données censurées à droite et en présence de covariable fonctionnelle. Nous établissons la convergence uniforme presque complète de ces estimateurs à des vitesses appropriées. Ce qui généralise la convergence presque sûre obtenue dans la littérature.

Key words: Nonparametric estimation; conditional hazard rate maximum value; censored data; small ball probabilities; functional random variable.

AMS 2010 Mathematics Subject Classification : 62N01; 62N02; 62P10; 62P12

* Komi Agbokou : ffomestein@gmail.com
Kossi E. Gneyou : kgneyou@univ-lome.tg

1. Introduction

The hazard rate is one of the tools in survival analysis. If T is a random variable associated with a lifetime (that is, a random variable with values in \mathbb{R}^+), then the hazard rate $\lambda(t)$ (sometimes also called hazard function, failure rate or survival rate) is defined at the point t as the instantaneous probability that this lifetime ends at the time t . More specifically, it is defined as follows : for all $t \geq 0$,

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{P}(t \leq T < t + \Delta t/T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}, \quad S(t) > 0 \quad (1.1)$$

where $f(t)$ is the probability density function, $S(t) = 1 - F(t)$ the survival function of and $F(t)$ is the cumulative distribution function of the random variable T .

From a mathematical point of view, the hazard rate may be sufficient to characterize the distribution of a positive random variable. In applications, it is more often used to analyze duration data, for example, for medical follow-up, industrial reliability, unemployment treatment in socio-economic problems or earthquakes study. In the latter case, $\lambda(t)$ is used to measure the instantaneous risk of a replica or to predict the maximum risk in the event of an earthquake. In most situations, the hazard rate depends on one or more covariates. This is the case, for example, when the event of interest is the survival time of a patient, which is influenced by the age and/or gender. In many practical situations, we can have an explanatory variable X and the question becomes how to estimate the conditional hazard rate $\lambda(t/x)$ of T given $X = x$ defined for $t \geq 0$ by:

$$\lambda(t/x) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{P}(t \leq T < t + \Delta t/T \geq t, X = x)}{\Delta t}, \quad (1.2)$$

which is also naturally written from the conditional density function $f(. / x)$ and the conditional distribution function $F(. / x)$ or the conditional survival function $S(. / x) = 1 - F(. / x)$ of T given $X = x$, in the form:

$$\lambda(t/x) = \frac{f(t/x)}{S(t/x)}, \quad (1.3)$$

for all t such that $S(t/x) > 0$.

The study of the hazard rate function $\lambda(.)$ or the conditional hazard rate $\lambda(. / x)$ is obviously of interest in many fields of science such as biology, medicine, reliability, seismology, econometrics, etc., and many of authors were interested in the nonparametric estimation of the hazard rate λ . Historically, the nonparametric estimation of the hazard rate was introduced for the first time in the statistical literature by [Wason and Leadbetter \(1964a\)](#). One of the most common techniques for constructing estimators of $\lambda(.)$ and $\lambda(. / x)$ is based on the form (1.1) and similarly on the form (1.3). It consists in studying a quotient between an estimator of $f(.)$ (respectively $f(. / x)$) and an estimator of $S(.)$ (respectively $S(. / x)$). The paper by [Patil and al. \(1994\)](#) gives a general overview of these estimation techniques. Nonparametric methods based on convolution by kernels ideas, which are known for their good behavior in problems of probability density estimation (conditional or not), are also widely used in the nonparametric estimation of the hazard function. Mention may in particular be made of the recent articles by [Gneyou \(1997\)](#), [Gefeller and Michels \(1992\)](#),

Nassari *and al.* (2000), Estévez (2002), Estévez *and al.* (2002), Pascu and Vadura (2003), Quintela (2006), Laksaci and Mechab (2010), Dupuy and Gneyou (2011), as examples and references therein. A wide range of literature in this field is provided by Tanner and Wong (1983), Singpurwalla and Wong (1983), Hassani *and al.* (1986) or by Ferraty *and al.* (2008). The works of Ramsay and Silverman (2005) and Ferraty and Vieu (2006) propose another wide range of recently developed statistical methods, both parametric and non-parametric, to deal with various estimation problems involving functional random variables (that is, with values in a space of infinite dimension).

In the presence of an explanatory variable X with values in a functional space, one of the problem also concerns the estimation of the maximum of the hazard rate. Given $X = x$, it is assumed that there exists an interval $[a_x; b_x] \subset \mathbb{R}^+$ and a unique $\theta \in [a_x; b_x]$ such that

$$\lambda(\theta/x) = \sup_{t \in [a_x, b_x]} \lambda(t/x). \tag{1.4}$$

Gneyou (2013) and Gneyou (2014) established almost sure representations and the asymptotic normality of a maximum risk estimator in the right-censored data model. This estimator is obtained via the conditional cumulative hazard rate by convolution with a kernel. Rabhi *and al.* (2015) established the almost complete uniform convergence and the asymptotic normality of the maximum conditional risk estimator under the condition of dependence data.

In this paper, we propose a direct estimator of the conditional hazard rate obtained by taking the quotient of two nonparametric estimators : an estimator of the conditional cumulative distribution function and an estimator of the conditional probability density function given $X = x$, based on a right-censored data model. We establish the almost complete convergence results with appropriate rates. We mention that part of our results was treated in Ferraty *and al.* (2008). The difference lies that, in the current paper, we focus on the high risk point.

This paper is organized as follows. In Section 2 we give the definitions and determine the nonparametric estimators of the conditional hazard rate function and its maximum point. In Section 3, we set out the assumptions under which we derive the results, which we state. The detailed proofs are relegated in Section 4. In Section 5, we present numerical studies and simulations.

2. Nonparametric estimate of the conditional hazard rate

Let us consider a conditional model in which the explanatory variable X is not necessarily real or multidimensional but assumed to be of values in an abstract space \mathcal{F} equipped with a semi-metric d . As in any nonparametric estimation problem, the dimension of the space \mathcal{F} plays an important role in the concentration properties of the random variable T . Thus, when this dimension is not necessarily finite, one defines some functions so-called small balls probabilities. A function of small balls probabilities $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is defined as follows :

$$\begin{aligned} \phi_x(h) &= \mathbb{P}[X \in B(x, h)] = \mathbb{P}[X \in \{x' \in \mathcal{F} : d(x, x') < h\}] \\ &= \mathbb{E} [1_{B(x, h)}(X)] = \mathbb{E} \left[1_{[0,1]} \left(\frac{d(x, X)}{h} \right) \right], \end{aligned} \tag{2.1}$$

where $B(x, h)$ is the ball of center x and radius h with $x \in \mathcal{F}$ fixed and h is a positive real number.

Note that the probability of the ball $B(x, h)$ appears clearly in the normalization. The function ϕ_x is called the "small balls probabilities" because the smoothing parameter h (also called the bandwidth) decreases with the size of the sample of the functional variable, more precisely h tends to 0 if n tends to $+\infty$. Thus, when we take n very large, h is close to 0 (zero) and then the ball $B(x, h)$ is considered as a small ball and $\mathbb{P}[X \in B(x, h)]$ as a small balls probability. This notion of small balls probabilities plays a major role both from the theoretical and practical points of view. Because the notion of ball is strongly related to the semi-metric d , the choice of this semi-metric will become an important issue. Thus the function ϕ_x intervenes directly in the asymptotic behavior of any functional nonparametric estimator.

Let $T \in \mathbb{R}^+$ be the lifetime of interest and $X \in \mathcal{F}$ a random covariate. We assume that T and X are absolutely continuous random variables with conditional distribution function $F(t/x)$ and conditional probability density $f(t/x)$ of T given $X = x$. Let C be a right censoring variable with conditional distribution function $G(t/x)$ and probability density $g(t/x)$ such that C and T are conditionally independent given X . Define $Y = \min(T, C)$ and $\delta = 1_{\{T \leq C\}}$ where $1_{\{A\}}$ denotes the indicator function: $\delta = 1$ if $T \leq C$ and $\delta = 0$ if $T \geq C$, in which case the duration T is said to be censored to the right by C .

Let $H(t/x) = \mathbb{P}[Y \leq t/X = x]$ be the conditional distribution function of Y given $X = x$,

$$H_1(t/x) = P[Y \leq t, \delta = 1/X = x] = \int_0^t (1 - G(u/x))dF(u/x), \quad (2.2)$$

the conditional sub-distribution function of the uncensored observation $(Y, \delta = 1)$ given $X = x$ and let $f^*(t/x) = f(t/x)(1 - G(t/x))$ be its corresponding conditional sub-density function. By the independence condition, we have

$$1 - H(t/x) = (1 - F(t/x))(1 - G(t/x)). \quad (2.3)$$

Hence, from the relation (1.3), we have the following form of $\lambda(t/x)$ which takes account of the censorship mechanism :

$$\lambda(t/x) = \frac{f(t/x)(1 - G(t/x))}{(1 - F(t/x))(1 - G(t/x))} = \frac{f^*(t/x)}{1 - H(t/x)}, \quad H(t/x) \neq 1. \quad (2.4)$$

Let K be a kernel on \mathbb{R} with support in $[0, 1]$. Define the function K_h by $K_h(x) = \frac{1}{h}K\left(\frac{x}{h}\right)$, $x \in [0, 1]$, $h \in \mathbb{R}_+^*$. Let $(Y_1, \delta_1), \dots, (Y_n, \delta_n)$ be a sample of observed durations of size n and $h = h_{K,n}$ a sequence of positive real numbers (smoothing parameter) decreasing to 0 (zero) as n tends to $+\infty$. Then a nonparametric Nadaraya-Watson type estimator of the conditional distribution function $H(t/x)$ is given by:

$$H_n(t/x) = \sum_{i=1}^n W_i(x, h_{K,n})1_{\{Y_i \leq t\}}, \quad (2.5)$$

where for $i = 1, \dots, n$,

$$W_i(x, h_{K,n}) = \frac{K_{h_{K,n}}(d(x, X_i))}{\sum_{i=1}^n K_{h_{K,n}}(d(x, X_i))} = \frac{K\left(\frac{d(x, X_i)}{h_{K,n}}\right)}{\sum_{i=1}^n K\left(\frac{d(x, X_i)}{h_{K,n}}\right)}. \quad (2.6)$$

The quantities W_i are the so-called Nadaraya-Watson weights (Nadaraya (1964), Watson and Leadbetter (1964b)).

It is easy to construct a smooth version of this naive estimator (see also in Roussas (1969) or in Samanta (1989)). To do so, it suffices to change the basic indicator function $(1_{\{Y_i \leq t\}})$ in the conditional distribution function $H(.|x)$ given by (2.5) into a smooth conditional distribution function L called "Integrated Kernel" or "cumulative Kernel". To fix the ideas, let us consider $K_0 : \mathbb{R} \rightarrow \mathbb{R}_+$ a symmetric kernel with compact support in $[0, 1]$ and define

$$L(t) = \int_{-\infty}^t K_0(u) du.$$

In this case, L is a conditional distribution function and the quantity $L\left(\frac{t - Y_i}{h_{L,n}}\right)$, (where $h_{L,n}$ is similarly defined as $h_{K,n}$ with K replaced by L), acts as a local weighting : when Y_i is less than t the quantity $L\left(\frac{t - Y_i}{h_{L,n}}\right)$ is large and the more Y_i is above t , the smaller the quantity $L\left(\frac{t - Y_i}{h_{L,n}}\right)$. Moreover, we can write:

$$L\left(\frac{t - Y_i}{h_{L,n}}\right) = \begin{cases} 0 & \text{si } t \leq Y_i - h_{L,n}, \\ 1 & \text{si } t \geq Y_i + h_{L,n}. \end{cases}$$

Thus, we define the kernel estimators of conditional distribution and sub-distribution functions $H(t/x)$ and $f^*(t/x)$, respectively as follows:

$$\hat{H}_n(t/x) = \sum_{i=1}^n W_i(x, h_{K,n}) L\left(\frac{t - Y_i}{h_{L,n}}\right),$$

and

$$\hat{f}_n^*(t/x) = \frac{1}{h_{L,n}} \sum_{i=1}^n W_i(x, h_{K,n}) \delta_i K_0\left(\frac{t - Y_i}{h_{L,n}}\right)$$

The final estimator of the conditional hazard rate $\lambda(t/x)$ is then given, for all $t \geq 0$ and $F(t/x) \neq 1$, by

$$\hat{\lambda}_n(t/x) = \frac{\hat{f}_n^*(t/x)}{1 - \hat{H}_n(t/x)} = \frac{\frac{1}{h_{L,n}} \sum_{i=1}^n W_i(x, h_{K,n}) \delta_i K_0\left(\frac{t - Y_i}{h_{L,n}}\right)}{1 - \sum_{j=1}^n W_j(x, h_{K,n}) L\left(\frac{t - Y_j}{h_{L,n}}\right)}. \quad (2.7)$$

A natural estimator $\hat{\theta}_n$ of θ is finally given by

$$\hat{\theta}_n(x) = \text{Argmax}_{a_x \leq t \leq b_x} \hat{\lambda}_n(t/x). \quad (2.8)$$

3. Asymptotic properties of estimators

Given x in the functional space \mathcal{F} , we denote by V_x a neighborhood of x in \mathcal{F} and for any conditional distribution function $L(\cdot/x)$, set $\tau_L = \sup \{t \in \mathbb{R}_+ / L(t/x) < 1\}$.

Choose $0 < \tau \leq \tau_H = \min(\tau_F, \tau_G)$, naturally τ_F, τ_G and therefore τ depend on the covariate x . For the rest of the paper, we set $S = [0, \tau]$. We need the following assumptions.

3.1. Assumptions

General Assumptions

$$(A_1) \quad \forall h > 0, \quad \mathbb{P}[X \in B(x, h)] = \phi_x(h) > 0,$$

(A₂) Conditionally to $X = x$, the random variables T and C are independent;

Assumptions about the regularity of conditional distribution functions

$$(F_1) \quad \exists C_1 > 0, C_2 > 0, b_1 > 0 \text{ and } b_2 > 0, \quad \forall (x_1, x_2) \in V_x^2, \forall (t_1, t_2) \in [0, \tau]^2, \quad :$$

- i) $|H(t_1/x_1) - H(t_2/x_2)| \leq C_1(d^{b_1}(x_1, x_2) + |t_1 - t_2|^{b_2});$
- ii) $|f^*(t_1/x_1) - f^*(t_2/x_2)| \leq C_2(d^{b_1}(x_1, x_2) + |t_1 - t_2|^{b_2}).$

$$(F_2) \quad \exists \mu < \infty, \quad f^*(t/x) < \mu, \quad \forall (t, x) \in [0, \tau] \times V_x,$$

$$(F_3) \quad \exists \eta > 0, \quad 1 - H(t/x) \geq \eta, \quad \forall (t, x) \in [0, \tau] \times V_x.$$

(F₄) The function $t \mapsto \lambda(t/x)$ has a continuous second derivative with respect to t and satisfies:

- i) $\lambda'(\theta/x) = 0;$
- ii) $d_x = \inf_{t \in [0, \tau]} |\lambda''(t/x)| > 0.$

(F₅) The function $t \mapsto H(t/x)$ has a continuous first derivative with respect to t denoted $H'(t/x)$ and uniformly bounded on S .

Assumptions on the Kernels

(K₁) The cumulative kernel L is hölderian of order a with $a \in]0, 1]$.

(K₂) The cumulative kernel L is differentiable and its derivative $L' = K_0$ satisfies:

- i) K_0 has compact support $[-1, 1]$ and $K_0(t) > 0, \forall t \in [-1, 1];$
- ii) $\exists C_5 > 0, \forall (t_1, t_2) \in [-1, 1]^2, |K_0(t_1) - K_0(t_2)| \leq C_5|t_1 - t_2|.$

(K₃) The functional kernel K has compact support $[0, 1]$, in particular $\exists M > 0, M' > 0, \forall m \geq 1, \forall t \in [0, 1], M \leq K^m(t) \leq M'.$

Assumptions on smoothing parameters

(H₁) The functional bandwidth $h_{K,n}$ satisfies the following conditions:

$$\lim_{n \rightarrow +\infty} h_{K,n} = 0 \quad \text{and} \quad \lim_{n \rightarrow +\infty} \frac{\log n}{n\phi_x(h_{K,n})} = 0.$$

(H₂) The cumulative kernel bandwidth $h_{L,n}$ satisfies :

$$\lim_{n \rightarrow +\infty} h_{L,n} = 0 \quad \text{and} \quad \exists \alpha > 0, \quad \lim_{n \rightarrow +\infty} n^\alpha h_{L,n} = \infty.$$

(H₃) $\frac{nh_L}{\log n} \rightarrow +\infty$ and $\lim_{n \rightarrow +\infty} \frac{\log n}{nh_L\phi_x(h_K)} = 0.$

Despite their number, our assumptions are not very restrictive. On the one hand, they are not specific to the problem of estimating the conditional hazard rate and its maximum risk. On the other hand, they correspond to the assumptions usually made in the context of non-functional random variables. The assumptions on the kernels and the assumptions on the smoothing parameters ensure the correct behavior of the $H_n(\cdot/x)$ and $\hat{f}_n^*(\cdot/x)$ estimators and hence $\hat{\lambda}_n(\cdot/x)$ (see in Ferraty and Vieu (2006)).

3.2. Results

Recall that a sequence $(Z_n)_{n \in \mathbb{N}}$ of almost everywhere (*a.e*)-finite real random variables converges almost completely (*p.co*) to an *a.e.* finite real random variables Z if for every $\epsilon > 0$, $\sum_{n \in \mathbb{N}} P[|Z_n - Z| > \epsilon] < +\infty$. Hence by the Borel-Cantelli Lemma the almost complete convergence implies the almost sure convergence (*p.s.*) of X_n to X . Similarly, given a sequence $(U_n)_{n \in \mathbb{N}}$ of real numbers, we say that $X_n = O(U_n)$ *p.co* if there exists a constant $M > 0$ such that $\limsup_{n \rightarrow +\infty} | \frac{X_n}{U_n} | \leq M$ *p.co* (see also (Lo and al., 2016), Section 5, page 121).

Theorem 1. *Assume that the assumptions (A_1) , (A_2) , $(F_1) - (F_3)$ and (F_5) , $(K_1) - (K_3)$ and $(H_1) - (H_3)$ are satisfied. Then for all $x \in \mathcal{F}$ fixed, we have:*

$$\sup_{t \in S} |\hat{\lambda}_n(t/x) - \lambda(t/x)| = \mathcal{O}\left(h_{K,n}^{b_1} + h_{L,n}^{b_2}\right) + \mathcal{O}\left(\sqrt{\frac{\log n}{nh_{L,n}\phi_x(h_{K,n})}}\right) \quad p.co \quad (3.1)$$

Proof.

The proof of this theorem is based on the following decomposition valid for all $t \in S$:

$$\begin{aligned} & \hat{\lambda}_n(t/x) - \lambda(t/x) \\ &= \frac{1}{1 - \hat{H}_n(t/x)} \left[\hat{f}_n^*(t/x) - f^*(t/x) \right] + \frac{\lambda(t/x)}{1 - \hat{H}_n(t/x)} [H_n(t/x) - H(t/x)]. \end{aligned}$$

Considering the assumptions (F_1) which ensure the uniform continuity of $\lambda(\cdot/x)$ on the compact interval S and the hypothesis (F_3) , we immediately obtain the inequality

$$\begin{aligned} \sup_{t \in S} |\hat{\lambda}_n(t/x) - \lambda(t/x)| &\leq cte. \frac{\sup_{t \in S} |\hat{f}_n^*(t/x) - f^*(t/x)| + \sup_{t \in S} |\hat{H}_n(t/x) - H(t/x)|}{\inf_{t \in S} \liminf_{n \rightarrow +\infty} |1 - \hat{H}_n(t/x)|} \\ &\leq M_0 \left[\sup_{t \in S} |\hat{f}_n^*(t/x) - f^*(t/x)| + \sup_{t \in S} |\hat{H}_n(t/x) - H(t/x)| \right], \end{aligned} \quad (3.2)$$

for a constant $M_0 > 0$. The complete proof of the Theorem 1 finally follows from the following lemmas:

Lemma 1. *Let the assumptions (A_1) , (A_2) , $(F_1) - (F_3)$ and (F_5) , $(K_1) - (K_3)$ and $(H_1) - (H_2)$ hold. We have:*

$$\sup_{t \in S} |\hat{H}_n(t/x) - H(t/x)| = \mathcal{O} \left(h_{K,n}^{b_1} + h_{L,n}^{b_2} \right) + \mathcal{O} \left(\sqrt{\frac{\log n}{n \phi_x(h_{K,n})}} \right), \quad p.co. \quad (3.3)$$

Lemma 2. *Under the conditions of Theorem 1, we have:*

$$\sup_{t \in S} |\hat{f}_n^*(t/x) - f^*(t/x)| = \mathcal{O} \left(h_{K,n}^{b_1} + h_{L,n}^{b_2} \right) + \mathcal{O} \left(\sqrt{\frac{\log n}{nh_{L,n} \phi_x(h_{K,n})}} \right), \quad p.co. \quad (3.4)$$

To simplify the demonstrations of the above results, we introduce the following notations:

- (a) $K_i(x) = K \left(\frac{d(x, X_i)}{h_{K,n}} \right); \quad L_i(t) = L \left(\frac{t - Y_i}{h_{L,n}} \right); \quad K_{0,i}(t) = L'_i(t) = L' \left(\frac{t - Y_i}{h_{L,n}} \right), \quad i = 1, \dots, n$
- (b) $\Theta_i = \frac{K_i(x)}{\mathbb{E}K_1(x)}, \quad \Delta_i = \frac{K_i(x)L_i(t)}{\mathbb{E}K_1(x)}, \quad \Omega_i = \frac{K_i(x)K_{0,i}(t)\delta_i}{h_{L,n}\mathbb{E}K_1(x)}, \quad i = 1, \dots, n,$
- (c) $\hat{\Psi}_1(x) = \frac{1}{n} \sum_{i=1}^n \Theta_i = \frac{1}{n} \sum_{i=1}^n \frac{K_i(x)}{\mathbb{E}K_1(x)} = \frac{1}{n\mathbb{E}K_1(x)} \sum_{i=1}^n K_i(x),$
- (d) $\hat{\Psi}_2(t, x) = \frac{1}{n} \sum_{i=1}^n \Delta_i = \frac{1}{n} \sum_{i=1}^n \frac{K_i(x)L_i(t)}{\mathbb{E}K_1(x)} = \frac{1}{n\mathbb{E}K_1(x)} \sum_{i=1}^n K_i(x)L_i(t),$
- (e) $\hat{\Psi}_3(t, x) = \frac{1}{n} \sum_{i=1}^n \Omega_i = \frac{1}{n} \sum_{i=1}^n \frac{K_i(x)K_{0,i}(t)\delta_i}{h_{L,n}\mathbb{E}K_1(x)} = \frac{1}{nh_{L,n}\mathbb{E}K_1(x)} \sum_{i=1}^n K_i(x)K_{0,i}(t)\delta_i.$

These notations lead to the following representations:

$$\hat{H}_n(t/x) = \frac{\sum_{i=1}^n K \left(\frac{d(x, X_i)}{h_{K,n}} \right) L \left(\frac{t - Y_i}{h_{L,n}} \right)}{\sum_{i=1}^n K \left(\frac{d(x, X_i)}{h_{K,n}} \right)} = \frac{\hat{\Psi}_2(t, x)}{\hat{\Psi}_1(x)}. \quad (3.5)$$

$$\hat{f}_n^*(t/x) = \frac{\sum_{i=1}^n K \left(\frac{d(x, X_i)}{h_{K,n}} \right) K_0 \left(\frac{t - Y_i}{h_{L,n}} \right) \delta_i}{h_{L,n} \sum_{i=1}^n K \left(\frac{d(x, X_i)}{h_{K,n}} \right)} = \frac{\hat{\Psi}_3(t, x)}{\hat{\Psi}_1(x)}. \quad (3.6)$$

Proof of Lemma 1.

The proof of this lemma is based on the following decomposition:

$$\begin{aligned} & \hat{H}_n(t/x) - H(t/x) \\ = & \frac{\left[\hat{\Psi}_2(t, x) - \mathbb{E}\hat{\Psi}_2(t, x) \right] - \left[H(t/x) - \mathbb{E}\hat{\Psi}_2(t, x) \right] + H(t/x) \left[\mathbb{E}\hat{\Psi}_1(x) - \hat{\Psi}_1(x) \right]}{\hat{\Psi}_1(x)} \end{aligned} \tag{3.7}$$

By the Strong Law of Large Numbers (SLLN), we almost surely have $\lim_{n \rightarrow +\infty} \hat{\Psi}_1 = 1$. On the other hand, according to the notations in (c) and (d), we can write

$$\hat{\psi}_1 - \mathbb{E}\hat{\psi}_1 = \frac{1}{n} \sum_{i=1}^n (\Theta_i - \mathbb{E}\Theta_i), \quad \hat{\psi}_2 - \mathbb{E}\hat{\psi}_2 = \frac{1}{n} \sum_{i=1}^n (\Theta_i L_i(t) - \mathbb{E}\Theta_i L_i(t)) \quad \text{and}$$

$$\begin{aligned} \mathbb{E}\hat{\Psi}_2(t, x) - H(t/x) &= \mathbb{E}\Theta_1 L_1(t) - H(t/x) = \mathbb{E}[\Theta_1 (\mathbb{E}(L_1/X) - H(t/x))] \\ &= \mathbb{E}[\Theta_1 1_{B(x, h_K)}(X) (\mathbb{E}(L_1/X) - H(t/x))]. \end{aligned}$$

In the sequel, we will write simply h_K (resp. h_L) to designate the sequence $h_{K,n}$ (respectively $h_{L,n}$).

Asymptotic behavior of $\hat{\psi}_1 - \mathbb{E}\hat{\psi}_1$

$$\begin{aligned} \hat{\psi}_1 - \mathbb{E}\hat{\psi}_1 &= \frac{1}{n} \sum_{i=1}^n (\Theta_i - \mathbb{E}\Theta_i). \text{ Let us consider } Z_i = \Theta_i - \mathbb{E}\Theta_i. \text{ For all } m \geq 1, \text{ we have} \\ \mathbb{E}\Theta_i^m &= \frac{1}{(\mathbb{E}K_i(x))^m} \mathbb{E}K_i^m(x). \text{ Under the assumption } (K_3) \text{ we check that} \end{aligned}$$

$$M\phi_x(h_K) \leq \mathbb{E}K_i^m(x) \leq M'\phi_x(h_K).$$

This leads to

$$\frac{M}{\phi_x^{m-1}(h_K)} \leq \mathbb{E}\Theta_i^m \leq \frac{M'}{\phi_x^{m-1}(h_K)}. \tag{3.8}$$

Therefore

$$\mathbb{E}|\Theta_i^m| = \mathcal{O}(\phi_x^{-m+1}(h_K)).$$

But

$$Z_i^m = (\Theta_i - \mathbb{E}\Theta_i)^m = (\Theta_i - 1)^m = \sum_{k=0}^m C_m^k \Theta_i^k (-1)^{m-k} \text{ with } C_m^k = \frac{m!}{k!(m-k)!},$$

and we can therefore write

$$\begin{aligned} \mathbb{E}|Z_i^m| &= \mathbb{E}|(\Theta_i - \mathbb{E}\Theta_i)^m| = \mathbb{E}|(\Theta_i - 1)^m| \leq M \sum_{k=0}^m \mathbb{E}|\Theta_i^k| \\ &\leq M.m \max_{k=0; \dots; m} \mathbb{E}|\Theta_i^k| \\ &\leq C_m \phi_x^{-m+1}(h_K) \end{aligned} \tag{3.9}$$

with $C_m = M.m$. It is therefore established that for every $m \geq 2$,

$$\mathbb{E}|Z_i^m| \leq C_m \phi_x^{-m+1}(h_K) = C_m a^{2(m-1)} \quad \text{with } a^2 = \frac{1}{\phi_x(h_K)}. \quad (3.10)$$

Under Assumption (H_1) the conditions of Corollary A.8 in Ferraty and Vieu (2006) page 234 are satisfied. From this corollary we have

$$|\hat{\psi}_1(x) - \mathbb{E}\hat{\Psi}_1(x)| = \mathcal{O} \left(\left(\frac{\log n}{n\phi_x(h_K)} \right)^{1/2} \right), p.co \quad (3.11)$$

Asymptotic behavior of $\hat{\psi}_2 - \mathbb{E}\hat{\psi}_2$

The compactness of the interval $S = [0, \tau]$ allows to cover it by μ_n disjoint intervals so that

$$\mathcal{S} \subset \bigcup_{k=1}^{\mu_n} [\tau_k - l_n, \tau_k + l_n[\quad (3.12)$$

where $\tau_1, \dots, \tau_{\mu_n}$ are points of \mathcal{S} and where l_n and μ_n are chosen such that

$$\exists C_\mu > 0, \exists c, l_n = C_\mu \mu_n^{-1} = n^{-c}. \quad (3.13)$$

For each $t \in \mathcal{S}$, we denote by

$$\tau_t = \operatorname{argmin}_{\tau \in \{\tau_1, \dots, \tau_{\mu_n}\}} |t - \tau|.$$

This allow to write

$$\sup_{t \in \mathcal{S}} |\hat{\Psi}_2(t, x) - \mathbb{E}\hat{\Psi}_2(t, x)| = D_1 + D_2 + D_3, \quad (3.14)$$

where

$$D_1 = \sup_{t \in \mathcal{S}} |\hat{\Psi}_2(t, x) - \hat{\Psi}_2(\tau_t, x)|$$

$$D_2 = \sup_{t \in \mathcal{S}} |\hat{\Psi}_2(\tau_t, x) - \mathbb{E}\hat{\Psi}_2(\tau_t, x)|$$

$$D_3 = \sup_{t \in \mathcal{S}} |\mathbb{E}\hat{\Psi}_2(t, x) - \mathbb{E}\hat{\Psi}_2(\tau_t, x)|$$

Consider D_1 and D_3 . Under Assumption (K_1) , we have:

$$\begin{aligned} |\hat{\Psi}_2(t, x) - \hat{\Psi}_2(\tau_t, x)| &= \frac{1}{n} \sum_{i=1}^n \Theta_i \left| L \left(\frac{t - Y_i}{h_L} \right) - L \left(\frac{\tau_t - Y_i}{h_L} \right) \right| \\ &\leq \frac{a}{n} \sum_{i=1}^n \Theta_i \frac{|t - \tau_t|}{h_L} \\ &\leq a \frac{l_n}{h_L} \\ &\leq \frac{M}{n^c h_L}. \end{aligned}$$

Hence

$$\sup_{t \in \mathcal{S}} |\hat{\Psi}_2(t, x) - \hat{\Psi}_2(\tau_t, x)| \leq \frac{M}{n^c h_L}$$

and by the Assumption (H_2) we get

$$\lim_{n \rightarrow \infty} D_1 = 0, p.co.$$

Similar reasoning makes possible to establish that

$$|\mathbb{E}\hat{\Psi}_2(t, x) - \mathbb{E}\hat{\Psi}_2(\tau_t, x)| \leq \frac{M}{n^c h_L}.$$

Consequently

$$\lim_{n \rightarrow \infty} D_3 = 0, p.co.$$

We are going to handle D_2 . We can write that for every $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P} \left[\sup_{t \in \mathcal{S}} |\hat{\Psi}_2(\tau_t, x) - \mathbb{E}\hat{\Psi}_2(\tau_t, x)| > \varepsilon \right] &= \mathbb{P} \left[\max_{j=1, \dots, \mu_n} |\hat{\Psi}_2(\tau_j, x) - \mathbb{E}\hat{\Psi}_2(\tau_j, x)| > \varepsilon \right] \\ &\leq \mu_n \max_{j=1, \dots, \mu_n} \mathbb{P} \left[|\hat{\Psi}_2(\tau_j, x) - \mathbb{E}\hat{\Psi}_2(\tau_j, x)| > \varepsilon \right]. \end{aligned}$$

In other hand, for all $j = 1, \dots, \mu_n$ we have

$$\hat{\Psi}_2(\tau_j, x) - \mathbb{E}\hat{\Psi}_2(\tau_j, x) = \frac{1}{n} \sum_{i=1}^n [\Theta_i L_i(\tau_j) - \mathbb{E}\Theta_i L_i(\tau_j)] = \frac{1}{n} \sum_{i=1}^n [T_{ji} - \mathbb{E}T_{ji}]$$

with $T_{ji} = \Theta_i L_i(\tau_j)$.

We will therefore use the same techniques as before. To do so, let us consider $Z_{ji} = T_{ji} - \mathbb{E}T_{ji}$. Since T_{ji} is positive, and given that the kernel L_i is bounded (Assumption (K_3)), we have

$$|Z_{ji}| \leq |T_{ji}| = |\Theta_i L_i(\tau_j)| \leq M |\Theta_i|.$$

Thus, by reasoning as before, we obtain similar inequalities. By using the relation (3.13) with ε_0 large enough, we have

$$\sum_{j=1}^{\infty} \mathbb{P} \left[\sup_{t \in \mathcal{S}} |\hat{\Psi}_2(\tau_j, x) - \mathbb{E}\hat{\Psi}_2(\tau_j, x)| > \varepsilon_0 \sqrt{\frac{\log n}{n \phi_x(h_K)}} \right] < \infty.$$

Therefore

$$D_2 = \mathcal{O} \left(\left(\frac{\log n}{n \phi_x(h_K)} \right)^{1/2} \right), p.co.$$

and ultimately

$$\sup_{t \in \mathcal{S}} |\hat{\Psi}_2(t, x) - \mathbb{E}\hat{\Psi}_2(t, x)| = \mathcal{O} \left(\left(\frac{\log n}{n\phi_x(h_K)} \right)^{1/2} \right) \text{ p.co.} \quad (3.15)$$

Asymptotic behavior of $H - \mathbb{E}\hat{\Psi}_2$:

We have

$$\begin{aligned} \mathbb{E}\hat{\Psi}_2(t, x) - H(t/x) &= \mathbb{E}\Theta_1 L_1(t) - H(t/x) \\ &= \mathbb{E} [\Theta_1 (\mathbb{E}(L_1/X) - H(t/x))] \\ &= \mathbb{E} [\Theta_1 1_{B(x, h_K)}(X) (\mathbb{E}(L_1/X) - H(t/x))] \end{aligned}$$

ince $L' = K_0$, we have

$$\begin{aligned} \mathbb{E}(L_1(t)/X) &= \int_{\mathbb{R}} L \left(\frac{t-u}{h_L} \right) d\mathbb{P}(u/X) \\ &= \int_{\mathbb{R}} \int_{-\infty}^{\frac{t-u}{h_L}} K_0(v) dv d\mathbb{P}(u/X) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} K_0(v) 1_{[v, +\infty[} \left(\frac{t-u}{h_L} \right) dv d\mathbb{P}(u/X) \\ &= \int_{\mathbb{R}} K_0(v) \left(\int_{\mathbb{R}} 1_{[v, +\infty[} \left(\frac{t-u}{h_L} \right) d\mathbb{P}(u/X) \right) dv \quad (\text{by Fubini's Theorem}) \\ &= \int_{\mathbb{R}} K_0(v) H(t - vh_L/X) dv. \end{aligned}$$

ince K_0 is a probability density, we can write

$$\begin{aligned} \mathbb{E}(L_1/X) - H(t/x) &= \int_{\mathbb{R}} K_0(v) H(t - vh_L/X) dv - \int_{\mathbb{R}} K_0(v) H(t/x) dv \\ &= \int_{\mathbb{R}} K_0(v) [H(t - vh_L/X) - H(t/x)] dv. \end{aligned} \quad (3.16)$$

We have

$$|H(t - vh_L/X) - H(t/x)| \leq |H(t - vh_L/X) - H(t - vh_L/x)| + |H(t - vh_L/x) - H(t/x)| \quad (3.17)$$

and by hypothesis $(F_1)i$ for all $t \in \mathcal{S}$, we have

$$\sup_{v \in [-1, 1]} 1_{B(x, h_K)}(X) |H(t - vh_L/X) - H(t - vh_L/x)| \leq M h_K^{b_1}$$

and

$$\sup_{v \in [-1, 1]} |H(t - vh_L/x) - H(t/x)| \leq M' h_L^{b_2}.$$

Let $M_{\max} = \max(M, M')$. Then we get via (3.16), (3.17) and the above increments

$$\sup_{t \in \mathcal{S}} 1_{B(x, h_K)}(X) |\mathbb{E}(L_1(t)/X) - H(t/x)| \leq M_{\max} (h_K^{b_1} + h_L^{b_2}).$$

ince $\mathbb{E}\Theta_1 = 1$, we finally get

$$\sup_{t \in \mathcal{S}} |\mathbb{E}\hat{\Psi}_2(t, x) - H(t/x)| = \mathcal{O} (h_K^{b_1} + h_L^{b_2}). \tag{3.18}$$

In view of the relations (3.7), (3.11), (3.15) and (3.18) Lemma 1 is proved.

Proof of Lemma 2

The proof of this lemma is based on the following decomposition:

$$\begin{aligned} & \hat{f}_n^*(t/x) - f^*(t/x) \\ = & \frac{[\hat{\Psi}_3(t, x) - \mathbb{E}\hat{\Psi}_3(t, x)] - [f^*(t/x) - \mathbb{E}\hat{\Psi}_3(t, x)] + f^*(t/x) [\mathbb{E}\hat{\Psi}_1(x) - \hat{\Psi}_1(x)]}{\hat{\Psi}_1(x)}. \end{aligned}$$

The result of Lemma 2 follows directly from the results below demonstrated in the same way as before under the assumptions (H_1) - (H_3)

$$|\hat{\Psi}_1(x) - \mathbb{E}\hat{\Psi}_1(x)| = \mathcal{O} \left(\left(\frac{\log n}{n\phi_x(h_{K,n})} \right)^{\frac{1}{2}} \right), \text{ p.co.} \tag{3.19}$$

$$\sup_{t \in \mathcal{S}} |\hat{\Psi}_3(t, x) - \mathbb{E}\hat{\Psi}_3(t, x)| = \mathcal{O} \left(\left(\frac{\log n}{nh_{L,n}\phi_x(h_{K,n})} \right)^{\frac{1}{2}} \right), \text{ p.co.} \tag{3.20}$$

$$\sup_{t \in \mathcal{S}} |f^*(t/x) - \mathbb{E}\hat{\Psi}_3(t, x)| = \mathcal{O} (h_{K,n}^{b_1} + h_{L,n}^{b_2}), \text{ p.co.} \tag{3.21}$$

Proof of Theorem 1

We apply the Lemma 1 and Lemma 2 to each of the terms of the decomposition (3.7), noting that

$$\frac{\log n}{n\phi_x(h_{K,n})} = o \left(\frac{\log n}{nh_{L,n}\phi_x(h_{K,n})} \right), \quad n \rightarrow +\infty$$

and Theorem 1 is proved.

We apply the above results to estimate the maximum risk point.

4. Maximum risk point estimation

Given $x \in \mathcal{F}$ fixed, we assume that the conditional hazard function $\lambda(\cdot/x)$ has a unique maximum in an interval $[a_x, b_x] \subset S = [0, \tau]$ and its maximum risk point denoted by $\theta = \theta(x)$ is defined by :

$$\lambda(\theta/x) := \sup_{t \in S} \lambda(t/x). \tag{4.1}$$

A kernel estimator of θ is then defined by the random variable $\hat{\theta}_n := \hat{\theta}_n(x)$ which maximizes the kernel estimator $\hat{\lambda}_n(\cdot/x)$ of $\lambda(\cdot/x)$. In other words,

$$\hat{\lambda}_n(\hat{\theta}_n/x) := \hat{\lambda}_n(\hat{\theta}_n/x) = \sup_{t \in S} \hat{\lambda}_n(t/x). \tag{4.2}$$

More precisely $\hat{\theta}_n$ is defined by :

$$\hat{\theta}_n = \inf \left\{ t \in S : \hat{\lambda}_n(t/x) = \sup_{s \in S} \hat{\lambda}_n(s/x) \right\}. \tag{4.3}$$

Theorem 2. *In addition to the hypotheses of Theorem 1, if the hypothesis (F_4) is also satisfied, then for any $x \in \mathcal{F}$ fixed,*

$$|\hat{\theta}_n(x) - \theta(x)| = \mathcal{O} \left((h_{K,n}^{b_1} + h_{L,n}^{b_2})^{\frac{1}{2}} \right) + \mathcal{O} \left(\left(\frac{\log n}{nh_{L,n} \phi_x(h_{K,n})} \right)^{\frac{1}{4}} \right), \quad p.co. \tag{4.4}$$

Proof of Theorem 2

Under the hypothesis (F_4) *i*), the Taylor expansion of the function $\lambda(\cdot/x)$ in the neighborhood of the θ gives:

$$\lambda(\hat{\theta}_n/x) = \lambda(\theta/x) + \left(\hat{\theta}_n - \theta \right)^2 \frac{\lambda''(\theta^*/x)}{2!}, \tag{4.5}$$

where θ^* is between θ and $\hat{\theta}_n$ in the compact $[a_x, b_x] \subset S = [0, \tau]$. Thus we have :

$$|\hat{\theta}_n(x) - \theta(x)|^2 \leq \frac{2!}{\min_{t \in S} |\lambda''(t/x)|} |\lambda(\hat{\theta}_n(x)/x) - \lambda(\theta(x)/x)|. \tag{4.6}$$

Now by the triangular inequality, we have the following increase:

$$\begin{aligned} |\lambda(\hat{\theta}_n(x)/x) - \lambda(\theta(x)/x)| &\leq |\lambda(\hat{\theta}_n(x)/x) - \hat{\lambda}_n(\theta(x)/x)| + |\hat{\lambda}_n(\theta(x)/x) - \lambda(\theta(x)/x)| \\ &\leq 2 \sup_{t \in S} |\hat{\lambda}_n(t/x) - \lambda(t/x)|. \end{aligned}$$

Thus (4.6) becomes:

$$|\hat{\theta}_n(x) - \theta(x)|^2 \leq \frac{4}{\min_{t \in S} |\lambda''(t/x)|} |\hat{\lambda}_n(t/x) - \lambda(t/x)|. \tag{4.7}$$

Applying Theorem 1 and the hypothesis (F_4) *ii*) to the right side of the inequality (4.7), we get the result. This completes the proof of this theorem.

5. Simulations

5.1. Empirical Validation of estimators

In this empirical study, we consider a usual law that is often found in parametric models of survival times: the Log-Normal distribution with probability density function $q(t)$ defined on \mathbb{R}_+^* by

$$q(t) = \frac{1}{\sigma 2\pi t} \exp\left(-\frac{[\log(t) - \mu]^2}{2\sigma^2}\right)$$

We simulate n (for $n \in \{1000, 2000, 3000, 5000, 7000, 10000\}$) independent observations representing the durations T_i , $1 \leq i \leq n$. We also simulate n independent observations which represent the values of the random censorship C . For the functional covariate X , we also simulate n independent observations such that $X(t) = \cos(2\pi Zt)$ for any $t \in [0, 1]$ where the variable Z is uniformly distributed on $[0.5, 1]$. Moreover, the conditional distribution of T , given X is a Log-Normal distribution of mean $\mu(X) = \|X\|_2$ and standard deviation $\sigma(X) = 1.5$ with

$$\|X\|_2^2 = \int_0^1 X^2(t)dt = \frac{1}{2} \left[1 + \frac{\sin(4\pi Z)}{4\pi Z} \right],$$

and the conditional distribution of C given X is an exponential distribution of parameter $\alpha(X) = (1 + \nu)\|X\|_2 + 1.5$ where the parameter ν will be varied to obtain the percentage of censorship. With a value of the parameter ν set near 3.5, we observed an average censorship rate of about 30% for each of the three values of the functional covariate x ($x \in \{0.7, 0.8, 0.9\}$). The choice of the semi-metric d is a recurrent question in the functional estimation [Ferraty and Vieu \(2006\)](#). In this paper we consider the semi-metric d defined by (see in [Girard and Gardes \(2013\)](#)):

$$d(s, t) = \left| \|s\|_2^2 - \|t\|_2^2 \right|.$$

The estimator $\tilde{\lambda}_n$ depends on the choice of two kernels K and K_0 and two windows h_K and h_L . The choice of kernels has little effect on the performance of the estimator, which has led us to favor the classical Gaussian kernel solution for K_0 and the kernel of Epanechnikov for K . The choice of the smoothing parameters h_K and h_L is, on the other hand, crucial. As an optimal window, we choose the value that minimizes the Mean Integrated Square Error (MISE). In fact, since the theoretical calculation of the MISE is too complicated for censored data, we repeat the previous simulations 100 times in order to obtain an empirical series $\{\tilde{\lambda}_n^j\}_{j=1, \dots, 100}$. It is possible to estimate naturally the empirical MISE on a compact $[\min(Y), \max(Y)]$:

$$MISE(h) = \frac{1}{100} \sum_{j=1}^{100} \int_{\min(Y)}^{\max(Y)} \left[\tilde{\lambda}_n^j(t/x) - \lambda(t/x) \right]^2 dt,$$

for a given bandwidth h and for x fixed. Thus, a scan on the values of h_K and h_L allows to determine the optimal values h_K^{opt} and h_L^{opt} of the bandwidths h_K and h_L which minimize the empirical MISE:

$$(h_K^{opt}, h_L^{opt}) = \arg \min_{h_K, h_L} \widehat{MISE}(h_K, h_L).$$

5.2. Results

In this subsection, we represent the various estimators $\hat{\lambda}_n$ and $\hat{\theta}_n$ of the conditional hazard function and the maximum risk point for each of the values of x ($x \in \{0.7; 0.8; 0.9\}$) for 6 samples of different sizes. These estimators, represented respectively by the curve in bold and orange color (for the hazard rate) and by the vertical line in dotted and blue color (for the maximum risk point), are calculated using the optimal windows h_K^{opt} and h_L^{opt} , for a set of n different data Simulated; The other curve in blizzard blue color and vertical line in red color correspond respectively to the hazard function and the maximum risk point of the theoretical hazard function.

The results of the estimated MISE and the optimal bandwidths values derived therefore are summarized in Tables 1 to 3 for each of the values of x and for each sample of size n .

TABLE 1: Maximum risk point nonparametric estimation of the conditional hazard rate with $x = 0.7$

covariate	$x = 0.7$						
	$\theta(x)$	$\hat{\theta}_n(x)$	$\lambda(\theta(x)/x)$	$\tilde{\lambda}_n(\hat{\theta}_n(x)/x)$	Mise	h_K^{opt}	h_L^{opt}
1000	0.28	0.50	0.6622886	0.6091518	0.01926373	0.5011872	0.25118860
2000	0.28	0.45	0.6622886	0.6214251	0.01360150	0.4676242	0.21867240
3000	0.28	0.42	0.6622886	0.6303655	0.01252181	0.4490429	0.20163960
5000	0.28	0.39	0.6622886	0.6396525	0.01088368	0.4266807	0.18205640
7000	0.28	0.36	0.6622886	0.6510653	0.00966699	0.4125629	0.17020820
10000	0.28	0.32	0.6622886	0.6546703	0.00926452	0.3981072	0.15848930

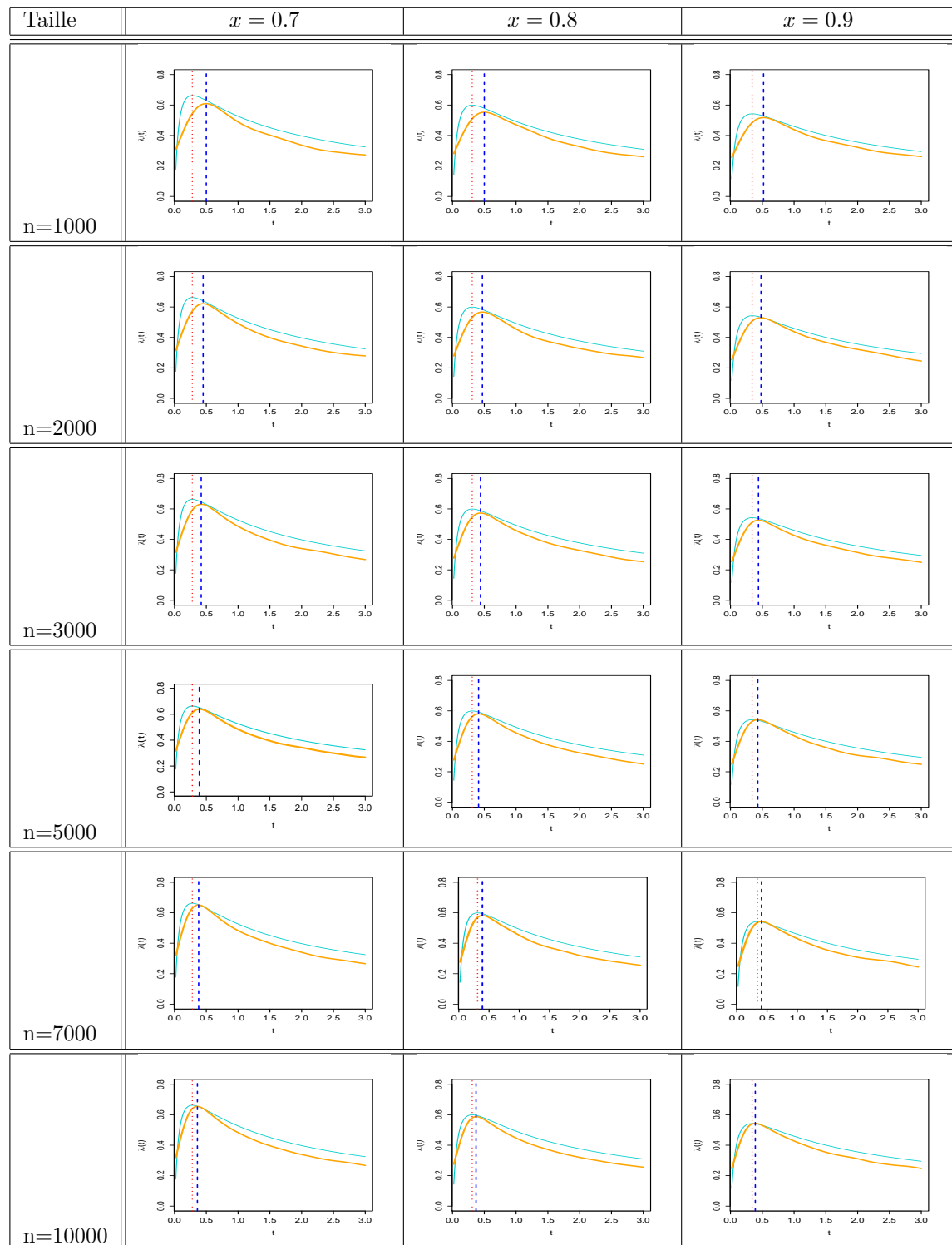
TABLE 2: Maximum risk point nonparametric estimation of the conditional hazard rate with $x = 0.8$

covariate	$x = 0.8$						
	$\theta(x)$	$\hat{\theta}_n(x)$	$\lambda(\theta(x)/x)$	$\tilde{\lambda}_n(\hat{\theta}_n(x)/x)$	Mise	h_K^{opt}	h_L^{opt}
1000	0.31	0.50	0.5992638	0.5530352	0.01515300	0.5011872	0.25118860
2000	0.31	0.47	0.5992638	0.5667058	0.01276774	0.4676242	0.21867240
3000	0.31	0.44	0.5992638	0.5726265	0.01095756	0.4490429	0.20163960
5000	0.31	0.41	0.5992638	0.5804458	0.01000052	0.4266807	0.18205640
7000	0.31	0.36	0.5992638	0.5820060	0.00943976	0.4125629	0.17020820
10000	0.31	0.34	0.5992638	0.5874231	0.00894696	0.3981072	0.15848930

TABLE 3: Maximum risk point nonparametric estimation of the conditional hazard rate with $x = 0.9$

covariate	$x = 0.9$						
	$\theta(x)$	$\hat{\theta}_n(x)$	$\lambda(\theta(x)/x)$	$\tilde{\lambda}_n(\hat{\theta}_n(x)/x)$	Mise	h_K^{opt}	h_L^{opt}
1000	0.34	0.52	0.5422315	0.5173909	0.01382277	0.5011872	0.2511886
2000	0.34	0.48	0.5422315	0.5295975	0.01012492	0.4676242	0.2186724
3000	0.34	0.44	0.5422315	0.5255532	0.00958982	0.4490429	0.2016396
5000	0.34	0.43	0.5422315	0.5420404	0.00765708	0.4266807	0.1820564
7000	0.34	0.40	0.5422315	0.5422305	0.00722211	0.4125629	0.1702082
10000	0.34	0.36	0.5422315	0.5422315	0.00682938	0.3981072	0.1584893

Curve of $\hat{\lambda}_n(t|x)$ in bold and orange color, that of $\lambda(t|x)$, the line in blizzard blue color, the
point $\hat{\theta}_n(x)$ in blue color and that of $\theta(x)$ in red color



Conclusion

We considered a nonparametric kernel estimator of the conditional hazard rate, the covariate taking values in a functional space. We have established its uniform (with respect to time) almost complete convergence. We applied the result to the nonparametric estimation of the high-risk point dependent on the covariate. This convergence can also be uniform in both time and space (covariate) by restricting the covariate to vary in a compact subspace of \mathcal{F} (see in Gneyou (2013, 2014)). We undertook a numerical study. The computational capabilities of our computers did not allow us to replication for larger values of N . This could lead numerically to the convergence and therefore to the suitability of the curves of $\lambda(t|x)$ and its nonparametric estimator $\hat{\lambda}_n(t|x)$. So we can say that, what we did not observe is good regardless to the theoretical aspect.

Acknowledgments: The authors want to thank to the Editor and one anonymous referee for a careful reading of the manuscript. Their valuable comments and suggestions during the preparation of this article are gratefully acknowledged.

References

- Dupuy, J.-F., Gneyou, K. (2011). A wavelet estimator of the intensity function with censored data. *Qual. Technol. Quant. Manage.*, 8(4), 401-410.
- Estévez-Pérez, G. (2002). On convergence rates for quadratic errors in kernel hazard estimation. *Statist. Probab. Lett.*, 57, 231-241.
- 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.
- Ferraty, F., Vieu, P. (2006). Nonparametric Functional Data Analysis. Theory and Practice. *Springer Series in Statistics*. Springer, New York.
- Ferraty, F., Rabhi, A. and Vieu, P. (2008). Estimation de la fonction de hasard avec variable explicative fonctionnelle. *Rev. Roumaine Math. Pures Appl.*, 53(1), 1-18.
- Gefeller, O., and Michels, P. (1992). A review on smoothing methods for the estimation of the hazard rate based on kernel functions. In: *Y. Dodge and J. Whittaker (Eds.), Computational Statistics*, pp. 459-464. Physica-Verlag.
- Girard, S., Gardes, L. (2013). Functional Kernel estimators of large conditional quantiles. *Int. Math. Stat.*, 6, 1715-1744.
- Gneyou, K. E. (1997). A functional law of iterated logarithm for a hazard rate process from censored data. *J. Rech. Sci. Univ. Bénin*, 1(2), 44-48.
- Gneyou, K. E. (2013). A central limit theorem for a nonparametric maximum conditional hazard rate in presence of right censoring. *Int. J. Stat. Probab.*, 2(3): 110-124.
- Gneyou, K. E. (2014). A strong linear representation for the maximum conditional hazard rate estimator in survival analysis. *Journal of Multivariate Analysis.*, 128: 10-18.
- Hassani, S., Sarda, P., and Vieu, P. (1986). Approche non-paramétrique en théorie de la fiabilité: revue bibliographique. *Rev. Statist. Appl.*, 35: 27-41.
- Laksaci A., Mechab B. (2010). Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle cas des données spatiales. *Rev. Roumaine Math. Pures*

- Appl.*, 55: 35-51.
- LO G.S., Ngom N. and Kpanzou T. A. (2016). Weak convergence (IA), sequences of random vectors. *Statistics and Probability African Society (SPAS) Books Series. Calgary, Alberta. 2016. DOI : <https://arxiv.org/pdf/1610.05415.pdf>*
- Nadaraya, E., A. (1964). On estimating regression. *Theory of probability and its applications*, 9(1), 141-142.
- Nassiri, A., Delecroix, M. and Bonneu, M. (2000). Estimation non-paramétrique du taux de hasard: applications à des durées de chômage censurées à droites. *Annales d'économie et de Statistiques*, 58: 215-232.
- Pascu, M., Vaduva, I. (2003). Nonparametric estimation of the hazard rate: a survey. *Rev. Roumaine Math. Pures Appl.*, 48, 173-191.
- Patil, P. N., Wells, M. T. and Maron, J. S. (1994). Some heuristics of kernel based estimators of ratio functions. *J. Nonparametric Statist.*, 4: 203-209.
- Quintela-del-Rio, A. (2006). Nonparametric estimation of the maximum hazard under dependence conditions. *Statist. Probab. Lett.*, 76: 1117-1124.
- Rabhi A., Hammou Y. and Djebbouri, T. (2015). Nonparametric estimation of the maximum of conditional hazard function under dependence conditions for functional data. *Afrika Statistika*, 10(1): 777-794.
- Ramsay, J., Silverman, B. (2005). Functional data Analysis, 2nd Ed. *Springer series in statistics*, Springer, New York.
- Roussas, G. (1969). Nonparametric estimation of the transition distribution function of a Markov process. *Annals of Mathematical Statistics*, 40: 1386-1400.
- Samanta, M. (1989). Nonparametric estimation of conditional quantile. *Statist. Proba. Letters*, 7 : 407-412.
- Singpurwalla, N., Wong, M., Y. (1983). Estimation of the failure rate - a survey of non-parametric models. Part I: Non-Bayesian methods. *Comm. Statist. Theory Methods*, 12, 559-588.
- Tanner, M., Wong, W. H. (1983). The estimation of the hazard function from randomly censored data by the kernel methods. *Ann. Statist.*, 11, 989-993.
- van Keilegom, I., Veraverbeke, N. (2001). Hazard rate estimation in nonparametric regression with censored data. *Ann. Inst. Statist. Math.*, 53: 730-745.
- Watson, G. S., Leadbetter, M. R. (1964). Hazard analysis. *Sankhyia, The Indian Journal of Statistics* 26: 101-116.
- Watson, G. S. (1964). Smooth regression analysis. *Sankhyia, The Indian Journal of Statistics (Series A)*, 26(4), 359-372.