Journal Afrika Statistika
ISSN 0825-0305

Journal Afrika Statistika
Vol. 6, 2011, pages 316–334.
DOI: http://dx.doi.org/10.4314/afst.v6i1.2

Strong and weak convergence of nonparametric estimator of regression in a competing risks model

Laurent Bordes[†] **and Kossi Essona Gneyou**[‡]

[†]Laboratoire de Mathématiques et de leurs Applications - UMR CNRS 5142, Université de Pau et des Pays de l'Adour - France
[‡]Département de Mathématiques - Université de Lomé B.P. 1515 Lomé - Togo

Received 31 January 2011; Accepted 20 September 2011

Abstract. In this paper we consider a competing risks model including covariates in which the observations are subject to random right censoring. Without any assumption of independence of the competing risks, and based on a nonparametric kernel-type estimator of the incident regression function, an estimator of the conditional regression function is proposed. We show that at a given covariate value and under suitable conditions the nonparametric estimator of the regression function is asymptotically normal. A simulation study is provided showing that our estimators have good behaviour for moderate sample sizes.

Résumé. Nous considérons dans ce papier un modèle de risques compétitifs dans lequel les observations sont soumises à une censure aléatoire à droite en présence de covariables. Sans aucune hypothèse d'indépendance sur les risques compétitifs, un estimateur non paramétrique de la fonction de répartition conditionnelle incidente est proposé. Cet estimateur est obtenu via celui d'un estimateur non paramétrique de type noyau de la fonction de régression incidente. Nous démontrons que pour une valeure fixée de la covariable, et sous certaines conditions, l'estimateur non paramétrique de la fonction de régression incidente est asymptotiquement normal. Des simulations illustrent le bon comportement de nos estimateurs pour des tailles modérées d'échantillons.

Key words: Competing risks; Nonparametric estimation; Kernel method; Regression function; right censoring.
AMS Mathematics Subject Classification : 62N02, 62G05, 62G20, 62H12.

Laurent Bordes: laurent.bordes@univ-pau.fr
Kossi Essona Gneyou: kgneyou@tg.refer.org

1. Introduction

Competing risks arise in medical, reliability or finance follow up involving multiple causes of failure but only the smallest failure time and its cause are observed. In competing risks mechanism, several failure times are right censored by observed failure time in informative manner and each failure time can be right censored by an event in non informative manner. For example, in order to determine the incidence of death due to breast cancer among breast cancer patients, every patient will be followed from a baseline time (e.g. date of diagnosis or of surgery) until the date of death due to breast cancer or study closing date. A patient who dies of breast cancer during the study period would be considered to have an 'event' at his date of death. A patient who is alive at the end of the study is considered as 'censored'. However, a patient can undergo a different event from the event of interest (e.g. death due to causes unrelated to the breast cancer disease). Such events are said competing risks events. As another practical example, consider women who start using an intrauterine device (IUD) (see Kalbfleisch and Prentice, [16]). They are subject to several risks, including accidental pregnancy, expulsion of the device, removal for medical reasons and removal for personal reasons. Kalbfleisch and Prentice discussed the areas of interest as a model of competing risks of IUD discontinuation. The events of competing risks are not independent a priori, hence they cannot be dealt with the statistical inference of standard censoring models (Kaplan and Meier, [17]). Before carrying on, we need to introduce the following notations.

Let $T_1, \ldots, T_m$ be failure times, due to any cause $j \in J = \{1, \ldots, m\}$. The indicator of failure cause will be denoted by η. Let $X = \min(T_1, \ldots, T_m)$ be the observed failure time with $X = T_j$ if and only if $\eta = j$. Assume that X is in turn, at risk of independent right-censoring by a non-negative random variable C. Set $Y = \min(X, C)$, $\delta = I(X \leq C)$, where $I(A)$ is the indicator function of any set A, $\delta = 0$ if X is right-censored by C and $\delta = 1$ otherwise. Moreover, assume that each individual or entity is characterized by a $\mathbb{R}^d$-valued covariate Z, independent of C and that the random vector $(T_1, \ldots, T_m, Z, C)$ is absolutely continuous with respect to the Lebesgue measure. For statistical applications, the observable random variables will be (Y, ξ, Z) where $\xi = \delta\eta$.

Denote by F_X (resp. G) the distribution function of the random variable X (resp. C), $S_X = 1 - F_X$ the survival function of X, λ_X (resp. Λ_X) its hazard (resp. cumulative hazard) function:

$$\lambda_X(t) = \lim_{\Delta \searrow 0} \frac{\mathbb{P}[t \leq X < t + \Delta | X > t]}{\Delta} = \frac{f_X(t)}{S_X(t)}$$

and

$$\Lambda_X(t) = \int_0^t \lambda_X(s) ds,$$

where f_X is the density function of X. Note that without specific assumptions, the joint or marginal distribution functions of the underlying failure times together with the previous hazard functions are not identifiable (Tsiatis, [22]). Nevertheless, if each individual is characterized by a 'sufficiently informative' set of covariates, these distribution functions are identifiable under some regularity conditions (Heckman and Honor, [15]). The problem of identifiability discussed in literature incite to concentrate no more on the distributions but on cause specific functions which are expressed in terms of observable functions of failure times:

– the cumulative incidence function

$$F_j(t) = \mathbb{P}[X \leq t, \eta = j],$$

– the cause-specific hazard rate function of type j

$$\lambda_j(t) = \lim_{\Delta \searrow 0} \frac{\mathbb{P}[t \leq X < t + \Delta t, \eta = j | X > t]}{\Delta} = \frac{f_j(t)}{S_X(t)},$$

where f_j is the subdensity function corresponding to F_j,

– the jth cause-specific cumulative hazard function

$$\Lambda_j(t) = \int_0^t \lambda_j(s) ds.$$

The incident functions are related by the following equations:

$$F_j(t) = \int_0^t f_j(s) ds = \int_0^t S_X(s) \lambda_j(s) ds = \int_0^t S_X(s) d\Lambda_j(s).$$

Further, we shall deal with conditional versions given $Z = z$ of the previous functions which will be denoted by $S_X(\cdot|z)$, $\lambda_X(\cdot|z)$, $\Lambda_X(\cdot|z)$, $F_j(\cdot|z)$, $f_j(\cdot|z)$ and $\Lambda_j(\cdot|z)$, and by independence of Z and C, $G(\cdot|z) = G$.

The parametric or nonparametric estimation of the previous underlying latent variables distributions has been considered in the literature. For example, Kwan and Singh [20] considered nonparametric estimates of the distribution function for every latent risk, assuming they are mutually independent. Fermanian [9], extending Heckman and Honor [15], considered a model involving nonparametric estimation of all unknown sub-distributions which together yield an estimator of the joint conditional distribution of the competing failure times. Geffray [10] considered the latent risks with independent censorship and established strong approximation results with statistical applications.

Most models make parametric assumptions on the joint distribution function of the failure times or assume their independence in order to avoid the non identifiability problem. When no such assumptions are made, the quantities usually estimated are the cause specific functions instead of the overall or latent distribution functions.

In this paper, we consider the problem of estimating the competing risks regression functions

$$r_j(z) = \mathbb{E}[\psi(X) I(\eta = j) | Z = z], \quad j = 1, \ldots, m, \tag{1}$$

based on n independent and identically distributed observations of (Y, ξ, Z), where ψ belongs to a class of measurable functions on $\mathbb{R}^+$ such that $\mathbb{E}[|\psi^p(X)|] < +\infty$, for $p = 1, 2$, without any parametric or independence assumption on competing lifetimes. For example, when $\psi(x) = x$, $\psi(x) = x^p$, or $\psi(x) = \psi_s(x) = I(x \leq s)$, $r_j(z)$ is the incident regression function, the p-th conditional moment, or the conditional distribution function of $X_j^* = XI(\eta = j)$ given $Z = z$: $r_j(z) = \mathbb{E}[X_j^* | Z = z]$, $r_j(z) = \mathbb{E}[X_j^{*p} | Z = z]$, or $r_j(z) = F_j(s|z) = \mathbb{E}[I(X_j^* \leq s) I(\eta = j) | Z = z]$ respectively.

The problem of estimating regression functions has been considered in the literature by several authors in non-censored as well as censored frameworks. In non-censored case, we may cite Nadaraya [19], Watson [25], Collomb [5], Beran [2], Greblicki et al. [13], Bosq and Lecoutre [3], Haerdle et al. [14], Carbonez et al. [4], Einmahl and Mason [8] Derzko and Deheuvels [7]. In censored models we can cite Dabrowska [6], Kohler and Math [18], Gneyou [12] and references therein.

Our paper is organized as follows. In Section 2, we give explicit kernel-type estimates of the competing risks regression functions conditional on $Z = z$. In Section 3 we state that for a given z, our estimator of $r_j(z)$ is consistent and fulfils a central limit theorem. In Section 4 some simulation results are given while some proofs are relegated to the appendix. Concluding remarks are given in Section 5.

2. Estimating regression functions

Let us define the following conditional distributions

$$\begin{aligned} H(t|z) &= \mathbb{P}[Y \le t|Z = z] \\ H_j(t|z) &= \mathbb{P}[Y \le t, \xi = j|Z = z], \; j = 1, \cdots, m \end{aligned}$$

Note that

$$\bar{H}(\cdot|z) = 1 - H(\cdot|z) = \bar{G}S_X(\cdot|z),$$

where $\bar{G}$ is the survival function of C. We have

$$\Lambda_j(t|z) = \int_0^t \lambda_j(s|z)ds = \int_0^t \frac{dF_j(s|z)}{S_X(s^-|z)} = \int_0^t \frac{dH_j(s|z)}{\bar{H}(s^-|z)}.$$

Note also that the competing risk regression functions $r_j(z)$ given in (1) can be written, provided the integral exists, in the form

$$\begin{aligned} r_j(z) &= \int_0^{\tau_z} \psi(t) f_j(x|z)dt \\ &= \int_0^{\tau_z} \psi(t) S_X(t|z) d\Lambda_j(t|z) \\ &= \int_0^{\tau_z} \frac{\psi(t) S_X(t|z)}{\bar{H}(t|z)} dH_j(t|z) \\ &= \int_0^{\tau_z} \frac{\psi(t)}{\bar{G}(t)} dH_j(t|z), \end{aligned} \tag{2}$$

where for technical reasons the interval of study is reduced to $[0, \tau_z]$ that will be specified later. Henceforth, the existence of the regression $r_j(z)$ holds since we assumed that $\mathbb{E}[|\psi(X)|] < +\infty$ in the previous section. A natural way to estimate $r_j(z)$ is to replace $\bar{G}$ and $H_j(\cdot|z)$ in (2) by their appropriate estimators. The survival function $\bar{G}$ is naturally estimated by the following product-limit estimators

$$\bar{G}_n(t) = \prod_{s \le t} (1 - \Delta\Lambda_n(s)),$$

where $\Delta\Lambda_n(t) = \Lambda_n(t) - \Lambda_n(t-)$ and

$$\Lambda_n(t) = \int_0^t \frac{dN(s)}{Y(s)},$$

where $N(s) = \sum_{i=1}^n I(Y_i \leq s; \xi_i = 0)$ and $Y(s) = \sum_{i=1}^n I(Y_i \geq s)$. Let $\boldsymbol{K}$ be a kernel function on $\mathbb{R}^d$, $(h_n)_{n\geq 0}$ be a sequence of positive real numbers tending to 0 as n tends to infinity (bandwidth) and set $\boldsymbol{K}_{h_n}(x) = h_n^{-d}\boldsymbol{K}(h_n^{-1}x)$. The conditional sub-distribution functions $H_j(\cdot|z)$ are estimated non-parametrically using kernel-type estimators with Nadaraya-Watson weights by

$$H_{jn}(t|z) = \frac{1}{nf_n(z)} \sum_{i=1}^n I(Y_i \leq t, \xi_i = j)\boldsymbol{K}_{h_n}(z - Z_i)$$

where

$$f_n(z) = \frac{1}{n}\sum_{i=1}^n \boldsymbol{K}_{h_n}(z - Z_i)$$

is the usual kernel density estimator of the marginal density function f of Z. The final estimator $\hat{r}_j(z)$ we propose is therefore

$$\begin{aligned} \hat{r}_{jn}(z) &= \int_0^{\tau_z} \frac{\psi(t)}{\bar{G}_n(t)} dH_{jn}(t|z) \\ &= \frac{1}{nf_n(z)} \sum_{i=1}^n \frac{\psi(Y_i)I(Y_i \leq \tau_z)I(\xi_i = j)\boldsymbol{K}_{h_n}(z - Z_i)}{\bar{G}_n(Y_i)}. \end{aligned} \tag{3}$$

In the next section we prove that under suitable conditions $\hat{r}_{jn}(z)$ is consistent and as $n \to +\infty$
$\sqrt{nh_n^d}(\hat{r}_{jn}(z) - r_j(z))$ converges in distribution to a centred normal distribution with variance consistently estimated by

$$\begin{aligned} \hat{\sigma}^2_{jn}(z) &= \frac{\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)}}{f_n(z)} \left(\hat{r}^2_{jn}(z) + 2\hat{r}_{jn}(z) \int_0^{\tau_z} \frac{\psi(t)}{\bar{G}^2_n(t)} H_{jn}(t|z) dK_{jn}(t|z) \right. \\ & \left. + \int_0^{\tau_z}\int_0^{\tau_z} \frac{\psi(s)\psi(t)}{\bar{G}^2_n(s)\bar{G}^2_n(t)} H_{jn}(s \wedge t|z) dK_{jn}(s|z) dK_{jn}(t|z) \right), \end{aligned} \tag{4}$$

where $K_{jn}(\cdot|z) = f_n(z)H_{jn}(\cdot|z)$.

Note that in practice τ_z is often taken equal to infinity, so that the indicator functions
$I(Y_i \leq \tau_z)$ can be deleted in formula (3). The choice of the bandwidth h_n is discussed in the simulation study. We note however that deleting the bias of $\hat{r}_{jn}(z)$ at the rate $\sqrt{nh_n^d}$ is not compatible with the usual bandwidth rate that allows to minimize the asymptotic mean square error of $\hat{r}_{jn}(z)$. Since we need to estimate carefully the density f we give priority to a bandwidth's choice that minimizes the asymptotic mean square error of f_n in our simulation study. In the remainder of the paper, $D[0,T]$ denotes the Skorohod space of all

right-continuous functions on $[0,T]$ with left limits, endowed with the metric induced by the supremum norm and $\ell^\infty[0,T]$ denotes the space of all bounded functions from the interval $[0,T]$ to $\mathbb{R}$ equipped with the supremum norm. The sequence X_n converges weakly to X is written as $X_n \rightsquigarrow X$.

3. Asymptotics

Let us consider the following assumptions.

(A). The upper bound τ_z of the interval of study satisfies $\bar{H}(\tau_z|z) > 0$, $G(\tau_z) > 0$ and $F_X(\tau_z) < 1$.
(B). The density function f is continuous at z and $f(z) > 0$.
(C). The application $s \mapsto H_j(t|s)$ is continuous at z, uniformly in $t \in [0, \tau_z]$.
(D). $\boldsymbol{K} = \varphi \circ p$, p being a polynomial and φ a positive bounded real function of bounded variation. The kernel function $\boldsymbol{K}$ has support in $[-1,1]^d$ and satisfies:

$$(i) \int_{\mathbb{R}^d} \boldsymbol{K}(s)ds = 1, \qquad (ii) \int_{\mathbb{R}^d} s\boldsymbol{K}(s)ds = 0.$$

(E). The bandwidth h_n satisfies $h_n = cn^{-\alpha}$ with $\alpha \in \left((5d)^{-1}, d^{-1}\right)$ and c is a constant.
(F). Functions f and $s \mapsto H_j(t|s)$ (for all $t \in [0,\tau_z]$) are twice continuously differentiable at z, and the second derivative of $s \mapsto H_j(t|s)f(s)$ is continuous at z, uniformly in $t \in [0, \tau_z]$.

The above assumptions are standard. Assumption A allows to obtain the uniform consistency and the weak convergence in $D[0,\tau_z]$ of the Kaplan-Meier estimator of G . Assumptions B and C are necessary to obtain consistency of kernel-type estimators since these estimators have an asymptotic bias that disappears if the target function is regular enough. The same problem holds when establishing a central limit theorem, but since the rate is $(nh_n^d)^{1/2}$ the conditions to make the asymptotic bias disappear are the stronger ones given in Assumptions F. The assumptions made on the kernel function $\boldsymbol{K}$ and the bandwidth are quite standard. However these assumptions on $\boldsymbol{K}$ and h_n could be less restrictive but it would involve more technicalities in our proofs. The bandwidth rates in Assumption E are simplified in order to satisfy:

$$(i) h_n \to 0, \qquad (ii) nh_n^d \to +\infty, \qquad (iii) nh_n^{5d} \to 0 \qquad \text{as } n \to +\infty$$

(i) and (ii) allows to obtain the consistency whereas (iii) is necessary to make the asymptotic bias disappear in the central limit theorem. In Gin and Guillou [11], Conditions (2.11) and (K_1), are given finest conditions (satisfied under Assumptions D and E) under which rates of strong uniform consistency results are obtained. Recall that

$$r_j(z) = \phi(f(z), \bar{G}, K_j(\cdot|z)) \equiv \frac{1}{f(z)} \int_0^{\tau_z} \frac{\psi(t)}{\bar{G}(t)} dK_j(t|z),$$

where we set $K_j(t|z) = H_j(t|z)f(z)$. Setting $K_{jn}(t|z) = H_{jn}(t|z)f_n(z)$ we have

$$\hat{r}_{jn}(z) = \phi(f_n(z), \bar{G}_n, K_{jn}(\cdot|z)).$$

Therefore the asymptotic behaviour of $(nh_n^d)^{1/2}\,(\hat{r}_{jn}(z) - r_j(z))$ can be expected by establishing the asymptotic behaviour of $(nh_n^d)^{1/2}\left(f_n(z) - f(z), \bar{G}_n - \bar{G}, K_{jn} - K_j\right)$ and using the functional delta-method with ϕ.

Lemma 1. *Under Assumptions A–E, we have as $n \to +\infty$*

(i). $f_n(z) \to f(z)$ *a.s.*
(ii). $\sup_{t\in[0,\tau_z]} |\bar{G}_n(t) - \bar{G}(t)| \to 0$ *a.s.*
(iii). $\sup_{t\in[0,\tau_z]} |K_{jn}(t|z) - K_j(t|z)| \to 0$ *a.s.*

Lemma 2. *Under Assumptions A–F, we have as $n \to +\infty$*

$$\sqrt{nh_n^d}\begin{pmatrix} f_n(z)-f(z) \\ \bar{G}_n - \bar{G} \\ K_{jn}(\cdot|z) - K_j(\cdot|z)\end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathcal{N}_z \\ 0 \\ \mathcal{G}_z \end{pmatrix}, \quad \text{in } \mathbb{R}\times(\ell^\infty[0,\tau_z])^2,$$

where $\mathcal{N}_z$ is a centred gaussian random variable and $\mathcal{G}_z$ a tight centred gaussian process. Moreover we have

$$\begin{aligned} \mathbb{E}[\mathcal{N}_z^2] &= f(z)\|\boldsymbol{K}\|^2_{L(\mathbb{R}^d)}, \\ \mathbb{E}[\mathcal{N}_z\mathcal{G}_z(t)] &= H_j(t|z)f(z)\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)}, \quad \textit{for all } t\in[0,\tau_z], \\ \mathbb{E}[\mathcal{G}_z(s)\mathcal{G}_z(t)] &= H_j(s\wedge t|z)f(z)\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)}, \quad \textit{for all } (s,t)\in[0,\tau_z]^2. \end{aligned}$$

The proof of this lemma mainly uses Theorem 19.28 in van der Vaart [24], it is given in the appendix. The next lemma, whose proof is also relegated to the appendix, gives the Hadamard derivative of ϕ at $(f(z), \bar{G}, H_j(\cdot|z))$.

Lemma 3. *The function ϕ is Hadamard-differentiable at $(f(z), \bar{G}, K_j(\cdot|z))$ with derivative*

$$\begin{aligned} \phi'_{(f(z),\bar{G},K_j(\cdot|z))}(h_1,h_2(\cdot),h_3(\cdot)) &= -\frac{h_1}{f(z)}\phi(f(z),\bar{G},K_j(\cdot|z)) \\ &+\frac{1}{f(z)}\left(\left[\frac{h_2\psi}{\bar{G}}\right]_0^{\tau_z} - \int_0^{\tau_z} h_2^-(t)d\left(\frac{\psi(t)}{\bar{G}(t)}\right) - \int_0^{\tau_z}\frac{\psi(t)}{\bar{G}^2(t)}h_3(t)dK_j(t|z)\right), \end{aligned}$$

where for any function ℓ, $[\ell]_0^{\tau_z} = \ell(\tau_z) - \ell(0)$ and $\ell^-(t) = \lim_{s\nearrow t}\ell(s)$.

Theorem 1. *Under conditions A–E, we have $\hat{r}_{jn}(z) \to r_j(z)$ almost surely as $n \to +\infty$.*

Proof. By Lemma 3 the function ϕ is continuous in $\mathbb{R}\times\ell^\infty[0,\tau_z]\times\ell^\infty[0,\tau_z]$ at point $(f(z),\bar{G},K_j(\cdot|z))$. Moreover, by Lemma 1 $(f_n(z),\bar{G}_n,K_{jn}(\cdot|z))$ converges almost surely to $(f(z),\bar{G},K_j(\cdot|z))$ in $\mathbb{R}\times\ell^\infty[0,\tau_z]\times\ell^\infty[0,\tau_z]$ as n tends to infinity. Hence by the continuous mapping theorem (see e.g. van der Vaart, [24], Theorem 18.11) $\phi(f_n(z),\bar{G}_n,K_{jn}(\cdot|z))$ converges almost surely to $\phi(f(z),\bar{G},K_j(\cdot|z))$ in $\mathbb{R}$ as $n\to+\infty$. □

Theorem 2. *Under conditions A–F, we have as n tends to infinity*

$$(nh_n^d)^{1/2}(\hat{r}_{jn}(z) \to r_j(z)) \rightsquigarrow \mathcal{N}\left(0,\sigma_j^2(z)\right),$$

where

$$\sigma_j^2(z) = \frac{\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)}}{f(z)} \left(r_j^2(z) + 2r_j(z) \int_0^{\tau_z} \frac{\psi(t)}{\bar{G}^2(t)} H_j(t|z) dK_j(t|z) + \int_0^{\tau_z}\int_0^{\tau_z} \frac{\psi(s)\psi(t)}{\bar{G}^2(s)\bar{G}^2(t)} H_j(s \wedge t|z) dK_j(s|z) dK_j(t|z) \right).$$

Moreover $\hat{\sigma}^2_{jn}(z)$ *defined by (4) converges almost surely to* $\sigma_j^2(z)$.

Proof. Lemmas 2 and 3 allow a straightforward application of the δ-method (Theorem 20.8 in van der Vaart, [24]). We obtain

$$(nh_n^d)^{1/2}(\hat{r}_{jn}(z) \to r_j(z)) \rightsquigarrow T_z = \phi'_{(f(z),\bar{G},K_j(\cdot|z))}(\mathcal{N}_z, 0, \mathcal{G}_z) = -\frac{1}{f(z)} \left(r_j(z)\mathcal{N}_z + \int_0^{\tau_z} \frac{\psi(t)\mathcal{G}_z(t)}{\bar{G}^2(t)} dK_j(t|z) \right).$$

T_z is gaussian as a linear form on a gaussian process and it is easy to check that $\mathbb{E}[T_z] = 0$ since both $\mathcal{N}_z$ and $\mathcal{G}_z$ are centred. It is also easy to calculate the variance $\sigma_j^2(z)$ of T_z since

$$\begin{aligned} \sigma_j^2(z) &= \frac{1}{f^2(z)} \mathbb{E}\left[\left(r_j(z)\mathcal{N}_z + \int_0^{\tau_z} \frac{\psi(t)\mathcal{G}_z(t)}{\bar{G}^2(t)} dK_j(t|z) \right)^2 \right] \\ &= \frac{1}{f^2(z)} \left(r_j^2(z)\mathbb{E}\left[\mathcal{N}_z^2\right] + 2r_j(z) \int_0^{\tau_z} \frac{\psi(t)}{\bar{G}^2(t)} \mathbb{E}[\mathcal{N}_z \mathcal{G}_z(t)] dK_j(t|z) \right. \\ &\quad \left. + \int_0^{\tau_z}\int_0^{\tau_z} \frac{\psi(s)\psi(t)}{\bar{G}^2(s)\bar{G}^2(t)} \mathbb{E}[\mathcal{G}_z(s)\mathcal{G}_z(t)] dK_j(s|z) dK_j(t|z) \right). \end{aligned}$$

The final formula of $\sigma_j^2(z)$ is obtained by replacing the expectations in the right hand side of the above equality by values provided in Lemma 2. Proving the strong consistency of $\hat{\sigma}_{jn}$ is again an application of the continuous mapping theorem, since $\hat{\sigma}_{jn}$ can be written as a continuous function of $(f_n(z), \bar{G}_n, K_{jn})$ which by Lemma 1 converges almost surely to $(f(z), \bar{G}, K_j))$ in $\mathbb{R} \times \ell^\infty[0, \tau_z] \times \ell^\infty[0, \tau_z]$. □

4. Simulation study

We consider that (T_1, T_2), conditional on $Z = z$, has an exponential bivariate distribution defined by the joint survival function

$$S(t_1, t_2|z) = \exp(-e^z(\lambda_1 t_1 + \lambda_2 t_2 + \theta t_1 t_2)),$$

for $t_1, t_2 \geq 0$ and $0 \leq \theta < \lambda_1\lambda_2$. It is easy to check that both T_1 and T_2 are marginally exponentially distributed with respective parameters $e^z\lambda_1$ and $e^z\lambda_2$. The joint density f_{12} of (T_1, T_2), conditional on $Z = z$, is therefore defined by

$$f_{12}(t_1, t_2|z) = e^z \left(e^z(\lambda_1 + \theta t_2)(\lambda_2 + \theta t_1) - \theta\right) \exp(-e^z(\lambda_1 t_1 + \lambda_2 t_2 + \theta t_1 t_2)),$$

for $t_1, t_2 \geq 0$. We show that conditional on $(T_1, Z) = (t_1, z)$ the distribution function of T_2 is given by

$$F_{2|1}(t_2|t_1, z) = 1 - (1 + \theta t/\lambda_1) \exp(-(\lambda_2 + \theta t_2 t_1)t_2),$$

for $t_2 \geq 0$. Simulated data are obtained by the inversion method: first Z is simulated following a standard normal distribution, then we simulate T_1 conditional on z and finally we simulate T_2 conditional on (T_1, Z) (by inverting numerically $F_{2|1}(\cdot|t_1, z)$). The censoring variable C is exponentially distributed with rate λ_C. The sample size is denoted by n and other parameters are set to

$$\lambda_1 = 0.1, \quad \lambda_2 = 0.15, \quad \lambda_C = 0.35, \quad \text{and} \quad \theta = 0.01.$$

With such parameters, the expected number of failure of type 1 and 2 are approximately equal to 42% and 38% respectively whereas the censoring rate is about 20%. $\boldsymbol{K}$ is the Epanechnikov kernel and the bandwidth is chosen to be equal to

$$h_n = \hat{\sigma}_Z \left(\frac{4}{3n}\right)^{1/5}$$

where $\hat{\sigma}_Z$ is the standard error of the Z_i's.

We show that the sub-density functions f_1 and f_2, conditional on $Z = z$, are defined by:

$$f_i(t|z) = e^z(\lambda_i + \theta t)\exp(-e^z(\lambda_1 + \lambda_2 + \theta t)t),$$

for $t \geq 0$ and $i = 1, 2$. Then we have for $j = 1, 2$

$$r_j(z) = \int_{\mathbb{R}^+} \psi(t) f_j(t|z)dt \equiv \mathbb{E}[\psi(X)I(\eta = j)].$$

The above quantities are calculated numerically and compared with $\hat{r}_{jn}(z)$ for various integrable functions ψ and for various values of z. In Fig. 1 for $\psi(t) = t$, we compare for several values of n, $z \mapsto r_1(z) = \mathbb{E}(T_1 I(\eta = 1)|z)$ $(z \in [0, 2])$ with its estimate $z \mapsto \hat{r}_{1n}(z)$. We can check that the expected consistency property is satisfied. It is also interesting to see in Tab. 1 that our estimators are consistent and that both the bias and the standard deviations of our estimators tend to 0 as the sample size increases.

z	0	1	2
$\mathbb{E}_1[T_1\|z]$	1.399	0.556	0.212
$n = 200$	1.420 (0.526)	0.575 (0.224)	0.219 (0.194)
$n = 500$	1.439 (0.363)	0.573 (0.161)	0.223 (0.131)
$n = 1000$	1.380 (0.270)	0.568 (0.121)	0.221 (0.101)

Table 1. Estimation of $\mathbb{E}(T_1 I(\eta = 1)|z)$ for $z \in \{0, 1, 2\}$: mean and standard deviation (within parenthesis) of $N = 1000$ estimates for various sample sizes n.

As we said in the introduction, taking a family of functions $\{\psi_s; s \geq 0\}$ such that $\psi_s(x) = I(x \leq s)$ we have $\mathbb{E}[\psi_s(X)I(\eta = j)|z] = F_j(s|z)$. Then, it follows by (3) that

$$\frac{1}{f_n(z)} \sum_{\{i; Y_i \leq s\}} \frac{I(\xi_i = j)\boldsymbol{K}_{h_n}(z - Z_i)}{\bar{G}_n(Y_i)}$$

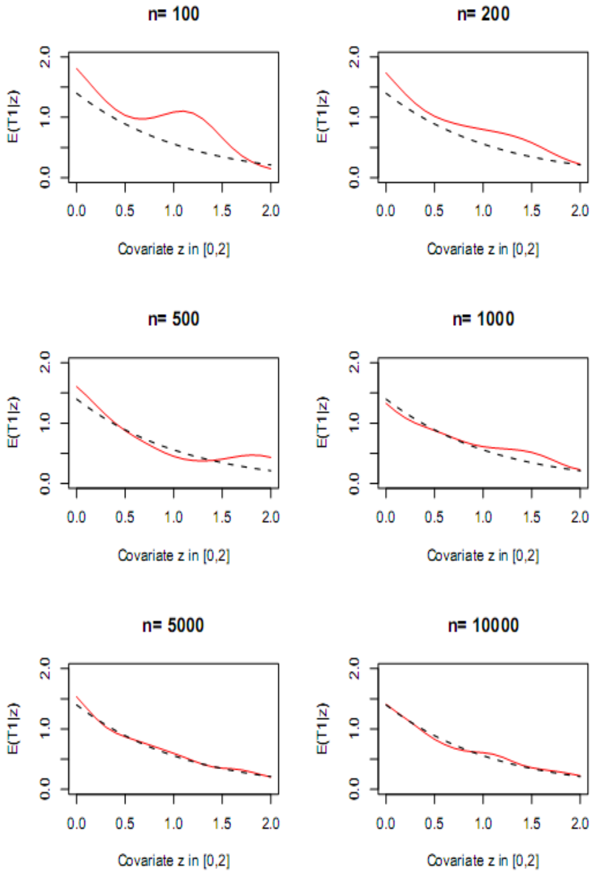

Fig. 1. Comparison of the conditional expectation $z \mapsto \mathbb{E}(T_1 I(\eta = 1)|z)$ (dashed line) and its estimates (solid line) for various values of n.

is a consistent estimator of $F_j(s|z)$. In Fig. 2–4 are given the estimates of $F_1(\cdot|z)$ for $z \in \{0, 1, 2\}$ and $n \in \{200, 500, 1000, 5000\}$. We can check that the the estimator behaves well for large sample size. For small or moderate sample size estimates can be more or less precise when t is large especially when z is such that they are very few observations satisfying $Z_i \in [z - h_n, z + h_n]$; for $z = 2$ and $n = 1000$ there is (in mean) about 30 observations accounted to estimate $F_1(\cdot|2)$. This explains the poor performances of our estimators in regions where the density f of Z has small values.

5. Concluding remarks

The estimation method we proposed is easy to implement and behaves quite well for moderate sample size. In some applications the censoring mechanism may depend on some covariates. If the later is true we need to replace $\bar{G}$ by $\bar{G}(\cdot|z)$ with an appropriate kernel type estimator, but in this case $(nh_n^d)^{1/2}(\bar{G}_n(\cdot|z) - \bar{G}(\cdot|z))$ is no longer asymptotically negligible and its asymptotic behaviour plays a part in the asymptotic behaviour of $\hat{r}_{jn}(z)$. This

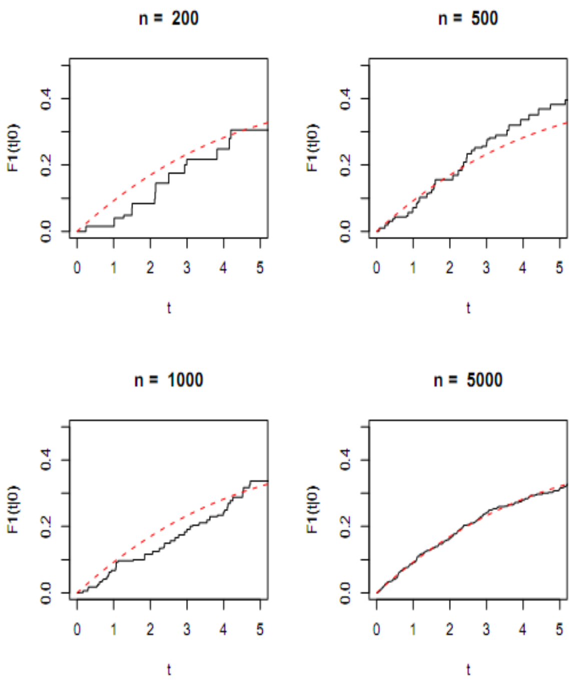

Fig. 2. Comparison of $t \mapsto F_1(t|0)$ (dashed) and its estimates (solid) for various values of n.

case will be considered in further work together with convergence rate for uniform (in z) consistency results.

Acknowledgements

The authors are grateful to the referee for a careful reading of the manuscript and making some useful suggestions. They thank Simplice Dossou-Gbété and Jean Bretagnolle for several conversations and comments on this article.

Appendix A: Proofs of technical results

Lemma 4. *Let $\ell : \mathbb{R}^d \to \mathbb{R}$ be a function continuous at z. Then, under Conditions D and E we have*

$$\int_{\mathbb{R}^d} \boldsymbol{K}_{h_n}(z-s)\ell(s)ds \to \ell(z), \quad \text{as } n \to +\infty.$$

Proof. Remark that

$$\begin{aligned}\left|\int_{\mathbb{R}^d} \boldsymbol{K}_{h_n}(z-s)\ell(s)ds - \ell(z)\right| &= \left|\int_{\mathbb{R}^d} \boldsymbol{K}(s)\left(\ell(z+h_n s) - \ell(z)\right) ds\right| \\ &\leq \sup_{\|s\|_\infty \leq 1} |\ell(z+h_n s) - \ell(z)| \sup_{s\in\mathbb{R}^d} |\boldsymbol{K}(s)|\end{aligned}$$

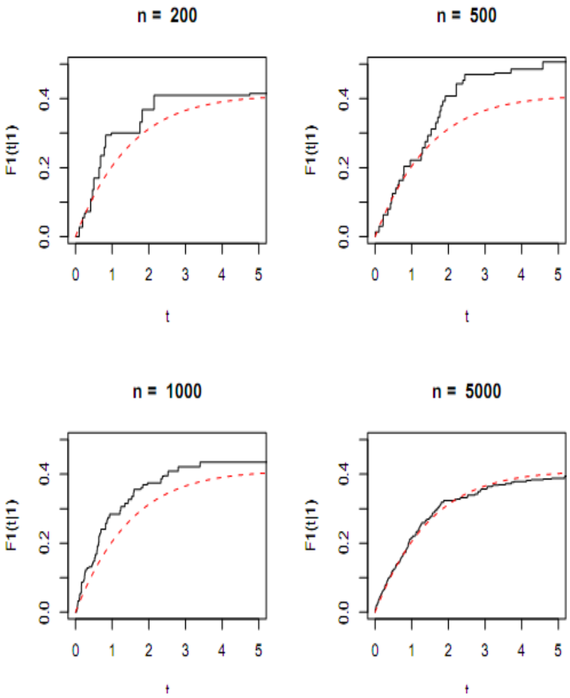

Fig. 3. Comparison of $t \mapsto F_1(t|1)$ (dashed) and its estimates (solid) for various values of n.

$$\to 0, as\ n \to +\infty$$

since ℓ is continuous at z, $h_n \to 0$ and $\boldsymbol{K}$ is bounded with support included in $[-1,1]^d$. $\square$

Lemma 5. *Let $\ell : \mathbb{R}^d \to \mathbb{R}$ be a function twice continuously differentiable at z. Then, under Conditions D and E we have*

$$\left| \int_{\mathbb{R}^d} \boldsymbol{K}_{h_n}(z-s)\ell(s)ds - \ell(z) \right| = O(h_n^{2d}), \quad as\ n \to +\infty.$$

Proof. Remark that under assumptions D (ii) and F, since $\|u_n(z) - z\| \to 0$ as $n \to \infty$, we have for n large enough

$$\begin{aligned}
\left| \int_{\mathbb{R}^d} \boldsymbol{K}_{h_n}(z-s)\ell(s)ds - \ell(z) \right| &= \left| \int_{\mathbb{R}^d} \boldsymbol{K}(s) \left(\ell(z + h_n s) - \ell(z)\right) ds \right| \\
&\leq \left| \int_{\mathbb{R}^d} \boldsymbol{K}(s) \left(h_n^d \frac{\partial \ell}{\partial s^t}(z)s + \frac{h_n^{2d}}{2} s^t \frac{\partial^2 \ell}{\partial s \partial s^t}(u_n(z))s \right) ds \right|, \\
&\leq h_n^{2d} C_0 \int_{\mathbb{R}^d} \|s\|^2 \boldsymbol{K}(s) ds,
\end{aligned}$$

where C_0 is a constant. The result follows from the last inequality. $\square$

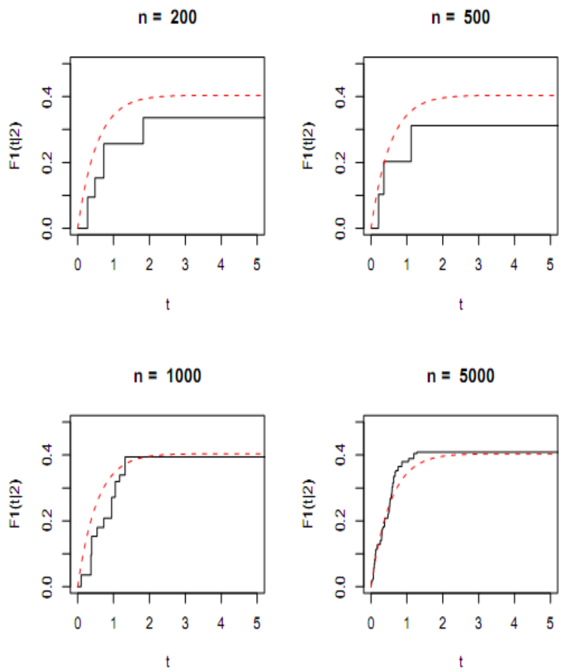

Fig. 4. Comparison of $t \mapsto F_1(t|2)$ (dashed) and its estimates (solid) for various values of n.

Proof of Lemma 1

Proof of (i). Under assumptions B, D and E and applying Lemma 4 with $\ell = f$, we have for n large enough $\mathbb{E}[f_n(z)] = f(z) + o(1)$. Moreover, by assumptions D and E, Conditions (2.11) and (K_1) of Gin and Guillou (2002) are fulfilled. Applying their Theorem 2.3 we obtain the strong convergence of $f_n(z) - \mathbb{E}[f_n(z)]$ to 0 as $n \to +\infty$. It follows that $f_n(z) - f(z)$ converges almost surely to 0 as $n \to +\infty$.

Proof of (ii). By assumption A we have $G(\tau_z) > 0$ and $F_X(\tau_z) < 1$. Moreover C and X are independent and their cumulative distribution functions do not have jumps in common. Hence this is a consequence of Corollary 1.3 of Stute and Wang (1993).

Proof of (iii). For fixed $z \in \mathbb{R}^d$, let $f_{n,t}$ be defined by

$$f_{n,t}(y,x,s) = I(y \le t, x = j)\boldsymbol{K}_{h_n}(z-s), \quad j \in \{1, \cdots, m\}, \; s \in \mathbb{R}^d$$

and

$$\mathcal{F}_n = \{f_{n,t}; t \in [0, \tau_z]\}.$$

Let us consider the following brackets $[f_{n,t_{i-1}}, f_{n,t_i^-}]$ with

$$f_{n,t^-}(y,x,z) = I(y < t, x = j)\boldsymbol{K}_{h_n}(z-s), \quad j \in \{1, \cdots, m\}, \; s \in \mathbb{R}^d$$

and $H_j(t_i^-|z) - H_j(t_{i-1}|z) < \varepsilon$. We show (see the proof of Lemma 2) that for n large enough we have

$$\mathbb{E}\left[f_{n,t_i}(Y,\xi,Z) - f_{n,t_{i-1}^-}(Y,\xi,Z)\right] \le 2\varepsilon f(z),$$

then

$$N_{[\,]}\left(2\varepsilon f(z),\mathcal{F}_n,L^1(\mathbb{P})\right)\leq\frac{2}{\varepsilon}.$$

Consequently, following the lines of the proof of Theorem 2.4.1 in van der Vaart and Wellner (1996) we show that $\mathcal{F}_n$ is $\mathbb{P}$–Glivenko-Cantelli. Finally, we have

$$\begin{aligned}K_{jn}(t|z)-K_j(t|z)&=\frac{1}{n}\sum_{i=1}^{n}\left(f_{n,t}(Y_i,\xi_i,Z_i)-\mathbb{E}\left[f_{n,t}(Y_i,\xi_i,Z_i)\right]\right)\\&\quad+\mathbb{E}\left[f_{n,t}(Y,\xi,Z)\right]-K_j(t|z),\end{aligned}$$

with

$$\begin{aligned}&|\mathbb{E}\left[f_{n,t}(Y,\xi,Z)\right]-K_j(t|z)|\\&\qquad\leq\int_{\mathbb{R}^d}|H_j(t|s)f(s)-H_j(t|z)f(z)|\,\boldsymbol{K}_{h_n}(z-s)ds\\&\qquad\leq\sup_{t\in[0,\tau_z]}\sup_{\|u\|_\infty\leq h_n}|H_j(t|z+u)f(z+u)-H_j(t|z)f(z)|\int_{\mathbb{R}^d}\boldsymbol{K}(u)du\\&\qquad\leq\sup_{t\in[0,\tau_z]}\sup_{\|u\|_\infty\leq h_n}|H_j(t|z+u)f(z+u)-H_j(t|z)f(z)|=o(1),\end{aligned}$$

by assumptions B–E. Hence we obtain

$$\sup_{t\in[0,\tau_z]}|\mathbb{E}\left[f_{n,t}(Y,\xi,Z)\right]-K_j(t|z)|\to 0\ as\ n\to+\infty.$$

This together with the fact that $\mathcal{F}_n$ is $\mathbb{P}$–Glivenko-Cantelli lead to the wanted result. $\square$

Proof of Lemma 2

Since the first component does not depend on t and the weak limit of the second component is degenerated, we obtain the weak convergence of the whole vector by proving that each component of the vector converges weakly.

Convergence of the first component. Applying the Lindeberg-Feller theorem (see e.g. van der Vaart, 1998) to $\sqrt{nh_n^d}(f_n(z)-\mathbb{E}[f_n(z)])$ we derive the weak limit $\mathcal{N}_z$. To prove the result we need in addition to show that $\sqrt{nh_n^d}(\mathbb{E}[f_n(z)]-f(z))\to 0\ as\ n\to+\infty$, which is obtained by applying Lemma 5 with $\ell=f$ under assumptions B–F. Finally we obtain the variance of $\mathcal{N}_z$ by calculating, using Lemma 4, the limit of $nh_n^d\mathrm{Var}(f_n(z))$.

Convergence of the second component. Under Assumption A, conditions on G and independence conditions on $(T_1,\ldots,T_m,Z,C)$ we have (see Andersen et al., 1993)

$$\sqrt{n}\left(\bar{G}_n-\bar{G}\right)\rightsquigarrow\mathcal{B},\quad\text{in } D[0,\tau_z],\ as\ n\to+\infty$$

where $\mathcal{B}$ is a centred gaussian process on $[0,\tau_z]$. By the continuous mapping theorem we have for n large enough $\sup_{t\in[0,\tau_z]}\left|\sqrt{n}\left(\bar{G}_n(t)-\bar{G}(t)\right)\right|=O_P(1)$, and then

$$\sup_{t\in[0,\tau_z]}\left|\sqrt{nh_n^d}\left(\bar{G}_n(t)-\bar{G}(t)\right)\right|=O_P(h_n^{d/2})=o_P(1),$$

since by Assumption E we have $h_n \searrow 0$. It follows that $\sqrt{nh_n^d}\left(\bar{G}_n - \bar{G}\right) \rightsquigarrow 0 \; as \; n \to +\infty$.

Convergence of the third component. For fixed $z \in \mathbb{R}^d$, let $f_{n,t}$ be defined by

$$f_{n,t}(y,x,s) = \frac{1}{h_n^{d/2}} I(y \le t, x = j) \boldsymbol{K}\left(\frac{z-s}{h_n}\right), \quad j \in \{1, \cdots, m\},\; s \in \mathbb{R}^d$$

and

$$\mathcal{F}_n = \{f_{n,t}; t \in [0, \tau_z]\}.$$

Since $\boldsymbol{K}$ is positive we have

$$\sup_{t \in [0,\tau_z]} |f_{n,t}| \le F_n = \frac{1}{h_n^{d/2}} \boldsymbol{K}\left(\frac{z-s}{h_n}\right).$$

Moreover since by Assumption E $h_n \to 0$ as $n \to +\infty$ and since by Assumption B f is continuous at z, applying Lemma 4 with kernel $\boldsymbol{K}^2/\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)}$ we have

$$\begin{aligned}
\mathbb{E}[F_n^2] &= \int_{\mathbb{R}^d} \frac{1}{h_n^d} \boldsymbol{K}^2\left(\frac{z-s}{h_n}\right) f(s)ds \\
&= \int_{\mathbb{R}^d} \frac{1}{h_n^d} \boldsymbol{K}^2(s) f(z + h_n s)ds \\
&= \|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)} f(z) + o(1), n \to +\infty.
\end{aligned}$$

Now, since $\boldsymbol{K}$ is bounded and by Assumption E $nh_n^d \to +\infty$, we have for $\varepsilon > 0$

$$\begin{aligned}
\mathbb{E}[F_n^2 I(F_n > \sqrt{n}\varepsilon] = \int_{\mathbb{R}^d} \frac{1}{h_n^d} \boldsymbol{K}^2\left(\frac{z-s}{h_n}\right) f(s) I\left(\boldsymbol{K}\left(\frac{z-s}{h_n}\right) > \sqrt{nh_n^d}\varepsilon\right) ds \\
\to 0,\; as\; n \to +\infty.
\end{aligned}$$

Let us consider the pseudo-distance ρ_j defined by

$$\rho_j(s,t) = |H_j(s|z) - H_j(t|z)|.$$

For $0 \le s \le t \le \tau_z$ and $\rho_j(s,t) \le \rho_n$, we have by Assumption F and the fact that $h_n \to 0$ under E

$$\begin{aligned}
&\mathbb{E}\left[(f_{n,t}(Y,\xi,Z) - f_{n,s}(Y,\xi,Z))^2\right] \\
&\quad = \int_{\mathbb{R}^d} \frac{1}{h_n^d} \mathbb{E}\left[I(Y \in (s,t]; \xi = j)\right] \boldsymbol{K}^2\left(\frac{z-u}{h_n}\right) f(u)du \\
&\quad = \int_{\mathbb{R}^d} \left(H_j(t|u) - H_j(s|u)\right) \frac{1}{h_n^d} \boldsymbol{K}^2\left(\frac{z-u}{h_n}\right) f(u)du \\
&\quad \le \left(2 \sup_{\|u\|_\infty \le 1} \sup_{t \in [0,\tau_z]} |H_j(t|z + h_n u) - H_j(t|z)| + \rho_j(s,t)\right) \\
&\qquad \int_{\mathbb{R}^d} \boldsymbol{K}^2(u) f(z + h_n u)du \\
&\quad \le (o(1) + \rho_n) O(1) \to 0, as\; n \to +\infty
\end{aligned}$$

if $\rho_n \to 0$. Now let us show that the integral entropy $J_{[\,]}(\delta_n, \mathcal{F}_n, L^2(\mathbb{P})) \to 0$ as $\delta_n \to 0$. Let us consider the following brackets

$$[f_{n,t_{i-1}}, f_{n,t_i^-}]$$

with

$$f_{n,t^-}(y,x,z) = I(y < t, x = j)\frac{1}{h_n^{d/2}}\boldsymbol{K}\left(\frac{z-s}{h_n}\right),$$

and $H_j(t_i^-|z) - H_j(t_{i-1}|z) < \varepsilon$. After some calculations similar to the above ones we obtain

$$\mathbb{E}\left[\left(f_{n,t_i^-}(Y,\xi,Z) - f_{n,t_{i-1}}(Y,\xi,Z)\right)^2\right] \leq 2\varepsilon f(z)\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)},$$

for n large enough. Then it is straightforward that for n large enough

$$N_{[\,]}(2\varepsilon f(z)\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)}, \mathcal{F}_n, L^2(\mathbb{P})) \leq \frac{2}{\varepsilon},$$

which leads to

$$N_{[\,]}(\varepsilon, \mathcal{F}_n, L^2(\mathbb{P})) \leq \frac{4f(z)\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)}}{\varepsilon}.$$

Finally, because of the above upper bound for $N_{[\,]}(\varepsilon, \mathcal{F}_n, L^2(\mathbb{P}))$, we have

$$J_{[\,]}(\delta_n, \mathcal{F}_n, L^2(\mathbb{P})) = \int_0^{\delta_n} \sqrt{\log N_{[\,]}(\varepsilon, \mathcal{F}_n, L^2(\mathbb{P}))}d\varepsilon \to 0,$$

as $n \to +\infty$. All assumptions of Theorem 19.28 (van der Vaart, 1998) being satisfied we obtain that the empirical process

$$\mathbb{G}_{jn}(t) = \frac{1}{\sqrt{n}}\left(\sum_{i=1}^n f_{n,t}(Y_i,\xi_i,Z_i) - \mathbb{E}f_{n,t}(Y_i,\xi_i,Z_i)\right)$$

satisfies

$$\mathbb{G}_{jn} \rightsquigarrow \mathcal{G}_j \quad as\ n \to +\infty$$

where $\mathcal{G}_j$ is a tight centred gaussian process on $[0,\tau_z]$ with correlation function defined for $s,t \in [0,\tau_z]$ by

$$\mathbb{E}\left[\mathcal{G}_j(t)\mathcal{G}_j(s)\right] = \lim_{n\to\infty} \mathbb{E}[f_{n,t}f_{n,s}] - \mathbb{E}[f_{n,t}]\mathbb{E}[f_{n,s}].$$

By using repeatedly Lemmas 4 and 5 we have as $n \to +\infty$

$$\mathbb{E}[f_{n,s}f_{n,t}] = \int_{\mathbb{R}^d} H_j(s\wedge t|u)\frac{1}{h_n^d}\boldsymbol{K}^2\left(\frac{z-u}{h_n}\right)f(u)du \to H_j(s\wedge t|z)f(z)\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)}.$$

It is easy to check that $\mathbb{E}[f_{n,t}(Y,\xi,Z)] = O(h_n^{d/2}) \to 0\ as\ n \to +\infty$. We have

$$\mathbb{E}[\mathcal{G}_j(s)\mathcal{G}_j(t)] = H_j(s\wedge t|z)f(z)\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)}.$$

It remains to study the bias term. Following the lines of Lemma 5 proof we have for n large enough

$$\begin{aligned}
&\sqrt{nh_n^d}\,|\mathbb{E}\left[I(Y \le t; \xi = j)\boldsymbol{K}_{h_n}(z-Z)\right] - H_j(t|z)f(z)| \\
&\quad = \sqrt{nh_n^d}\left|\int_{\mathbb{R}^d} H_j(t|s)f(s)\boldsymbol{K}_{h_n}(z-s)ds - \int_{\mathbb{R}^d} H_j(t|z)f(z)\boldsymbol{K}_{h_n}(s)ds\right| \\
&\quad \le \sqrt{nh_n^d}\left|\int_{\mathbb{R}^d} \boldsymbol{K}(u)\left(h_n^d \frac{\partial}{\partial s^t} H_j(t|s)f(s)_{|u=z} u \; + \right.\right. \\
&\qquad\qquad \left.\left. \frac{h_n^{2d}}{2} u^t \frac{\partial^2}{\partial s \partial s^t} H_j(t|s)f(s)_{|u_n(z)} u\right) du\right| \\
&\quad \le \sqrt{nh_n^{5d}} C_0 \int_{\mathbb{R}^d} \|u\|^2 \boldsymbol{K}(u)du = o(1),
\end{aligned}$$

where $\|u_n(z) - z\| \to 0$ as $n \to \infty$, and where C_0 does not depend on t since the second derivative of $s \mapsto H_j(t|s)f(s)$ is continuous at z uniformly in $t \in [0, \tau_z]$ by Assumption F. Thus the bias is asymptotically negligible since $nh_n^{5d} \to 0 as\ n \to +\infty$ by Assumption E.

Correlation between the first and third components. By Lemma 4 we have for n large enough $\mathbb{E}[f_n(z)] = f(z) + o(1)$ and $\mathbb{E}[K_{jn}(t|z)] = K_j(t|z) + o(1)$, then we can write

$$\begin{aligned}
&nh_n^d \mathbb{E}\left\{(f_n(z) - \mathbb{E}[f_n(z)])\,(K_{jn}(t|z) - \mathbb{E}[K_{jn}(t|z)])\right\} \\
&\quad = \mathbb{E}\left[I(\overset{2}{Y} \le t, \xi = j)\boldsymbol{K}_{h_n}(z-Z)\right] - h_n^d \mathbb{E}[f_n(z)]\mathbb{E}[K_{jn}(t|z)] \\
&\quad = \int_{\mathbb{R}^d} H_j(t|s)f(s)\frac{1}{h_n^d}\boldsymbol{K}^2\left(\frac{z-s}{h_n}\right)ds - h_n^d\left\{f(z) + o(1)\right\}\left\{K_j(t|z) + o(1)\right\} \\
&\quad = K_j(t|z)\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)} + o(1) + O(h_n^d),
\end{aligned}$$

where the last equality holds applying Lemma 4 with kernel $\boldsymbol{K}^2/\|\boldsymbol{K}\|^2_{L^2(\mathbb{R}^d)}$ instead of $\boldsymbol{K}$. □

Proof of Lemma 3

Recall that $\phi : (0, +\infty) \times D[0, \tau_z] \times BV_c[0, \tau_z] \to \mathbb{R}$ with

$$\phi(x, u, v) = \frac{1}{x}\int_0^{\tau_z} \frac{\psi(s)}{u(s)} dv(s),$$

where $BV_c[0, \tau_z]$ is the set of real-valued functions defined on $[0, \tau_z]$ with variation bounded by c. Let us consider as $t \downarrow 0$, $h_{1t} \to h_1$ in $\mathbb{R}$, and $h_{2t} \to h_2$, $h_{3t} \to h_3$ in $D[0, \tau_z]$ with $x_t = x + th_{1t}$, $u_t = u + th_{2t}$ and $v_t = v + th_{3t}$ such that $v_t \in BV_c[0, \tau_z]$. We need to calculate the limit for $t \downarrow 0$ of

$$\begin{aligned}
\frac{1}{t}\left(\phi(x_t, u_t, v_t) - \phi(x, u, v)\right) &= -\frac{h_{1t}}{x_t}\phi(x, u, v) \\
&\quad + \frac{1}{x_t}\frac{1}{t}\left[\int_0^{\tau_z} \frac{\psi(s)}{u_t(s)} dv_t(s) - \int_0^{\tau_z} \frac{\psi(s)}{u(s)} dv(s)\right].
\end{aligned}$$

It is obvious that the first term in the right hand side of the above equality tends to $-h_1\phi(x,u,v)/x$ whereas the limit of the second term is obtained following the lines of the proof of Theorem 20.10 in van der Vaart (1998) and is equal to

$$\frac{1}{x}\left(\left[\frac{h_2(s)\psi(s)}{u(s)}\right]_0^{\tau_z} - \int_0^{\tau_z} h_2^-(s)d\left(\frac{\psi(s)}{u(s)}\right) - \int_0^{\tau_z}\frac{\psi(s)h_3(s)}{u^2(s)}dv(s)\right).$$

Replacing (x,u,v) by $(f(z),\bar{G},K_j(\cdot|z))$ in the above limits gives the Hadamard derivative.
□

References

[1] Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N., 1993. *Statistical Models Based on Counting Processes.* Springer, New York.

[2] Beran, R., 1981. Nonparametric regression with randomly censored data. Technical report, *Univ. Carolina, Berkeley.*

[3] Bosq, D. and Lecoutre, J. P., 1987. *Thorie de l'Estimation Fonctionnelle.* Economica, Paris.

[4] Carbonez, A., Gyrfi, L. and van der Meulen, E. C., 1995. Partition-estimates of a regression function under random censoring. *Statist. Decisions*, **13**, 21-37.

[5] Collomb, G., 1981. Estimation non-paramtrique de la rgression: revue bibliographique. *Inst. Statist. Rev.*, **49**, 75-93.

[6] Dabrowska, D. M., 1987. Nonparametric regression with censored survival data. *Scand. J. Statist.*, **14**, 181-197.

[7] Derzko, G. and Deheuvels, P., 2002. Estimation non-paramtrique de la rgression dichotomique - Application biomdicale. *C.R. Acad. Sci. Paris, Ser. I*, **334**, 59-63.

[8] Einmahl, U. and Mason, D. M., 2000. An empirical processes approach to the uniform consistency of kernel-type function estimators. *J. Theoret. Probab.*, **13**, 1-37.

[9] Fermanian, J-D., 2003. Nonparametric estimation of competing risks models with covariates. *J. Multivariate Anal.*, **85**, 156–191.

[10] Geffray, S., 2008. Strong approximations for dependent competing risks with independent censoring. *Test*, to appear.

[11] Gin, E. and Guillou, A., 2002. Rates of strong uniform consistency for multivariate kernel density estimators. *Ann. I. H. Poincar*, **38**, 907-921.

[12] Gneyou, K. E., 2005. Vitesse de convergenge de certains estimateurs de Kaplan-Meier de la rgression. *Afrika Statistika*, **1**, 77-92.

[13] Greblicki, W., Krzyzl, A. and Pawlak, M., 1984. Distribution-free pointwise consistency of kernel regression estimate. *Ann. Statist.*, **12**, 1570-1575.

[14] Haerdle, W., Janssen, P. and Serfling, R., 1988. Strong uniform consistency rates for estimators of conditional functionals. *Ann. Statist.*, **16**, 1428-1449.

[15] Heckman, J. J. and Honor, B., 1989. The identifiability of competing risks model. *Biometrika*, **76**, 325-330.

[16] Kalbfleisch, J. D. and Prentice, R. L., 1980. *The Statistical Analysis of Failure Time Data*, Wiley, New York.

[17] Kaplan, E. K. and Meier, P. 1958. Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc*, **53**, 457-481.

[18] Kohler, M. and Math, K., 2002. Prediction from randomly right censored data. *J. Multivariate Anal.*, **1**, 73-100.

[19] Nadaraya, E. A. 1964. On estimating regression. *Theo. Prob. Appl.*, **9**, 141-142.

[20] Kwan, P. H. and Singh, H. 2001., On nonparametric estimation of the survival function with competing risks. *Scand. J. Statist*, **28**, 715-724.

[21] Stute W. and Wang J-L., 1993. The strong law under random censorship. *Ann. Statist.*, **21**, 1591-1607.

[22] Tsiatis, A., 1975. A nonidentifiability aspect of the problem of competing risks. *Proc. Natl. Acad. Sci. USA*, **72**, 20-22.

[23] van der Vaart, A. W. and Wellner, J., 1996. *Weak convergence and empirical processes with applications to statistics*, Springer-Verlag, New York.

[24] van der Vaart, A. W., 1998. *Asymptotic Statistics*, Cambridge University Press, New York.

[25] Watson, G. S., 1964. Smooth regression analysis. *Sankhy Ser. A*, **26**, 359-372.