



HAL
open science

Uniform convergence of nonparametric regressions in competing risks model with right censoring

Laurent Bordes, Kossi Essona Gneyou

► **To cite this version:**

Laurent Bordes, Kossi Essona Gneyou. Uniform convergence of nonparametric regressions in competing risks model with right censoring. *Statistics and Probability Letters*, 2011, 81 (11), pp.1654. 10.1016/j.spl.2011.06.008 . hal-00789865

HAL Id: hal-00789865

<https://hal.science/hal-00789865>

Submitted on 19 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

Uniform convergence of nonparametric regressions in competing risks model with right censoring

Laurent Bordes, Kossi Essona Gneyou

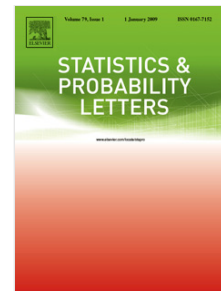
PII: S0167-7152(11)00210-0
DOI: [10.1016/j.spl.2011.06.008](https://doi.org/10.1016/j.spl.2011.06.008)
Reference: STAPRO 6022

To appear in: *Statistics and Probability Letters*

Received date: 13 October 2010
Revised date: 8 June 2011
Accepted date: 9 June 2011

Please cite this article as: Bordes, L., Gneyou, K.E., Uniform convergence of nonparametric regressions in competing risks model with right censoring. *Statistics and Probability Letters* (2011), doi:10.1016/j.spl.2011.06.008

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Uniform convergence of nonparametric regressions in competing risks model with right censoring

Laurent BORDES

Laboratoire de Mathématiques et de leurs Applications - UMR CNRS 5142

Université de Pau et des Pays de l'Adour -France

Kossi Essona GNEYOU

Université de Lomé -Département de Mathématiques BP 1515 Lomé -Togo

Abstract

We consider, in the presence of covariates, non independent competing risks that are subject to right censoring. We define a nonparametric estimator of the incident regression function through the generalized product-limit estimator of the conditional censorship distribution function. Under suitable conditions we establish the almost sure uniform convergence of those estimators with appropriate rate.

Keywords: *Competing risks, nonparametric regression function, right censoring, generalized product-limit estimator, convergence rate.*

AMS Subject Classification: *62N02; 62G05; 62G20, 62H12*

1 Introduction

The model of competing risks has been widely studied in the literature (see e.g. Kalbfleisch and Prentice (1980), Heckman and honoré (1989), Kwan and Singh (2001), Fermanian (2003), El Barmi and Mukerjee (2006), Geffray (2009), Bordes and Gneyou (2009)). Competing risks arise in medical, reliability or finance follow up involving multiple causes of failure when only the smallest failure time and the associated cause type are observed. In this mechanism, several failure times are right censored by the observed failure time in an informative way but in addition each failure time may be right censored by an event in a non informative manner. In many approaches, the competing risks are assumed to be either all independent or not. Here, we consider a population in which each individual is exposed to m mutually exclusive competing risks of failure eventually dependent. We study the strong uniform consistency of nonparametric estimators of classes of incident regression functions.

Let us denote by T_j the failure time from the j th cause with $j \in \{1, \dots, m\}$ and $m \geq 2$. Assume that each individual or entity is characterized by a \mathbb{R}^d -valued covariate Z

and denote by $X = \min(T_1, \dots, T_m)$ the smallest failure time, η the indicator of failure cause equal j if and only if $X = T_j$ where $1 \leq j \leq m$. Assume that X is also at risk of being right-censored by a continuous random variable C , independent of X given Z . Set $Y = \min(X, C)$ and $\delta = I(X \leq C)$ where $I(A)$ denotes the indicator function of any event A . Hence $\delta = 0$ if X is right-censored by C and $\delta = 1$ otherwise. In addition we define $\xi = \eta\delta$ satisfying $\xi = 0$ if X (and then all durations T_j s) is right-censored by C and $\xi = j$ if $X = T_j \leq C$. In statistical applications, a sample $\{(Y_i, \xi_i, Z_i)\}_{1 \leq i \leq n}$ of n independent copies of (Y, ξ, Z) is observed.

In this paper our aim is to estimate with appropriate almost sure uniform convergence rate with respect to $Z = z$ over a subset Δ of \mathbb{R}^d , the competing risk regression function r defined by

$$r(z) = \mathbb{E}[\psi(X)|Z = z],$$

where ψ belongs to a family of real-valued measurable functions on \mathbb{R}^+ , without any parametric or independence assumption. For example, nonparametric estimation of the conditional distribution $F_X(t|z) = \mathbb{P}[X \leq t|Z = z]$ is obtained for $\psi(x) = \psi_t(x) = I(x \leq t)$. For the case of a single duration ($m = 1$) Dabrowska (1989) gave some uniform convergence results with rates.

Unfortunately, in the competing risks model, without specific assumptions, the joint or marginal distribution functions, together with the related probability densities and hazard functions of the underlying failure times and the previous regression function are not identifiable (Tsiatis, 1975). In order to avoid the non identifiability problem, most models make parametric assumptions on the joint distribution function of the failure times or assume their independence. When no such assumptions are made, the quantities usually estimated are the cause specific functions instead of the overall or latent distribution functions. However, if each individual is characterized by a 'sufficiently informative' set of covariates, these distribution functions are identifiable under some regularity conditions (Heckman and Honoré, 1989). The problem of identifiability discussed in literature leads to concentrate no more on the latter regression function but on cause specific regression functions which are expressed in terms of observable functions of failure times given by

$$(1.1) \quad r_j(z) = \mathbb{E}[\psi(X)I(\eta = j)|Z = z], \quad j = 1, \dots, m,$$

where in order to insure the existence of $r_j(z)$, we assume that $\mathbb{E}|\psi(X)| < +\infty$.

The problem of estimating regression functions has been considered in the literature in non censored as well as censored frameworks (see e.g. Beran (1981), Dabrowska (1987, 1989), Haerdle et al. (1988), Derzko and Deheuvels (2000), Einmahl and Mason (2000), Kohler and Mathé (2002), Sun (2003), Gneyou (2005), Bordes and Gneyou (2009) and references therein).

In this paper, we propose a kernel-type estimator $\hat{r}_{jn}(z)$ of the incident regression function r_j defined in (1.1) and we establish that under suitable conditions it converges uniformly on Δ with some rates that are given in Section 3.

2 Definitions and nonparametric estimators

Recall that $\xi = \eta\delta$ where η and δ are respectively the failure cause and censoring indicators. Let us define the following conditional distribution functions:

$$\begin{aligned} F_X(t|z) &= \mathbb{P}[X \leq t|Z = z], \\ G(t|z) &= \mathbb{P}[C \leq t|Z = z], \\ H(t|z) &= \mathbb{P}[Y \leq t|Z = z], \end{aligned}$$

and for $1 \leq j \leq m$

$$F^{(j)}(t|z) = \mathbb{P}[X \leq t, \eta = j|Z = z].$$

The conditional subdistribution functions are defined by:

$$H^{(j)}(t|z) = \mathbb{P}[Y \leq t, \xi = j|Z = z], \text{ for } j = 0, \dots, m.$$

Note that

$$H^{(0)}(t|z) = \mathbb{P}[Y \leq t, \xi = 0|Z = z] = \mathbb{P}[Y \leq t, \delta = 0|Z = z],$$

and since conditionally on Z the random variables X and C are independent, we have

$$1 - H(t|z) = (1 - F_X(t|z))(1 - G(t|z)).$$

The connections between the observable incident cumulative distribution functions $F^{(j)}$ and $H^{(j)}$ ($0 \leq j \leq m$), and the unobservable cumulative distribution functions F_X , H and $H^{(0)}$ are given by:

$$\begin{aligned} H^{(j)}(t|z) &= \int_0^t \bar{G}(s^-|z) dF^{(j)}(s|z) \quad \text{for } 1 \leq j \leq m, \\ H^{(0)}(t|z) &= \int_0^t \bar{F}_X(s^-|z) dG(s|z), \end{aligned}$$

and $F_X = \sum_{j=1}^m F^{(j)}$, $H = \sum_{j=0}^m H^{(j)}$ where for a real function L we define $L(s^-) = \lim_{u \nearrow s} L(u)$ and $\bar{L} = 1 - L$.

Our nonparametric estimators of competing risks regression functions will be based on empirical versions of the j th cause specific conditional cumulative hazard function $\Lambda^{(j)}(t|z)$ and the censoring conditional cumulative hazard function $\Lambda^{(c)}(t|z)$ which are defined by:

$$\begin{aligned} \Lambda^{(j)}(t|z) &= \int_0^t \frac{dF^{(j)}(s|z)}{\bar{F}_X(s^-|z)} = \int_0^t \frac{dH^{(j)}(s|z)}{\bar{H}(s^-|z)}, \quad j = 1, \dots, m \\ (2.2) \quad \Lambda^{(c)}(t|z) &= \int_0^t \frac{dG(s|z)}{\bar{G}(s^-|z)} = \int_0^t \frac{dH^{(0)}(s|z)}{\bar{H}(s^-|z)}. \end{aligned}$$

Instead of considering $r_j(z)$ defined by (1.1) we consider $\bar{r}_j(z)$ defined by

$$(2.3) \quad \bar{r}_j(z) = \int_0^{\tau_z} \psi(t) \bar{F}_X(t^-|z) d\Lambda^{(j)}(t|z) = \int_0^{\tau_z} \frac{\psi(t)}{\bar{G}(t^-|z)} dH^{(j)}(t|z),$$

where for each $z \in \Delta$ and a given (small) real number $\gamma > 0$, $\tau_z = \inf\{t \geq 0; \bar{H}(t|z) \geq \gamma\}$. Note that whenever r_j exists, the smaller will be γ the closer will be the competing risks regression functions \bar{r}_j and r_j .

As a consequence we can estimate $\bar{r}_j(z)$, by replacing $H^{(j)}(t|z)$ and $\bar{G}(t|z)$ in (2.3) by some appropriate estimators. The cumulative distribution function $H(t|z)$ and the sub-cumulative distribution functions $H^{(j)}(t|z)$ ($0 \leq j \leq m$) can be respectively estimated by:

$$H_n(t|z) = \sum_{i=1}^n I(Y_i \leq t)W_i(h_n, z),$$

and

$$H_n^{(j)}(t|z) = \sum_{i=1}^n I(Y_i \leq t, \xi_i = j)W_i(h_n, z),$$

where for $1 \leq i \leq n$ the Nadaraya-Watson weights are defined by

$$W_i(h, z) = \frac{K_h(z - Z_i)}{\sum_{j=1}^n K_h(z - Z_j)}.$$

In the above formula K is a kernel function on \mathbb{R}^d , $K_h(\cdot) = h^{-d}K(\cdot/h)$, and $(h_n)_{n \geq 1}$ is a bandwidth sequence of non-increasing positive real numbers tending to 0. The censorship conditional survival function $\bar{G}(t|z)$ satisfies the one-to-one map relation (product integral mapping)

$$\bar{G}(t|z) = \prod_{s \leq t} \left(1 - \Lambda^{(c)}(ds|z)\right).$$

Because of relation (2.2), the conditional cumulative hazard function $\Lambda^{(c)}$ associated to $\bar{G}(t|z)$ is naturally estimated by

$$\Lambda_n^{(c)}(t|z) = \int_0^t \frac{dH_n^{(0)}(s|z)}{\bar{H}_n(s^-|z)} = \sum_{i=1}^n \frac{I(\xi_i = 0)I(Y_i \leq t)W_i(h_n, z)}{N(Y_i, z)},$$

where $N(t, z) = \sum_{i=1}^n I(Y_i \geq t)W_i(h_n, z)$. This leads to the Beran's (see Beran (1981) or Dabrowska (1989)) estimator \bar{G}_n of \bar{G} defined by

$$\bar{G}_n(t|z) = \prod_{s \leq t} \left(1 - \Delta \Lambda_n^{(c)}(s|z)\right) = \prod_{i=1}^n \left(1 - \frac{I(\xi_i = 0)I(Y_i \leq t)W_i(h_n, z)}{N(Y_i, z)}\right),$$

where $\Delta \Lambda_n^{(c)}(s|z) = \Lambda_n^{(c)}(s|z) - \Lambda_n^{(c)}(s^-|z)$.

The final nonparametric estimator of $\bar{r}_j(z)$ is therefore defined by

$$(2.4) \quad \hat{r}_{jn}(z) = \sum_{i=1}^n \frac{\psi(Y_i)I(Y_i \leq \tau_z)I(\xi_i = j)W_i(h_n, z)}{\bar{G}_n(Y_i|z)},$$

for $1 \leq j \leq m$.

Note that if C is independent of the covariate Z then $\bar{G}(t|z) = \bar{G}(t)$ and then the weights $W_i(h_n, z)$ are replaced by $1/n$ in $\bar{G}_n(\cdot|z)$ and the estimator in (2.4) reduces to the estimator of Bordes and Gneyou (2009). The later authors established strong consistency of their estimator for fixed z in \mathbb{R}^d and, using the delta method, they gave a central limit theorem for their estimator. In the next section, we show that under suitable assumptions, the strong convergence of \hat{r}_{jn} to \bar{r}_j holds uniformly over a compact subset $\Delta \subset \mathbb{R}^d$.

3 Strong uniform consistency

Let f be the marginal probability density function of the covariate Z and $\Delta \subset \text{supp} f \subset \mathbb{R}^d$ be a compact subset of \mathbb{R}^d . Our asymptotic results are obtained under smoothness conditions on the conditional subdistribution function given in the previous section and some conditions on both the kernel function and the bandwidth.

F1. $\inf_{z \in \Delta} \bar{H}(\tau_z|z) = \gamma > 0$.

F2. The marginal density function f is continuous on Δ . We note $\alpha = \inf_{z \in \Delta} f(z) > 0$.

F3. Functions f and $z \mapsto H^{(j)}(t|z)$ (for all $t \in [0, \tau_z], j = 0, \dots, m$) are twice continuously differentiable with respect to z , and the second derivative of $z \mapsto K^{(j)}(z) = H^{(j)}(t|z)f(z)$ is continuous on Δ uniformly in $t \in [0, \tau_z]$.

F4. Let P be a polynomial and ϕ a positive bounded real function of bounded variation such that $K = \phi \circ P$ is a kernel function satisfying:

$$(i) \int_{\mathbb{R}^d} K(x)dx = 1, \quad (ii) \int_{\mathbb{R}^d} xK(x)dx = 0, \quad (iii) \int_{\mathbb{R}^d} xx^T K(x)dx \text{ is positive definite.}$$

F5. $\sup_{z \in \Delta} \int_0^{\tau_z} |d\psi(s)| \leq M < +\infty$.

F6. The sequence of bandwidth $(h_n)_{n \geq 1}$ satisfies:

$$(i) h_n \rightarrow 0, \quad (ii) \frac{nh_n^d}{|\log h_n|} \rightarrow +\infty, \quad (iii) \frac{|\log h_n|}{\log \log n} \rightarrow +\infty, \quad (iv) h_n^d \leq ch_{2n}^d$$

for some $c > 0$.

Assumption F1 allows to ensure an uniform observation rate in $z \in \Delta$ and $t \in [0, \tau_z]$, while F2 allows to control the denominator of the estimator \hat{r}_{nj} of \bar{r}_j . Assumptions F3 and F4 (i)–(iii) allows to control the sup-norm distance between any involved function L and its regularized version $L * K_h$. The class of kernel functions we consider in F4 was introduced by several authors (see e.g. Giné and Guillou, 2002) it ensures that some classes of function are \mathcal{VC} -classes. Concerning F5 it is easy to check that together with F1 it leads to:

$$(3.5) \quad \sup_{z \in \Delta} \int_0^{\tau_z} \left| d \frac{\psi(s)}{G(s|z)} \right| \leq M' < +\infty.$$

As a consequence it guarantees that $\bar{r}_j(z)$ exists for all $z \in \Delta$. Assumptions in F6 are those given by Giné and Guillou (2002). Under such assumptions these authors obtain uniform consistency of kernel estimators for multivariate densities. Hereinafter we show almost sure uniform consistency results on the set $\{(t, z); t \in [0, \tau_z], z \in \Delta\} \subset \mathbb{R}^{d+1}$. Let us begin with the following lemmas.

Lemma 3.1 *Let Γ be a compact subset of $\mathbb{R} \times \mathbb{R}^d$, Δ the restriction of Γ to \mathbb{R}^d , and $\ell : \Gamma \rightarrow \mathbb{R}$ be a function such that $z \mapsto \ell(t, z)$ is twice continuously differentiable on the compact set Δ and the partial derivative $(t, z) \mapsto \frac{\partial^2 \ell}{\partial z \partial z^T}(t, z)$ is continuous on Γ . Then under assumptions F4 and F6-(i)*

$$\sup_{(t,z) \in \Gamma} \left| \int_{\mathbb{R}^d} K_{h_n}(z-s)\ell(t,s)ds - \ell(t,z) \right| = \mathcal{O}(h_n^{2d}).$$

The proof of this lemma is a straightforward extension of the proof of Lemma A.2 of Bordes and Gneyou (2009), hence it is omitted. For simplicity from now on we set $\Gamma = \{(t, z); t \in [0, \tau_z], z \in \Delta\} \subset \mathbb{R}^{d+1}$.

Lemma 3.2 *Assume that F1–F4, and F6 hold. Then for $0 \leq j \leq m$ we have*

$$\sup_{(t,z) \in \Gamma} \left| H_n^{(j)}(t|z) - H^{(j)}(t|z) \right| = \mathcal{O} \left(\left(|\log h_n| / nh_n^d \right)^{1/2} \right) + \mathcal{O} \left(h_n^{2d} \right) \quad a.s.$$

Proof. For all $0 \leq j \leq m$ and all $(t, z) \in \Gamma$ we have

$$\begin{aligned} H_n^{(j)}(t|z) - H^{(j)}(t|z) &= \frac{1}{nf_n(z)} \sum_{k=1}^n I(Y_k \leq t, \xi_k = j) K_{h_n}(z - Z_k) - H^{(j)}(t|z) \\ &= \frac{1}{nf_n(z)} \sum_{k=1}^n \left[I(Y_k \leq t, \xi_k = j) K_{h_n}(z - Z_k) - H^{(j)}(t|z) f(z) \right] \\ &\quad + H^{(j)}(t|z) \frac{f(z) - f_n(z)}{f_n(z)}, \end{aligned}$$

where $f_n(z) = \frac{1}{n} \sum_{k=1}^n K_{h_n}(z - Z_k)$ is the usual kernel estimator of the marginal probability density function f of Z . It readily follows that

$$(3.6) \quad \left| H_n^{(j)}(t|z) - H^{(j)}(t|z) \right| \leq \left(\inf_{z \in \Delta} f_n(z) \right)^{-1} (A_n + B_n + C_n),$$

where

$$\begin{aligned} A_n &= \sup_{(t,z) \in \Gamma} \left| \frac{1}{n} \sum_{k=1}^n [I(Y_k \leq t, \xi_k = j) K_{h_n}(z - Z_k) - \mathbb{E}(I(Y_k \leq t, \xi_k = j) K_{h_n}(z - Z_k))] \right|, \\ B_n &= \sup_{(t,z) \in \Gamma} \left| \mathbb{E}(I(Y_k \leq t, \xi_k = j) K_{h_n}(z - Z_k)) - H^{(j)}(t|z) f(z) \right|, \\ C_n &= \sup_{(t,z) \in \Gamma} \left| H^{(j)}(t|z) (f(z) - f_n(z)) \right|. \end{aligned}$$

By assumptions F2, F4, and F6, applying Theorem 2.3 of Giné and Guillou (2002), we have

$$(3.7) \quad \sup_{z \in \Delta} |f_n(z) - \bar{f}_n(z)| = \mathcal{O} \left(\left(|\log h_n| / nh_n^d \right)^{1/2} \right) \quad a.s.$$

where

$$\bar{f}_n(z) = f \star K_{h_n}(z) = \int_{\mathbb{R}^d} K_{h_n}(z - s) f(s) ds.$$

By assumptions F3 and F4, applying Lemma 3.1 we obtain

$$(3.8) \quad \sup_{z \in \Delta} |\bar{f}_n(z) - f(z)| = \mathcal{O} \left(h_n^{2d} \right).$$

As a straightforward consequence of (3.7) and (3.8) and because of assumptions F2 and F6 (i)–(ii) we have

$$(3.9) \quad \inf_{z \in \Delta} f_n(z) = \alpha + o(1) \quad a.s.$$

Moreover since

$$C_n \leq \sup_{z \in \Delta} |f_n(z) - \bar{f}_n(z)| + \sup_{z \in \Delta} |\bar{f}_n(z) - f(z)|$$

we have by (3.7) and (3.8) the following result

$$(3.10) \quad C_n = \mathcal{O} \left(\left(|\log h_n| / nh_n^d \right)^{1/2} \right) + \mathcal{O} \left(h_n^{2d} \right) \quad \text{a.s.}$$

Let us now consider B_n . First remark that

$$\begin{aligned} \mathbb{E} [I(Y_k \leq t, \xi_k = j) K_{h_n}(z - Z_k)] &= \mathbb{E} [K_{h_n}(z - Z_k) \mathbb{E} [I(Y_k \leq t, \xi_k = j) | Z_k]] \\ &= \int_{\mathbb{R}^d} K_{h_n}(z - s) H^{(j)}(t|s) f(s) ds. \end{aligned}$$

Applying Lemma 3.1 to $z \mapsto K_j(t, z) = H^{(j)}(t|z) f(z)$ we obtain under assumptions F3–F4, and F6 that

$$(3.11) \quad B_n = \mathcal{O} \left(h_n^{2d} \right).$$

For the remaining term A_n , for $h > 0$ and $(t, z) \in \Gamma$ we define functions

$$g_{t,z,h}(y, x, s) = I(y \leq t, x = j) K \left(\frac{z - s}{h} \right).$$

Let P be the probability measure generated by (Y, ξ, Z) , we note $Pg_{t,z,h_n} = \mathbb{E} [g_{t,z,h_n}(Y, \xi, Z)]$. Considering the empirical measure $\mathbb{P}_n = \sum_{i=1}^n \delta_{(Y_i, \xi_i, Z_i)}$ where δ_x denotes the Dirac measure at x , we have

$$\mathbb{P}_n g_{t,z,h_n} = \sum_{i=1}^n I(Y_i \leq t, \xi_i = j) K \left(\frac{z - Z_i}{h_n} \right).$$

By Lemma 2.6.16 and 2.6.18 of van der Vaart and Wellner (1996)

$$\{(y, x) \mapsto I(y \leq t, x = j); t \geq 0\}$$

is a bounded \mathcal{VC} -class of measurable functions. Moreover, by Giné and Guillou (2002), under assumption F4 the class of functions

$$\mathcal{F} = \left\{ s \mapsto K \left(\frac{z - s}{h} \right); z \in \mathbb{R}^d, h \in (0, +\infty) \right\}$$

is a bounded \mathcal{VC} -class of measurable functions. Applying Lemma A.1 of Einmahl and Mason (2000), we see that under assumption F4, the class of functions

$$\mathcal{G} = \left\{ (y, x, z) \mapsto g_{t,s,h}(y, x, z) = I(y \leq t, x = j) K \left(\frac{z - s}{h} \right); h > 0, t \geq 0, s \in \mathbb{R}^d \right\}$$

is a \mathcal{VC} -class of bounded measurable functions satisfying for all probability measures Q on the Borel subsets of \mathbb{R}^{d+2}

$$\mathcal{N}(\varepsilon \|K\|_\infty, \mathcal{G}, L^2(Q)) \leq \left(\frac{A}{\varepsilon} \right)^\nu \quad 0 < \varepsilon < 1,$$

where A and ν are suitable constants. The measurability follows from the continuity of the kernel function and the measurability of the indicator functions. Let us consider \mathcal{G}_k , the \mathcal{VC} -subclasses of \mathcal{G} defined by

$$\mathcal{G}_k = \{g_{t,z,h} \in \mathcal{G}; (t, z) \in \Gamma, h_{2^k} < h \leq h_{2^{k-1}}\}$$

for $k \geq 1$, we have:

$$\sup_{g \in \mathcal{G}_k} \|g\|_\infty \leq \|K\|_\infty = u_k,$$

and

$$\sup_{g \in \mathcal{G}_k} \text{Var}[g_{t,z,h}(Y, \xi, Z)] \leq \int_{\mathbb{R}^d} K^2 \left(\frac{z-s}{h} \right) f(s) ds \leq \|f\|_\infty \|K\|_{2^2}^2 h_{2^{k-1}}^d = \sigma_k^2.$$

The end of the proof follows the lines of the proof of Theorem 2.3 of Giné and Guillou (2002). Indeed

$$A_n = \frac{1}{nh_n^d} \sup_{(t,z) \in \Gamma} |(\mathbb{P}_n - P)g_{t,z,h_n}|$$

where

$$(\mathbb{P}_n - P)g = \sum_{i=1}^n (g(Y_i, \xi_i, Z_i) - \mathbb{E}g(Y_i, \xi_i, Z_i)).$$

Hence by the Montgomery-Smith's maximal inequality (Montgomery-Smith, 1993) we have for some finite positive constant $C > 0$

$$\begin{aligned} & \mathbb{P} \left(\max_{2^{k-1} < n \leq 2^k} \sqrt{\frac{nh_n^d}{|\log h_n|}} A_n > C \right) \\ & \leq \mathbb{P} \left(\max_{2^{k-1} < n \leq 2^k} \frac{1}{\sqrt{nh_n^d |\log h_n|}} \sup_{(t,z) \in \Gamma} |(\mathbb{P}_n - nP)g_{t,z,h_n}| > C \right) \\ & \leq 9\mathbb{P} \left(\sup_{\substack{(t,z) \in \Gamma \\ h_{2^{k-1}} \leq h \leq h_{2^k}}} |(\mathbb{P}_{2^k} - P)g| > \frac{C}{30} \sqrt{2^{k-1} h_{2^k}^d |\log h_{2^k}|} \right) \\ & \leq 9\mathbb{P} \left(\sup_{g \in \mathcal{G}_k} |(\mathbb{P}_{2^k} - P)g| > \frac{C}{30} \sqrt{2^{k-1} h_{2^k}^d |\log h_{2^k}|} \right). \end{aligned}$$

It is easy to check that $\sigma_k < u_k/2$ and $\sqrt{2^k} \sigma_k \geq u_k \sqrt{\log(u_k/\sigma^k)}$ for k large enough, thus the version of the exponential inequality from Talagrand (1996) given in Giné and Guillou (2002, Corollary 2.2) may be applied for some constant $C > 0$:

$$(3.12) \quad \mathbb{P} \left(\sup_{g \in \mathcal{G}_k} \|(\mathbb{P}_{2^k} - P)g\| > \frac{C (2^{k-1} h_{2^k}^d |\log h_{2^k}|)^{1/2}}{30} \right) \leq K \exp \left(-K' \log \left(\frac{u_k}{\sigma_k} \right) \right)$$

where K and K' are positive constants that do not depend on k . Because of F6 (iii) we have

$$\frac{\log(u_k/\sigma_k)}{\log k} \rightarrow +\infty,$$

thus for k large enough the right hand side of inequality (3.12) is less than K''/k^α with $\alpha > 1$ and K'' a positive constant. Hence by the Borel-Cantelli lemma, we obtain

$$(3.13) \quad A_n = \mathcal{O} \left(\left(|\log h_n| / nh_n^d \right)^{1/2} \right) \quad \text{a.s.}$$

Lemma 3.2 follows from (3.6), (3.9), (3.10), (3.11) and (3.13). \square

Before proving the next theorem we need the following Lemma.

Lemma 3.3 *Under assumptions F1-F4 and F6 we have*

$$\sup_{z \in \Delta} \frac{\sum_{i=1}^n W_i^2(h_n, z)}{N^2(\tau_z, z)} = \mathcal{O} \left((nh_n^d)^{-1} \right) \quad \text{a.s.}$$

Proof. For $(t, z) \in \Gamma$ we have $N(\tau_z, z) = \bar{H}_n(\tau_z^- | z) = \sum_{j=0}^m H_n^{(j)}(\tau_z^- | z)$ then by Lemma 3.2 we have $\sup_{(t,z) \in \Gamma} |N(t, z) - \bar{H}(t^- | z)| = o(1)$ a.s. Both N and \bar{H} are $[0,1]$ -valued then for n large enough

$$\begin{aligned} \frac{1}{N^2(\tau_z, z)} &\leq \frac{1}{\bar{H}^2(\tau_z^- | z) + (N^2(\tau_z, z) - \bar{H}^2(\tau_z^- | z))} \\ &\leq \frac{1}{\gamma^2 + (N(\tau_z, z) + \bar{H}(\tau_z^- | z))(N(\tau_z, z) - \bar{H}(\tau_z^- | z))} \\ &\leq \frac{1}{\gamma^2 - 2 \sup_{(t,z) \in \Gamma} |N(t, z) - \bar{H}(t^- | z)|} = \mathcal{O}(1) \quad \text{a.s.} \end{aligned}$$

Moreover we have

$$\begin{aligned} \sum_{i=1}^n W_i^2(h_n, z) &= \frac{h_n^{-2d} \sum_{i=1}^n K^2 \left(\frac{z - Z_i}{h_n} \right)}{\left(\sum_{i=1}^n K_{h_n}(z - Z_i) \right)^2} \\ &= \frac{n^{-1} \sum_{i=1}^n K_{h_n}^*(z - Z_i)}{\left(n^{-1} \sum_{i=1}^n K_{h_n}(z - Z_i) \right)^2} (nh_n^d)^{-1} \int_{\mathbb{R}^d} K^2(x) dx \\ &= \frac{f_n^*(z)}{f_n^2(z)} (nh_n^d)^{-1} \int_{\mathbb{R}^d} K^2(x) dx, \end{aligned}$$

where $K_h^*(x) = h^{-d} K^2(x/h) / \int_{\mathbb{R}^d} K^2(x) dx$ is a kernel function satisfying conditions required in Theorem 3.2 in Giné and Guillou (2002). Because f is continuous on Δ and the bandwidth satisfies F6 we obtain by Theorem 3.2 in Giné and Guillou (2002) that both f_n^* and f_n converge almost surely to f uniformly on Δ . As a consequence we have $\sup_{z \in \Delta} \sum_{i=1}^n W_i^2(h_n, z) = \mathcal{O}((nh_n^d)^{-1})$ a.s. which proves the lemma. \square

Theorem 3.1 *Assume that assumptions F1-F4 and F6 hold. Then for n large enough we have*

$$\sup_{(t,z) \in \Gamma} |G_n(t|z) - G(t|z)| = \mathcal{O} \left(\left(|\log h_n| / nh_n^d \right)^{1/2} \right) + \mathcal{O} \left(h_n^{2d} \right) \quad \text{a.s.}$$

Proof. Let us introduce $\tilde{\Lambda}_n^{(c)}$, a perturbation of $\Lambda_n^{(c)}$ defined by

$$\tilde{\Lambda}_n^{(c)}(t|z) = \sum_{i=1}^n \frac{I(\xi_i = 0, Y_i \leq t)W_i(h_n, z)}{N(Y_i, z) + W_i(h_n, z)},$$

with the corresponding product limit estimator \tilde{G}_n defined by

$$\tilde{G}_n(t|z) = \prod_{i=1}^n \left(1 - \frac{I(\xi_i = 0, Y_i \leq t)W_i(h_n, z)}{N(Y_i, z) + W_i(h_n, z)} \right).$$

Since $|e^{-x} - e^{-y}| \leq |x - y|$ for all $x, y \geq 0$ we have

$$\begin{aligned} |\bar{G}_n(t|z) - G(t|z)| &\leq \left| \bar{G}_n(t|z) - \tilde{G}_n(t|z) \right| + \left| \exp\left(-\tilde{\Lambda}_n^{(c)}(t|z)\right) - \exp\left(-\Lambda^{(c)}(t|z)\right) \right| \\ &\quad + \left| \exp\left(\log \tilde{G}_n(t|z)\right) - \exp\left(-\tilde{\Lambda}_n^{(c)}(t|z)\right) \right| \\ &\leq A_n(t, z) + B_n(t, z) + C_n(t, z), \end{aligned}$$

where $A_n(t, z) = \left| \bar{G}_n(t|z) - \tilde{G}_n(t|z) \right|$, $B_n(t, z) = \left| \tilde{\Lambda}_n^{(c)}(t|z) - \Lambda^{(c)}(t|z) \right|$ and $C_n(t, z) = \left| \log \tilde{G}_n(t|z) + \tilde{\Lambda}_n^{(c)}(t|z) \right|$.

Because $|\prod_{i=1}^n a_i - \prod_{i=1}^n b_i| \leq \sum_{i=1}^n |a_i - b_i|$ whenever $|a_i| \leq 1$ and $|b_i| \leq 1$ for $1 \leq i \leq n$ we have

$$\begin{aligned} A_n(t, z) &\leq \sum_{i=1}^n \left| \frac{W_i(h_n, z)}{N(Y_i, z)} - \frac{W_i(h_n, z)}{N(Y_i, z) + W_i(h_n, z)} \right| \\ &\leq \frac{\sum_{i=1}^n W_i^2(h_n, z)}{N^2(\tau_z, z)}. \end{aligned}$$

Applying Lemma 3.3 we obtain that $\sup_{(t,z) \in \Gamma} A_n(t, z) = \mathcal{O}((nh_n^d)^{-1})$ a.s. Now using the fact that $0 \leq -\log(1-x) - x \leq x^2(1-x)^{-1}$ for $x \in [0, 1)$ we have

$$\begin{aligned} C_n(t, z) &\leq \sum_{i=1}^n \left| \log \left(1 - \frac{I(\xi_i = 0, Y_i \leq t)W_i(h_n, z)}{N(Y_i, z) + W_i(h_n, z)} \right) + \frac{I(\xi_i = 0, Y_i \leq t)W_i(h_n, z)}{N(Y_i, z) + W_i(h_n, z)} \right| \\ &\leq \sum_{i=1}^n \frac{W_i^2(h_n, z)}{(N(Y_i, z) + W_i(h_n, z))^2} \left(1 - \frac{W_i(h_n, z)}{N(Y_i, z) + W_i(h_n, z)} \right)^{-1} \\ &\leq \frac{\sum_{i=1}^n W_i^2(h_n, z)}{N^2(\tau_z, z)}. \end{aligned}$$

By Lemma 3.3 we conclude that $\sup_{(t,z) \in \Gamma} C_n(t, z) = \mathcal{O}((nh_n^d)^{-1})$ a.s. It remains to study the convergence rate of B_n . We have

$$B_n(t, z) \leq \left| \tilde{\Lambda}_n^{(c)}(t|z) - \Lambda_n^{(c)}(t|z) \right| + \left| \Lambda_n^{(c)}(t|z) - \tilde{\Lambda}^{(c)}(t|z) \right|$$

it is straightforward to see that applying again Lemma 3.3 we have

$$\sup_{(t,z) \in \Gamma} \left| \tilde{\Lambda}_n^{(c)}(t|z) - \Lambda_n^{(c)}(t|z) \right| = \mathcal{O}((nh_n^d)^{-1}).$$

Furthermore note that from Section 2 $\Lambda_n^{(c)} - \Lambda^{(c)}$ can be written

$$\Lambda_n^{(c)}(t|z) - \Lambda^{(c)}(t|z) = D_n(t, z) + E_n(t, z)$$

where

$$\begin{aligned} D_n(t, z) &= \int_0^t \frac{d(H_n^{(0)}(s|z) - H^{(0)}(s|z))}{1 - H(s^-|z)}, \\ E_n(t, z) &= \int_0^t \frac{(H_n(s^-|z) - H(s^-|z))}{(1 - H(s^-|z))(1 - H_n(s^-|z))} dH_n^{(0)}(s|z). \end{aligned}$$

By integration by parts formula and assumption F1 we have

$$\begin{aligned} D_n(t, z) &= \frac{H_n^{(0)}(t|z) - H^{(0)}(t|z)}{1 - H(t^-|z)} - \int_0^t \left(H_n^{(0)}(s|z) - H^{(0)}(s|z) \right) d(\bar{H}(s^-|z)^{-1}) \\ &\leq \sup_{(t,z) \in \Gamma} \left| H_n^{(0)}(t|z) - H^{(0)}(t|z) \right| \times \frac{2 - \gamma}{\gamma}. \end{aligned}$$

Hence by Lemma 3.2 we obtain

$$(3.14) \quad \sup_{(t,z) \in \Gamma} |D_n(t, z)| = \mathcal{O} \left(\left(|\log h_n| / nh_n^d \right)^{1/2} \right) + \mathcal{O} \left(h_n^{2d} \right).$$

Remark that

$$(3.15) \quad \int_0^t \left| dH_n^{(j)}(s^-|z) \right| = \sum_{i=1}^n W_i(h_n, z) \leq 1$$

then for n large enough we have with probability one

$$\begin{aligned} |D_n(t, z)| &\leq \sup_{(t,z) \in \Gamma} \left| H_n^{(0)}(t|z) - H^{(0)}(t|z) \right| \times \sup_{(t,z) \in \Gamma} \int_0^t \left| dH_n^{(j)}(s^-|z) \right| \\ &\quad \times \frac{1}{\gamma(\gamma - \sup_{(t,z) \in \Gamma} \left| H_n^{(0)}(t|z) - H^{(0)}(t|z) \right|)} \\ &\leq \mathcal{O}(1) \times \sup_{(t,z) \in \Gamma} \left| H_n^{(0)}(t|z) - H^{(0)}(t|z) \right|, \end{aligned}$$

thus by Lemma 3.2 we obtain

$$(3.16) \quad \sup_{(t,z) \in \Gamma} |E_n(t, z)| = \mathcal{O} \left(\left(|\log h_n| / nh_n^d \right)^{1/2} \right) + \mathcal{O} \left(h_n^{2d} \right).$$

The theorem follows from results on A_n, B_n, C_n, D_n and E_n . \square

Note that analogous strong convergence results of the nonparametric estimator of the conditional distribution function distribution has been established earlier by Dabrowska (1987, 1989), and more recently by Ghouch and van Keilegom (2008) for the conditional censoring distribution in the dependent data setup.

Theorem 3.2 Suppose that assumptions F1–F6 are fulfilled, then for $1 \leq j \leq m$, as n tends to infinity we have:

$$(3.17) \quad \sup_{z \in \Delta} |\hat{r}_{jn}(z) - \bar{r}_j(z)| = \mathcal{O} \left(\left(|\log h_n| / nh_n^d \right)^{1/2} \right) + \mathcal{O} \left(h_n^{2d} \right) \quad a.s.$$

Proof. Let us consider $j \in \{1, \dots, m\}$ and $z \in \Delta$. We have

$$(3.18) \quad \begin{aligned} \hat{r}_{jn}(z) - \bar{r}_j(z) &= \int_0^{\tau_z} \frac{\psi(s)}{\bar{G}_n(s^-|z)} dH_n^{(j)}(s|z) - \int_0^{\tau_z} \frac{\psi(s)}{\bar{G}(s^-|z)} dH^{(j)}(s|z) \\ &= Q_n(\tau_z, z) + R_n(\tau_z, z), \end{aligned}$$

where

$$(3.19) \quad \begin{aligned} Q_n(t, z) &= \int_0^t \frac{\psi(s)}{\bar{G}(s^-|z)} d(H_n^{(j)}(s|z) - H^{(j)}(s|z)), \\ R_n(t, z) &= \int_0^t \frac{\psi(s)(G_n(s^-|z) - G(s^-|z))}{\bar{G}_n(s^-|z)\bar{G}(s^-|z)} dH_n^{(j)}(s|z). \end{aligned}$$

Using F5 and the integration by parts formula we have

$$Q_n(t, z) = \frac{\psi(t)(H_n^{(j)}(t|z) - H^{(j)}(t|z))}{\bar{G}(t|z)} - \int_0^t (H_n^{(j)}(s|z) - H^{(j)}(s|z)) d \left(\frac{\psi(s)}{\bar{G}(s^-|z)} \right).$$

Because of assumption F1 and (3.5) derived from F5 we have

$$\sup_{z \in \Delta} |Q_n(\tau_z, z)| \leq \|\psi\|_\infty (\gamma^{-1} + M') \sup_{(t,z) \in \Gamma} |H_n^{(j)}(t|z) - H^{(j)}(t|z)|,$$

then by Lemma 3.2 we derive

$$(3.20) \quad \sup_{z \in \Delta} |Q_n(t, z)| = \mathcal{O} \left(\left(|\log h_n| / nh_n^d \right)^{1/2} \right) + \mathcal{O} \left(h_n^{2d} \right) \quad a.s.$$

By assumptions F1 and F5 and (3.15) we have

$$\begin{aligned} &\sup_{z \in \Delta} |R_n(\tau_z, z)| \\ &\leq \frac{\|\psi\|_\infty}{\gamma} \sup_{(t,z) \in \Gamma} \frac{1}{\bar{G}_n(t|z)} \sup_{(t,z) \in \Gamma} |G_n(t|z) - G(t|z)| \\ &\leq \frac{\|\psi\|_\infty}{\gamma} \frac{1}{\gamma - \sup_{(t,z) \in \Gamma} |\bar{G}_n(t|z) - G(t|z)|} \sup_{(t,z) \in \Gamma} |G_n(t|z) - G(t|z)|. \end{aligned}$$

By Theorem 3.1 for n large enough we have

$$\frac{1}{\gamma - \sup_{(t,z) \in \Gamma} |\bar{G}_n(t|z) - G(t|z)|} = \mathcal{O}(1) \quad a.s.$$

then it follows that

$$(3.21) \quad \sup_{z \in \Delta} |R_n(\tau_z, z)| = \mathcal{O} \left(\left(|\log h_n| / nh_n^d \right)^{1/2} \right) + \mathcal{O} \left(h_n^{2d} \right) \quad a.s.$$

Finally by (3.18), (3.20) and (3.21) we obtain the expected uniform convergence rate for $\hat{r}_{jn} - \bar{r}_j$. \square

4 Concluding remarks

The estimation method we proposed is quite general and allows to estimate many quantities like the classical incident regression function

$$r_j(z) = \mathbb{E}(XI(\eta = j)|Z = z),$$

which is obtained for $\psi(x) = x$. Function $\psi(x) = x^2$ yields the nonparametric estimator of the incident conditional variance. Because the incident conditional distribution function $F^{(j)}$ may be estimated at $(t, z) \in \Gamma$ by

$$\hat{F}_n^{(j)}(t|z) = \hat{r}_{jn}(t|z)$$

where for a given $z \in \Delta$ we replace the function ψ by a function $\psi_t(x) = I(x \leq t)$ indexed by $t \in [0, \tau_z]$. Following the lines of the proof of Theorem 3.2 it is straightforward to obtain the following convergence rate.

Corollary 4.1 *Under assumptions of Theorem 3.1 we have*

$$\sup_{(t,z) \in \Gamma} \left| \hat{F}_n^{(j)}(t|z) - F^{(j)}(t|z) \right| = \mathcal{O} \left(\left(|\log h_n| / nh_n^d \right)^{1/2} \right) + \mathcal{O} \left(h_n^{2d} \right) \quad a.s.$$

Some simulation results are provided in Bordes and Gneyou (2009) where asymptotic results deals with consistency and central limit theorem for $r_{jn}(z)$ given a fixed value of $z \in \Delta$. Here convergence results are obtained uniformly in $z \in \Delta$ when moreover the censoring variables may depend on the covariate Z . The convergence rates we obtain depend on the bandwidth h_n . Choosing $h_n = cn^{-\alpha}$ it is easy to see that F6 is satisfied whenever $\alpha \in (0, 1/d)$ and the best rate is obtained for $\alpha = 1/5d$. For the same bandwidth the convergence rate we obtain for the conditional Kaplan-Meier estimator is optimal and is the same as in Dabrowska (1989). More generally the $\mathcal{O}(h_n^{2d})$ term involved in the rates come from the regularity of the function to be estimated and this rate could be improved by assuming more regularity on these functions. Concerning the choice of the kernel function K the multivariate gaussian kernel function

$$K(z) = \frac{1}{(2\pi)^{d/2}} \exp \left(-\frac{1}{2} \sum_{k=1}^d z_k^2 \right) \quad z \in \mathbb{R}^d,$$

fulfills the condition F4. However, as discussed in Giné and Guillou (2002) many other kernel functions satisfying their condition (K₁) are possible, like for example the uniform kernel on $[-1, 1]^d$.

Acknowledgments

We are grateful to one anonymous referee for numerous helpful comments during the preparation of this article.

References

- [1] BERAN, R. (1981). Nonparametric regression with randomly censored data. Technical report, *Univ. Carolina, Berkeley*.
- [2] BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and product-limit estimates under random censorship. *The Annals of Statistics* **2(3)**, 437–453.
- [3] BORDES, L. and GNEYOU, K. (2009). A central limit theorem for nonparametric regression functions in competing risks model under right censoring. Submitted.
- [4] DABROWSKA, D. M. (1987). Nonparametric regression with censored survival data. *Scandinavian Journal of Statistics* **17**, 1157–1167.
- [5] DABROWSKA, D. M. (1989). Uniform consistency of the kernel conditional Kaplan-Meier estimate. *The Annals of Statistics* **17(3)**, 1157–1167.
- [6] DERZKO, G. and DEHEUVELS, P. (2002). Estimation non paramétrique de la régression dichotomique - Application biomédicale. *Comptes-Rendus de l'Académie des Sciences - Paris - Série I* **334**, 59–63.
- [7] EINMAHL, U. and MASON, D. M. (2000). An empirical process approach to the uniform consistency of kernel-type function estimators. *Journal of Theoretical Probability* **13**, 1–37.
- [8] EL BARMÍ, H. and MUKERJEE, H. (2006). Restricted estimation of the cumulative incidence functions corresponding to competing risks. *Institute of Mathematical Statistics Lecture Notes - Monograph Series*, 241–252.
- [9] FERMANIAN, J-D. (2003). Nonparametric estimation of competing risks models with covariates. *Journal of Multivariate Analysis* **85**, 156–191.
- [10] GEFFRAY, S. (2009). Strong Approximations for dependent competing risks with independent censorship. *Test* **18**, 76–95.
- [11] GNEYOU, K. E. (2005). Vitesse de convergence de certains estimateurs de Kaplan-Meier de la régression. *Afrika Statistika* **1(1)**, 77–92.
- [12] GINÉ, E. and GUILLOU, A. (2001). On consistency of kernel density estimators for randomly censored data : rates holding uniformly over adaptive intervals. *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistique* **37(4)**, 503–522.
- [13] GINÉ, E. and GUILLOU, A. (2002). Rate of strong uniform consistency for multivariate kernel density estimators. *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistique* **38(6)**, 907–921.
- [14] GHOUGH, A. and VAN KEILEGOM, I. (2008). Nonparametric regression with dependent censored data. *Scandinavian Journal of Statistics* **35**, 228–237.
- [15] HAERDLE, W., JANSSEN, P. and SERFLING, R. (1988). Strong uniform consistency rates for estimators of conditional functionals. *The Annals of Statistics* **16**, 1428–1449.
- [16] HECKMAN, J. J. and HONORÉ, B. (1989). The identifiability of competing risks model. *Biometrika* **76**, 325–330.
- [17] KALBFLEISCH, J. D. and PRENTICE R.L. (1980). *The Statistical Analysis of Failure Time Data*. John Wiley and Sons, New York.

- [18] KOHLER, M. and MATHÉ, K. (2001). Prediction from randomly right censored data. *Journal of Multivariate Analysis* **1**, 73–100.
- [19] KWAN, P. H. and SINGH, H. (2001). On nonparametric estimation of the survival function with competing risks. *Scandinavian Journal of Statistics* **28**, 715–724.
- [20] MONTGOMERY-SMITH, S. J. (1993). Comparison of sums of independent identically distributed random vectors. *Probability and Mathematical Statistics* **14**, 281–285.
- [21] STUTE, W. (1994). Strong and weak representations of cumulative hazard function and Kaplan-Meier estimators on increasing sets. *Journal of Statistical and Planning Inference* **42**, 315–329.
- [22] SUN, L. (2003). Fixed design nonparametric regression with truncated and censored data. *Acta Mathematicae Applicatae Sinica* **29(2)**, 229–238.
- [23] TALAGRAND, M. (1996). New concentration inequalities in product spaces, *Inventiones Mathematicae* **126**, 505–563.
- [24] TSIATIS, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences USA* **72**, 20–22.
- [25] VAN DER VAART, A. and WELLNER, J. (1996). *Weak convergence and empirical processes, with applications to statistics*. Springer-Verlag, New-York.