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 honé (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 $j \in \{1, \ldots, m\}$ and $m \geq 2$. Assume that each individual or entity is characterized by a $\mathbb{R}^d$-valued covariate $Z$
and denote by $X = \min(T_1, \ldots, T_m)$ the smallest failure time, $\eta$ 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 $(Y, \xi, Z)$ 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 $\Delta$ of $\mathbb{R}^d$, the competing risk regression function $r$ defined by

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

where $\psi$ 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 conduces 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

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

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 $\Delta$ with some rates that are given in Section 3.
2 Definitions and nonparametric estimators

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

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

and for $1 \leq j \leq m$

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

The sub-conditional distribution functions are defined by:

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

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:

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

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

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 $\tilde{\Lambda}(t|z)$ which are defined by:

$$\Lambda^{(j)}(t|z) = \int_0^t \frac{dF^{(j)}(s|z)}{F_X(s^0|z)} = \int_0^t \frac{dH^{(j)}(s|z)}{H(s^0|z)}, \quad j = 1, \ldots, m,$$
$$\tilde{\Lambda}(t|z) = \int_0^t \frac{dG(s|z)}{G(s^0|z)} = \int_0^t \frac{dH^{(0)}(s|z)}{H(s^0|z)}. \quad \text{(2.2)}$$

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

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

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


As a consequence we can estimate \( \hat{r}_j(z) \), by replacing \( H^{(j)}(t|z) \) and \( \hat{G}(t|z) \) 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

\[
W_i(h, z) = \frac{K_h(z - Z_i)}{\sum_{i=1}^{n} K_h(z - Z_i)}.
\]

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

\[
\hat{G}(t|z) = \prod_{s \leq t} \left( 1 - \hat{\Lambda}(ds|z) \right).
\]

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

\[
\hat{\Lambda}_n(t|z) = \int_0^t \frac{dH_n^{(0)}(s|z)}{H_n(s^-|z)} = \sum_{i=1}^{n} \frac{I(\xi_i = 0)I(Y_i \leq t)W_i(h_n, z)}{N(Y_i, z)},
\]

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

\[
\hat{G}_n(t|z) = \prod_{s \leq t} \left( 1 - \Delta \hat{\Lambda}_n(s|z) \right) = \prod_{i=1}^{n} \left( 1 - \frac{I(\xi_i = 0)I(Y_i \leq t)W_i(h_n, z)}{N(Y_i, z)} \right),
\]

where \( \Delta \hat{\Lambda}_n(s|z) = \hat{\Lambda}_n(s|z) - \hat{\Lambda}_n(s^-|z) \).

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

\[
(2.4) \quad \hat{r}_{jn}(z) = \sum_{i=1}^{n} \psi(Y_i)I(Y_i \leq \tau_z)I(\xi_i = j)W_i(h_n, z),
\]

for \( 1 \leq j \leq m \).

Note that if \( C \) is independent of the covariate \( Z \) then \( \hat{G}(t|z) = \hat{G}(t) \) and then the weights \( W_i(h_n, z) \) are replaced by \( 1/n \) in \( \hat{G}_n(.|z) \) 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 \( \hat{r}_j \) 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 sub-conditional distribution function given in the previous section and some conditions on both the kernel function and the bandwidth.

(F1) $\inf_{z \in \Delta} H(\tau_z | z) = \eta > 0$.
(F2) The marginal density function $f$ is continuous on $\Delta$. We note $\alpha = \inf_{z \in \Delta} f(z) > 0$.
(F3) Functions $f$ and $z \mapsto H^{(j)}(t|z)$ (for all $t \in [0, \tau_z], j = 0, \ldots, m$) 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 $\Delta$ uniformly in $t \in [0, \tau_z]$.
(F4) Let $P$ be a polynomial and $\phi$ a positive bounded real function of bounded variation such that $\bar{K} = \phi \circ P$ is a kernel function satisfying:

\begin{enumerate}[(i)]
  \item $\int_{\mathbb{R}^d} K(x)dx = 1$,\
  \item $\int_{\mathbb{R}^d} xK(x)dx = 0$,\
  \item $\int_{\mathbb{R}^d} x x^T K(x)dx$ is positive definite.
\end{enumerate}

(F5) $\sup_{z \in \Delta} \int_{[0, \tau_z]} |d\psi(s)| \leq M < +\infty$.
(F6) The sequence of bandwidth $(h_n)$ satisfies:

\begin{enumerate}[(i)]
  \item $h_n \to 0$,\
  \item $\frac{nh_n}{\log h_n} \to +\infty$,\
  \item $\frac{\log h_n}{\log \log n} \to +\infty$,\
  \item $h_n \leq ch_n + 2n$
\end{enumerate}

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 $\hat{r}_j$. Assumptions (F3) and (F4) (i)–(iii) allows to control the sup-norm distance between any involved function $L$ and its regularized version $L \ast K_h$. 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:

$$\sup_{z \in \Delta} \int_{[0, \tau_z]} \left| d\frac{\psi(s)}{G(s | z)} \right| \leq M < +\infty.$$  \hfill (3.5)

As a consequence it guaranties that $\hat{r}_j(z)$ 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 $\Gamma$ be a compact subset of $\mathbb{R} \times \mathbb{R}^d$, $\Delta$ the restriction of $\Gamma$ to $\mathbb{R}^d$, and $\ell : \Gamma \to \mathbb{R}$ be a function such that $z \mapsto \ell(t, z)$ is twice continuously differentiable on the compact set $\Delta$ and the partial derivative $(t, z) \mapsto \frac{\partial^2 \ell}{\partial z^2}(t, z)$ is continuous on $\Gamma$. Then under assumption (F4)\textsuperscript{\textasteriskcentered}.

$$\sup_{(t, z) \in \Gamma} \left| \int_{\mathbb{R}^d} K_{h_n}(z - s)\ell(t, s)ds - \ell(t, z) \right| = O(h_n^{2d}).$$
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 Assumptions (F1)-(F4), and (F6) hold. Then for $0 \leq j \leq m$ we have

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

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

$$H_n^{(j)}(t|z) - H(t|z) = \frac{1}{nf_n(z)} \sum_{k=1}^n I(Y_k \leq t, \xi_k = j)K_{h_n}(z - Z_k) - H(t|z)$$

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

$$+ H(t|z)f(z)f_n(z) - f_n(z),$$

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

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

where

$$A_n = \sup_{(t,z) \in \Gamma} \left| \frac{1}{n} \sum_{k=1}^n I(Y_k \leq t, \xi_k = j)K_{h_n}(z - Z_k) - \mathbb{E}(I(Y_k \leq t, \xi_k = j)K_{h_n}(z - Z_k)) \right|,$$

$$B_n = \sup_{(t,z) \in \Gamma} \left| \mathbb{E}(I(Y_k \leq t, \xi_k = j)K_{h_n}(z - Z_k)) - H(t|z)f(z) \right|,$$

$$C_n = \sup_{(t,z) \in \Gamma} \left| H(t|z)(f(z) - f_n(z)) \right|.$$

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

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

where

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

By Assumption (F3) and (F4), applying Lemma 3.1 we obtain

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

As a straightforward consequence of (3.7) and (3.9) and because of Assumption (F2) we have

$$\inf_{z \in \Delta} f_n(z) = \alpha + o(1) \text{ a.s.}$$
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
\[ (3.10) \quad C_n = O \left( \left( \frac{\log h_n}{nh_n^d} \right)^{1/2} \right) + O \left( h_n^{2d} \right) \text{ a.s.} \]

Let us now consider \( B_n \). First remark that
\[ E[I(Y_k \leq t, \xi_k = j) K_{h_n}(z - Z_k)] = E[K_{h_n}(z - Z_k) E[I(Y_k \leq t, \xi_k = j)|Z_k]] \]
\[ = \int_{\mathbb{R}^d} K_{h_n}(z - s) H^{(j)}(t|s)f(s)ds. \]
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
\[ (3.11) \quad B_n = O \left( h_n^{2d} \right). \]

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 \((Y, \xi, Z)\), we note \( P_{g_{t,z,h}} = E[g_{t,z,h}(Y,\xi,Z)] \). Considering the empirical measure \( P_n = \sum_{i=1}^n \delta_{(Y_i,\xi_i,Z_i)} \) where \( \delta_x \) denotes the Dirac measure at \( x \), we have
\[ P_ng_{t,z,h} = \sum_{i=1}^n I(Y_i \leq t, \xi_i = j)K \left( \frac{z-Z_i}{h_n} \right). \]
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 \( \mathbb{R}^{d+2} \),
\[ \mathcal{N}(\varepsilon\| K_{\infty}, \mathcal{G}, L^2(Q)) \leq \left( \frac{A}{\varepsilon} \right)^{\nu} \quad 0 < \varepsilon < 1, \]

7
where \( A \) and \( \nu \) are suitable constants. The measurability follows from the continuity of the kernel function and the measurability of the indicator functions. Let us consider \( G_k \), the VC-subclasses of \( G \) defined by

\[
G_k = \{ g_{t,z,h} \in G; \ (t,z) \in \Gamma, \ h_{2^{k}} < h \leq h_{2^{k-1}} \}
\]

for \( k \geq 1 \), we have:

\[
\sup_{g \in G_k} \| g \|_\infty \leq \| K \|_\infty = u_k,
\]

and

\[
\sup_{g \in G_k} \text{Var}[g_{t,z,h}(Y,\xi,Z)] \leq \int_{\mathbb{R}^d} K^2 \left( \frac{z-s}{h} \right) f(s) ds \leq \| f \|_\infty \| K \|_2^2 h_{2^{k-1}}^d = \sigma_k^2.
\]

The end of the proof follows the lines of the proof of Theorem 2.3 of Giné and Guillou (2002). First by the Montgomery-Smith’s maximal inequality (Montgomery-Smith, 1993; see (2.9) in Giné and Guillou, 2002) we have for \( C > 0 \) and \( a_n = (nh_n^d/|\log h_n|)^{1/2} \):

\[
\mathbb{P} \left( \max_{2^{k-1} < n \leq 2^k} a_n \sup_{(t,z) \in \Gamma} |(nh_n^d)^{-1}(P_n - nP)g_{t,z,h_n}| > C \right)
\]

\[
\leq \mathbb{P} \left( \max_{2^{k-1} < n \leq 2^k} a_n \sup_{g \in G_k} |(nh_n^d)^{-1}(P_n - nP)g| > C \right)
\]

\[
\leq 9 \mathbb{P} \left( \sup_{g \in G_k} \| (P_{2^k} - 2^k P)g \| > \frac{C (2^{k-1}h_{2^k}^d |\log h_{2^k}|)^{1/2}}{30} \right).
\]

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 \( C > 0 \):

\[
(3.12) \quad \mathbb{P} \left( \sup_{g \in G_k} \| (P_{2^k} - P)g \| > \frac{C (2^{k-1}h_{2^k}^d |\log h_{2^k}|)^{1/2}}{30} \right) \leq K \exp \left( -K' \log \left( \frac{u_k}{\sigma_k} \right) \right)
\]

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

\[
\frac{\log(u_k/\sigma_k)}{\log k} \to +\infty,
\]

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

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

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

**Theorem 3.1** Assume that Assumptions (F1)-(F4) and (F6) hold. Then for \( n \) large enough we have

\[
\sup_{(t,z) \in \Gamma} |G_n(t|z) - G(t|z)| = O \left( \left( \frac{|\log h_n|/nh_n^d}{\log h_n} \right)^{1/2} \right) + O \left( h_n^{2d} \right) \quad \text{a.s.}
\]
Proof. Let us introduce the following modified product-limit estimator of $\tilde{G}(t|z)$

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

Under Assumptions (F2), (F4), and (F6) (i)-(ii), it is easy to see that

$$\sum_{i=1}^n W_i^2(h_n, z) = \frac{1}{nh_n^d} \left( \frac{1}{f(z)} \int_{\mathbb{R}^d} K^2(x)dx + O(1) \right) = O\left((nh_n^d)^{-1}\right) \text{ a.s.}$$

and hence applying to $G_n(t|z) - \tilde{G}_n(t|z)$ the inequality

$$\left| \prod_{i=1}^n a_i - \prod_{i=1}^n b_i \right| \leq \sum_{i=1}^n |a_i - b_i|$$

for any real numbers $|a_i| \leq 1$ and $|b_i| \leq 1$, it follows that

$$\sup_{(t,z) \in \mathcal{I}} \left| G_n(t|z) - \tilde{G}_n(t|z) \right| \leq \frac{1}{\inf_{z \in \Delta} N^2(\tau_z, z)} \sum_{i=1}^n W_i^2(h_n, z).$$

But by (F1), the fact that $N(t, z) = \bar{H}_n(t^-, z)$ and Lemma 3.2 we have for $n$ large enough

$$\frac{1}{N^2(\tau_z, z)} = \frac{1}{H^2(\tau_z|z) + (N^2(\tau_z, z) - H^2(\tau_z|z))} \leq \frac{1}{H^2(\tau_z|z) - 2\sup_{z \in \Delta} |H_n(\tau_z, z) - H(\tau_z)|} \leq \frac{1}{\eta^2 + o(1)} = O(1) \text{ a.s.}$$

By (3.14), (3.15) and (3.16) we obtain

$$\sup_{(t,z) \in \mathcal{I}} \left| G_n(t|z) - \tilde{G}_n(t|z) \right| = O\left((nh_n^d)^{-1}\right) \text{ a.s.}$$

As in Breslow and Crowley (1974) (see also Stute (1994, p 328), and Sun (2003)), $\tilde{G}_n(t|z) - G(t|z)$ can be represented in the form

$$\tilde{G}_n(t|z) - G(t|z) = G(t|z)[\Lambda_n(t|z) - \tilde{\Lambda}(t|z)] + \frac{1}{2} \exp(-\Lambda^*(t, z))[\Lambda_n(t|z) - \tilde{\Lambda}(t|z)]^2 \quad \text{and} \quad B_n(t, z) = \log(1 - \tilde{G}_n(t|z)) + \tilde{\Lambda}_n(t|z).$$

where $\Lambda^*(t, z)$ lies between $\tilde{\Lambda}(t|z)$ and $\Lambda_n(t|z)$, $\Lambda^{**}(t, z)$ lies between $-\log(1 - \tilde{G}_n(t|z))$ and $\tilde{\Lambda}_n(t|z)$. The remainder of the proof consists in approximating $A_n(t, z) = \tilde{\Lambda}_n(t|z) - \Lambda(t|z)$ and $B_n(t, z) = \log(1 - \tilde{G}_n(t|z)) + \Lambda_n(t|z)$. Recall that

$$\tilde{\Lambda}_n(t|z) = \int_0^t \frac{dH_n^{(0)}(s|z)}{1 - H_n(s^{|z}|)} = \sum_{i=1}^n \frac{I(\xi_i = 0)I(Y_i \leq t)W_i(h_n, z)}{1 - H_n(Y_i^{|z})}$$

$$\sum_{i=1}^n \frac{I(\xi_i = 0)I(Y_i \leq t)W_i(h_n, z)}{N(Y_i, z) + W_i(h_n, z)}.$$
Thus using the inequality \(-\log(1 - y) - y \leq y^2(1 - y)^{-1}\) that holds for \(y \in (0, 1)\), it is readily seen that
\[
\log(1 - \tilde{G}_n(t|z)) + \tilde{\Lambda}_n(t|z) \leq \frac{1}{\inf_{z \in \Delta} N^2(\tau_z, z)} \sum_{i=1}^n W_i^2(h_n, z).
\]
Hence by (3.14) and (3.16),
\[
\sup_{(t,z) \in \Gamma} |B_n(t, z)| = O\left((nh_n^d)^{-1}\right) \quad \text{a.s.}
\]
Furthermore we check that \(\tilde{\Lambda}_n(t|z) - \tilde{\Lambda}(t|z)\) can be written in the form
\[
(3.18) \quad \tilde{\Lambda}_n(t|z) - \tilde{\Lambda}(t|z) = C_n(t, z) + D_n(t, z),
\]
where
\[
C_n(t, z) = \int_0^t \frac{d(H_n^0(s|z) - H(0)(s|z))}{1 - H(s^-|z)},
\]
\[
D_n(t, z) = \int_0^t \frac{(H_n(s^-|z) - H(s^-|z)) - \int_0^t (H_n^0(s|z) - H(0)(s|z)) d(H(s^-|z)^{-1})}{(1 - H(s^-|z))(1 - H_n(s^-|z))} dH_n^0(s|z).
\]
By the integration by parts formula and by Assumption (F1) we have
\[
C_n(t, z) = \frac{H_n^0(t|z) - H(t|z)}{1 - H(t^-|z)} - \int_0^t \left( H_n^0(s|z) - H(0)(s|z) \right) d(H(s^-|z)^{-1})
\leq \sup_{(t,z) \in \Gamma} \left| H_n^0(t|z) - H(0)(t|z) \right| \times \frac{2 - \eta}{\eta}.
\]
Hence by Lemma 3.2 we obtain
\[
(3.19) \quad \sup_{(t,z) \in \Gamma} |C_n(t, z)| = O\left((\log h_n / nh_n^d)^{1/2}\right) + O\left(h_n^2d\right) \quad \text{a.s.}
\]
Remark that
\[
(3.20) \quad \sup_{(t,z) \in \Gamma} \int_0^t |dH_n^{(j)}(s^-|z)| \leq \sup_{(t,z) \in \Gamma} \sum_{i=1}^n W_i(h_n, z) \leq 1.
\]
Moreover it is easy to see that for \(n\) large enough
\[
|D_n(t, z)| \leq \sup_{(t,z) \in \Gamma} \left| H_n^{(0)}(t|z) - H^{(0)}(t|z) \right| \times \sup_{(t,z) \in \Gamma} \int_0^t |dH_n^{(j)}(s^-|z)|
\times \frac{1}{\eta(\eta - \sup_{(t,z) \in \Gamma} |H_n^{(0)}(t, z) - H^{(0)}(t, z)|)} \quad \text{a.s.}
\leq O(1) \times \sup_{(t,z) \in \Gamma} \left| H_n^{(0)}(t|z) - H^{(0)}(t|z) \right| \quad \text{a.s.}
\]
and applying again Lemma 3.2 we obtain
\[
(3.21) \quad \sup_{(t,z) \in \Gamma} |D_n(t, z)| = O\left((\log h_n / nh_n^d)^{1/2}\right) + O\left(h_n^2d\right) \quad \text{a.s.}
\]
Finally (3.18), (3.19) and (3.21) yield
\[
\sup_{(t,z) \in \Gamma} |A_n(t, z)| = O \left( \left( \log h_n / nh_n^d \right)^{1/2} \right) + O \left( h_n^{2d} \right) \quad \text{a.s.}
\]
and the proof of the theorem is complete. \qed

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 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:

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

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

\[
\hat{r}_{jn}(z) - \bar{r}_j(z) = \int_0^{\tau_z} \frac{\psi(s)}{G_n(s^-|z)} dH_n^{(j)}(s|z) - \int_0^{\tau_z} \frac{\psi(s)}{G(s^-|z)} dH^{(j)}(s|z)
\]

\[
\quad = Q_n(\tau_z, z) + R_n(\tau_z, z),
\]

where

\[
Q_n(\tau_z, z) = \int_0^{\tau_z} \frac{\psi(s)}{G(s^-|z)} d(H_n^{(j)}(s|z) - H^{(j)}(s|z)),
\]

\[
R_n(\tau_z, z) = \int_0^{\tau_z} \frac{\psi(s)(G_n(s^-|z) - G(s^-|z))}{G_n(s^-|z)G(s^-|z)} dH_n^{(j)}(s|z).
\]

Using (F5) and the integration by parts formula we have

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

and

\[
R_n(t, z) = \int_0^t \frac{\psi(s)(G_n(s^-|z) - G(s^-|z))}{G_n(s^-|z)G(s^-|z)} dH_n^{(j)}(s|z).
\]

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

\[
\sup_{z \in \Delta} |Q_n(t, z)| \leq ||\psi||_{\infty} (\eta^{-1} + M') \sup_{z \in \Delta} \left| H_n^{(j)}(t|z) - H^{(j)}(t|z) \right|,
\]

then by Lemma 3.2 we derive

\[
\sup_{z \in \Delta} |Q_n(t, z)| = O \left( \left( \log h_n / nh_n^d \right)^{1/2} \right) + O \left( h_n^{2d} \right) \quad \text{a.s.}
\]
By Assumptions (F1) and (F5) and (3.20) we have

\[
\sup_{z \in \Delta} |R_n(t, z)| \\
\leq \frac{\|\psi\|_{\infty}}{\eta} \sup_{(t, z) \in \Gamma} \frac{1}{G_n(t, |z|)} \sup_{z \in \Delta} |G_n(t|z) - G(t|z)| \\
\leq \frac{\|\psi\|_{\infty}}{\eta} \sup_{\eta - \sup_{(t, z) \in \Gamma} |G_n(t, |z|) - G(t, |z|)|} \sup_{(t, z) \in \Gamma} |G_n(t|z) - G(t|z)|.
\]

By Theorem 3.1 for \( n \) large enough we have

\[
\frac{1}{\eta - \sup_{(t, z) \in \Gamma} |G_n(t, |z|) - G(t, |z|)|} = O(1) \quad a.s.
\]

then it follows that

\[
(3.25) \quad \sup_{z \in \Delta} |R_n(t, z)| = O \left( \left( \frac{|\log h_n|}{nh_n^d} \right)^{1/2} \right) \quad a.s.
\]

Finally by (3.23), (3.24) and (3.25) we obtain the expected uniform convergence rate for \( \hat{r}_{jn} - \bar{r}_j \). \( \Box \)

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 \( \psi \) 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

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

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 $O(n^{-d/2})$ term involved in the rates come from the regularity of the function we want to estimate 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 ($K_1$) are possible, like for example the uniform kernel on $[-1,1]^d$.

References


